

OPTIMIZING PLANNING AND DELIVERY OF HIGH-PRECISION ROBOTIC RADIOTHERAPY AND INTENSITY-MODULATED PROTON THERAPY

Optimizing Planning and Delivery of High-Precision Robotic Radiotherapy and Intensity-Modulated Proton Therapy

© S. van de Water, Rotterdam 2015

ISBN: 978-94-6108-979-3

Lay-out: Zink Typografie (www.zinktypografie.nl)

Cover design: sophie ontwerpt// Printed by: Gildeprint Enschede

The research described in this thesis was performed at the Erasmus MC Cancer Institute (chapters 2–7 and 9) and the Paul Scherrer Institute, Switzerland (chapter 8).

Financial support for the printing of this thesis was provided by Accuray.

Optimizing Planning and Delivery of High-Precision Robotic Radiotherapy and Intensity-Modulated Proton Therapy

Optimalisatie van de 'planning' en afgifte van hoge-precisie robotbestraling en intensiteitsgemoduleerde protonentherapie

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Erasmus Universiteit Rotterdam op gezag van de rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties. De openbare verdediging zal plaatsvinden op

dinsdag 9 juni 2015 om 11.30 uur

Steven van de Water geboren te Veghel

Erasmus University Rotterdam

(zafus

Promotiecommissie

Prof.dr. B.J.M. Heijmen Promotor:

Overige leden: Prof.dr. J.P. Pignol Prof.dr. R.J. Baatenburg de Jong

Prof.dr. D. De Ruysscher

Dr. M.S. Hoogeman Copromotor:



Contents

1	introduction	1
Paı	rt I: High-precision robotic radiotherapy using the CyberKnife	
2	Variable circular collimator in robotic radiosurgery: A time-efficient alternative to a mini-multileaf collimator?	13
3	Shortening treatment time in robotic radiosurgery using a novel node reduction technique	31
4	Intrafraction prostate translations and rotations during hypofractionated robotic radiation surgery: Dosimetric impact of correction strategies and margins	49
Paı	rt II: Intensity-modulated proton therapy	
5	Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams	67
6	Shortening delivery times of intensity-modulated proton therapy by reducing proton energy layers during treatment plan optimization	85
7	Dose uncertainties in IMPT for oropharyngeal cancer in the presence of errors	99
8	Tumor tracking with scanned proton beams: assessing the accuracy and practicalities	115
9	The price of robustness; impact of worst-case optimization on organ-at-risk dose in intensity-modulated proton therapy for oropharyngeal cancer patients	135
10	Discussion	149

viii | Contents

Appendices	163
References	193
List of publications	205
Summary	207
Samenvatting	213
Dankwoord	219
Curriculum vitae	221
PhD portfolio	223

chapter 1

Introduction

1 1 Radiotherapy

Radiotherapy is a treatment for cancer that kills tumor cells by means of ionizing radiation. Radiation damages the DNA of cells and thereby stops further cell division. Approximately 50% of cancer patients are treated with radiotherapy, often in combination with surgery or chemotherapy, and either with curative or palliative intent (1). Patients treated with curative intent generally receive the radiation dose in multiple treatment fractions (approximately 20 fractions on average) as a means to reduce the damage in healthy tissues and 'organs-at-risk'. Prior to treatment, a personalized treatment plan is made using a 'planning' CT-scan of the patient in treatment position. The physician contours the tumor and organs-at-risk in this CT-scan and provides the dose prescriptions, after which the treatment plan can be generated. This plan specifies the delivery settings (e.g. beam directions, field shapes and beam intensities) with which the patient is irradiated and describes the anticipated dose distribution inside the patient. A treatment plan is typically generated once before the start of the first treatment fraction and is applied in all subsequent fractions. The most favorable delivery settings are obtained by performing a mathematical optimization, which aims at finding a balance between the tumor dose and healthy tissue dose. The tumor should receive a sufficiently high dose to achieve local tumor control, whereas the dose in healthy tissues and organs-at-risk should be kept as low as possible or below a certain tolerance level in order to minimize the risk of treatment-related side effects.

In recent decades, radiotherapy has improved greatly by the introduction of techniques that enabled healthy tissue doses to be reduced. Such techniques include intensity-modulated radiotherapy, image-guided radiotherapy and rotational delivery techniques in which the patient is irradiated from many directions (e.g. tomotherapy or volumetric modulated arc therapy) (2–9). Two examples of relatively new radiotherapy techniques,

particularly aiming at a highly localized dose delivery and sparing of surrounding healthy tissues, are high-precision robotic radiotherapy and intensity-modulated proton therapy (IMPT) (10-13). For both treatment modalities, promising treatment outcomes have been obtained for a selection of tumor sites (14-18). However, the application of robotic radiotherapy and IMPT for a wide range of treatment sites is currently limited by a number of obstacles concerning the treatment accuracy and the efficiency of treatment planning and delivery.

The aim of this thesis is to address these obstacles for both treatment modalities. The core of this thesis is divided into two parts: Part I (chapters 2-4) deals with improving the delivery efficiency and accuracy of high-precision robotic radiotherapy using the CyberKnife, and Part II (chapters 5-9) deals with improving the efficiency of treatment planning and delivery for IMPT and addresses the high sensitivity of IMPT treatment plans to treatment uncertainties. This thesis is concluded by a general discussion covering both treatment modalities in chapter 10.

1 7 High-precision robotic radiotherapy using the CyberKnife

The CyberKnife robotic radiosurgery system consists of a small linear accelerator, producing 6-MV photons (X-rays), which is mounted on a robotic manipulator as depicted in figure 1.1. This configuration allows the CyberKnife to move relatively freely around the patient and irradiate the patient from many directions. The robotic manipulator can be positioned at so-called 'node positions', which are approximately 120 pre-defined locations distributed semi-spherically around the patient. From each node position, multiple fields can be delivered by adjusting the orientation of the linear accelerator. The system is equipped with an integrated stereoscopic imaging system, which allows for positioning errors to be corrected in between and during treatment fractions. Irradiating the patient from many (non-coplanar) beam directions using hundreds of small beams while performing image-guidance enables adequate irradiation of the tumor and a very good sparing of healthy tissue and organs-at-risk. As a result, excellent clinical outcomes have been obtained in hypofractionated treatments (typically 3-4 fractions) for stage I non-small-cell lung cancer patients and for low to intermediate risk prostate cancer patients (14, 15).

Although clinical outcomes have been very promising, CyberKnife treatments are also associated with a number of challenges. Firstly, the treatment time per fraction is relatively long compared with conventional radiotherapy techniques (19). This affects patient comfort, treatment accuracy and may reduce the biological effectiveness of a

4 | Introduction



Figure 1.1: The CyberKnife robotic radiosurgery system at the Erasmus MC Cancer Institute in Rotterdam.

treatment. Moreover, the long fraction duration causes the CyberKnife to be unsuitable for conventionally fractionated treatments, as this would result in the overall treatment time to become unacceptably long. A substantial shortening of the fraction duration will facilitate a more widespread use of the CyberKnife, as it enables tumor sites to be treated for which hypofractionation cannot be applied. A second issue concerns the accuracy of 'tumor tracking' using the integrated imaging system. During tumor tracking, images are acquired with a certain time interval (not continuously) and tumor displacements are subsequently corrected for by adjusting the position of the robotic manipulator (up to a certain extent). Prostate tumors, however, were shown to display very rapid excursions of considerable magnitude (20, 21). Prostate displacements can therefore remain undetected in between the acquisition of successive images and might not be fully accounted for by robot corrections. The resulting dosimetrical accuracy of tumor tracking and the optimal imaging and correction settings have not been determined before for CyberKnife prostate treatments. We address these issues regarding treatment efficiency and treatment accuracy of the CyberKnife in Part I of this thesis. More specifically, we considered the following research objectives:

Improving treatment efficiency by using a mini-multileaf collimator

Traditionally, CyberKnife systems are equipped with collimators that produce circular fields. The field diameter can be specified either by applying a certain fixed-aperture collimator or by using the 'Iris' variable-aperture collimator. Alternatively, the use of a mini-multileaf collimator (mMLC) would allow the fields to be shaped freely and to match the tumor shape. A mMLC does not restrict the fields to a circular shape and is therefore expected to enable more efficient dose delivery. In chapter 2, we present the first treatment planning study that was performed to compare the Iris collimator with a mMLC, evaluating the expected delivery time and plan quality for early-stage lung cancer patients.

Improving treatment efficiency by reducing the number of node positions in the treatment plan

During a CyberKnife treatment, the time required for the robotic manipulator to travel between node positions takes up the largest part of the delivery time per fraction. A reduction of the number of node positions in the treatment plan is therefore likely to result in a shortening of the fraction duration. A method to reduce the number of node positions while preserving plan quality is described in chapter 3. The method was evaluated for two lung cases, a prostate case and a head-and-neck case. A similar approach is presented in chapter 6 to reduce the number of energy layers in IMPT treatment plans.

Assessing the impact of intrafraction prostate motion and the effectiveness of tumor tracking

Prostate tumors can display considerable motion during a treatment fraction (intrafraction motion), for example due to organ filling and peristaltic motion. During CyberKnife prostate treatments, fiducial marker tracking is used to correct for tumor displacements and an additional safety margin of 3 mm around the prostate is applied to account for any residual errors. However, the effectiveness of these measures to account for intrafraction motion has not been studied before. In chapter 4, we assessed the dosimetric impact of intrafraction prostate motion and we investigated the effect of the frequency and extent of robot corrections, and of the applied safety margin.

1 3 Intensity-modulated proton therapy

IMPT is a radiotherapy treatment in which the patient is irradiated using protons instead of the traditional photons. Protons are positively charged particles that interact

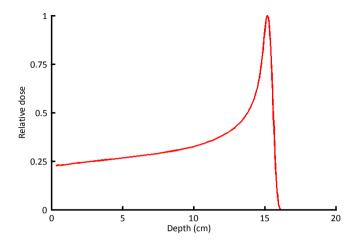


Figure 1.2: Integral dose of a 148 MeV proton beam as function of the depth.

differently with tissue than photons. Protons have a finite range in tissue and deposit a high peak-dose just before the end-of-range. This so-called 'Bragg peak' is depicted in figure 1.2. The range of the protons and the resulting depth of the Bragg peak depend on the initial energy of the proton beam. A low-dose plateau can be observed in front of the Bragg peak and the dose decreases rapidly to zero beyond the peak. As a consequence, the Bragg peak is very suitable to deposit a high dose inside the tumor while keeping the dose in the surrounding healthy tissues relatively low. In IMPT, thousands of small proton pencil beams (or 'spots') are delivered to the tumor one-by-one, each time adjusting the lateral position and/or energy of the proton beam (12, 22). The lateral spot position can be adjusted rapidly by magnetically deflecting the positively charged protons. By varying the intensity of each spot, the dose can be modulated such that it conforms accurately to the tumor shape. Promising clinical outcomes have been reported for a number of tumor sites, including skull-base tumors and tumors in pediatric patients (16–18), but a more widespread application of IMPT is currently limited by a number of issues.

Firstly, the use of thousands of small beams to achieve a highly localized dose deposition puts pressure on the time-efficiency of treatment planning and treatment delivery. The time required to generate a high-quality IMPT treatment plan is relatively long, because the optimization time scales polynomially with the degrees of freedom (i.e. the number of proton pencil beams). The delivery time per fraction can become relatively long as well, especially when compared with time-efficient rotational delivery techniques in photon therapy (23). In particular, the time spent on adjusting the energy of the proton



Figure 1.3: Dose distributions for a non-robustly optimized IMPT treatment plan. The tumor is delineated in white and organs-at-risk are delineated in yellow. The colors indicate isodose levels (blue = low dose, red = high dose). The left panel shows the intended (planned) dose distribution without treatment errors. The right panel depicts a degraded (delivered) dose distribution resulting from errors in patient setup and in proton range. The isodose lines nicely conform to the target volumes in the left panel, whereas they display a disturbed pattern in the right panel.

beam (to adjust the spot location in depth) can take up a large proportion of the delivery time. Improved efficiency of planning and delivery will improve plan quality, patient comfort and treatment accuracy, and maybe most importantly, it will contribute to a reduction of treatment costs. IMPT treatment costs (both investment and operational) are higher than in conventional radiotherapy. As a result, an IMPT treatment fraction is on average approximately 3 times more expensive than a conventional treatment fraction (24). Improved time-efficiency might make IMPT a more attractive and affordable treatment modality for a wider range of treatment sites.

The second issue concerns the fact that IMPT is very sensitive to treatment uncertainties. The highly localized dose deposition improves healthy tissue sparing on the one hand, but it makes proton therapy also very sensitive to motions of the tumor and surrounding tissues that may occur on all relevant time scales (from seconds to days to weeks). This sensitivity can severely limit the targeting accuracy and affect the efficacy and safety of IMPT. As a result, it is not trivial that the high plan quality obtained during treatment plan generation can be realized during actual treatment delivery. This is illustrated by the dose distributions depicted in figure 1.3. The figure shows that the intended dose distribution can degrade substantially due to errors in patient setup and proton range. In order to unleash the full potential of IMPT, it is of great importance to accurately quantify the degradation of the intended dose distribution and to design strategies to mitigate these dose degradations, thereby ensuring an effective and safe delivery of IMPT. We address the abovementioned issues regarding IMPT in Part II of this thesis. More specifically, we considered the following research objectives:

Improving efficiency of high-quality treatment plan generation

The time required to generate a high-quality IMPT treatment plan can be relatively long, due to the large number of beams that is optimized. In chapter 5, we present a new treatment planning approach for IMPT called 'pencil beam resampling'. Resampling was originally developed to improve the efficiency of CyberKnife treatment planning. The method aims at reducing optimization times by dividing the large optimization problem into multiple smaller sub-problems, and by allowing the proton pencil beams (the Bragg peaks) to be distributed more freely within the target volume. We evaluated the method in terms of optimization time and corresponding plan quality for 5 head-and-neck cancer patients.

Improving treatment efficiency by reducing the number of energy layers in the treatment plan

In current commercial IMPT systems, the time required to adjust the energy of the proton beam is the major component of the treatment time. Treatments can be made more time-efficient by shortening the energy switching times, but also by reducing the number of energy layers in an IMPT treatment plan. A method to reduce the number of energy layers in robust IMPT treatment plans while preserving plan quality is presented and evaluated in chapter 6. The method is similar to the node reduction technique as presented in chapter 3 for the CyberKnife.

Assessing the sensitivity of IMPT in the presence of treatment uncertainties

In chapter 7, we performed extensive treatment simulations to investigate the impact of uncertainties in patient position, patient anatomy and proton range on the delivered dose for high-quality IMPT treatment plans of 10 head-and-neck cancer patients. The treatment plans were generated using the conventional approach of applying a safety margin around the tumor in order to account for treatment uncertainties. In addition, we investigated whether the dose uncertainty could be reduced by generating a new

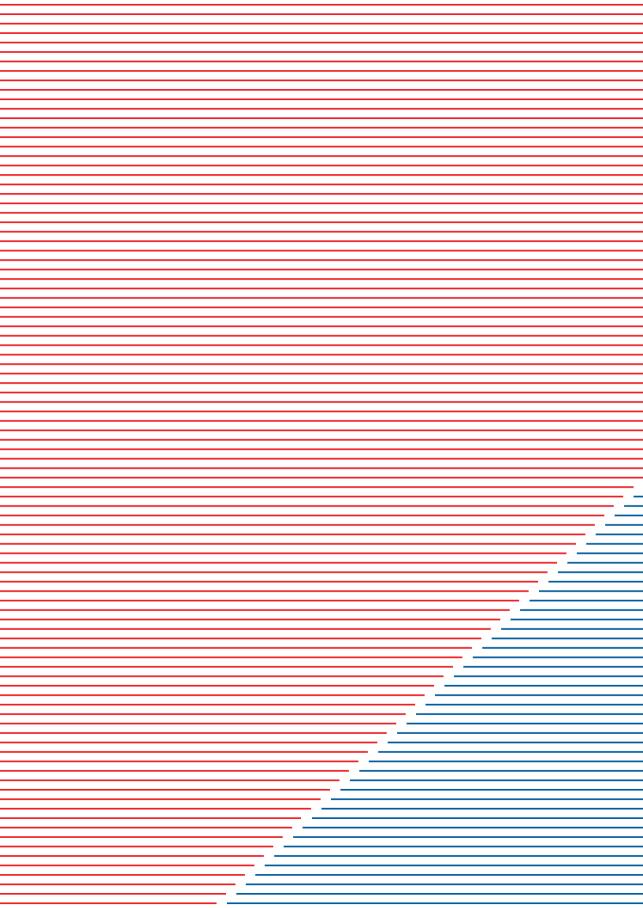
treatment plan halfway the treatment course (adaptive planning) or by increasing the number of beam directions.

Assessing impact of respiratory motion and the effectiveness of tumor (re) tracking

Dose uncertainties due to respiratory motion can be reduced either by minimizing position errors during dose delivery or by reducing the impact of the position errors on the delivered dose. The first can be achieved by performing 'real-time tumor tracking', in which the moving tumor is continuously followed by magnetically adjusting the position of the proton pencil beam. The latter can be achieved by performing 'rescanning', in which the position errors are averaged out by scanning the tumor volume several times while delivering a fraction of the dose in each rescan. Combining the two methods results in 'retracking' and might further reduce dose uncertainties. In chapter 8, we investigated the effectiveness of real-time tumor (re)tracking to deal with respiratory motion during IMPT treatments.

Assessing the price of robust treatment plan generation

The use of traditional safety margins around the tumor is considered inadequate to deal with treatment uncertainties in IMPT, as margins do not account for the additional uncertainty in proton range. Instead, so-called 'robust' treatment planning should be performed, in which possible error scenarios (errors in patient position and proton range) are explicitly included in the plan optimization. This will minimize the dose uncertainty in the tumor and healthy tissues for treatment errors up to a certain extent. However, the magnitude of treatment errors accounted for during robust treatment planning is likely to have an effect on the dose received by surrounding healthy tissues and organs-at-risk. We assessed the effect of the degree of robustness on the dose received by organs-at-risk and the resulting risk of treatment-related side effects in **chapter 9**.



PART I HIGH-PRECISION ROBOTIC RADIOTHERAPY **USING THE CYBERKNIFE**

chapter 2

Variable circular collimator in robotic radiosurgery: A time-efficient alternative to a mini-multileaf collimator?

International Journal of Radiation Oncology Biology Physics, Vol. 81, No. 3, pp. 863–870, 2011.

> Steven van de Water¹ Mischa Hoogeman¹ Sebastiaan Breedveld¹ Joost Nuyttens¹ Dennis Schaart² Ben Heijmen¹.

 Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands
 Faculty of Applied Sciences, Section Radiation Detection and Medical Imaging, Delft University of Technology, Delft, the Netherlands

Abstract

Purpose: Compared with many small circular beams used in CyberKnife treatments, beam's eye view-shaped fields are generally more time-efficient for dose delivery. However, beam's eye view-shaping devices, such as a mini-multileaf collimator (mMLC), are not presently available for CyberKnife, although a variable-aperture collimator (Iris, 12 field diameters; 5–60 mm) is available. We investigated whether the Iris can mimic noncoplanar mMLC treatments using a limited set of principal beam orientations (nodes) to produce time-efficient treatment plans.

Methods and Materials: The data from 10 lung cancer patients and the beam-orientation optimization algorithm 'Cycle' were used to generate stereotactic treatment plans (3 × 20 Gy) for a CyberKnife virtually equipped with a mMLC. Typically, 10−16 favorable beam orientations were selected from 117 available robot node positions using beam's eye view-shaped fields with uniform fluence. Second, intensity-modulated Iris plans were generated by inverse optimization of non-isocentric circular candidate beams targeted from the same nodes selected in the mMLC plans. The plans were evaluated using the mean lung dose, lung volume receiving ≥20 Gy, conformality index, number of nodes, beams, and monitor units, and estimated treatment time.

Results: The mMLC plans contained an average of 12 nodes and 11690 monitor units. For a comparable mean lung dose, the Iris plans contained 12 nodes, 64 beams, and 21990 monitor units. The estimated fraction duration was 12.2 min (range, 10.8–13.5) for the mMLC plans and 18.4 min (range, 12.9–28.5) for the Iris plans. In contrast to the mMLC plans, the treatment time for the Iris plans increased with an increasing target volume. The Iris plans were, on average, 40% longer than the corresponding mMLC plans for small targets (<80 cm³) and \leq 121% longer for larger targets. For a comparable conformality index, similar results were obtained.

Conclusion: For stereotactic lung irradiation, time-efficient and high-quality plans were obtained for robotic-controlled non-coplanar treatments using a mMLC. Iris is a time-efficient alternative for small targets, with similar or better plan quality.

2.1 Introduction

Excellent local tumor control has been achieved in patients with early-stage non-smallcell lung cancer treated using the CyberKnife robotic radiosurgery system (Accuray, Sunnyvale, CA). The CyberKnife consists of a compact 6-MV linear accelerator mounted on a robotic manipulator and performs real-time tumor tracking to compensate for respiratory motion. Despite the high local tumor control rates, the long treatment time per fraction is an area for improvement (19).

The fraction duration mainly consists of the beam-on time, robot motion time, and image guidance time and can be reduced by developing time-efficient treatment plans. The beam-on time can be reduced by reducing the number of monitor units (MUs). The robot motion time —the prevalent factor— can be reduced by reducing the number of node positions and beams. The node positions are the preset locations where the robotic manipulator can position the focal spot of the linear accelerator (i.e. the source of the X-ray beam). From these node positions, multiple beams can be targeted at various locations in the tumor by adjusting the orientation of the linear accelerator. Traveling between node positions takes longer than a reorientation of the linear accelerator between beams at a node position. The imaging time is reduced automatically by reducing the overall treatment time, because fewer images are required to ensure position and tracking accuracy.

The number of node positions, beams, and MUs in a treatment plan can be reduced by increasing the degrees of freedom of beam collimation. The CyberKnife at our institute has 12 collimators with fixed circular apertures ranging from 5 to 60 mm in diameter. Generally, only one or two collimators are used per treatment. Clinical treatment plans typically contain 60 node positions, 125 beams, and 35000 MUs. Pöll et al. (19) showed that the number of MUs in lung cancer treatments could be reduced by an average of 31% if two collimator sizes were used instead of one. An even larger reduction in the required number of MUs (~60% in lung cancer treatments) was achieved when a variable-aperture collimator was used (25). This variable-aperture collimator is called the Iris Variable-Aperture Collimator (Accuray) and allows 12 field diameters to be used without the manual exchange of collimators.

A mini-multileaf collimator (mMLC) mounted on the CyberKnife could lead to even more time-efficient treatment plans, because it does not restrict the field to a circular field shape as does the Iris collimator. It allows for three-dimensional conformal radiotherapy (3D-CRT) to be performed using fields shaped according to the beam's eye view projection of the target. This will make dose painting with a large number of circular beams superfluous, thereby reducing the required number of beams and MUs. Moreover, a mMLC is also expected to require a limited set of node positions, because a feasible number of beam directions is 10-15 in stereotactic radiotherapy for lung and liver lesions using a

MLC (26, 27). A CyberKnife equipped with a mMLC could even be more efficient in delivering non-coplanar beams than gantry-based modalities, because time-consuming manual couch adjustments would not be required. However, a mMLC is not yet available for the CyberKnife.

The goal of the present study was first to assess the plan quality and delivery time for computer-optimized non-coplanar 3D-CRT plans designed for a CyberKnife theoretically equipped with a mMLC. The second goal was to compare these mMLC plans with intensity-modulated plans developed for the currently available Iris collimator, only using those node positions selected in the mMLC plans to generate time-efficient treatment plans. To generate the Iris plans, a dedicated inverse planning algorithm was developed in-house for the present study.

7 Methods and materials

Patient group and dose prescriptions

The data from 10 lung cancer patients treated at our clinic with the CyberKnife were used. These patients were also included in the study by Pöll et al. (19). The gross target volume ranged from 5.9 to 78.0 cm³ (median, 14.1) and was delineated using the lung level and window settings on computed tomography (CT). A gross target volume to planning target volume (PTV) margin of 5 mm was used to include microscopic extension of the tumor and to account for inaccuracies of the Synchrony Respiratory Tracking System (28). The PTV ranged from 20.8 to 152.3 cm³ (median, 40.5).

The PTV was prescribed a dose of 60 Gy at an isodose level of ≥80%, to be delivered in three fractions. At least 95% of the PTV had to receive ≥60 Gy. The dose constraints for the organs-at-risk were mainly defined by the Radiation Therapy Oncology Group 0236 protocol for stereotactic radiotherapy for lung cancer (29). These constraints are listed in table 2.1. A conformality index (CI) of ≤1.20 was required, defined as the ratio of the volume receiving the prescription dose (≥60 Gy) and the PTV.

CyberKnife characteristics

One of the standard CyberKnife extracranial node sets was used, which included 117 node positions. The node positions are distributed semispherically (non-coplanar) around the patient at a distance to the imaging center of 800-1000 mm. Each node

Organ	Volume	Dose
Spinal Cord	Any point	6 Gy/fraction
Liver	Any point	8 Gy/fraction
Ipsilateral Brachial Plexus	Any point	8 Gy/fraction
Esophagus	Any point	9 Gy/fraction
Heart	Any point	10 Gy/fraction
Trachea and main Bronchus	Any point	10 Gy/fraction
Ribs	Any point	20 Gy/fraction
Lung	< 10% of total volume	20 Gy in total
Conformality constraint	< 1.2 × PTV	60 Gy in total

Table 2.1: Maximal dose constraints for organs-at-risk and conformality constraint.

Abbreviation: PTV = planning target volume

position effectively acts as a focal spot position. The Iris variable-aperture collimator allows 12 field diameters to be used: 5, 7.5, 10, 12.5, 15, 20, 25, 30, 35, 40, 50, and 60 mm defined at 800 mm from the focal spot (25). The 6-MV linear accelerator delivers an unflattened beam and is calibrated such that 1 MU corresponds to a dose of 1 cGy in a reference point located on the central axis of a 60-mm field, at 800 mm from the focal spot and a 15 mm depth in water.

Step 1: mMLC plans generated using 'Cycle'

The mMLC plans for the CyberKnife were developed using Cycle, an algorithm developed in-house for simultaneous optimization of beam orientations, shapes, and weights (30). Cycle generates treatment plans by sequential selection of beams from a set of user-defined candidate beam orientations. Sequential selection means that the planning process starts with an empty plan and that beams are added one-by-one to the plan. In each iteration, the optimal beam is selected from the candidate beam orientations using a weighted-sum score function, which takes into account the imposed dose constraints and the dose given by the beams previously added to the plan. Favorable beam orientations can be selected more than once, increasing the beam weight. The iterations are performed until a feasible solution has been found (i.e. a solution that satisfies all dose constraints). The minimization or maximization of plan parameters can be achieved by repetitive application of Cycle, each time adjusting the constraint until a feasible solution can no longer be found.

In the present study, the candidate beam orientations were defined by the 117 node positions of the CyberKnife and the center of mass of the tumor as the isocenter. The 3D- CRT fields were shaped, assuming the use of a mMLC with a resolution of 2.5 mm at the isocenter. The beam's eye view projection of the PTV was taken as the initial field shape for each beam. This field shape was subsequently optimized by extending or retracting the edges of the field in four perpendicular lateral directions, according to de Pooter et al. (31). Performing 3D-CRT requires the CyberKnife to be additionally equipped with a flattening filter. Because such a CyberKnife does not exist, the beam data from a gantrybased 6-MV linear accelerator equipped with a mMLC and flattening filter was used, normalized according to the CyberKnife conventions. The field intensity was uniform, and the penumbra was constructed by convolution of the field intensity matrix with an off-axis kernel. The dose was calculated on a rectangular dose grid with a grid spacing of $1 \times 1 \times 3$ mm or $3 \times 3 \times 2$ mm, depending on the CT slice spacing. An equivalent path length algorithm was used to correct for density heterogeneities. After obtaining the final treatment plan, the dose was recalculated using a grid spacing of $1 \times 1 \times 1.5$ mm or 1×1.5 1×2 mm, depending on the CT slice spacing.

Step 2: Iris plans generated using inverse optimization

The treatment plans for the Iris collimator (Iris plans) were generated using a new treatment planning algorithm for the Cyber-Knife that has been developed in-house for the present study. A flow chart of the planning algorithm is depicted in figure 2.1.

The planning process starts with the selection of candidate beams, defined by a node position, a target point, and a collimator size. For each node position, a projection of the PTV was made onto a reference plane at 800 mm from the node position. The target points were regularly distributed over the reference plane with user-defined spacing. For all target points within the projected PTV contour, all 12 collimator sizes were selected as candidate beams. The target point spacing was set to 7-14 mm, depending on the tumor size and number of node positions, to obtain 3000-4000 candidate beams. For each patient, only those node positions selected by Cycle in the mMLC plan were used for candidate beam selection.

The next step is the construction of an individual dose matrix for each structure involved in the optimization, containing the dose deposited (in Gy/MU) in every voxel of the structure by every candidate beam. It was constructed by performing dose calculations for each element of the matrix, using the tissue maximal ratio/off-axis ratio formalism and beam data from the CyberKnife. Density heterogeneities were accounted for by calculating the equivalent path length for each voxel. The use of separate dose matrices for each structure allowed the dose grid spacing to be adjusted to the size of the structure. The default grid spacing was set to $3 \times 3 \times 3$ mm or $3 \times 3 \times 2$ mm, depending on the CT slice spacing. Smaller grid spacing was used to ensure sufficient coverage or

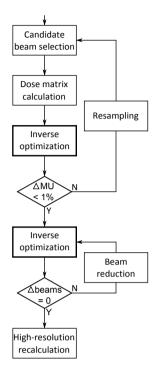


Figure 2.1: Flow chart of inverse planning algorithm for Iris plans.

sparing of small structures. Larger grid spacing was used for large structures such as the lung. The final dose distribution in the entire patient was recalculated using a dose grid spacing of $1 \times 1 \times 1.5$ mm or $1 \times 1 \times 2$ mm, depending on the CT slice spacing.

The core of the planning algorithm is the inverse multi-criteria optimization algorithm developed in-house by Breedveld et al. (32). Using dose matrices and prescriptions as input, it is capable of determining the optimal weights (in MU) of the candidate beams. It makes a distinction in the prescriptions between the constraints and objectives. The constraints are prescriptions that must be fulfilled, otherwise the solution is invalid. Objectives are prescriptions for which the optimal value (minimum or maximum) is searched without violating the constraints. Minimization of the total number of MUs was the only objective in the optimization of the Iris plans. Because the linearity of the dose delivery cannot be guaranteed at <5 MU/fraction, the candidate beams should either be ascribed ≥15 MU (in three fractions) or not selected at all. This was achieved by repeating the inverse optimization, typically one to three times, each time excluding the beams that

had been ascribed ≤15 MU, until all remaining candidate beams had a weight of ≥15 MU. Typically, only a small fraction of the candidate beams (~60-360 beams) was not excluded. Because the CI constraint could not be directly included in the optimization, a maximal dose constraint on a shell structure around the PTV was used as a surrogate.

As can be seen in figure 2.1, the steps described are embedded in the so-called resampling loop. Resampling is a method to improve the optimality of a treatment plan by iteratively increasing the total number of candidate beams (33). It makes use of the fact that most of the candidate beams are excluded by the inverse optimization, owing to the minimal beam weight of 15 MU. In each resampling run, all excluded candidate beams were replaced by newly selected candidate beams, after which the inverse optimization was repeated. The new candidate beams were selected by repeating the candidate beam selection using a target point spacing lowered by 1 mm. A decreasing target point spacing results in an increasing number of candidate beams and resampling was terminated when the maximum of 10000 candidate beams was exceeded. Resampling was also terminated when the improvement in the number of MUs was <1%.

The final optimization step is the minimization of the total number of beams in the treatment plan, indicated by the beam reduction loop in figure 2.1. During beam reduction, the inverse optimization was performed repeatedly, each time excluding beams with the lowest contribution, until a feasible solution could no longer be found. The contribution of a beam was defined as the dose delivered to the isocenter, if the beam would have been targeted at the isocenter.

Plan evaluation

All treatment plans were inspected by a physician and only clinically acceptable plans were included. The quality of the treatment plans was assessed using the biologic equivalent mean lung dose of 2 Gy fractions (MLD), total lung volume receiving ≥20 Gy, CI, number of node positions, number of beams, total number of MUs, and estimated treatment time for each fraction. The MLD is the most accurate parameter to predict the incidence of radiation pneumonitis and was calculated using the linear quadratic model $(\alpha/\beta = 3)$ Gy for the lung) (34). The estimated treatment time for each fraction was calculated using a dedicated algorithm that included the beam-on time, robot motion time, and time needed for imaging during treatment. It assumed a dose rate of 800 MU/min, a CyberKnife G4 robot speed, and approximately 5 s/image pair, acquired every three beams. Because the CyberKnife is not equipped with a flattening filter, the beam-on time of the mMLC plans was increased by 20% to account for flattening of the fields, according to the off-axis profile of the uncollimated CyberKnife beam. The treatment time estimations did not include the time needed for patient setup or for building a

correlation model between the positions of the implanted markers and external markers on the patient's chest (28).

The MLD was used as primary objective in the comparison of the mMLC and Iris plans. For each patient, the MLD was minimized in the mMLC plan and an equal MLD was aimed at in the MLD-equivalent Iris plan. In addition, several Iris plans with varying MLD constraints were developed for each patient. A relaxed MLD constraint (15 Gy) was prescribed for the first plan and was lowered by 0.5 Gy in successive Iris plans until a feasible treatment plan could no longer be found, thereby, resulting in the range of MLD values that can be obtained within other constraints. From all Iris plans, a CI-equivalent Iris plan was selected, having a CI closest to that of the mMLC plan. Finally, the physician selected the one preferred Iris plan from all Iris plans for each patient that in the physician's opinion provided the optimal balance between the plan quality and treatment time.

2.3 Results

Table 2.2 lists the average plan parameters for the mMLC plans, the MLD-equivalent Iris plans, the CI-equivalent Iris plans, and the preferred Iris plans. Cycle selected 10–16 node positions in the mMLC plans, depending on the patient. These node positions were subsequently used in the generation of the Iris plans. The average estimated treatment time/fraction was 12.2 min for the mMLC plans and 18.4 min for the MLD-equivalent Iris plans. When the CI was used as similarity measure, the results differed little from

Table 2.2: Average plan parameters of mMLC plans, MLD-equivalent Iris plans, CI-equivalent Iris plans, and
preferred Iris plans.

Parameter	mMLC	MLD-equivalent Iris	CI-equivalent Iris	Preferred Iris
MLD (Gy)	8.0 (4.8-12.3)	7.8 (4.3–12.3)*	7.9 (4.0–12.2)	7.4 (4.0-11.4)*
V ₂₀ (%)	3.9 (2.4-6.7)	3.8 (2.1-6.4)	3.9 (2.0-6.4)	3.6 (2.0-6.0)*
CI	1.20 (1.14-1.27)	1.17 (1.11-1.20)	1.17 (1.11-1.20)	1.15 (1.09-1.20)
Node positions (n)	12.4 (10-16)	11.7 (10-14)	11.8 (10-15)	11.5 (9-14)*
Beams (n)	12.4 (10-16)	63.5 (31-120)*	59.2 (34-120)*	59.6 (31-127)*
MUs (n)	11690 (10764-13554)	21990 (14692-33358)*	21215 (14692-31217)*	21961 (16695-32849)*
Fraction duration (min)	12.2 (10.8-13.5)	18.4 (12.9-28.5)*	17.8 (12.9-28.5)*	18.1 (14.0-29.8)*

Abbreviations: mMLC = mini-multileaf collimator; MLD = biologically equivalent mean lung dose of 2 Gy fractions; V_{20} = percentage of total lung volume receiving \geq 20 Gy; CI = conformality index (ratio of volume receiving prescription dose [\geq 60 Gy] and PTV); MU = monitor unit; PTV = planning target volume.

^{*} Significantly different (p < 0.05) from mMLC plans (Wilcoxon signed-rank test).

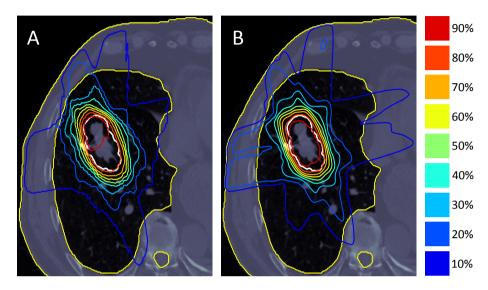


Figure 2.2: Dose distributions of (A) mini-multileaf collimator plan and (B) mean lung dose-equivalent Iris plan for patient 4 (isocenter slice). Thick white lines indicate planning target volume contour. Depicted isodose lines ranged from 10% to 90% of maximal planning target volume dose.

those using the MLD; sometimes, the same plan was selected as being equivalent. The preferred Iris plans selected by the physician had a lower MLD than the MLD-equivalent Iris plans, and a treatment time/fraction of 18.1 min on average.

During Iris plan generation, the inverse planning algorithm performed an average of 7 resampling runs, thereby increasing the total number of candidate beams from 3300 to 31000 on average. The number of MUs was consequently reduced by an average of 28% compared with treatment planning without resampling. The beam reduction resulted, on average, in the use of 42% fewer beams at the cost of an increase in the number of MUs of only 5%. In 2 patients, the tumor was located within 2 cm of the spinal cord. The conformality constraint was sacrificed in the mMLC plans of these 2 patients to improve the sparing of the spinal cord. Moreover, the spinal cord constraint was relaxed to 8 Gy/fraction in one of these patients. The rib constraint was relaxed in the mMLC and Iris plans of 3 patients who had one or two ribs located partially within the PTV, because the PTV coverage was given the greatest priority. The dose distributions of the mMLC plan and MLD-equivalent Iris plan of patient 4 are given in figure 2.2.

The variation in the fraction durations of the mMLC plans was small, but they varied greatly for the Iris plans (figure 2.3). Figure 2.3 shows that the fraction duration increased

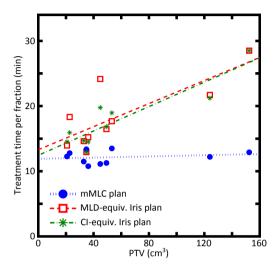


Figure 2.3: Estimated treatment time/fraction as function of planning target volume for all 10 patients. Lines indicate linear fits to data (mini-multileaf collimator, $R^2 = 0.04$, p = 0.6; mean lung dose-equivalent Iris, $R^2 = 0.62$, p = 0.007; conformity index-equivalent Iris, $R^2 = 0.81$, p < 0.001; t test, null-hypothesis, slope = 0).

with an increasing target volume when using circular fields, similar for MLD-equivalent Iris plans (p = 0.007) and CI-equivalent Iris plans (p < 0.001). When conformally shaped fields were applied, the treatment time for each fraction was not significantly influenced by the target volume (p = 0.6). Thus, the fraction duration of MLD-equivalent plans was, on average, 40% longer than that of the mMLC plans for tumors <80 cm³ and \leq 121% longer for larger targets. For 1 patient, the Iris plan (same plan as for the MLD-equivalent and CI-equivalent plans) was more time-efficient than the mMLC plan. For this patient, this could be explained by the fact that the Iris plan contained only 12 of the 16 node positions of the mMLC plan. Although Iris plan generation started with all the node positions of the mMLC plan, some node positions were not selected by the inverse optimization algorithm. This was also observed in 5 other patients, but to a much lesser extent.

In figure 2.4, the MLD is plotted as a function of the estimated fraction duration for the mMLC plan (with a minimized MLD) and all Iris plans (with various MLD constraints) of each patient. The solutions in the lower part of each Iris graph are Paretoefficient in terms of the MLD and fraction duration. Thus, a reduction in MLD was inevitably accompanied by an increase in the fraction duration and vice versa. How-

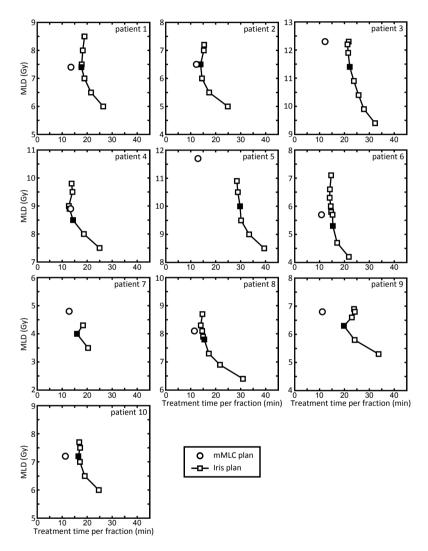


Figure 2.4: Mean lung dose (MLD) as function of estimated fraction duration of mini-multileaf collimator plan and Iris plans. Each marker represents single treatment plan. Solid black markers indicate preferred Iris plans selected by physician.

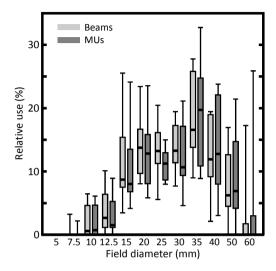


Figure 2.5: Box plots of relative use of field sizes in mean lung dose-equivalent Iris plans of all patients, measured in number of beams and monitor units delivered. Whiskers indicate extreme values.

ever, this Pareto-efficiency no longer holds for a greater MLD, because the solutions are restricted by different constraints in both regions of the graphs. The prescribed MLD is the limiting constraint in the lower part of the graphs, and the CI constraint is limiting in the upper part. The CI constraint also restricts the maximal MLD that can be obtained, which, in 2 patients, was even lower than the minimized MLD of the mMLC plan. The lowest MLD in all patients was obtained using the Iris collimator.

On average, nine collimator sizes were used in the MLD-equivalent Iris plans. Figure 2.5 shows the relative use of the collimator sizes, measured in the number of beams and in the number of MUs delivered. The smallest collimator size was never used, while beams with field diameters of 15–40 mm were present in every treatment plan. A collimator size of 35 mm was, on average, most frequently used. The four largest collimator sizes delivered more MU/beam than the smaller collimators. Figure 2.6 shows the targeting of beams in the MLD-equivalent Iris plan of patient 4. The large fields were typically aimed at the center of the PTV, and the smaller beams were used to deliver the dose at the PTV periphery. The corresponding fluence profiles of each node position are given in figure 2.7. Considerable fluence variation (i.e. intensity modulation) was observed within almost all fields delivered from each node position.

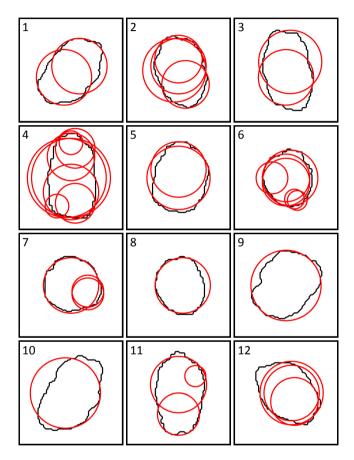


Figure 2.6: Targeting of beams from node positions in mean lung dose-equivalent Iris plan of patient 4. For each node position, beam's eye view projection of planning target volume given in black, and field contour (50% isodose) of each beam in red.

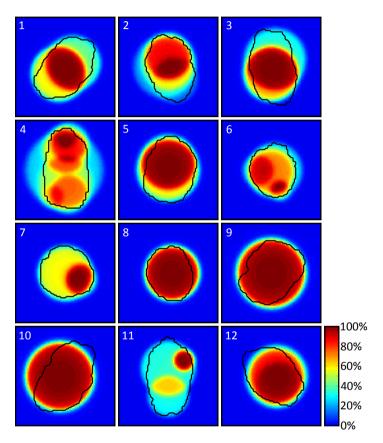


Figure 2.7: Delivered fluence corresponding to figure 2.6. Beam's eye view projection of planning target volume given in black. Fluences normalized to maximal fluence for each node position.

2.4 Discussion

The use of a mMLC on the CyberKnife was found to give time-efficient treatment plans in lung cancer patients. Iris plans generally had a greater treatment time/fraction, especially in large tumor volumes. Both the mMLC and Iris plans were much more time-efficient than the traditional fixed-collimator plans, which contain, for these patients, approximately 60 nodes positions, 125 beams, and 35000 MUs. The Iris plans are also likely to be more time-efficient than the two-collimator plans from the study by Pöll *et al.* (19) for the same patient group. The Iris plans contained an average of 47% fewer beams and 2% more MUs. Although the number of node positions in the two-collimator plans was not reported, it is evidently expected to be much greater than in the Iris plans. The average MLD and CI were also considerably lower in the Iris plans.

The mMLC plans and Iris plans were generated using two different planning algorithms. Therefore, our observations could be biased by differences in these algorithms. As both algorithms were developed in-house, the parts of the algorithms such as structure segmentation, heterogeneity correction, dose grids, and plan evaluation, were deliberately kept similar. However, some differences were inevitable, because the algorithms had to serve a different purpose (beam-angle optimization vs. inverse optimization for given node positions) and had to manage different field characteristics (isocentric conformal uniform beams vs. non-isocentric circular non-uniform beams). Another limitation of the present study was that a CyberKnife equipped with a mMLC and a flattening filter does not exist. The beam model, derived from a 6-MV gantry-based stereotactic linear accelerator, could therefore not be verified with a real machine. We furthermore assumed the use of an idealized mMLC. Instead of performing actual leaf segmentation, an arbitrarily orientated collimator grid was used with 2.5-mm resolution. Moreover, the mMLC was assumed to perfectly block the beam.

In the present study, we used flattened fields for the mMLC plans. Unflattened fields are becoming more common but require intensity modulation to cope with the nonuniform beam profiles, especially for larger tumors. The use of intensity-modulated radiotherapy (IMRT) on conventional gantries caused the delivery time to increase considerably compared with that for 3D-CRT. Because the aim of the present study was to improve the time-efficiency of the CyberKnife treatments, we benchmarked the Iris plans with flattened non-intensity-modulated fields. To account for a flattening filter, the beam-on time was increased by 20%. The estimated mMLC fraction duration was, however, rather insensitive to this assumption. If a dose rate loss of 30% was assumed instead, the average mMLC treatment time would increase by only 4%. The use of flattened and unflattened fields in relation to plan quality and treatment time is the subject of future research.

The mMLC plans were not reproduced exactly by the Iris plans. This was, for example, illustrated by the dose distributions in figure 2.2 and by the fact that not all node positions of the mMLC plans were selected in the Iris plans. Moreover, figure 2.7 shows that the fluence profiles were highly modulated in the Iris plans. In contrast, a uniform fluence was used in the mMLC plans. Thus, the dosimetric plan quality of the Iris plans could be superior to that of the mMLC plans. However, it raises the question of whether the node positions selected by Cycle were optimal for use in the Iris plans. De Pooter et al. (27) used a similar approach for liver tumors, in which the selection of beam orientations by Cycle was followed by inverse IMRT optimization. The two-step optimization of beam orientations and IMRT profiles was, in some cases, found to result in suboptimal treatment plans. It is, therefore, likely that the quality of the Iris plans can be improved further if the beam orientation selection is integrated into the inverse planning algorithm. Simultaneous orientation and profile optimization for IMRT and CyberKnife are currently being developed at our institute.

2.5 Conclusion

Non-coplanar robotic stereotactic radiotherapy using a mMLC was found to give very time-efficient treatment plans in lung cancer patients. For small target volumes (<80 cm³), the Iris variable-aperture collimator is a time-efficient alternative to a mMLC. The fraction duration of these Iris plans (MLD-equivalent) was, on average, 40% longer than mMLC plans, with similar or better plan quality. For larger tumors, the treatment time of the Iris plans was ≤121% longer.

Acknowledgments

The authors would like to thank Warren Kilby for his contribution to our treatment time estimator and general discussions.

chapter 3

Shortening treatment time in robotic radiosurgery using a novel node reduction technique

Medical Physics, Vol. 38, No. 3, pp. 1397-1405, 2011.

Steven van de Water, Mischa Hoogeman, Sebastiaan Breedveld, Ben Heijmen.

Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

Abstract

Purpose: The fraction duration of robotic radiosurgery treatments can be reduced by generating more time-efficient treatment plans with a reduced number of node positions, beams, and monitor units (MUs). Node positions are preprogramed locations where the robot can position the focal spot of the x-ray beam. As the time needed for the robot to travel between node positions takes up a large part of the treatment time, the aim of this study was to develop and evaluate a node reduction technique in order to reduce the treatment time per fraction for robotic radiosurgery.

Methods: Node reduction was integrated into the inverse planning algorithm, developed in-house for the robotic radiosurgery modality. It involved repeated inverse optimization, each iteration excluding low-contribution node positions from the planning and resampling new candidate beams from the remaining node positions. Node reduction was performed until the exclusion of a single node position caused a constraint violation, after which the shortest treatment plan was selected retrospectively. Treatment plans were generated with and without node reduction for two lung cases of different complexity, one oropharyngeal case and one prostate case. Plan quality was assessed using the number of node positions, beams and MUs, and the estimated treatment time per fraction. All treatment plans had to fulfill all clinical dose constraints. Extra constraints were added to maintain the low-dose conformality and restrict skin doses during node reduction.

Results: Node reduction resulted in 12 residual node positions, on average (reduction by 77%), at the cost of an increase in the number of beams and total MUs of 28% and 9%, respectively. Overall fraction durations (excluding patient setup) were shortened by 25% (range of 18%–40%) on average. Dose distributions changed only little and dose in low-dose regions was effectively restricted by the additional constraints.

Conclusions: The fraction duration of robotic radiosurgery treatments can be reduced considerably by node reduction with minimal changes in dosimetrical plan quality. Additional constraints are required to guarantee low-dose conformality and to avoid unacceptable skin dose.

Introduction

Design of time-efficient treatment plans is attractive for reduction of the relatively long treatment times of the CyberKnife robotic radiosurgery system (RRS) (Accuray Inc., Sunnyvale, CA, USA) (10, 19). Time-efficient treatment plans can be created by exploiting the degeneracy of the planning problem, i.e. different beam setups can result in similar dose distributions (35, 36). This degeneracy can be used to reoptimize a treatment plan such that the plan requires less time to be delivered, but without compromising dosimetrical plan quality.

Treatment delivery using the RRS is schematically illustrated in figure 3.1. At the start of the treatment, the robotic manipulator moves from its rest position to the first node position. Node positions are preprogramed locations where the robot can position the focal spot of the x-ray beam. By adjusting the linac orientation and the collimator size, multiple beams can be delivered from a single node position. The time needed for linac reorientation between beams delivered from a single node position is much shorter than the robot motion time between node positions. The robot moves on to the next position,

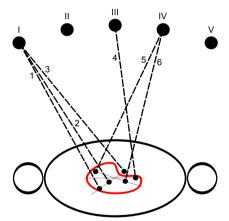


Figure 3.1: Schematic illustration of the RRS delivery sequence. The cross section of a patient (trunk and arms) is depicted with the target volume located in the center. The node positions are indicated by Roman numerals. Multiple beams (beam axes indicated by dashed lines, drawn from node position to target point) can be delivered from each node position. Beam numbers indicate the order in which the beams are delivered. The treatment starts at the first node position and when all beams of a node position are delivered, the robot moves on to the next node position. Node positions can be skipped if there are no beams to be delivered. Note that in reality the node position distribution is non-coplanar.

when all beams from a node position have been delivered. In between beam delivery, orthogonal kV images are acquired for position verification and correction.

To deliver a 20 Gy fraction to a lung tumor, ~45% of the fraction duration (excluding patient setup) is spent on robot motion between node positions, ~35% on beamon time, ~10% on linac reorientation between beams, and ~10% on image acquisition. These numbers will differ for different treatment sites, protocols, fractionation schemes, complexity of cases, size of cases, and versions of the RRS, but it illustrates that the fraction duration is greatly influenced by the number of node positions. Reducing the number of node positions in a treatment plan is therefore essential in order to minimize the treatment time per fraction. Previously, our group proposed a two-step process to generate time-efficient RRS treatment plans with a low number of node positions (37). In the first step, a beam-angle optimization algorithm developed in-house (Cycle) (38, 39) was used to establish a small number of favorable node positions for a virtual RRS plan using conformal uniform fields. In the second step, inverse planning was performed using circular RRS beams with varying orientation and diameter from these favorable node positions. Because such a sequential approach is likely to result in suboptimal solutions, simultaneous node reduction integrated into the inverse planning algorithm is to be preferred (27).

In this study, we developed and evaluated such an algorithm based on iteratively reducing an initially large number of node positions to a few node positions, while maintaining plan quality. No concessions are made on the imposed dose prescriptions and the iterative node reduction stops if further reduction would result in a constraint violation. The treatment plan with the lowest treatment time per fraction is selected retrospectively, which is not necessarily the plan with the lowest number of node positions. The effect of node reduction on the number of node positions, beams and MUs, and more importantly on the fraction duration was studied for different sites (simple lung, complex lung, headand-neck, and prostate) by comparing treatment plans with and without node reduction.

3.7 Methods and materials

Patient group and dose prescriptions

The developed algorithm for plan optimization with integrated node reduction was evaluated using treatment data and CT scans of four patients previously treated at our institute: Two non-small-cell lung cancer patients of different complexity, one oropharyngeal cancer patient (tonsillar fossa carcinoma), and one prostate cancer patient. The gross target volume to planning target volume (PTV) margin was 5 mm for the lung cases and 3 mm for the oropharyngeal case and prostate case. The resulting PTVs were 20.8 cm³ (simple lung), 152.3 cm³ (complex lung), 134.8 cm³ (oropharynx) and 62.3 cm³ (prostate). The complex lung case mainly differed from the simple lung case in the fact that the PTV was located within 2 cm of the spinal cord, requiring a very steep dose gradient in between these structures. The prostate case received a virtual high-doserate brachytherapy treatment, implying a heterogeneous target dose distribution and sparing of the urethra (40). The seminal vesicles were not included in the PTV. For the oropharyngeal case, the RRS was used for a boost irradiation of the primary tumor, after having delivered 46 Gy (in 2 Gy fractions) to the primary tumor and elective nodes using intensity-modulated radiotherapy (IMRT). The planning constraints for each treatment site are given in table 3.1. To ensure high-dose conformality, a conformality index of 1.2 or lower was required for all patients. The conformality index was defined as the volume receiving the prescription dose divided by the PTV. This constraint was enforced by a constraint on the lung dose in the lung tumor cases or by a maximum-dose constraint on a shell structure around the PTV at 15 mm distance in the oropharyngeal and prostate cases.

Since reduction in the number of node positions might result in increased skin doses or degraded low-dose conformality, structures to constrain these parameters were added. The skin structure was generated by randomly sampling 3000 dose calculation points at 10 mm distance from the skin in regions where beams could enter the patient. As the PTV was located close to the skin in the oropharyngeal case, the skin structure was not generated in the vicinity (<25 mm) of the PTV to avoid underdosage of the target. Low-dose conformality was achieved using a shell structure around the PTV at 25 mm distance. The maximum tolerated doses to the skin and shell structure are also listed in table 3.1.

Robotic radiosurgery system characteristics

A standard node set for extracranial treatments was used in this study. The node set contained 117 node positions distributed non-coplanar around the imaging center at 800-1000 mm distance. The robotic manipulator of the RRS travels along a semifixed path that connects all the available node positions in a specific order. If there are no beams to be delivered from a certain node position, it can be skipped and the robot immediately travels to the next node position. However, some node positions cannot be skipped to ensure safe traveling of the robot and to avoid the cables becoming entangled.

The 6-MV photon beam was assumed to be circularly shaped by the Iris™variableaperture collimator (Accuray Inc.) (25). The Iris collimator can automatically adjust the

Structure Constraint Lung (3 \times 20 Gy at 80% isodose): $D_{95\%} > 60 \text{ Gy}$ High-dose conformality index (60 Gy) < 1.2 Low-dose conformality (shell at 25 mm) $D_{max} < 25 Gy$ Skin $D_{\text{max}} < 18 \text{ Gy}$ Spinal Cord $D_{max} < 18 Gy$ Luna $V_{20GV} < 10\%$ Oropharynx (3 \times 5.5 Gy at 80% isodose): D_{95%} > 16.5 Gy High-dose conformality index (16.5 Gy) < 1.2 Low-dose conformality (shell at 25 mm) $D_{max} < 7 Gy$ Skin $D_{max} < 6.5 Gy$ Spinal Cord $D_{\text{max}} < 10 \text{ Gy}$ Ipsilateral parotid gland $D_{max} < 15 Gy$ Contralateral parotid gland $D_{max} < 2 Gy$ $D_{\text{max}} < 12 \text{ Gy}$ Contralateral submandibular gland $D_{max} < 8 Gy$ Prostate (4 \times 9.5 Gy at 60% isodose): PTV $D_{95\%} > 38 \text{ Gy}$ D_{mean} > 50 Gy High-dose conformality index (38 Gy) Low-dose conformality (shell at 25 mm) $D_{max} < 18 Gy$ $D_{max} < 18 Gy$ Bladder $D_{max} < 42 Gy$ Bladder $D_{1cc} < 38 \text{ Gy}$ Rectum wall $D_{max} < 38 \text{ Gy}$ Rectum wall $D_{1cc} < 33 \text{ Gy}$ $D_{max} < 28.5 \, \text{Gy}$ Rectal mucosa D_{5%} < 45.5 Gy Urethra Urethra $D_{10\%} < 42 \text{ Gy}$ $D_{50\%} < 40 \text{ Gy}$ Urethra

Table 3.1: Planning constraints considered for each treatment site.

field diameter of each beam to 5, 7.5, 10, 12.5, 15, 20, 25, 30, 35, 40, 50, or 60 mm (12 field sizes), defined at 800 mm from the focal spot. The linac was calibrated such that 1 MU corresponded to a dose of 1cGy in a reference point located on the central axis of a 60 mm field, at 800 mm from the focal spot, and at 15 mm depth in water.

Inverse planning with automated node reduction

An inverse planning algorithm for the RRS was developed to generate treatment plans with short treatment times, largely achieved by reduction in the used node positions. Figure 3.2 shows a flow chart of the algorithm. In this section, we will first briefly explain the main components of the inverse planning algorithm with automated node reduction.

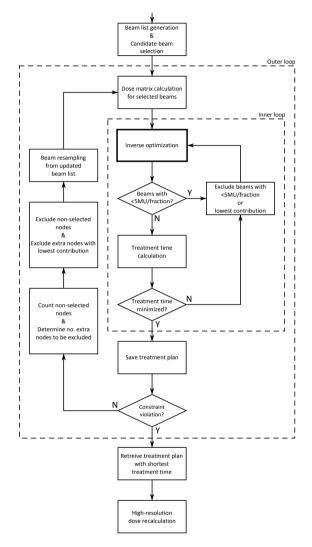


Figure 3.2: Flow chart of the inverse planning algorithm with node reduction, developed in-house for the RRS. See text for a detailed explanation.

Detailed information on the various subcomponents will be provided in subsequent sections and in Appendixes 3.A-3.C.

The treatment planning process started with the generation of the so-called 'beam list' containing 1.6-5.9 million candidate beams that could be used for treatment planning (see Appendix 3.A). An initial set of 5000 candidate beams was randomly selected from this list. For each selected candidate beam, dose calculations were performed to construct a dose matrix for each structure of interest (see Appendix 3.B). These dose matrices served, together with the dose prescriptions, as input for the inverse optimization.

The inverse optimization was embedded in two iterative loops, the inner 'inverse optimization loop' and the outer 'node reduction and beam resampling loop'. In the inner loop, a feasible treatment plan was created, followed by the exclusion of low-MU beams and the calculation of the treatment time (see section 'Inner loop: Inverse optimization'). This loop was repeated as long as the treatment time decreased with the exclusion of low-MU beams.

The outer loop, i.e. the 'node reduction and beam resampling loop', aimed at shortening the treatment time by the iterative reduction in the number of used node positions (see section 'Outer loop: Node reduction and beam resampling'). At each iteration, node positions that contributed the least to the plan were deleted from the treatment plan and from the beam list (node reduction) and new candidate beams were selected from this reduced beam list (beam resampling). These newly selected candidate beams were added to the beams selected in the previous treatment plan, complementing them to again 5000 beams, after which dose calculation and inverse optimization were repeated. The treatment plan was stored after each iteration. The outer loop was performed as long as a feasible plan could be obtained that fulfilled the imposed constraints. Finally, the stored treatment plan with the shortest treatment time was selected retrospectively as the final treatment plan, for which a high-resolution $(1 \times 1 \times 1.5 \text{ mm}^3)$ dose calculation in the entire patient was performed.

Inner loop: Inverse optimization

The inner loop used the inverse multi-criteria optimization routine from the treatment planning system 'YARTOS' (indicated by the 'inverse optimization' block in figure 3.2), which was developed in-house by Breedveld et al. (32). The rest of the planning infrastructure (inner and outer loop) was developed exclusively for the generation of timeefficient RRS treatment plans. Using dose matrices and dose prescriptions as input, the inverse optimization algorithm constructs a treatment plan by determining the optimal weights of the candidate beams. It makes a distinction in dose prescriptions between constraints and objectives. Constraints are prescriptions that have to be fulfilled, otherwise the solution is invalid. Objectives are prescriptions for which the optimal value (minimum or maximum) is searched in a prioritized order without violating the constraints (41). In this study, however, minimization of the total number of MUs was the only objective. The constraints that were used for each patient are listed in table 3.1. The deliverability and the time-efficiency of the treatment plan were improved by repeating the inverse optimization several times, each time excluding all beams with a weight of less than 5 MU/fraction or excluding beams with the lowest dose contribution to the PTV. A minimum beam weight of 5 MU was required to guarantee the linearity of the dose delivery and the low-contribution beams were excluded until the shortest treatment time per fraction was obtained (see Appendix 3.C for treatment time calculation). The contribution of a beam was defined as the dose delivered to the isocenter if the beam would have been targeted at the isocenter. Of the 5000 candidate beams used, typically 50–200 beams were selected in the optimized treatment plan.

Outer loop: Node reduction and beam resampling

Removal of node positions from the treatment plan and beam list was performed in two steps. First, node positions without selected beams in the optimized treatment plan were excluded (unforced exclusion). Second, node positions that contributed the least to the PTV dose were excluded (forced exclusion). The number of node positions to be excluded per iteration (reduction goal) was defined such that a plan with 15 node positions should be achieved after ten outer loop iterations, assuming a linear reduction in the node positions with the number of iterations. The number of unforced excluded node positions could vary between iterations. If the number of unforced exclusions was less than the reduction goal, then forced exclusion of node positions was performed to complement the node reduction. If the number of unforced excluded node positions was more than required, then these node positions were removed from the beam list anyway, but the reduction goal was lowered for the subsequent iterations. When a plan with at most 15 node positions had been obtained, the node reduction goal was kept fixed while node reduction continued. Toward the end of the node reduction process (<20 node positions in the treatment plan), the number of node positions excluded in one iteration was considered too large when a feasible solution could not be found or the fraction duration increased. In that case, the last exclusion was undone and the iteration was repeated with the reduction goal reduced by half. However, at least one node position had to be excluded per iteration. The node reduction and beam resampling loop was exited when a feasible plan could not be found after the exclusion of a single node position.

Beam resampling is a method to improve the quality of a treatment plan by iteratively increasing the total number of used candidate beams (33). It makes use of the fact that the majority of the 5000 candidate beams is not selected in the inverse optimization loop or is removed by node reduction and therefore can be replaced by newly selected candidate beams, randomly sampled from the updated beam list. These new candidate beams were added to the beams selected in the previously optimized treatment plan. The inverse

optimization was always performed with a total of 5000 candidate beams, so the number of newly selected candidate beams equaled 5000 minus the number of beams selected in the previous treatment plan.

The efficiency of the loop was improved by two measures. First, dose constraints that were far from being violated were switched off to speed up the inverse optimization. If the dose in a structure was approaching the maximum tolerable dose in subsequent iterations, the constraint was turned on again. Second, beams that were considered unlikely to contribute to the final treatment plan were excluded from the beam list. This included previously used candidate beams and beams with a collimator size smaller than the smallest collimator size selected in the previously optimized treatment plan. Moreover, beams with an area outside the PTV projection larger than the maximum area outside the PTV projection of the beams in the initial treatment plan were excluded as well.

Evaluation

A comparison was made for each patient between treatment plans generated with and without node reduction. For this purpose, the algorithm was adjusted so that beam resampling could also be performed without excluding any node position from the beam list during treatment planning. For each patient, the plan with node reduction was generated first and the plan without node reduction was generated subsequently using the same number of iterations as was required to obtain the shortest node reduction plan. In this way, the total number of candidate beams used to generate both treatment plans was approximately equal, allowing for a fair comparison. The no node reduction plan with the shortest fraction duration was selected in retrospect for evaluation. The quality of the treatment plans was assessed using the number of node positions, beams and MUs, and the estimated treatment time per fraction. A visual inspection of the dose distribution was performed to ensure the absence of any undesirably high doses in unspecified tissues.

3.3 Results

The results of the final treatment plans generated with and without node reduction are given in table 3.2 for all four patients. All treatment plans fulfilled the clinical dose constraints as listed in table 3.1 and were approved after visual inspection. Node reduction causes the number of node positions to be reduced by 77%, on average, compared to no

Table 3.2: Results for the treatment plans generated without and with node reduction for the four patient cases. The percentages between the parentheses indicate the relative difference between node reduction and no node reduction plans.

	Simple lung	Complex lung	Oropharynx	Prostate
No node reduction:				
Node positions	36	66	52	53
Beams	47	121	100	112
MUs	17887	32856	8333	50518
Fraction duration (min)	19.3	36.3	22.1	34.1
Node reduction:				
Node positions	10 (-72%)	15 (-77%)	8 (-85%)	14 (-74%)
Beams	59 (+26%)	158 (+31%)	123 (+23%)	150 (+34%)
MUs	19983 (+12%)	33194 (+1%)	9820 (+18%)	53352 (+6%)
Fraction duration (min)	15.9 (-18%)	27.1 (-25%)	13.2 (-40%)	27.9 (-18%)

node reduction, while increasing the number of beams (+28%) and MUs (+9%). As a result, the fraction duration is, on average, 25% shorter in the treatment plans generated with node reduction. The amount of treatment time reduction differs greatly between cases (range of 18%-40%) and is not clearly related to the complexity of the case, as, for example, the lowest reduction is found for the simple lung case. The total treatment time reduction from the very first optimization run to the final treatment plan is, on average, 54% (range of 47%-59%) for node reduction and 38% (range of 32%-43%) for no node reduction.

Figure 3.3 shows for all cases the dose distributions of the treatment plans generated with and without node reduction. Doses are very similar in both plans and differences can mainly be observed in the low-dose regions. Dose-volume histograms for the prostate case are shown in figure 3.4. The target coverage of both plans is nearly identical and small differences can be observed for the organs-at-risk.

The number of node positions as a function of the outer loop iteration is depicted in figure 3.5 for the reduction and no-reduction plans of all four cases. Each marker represents a treatment plan that was saved during the planning process. The treatment planning algorithm required 11 or 13 iterations to obtain the treatment plan with the shortest fraction duration. The number of node positions could be reduced further to a minimum of 6-11 without violation of the dose constraints. Interestingly, the figure shows that beam resampling without explicit node reduction already leads to a considerable reduction in the number of node positions.

The treatment time per fraction as a function of the iteration is given in figure 3.6. The figure illustrates that after an initial decrease in the treatment time, the treatment time eventually increases when using node reduction due to an increase in the number of beams and MUs.

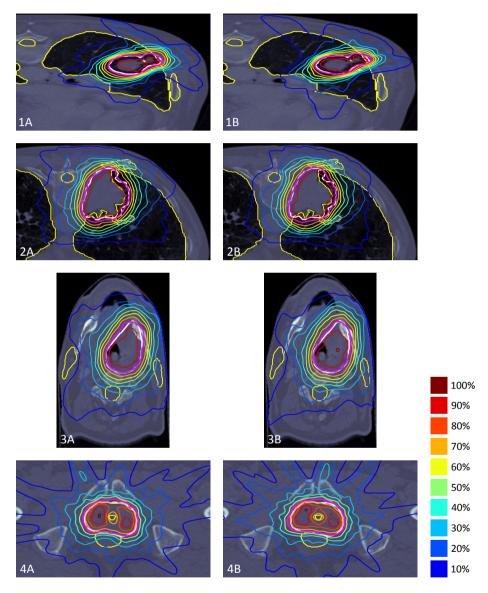


Figure 3.3: Dose distributions of the treatment plans generated without node reduction (a) and with node reduction (b) for the simple lung case (1), complex lung case (2), oropharyngeal case (3) and prostate case (4). The PTV contour is given by the thick white line. Depicted isodose lines range from 10% to 100% of the maximum PTV dose. The prescription isodose line is indicated by the thick isodose line (80% for lung cases and oropharyngeal case, 60% for prostate case).

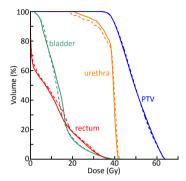


Figure 3.4: Dose-volume histograms for the treatment plans of the prostate case generated without node reduction (solid lines) and with node reduction (dashed lines).

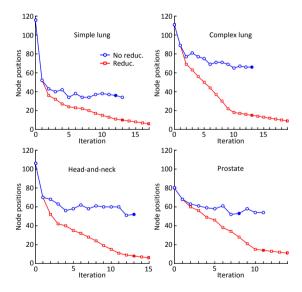


Figure 3.5: The number of node positions as a function of the iteration during treatment planning for the four patient cases, with node reduction (squares) or without node reduction (circles). Each marker represents a stored treatment plan. At iteration 0, the number of available node positions is indicated. This number is the maximum number of node positions (117) minus the disabled node positions. Solid markers indicate the treatment plans with the shortest fraction duration (see figure 3.6), which were selected for evaluation.

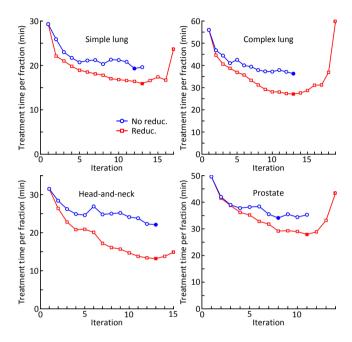


Figure 3.6: Treatment time per fraction as a function of the iteration during treatment planning for the four patient cases, with node reduction (squares) or without node reduction (circles). Each marker represents a stored treatment plan. Solid markers indicate the treatment plans with the shortest fraction duration, which were selected for evaluation.

3.4 Discussion

This study showed that the number of node positions in RRS treatment plans could be reduced greatly without violating imposed dose constraints by using the proposed node reduction technique. As a consequence, the treatment time per fraction required to deliver the treatment plans shortened considerably. Not only node reduction, but also beam resampling by itself and the exclusion of low-contribution beams, helped to shorten fraction duration. It should be noted that the treatment time could not optimized directly using our inverse optimization algorithm. The treatment time was only indirectly optimized through a reduction the number of node positions, beams, and MUs. As a consequence, the optimality of the obtained treatment times is not guaranteed. The treatment times are only minimal for the reduction schemes used in this study (see section 'Outer loop: Node reduction and beam resampling'). The heuristics in this study were mainly applied to speed up the planning process. Omitting these heuristic rules would require more planning iterations and consequently increase the planning time, but might result in shorter fraction durations as candidate beams are more intensely sampled. Note that for this study, only one set of heuristic rules was used irrespective of the treatment site. The approximated time needed to generate a treatment plan with node reduction ranged from 2.0 to 5.6 h on a modern 8-core computer, mainly depending on the patient case (number of constraints and voxels considered in the inverse optimization) and on the number of iterations. The code was written in MATLAB and C and was run using MAT-LAB 7 (The MathWorks, Inc., Natick, MD). The number of iterations can be reduced by using a different stopping criterion for the outer loop, instead of continuing until the minimum number of node positions has been obtained. Terminating node reduction when the exclusion of a single node position caused the treatment time to increase would in this study have resulted in approximately 15% less iterations, without losing the gain in treatment time.

Intuitively, a reduction in the number of node positions is associated with degrading plan quality. A gain in fraction duration is often believed to come at a price in terms of organ sparing. However, because of the degeneracy of the optimization problem, this is not necessarily true. Of course, node reduction does alter the dose distribution in some regions, but almost solely in regions in which no dose constraints are imposed. Still, one should be cautious that doses can change undesirably. Skin dose and low-dose conformality, for example, can be acceptable without the use of specific constraints in treatment plans without node reduction but can deteriorate when applying node reduction. This can, however, be solved easily by adding constraints in these regions of concern. Despite the extra constraints, we experienced that sufficient freedom was left for the algorithm to reduce the number of node positions by altering the dose in other, unconstrained regions. So, node reduction does not necessarily lead to degraded plan quality, but it might be required to further constrain the optimization problem. Moreover, a reduction in the fraction duration is by itself an improvement of the plan quality. Not only does it improve patient comfort and patient throughput, it may also improve position accuracy during a treatment fraction (28) and increase the biological effectiveness of the treatment (42, 43).

In previous work, comparing treatment plans for a theoretical mini-multileaf collimator and the Iris collimator, we used a two-step approach to reduce the number of node positions (37). Favorable node positions were selected first by beam-angle optimization using conformal uniform fields, followed by inverse optimization for the RRS only using those node positions. As was also indicated by de Pooter et al. (27), an approach like this is likely to result in suboptimal results, as the fields used for selection of node positions differ much from the fields that are actually delivered. The novel node reduction technique presented in the current study, in which the selection of favorable

node position is integrated in the inverse optimization process, is therefore expected to result in a more optimal set of node positions and thus in shorter treatment times. This can be illustrated by comparing the treatment times obtained by both approaches for the two lung cases, which were also included in our previous work. The integrated method reduced treatment times by 20%, on average, compared to the two-step approach. Besides the shorter treatment times, the integrated node reduction technique also provides a more practical implementation of node reduction compared to the implementation that uses two distinct algorithms.

A method to reduce the treatment time per fraction in robotic radiosurgery was also recently developed and investigated by Dooley et al. (44), representing the vendor of the CyberKnife (Accuray Inc.). In their investigation, low-utility node positions and beams were also iteratively removed from the treatment plan while resampling new candidate beams. In addition to differences in the reduction and resampling heuristics, Dooley et al. used a preselected set of 23 or 38 node positions to obtain a node-reduced plan instead of the full set of 117 node positions. Another major difference lies in the inverse optimization routine. The vendor's planning system (Multiplan v4.0, Accuray Inc.) uses a 'sequential optimization' algorithm, which is not capable of imposing a minimum dose constraint on the PTV (45). As a result, target coverage and dose homogeneity are objectives, which tended to be compromised by the treatment time reduction. The user can set the desired fraction duration for a treatment plan, but has to check whether the plan still fulfills the imposed dose prescriptions. In our method, all dose parameters can be constrained and are not compromised by the node reduction. For five lung cases planned according to a similar protocol as used in our study, Dooley et al. found a reduction in treatment time of 66%, on average, compared to a total treatment time reduction of 52% (simple lung) and 57% (complex lung) using our method. The results of Dooley et al. are, however, difficult to compare as the patient database was different except for one patient. For this patient (simple lung case) the data were provided to the vendor by our group. To make a rudimentary comparison, we recalculated our simple lung case treatment time for their CyberKnife characteristics. Dooley et al. obtained a slightly shorter fraction duration of 11.5 min vs 12.2 min, but this was accompanied by poorer plan quality when measured by plan parameters such as the high-dose conformality index (1.24 vs 1.09), the maximum spinal cord dose (4.1 Gy vs 0.22 Gy), and the lung volume receiving 20 Gy or more (4.3% vs 2.9%). It would be interesting to compare the performance of both algorithms in terms of treatment time reduction and planning time using equally constrained dose prescriptions for both algorithms in order to uphold plan quality throughout the reduction process.

The node reduction technique is basically a beam-angle optimization technique. From a set of candidate node positions, a small set of node positions is selected to be used in the treatment. Several beam-angle optimization algorithms for IMRT using conventional radiotherapy modalities can be found in literature. For example, Llacer et al. (46) proposed a method that selects, using a full combinatorial analysis of automatically preselected groups of beam directions, favorable beam orientations based on the geometrical PTV coverage and OAR avoidance. The dosimetrical method reported by Das et al. (47) selects a number of individually optimal beam directions first, after which a more optimal set of beam orientations is obtained by substituting the selected beams one-by-one with beams from the angular vicinity until no further improvements can be found. Rowbottom et al. (48) combined a fast simulated annealing (FSA) algorithm and a downhill simplex algorithm to optimize the beam arrangement and the fluence profiles, respectively, starting from a randomly selected set of individually acceptable beams and sampling new gantry and couch angles for each beam in each FSA iteration. Woudstra et al. reported on a 'sequential' selection method. Treatment planning starts with an empty plan and one-by-one favorable beam orientations are added to the plan, taking into account the dose delivered by previously added beams (39). A similar approach was used by Meedt et al. (35), in which the sequential selection is followed by routines that aim at replacing redundant beams and avoiding local minima. Breedveld et al. (49) extended the sequential selection approach to perform a prioritized multi-criteria optimization. The node reduction technique developed in this study works differently from the methods mentioned above. Instead of selecting the desired number of favorable beam directions, it starts treatment planning with all candidate beam directions (node positions) and iteratively excludes the least favorable beam directions. Irrespective of these two approaches, nearly all beam-angle optimization algorithms suffer from the fact that the selection of a beam direction is based on an intermediate dose distribution and not on the final dose distribution. A beam direction that is added or excluded at a certain moment in the planning process could become redundant or favorable in the final treatment plan. It would be interesting to investigate in which of the two approaches (including vs excluding) the treatment plan is affected most by this issue and which approach will give the best planning solution. For this purpose, the node reduction technique could be adapted to beam-angle optimization for gantry-based IMRT. A method analogous to resampling could be implemented by varying the collimator grid resolution throughout the planning iterations. Treatment planning starts with all available candidate beam directions, using a few beamlets per beam direction in order to avoid the problem becoming too large. As the planning proceeds and beam directions are excluded in subsequent iterations, increasingly finer collimator grids with more beamlets per beam direction can be used in the optimization.

3.5 Conclusion

This study showed that the number of node positions in a CyberKnife treatment plan could be reduced by 77%, on average, without compromising the plan quality for the four clinical cases considered. No concessions were made on the imposed dose prescriptions compared to the no node reduction plans. The node reduction technique resulted in an average reduction in the treatment time per fraction of 25%. The use of node reduction requires additional constraints to avoid dose violations in unspecified tissues.

chapter 4

Intrafraction prostate translations and rotations during hypofractionated robotic radiation surgery: Dosimetric impact of correction strategies and margins

International Journal of Radiation Oncology Biology Physics, Vol. 88, No. 5, pp. 1154–1160, 2014

> Steven van de Water¹, Lorella Valli^{1,2}, Shafak Aluwini¹, Nico Lanconelli², Ben Heijmen¹, Mischa Hoogeman¹.

Department of Radiation Oncology,
 Erasmus MC Cancer Institute, Rotterdam, The Netherlands
 Department of Physics and Astronomy,
 Alma Mater Studiorum, Bologna University, Bologna, Italy

Steven van de Water and Lorella Valli contributed equally to this work.

Abstract

Purpose: To investigate the dosimetric impact of intrafraction prostate motion and the effect of robot correction strategies for hypofractionated CyberKnife treatments with a simultaneously integrated boost.

Methods and materials: A total of 548 real-time prostate motion tracks from 17 patients were available for dosimetric simulations of CyberKnife treatments, in which various correction strategies were included. Fixed time intervals between imaging/correction (15, 60, 180, and 360 s) were simulated, as well as adaptive timing (i.e. the time interval reduced from 60 to 15 s in case prostate motion exceeded 3 mm or 2° in consecutive images). The simulated extent of robot corrections was also varied: no corrections, translational corrections only, and translational corrections combined with rotational corrections up to 5°, 10°, and perfect rotational correction. The correction strategies were evaluated for treatment plans with a 0-mm or 3-mm margin around the clinical target volume (CTV). We recorded CTV coverage (V_{100%}) and dose-volume parameters of the peripheral zone (boost), rectum, bladder, and urethra.

Results: Planned dose parameters were increasingly preserved with larger extents of robot corrections. A time interval between corrections of 60 to 180 s provided optimal preservation of CTV coverage. To achieve 98% CTV coverage in 98% of the treatments, translational and rotational corrections up to 10° were required for the 0-mm margin plans, whereas translational and rotational corrections up to 5° were required for the 3-mm margin plans. Rectum and bladder were spared considerably better in the 0-mm margin plans. Adaptive timing did not improve the delivered dose.

Conclusions: Intrafraction prostate motion substantially affected the delivered dose but was compensated for effectively by robot corrections using a time interval of 60 to 180 s. A o-mm margin required larger extents of additional rotational corrections than a 3-mm margin but resulted in lower doses to rectum and bladder.

4.1 Introduction

The mobility of the prostate has long since been recognized as a potential source of error during radiation therapy for prostate cancer patients. Several studies have described the characteristics of prostate motion between treatment fractions (50-52). More recently, with the advancement of prostate monitoring techniques, studies have investigated prostate motion during treatment fractions (20, 21, 53-57). For conventional treatment regimens (i.e. homogeneous target doses in 2 Gy fractions), the effect of intrafraction prostate motion has been addressed in previous studies (58, 59). Moreover, a technique to perform intrafraction corrections has been implemented clinically by Mutanga et al. (60).

The low α/β ratio of prostate cancer has triggered a growing interest in the application of hypofractionated radiation therapy for prostate cancer. One of the most extreme hypofractionation regimens currently applied aims at delivering 4 fractions of 9.5 Gy using the CyberKnife robotic radiosurgery system (Accuray, Inc, Sunnyvale, CA), emulating high dose rate (HDR) brachytherapy (15, 40, 61). The use of extreme hypofractionation might increase the impact of intrafraction motion, due to longer treatment fraction time and reduced statistical averaging of position errors. Moreover, the use of a simultaneously integrated boost (SIB) to deliver a higher dose to the peripheral zone of the prostate could increase the susceptibility to uncorrected rotations. The effect of intrafraction prostate translations and rotations on these hypofractionated treatments with a SIB has, to our knowledge, not been investigated before.

To correct intrafraction prostate motion, the CyberKnife is equipped with an automated marker tracking system (11). The system acquires orthogonal kV images with a user-defined interval and automatically adjusts the position and orientation of the robotic manipulator or treatment couch during treatment fractions. The system also features an adaptive timing mode in which the time interval between corrections is reduced when large displacements are observed. To account for residual motion between corrections, a planning target volume (PTV) is constructed in our clinic by applying a 3-mm margin around the clinical target volume (CTV) (15, 61). The relevance of intrafraction corrections has been reported previously for other hypofractionated treatments (62-65), but the effectiveness of these measures to account for intrafraction prostate motion during CyberKnife treatments has not been investigated before. In relation to this, it has also not been investigated before whether more frequent and more accurate robot corrections would allow for a reduction of the PTV margin.

The aim of this study was to quantify the dosimetric impact of the timing and extent of translational and rotational corrections during hypofractionated CyberKnife treatments and of the applied CTV-to-PTV margins. For this purpose, we performed

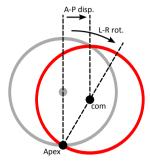


Figure 4.1: Schematic illustration showing the conversion of center-of-mass (com) displacements (disp.) in anterior-posterior (A-P) direction into rotations (rot.) around the left-right (L-R) axis.

treatment simulations using different time intervals between corrections, using various degrees of robot corrections and CTV-to-PTV margins of o and 3 mm.

4 7 Methods and materials

Prostate motion data

Real-time prostate motion data was collected by using an electromagnetic tracking device (Calypso Medical Technologies, provided by the Department of Radiation Oncology of the M.D. Anderson Cancer Center [Orlando, FL]) (20, 21). The device measured the position of implanted transponders with a frequency of 10 Hz. A total of 548 prostate motion tracks of 17 patients (denoted by 'motion data patients') were at our disposal. The average length of the motion tracks was 10.1 \pm 2.0 min (1 SD). Prostate rotations were not recorded by the tracking device but should also be taken into account to adequately determine the impact of intrafraction motion on the delivered dose (53). Several studies of interfraction prostate motion showed that rotations around the left-right axis are predominant and that prostate motion can be approximated in the sagittal plane by rotations around the apex (50-52). Based on these studies, we derived left-right rotations by converting the anterior-posterior translations into rotations around the prostate apex as shown in figure 4.1. Rotations around the anterior-posterior and superior-inferior axes are generally small and were not taken into account in this study.

Constraints			
	Structure	Туре	Limit
	PTV	Minimum	0.99 × 38 Gy
	PTV	Maximum	1.50 × 38 Gy
	CTV	Minimum	1.02 × 38 Gy
	Peripheral zone	Minimum	1.19 × 38 Gy
	PTV-ring 10 mm	Maximum	30 Gy
	PTV-ring 20 mm	Maximum	19 Gy
	PTV-ring 30 mm	Maximum	17 Gy
	Skin	Maximum	17 Gy
	Urethra	Maximum	41 Gy
	Penis - scrotum	Maximum	4 Gy
Objectives			
Priority	Structure	Type	Goal
1	Rectum	gEUD-12	1 Gy
2	Bladder	gEUD-12	1 Gy
3	Rectal mucosa	gEUD-12	1 Gy
4	Rectum	Mean	1 Gy
5	Urethra	Mean	1 Gy
6	Bladder	Mean	1 Gy
7	Cumulative beam weight	Sum	1 MU

Table 4.1: Wish-list containing constraints and objectives used for plan generation for the prostate cases in this study.

Abbreviations: PTV = planning target volume; CTV = clinical target volume; gEUD-12 = generalized equivalent uniform dose using a biological parameter of 12; MU = monitor unit. Priority numbers indicate the order in which the objectives were optimized; a low number corresponds to a high-priority objective. A constraint must always be fulfilled during treatment planning. Dose limits for the targets (PTV, CTV, and peripheral zone) are expressed relative to the prescribed dose (38 Gy).

Treatment plans

Treatment plans were generated using Erasmus-iCycle, a treatment planning system developed in-house (66) that was extended with CyberKnife treatment planning (67, 68). The planning system performs 'prioritized' optimization, which means that objectives are optimized successively according to their assigned priorities. The constraints and the objectives with their priorities are defined by the user in the so-called wish-list. Typically, the same wish-list can be used for an entire patient group, thereby allowing for fully automated plan generation (69, 70). For CyberKnife treatments using circular aperture collimators, Erasmus-iCycle features a node reduction technique to generate time-efficient treatment plans (67). In this study, we assumed that CyberKnife beam collimation was performed using a variable circular aperture collimator (25).

Treatment plans were generated for 3 prostate cases, with CTVs ranging from 52 to 66 ml. The CTV consisted of the prostate without seminal vesicles. Treatment planning was performed according to a virtual HDR brachytherapy protocol (15, 40, 61). We required at least 95% of the PTV and 99% of the CTV to receive 38 Gy (4 fractions, prescribed at 67% isodose) and at least 95% of the peripheral zone to receive 120% of the prescribed dose (45.6 Gy). Urethra dose was restricted, and doses to surrounding organs-at-risk (OARs) were minimized. The wish-list used in this study to achieve these aims is given in table 4.1. For each prostate case, treatment plans were generated using 2 CTV-to-PTV margins: the clinically applied margin of 3 mm and a reduced margin of 0 mm. For each margin, 2 treatment plans were generated: 1 using the standard prostate node set and 1 using the adaptive timing prostate node set. A node set contains the principal beam directions from which the CyberKnife can irradiate. The standard prostate node set contains 106 node positions. The adaptive timing node set contains only 72 node positions, as node positions that block the imaging system are left out.

Treatment simulations

We simulated realistic 4-fraction CyberKnife treatments by combining real-time prostate motion data with the CyberKnife delivery characteristics (i.e. robot trajectory and speed, linear accelerator output, and correction strategy). As treatment fractions were longer than the individual motion tracks, available tracks were randomly combined by connecting the end point of one track with the starting point of another track. Only tracks belonging to the same motion data patient were combined, and connection points during beam-on time were not allowed to avoid unrealistic changes in prostate motion. Because the prostate motion data only described intrafraction displacements, we intrinsically simulated perfect patient set up at the start of each treatment fraction. For each treatment fraction, the time points were determined when beams were switched on and off and when imaging and subsequent correction were performed. Imaging was simulated to be performed when the next beam could not be delivered within a user-defined time interval since the last image acquisition. At each imaging/correction time-point, the robot correction was determined according to the simulated correction strategy. The residual motion was then calculated as the difference between the prostate displacements and the applied robot corrections. A detailed description of the combined prostate motion tracks is provided in Appendix 4.A.

Using the residual motion track as input, a 4-dimensional (4D) dose calculation was performed by an algorithm developed in-house. The static dose distribution was calculated (at CT grid resolution) for each beam individually, inside a region of interest enclosing all possible locations and orientations of the CTV and high-dose regions of the OARs. The delivered dose of each beam was then obtained by convoluting the static dose with the residual motion track corresponding to the beam-on-time. The residual

motion tracks were down-sampled to 1 data point per 5 s, as this provided a reasonable trade-off between calculation time and simulation accuracy. The total dose distribution was subsequently calculated by adding the convoluted dose distributions of all beams for all 4 treatment fractions.

For the standard treatment plans, simulations were performed using fixed time intervals between corrections of 15, 60, 180, and 360 s. For the adaptive timing treatment plans, simulations were performed with the adaptive timing mode both switched on and off. During adaptive imaging/correction, the time interval was reduced from 60 s to 15 s when translations larger than 3 mm or rotations larger than 2° were observed between 2 consecutive images. For all 6 timing strategies (15, 60, 180, and 360 s; adaptive imaging on and adaptive imaging off), 4 different robot correction scenarios were simulated: (1) full translational corrections only, and full translational corrections together with rotational corrections up to (2) 5°; (3) 10°; and (4) also fully corrected. Rotational corrections of 5° and 10° corresponded to clinical scenarios in which maximum robot rotational corrections (5°) and an additional maximum robotic couch rotation (5°), respectively, were applied. Treatment simulations were additionally performed for the standard and adaptive timing treatment plans assuming no robot corrections and no image acquisition accordingly. For all 3 prostate cases, 52 combinations of time interval, correction strategy, and margin were simulated. For each combination, we simulated 50 complete 4-fraction treatments for all 17 motion data patients, resulting in a total of 132600 treatment simulations.

Evaluation

The dosimetric parameters used for target evaluation were the CTV coverage (volume receiving the prescribed dose, V_{100%}) and the dose received by 98% of the peripheral zone volume (PZ $D_{98\%}$). For the rectum and bladder, the dose received by 1 cc (D_{1cc}) was evaluated. The urethra dose was assessed using the dose received by 5% $(D_{5\%})$ and 10% ($D_{10\%}$) of the volume. The near-minimum dose received by the CTV (CTV $D_{98\%}$) is reported in Appendix 4.C.

We evaluated the percentage of treatments for which the CTV coverage was higher than 98%. The correction strategy and margin were considered effective when this requirement was achieved in at least 98% of the treatments. Statistical analysis was performed comparing planned and simulated dose parameters (Wilcoxon signed-rank test), and differences between correction strategies and margins (Wilcoxon rank-sum test). Pvalues less than 0.05 were considered statistically significant.

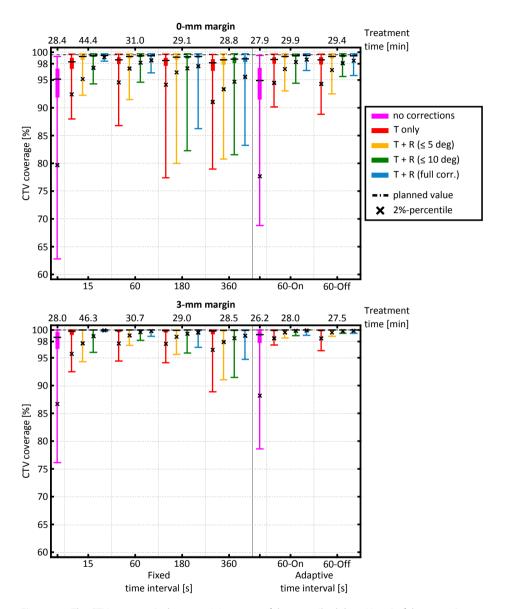


Figure 4.2: The CTV coverage (volume receiving 100% of the prescribed dose, $V_{100\%}$) of the 0-mm (upper graph) and 3-mm (lower graph) margin plans is shown for different correction strategies. Each box plot indicates the quartiles. Average simulated treatment times are shown for each time interval. CTV = clinical target volume; T = translational correction; R = rotational correction.

4.3 Results

60-Off

Figure 4.2 shows the CTV coverage of the o-mm and 3-mm margin plans for the different correction strategies simulated. Each box plot displays quartiles of the outcomes of 2550 treatment simulations (i.e. 3 prostate cases, 17 prostate motion patients, 50 simulations). For the standard o-mm margin plans, the CTV coverage decreased from the average planned value of 99.5% to 93.6% (SD: 4.8%; range: 62.9%-99.2%; p < 0.001) when no motion correction was applied. The standard 3-mm margin plans were found to be more robust against intrafraction prostate motion, resulting in the CTV coverage decreasing from the average planned value of 100% to 97.4% (SD: 3.4%; range: 76.2%-100%; p < 0.001) when no motion correction was applied. The CTV coverage was better preserved with increasing extents of robot corrections and in general for shorter time intervals between corrections. However, the beneficial effect of reducing time intervals was marginal for intervals smaller than 60 to 180 s. The effect of enabling or disabling adaptive timing (i.e. allowing it to switch to a 15-s time interval in case of large prostate displacements) was found to be very small, but the 3-mm adaptive timing plans were less sensitive than the standard treatment plans.

Percentages of treatments with more than 98% CTV coverage are listed in table 4.2. For the o-mm margin plans, this requirement was met in more than 98% of treatments when applying translational corrections and rotational corrections up to 10° (or more), using a time interval of 60 s. Translational corrections and rotational corrections up to

Time 0-mm standard plan 3-mm standard plan interval [s] No corr Tonly T+R(<5°) T+R(<10°) T+R(full) No corr Tonly T+R(<5°) T+R(<10°) T+R(full) no imaging 60 99.7 92.2 96.6 100.0 100.0 180 95.8 99.6 99.8 99.9 Q1 1 97.7 360 99.6 0-mm adaptive plan 3-mm adaptive plan No corr. Tonly T+R(≤5°) T+R(≤10°) T+R(full) No corr Tonly T+R(≤5°) T+R(≤10°) T+R(full) no imaging 60-On 93.4 99.5 100.0 100.0

Table 4.2: Percentage of treatments with at least 98% CTV coverage (V_{100%} ≥ 98%) for different correction strategies and margins*.

Abbreviations: T = translational corrections; R = rotational corrections; 60-On = initial time interval of 60 s with adaptive timing switched on; 60-Off = initial time interval of 60 s with adaptive timing switched off.

100.0

100.0

100.0

When the percentage of treatments is higher than 98%, the correction strategy is displayed in green, between 95% and 98% in yellow, between 90% and 95% in orange, and lower than 90% in red.

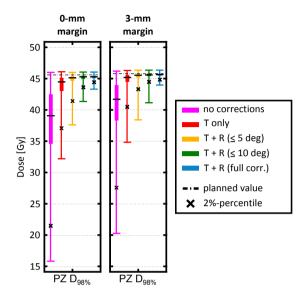


Figure 4.3: The dose received by 98% of peripheral zone (PZ D_{98%}) is shown for the 0-mm (left graph) and 3-mm (right graph) margin plans and for different correction strategies, using a time interval of 60 s. Each box plot indicates quartiles. T = translational corrections; R = rotational corrections.

5° (or more) together with a time interval of 60 s or 180 s were required for the standard 3-mm margin plans to achieve similar CTV coverage in at least 98% of the treatments.

Figure 4.3 shows the effect of intrafraction prostate motion on the boost dose in the PZ when using a time interval of 60 s. The D_{98%} of the PZ decreased significantly for both CTV-to-PTV margins, when no motion correction was applied, from an average planned value of 45.6 Gy to a simulated value of 37.7 Gy (SD: 6.0 Gy; range: 15.9-46.0 Gy; p < 0.001) for the o-mm margin plans and from 45.8 Gy to 40.4 Gy (SD: 4.6 Gy; range: 20.3-46.2 Gy; p < 0.001) for the 3-mm margin plans. Translational and rotational corrections should be applied for both CTV-to-PTV margins when the D_{98%} is required to be, for example, at least 110% of the prescribed dose in more than 98% of the treatments.

The dose received by the OARs when using a time interval of 60 s is shown in figure 4.4 for the 3 prostate cases separately. OAR doses were generally found to increase due to uncorrected intrafraction motion and to be better preserved with increasing extents of robot corrections. However, rectum D_{1cc} values decreased due to uncorrected motion by, on average, 1.9 \pm 5.5 Gy (1 SD) for the o-mm plans and by 3.4 \pm 5.4 Gy for the 3-mm plans. The prostate tended to move dorsally in the prostate motion data,

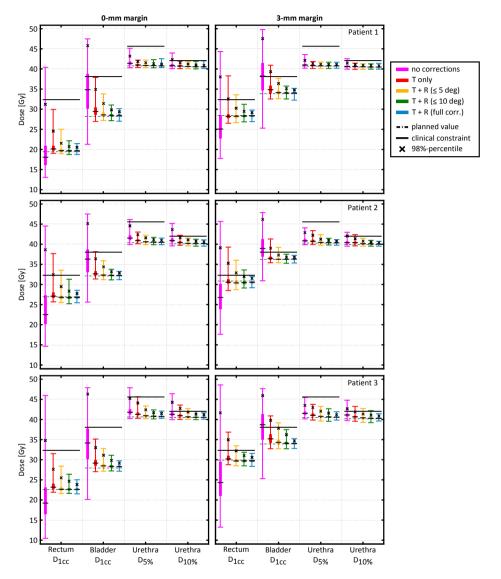


Figure 4.4: OAR dose parameters of the 0-mm (left column) and 3-mm (right column) margin plans of the individual prostate cases (rows) are shown for different correction strategies, using a time interval of 60 s. Each box plot indicates quartiles. OAR = organs-at-risk; T = translational correction; R = rotational correction.

and the rectum was simulated to displace accordingly, thereby moving into a region with lower doses. For the rectum and bladder, planned D_{1cc} values were lower in the o-mm margin plans than in the 3-mm margin plans by, on average, 6.8 Gy (range: 4.0-9.0 Gy) and 5.2 Gy (range: 4.1-5.9 Gy), respectively. As a result, rectum and bladder sparing were found to be considerably better for the o-mm margin plans. When using the correction strategies to achieve 98% CTV coverage in at least 98% of treatments, rectum doses differed from the planned values by, on average, 0.1 \pm 3.0 Gy and 0.2 \pm 1.2 Gy for the o-mm and 3-mm margin plans, respectively, whereas bladder D_{1cc} values differed by, on average, 0.3 ± 1.9 Gy and 0.4 ± 1.2 Gy, respectively. The effect of intrafraction prostate motion on urethra dose was generally small.

4.4 Discussion

In this study, the actual delivered dose in the presence of intrafraction motion was simulated for hypofractionated CyberKnife prostate treatments with a SIB. To achieve the 98% CTV coverage in at least 98% of the treatments, rotational corrections of up to 10° and up to 5° (in addition to full translational corrections) were required for a o-mm and 3-mm margin, respectively. This requirement may be perceived as very strict but is in line with the excellent treatment outcomes that should be considered standard in this patient group (71). It is important to note that other treatment uncertainties (e.g. beam delivery inaccuracy and delineation uncertainty) were not considered in the current study and that additional margins should be applied to account for these treatment uncertainties.

An important finding of the current study is the fact that more frequent imaging/correction (e.g. 15-s time interval) does not necessarily result in improved CTV coverage. As image acquisition and correction take time, the fraction duration increased with a decreasing time interval (figure 4.2), which resulted in larger residual errors when displacements were not fully corrected (see Appendix 4.A). This might also explain the improved robustness of the 3-mm adaptive timing plans, which had considerably shorter treatment times than the standard treatment plans. The optimal time interval and to what extent it can further improve the delivered dose should be established in future research. Xie et al. (54) performed a geometrical analysis and advised a time interval of approximately 40 s, which is considerably shorter than 60 to 180 s recommended in this study. The current study shows that short time intervals should be used with caution.

Prostate motion data measured by the Calypso system was previously used to perform a dosimetric evaluation of helical tomotherapy (HT) and step-and-shoot intensitymodulated radiation therapy (IMRT) (58, 59). In contrast to the findings in the current study, the effect of intrafraction motion was concluded to be generally low for both delivery techniques. This is probably due to the fact that these studies only included prostate translations and due to the use of larger PTV margins (5 mm, 3 mm in posterior direction), homogeneous dose distributions and 2 Gy fractions in HT and IMRT treatments. Whether the findings of the current study also hold for hypofractionated prostate treatments delivered using different treatment machines, but with a similar SIB scheme and tight margins, should be established in future studies.

We used CT data from 3 prostate cancer patients previously treated in our clinic using the CyberKnife, as the CT data corresponding to the Calypso motion tracks were not at our disposal. Correlation between patient anatomy and prostate motion was therefore missing. Although rectum and bladder doses varied considerably (figure 4.4), we observed very similar patterns for the CTV dose for all 3 patients. Based on these results, we speculate that the findings of the current study also hold for other patients.

In the current study, rotations around the left-right axis were derived from the measured anterior-posterior translations, and we randomly combined motion tracks to obtain tracks of sufficient length. These operations might have introduced errors compared with actual prostate motion during a CyberKnife treatment fraction. This is, however, difficult to verify, as intrafraction motion data for fractions of comparable length are not reported in the literature. A detailed description of the simulated prostate motion is provided in Appendix 4.A, which can be used for comparison and to relate the outcomes of this study to CyberKnife treatments at other institutes. A further limitation of this study is the assumption of rigid organ motion, which can be realistically assumed for the prostate (without vesicles) and urethra, but is likely to be inadequate for the rectum and bladder. However, as the high-dose regions in these organs are located very close to the prostate, rigid motion seems a reasonable assumption for the D_{1cc} -parameter that we evaluated. Finally, we assumed the densities encountered by each beam along its path to be constant during the entire treatment.

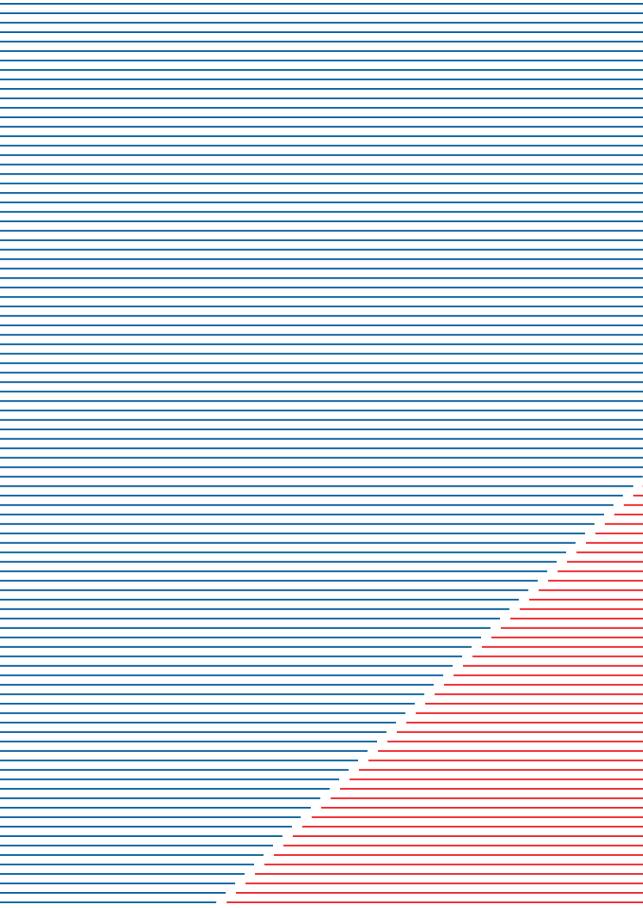
4.5 Conclusions

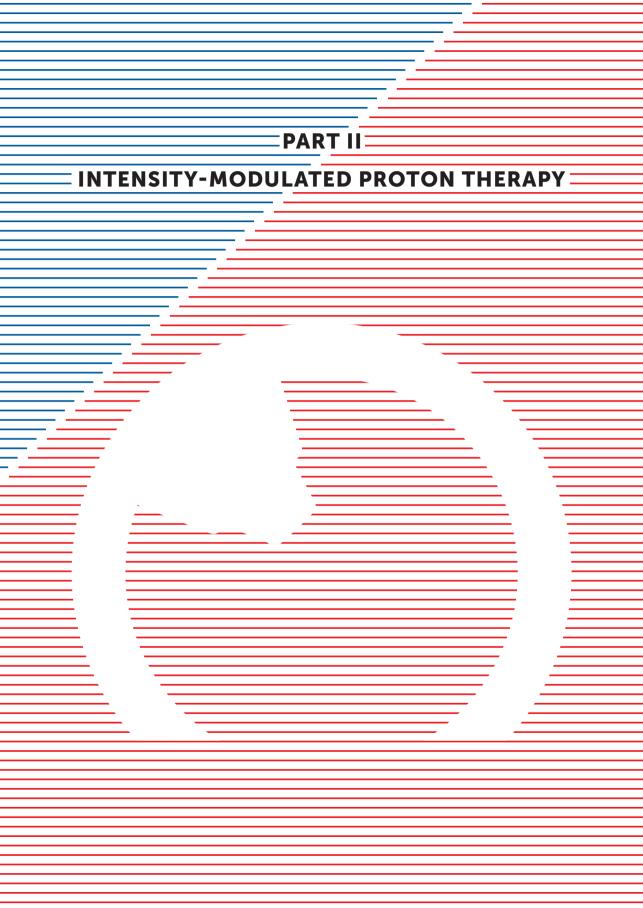
For hypofractionated CyberKnife prostate treatments with a simultaneously integrated boost, intrafraction motion can have a substantial impact on doses delivered to the CTV, boost volume, rectum and bladder. Applying robot corrections with a time interval of 60 to 180 s was found to be an effective way to account for prostate motion. To ensure

98% CTV coverage in more than 98% of the treatments, translational and rotational corrections up to 10° are recommended when using a 0-mm margin, while rotational corrections up to 5° are advised when using a 3-mm margin. Rectum and bladder were spared considerably better by using a o-mm margin. The adaptive time interval feature did not improve delivered dose distributions.

Acknowledgments

The authors would like to thank Katja Langen for sharing the prostate motion data set.





chapter 5

Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams

Physics in Medicine and Biology, Vol. 58, pp. 6969-6983, 2013

Steven van de Water¹,

Aafke Kraan¹,

Sebastiaan Breedveld¹,

Wilco Schillemans¹,

David Teguh¹,

Hanne Koov²,

Tom Madden²,

Ben Heijmen 1,

Mischa Hoogeman¹

 ¹ Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands
 ² Department of Radiation Oncology, F H Burr Proton Therapy Center, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Abstract

This study investigates whether 'pencil beam resampling', i.e. iterative selection and weight optimization of randomly placed pencil beams (PBs), reduces optimization time and improves plan quality for multi-criteria optimization in intensity-modulated proton therapy, compared with traditional modes in which PBs are distributed over a regular grid. Resampling consisted of repeatedly performing: (1) random selection of candidate PBs from a very fine grid, (2) inverse multi-criteria optimization, and (3) exclusion of low-weight PBs. The newly selected candidate PBs were added to the PBs in the existing solution, causing the solution to improve with each iteration. Resampling and traditional regular grid planning were implemented into our in-house developed multi-criteria treatment planning system 'Erasmus-iCycle'. The system optimizes objectives successively according to their priorities as defined in the so-called 'wish-list'. For five headand-neck cancer patients and two PB widths (3 and 6 mm sigma at 230 MeV), treatment plans were generated using: (1) resampling, (2) anisotropic regular grids and (3) isotropic regular grids, while using varying sample sizes (resampling) or grid spacings (regular grid). We assessed differences in optimization time (for comparable plan quality) and in plan quality parameters (for comparable optimization time). Resampling reduced optimization time by a factor of 2.8 and 5.6 on average (7.8 and 17.0 at maximum) compared with the use of anisotropic and isotropic grids, respectively. Doses to organs-at-risk were generally reduced when using resampling, with median dose reductions ranging from o.o to 3.0 Gy (maximum: 14.3 Gy, relative: 0%-42%) compared with anisotropic grids and from -0.3 to 2.6 Gy (maximum: 11.4 Gy, relative: -4%-19%) compared with isotropic grids. Resampling was especially effective when using thin PBs (3 mm sigma). Resampling plans contained on average fewer PBs, energy layers and protons than anisotropic grid plans and more energy layers and protons than isotropic grid plans. In conclusion, resampling resulted in improved plan quality and in considerable optimization time reduction compared with traditional regular grid planning.

Introduction 5.1

In spot-scanned intensity-modulated proton therapy (IMPT), discrete proton pencil beams (PBs) are successively delivered to the target volume. Each PB (or 'spot') has a beam direction, beam width, lateral position and energy, and is individually weighted during inverse treatment planning, Conventionally, in IMPT treatment planning, candidate PBs of a certain beam width are distributed over the target volume using a regular grid for each beam direction (12). This grid can be isotropic (Cartesian grid) or anisotropic. For the latter, the spacing between energy layers is typically smaller than the lateral spot spacing or non-uniformly depending on the Bragg peak width in depth (72, 73). The lateral spot spacing and the spacing between energy layers must be defined by the user before the optimization can start. Smaller grid spacing theoretically results in improved plan quality, as the degrees of freedom during inverse optimization increase (73). However, the use of very fine grids is limited during treatment planning by the optimization time. Optimization times will increase when using finer grids, due to an increase in the number of PBs. Selecting a grid spacing thus involves making a trade-off between plan quality and planning time (73). This is particularly relevant for large target volumes.

To improve IMPT treatment planning efficiency, an alternative to regular grid planning is presented in this study, called 'PB resampling' (also denoted as 'resampling'). The principle of resampling is to repeat the inverse optimization, while randomly selecting in each iteration a relatively small set of candidate PBs from a very fine grid. This has two potential advantages. Firstly, the optimization problem is divided into multiple smaller sub-problems. Since optimization times typically grow polynomially with the number of PBs, this is very likely to reduce optimization time. Secondly, more optimal PB placement may be possible, as the PBs are not confined to a coarse regular grid. Resampling was originally developed for treatment planning in robotic radiosurgery using the CyberKnife, where it was shown to successfully improve CyberKnife treatment planning efficiency and to additionally allow for delivery time reductions of up to approximately 60% (33, 67).

The aim of this study is to compare the performance of PB resampling for multicriteria IMPT treatment planning with traditional isotropic and anisotropic regular grid planning, in terms of plan quality and optimization time. This comparison will be made for two PB widths and five oropharyngeal cancer patients. This treatment site is characterized by relatively large target volumes and many organs-at-risk (OARs) surrounding the target.

5.2 Methods and materials

Treatment planning system

Treatment plans for IMPT were generated using 'Erasmus-iCycle', an in-house developed multi-criteria treatment planning system (66), which was extended with proton PB scanning. The system performs 'prioritized' or 'lexicographic' optimization. This means that objectives are optimized successively according to their pre-selected priorities instead of optimizing a single objective function in which the relative importance of each objective is a priori defined by a weighting factor (32). Erasmus-iCycle switches to optimizing the next objective with a lower priority when the desired or optimal value for the current objective is obtained and constrained. This procedure is continued until none of the objectives can be optimized any further. In this way, Erasmus-iCycle always produces treatment plans that are pareto-optimal with respect to the considered objectives. The hard constraints and the objectives with their priorities are defined by the user in the so-called 'wish-list'. Typically, the same wish-list can be used for an entire patient group, thereby allowing for fully automated treatment planning (70). This also enables an objective treatment planning comparison between planning techniques, as user-related variability is not present.

The proton dose calculation algorithm used by Erasmus-iCycle was developed at the Massachusetts General Hospital and Harvard Medical School, where it is implemented in the in-house developed treatment planning system ASTROID (23). The dose calculation algorithm uses a superposition-convolution approach in order to account for density heterogeneities. Dose calculation voxels of $3 \times 3 \times 3$ mm³ were used (73), which provided a reasonable trade-off between dose accuracy and calculation time.

Pencil beam resampling

Resampling and traditional regular grid planning were implemented in the ErasmusiCycle treatment planning system. Figure 5.1 shows the flow charts for each technique. Resampling mainly differed from regular grid planning on two aspects, which are highlighted in the flow chart. Firstly, candidate PBs were selected by randomly sampling a user-defined number of PBs from a very fine grid. Secondly, a loop was introduced in which the candidate PB selection, dose matrix calculation, inverse optimization, and PB reduction were repeated. In this loop, newly selected candidate PBs were added to the PBs in the existing solution, after which they were optimized all together. Since the previous solution was thereby available in the set of candidate PBs to be optimized, the solution

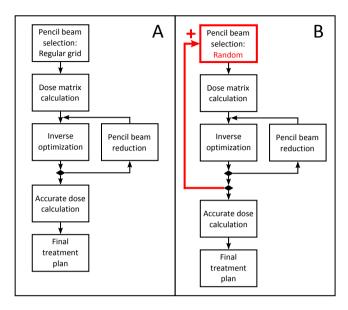


Figure 5.1: Flow charts of traditional regular grid planning (A) and resampling (B) as implemented in the treatment planning system. Main differences are highlighted by the thick red lines (in B). See section 5.2 for details and explanation of the 'pencil beam (PB) reduction' loop.

only improved throughout iterations. In this study, resampling was terminated after a user-defined number of iterations, but a stopping criteria based on a preset minimal plan parameter improvement could also be used.

A 'PB reduction' loop was included in each resampling iteration to avoid excessive growth in the set of candidate PBs. In this loop, PBs with a weight below the minimum weight (0.001 Giga-protons (Gp)) and 5% of the lowest contributing PBs were iteratively excluded from the solution after which the remaining PBs were reoptimized. The contribution of a PB was defined as the maximum dose delivered to the targets. All previously achieved dose parameters were constrained during the reoptimization, thereby maintaining plan quality, and only the sum of the beam weights was minimized. Beam reduction was terminated when the exclusion of the 5% lowest contributing PBs no longer resulted in a feasible solution and when all PBs had a weight higher than the minimum beam weight. In this way, an efficient and deliverable set of PBs was obtained. PB reduction was also performed at the end of regular grid planning to allow for a fair comparison between the optimization techniques.

Proton pencil beam model

Available proton energies ranged from 70 MeV to 230 MeV, corresponding to ranges in water from 40 mm to 325 mm, respectively. Two PB widths were evaluated. Thin PBs had a beam width (sigma in air at the isocenter) ranging from 3 to 7 mm, at 230 MeV and 70 MeV, respectively. Wide PBs had a sigma ranging from 6 to 10 mm, at 230 MeV and 70 MeV, respectively. In order to irradiate superficially located targets, a single range shifter of 75 mm water-equivalent thickness was available. The range shifter was modelled such that it could be inserted during the delivery of a field.

Patient group and dose prescriptions

Data of five oropharyngeal cancer patients were used in this study. We used this treatment site, as it combined large target volumes with many neighboring OARs. Dose was prescribed using a simultaneously integrated boost scheme, prescribing 66 Gy to the high-dose target and 54 Gy to the low-dose target, delivered in 30 fractions. The highdose clinical target volume (CTV) contained the primary tumor and positive neck levels, and the low-dose CTV consisted of elective neck levels. A CTV to planning target volume (PTV) margin of 5 mm in all directions was used to account for treatment uncertainties (74–76). The median total PTV (low-dose + high-dose) was 249 ml (range: 176–577 ml), while the high-dose PTV only had a median value of 107 ml (range: 19–267 ml).

The wish-list for this patient group is presented in table 5.1. We used a clinically accepted wish-list for IMRT in oropharyngeal patients (70), which was adjusted to IMPT treatment planning in close collaboration with a physician. The minimum PTV doses were constrained, implying that these were always fulfilled in the treatment plan. Within these constraints, the planning system successively optimized maximum PTV doses, high-dose conformality, OAR doses, low-dose conformality, unspecified tissue dose and cumulative beam weight (expressed in Gp). The dose values of the PTV constraints and objectives were chosen according to the aim of having 95% of the prescribed dose in at least 98% of the PTV ($V_{95\%} \ge 98\%$) and 107% of the prescribed dose in not more than 2% of the PTV ($V_{107\%} \le 2\%$) (75). To obtain a steep dose fall-off from the high-dose PTV to the low-dose PTV, the low-dose PTV was divided during treatment planning into a 5 mm transition region (PTV-intermediate) and the remaining low-dose PTV that excluded the PTV-intermediate (PTV-low'). The objective for the spinal cord and brainstem was a maximum dose of 15 Gy. If this objective was achieved, an extra optimization run was performed after optimizing all subsequent objectives to further minimize the dose below 15 Gy. The objective of 15 Gy seems very low compared with a typically used constraint of 50 Gy. The rationale behind this objective is that it accommodates future irradiations, while it can be achieved relatively easily using IMPT. Voxels that were located in an over-

Table 5.1: Wish-list containing the constraints and objectives for the head-and-neck patient group used in this study. The priority numbers indicate the order in which the objectives are optimized, a low number corresponds to a high-priority objective. A constraint has always to be fulfilled during treatment planning. PTV-intermediate is a 5 mm transition region (within PTV-low) at the PTV-high and PTV-low interface. PTVlow' was constructed by subtracting PTV-intermediate from PTV-low.

Objective	Type	Goal	Priority
PTV-high	minimum	0.98 × 66 Gy	constraint
PTV-intermediate	minimum	0.98 × 54 Gy	constraint
PTV-low'	minimum	$0.98 \times 54 \text{Gy}$	constraint
PTV-high	maximum	1.08 × 66 Gy	1
PTV-intermediate	maximum	1.08 × 66 Gy	1
PTV-low'	maximum	1.08 × 54 Gy	1
PTV-high ring 0-5 mm	maximum	1.05 × 66 Gy	2
PTV-low ring 0-5 mm	maximum	1.05 × 54 Gy	2
Parotid glands	mean	0 Gy	3
Submandibular glands	mean	0 Gy	4
Spinal cord	maximum	15 Gy	5
Brainstem	maximum	15 Gy	5
Swallowing muscles	mean	0 Gy	6
Larynx	mean	0 Gy	6
Oral cavity	mean	0 Gy	6
PTV shell 5 mm	maximum	0 Gy	7
PTV shell 10 mm	maximum	0 Gy	7
Unspecified tissue	maximum	0 Gy	8
Cumulative beam weight	sum	0 Gp	9

lap region between two structures were ascribed to both structures. Due to the prioritized optimization, the dose in an overlap region was primarily governed by the structure with the highest priority.

A three-beam coplanar arrangement as described in literature was used, with gantry angles of 50°, 180° and 310° (76). For each beam direction, candidate PBs were distributed over the PTV expanded by 5 mm to assure proper dose deposition at the target edges (12).

Treatment plan generation

Treatment plans were generated using: (1) resampling, (2) anisotropic regular grids and (3) isotropic regular grids. The characteristics of these three planning approaches are schematically illustrated in figure 5.2.

For resampling, treatment plans were generated using sample sizes of 1000, 3000 and 5000 PBs per resampling iteration. PBs were sampled from a grid with a spacing of 1 mm in both lateral directions and 2 mm between energy layers. It was possible to use a smaller energy layer spacing of 1 mm. However, this would hardly improve plan quality,

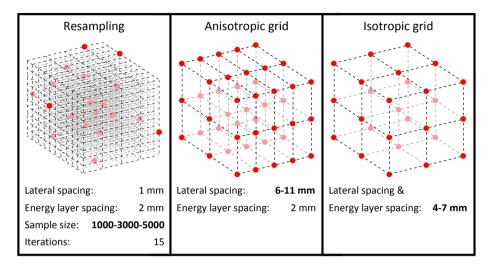


Figure 5.2: Schematic illustration of the different treatment planning approaches and their treatment planning settings: (1) resampling, (2) anisotropic regular grid planning and (3) isotropic regular grid planning. The settings that were varied during treatment plan generation are depicted in bold.

while it would result in treatment plans with more energy layers and consequently longer delivery times (see section 5.4 for considerations on the energy layer spacing, plan quality and treatment time). Resampling was terminated after 15 iterations, as outcomes had typically converged to a stable solution after this number of iterations (see the Appendix 5.A).

Anisotropic grid plans were generated using a fixed energy layer spacing of 2 mm, while varying the lateral grid spacing between 6 and 11 mm. As a consequence, the number of available energy layers in the anisotropic grid plans and in the resampling plans was similar. Isotropic grid plans were generated using a grid spacing ranging from 4 to 7 mm in both lateral directions and between energy layers. To avoid very long optimization times, the 6 and 7 mm anisotropic grids and the 4 mm isotropic grid were not used if the optimization time of a coarser grid spacing was already longer than the longest resampling plan. Note that as the energy layer spacing in isotropic grid plans was larger than in anisotropic grid plans, the lateral grid spacing could be smaller for similar optimization times.

These treatment plans were generated for all patients and for both PB widths, resulting in a total of 450 resampling plans, 52 anisotropic grid plans and 36 isotropic grid plans. Plan generation was performed on an 8-core computer using MATLAB (The MathWorks, Inc., Natick, MD).

Treatment plan evaluation

To evaluate and rank the overall plan quality of each treatment plan, we calculated the so-called 'plan-score', which was defined as the weighted-sum of the differences between the obtained value and the desired value for each objective. A lower plan-score thus corresponds to a treatment plan that better meets the objectives as defined in the wish-list. The weights were obtained by linearly converting the wish-list priority numbers into normalized weighting factors, where a high-priority (low-priority number) corresponded to a high weight. The plan-score was considered a practical measure, as the multi-criteria optimization does not provide a straightforward quantitative metric for assessing and ranking the plan quality. Besides plan-score, treatment plans were evaluated on optimization time, OAR doses and on the number of PBs, energy layers and protons. The expected delivery time was not part of the evaluation in the current study, but will be addressed in section 5.4.

We evaluated the efficiency of the planning techniques by comparing the optimization time required to achieve a certain plan quality, and the other way around, by comparing the plan quality achievable in a certain optimization time. Thus, in comparing optimization time, anisotropic and isotropic grid plans were compared with resampling plans with the smallest difference in plan-score. In comparing plan quality, anisotropic and isotropic grid plans were compared with resampling plans with the smallest difference in optimization time. For all patients, we only included treatment plans in the evaluation for which the constraints and objectives of the PTVs were fulfilled. PTV doses are in this way comparable for all treatment plans evaluated. Statistical analysis of the results was performed using the Wilcoxon signed-rank test. P-values lower than 0.05 were considered to be statistically significant.

5.3 Results

The plan-score as a function of the optimization time is given in figure 5.3 for all patients and both PB widths. In each panel, results are presented for resampling using sample sizes of 1000, 3000 and 5000 PBs, and for anisotropic and isotropic regular grid planning using different grid spacings. All graphs show that a better plan quality (lower plan-score) was achieved in shorter optimization times when using resampling instead of using traditional anisotropic or isotropic regular grid planning. This improvement is especially pronounced when using thin PBs. Anisotropic grids generally performed better than isotropic grids. Some resampling graphs show a minor increase of the planscore in successive iterations, suggesting a worsening of the plan quality, despite the

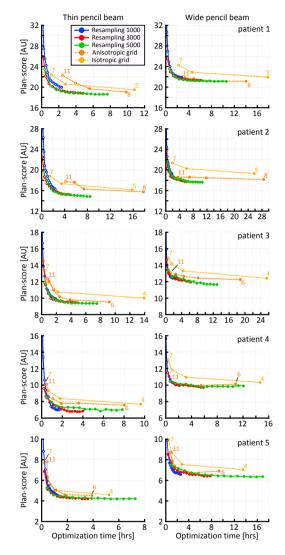


Figure 5.3: Graphs displaying for each patient (rows 1–5) the plan-score as a function of the optimization time. Thin PBs are displayed in the left column, wide PBs in the right column. Solid and open markers indicate treatment plans in which the PTV constraints and objectives are fulfilled or violated, respectively. Solid lines connect treatment plans that were obtained successively during a single optimization sequence (resampling), while dashed lines connect plans that were obtained after individual optimization runs (regular grid planning). Numbers next to regular grid plans indicate the smallest and largest lateral grid spacing used (equal to the energy layer spacing in isotropic grids).

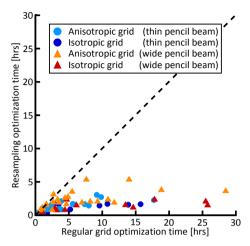


Figure 5.4: Scatter plot indicating the optimization time of anisotropic and isotropic regular grid plans (horizontal axis) and the optimization time of their corresponding resampling plans with the smallest difference in plan quality (vertical axis), all fulfilling the PTV constraints and objectives. For resampling, sample sizes were 5000 and 3000 for thin and wide PBs, respectively.

fact that the optimality of the solution should only improve during resampling. These small increases are an artefact of the plan-score definition as the weighted-sum of plan parameters, which is not directly optimized using a prioritized wish-list. For anisotropic grid planning, the plan quality incidentally worsened when using less coarse lateral grid spacing, probably due to unfavorable alignment of the grid at the target edge or with respect to OARs. To obtain treatment plans for all patients that fulfilled the constraints and objectives of the PTVs, a sample size of 5000 PBs per iteration was required when using thin PBs, while a sample size of 3000 PBs was sufficient for wide PBs. In further plan comparisons sample sizes of 5000 and 3000 PBs were used for thin and wide PBs, respectively. In the treatment plans that fulfilled the PTV constraints and objectives, the average $V_{95\%}$ was 99.5% (range: 98.4%–100%) for the high-dose PTV and 99.4% (range: 98.7%–99.9%) for the low-dose PTV. The $V_{107\%}$ of the high-dose PTV was 0.9% on average (range: 0.1%–2.2%).

In figure 5.4, the optimization times are depicted for the anisotropic and isotropic regular grid plans and their corresponding resampling plans with the smallest difference in plan-score, all fulfilling the PTV constraints and objectives. Resampling resulted in optimization times reduced by a factor of 2.8 on average (factor of 7.8 at maximum) compared with anisotropic grid planning and by a factor of 5.6 on average (factor of 17.0 at maximum) compared with isotropic grid planning, averaged over both PB widths.

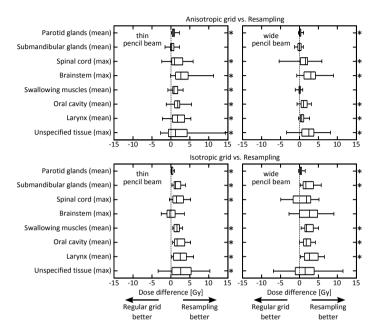


Figure 5.5: Box plots depicting the difference in dose parameters between regular grid plans and their corresponding resampling plans with the smallest difference in optimization time, all fulfilling the PTV constraints and objectives. The box plots display quartiles. Results for anisotropic grids are given in the top row, for isotropic grids in the bottom row, for thin PBs in the left column and for wide PBs in the right column. For resampling, sample sizes were 5000 and 3000 for thin and wide PBs, respectively. Significant differences are indicated by asterisks (Wilcoxon signed-rank test, p < 0.05).

The figure shows that the relative gain was especially large when the optimization times of regular grid plans were long.

Figure 5.5 shows for the OARs (in order of priority) the difference in dose parameters between the regular grid plans and their corresponding resampling plans with the smallest difference in optimization time, all fulfilling the PTV constraints and objectives. For most OARs, significant differences in favor of resampling were found. Median dose improvements ranged from 0.0 to 3.0 Gy (relative: 0%-42%) compared with anisotropic grids and from -0.3 to 2.6 Gy (relative: -4%-19%) compared with isotropic grids, depending on the OAR and PB width. The submandibular glands, swallowing muscles, oral cavity and larynx always received lower doses when using resampling instead of an isotropic grid, with improvements ranging from 0.0 to 6.5 Gy (relative: 0%-27%). Only for the brainstem in combination with an isotropic grid and thin PBs, regular grid plans

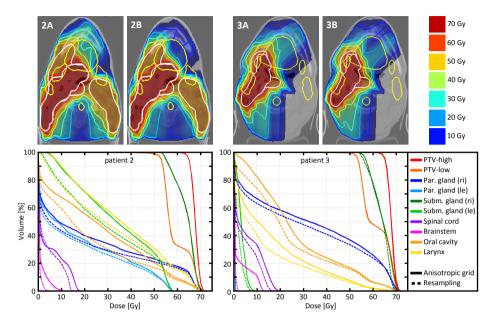


Figure 5.6: Dose distributions and dose-volume histograms (DVHs) of the $10 \times 10 \times 2 \text{ mm}^3$ anisotropic grid plans (dose distributions A, solid DVHs) and the corresponding resampling plans with comparable optimization time (dose distributions B, dashed DVHs) for patients 2 and 3, for thin PBs. PTV delineations are depicted in white and organs-at-risk in yellow. The DVHs of the swallowing muscles are not displayed for clarity.

were found to perform slightly better than the corresponding resampling plans. In this case however, the difference was not significant (p = 0.97).

The dose distributions and dose-volume histograms (DVHs) of anisotropic grid plans and the corresponding resampling plans with comparable optimization time are given in figure 5.6 for patients 2 and 3, for thin PBs. Dose distributions show lower doses in the oral cavity and improved conformality when using resampling. DVHs show overall dose reductions for the OARs in the resampling plans, while doses to the high-dose and lowdose PTVs are similar. Considerably lower doses were observed in patient 2 for the right parotid gland (mean dose: -2.4 Gy, -10.7%), the left submandibular gland (mean dose: -4.5 Gy, -13.4%), the oral cavity (mean dose: -5.4 Gy, -19.9%) and the larynx (mean dose: -4.9 Gy, -13.5%). For patient 3, resampling resulted in pronounced dose reductions for the brainstem (maximum dose: -11.2 Gy, -80.0%), the oral cavity (mean dose: -3.3 Gy, -12.0%) and the larynx (mean dose: -3.3 Gy, -19.5%).

Table 5.2: Average plan characteristics (and extreme values) of the regular grid plans and their corresponding resampling plans with the smallest difference in optimization time, all fulfilling the PTV constraints and objectives. Characteristics are given for plans generated using thin pencil beams (PBs). Significant differences between resampling and regular grid planning are indicated by asterisks (Wilcoxon signedrank test, p < 0.05).

	Anisotro	opic grid	Resampling		
Plan-score	10.5	(4.5 – 19.7)	9.7	(4.3–18.7)*	
Total number of pencil beams	1837	(1073 – 3387)	1541	(922-2769)*	
Total number of energy layers	186	(164 – 218)	185	(160-216)*	
Cumulative beam weight [Gp]	3874	(2157 – 7443)	3571	(2135–6388)*	
	Isotropic grid		Resampling		
Plan-score	11.3	(4.6 – 20.6)	10.4	(4.2–19.1)*	
Total number of pencil beams	1594	(934 – 2811)	1596	(952-2786)	
Total number of energy layers	80	(62 – 100)	186	(160-216)*	
Cumulative beam weight [Gp]	3609	(2121 – 6226)	3739	(2105–6526)*	

Table 5.2 lists for thin PBs the plan characteristics of the regular grid plans and their corresponding resampling plans with the smallest difference in optimization time, all fulfilling the PTV constraints and objectives. Resampling plans contained on average significantly fewer PBs, energy layers and protons compared with the anisotropic grid plans and more energy layers and protons compared with the isotropic grid plans (p < 0.05). In both cases however, the plan-score was significantly lower (better plan quality) in the resampling plans than in the regular grid plans (p < 0.05).

5.4 Discussion

This study showed that PB resampling in IMPT optimization can improve plan quality compared with traditional isotropic and anisotropic regular grid planning, resulting in considerable OAR dose reductions as depicted in figures 5.5 and 5.6. Moreover, resampling was found to dramatically shorten optimization times, which was illustrated in figure 5.4.

Resampling has two important additional advantages over traditional regular grid planning. Firstly, it allows making a real-time trade-off between plan quality and optimization time during treatment planning, as the inverse optimization is performed

repeatedly. The user can evaluate plan quality and corresponding optimization time after each resampling iteration and make the decision to continue plan optimization or to accept the current solution as the final treatment plan. This decision can be automated by defining a required minimal improvement for the plan-score or a specific OAR. If the improvement between resampling iterations is lower than this threshold, the optimization will be terminated. In this way, valuable optimization resources are only employed when deemed advantageous by the user. Secondly, resampling eliminates the important a priori choice of selecting a proper grid spacing. In practice, when using regular grid planning, the grid spacing is typically chosen based on previous experiences. Several studies systematically investigated the effect of the grid spacing on plan quality (73, 77). However, it is unclear how to apply their results to different patient anatomies, beam arrangements and dose prescriptions. When using resampling, the user has to choose a priori the number of PBs that will be randomly sampled in each iteration, but the effect of the sample size on the achievable plan quality was found to be low. This is illustrated in figure 5.3 by the general overlap of resampling graphs for different sample sizes. Still, the sample size can be too small to obtain a treatment plan that fulfils the constraints and objectives of the PTVs. Conversely, figure 5.3 also shows that sometimes smaller sample sizes resulted in better planning efficiency than larger sample sizes, i.e. an improved plan quality for comparable optimization times or shorter optimization times for a comparable plan-score. This suggests that there is a certain optimal sample size for each case, probably depending on target size and PB width. However, for sample sizes that did provide solutions fulfilling the PTV constraints and objectives, the final plan quality was found to be rather invariant to the sample size used, whereas the obtained plan quality could vary considerably with the grid spacing used during regular grid planning.

Although treatment plan optimization could be shortened considerably using resampling, absolute optimization times were still in the order of hours. This mainly resulted from the fact that our treatment planning system Erasmus-iCycle uses a prioritized optimization approach. The objectives are optimized successively according to their priority (instead of only once when using a single objective function), resulting in optimization times that depend approximately linearly on the number of priorities specified in the wish-list. Also the 'strictness' the wish-list can have an effect on the optimization time. A very challenging wish-list, with for example tight constraints on OAR doses, is likely to require fine grid spacings (for regular grid planning) or many iterations (for resampling) to obtain a feasible treatment plan, thereby resulting in relatively long optimization times. Long optimization times are associated with comparatively small differences in the plan quality between regular grid planning and resampling (see figure 5.3), but also with large gains in optimization time when using resampling (see figure 5.4). An important advantage of the prioritized approach is that it allows the treatment planning to be automated.

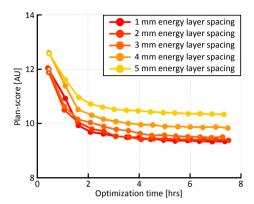


Figure 5.7: Plan-score as a function of the optimization time for patient 3. Resampling was performed using thin PBs, a sample size of 5000 PBs per iteration and an energy layer spacing ranging from 1 to 5 mm.

In the end, this can result in shorter planner hands-on time, as patient-specific tuning of weighting parameters by trial-and-error is not required (70).

In current proton therapy facilities, the delivery time is typically determined to a large extent by the number of energy layers in a treatment plan (78). A reduction of the energy layers can considerably reduce delivery times, but is also likely to result in a worsening of the plan quality. Figure 5.7 depicts the resampling outcomes for an energy layer spacing varying between 1 and 5 mm. The difference in plan-score is very small between 1 and 2 mm energy layer spacing (~1% difference), but increases for larger energy layer spacings. Based on these outcomes, we used an energy layer spacing of 2 mm for resampling and anisotropic regular grid planning. The expected delivery times of treatment plans were not explicitly evaluated in the current study. Instead, we compared resampling with anisotropic grid planning using the same energy layer spacing of 2 mm. This resulted in a comparison of treatment plans with comparable expected treatment times, in which resampling was shown to provide better plan quality and shorter optimization times. For isotropic grid planning, we expect shorter delivery times as the number of energy layers was lower. However, these plans were also associated with poorer plan quality and longer optimization times, compared with both resampling and anisotropic grid planning. Switching times between energy layers are very likely to be shortened in the future, reducing the impact of the number of energy layers on delivery times (78-80).

In this study, we additionally investigated the effect of the PB width on the planning outcomes. The use of thin PBs resulted in better plan quality and shorter optimization times (see figure 5.3). Optimization times for thin PBs were shorter due to the fact that they have a smaller 'volume-of-influence' (i.e. dose is deposited in fewer voxels), thereby

shortening dose-matrix operations performed during inverse optimization. Figure 5.3 also shows that the difference in plan quality and optimization times between resampling and anisotropic grid planning was larger for thin PBs than for wide PBs. Since the energy layer spacing was equal for resampling and anisotropic grid planning, differences in plan quality between these techniques are likely to have resulted mainly from differences in lateral positioning of the PBs. The effect of improving lateral PB placement is probably larger for thin PBs due to their steeper lateral dose fall-off. The effect of the PB width on plan quality was also investigated in other studies. Van de Water et al. generated IMPT treatment plans for ten head-and-neck cancer patients using two PB widths (76). When using thin PBs instead of wide PBs, they found mean dose reductions in the parotid glands and the oral cavity of on average 2.1 Gy and ~3 Gy, respectively. This is comparable to mean dose reductions of 3.6 and 4.2 Gy observed in the current study for treatment plans generated using a similar isotropic grid ($5 \times 5 \times 5$ mm³) and beam arrangement. In the submandibular glands however, they observed a mean dose reduction of 7.7 Gy on average (maximum of 24.6 Gy), which is much larger than the mean dose reduction of 3.1 Gy on average (maximum of 5.9 Gy) found in the current study. This might be caused by the fact that they varied the PB width by modelling the use of upstream degraders (thin PBs) or downstream degraders (wide PBs) for energy modulation, instead of varying the initial PB width in air (as in the current study). This probably resulted in different PB widths evaluated and in especially pronounced differences for certain beam energies and certain OARs. Also Widesott et al. found the plan quality to improve with smaller PB widths (77). For a head-and-neck case, they observed the target coverage and OAR sparing to be comparable to that of advanced photon techniques for PB widths equal to or smaller than 4 mm sigma.

We did not include robustness against setup and range errors in the treatment plan generation. Instead, we used a PTV margin of 5 mm to account for treatment uncertainties. It is very likely that robust treatment planning will be more effective in accounting for these treatment uncertainties (81, 82). We speculate that resampling is especially beneficial for large-scale optimization problems such as robust planning, as the relative gain in optimization time was found to be particularly large when the optimization times of regular grid plans were long. Future research should assess whether similar improvements between resampling and regular grid planning will indeed be observed for robust treatment planning.

5.5 Conclusion

Pencil beam resampling for IMPT, i.e. iterative selection and weight optimization of randomly placed pencil beams, resulted in improved plan quality and shorter optimization time compared with traditional regular grid planning. Optimization times were reduced by a factor of 2.8 on average (factor of 7.8 at maximum) compared with anisotropic grid planning and by a factor of 5.6 on average (factor of 17.0 at maximum) compared with isotropic grid planning. Improvements were especially pronounced when using thin pencil beams.

chapter 6

Shortening delivery times of intensity-modulated proton therapy by reducing proton energy layers during treatment plan optimization

International Journal of Radiation Oncology Biology Physics, accepted for publication on January 23, 2015

> Steven van de Water¹, Hanne Kooy², Ben Heijmen¹, Mischa Hoogeman¹.

 Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands
 Department of Radiation Oncology, F H Burr Proton Therapy Center, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Abstract

Purpose: To shorten delivery times of intensity-modulated proton therapy (IMPT) by reducing the number of energy layers in the treatment plan.

Methods and materials: We have developed an energy layer reduction method, which was implemented into our in-house developed multi-criteria treatment planning system 'Erasmus-iCycle'. The method consisted of two components: (1) minimizing the logarithm of the total spot weight per energy layer, and (2) iteratively excluding low-weighted energy layers. The method was benchmarked by comparing a robust 'time-efficient plan' (with energy layer reduction) with a robust 'standard clinical plan' (without energy layer reduction) for five oropharyngeal cases and five prostate cases. Both plans of each patient had equal robust plan quality, as the worst-case dose parameters of the standard clinical plan were used as dose constraints for the time-efficient plan. Worst-case robust optimization was performed, accounting for setup errors of 3 mm and range errors of 3% +1 mm. We evaluated the number of energy layers and the expected delivery time per fraction, assuming 30 s per beam direction, 10 ms per spot and 400 Giga-protons per minute. The energy switching time was varied from 0.1 to 5 s.

Results: The number of energy layers was on average reduced by 45% (range: 30-56%) for the oropharyngeal cases and by 28% (range: 25-32%) for the prostate cases. When assuming 1, 2 or 5 s energy switching time, the average delivery time was shortened from 3.9 to 3.0 min (25%), 6.0 to 4.2 min (32%) or 12.3 to 7.7 min (38%) for the oropharyngeal cases, and from 3.4 to 2.9 min (16%), 5.2 to 4.2 min (20%) or 10.6 to 8.0 min (24%) for the prostate cases.

Conclusions: Delivery times of IMPT can be reduced substantially without compromising robust plan quality. Shorter delivery times are likely to reduce treatment uncertainties and cost

6.1 Introduction

A reduction of the delivery time per fraction is likely to have a beneficial effect on patient comfort, treatment uncertainties and treatment costs. Especially the last two motives are important for intensity-modulated proton therapy (IMPT). Compared with photon therapy, IMPT is relatively sensitive to treatment uncertainties (83, 84) and it is associated with higher treatment costs (24, 85).

In spot-scanned IMPT, each proton pencil beam (or 'spot') is spatially characterized by its lateral position and its position in depth, the latter depending on the energy of the proton beam. Spots are delivered one-by-one, adjusting the lateral position and/or energy in between the delivery of each spot. In modern commercial double-scanning proton delivery systems (i.e. systems in which the proton beam is magnetically deflected in both lateral directions), the energy switching time is typically in the order of seconds, causing it to be the major component of the delivery time (78). A reduction of the number of energy layers in an IMPT treatment plan is therefore likely to result in a reduction of the delivery time. However, reducing energy layers by simply increasing the energy layer spacing was demonstrated to result in a worsening of the plan quality (73, 86).

In this paper we hypothesize that the number of energy layers can be reduced without affecting the plan quality, because of the degeneracy of IMPT treatment planning. Some treatment plans with a comparable plan quality may contain many energy layers, while others contain fewer. By taking the number of energy layers into account during treatment planning (which is typically not done), the solution can potentially be steered towards an outcome with fewer energy layers. For CyberKnife treatment planning, a similar approach resulted in treatment plans with fewer principal beam directions, thereby reducing delivery times in robotic radiosurgery (67).

The aim of this study is to reduce the delivery time per fraction of robust IMPT treatment plans without compromising the dosimetric plan quality. To this purpose, we developed a method to incorporate the reduction of energy layers in the treatment planning process. This method was implemented into our treatment planning system and tested by generating robust treatment plans for five oropharyngeal cancer patients and five locally-advanced prostate cancer patients.

6.2 Methods and materials

Treatment planning system

The novel energy layer reduction method was implemented into our in-house developed treatment planning system 'Erasmus-iCycle' (66, 86). The planning system does not optimize a single objective function that is the weighted-sum of all objectives, but it

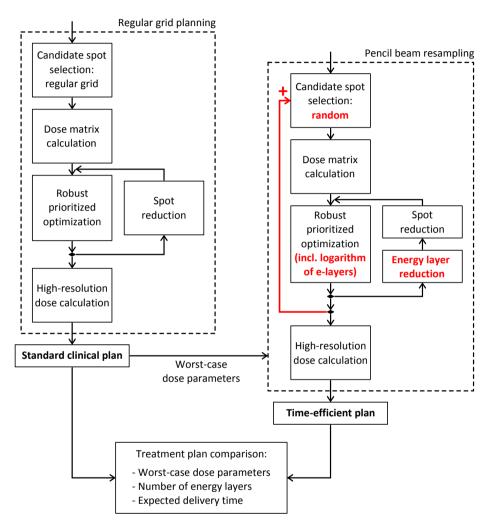


Figure 6.1: Study workflow describing treatment plan generation and comparison. The differences between the standard clinical plan and the time-efficient plan are highlighted in red.

performs 'prioritized' or 'lexicographic' multi-criteria optimization instead. This means that objectives are optimized one-by-one according to their priorities, which are defined a priori by the user in the so-called 'wish-list'. Generally, a single wish-list can be used per patient group, which enables fully automatic treatment planning. Automatic plan generation allows for an objective comparison between different planning strategies and was shown to result in superior plan quality compared with manual planning (70, 87).

Erasmus-iCycle features two planning methods for IMPT, which are schematically illustrated in figure 6.1. Firstly, treatment plans can be generated using the traditional planning method in which the candidate spots are distributed over the target volume using a regular grid with a spot spacing defined by the user (denoted as 'regular grid planning'). Secondly, treatment planning can be performed using the new and more efficient 'pencil beam resampling' method, as described by Van de Water et al. (86). This method uses an iterative approach in which the multi-criteria optimization is performed repeatedly, while adding in each iteration a new sample of randomly selected candidate spots to the obtained solution. The user has to define the sample size and a stopping criterion to terminate the iterative resampling process. Pencil beam resampling was shown to reduce optimization times by a factor of 2.8-5.6 on average, compared with traditional regular grid planning (86).

In Erasmus-iCycle, both planning methods feature a 'spot reduction' loop (indicated in figure 6.1). After each multi-criteria optimization, low-weighted spots and spots with a weight below the minimum deliverable spot weight are iteratively excluded from the solution, while constraining previously achieved dose parameters. The exclusion of spots is terminated when further reduction results in a violation of previously achieved dose parameters. In this way, Erasmus-iCycle will always generate deliverable plans that are efficient in terms of the number of spots (86).

The treatment planning system also features a robust planning mode to generate IMPT treatment plans that are robust against setup errors and range errors. ErasmusiCycle uses a 'minimax' approach for this purpose (88, 89). The nominal scenario and several error scenarios are simultaneously included in the multi-criteria optimization, optimizing the worst scenario for each objective. In this study, robust optimization was performed using a total of nine scenarios: the nominal scenario (1 scenario), 3 mm setup errors in positive and negative directions along three axes (6 scenarios), and a proton undershoot and overshoot of 3% +1 mm (2 scenarios).

The proton dose calculation algorithm used for IMPT treatment planning was developed at the Massachusetts General Hospital and Harvard Medical School, where it is implemented in their in-house developed radiotherapy platform 'ASTROID' (23). The algorithm accounts for density heterogeneities using a superposition-convolution approach. In this study, we used a dose grid with $3 \times 3 \times 3$ mm³ resolution. Available proton energies ranged from 70 to 230 MeV with corresponding beam widths ranging from 7 to 3 mm sigma (in-air at the isocenter), respectively. A range shifter of 75 mm water-equivalent thickness could be inserted during the delivery of a field. The minimum deliverable spot weight was 0.001 Giga-protons.

Energy layer reduction method

The energy layer reduction method that we developed in this study consisted of two components. The first component was the introduction of an additional objective, which aimed at minimizing the logarithm of the total spot weight per energy layer. The logarithm has a steeper gradient at lower values, causing the optimizer to focus on the minimization of spot weights in low-weighted energy layers, thereby redistributing them to relatively high-weighted energy layers. The second component was the iterative exclusion of low-weighted energy layers, which was implemented as an extension of the spot reduction loop described above. After each multi-criteria optimization, low-weighted energy layers were excluded from the solution first, before switching to the iterative exclusion of individual spots. The energy layer reduction was terminated when further exclusion resulted in the violation of previously achieved dose parameters. Excluded energy layers were not available during candidate spot selection in subsequent resampling iterations.

Patient data and dose prescription

Data of five oropharyngeal cancer patients and five locally-advanced prostate cancer patients were used in this study. The oropharyngeal cases (three unilateral and two bilateral cases) were prescribed a simultaneously integrated boost of 66 Gy to the primary tumor and positive neck levels, and 54 Gy to the elective neck levels, delivered in 30 fractions. The median high-dose clinical target volume (CTV) was 41 ml (range: 5-106 ml), while the median total high-dose and low-dose CTV was 95 ml (range: 67-229 ml). The prostate cases were prescribed a simultaneously integrated boost of 74 Gy to the prostate expanded by a margin of 4 mm, and 55 Gy to the seminal vesicles and pelvic lymph nodes, both expanded using a margin of 7 mm. These margins were applied to construct an internal target volume (ITV) in order to account for the organ-specific interfraction motion in prostate cases (90). The median high-dose ITV was 60 ml (range: 56-65 ml) and the median total high-dose and low-dose ITV was 875 ml (range: 810-957 ml).

The wish-lists containing the constraints and objectives for both patient groups are given in Appendix 6.A. The constraints and objectives of the target volumes were chosen such that the volume receiving 95% of the prescribed dose was higher than 99% $(V_{95\%} > 99\%)$ and the volume receiving 107% of the prescribed dose smaller than 2%

 $(V_{107\%} < 2\%)$ in all error scenarios included in the optimization. For the oropharyngeal cases, we included the salivary glands, spinal cord, brainstem, larynx, oral cavity and swallowing muscles as organs-at-risk (OARs) in the robust optimization, whereas the rectum, bladder and femur heads were considered for the prostate cases. The oropharyngeal cases were planned using a 3-beam arrangement (gantry angles: 60°, 180° and 300°), whereas the prostate cases were planned using a 2-beam arrangement (gantry angles: 90° and 270°) (76, 90).

Study design

To assess the effectiveness of the proposed energy layer reduction method, we compared for each patient a robust 'time-efficient' IMPT plan with a robust 'standard clinical' IMPT plan. The study workflow is depicted in figure 6.1. Firstly, the standard clinical plan was generated using traditional regular grid planning without energy layer reduction, resembling clinical practice in currently operating proton facilities. Regular grid planning was performed using a lateral spot spacing of 7 mm, corresponding to approximately two times the width (sigma) of the smallest spot (73). The spacing between energy layers depended on the proton energy and was set to the longitudinal width of the Bragg peak (at 80% of the peak height) (72, 73). Secondly, the time-efficient plan was generated using the new pencil beam resampling method with energy layer reduction. The worst-case dose parameters of each objective achieved in the standard clinical plan were used as dose constraints for the time-efficient plan, to ensure comparable robust plan quality in both treatment plans. Pencil beam resampling was performed using a sample size of 5000 spots per iteration, which were randomly selected from a grid with 1 mm lateral spot spacing and the same energy layer spacing as used for the standard clinical plan. Time-efficient plan generation was terminated when the number of energy layers could not be reduced any further in two consecutive resampling iterations.

We compared the worst-case dose parameters of the CTVs (oropharyngeal cases), ITVs (prostate cases) and OARs, considering all nine scenarios included in the robust optimization. Next to that, we evaluated the number of energy layers and the expected delivery time per fraction (i.e. time between the delivery of the first spot and the last spot, excluding patient setup), which was calculated assuming 30 s per beam direction (gantry rotation and beam setup), 10 ms per spot (lateral spot adjustment) and 400 Giga-protons per min (beam current). The energy switching time was varied between 0.1 and 5 s, as this was assumed to cover the variety in currently operating spot-scanning systems, from the PSI gantry 2 to a slow-switching system respectively (91, 92). Switching times of 1, 2 and 5 s were assumed to be typical values for commercial systems and are therefore explicitly mentioned in this paper.

Table 6.1: Average worst-case dose parameters and delivery parameters (and range) for the standard clinical plans and the time-efficient plans of the oropharyngeal and prostate cases. The conformity index was defined as the patient volume receiving 95% of the prescription dose divided by the target volume (CTV or ITV).

Parameter		Oropharyngeal cases (n=5)		-:	Dalatina difference		
		Standard clinical plan		Time-em	Time-efficient plan		Relative difference
CTV-low	V _{95%} [%]	99.5	(99.3-99.7)	99.6	(99.4-99.8)	0%	(0-0%)
CTV-high	V _{95%} [%]	99.4	(99.0-99.8)	99.5	(99.3-99.8)	0%	(0-0%)
CTV-high	V _{107%} [%]	0.3	(0.1-0.7)	0.3	(0.1-0.7)	31%	(-40-113%)
Parotid glands	mean [Gy]	14.1	(6.9-27.0)	14.1	(7.0-27.0)	0%	(0-2%)
Submandibular glands	mean [Gy]	34.7	(12.3–59.6)	34.7	(12.4–59.6)	0%	(-1-0%)
Spinal cord	max [Gy]	20.7	(19.7-21.3)	20.3	(19.6-20.9)	-2%	(-4-0%)
Brainstem	max [Gy]	19.4	(18.3-20.4)	19.4	(18.2-20.6)	0%	(-10-13%)
Larynx	mean [Gy]	18.0	(8.7-33.9)	18.0	(8.7-33.7)	0%	(-1-0%)
Oral cavity	mean [Gy]	19.5	(6.8-32.7)	19.5	(6.8-32.7)	0%	(-1-0%)
Swallowing muscles	mean [Gy]	23.9	(13.6-34.4)	23.8	(13.1-34.7)	-1%	(-3-1%)
CI-low		2.23	(1.99-2.48)	2.20	(1.93-2.43)	-2%	(-3-0%)
CI-high		2.19	(1.98–2.41)	2.13	(1.98-2.31)	-2%	(-4-0%)
Beam directions		3	(3–3)	3	(3-3)	0%	(0-0%)
Energy layers		126	(107-144)	71	(52-100)	-45%	(-5630%)
Spots		2099	(1123-3314)	1809	(1050-2841)	-12%	(-187%)
Total spot weight	[Gp]	3899	(2046-6411)	3855	(2069-6298)	-1%	(-3-1%)
Delivery time (1 s)	[min]	3.9	(3.5-4.5)	3.0	(2.6-3.6)	-25%	(-3018%)
Delivery time (2 s)	[min]	6.0	(5.3-6.9)	4.2	(3.4-5.3)	-32%	(-3922%)
Delivery time (5 s)	[min]	12.3	(10.6–14.1)	7.7	(6.1–10.3)	-38%	(-4826%)
			Prostate cases (n:	=5)			
Parameter		Standard c		•	cient plan	Relative o	difference
ITV-low	V _{95%} [%]	99.3	(99.1-99.4)	99.3	(99.1-99.4)	0%	(0-0%)
ITV-high	V _{95%} [%]	99.3	(98.7-99.7)	99.5	(98.8-99.8)	0%	(0-0%)
ITV-high	V _{107%} [%]	0.1	(0.0-0.1)	0.1	(0.0-0.3)	29%	(-67-190%)
Femur left	max [Gy]	50.9	(50.5-51.3)	50.8	(50.2-51.3)	0%	(-2-0%)
Femur right	max [Gy]	50.3	(48.7-51.0)	50.2	(47.8-51.1)	0%	(-2-1%)
Rectum	mean [Gy]	26.3	(16.3-38.9)	26.2	(16.3-38.9)	0%	(0-0%)
Bladder	mean [Gy]	39.2	(34.1-43.6)	39.2	(34.1-43.6)	0%	(0-0%)
Femur left	mean [Gy]	27.9	(20.9-31.8)	27.9	(20.9-31.8)	0%	(0-0%)
Femur right	mean [Gy]	30.9	(27.9-35.5)	30.9	(27.9-35.6)	0%	(0-0%)
CI-low		1.61	(1.55-1.69)	1.62	(1.56-1.71)	1%	(0-1%)
CI-high		1.54	(1.51–1.59)	1.56	(1.51–1.59)	1%	(-1-3%)
Beam directions		2	(2-2)	2	(2-2)	0%	(0-0%)
Energy layers		107	(101-112)	77	(72-83)	-28%	(-3225%)
Spots		3835	(3445-4360)	3649	(3069-4300)	-5%	(-15-8%)
Total spot weight	[Gp]	8326	(7867-9056)	8169	(7777-8839)	-2%	(-21%)
Delivery time (1 s)	[min]	3.4	(3.3-3.5)	2.9	(2.7-3.1)	-16%	(-1912%)
Delivery time (2 s)	[min]	5.2	(4.9-5.4)	4.2	(3.9-4.5)	-20%	(-2317%)

 $Abbreviations: \ CTV = clinical\ target\ volume;\ ITV = internal\ target\ volume;\ CTV/ITV-low = low-dose\ target\ volume;\ CTV/ITV-high = high-dose\ target\ volume;\ CI-low = low-dose\ conformity\ index;\ CI-high = high-dose\ conformity\ index;\ Gp = Giga = protons.$

6.3 Results

Table 6.1 lists the plan parameters of the standard clinical plans and the time-efficient plans for the oropharyngeal cases and the prostate cases. For both patient groups, the table shows that the worst-case dose parameters were nearly identical, indicating comparable robust plan quality for both types of treatment plans. The number of energy layers in the time-efficient treatment plans was on average reduced by 45% (range: 30-56%) for the oropharyngeal cases and by 28% (range: 25-32%) for the prostate cases, compared with the standard clinical plans. The corresponding delivery times per fraction are depicted in figure 6.2 as a function of the energy switching time. When assuming energy switching

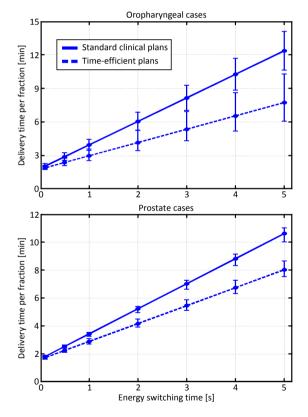


Figure 6.2: Average delivery time as a function of the energy switching time for standard clinical plans (solid lines) and time-efficient plans (dashed lines). Whiskers indicate minimum and maximum values.

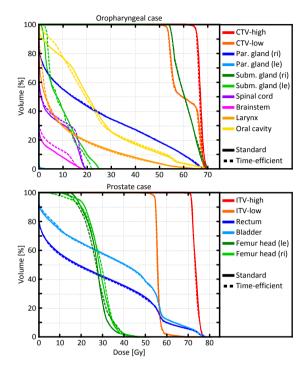


Figure 6.3: Worst-case dose-volume histograms for the standard clinical plan (solid lines) and the timeefficient plan (dashed lines) of one oropharyngeal case and one prostate case. Histograms depict the worstcase value for each dose-volume point.

times of 1, 2 or 5 s, the average delivery time of the oropharyngeal cases was shortened from 3.9 to 3.0 min (25%, range: 18-30%), from 6.0 to 4.2 min (32%, range: 22-39%) or from 12.3 to 7.7 min (38%, range: 26-48%). For the prostate cases, the delivery time was reduced from 3.4 to 2.9 min (16%, range: 12-19%), from 5.2 to 4.2 min (20%, range: 17-23%) or from 10.6 to 8.0 min (24%, range: 21-27%) on average, when assuming energy switching times of 1, 2 or 5 s, respectively. The (field-specific) plan parameters of each individual patient are provided in Appendix 6.B.

The worst-case dose-volume histograms (DVHs) of both treatment plans for one oropharyngeal case and one prostate case are displayed in figure 6.3. The worst-case DVHs were constructed by taking for each dose-volume point the worst value from the nine scenarios included in the optimization (nominal, 3 mm setup errors, 3% +1 mm range errors). The curves are generally located very close to each other, indicating a good agreement between the standard clinical plans and the time-efficient plans. Figure 6.4

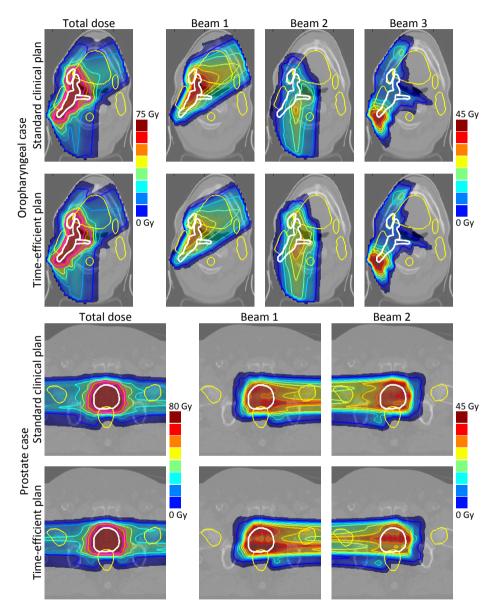


Figure 6.4: Dose distributions (total and per field) for the oropharyngeal and prostate cases depicted in figure 6.3. The CTV/ITV is delineated in white and organs-at-risk in yellow. The 95% prescription isodose lines (low-dose and high-dose) are indicated in pink.

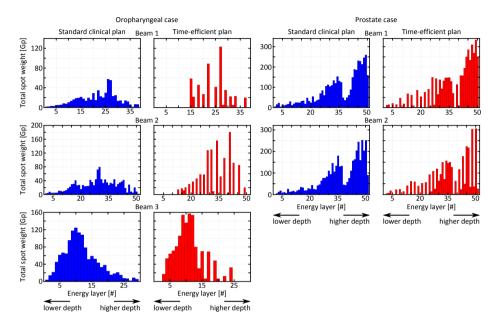


Figure 6.5: Total spot weight per energy layer for each beam direction of the standard clinical plan (blue) and the time-efficient plan (red) of the oropharyngeal and prostate cases depicted in figures 6.3 and 6.4.

shows for both cases the corresponding total dose distributions and the doses per field for the nominal scenario. The total dose distributions of the standard clinical plans and the time-efficient plans were very similar, especially for the oropharyngeal case. Larger differences can be observed in the dose distributions of the individual fields, illustrating the degeneracy of IMPT treatment plans. For the oropharyngeal case, the intensity of the individual fields could differ between both types of plans. For the prostate case, a change in the spatial distribution can be observed, indicating a redistribution of the spots within the remaining energy layers in the time-efficient plan.

The total spot weight per energy layer is depicted in figure 6.5 for the same cases as shown in figures 6.3 and 6.4. The energy layer reduction method generally resulted in a redistribution of spots from low-weighted energy layers to high-weighted energy layers, which was observed most strongly for the oropharyngeal case. The graphs show that the energy layers were not excluded evenly from all beam directions or uniformly across the proton energy range. Furthermore, the excluded low-weighted energy layers did not necessarily correspond to low proton energies, as illustrated by beam 3 of the oropharyngeal case.

6.4 Discussion

This study shows that the number of energy layers in robust IMPT treatment plans can be reduced considerably without affecting the dosimetric plan quality. This can result in a substantial shortening of the delivery time.

It should be noted that the delivery times reported in this study did not include the time required for patient setup. The relative gain in overall treatment efficiency will therefore depend on the patient setup time and on the number of treatment rooms that share the proton beam. In multi-room centers, patient setup can take place while patients in other treatment rooms are irradiated. If patient setup can be completed within the delivery time in other rooms, a shortening of the delivery time will reduce the beam waiting time and enable more efficient use of multiple treatment rooms. This can result in an increase in the number of treated patients and subsequently in a reduction of treatment costs. Regardless of the type of facility, delivery time reduction may also improve patient comfort and reduce treatment uncertainties, as intrafraction displacements of for example spine, prostate and liver tumors tend to increase with delivery time (21, 55, 62, 93).

Very recently, Cao et al. published an alternative method to reduce energy layers in IMPT treatment plans (94). They proposed a mixed-integer programming approach, iteratively reducing the number of energy layers until the plan quality degraded beyond a user-defined level. The number of energy layers in non-robust IMPT treatment plans was reduced by 14-19% for prostate cases, by 11% for a lung case and by 26% for a mesothelioma case, when allowing for the cost-function value to worsen by 5%. Our study shows that energy layer reduction can also be performed for robust IMPT treatment plans and without compromising dosimetric plan quality. The larger energy layer reduction observed in this study (25-56% vs. 11-26% by Cao et al.) might be explained by differences between both methods or by differences in patient and plan characteristics. More details on the performance of the different components of our method are provided in Appendix 6.C.

A reduction of the number of energy layers can also be achieved by an enlargement of the energy layer spacing, but this was shown to compromise plan quality (73, 86). We verified this for one oropharyngeal case by generating a standard clinical plan using an energy layer spacing increased by a factor of 2. The lateral spot spacing was reduced to keep the total number of optimized spots approximately the same. This resulted in an energy layer reduction (50%) which was comparable to the reduction in the time-efficient plan of this patient (52%). However, worst-case doses in the submandibular glands were increased by 7.1 Gy (+33%, mean dose), in the spinal cord by 3.3 Gy (+17%, maximum dose), in the larynx by 13.0 Gy (+149%, mean dose), in the oral cavity by 5.2 Gy (+30%,

mean dose) and in the swallowing muscles by 12.1 Gy (+68%, mean dose), compared with the time-efficient plan.

These results confirm that energy layers cannot be excluded before the start of treatment planning without compromising plan quality. Energy layer reduction should be incorporated in the planning process, as it makes use of the interchangeability of spots within and between the applied beam directions. This might also explain why fewer energy layers were excluded in the prostate cases. For these patients, it is more difficult to exchange spots between beam directions, as the pelvic lymph nodes on either side of the patient were mainly irradiated from a single beam direction to spare centrally located OARs. The effect of other patient and plan characteristics (e.g. tumor size, tissue heterogeneities, beam directions) on the performance of energy layer reduction should be investigated in future research.

The number of energy layers was incorporated in the mathematical optimization by means of the logarithm of the total spot weight per energy layer. However, the logarithm is a concave function and this can give problems during optimization. As a consequence, it required 16.7 hours on average to complete the fully automated energy layer reduction. It would therefore be useful to find a convex function to include the number of energy layers in the mathematical optimization.

6 5 Conclusions

The number of energy layers in robust IMPT treatment plans can be reduced considerably without affecting the dosimetric plan quality. The method presented in this study resulted in an average energy layer reduction of 45% and 28% for oropharyngeal and prostate cases, respectively. When assuming 1, 2 or 5 s energy switching time, the delivery time was on average shortened by 25%, 32% or 38% for the oropharyngeal cases, and by 16%, 20% or 24% for the prostate cases. Shorter delivery times are likely to reduce treatment uncertainties and costs.

Acknowledgments

The authors would like to thank Martijn Engelsman for providing valuable information regarding IMPT delivery.

chapter 7

Dose uncertainties in IMPT for oropharyngeal cancer in the presence of errors

International Journal of Radiation Oncology Biology Physics, Vol. 87, No. 5, pp. 888–896, 2013

> Aafke Kraan¹, Steven van de Water¹, David Teguh¹, Abrahim Al-Mamgani¹, Tom Madden², Hanne Kooy², Ben Heijmen¹, Mischa Hoogeman¹.

 ¹ Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands
 ² Department of Radiation Oncology, F H Burr Proton Therapy Center, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Abstract

Purpose: Setup, range, and anatomical uncertainties influence the dose delivered with intensity-modulated proton therapy (IMPT), but clinical quantification of these errors for oropharyngeal cancer is lacking. We quantified these factors and investigated treatment fidelity, that is, robustness, as influenced by adaptive planning and by applying more beam directions.

Methods and materials: We used an in-house treatment planning system with multicriteria optimization of pencil beam energies, directions, and weights to create treatment plans for 3-, 5-, and 7-beam directions for 10 oropharyngeal cancer patients. The dose prescription was a simultaneously integrated boost scheme, prescribing 66 Gy to primary tumor and positive neck levels (clinical target volume-66 Gy; CTV-66 Gy) and 54 Gy to elective neck levels (CTV-54 Gy). Doses were recalculated in 3700 simulations of setup, range, and anatomical uncertainties. Repeat computed tomography (CT) scans were used to evaluate an adaptive planning strategy using non-rigid registration for dose accumulation.

Results: For the recalculated 3-beam plans including all treatment uncertainty sources, only 69% (CTV-66 Gy) and 88% (CTV-54 Gy) of the simulations had a dose received by 98% of the target volume ($D_{98\%}$) > 95% of the prescription dose. Doses to organs-at-risk (OARs) showed considerable spread around planned values. Causes for major deviations were mixed. Adaptive planning based on repeat imaging positively affected dose delivery accuracy: in the presence of the other errors, percentages of treatments with $D_{98\%}$ > 95% increased to 96% (CTV-66 Gy) and 100% (CTV-54 Gy). Plans with more beam directions were not more robust.

Conclusions: For oropharyngeal cancer patients, treatment uncertainties can result in significant differences between planned and delivered IMPT doses. Given the mixed causes for major deviations, we advise repeat diagnostic CT scans during treatment, recalculation of the dose, and if required, adaptive planning to improve adequate IMPT dose delivery.

7.1 Introduction

Intensity-modulated proton therapy (IMPT) uses proton pencil beams with varying position and energy, whose intensities are optimized individually. Oropharyngeal tumors form an attractive site for IMPT because of their complex shape and proximity to organsat-risk (OARs). Improved OAR sparing is expected compared to intensity-modulated radiation therapy (IMRT) (75, 95), potentially resulting in reduced negative side effects. In IMPT, however, uncertainties may exist that cause the delivered dose to seriously deviate from planning and thus reduce the benefit of IMPT. These uncertainties include anatomical changes (e.g. organ motion, changes in air cavities, tumor regression, weight loss), patient setup errors and range uncertainties from uncertainties in computed tomography (CT) Hounsfield units (HU), conversion of HUs into stopping power and reconstruction artifacts (82–84, 96–98).

Various works discuss IMPT treatment uncertainties (82-84, 95, 98) and mitigation by robust optimization (81, 82, 88, 89, 99-103). Many studies focusing on dose impacts of treatment uncertainties (82-84, 98) show one or few patient cases, do not include oropharyngeal patients, or do not include combined effects of treatment uncertainties. Robust optimization studies (81, 82, 88, 89, 99-103) typically present results obtained with robust treatment plans (i.e. to confirm the effectiveness of a particularly robust algorithm). In this context, for head-and-neck cancer patients, a recent article by Liu et al. (99) focused on differences between various robust optimization strategies to compensate range and patient setup errors. However, the impact of anatomical changes and realistic simulation of all treatment-related uncertainties was not presented. Simone et al. (95) compared adaptive IMRT with IMPT for head-and-neck cancer patients but did not investigate the impact of range and setup errors. Despite the concerns for robustness of IMPT for head-and-neck cancer and the interest to treat head-and-neck cancer with IMPT, a realistic and comprehensive overview of the effects of treatment uncertainties on the delivered dose is still lacking. Furthermore, such studies are needed, a priori, to assess the required level of robustness mitigation in clinical practice.

This study presents a realistic and accurate (3700 simulations from 10 oropharyngeal cancer patients) analysis of the effects of treatment uncertainties on the delivered dose. First, we studied the impact of anatomy, range and setup errors, separately and combined, by simulating realistic treatments. Moreover, we evaluated the effect of range and setup errors of different magnitudes. Results for target and OARs are readily interpreted in terms of dose-volume histogram (DVH) parameters. Second, we investigated 2 approaches, not yet reported, to make treatments robust and applicable independent of robust optimization by posing the questions: (1) how is treatment precision affected by adaptive planning i.e. replanning based on repeated imaging, as recommended in IMRT (104)?; and (2) does plan robustness improve by choosing more beam directions, as postulated by Unkelbach *et al.* (82)?

7.2 Methods and materials

Patient group

We included 10 prospectively selected oropharyngeal cancer patients treated at Erasmus MC Cancer Institute in 2004 and 2005, aged 48 to 83 years old (mean, 60 years old) (table 7.1). Each patient had a planning and a repeat CT scan, the latter at 46 Gy. Scans were enhanced by administration of intravenous contrast (100 ml of Omnipaque, 647 mg of iohexol per ml, see Discussion). The primary tumor was delineated by a radiation oncologist on each CT. Contouring of OARs was assisted by atlas-based autosegmentation (106) and edited if needed. Artifact reduction software (107) was used for 5 patients with metal teeth implants. The dose prescription was a simultaneously integrated boost scheme, prescribing 66 Gy to the primary tumor and positive neck levels (CTV-66 Gy) and 54 Gy to elective neck levels (CTV-54 Gy) in 30 IMPT fractions. CTV volumes are shown in table 7.1. We defined the planning target volumes (PTV)-66 Gy and PTV-54 Gy as the CTV-66 Gy and CTV-54 Gy with 5-mm margins. Being aware that margin recipes are not fully valid for IMPT, we recall that this would be sufficient in IMRT to account for setup errors of 1.5 mm standard deviation on systematic and random components (75, 108). We quantified the amount of anatomical change between the 2 CTs as the average deformation of salivary glands and primary tumor measured in 6 directions, as suggested by Vasquez-Osorio *et al.* (105) (table 7.1, last column).

Table 7.1: Summary of patient characteristics.

Patient	Site	TNM staging	Planning CT: Volume CTV-66 Gy (cm ³)	Planning CT: Volume CTV-54 Gy (cm ³)	Repeat CT: Volume CTV-66 Gy (cm ³)	Repeat CT: Volume CTV-54 Gy (cm ³)	Average deformation* (mm)
1	Base of tongue	T1N2c	106	199	82	176	4.2
2	Base of tongue	T3N2a	99	313	73	268	3.3
3	Tonsil	T2N1	43	165	37	156	2.5
4	Tonsil	T2N0	11	77	9	75	3.1
5	Soft palate	T2N0	14	72	10	67	2.4
6	Base of tongue	T3N2a	68	221	47	191	6.6
7	Tonsil	T2N0	5	67	5	72	2.3
8	Tonsil	T1N1	41	95	35	87	2.5
9	Base of tongue	T3N3	178	343	132	296	6.6
10	Base of tongue	T1N2c	70	294	63	252	4.5
Patient average			63	185	49	164	3.8

Abbreviations: CT = computed tomography; CTV-66 Gy = Clinical Target Volume receiving 66 Gy (primary tumor and positive neck levels); CTV-54 Gy = Clinical Target Volume receiving 54 Gy (elective neck levels).

Average deformation was measured in 6 directions as suggested by Vasquez Osorio et al. (105).

Proton delivery system

We used the pencil beam scanning techniques from Hong et al. (109) and Koov et al. (23). The proton energy ranged from 45 MeV to 230 MeV. Range shifters were applied when needed for lower energies. The distance between energy layers was 2 mm in water. The full width at half the maximum (FWHM) of the pencil beam ranged from 9 mm at 230 MeV to 21 mm at 45 MeV (numbers for 5 cm of air excluding range shifters (23, 109)).

Planning technique

Treatment planning used an in-house developed IMPT treatment planning system. The proton dose calculation algorithm from Hong et al. (109) and Kooy et al. (23) was integrated in an automated plan generation platform called 'Erasmus-iCycle' (66, 110). It performed multi-criteria optimization, which optimized objectives sequentially according to a list of prioritized objectives and hard constraints, identical for all patients. The minimal PTV dose was a hard constraint, while prioritized objectives included (in order of priority) minimizing maximum PTV dose, target conformality, and minimization of dose to all OARs and unspecified tissues (see Appendix 7.A for details). We applied a newly developed method of iterative resampling of randomly placed pencil beams (86). The dose grid resolution was $3 \times 3 \times 3 \text{ mm}^3$.

For each patient, treatment plans with 3, 5, and 7 beam directions were created for planning and repeat CT. The 3-beam plan is the basis against which dose deviations and adaptive planning strategy were quantified and has coplanar beams at -50°, 50°, and 180° as published previously (75). We generated plans with 5 and 7 beam directions with uniformly spaced angles, starting at o°, to investigate robustness versus more beams. Examples of dose distributions are shown in figure 7.1.

Simulation and quantification of dose deviations

Anatomical uncertainties

These were modeled by calculating the treatment plan dose for fractions 1 to 15 on the planning CT and for fractions 16 to 30 on the repeat CT. Dose accumulation for OARs and target, both subject to deformation, of the first and second part of the treatment was performed with an in-house developed non-rigid registration method (105).

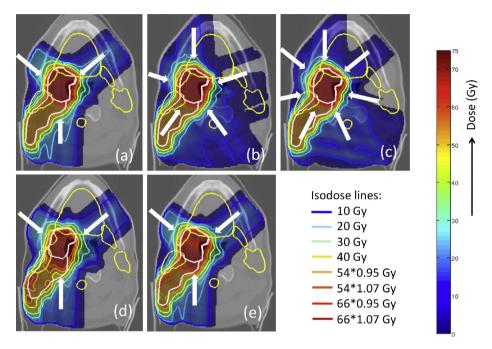


Figure 7.1: Dose distributions for patient 4. (Top) planned dose for 3 (a), 5 (b), and 7 (c) beam directions. (Bottom, d) calculated dose delivered to repeat CT when treated with original 3-beam plan, and (e) delivered dose when applying a 4-mm posterior isocenter shift to the original CT. Target is shown in white; organs-at-risk (cord, parotids, oral cavity) are in yellow.

Range uncertainties

These were simulated by recalculating the treatment plans on a CT scan with all CT numbers systematically increased or decreased by a given percentage of its value, as proposed by Lomax et al. (83). We investigated shifts in CT numbers ranging from -5% to +5% with steps of 1%. Negative (positive) values represent an overshoot (undershoot) of the Bragg peak.

Setup uncertainties

We simulated an online patient setup correction protocol with bony anatomy as reference. Ideally, setup errors are zero after applying online corrections. Residual errors, however, remain from intrafraction motion or uncorrected deformations of the bony structures (97). These were simulated by applying rigid isocenter shifts in anteriorposterior (AP), superior-inferior (SI), and right-left (RL) lateral directions. For each patient, we simulated 40 treatments of 30 fractions each. For each simulated treatment, we did the following:

- A systematic shift $(\delta_{AP}, \delta_{SI}, \delta_{RI})$ was applied, fixed for all fractions. The size of each δ (AP, SI, or RL) was determined randomly and independently from a normal distribution with standard deviation Σ (108).
- A random shift $(\Delta_{AP}, \Delta_{SI}, \Delta_{RL})$ was applied, varying for each fraction. The size of each Δ was determined randomly from a normal distribution with standard deviation σ (108). We assumed that $\sigma = \Sigma$ (108).
- The total isocenter shift for each fraction was determined, which was the vector sum of the above-described 2 contributions; and a dose matrix was calculated for each fraction.
- The resulting 30 dose matrices of all fractions were added to obtain the total dose for 1 treatment.

This was repeated 40 times, yielding 40 dose matrices per patient, representing 40 treatments. We repeated this procedure for $\Sigma = \sigma = 1.0, 1.5, 2.0, 2.5, \text{ and } 3.0 \text{ mm}.$

All uncertainties combined

We investigated the combined effect of all errors by including them simultaneously. We included a range shift of -3% (83) and setup errors $\Sigma = \sigma = 1.5$ mm (both following the procedure above for both planning and repeat CT), corresponding to the average residual setup error from bony anatomy deformation (97). We applied the treatment plan to the planning CT for fractions 1 to 15 and to the repeat CT for fractions 16 to 30. Nonrigid registration was used for dose accumulation between the CTs. We simulated 40 treatments of 30 fractions, yielding 40 dose distributions per patient.

Improving treatment quality

Adaptive planning

We investigated the effect of adaptive planning (based on repeat imaging) on treatment quality by repeating the simulations and including all errors as described above, but now applying for fractions 16 to 30 a newly created plan based on the repeat CT. Forty treatments were simulated per patient.

Increased beam directions

Analogue to the 3-beam plans, we simulated 40 treatments that included all errors for 5 and 7 beam directions (see section "All uncertainties combined" above) to investigate whether plan robustness improved when applying more beam directions.

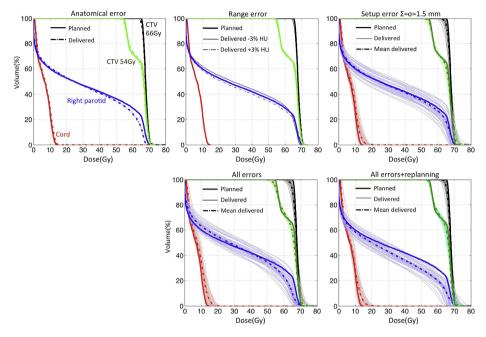


Figure 7.2: Example of dose-volume histograms of patient 1 with various errors included for 5 sets of simulations. For visualization, we show only the target regions and 2 organs-at-risk.

In summary, for each patient we performed 291 treatment simulations for 3-beam directions (nominal situation: 1 simulation; range error: 9 simulations; anatomical uncertainty: 1 simulation, setup error: $40 \times 5 = 200$ simulations; combined errors: 40 simulations; adaptive plan: 40 simulations), 40 simulations for 5-beams, and 40 simulations for 7-beams. For 10 patients this resulted in a total of 3710 treatment simulations.

Evaluation

Plan evaluation was based on DVHs following ICRU recommendations (111) (figure 7.2). For CTV-66 Gy and CTV-54 Gy, we considered the percentages of the target volumes receiving at least 95% and 107% of the prescribed dose ($V_{95\%}$ and $V_{107\%}$, respectively) (111) and dose received by 98% of the target volume (D_{98%}) (111). For cord and brainstem, we report dose to 2% of the volume (D_{2%}) for near-maximum dose (111), while for other OARs, we report mean doses (111). Statistical analysis of the differences

among the plans for 3, 5, and 7 beam directions used the ANOVA test (p-value < 0.05 was considered significant).

7.3 Results

Below we report the dose deviations in our simulated treatments and corroborate our findings with the treatment intent that 90% of the patient population should have $D_{98\%} \ge 95\%$ (108). We also highlight a few large dose deviations.

Ouantification of dose deviations

Anatomical errors

Figure 7.1d shows an example of the dose degradation that resulted from applying the treatment plan from the planning CT to the repeat CT. Figure 7.3 summarizes dose effects for all simulations. For the target (figure 7.3a, 'anatomical error'), the effect of anatomical errors was clearly visible in the $D_{98\%}$. The increase in target hot spots was small, except for patient 9 ($V_{107\%} = 14\%$, see Appendix 7.B). Anatomical errors had a negative impact, but $D_{98\%}$ for all patients was >95%. For CTV-66 Gy and CTV-54 Gy, the average $D_{98\%}$ reduction was -1.9% and -2.1%, respectively. For CTV-66 Gy, the $V_{107\%}$ increase due to anatomical errors was correlated with its size (Pearson correlation, 0.76; p = 0.011) and a trend toward a moderate correlation was observed for the average deformation (Pearson correlation, 0.64; p = 0.058). The biggest increase in OAR dose (figure 7.3b, "anatomical error") was in patient 6, with a 9.2 Gy increase in the brainstem (see Appendix 7.C).

Range error

Figure 7.4 (left panels) shows the impact of range errors. For CTV-66 Gy, we note that for range errors within 4%, more than 90% of the patient population has $D_{98\%} \ge 95\%$, whereas for CTV-54 Gy, even larger range errors do not cause a violation. Figure 7.3 shows the impact of a -3% range error on target (figure 7.3a) and OARs (figure 7.3b). For CTV-66 Gy and CTV-54 Gy the average $D_{98\%}$ reductions were -2.0% and -2.2%, respectively. The average increase for the 10 patients in $D_{2\%}$ (for cord, brainstem) and mean dose (for other OARs) was less than 1 Gy for all OARs.

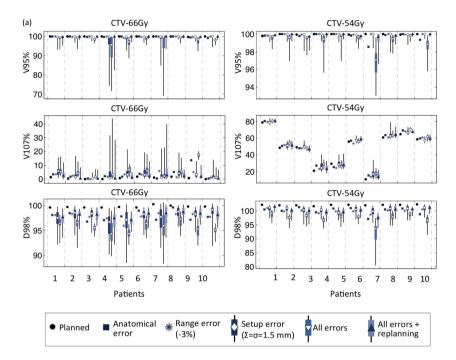


Figure 7.3a: For the 10 patients, impact of various uncertainties on target dose are shown. Solid markers represent medians, boxes are spread between the 25th and 75th percentiles, and lines extend to the minimum and maximum of the 40 simulations. Because CTV-54 Gy contains CTV-66 Gy region, the $V_{107\%}$ and D_{98%} of CTV-54 Gy are naturally high for all patients.

Setup error

Figures 7.1a and 7.1e demonstrate the impact of a 4-mm rigid isocenter shift on delivered dose. Figure 7.4 (right panels) shows the effects of setup errors. All parameters show an increase in target dose deterioration with increasing setup errors. At $\Sigma = \sigma = 2$ mm (CTV-66 Gy), the treatment intent that 90% of the population have $D_{98\%} > 95\%$, was not satisfied anymore. Figure 7.3a ('setup error') summarizes the dose effects for $\Sigma = \sigma$ = 1.5 mm for target. For both CTVs, the median reduction in $D_{98\%}$ was small, although with a large spread around the median. The largest $D_{98\%}$ spread in CTV-66 Gy was seen in patient 4 with a relatively small tumor. For OARs (figure 7.3b), we quantified the effect of the setup error $\Sigma = \sigma = 1.5$ mm by calculating the average, standard deviation, and maximum values across all simulations of the delivered minus the planned maximum dose (cord, brainstem) or mean dose (other OARs). The average dose increase across all

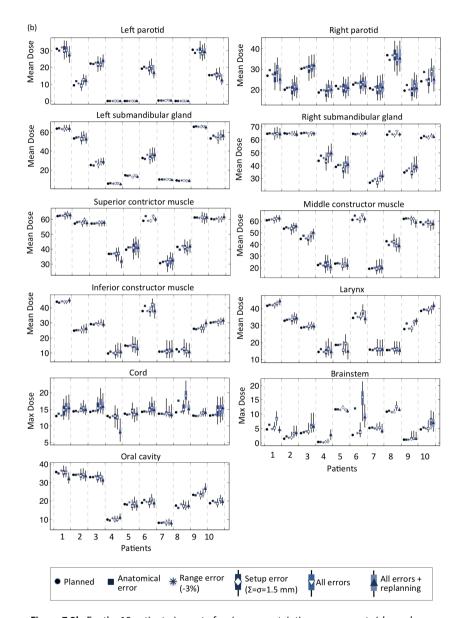


Figure 7.3b: For the 10 patients, impact of various uncertainties on organs-at-risk are shown.

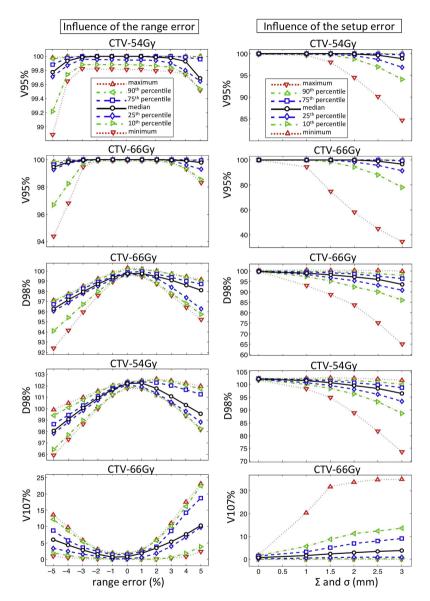


Figure 7.4: Influence of range (left) and setup (right) uncertainties on target coverage. For the range error, maximum, minimum, and percentiles refer to the statistics of the group of 10 patients. For the setup error, they refer to 400 simulations (40 simulations x 10 patients). Planned values correspond to an error of zero.

Patient	All e	rrors	Replanning		
	CTV-66 Gy	CTV-54 Gy	CTV-66 Gy	CTV-54 Gy	
1	88%	100%	92%	100%	
2	90%	95%	100%	100%	
3	75%	98%	100%	100%	
4	40%	90%	85%	100%	
5	28%	95%	98%	100%	
6	82%	100%	100%	100%	
7	68%	20%	90%	98%	
8	95%	98%	100%	100%	
9	60%	100%	100%	100%	
10	62%	82%	100%	100%	
All patients	69%	88%	96%	100%	

Table 7.2: Percentage of simulations satisfying $D_{98\%} \ge 95\%$ (or $V_{95\%} \ge 98\%$).

simulations was less than 1 Gy for each OAR, although some large variations between individual treatments were observed (figure 7.3a).

All errors combined

Figures 7.3a and 7.3b show that the combination of errors can lead to inferior treatments. For CTV-66 Gy and CTV-54 Gy, percentages of simulations with $D_{98\%} > 95\%$ were only 69% and 88%, respectively (table 7.2), which is clinically unsatisfying. Hot spots increase: the average increase in $V_{107\%}$ (i.e. delivered minus planned $V_{107\%}$) across all simulations was 5.1% in $V_{107\%}$ (CTV-66 Gy) and 2.9% (CTV-54 Gy). For OARs, we found average dose increases below 1 Gy but again large variations among treatments were present. The largest variation was in the right parotid, with a standard deviation in mean dose difference of 3.6 Gy and a maximum difference of 11.3 Gy (see Appendix 7.D). The spinal cord showed a standard deviation in the maximum dose difference of 2.2 Gy and a maximum difference of 9.5 Gy (see Appendix 7.E).

Treatment quality improvement

Adaptive planning

Figure 7.3 and table 7.2 demonstrate how adaptive planning positively affects dose delivery. Target coverage showed a clear increase. For CTV-66 Gy and CTV-54 Gy, 96% and 100%, respectively, of simulations had a $D_{98\%} > 95\%$, now safely complying with the treatment intent that 90% of the population has $D_{98\%} > 95\%$ (table 7.2). Target hot spots decreased compared to the scenario were no adaptive planning was applied: the delivered minus planned $V_{107\%}$ was now only 1.5% (CTV-66 Gy) and 2.0% (CTV-54 Gy).

For OARs, the dose spread was reduced. For patient 4, the spinal cord dose was reduced greatly due to the lower cord dose in the repeat CT plan.

More beam directions

Appendix 7.F gives dose effects for plans with 3, 5, and 7 beam directions. As shown in Appendix figure 7.F.1, plans with 5 or 7 beam directions were not significantly more robust than plans with 3 beam directions.

7.4 Discussion

An important observation is that the errors individually do not lead to serious target underdosage ($D_{98\%} < 95\%$) but that combined effects of the errors generate serious problems and cause the treatment intent ($D_{98\%} \ge 95\%$ for at least 90% of the patient population (108)) not to be fulfilled. We recommend for IMPT studies that simulations combine all uncertainties to study site-specific clinical robustness requirements.

Concerning anatomical deformations, 1 of the 10 patients had a significant increase in hot spots. This patient had substantial target volume shrinkage (178 to 132 ml for CTV-66 Gy, table 7.1), deformation (6.6 mm, table 7.1), and weight loss, resulting in incorrect spatial pencil beam matching in the target (see Appendix 7.B for details). Patient 6, with similar deformation, did not show serious target hot spots but instead showed a large dose increase to the brainstem (see Appendix 7.C). Small changes in $V_{107\%}$ were seen in patients 3, 5, 7, and 8, with small CTV volumes and small deformations (2.3-2.5 mm). Factors such as initial tumor size can help to estimate which patients risk large dose deviations, as shown by the significant correlation between the increase in V_{107%} and the tumor size. On the other hand, only a moderate correlation was found between increase in V₁₀₇% and deformation. Much higher patient statistics and a thorough classification of anatomy changes (tumor shrinkage, weight loss, swelling, and others) are needed to identify patients at risk. In the absence of accurate prognostic factors, repeat imaging, dose recalculation and, if required, adaptive planning are recommended to ensure sufficient target coverage and to avoid unwanted exposure of OARs. The timing and frequency of repeated imaging are subjects for future research.

We found no dramatic dose modifications when studying range errors separately. This is a significant clinical observation, as range errors are considered one of the primary concerns in robust optimization algorithms. A different spot placement technique could alter this, as shown by Lomax *et al.* (83).

In the simulations that include setup errors, median dose deviations are small, but

the dose spread is large both for OARs and target. Dose deteriorations in the target are caused by incorrect spatial matching of pencil beams, resulting in target hot spots, and by target misalignment (figures 7.1e and 7.3a). The impact on $V_{95\%}$ is seen to be largest for small tumors (table 7.1, patients 4 and 7), which is explained by the larger relative impact of a misalignment on these small volumes. The 5-mm CTV margin helps to account for target misalignment but cannot fully avoid dose impacts. Robust optimization could reduce the contribution of this error source (99). In addition, based on figure 7.4, it is highly recommended that the systematic and random setup errors (Σ and σ , respectively) are kept below 2 mm, for instance by the use of online image guidance. This would also work to deal with problems such as a different position of the neck at planning and during treatment.

Concerning combined errors, the simulations (figure 7.3) showed that the combined occurrence of errors can amplify negative dose effects. We saw for patient 6 a large dose increase in the brainstem (see Appendix 7.C), where the combined effect of the errors was larger than the sum of the effects separately (figure 7.3b). This was also observed for the spinal cord of patient 8 (see Appendix 7.E). Given the mixed causes for deviations, dose distortions are very difficult to predict without recalculating the dose with uncertainties included.

Treatment accuracy can be substantially improved by applying adaptive planning, increasing the percentage of treatments with $D_{98\%} > 95\%$ safely to above 90%. Because adaptive planning increases staff workload and costs, the optimal adaptive strategy would have to be defined. Treatment plans with more beam directions were not more robust, in contrast to what was suggested by by Unkelbach *et al.* (82). We suggest that the use of more beam directions does not necessarily solve dose deteriorations resulting from pencil beam mismatching or anatomical changes. We are aware that the choice of uniformly spaced beam angles may not be optimal and that beam-angle optimization (112) may result in a more optimal choice, which could alter our observation. Other ways to improve treatment robustness not investigated here are to use wider pencil beams (however, this is expected to worsen OAR sparing (76)) or to apply robust optimization to dampen the effect of range and patient setup errors (99).

A limitation of this study is that intrafraction setup errors (i.e. changes in patient position during the fraction) were not modeled separately and dose calculation errors were not included. Both are expected to have only minor dose impact (62, 83). We verified that the contrast fluid in our CTs, which led to a slight increase in inhomogeneities and thereby possibly influenced the simulations, did not alter our results.

7.5 Conclusions

Based on 3700 simulated treatments for oropharyngeal cancer patients, we quantified the dose differences between planned and delivered IMPT doses in targets and OARs. Without action against treatment uncertainties, the treatment intent ($D_{98\%} \geq 95\%$ for at least 90% of the patient population) was not fulfilled. Given the mixed causes for major deviations observed, we advise acquisition of repeat CT scans and dose recalculation to properly assess delivered dose. If required, adaptive planning is effective for mitigating the effect of treatment-related uncertainties. Applying more than 3 beam directions did not increase plan robustness.

chapter 8

Tumor tracking with scanned proton beams: assessing the accuracy and practicalities

Physics in Medicine and Biology, Vol. 54, pp 6549-6563, 2009

Steven van de Water^{1,2}, Rob Kreuger², Silvan Zenklusen¹, Eugen Hug¹, Tony Lomax¹.

¹ Center for Proton Radiation Therapy,
 Paul Scherrer Institute, Villigen-PSI, Switzerland
 ² Faculty of Applied Sciences, Section Radiation Detection and Medical Imaging, Delft University of Technology, Delft, the Netherlands

Abstract

The potential of tumor tracking for active spot-scanned proton therapy was assessed. Using 4D dose calculation and simulated target motion, a tumor tracking algorithm has been implemented and applied to a simple target volume in both homogenous and heterogeneous in-silico phantoms. For tracking and retracking (a hybrid solution combining tumor tracking and rescanning), three tracking modes were analyzed: 'no tracking' (uncorrected irradiation of a moving target), 'perfect tracking' (no time delays and exact knowledge of target position) and 'imperfect tracking' (simulated time delays or position prediction errors). For all plans, dose homogeneity in the target volume was assessed as the difference between $D_{5\%}$ and $D_{95\%}$ in the CTV. For the homogeneous phantom, perfect tracking could retrieve nominal dose homogeneity for all motion phases and amplitudes while severe deterioration of treatment outcomes was found for imperfect tracking. The use of retracking reduced the sensitivity to position errors significantly in the homogeneous phantom. In the heterogeneous phantoms (simulated rib proximal to target), the nominal dose homogeneity could not be obtained with perfect tracking. Adjustments in pencil beam positions could cause pencil beams to deform under the influence of the bone, resulting in loss of dose homogeneity. As retracking was not capable of reducing these effects, rescanning provided the best treatment outcomes for moving heterogeneous targets in this study.

8.1 Introduction

Proton therapy using scanned beams is a fast maturing technology (22, 113). Although predominantly still delivered in the research environment, more than 1000 patients have now been treated worldwide with either scanned proton or carbon ion beams. Initial clinical results are extremely encouraging (16, 17, 114), and based on this, many new and planned particle therapy facilities are currently working on, or have recently started, using scanned proton beams clinically. In contrast to passive scattering, the traditional approach to delivering particle beams for therapy, scanning generally allows for better dose conformation and reduced dose to all surrounding normal tissues (115). In addition, it provides the flexibility to deliver so-called intensity-modulated proton therapy (12, 13), which provides the power and flexibility similar to that provided through the introduction of IMRT techniques in photon therapy. Consequently, many believe that scanning is the future of particle therapy delivery.

However, as with any highly dynamic delivery technique, scanned particle beam therapy is extremely sensitive to organ motion. With this approach, narrow pencil beams of particles (typically only a few millimetres in width) are scanned across the target using scanning magnets and, in some cases, a motion of the treatment table (22). In addition, the depth of the Bragg peak is varied (116), either by varying the energy of the particle beam at a source (if using a synchrotron) or through the insertion of varying thicknesses of a preabsorber material (if using a cyclotron). Either way, any motion of the target or patient can significantly affect the form of the delivered dose, through the interplay of the motion of the beam with the motion of the target or patient (the so-called interplay effect). The sensitivity of scanned particle beams to motion was already pointed out by Phillips et al. (117), who showed that, with only a 2.5 mm motion of one pencil beam in relation to another (or equivalently, a motion of the target by 2.5 mm between the delivery of the two pencil beams), localized dose increases or decreases of up to 20% could be expected (117). This effect was subsequently verified experimentally at PSI using a moving table to simulate motion (118). Much work has also been done by the GSI group for carbon ion scanned beams, again confirming the deleterious effects of organ motion and the interplay effect on scanned particle beams (119). Due to these effects, scanned beams have up to now primarily been employed against static or quasi-static tumors, particularly tumors attached to bony or rigid soft tissue structures or in the cranium.

Similar interplay effects have also been studied for IMRT (120), and a number of solutions proposed, including gating (121) and tracking (122, 123). Similar efforts are now being made for scanned particle delivery, with some groups exploring the potential of tumor tracking using scanned carbon ion beams (124–127), and others looking into rescanning and gating (119, 120, 128). Indeed, in principle, using a scanned particle beam is ideally suited to tumor tracking, in that, once one knows where the tumor is or how it is

moving, the steering elements of the scanning delivery hardware are more than capable of being adjusted to 'follow' the tumor, with typical scanning times of 5 ms mm⁻¹, which is generally much faster than any motions in a patient due to breathing. However, there are nevertheless some question marks over the efficacy of tracking, given the highly sensitive nature of scanning to small spatial errors (117), the difficulty in obtaining a reliable signal of the tumor's whereabouts as it moves and lastly, the fact that, even if the location of the tumor is exactly known, there will be an inevitable delay between obtaining this signal, analyzing it and adjusting the scanning parameters or trajectory to adjust for this motion.

The goal of this study was to assess by simulation the influence of such tracking uncertainties on the treatment outcome for scanned proton beam therapy in both homogeneous and heterogeneous volumes. In addition, we introduce the concept of 'retracking', whereby the advantages of tracking are combined with the robustness of rescanning in an attempt to reduce the sensitivity of pure tracking to the inevitable uncertainties resulting from the tracking process.

8.7 Methods and materials

Gantry

The simulations were performed for treatments with the gantry 2 of PSI (79), which is currently being developed at PSI as the successor of gantry 1 (22). The coordinate system that was used to describe the position of the spot position relative to the gantry is displayed in figure 8.1. Gantry 2 is equipped with two sweeper magnets to perform fast beam translations in the t- and u-direction and the spot position can be adjusted continuously in depth (s-direction) with an energy degrader. The energy degrader consists of carbon wedges that can be inserted into the beam line. During beam position adjustments, the beam is switched off. For a spot translation of 5 mm, the adjustment times are approximately 150 ms * , 4 ms and 5 ms for the s-, t- and u-direction respectively. In order to minimize the total treatment time, the t-direction is set to be the primary scanning direction followed by the *u*-direction as the secondary scanning direction and

Since this work was performed, we have found that our new gantry 2 will be able to change energy at a rate of about 80 ms per 5 mm range step. We do not believe that this changes the validity of our results or conclusions, however.

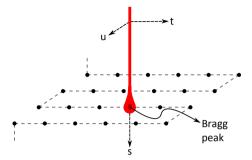


Figure 8.1: Illustration of the gantry/beam coordinate system. The black spots indicate the planned spot positions and the dotted line shows the order in which they are applied.

the s-direction as the tertiary scanning direction. This t-u-s scanning order is illustrated in figure 8.1. For all simulations, discrete (rather than continuous) scanning has been assumed.

Beam sizes in air used for the simulations were those of our existing treatment gantry at PSI, which are about 3 mm sigma. However, the actual beam size entering the phantom is a rather complex function of energy (determined by the insertion of different numbers of range shifters (22)) and the air gap between the last range shifter and the patient. A previous analysis of beam size calculated over many fields and patients has shown that, on average, the beam size at the Bragg peak is about 8 mm sigma (129), and it can be assumed that that is the average beam width at the Bragg peak for the plans analyzed in this work.

Target

Simulations were performed with simple mathematical phantoms. The homogeneous phantom consisted of a square box of water with a $30 \times 40 \times 50$ mm rectangular clinical target volume (CTV) located in the centre. For the planning target volume (PTV), a margin of 5 mm around the CTV was used. A heterogeneous phantom was constructed from the homogeneous phantom by modelling a single bone (elliptical in shape, with a maximum thickness of 40 mm) in the proximal region of the target. When a tumor moves relative to a bone, adjustments of the spot position in depth are only necessary when changes in the thickness of the bone (or boundaries) are present in the direction

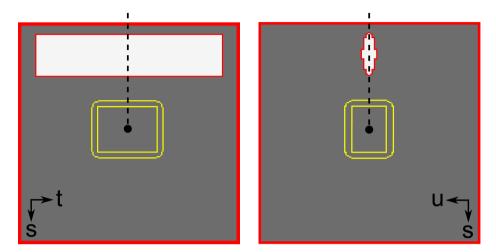


Figure 8.2: Two views of the heterogeneous phantom for target motion in the *u*-direction. The phantom is a square box of water (grey) with the CTV and PTV outlined in yellow. The bone is illustrated in white. The dotted line indicates the beam orientation.

of the target motion. To isolate the effect of the bone on tracking, the bone was modelled such that it had only changes in thickness and boundaries in the motion direction. Perpendicular to the direction of target motion, the thickness of the bone was constant and it fully overlapped the PTV as seen from BEV perspective. This approach required two heterogeneous phantoms for target motion in the t- and u-direction, with the bone rotated over 90°. In figure 8.2, the heterogeneous phantom for target motion in the udirection is depicted. The gantry and couch angles were both set to o° for all simulations, meaning that the incidence of the pencil beams was perpendicular to the upper side of the phantom.

For all simulations, a fixed CTV-to-PTV margin of 5 mm was used. Although ideally, this margin would be customized depending on the motion management method used, it has been deliberately left unchanged here in order to not distort the results by introducing another variable (i.e. a variable PTV margin depending on the motion and/or accuracy of the tracking). By adopting a single PTV for all cases, we wanted to study what would happen to CTV coverage for cases where perfect tracking was assumed but cannot be perfectly applied, i.e. under conditions of different tracking inaccuracies. In addition, as the same PTV was used for all simulations, it enables a more direct comparison of the results between these techniques.

Target motion

To simulate respiratory motion, the target was given uniform displacements in either tor u-direction according to the \cos^4 -function given in equation 8.1. The \cos^4 is a wellknown mathematical simulation of respiratory motion (130, 131):

$$dx(T) = A\left(-\frac{3}{8} + \cos^4\left(\frac{\pi}{T_0} - \varphi\right)\right) \tag{8.1}$$

with A the amplitude (peak-to-peak) (mm), T_o the period of respiration (s) and φ the starting phase.

The breathing period was set to be 5 s and the amplitude of the motion (peak-topeak) was varied from 0 to 15 mm. The phase of the motion at the start of the treatment was varied from 0 to 2π . The cos⁴-function was given a negative offset equal to its average, to make the average displacement equal to o during a breathing period.

In order to simulate the PTV moving relative to the bone, a partial motion field was used for the heterogeneous targets. The proximal part of the phantom containing the bone was kept static whereas the distal part containing the PTV was ascribed motion. This resulted in a discontinuity in the motion field and in the dose distribution as well. However, it did not affect the dose distribution within the target. Moreover, discontinuous motion fields can occur in real life, for example with the liver, which slides along the abdominal wall (132).

Motion in depth (s-direction) and non-uniform (non-rigid) motion were not simulated to avoid problems with compression of tissue and to keep the dose calculations simple.

Treatment plan

The treatment plans for the phantoms were made by PSIplan, the in-house developed treatment planning system for static treatments with gantry 1. A regular spot distribution within the target was made, and by inverse optimization, the optimal beam weights for each applied spot were determined. The spot spacing was 5 mm in the t-direction, 4 mm in the u-direction and 4.6 mm water equivalent distance in the s-direction (thickness of the range shifter plates in gantry 1). The treatment plans for the phantoms used in this study contained approximately 2100 spots. For a prescription dose of 0.6 Gy (equivalent to the dose of a single field of a three field plan), the treatment time was approximately 20.1 s of which roughly half was beam-on time and half was beam adjustment time.

Dose calculation

To simulate the effects of motion, a 4D dose calculation algorithm has been used to simulate the dose deposition in moving targets. It is based on a simple dose relation for a proton pencil beam in a static target. The dose d in a dose grid point at $(s_{dgp}, t_{dgp}, u_{dgp})$ delivered by a pencil beam with its spot position at $(s_{spot}, t_{spot}, u_{spot})$ can be calculated with equation 8.2:

$$d = \frac{D_w}{2\pi\sigma_w^2} e^{\frac{-(t_{\text{spot}} - t_{\text{dgp}})^2}{2\sigma_w^2}} e^{\frac{-(u_{\text{spot}} - u_{\text{dgp}})^2}{2\sigma_w^2}} e^{\frac{-(u_{\text{spot}} - u_{\text{dgp}})^2}{2\sigma_w^2}}$$
(8.2)

with D_w the integral dose at a radiological depth w (Gy mm²) and σ_w the beam width at a radiological depth w (mm).

The radiological depth w is calculated by taking the density integral along a straight line from the dose grid point to the source (gantry nozzle) and dividing it by the density of water. The integral dose D_w at w can be derived from a standard depth-dose curve for 250 MeV protons in water, taking into account the beam weight and range shifter settings, whereas the beam widths are calculated taking into account scatter in the patient and range shifters, plus the additional effect of the air gap between the range shifter and the patient surface. This dose calculation is repeated for all pencil beams and for all dose grid points within 3σ of each beam. Density heterogeneities are taken into account using a ray casting approach, as described in detail by Schaffner et al. (133).

In the 4D dose algorithm, equation 8.2 is adapted to cope with target motion by adding time-dependent displacements $dt_{dgp}(T)$ and $du_{dgp}(T)$ for motion in the t- and u-direction respectively. Tracking in the t- and u-direction was included similarly by adding spot displacements $dt_{spot}(T)$ and $du_{spot}(T)$ as can be seen in equation 8.3. For each beam adjustment in the t- or u-direction, the radiological depth of the spot at s_{spot} is calculated. If a difference in the radiological depth of the spot at the planned position and at the new position occurs, the range shifter settings are adapted to correct for this difference. The spot position in depth will therefore always be exactly the same as planned, and its radiological depth however can be adjusted in heterogeneous targets:

$$d = \frac{D_w}{2\pi\sigma_{vv}^2} e^{\frac{-((t_{\text{spot}} + dt_{\text{spot}}(T)) - (t_{\text{dgp}} + dt_{\text{dgp}}(T)))^2}{2\sigma_w^2}} e^{\frac{-((u_{\text{spot}} + du_{\text{spot}}(T)) - (u_{\text{dgp}} + du_{\text{dgp}}(T)))^2}{2\sigma_w^2}}$$
(8.3)

As the PTV moves relative to the bone in the heterogeneous targets, parts of the target will move into or out of the 'shadow' of the bone, resulting in changing radiological depths for these regions. In the dose calculation for a pencil beam applied at time T, this is taken into account by performing a recalculation of the radiological depth w for every displaced dose grid point. The new radiological depth is obtained by linear interpolation of the radiological depths of the dose grid points in the static target to the position of the displaced dose grid point. In this way, the dose *d* is calculated using the correct values for the integral dose D_w and the beam width σ_w , both depending on the radiological depth.

The time T (counted from the start of the treatment), at which a pencil beam is applied, is calculated in advance from the static treatment plan as developed by PSIplan. First, the order in which the pencil beams are applied is determined from the spot positions and the scanning order of the gantry. By adding up the irradiation times and the adjustment times of all previous beams, T is calculated for every pencil beam, starting with T = 0 for the first beam. The dose calculation assumes irradiation at the specific time instant T and does not take into account any motion during irradiation. However, as irradiation times are very short (average 5 ms), this is not likely to give significant errors in the dose distribution.

The time needed to make additional adjustments of the spot position to compensate for target motion during tracking is added to the 'treatment clock'. For beam adjustments in the t- and u-direction, this time depends on the size of the correction that has to be made. As this time is generally in the order of a couple of milliseconds, target displacements occurring during beam adjustments were neglected. Also, as the time to adjust the spot position in depth is assumed to take a fixed time of 150 ms for small adjustments (estimated for the PSI gantry 2), this time was added and adjustments were made for the target position 150 ms ahead in time.

For comparison purposes, pure rescanning with a rescan factor *n* has also been implemented by repeating the dose calculation described above n times. The spots were applied in the same order in each rescan and the irradiation time of each beam was reduced by a factor n. The time that is needed in between rescans to reposition the pencil beam to the first spot position was taken into account as well. It is important to assure that the rescan frequency and breathing frequency are not equal in order to avoid interference and loss of the rescan effect. For this target and breathing period of 5 s, interference was not observed.

Position error

To investigate the influence of a spot position error on the treatment outcome, simulations were performed with different tracking modes that are listed in table 8.1. Three main tracking modes were used: 'no tracking', 'perfect tracking' and 'imperfect tracking'. In the 'no tracking' and 'perfect tracking' modes, the target motion amplitude ranged

Table 8.1: Overview of the different tracking modes used for motion in either the tor u-direction. For imperfect tracking, a target motion amplitude of 15 mm was used. The time delay d was varied from 0 to 750 ms and the position error factor f ranged from 0 to 2.

Tracking mode	Target motion	Tracking motion	RMS position error
No tracking			
Static target	$dx_{dgp}(T) = 0$	$dx_{spot}(T) = dx_{dgp}(T)$	RMS = 0
Moving target	$dx_{dqp}(T) \neq 0$	$dx_{spot}(T) = 0$	RMS > 0
Perfect tracking	$dx_{dqp}(T) \neq 0$	$dx_{spot}(T) = dx_{dqp}(T)$	RMS = 0
Imperfect tracking	5.	· •	
Time delay (d)	$dx_{dqp}(T) \neq 0$	$dx_{spot}(T) = dx_{dqp}(T-d)$	RMS > 0
Amplitude factor (f)	$dx_{dgp}(T) \neq 0$	$dx_{spot}(T) = f \cdot dx_{dgp}(T)$	RMS > 0

from 0 to 15 mm (interval 5 mm) and several starting phases in the range from 0 to 2π (interval $\pi/3$) were used. The target was either not tracked ($d_{spot} = o$) or perfectly tracked $(d_{spot} = d_{dgp})$. In the 'imperfect tracking' mode, spot position errors were deliberately introduced by a time delay d ranging from 0 to 750 ms (interval 50 ms) or by a tracking error factor f that was varied from 0 to 2 (interval 0.1). This simulated errors in the tracking signal, with f = 1 representing perfect tracking and $f \neq 1$ representing various magnitudes of spatial errors in the tracking signal. The target motion was set to 15 mm and the starting phases were varied.

To obtain an indication of the average spot position error of all beams during a simulation, the root mean square (RMS) position error was calculated for every simulation. The RMS position error for a treatment with *n* pencil beams is given in equation 8.4:

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(dx_{\text{spot},i}(T_i) - dx_{\text{dgp},i}(T_i) \right)^2}$$
 (8.4)

Treatment quality

The treatment outcome was evaluated by calculation of the $D_{5\%}$ - $D_{95\%}$ value derived from the dose volume histogram (DVH) of the CTV. $D_{5\%}$ and $D_{95\%}$ are the doses, expressed in percentage prescription dose, received by 5% and 95% of the CTV volume respectively. The $D_{5\%}$ - $D_{95\%}$ value is a measure for the dose homogeneity in the CTV and a small value represents a homogeneous dose distribution. As this measure does not give any information about the dose level in a target, it is important to note that the prescription

dose was received by part of the CTV in all simulations. Given the simplicity of the target volume and geometry, we take as unacceptable any deviation of $D_{5\%}$ - $D_{95\%}$ greater than 5%.

8.3 Results

Homogeneous target

Figure 8.3 shows the $D_{5\%}$ - $D_{95\%}$ values as a function of the RMS position error for all simulations for the homogeneous target. The shaded areas indicate the bandwidth of

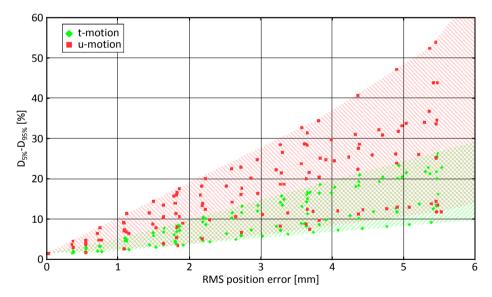


Figure 8.3: Treatment outcome ($D_{5\%}$ - $D_{95\%}$) as a function of the RMS position error for all simulations for the homogeneous target. The target was either moving parallel to the primary scanning direction of the gantry (t-direction, green points) or moving parallel to the secondary scanning direction (u-direction, red points). For static irradiation and perfect scanning the RMS position error is by definition equal to zero. Points with a RMS position error larger than zero were obtained by performing either no tracking or imperfect tracking while varying the motion amplitude, starting phase, time delay or position error factor. The shaded area approximates the bandwidth of the simulation points.

the results for the two target motion directions. No distinction was made between the different tracking modes (see table 8.1) as no clear differences in behavior were observed in the results. So although tracking can give a reduction in the position error of applied pencil beams, it does not alter the response to a certain position error.

The figure shows that treatment outcomes are equal for static irradiation and perfect tracking, both with RMS position errors of o by definition. The algorithm is thus capable of perfectly correcting for target motion if the exact location is known. Moreover, it can be seen that treatment outcomes are likely to degrade quickly as the RMS position error increases, as well as the spread of the outcomes. And finally, the figure reveals that spotscanned proton therapy is more sensitive to position errors in the secondary scanning direction (u-direction) than to position errors in the primary scanning direction (tdirection).

This phenomenon is the result of the use of a specific scanning order (t-u-s) and the fact that the position error was time dependent in all cases. Because all possible spots in the primary scanning direction are applied first, the time between the application of neighboring spots is always longer in the secondary scanning direction than in the primary scanning direction. As the position error itself is also a periodical function, the

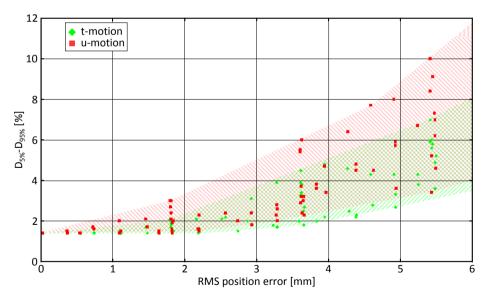


Figure 8.4: Treatment outcome (D_{5%}-D_{9%5}) as a function of the RMS position error for all simulations for the homogeneous target with a rescan factor of 4.

Table 8.2: $D_{5\%}$ - $D_{95\%}$ values (in percent) for static irradiation and perfect tracking of the heterogeneous targets. Perfect tracking was performed for target motion amplitudes of 5, 10 and 15 mm and a starting phase of 0°. Although RMS position errors are equal to 0 by definition for all simulations, treatment outcomes differ.

	t-	t-Motion (motion amplitude)				u-Motion (motion amplitude)			
Tracking mode	0 mm	5 mm	10 mm	15 mm	0 mm	5 mm	10 mm	15 mm	
No tracking	2.1	-	-	-	2.2	-	-	-	
Perfect tracking	-	2.2	2.9	3.4	-	3.8	4.6	5.0	

differences in the position errors of neighboring spots are likely to be relatively large in the secondary scanning direction, whereas they are small in the primary direction. These differences in position error give rise to deviations in the distance between neighboring spots, which affect the dose homogeneity in the target. Consequently, it can be stated that the dose inhomogeneity is mainly caused by changes in the relative spot positions in the secondary scanning direction. The observed effect is thus not directly caused by differences in scanning speed or spot spacing between scanning directions. These factors however do influence the choice for a particular scanning order.

To guarantee a $D_{5\%}$ - $D_{95\%}$ value below 5%, the RMS position error of a treatment should not be larger than approximately 1 mm and 0.5 mm in the primary and secondary scanning direction respectively. For an RMS position error larger than 3 mm, an acceptable treatment outcome is not to be expected.

In figure 8.4, the results of all simulations for the homogeneous target with a rescan factor of 4 are given. Simulations were performed for all (re)tracking modes, but with less variation in the starting phase compared to the simulations in figure 8.3.

Although fewer simulations were performed, a good resemblance with figure 8.3 can be observed. The treatment characteristics for retracking are similar as stated above; the $D_{5\%}$ - $D_{95\%}$ values are however reduced by roughly a factor of 4. This significant improvement in treatment results comes at the cost of an increase in treatment time with a factor of 2.5 for this target. For RMS position errors below 3 mm, acceptable treatment results can now be guaranteed.

Heterogeneous target

Table 8.2 gives the treatment results for static irradiation and perfect tracking of the heterogeneous targets in both t- and u-directions. For all simulations, the RMS position error was by definition equal to 0, as the spot position is perfectly corrected in the s-, t- and u-directions. In contrast to the homogeneous target, the results for perfect track-

Table 8.3: $D_{5\%}$ - $D_{95\%}$ values (in percent) for irradiations of the heterogeneous targets with different tracking modes and motion amplitudes and a starting phase of 0°. Perfect retracking was performed with a rescan factor of 4. Rescanning without beam adjustments was performed with rescan factors of 4 and 8. The treatment outcomes for uncorrected target motion are worse for motion in the t-direction (primary scanning direction) than for motion in the u-direction (secondary scanning direction). This observation contrasts the pattern found in the homogeneous target which showed that results are likely to be better for motion in the t-direction than for motion in the u-direction. This can be caused by the fact that two different (but almost identical) heterogeneous targets were used for motion in the two directions, but it can also be an accidental case where the pattern is not obeyed. However, when tracking and rescanning are applied, the pattern can again be observed in the results.

	t-Motion (motion amplitude)				u-Motion (motion amplitude)			
Tracking mode	0 mm	5 mm	10 mm	15 mm	0 mm	5 mm	10 mm	15 mm
No tracking	2.1	25.1	45.5	56.4	2.2	23.0	34.4	37.4
Perfect tracking	-	2.2	2.9	3.4	-	3.8	4.6	5.0
Perfect retracking 4x	-	2.6	2.4	2.8	-	3.0	4.8	5.4
Rescanning 4x	-	5.5	9.6	12.4	-	7.1	12.0	15.3
Rescanning 8x	-	1.6	1.7	2.9	-	1.8	2.5	4.5

ing are not equal to static irradiation and become worse for increasing target motion amplitude.

Table 8.3 gives in addition to table 8.2 the results for uncorrected irradiation, perfect retracking and rescanning of the heterogeneous targets. For perfect retracking, a rescan factor of 4 was used, while for rescanning rescan factors of 4 and 8 were used. It can be observed that all techniques result in dramatic improvements of the treatment compared to uncorrected irradiation of a moving heterogeneous target.

The use of 'pure' rescanning without beam position adjustments resulted in the most favorable treatment outcomes, which improved with increasing rescan factors. Remarkably, results were even observed to be better than static irradiations for some motion amplitudes. This might indicate that the treatment planning system has some troubles in determining the optimal plan for heterogeneous targets and that the imperfections are averaged out by the rescanning. The treatment outcomes for an amplitude of 15 mm might be negatively influenced by the fact that a PTV margin of 5 mm used in these targets is too small for this motion amplitude. The actual $D_{5\%}$ - $D_{95\%}$ values for the target are expected to be lower if proper margins are taken into account.

8.4 Discussion

The results of this study for the homogeneous target, showing that treatment outcomes are very sensitive to position errors and dependent on the motion direction, are in agreement with observations previously reported by Phillips *et al.* (117). For scanned carbon ions, Grözinger *et al.* presented similar results and also noted that dose inhomogeneity as most prominent in the distal region of the target (124). This was observed in our results as well. As explained earlier, the influence of the motion direction on the treatment outcome is caused by the fact that the position error is time dependent and that a specific scanning order is used. Differences in spot distance are therefore likely to be larger for target motion in the secondary scanning direction, resulting in lower dose homogeneity. If position errors were random in both directions, no dependence on the motion direction would have been observed. As this time dependence of the position error is also present in our modelling of residual tracking errors during 'imperfect tracking', no differences in sensitivity to position errors between 'no tracking' and 'imperfect tracking' (the two tracking modes with RMS > 0) were observed. For both tracking modes, the same pattern therefore holds, as depicted in figures 8.3 and 8.4.

However for 'imperfect tracking', especially the modelling of the prediction error by the use of an amplitude factor is questionable. It makes a prediction error equivalent to uncorrected irradiation of a target moving with a certain amplitude. The position errors are time dependent, proportional to target acceleration and have a harmonic probability density function (PDF). The latter contrasts observations by Vedam *et al.*, who found the PDF for the prediction error to be normally distributed (134). However, results found by Vedam *et al.* (134) and Hoogeman *et al.* (28) reveal that prediction errors are time dependent and that large errors are to be expected for large target accelerations. Although the amplitude factor model does have some correspondence to observed prediction errors, it is unclear whether its use is appropriate. Real-life target motion and prediction data are therefore recommended to be used to investigate the effect of prediction errors on the treatment outcome.

In this study, we have also looked into the effect of tracking and rescanning for a simple, static density heterogeneity behind which the target volume moves. An interesting observation from this analysis is that, even in the case of perfect tracking, small dose heterogeneities appear behind the density heterogeneity that are not there in the static case. This loss in dose homogeneity while the spot position is at exactly the same location within the target is caused by a change in the pencil beam shape. When a displaced pencil beam encounters a difference in the thickness of the heterogeneity, the amount of scatter and thus the beam width will change as well. So, if the pencil beam energy has to be adjusted during tracking to obtain the original spot position in depth, the pencil beam width will always alter. Moreover, differences of the heterogeneity thickness within the

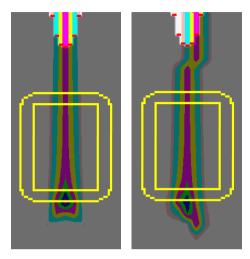


Figure 8.5: Dose distributions of two proton pencil beams in a static target (left) and in a displaced target (right). In the right figure, the moving target is displayed in its original position, resulting in a discontinuity in the dose distribution. The position of the Bragg peak on the central axis of the pencil beam within the target is exactly the same for both situations. The effect of the pencil beam travelling through a different part of the heterogeneity on the pencil beam shape can clearly be observed.

beam can additionally give rise to (asymmetrical) changes in the pencil beam shape. This is illustrated in figure 8.5, where the same pencil beam is depicted for a static irradiation and for a displaced target. The shape of the beam has clearly changed while the spot position within the target is equal for both situations.

The presence of differences in the pencil beam shape is of itself not a cause for dose inhomogeneity, as this can always be observed in treatment plans for heterogeneous targets. By optimizing the beam weights during treatment planning, a homogeneous dose distribution can be obtained from beams with different shapes. However if tracking is applied, these optimal beams' weights might no longer correspond to the shape of the displaced beam, resulting in dose inhomogeneity. This effect is illustrated in figure 8.6 for a target that was given a constant offset (no motion) from its original planning position. Although the target offset was perfectly corrected for by adjusting the spot positions of all pencil beams, dose inhomogeneity can clearly be observed in the region where spot adjustments in depth had to be made due to the bone. The hotspot is caused by beams that were heavily weighted to compensate for the beam widening by the bone in the original planning setup. These beam weights are too high for the displaced setup where no bone is encountered and the beams are not widened. The opposite holds for the region that

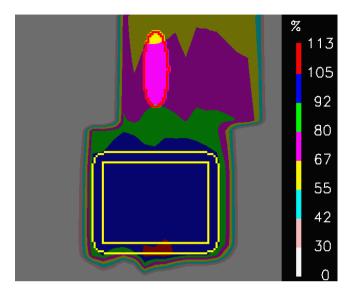


Figure 8.6: Dose distribution in a heterogeneous target that was given a position offset (no motion). The displaced target is displayed in its original position. A hotspot can be seen in the region that was originally behind the bone, but not during irradiation. Some underdosage of the PTV can be seen in the part of the target that moved behind the bone. Note that the effects are mainly present in the distal region of the target, where beam weights are the highest.

moves behind bone, resulting in slight underdosage of the PTV. It should be noted that the example given in figure 8.6 is an exaggeration of the effect that can be observed in a continuously moving target, where position errors are not constant and on average less.

These results suggest that for heterogeneous targets, beam weights are only optimal for the exact configuration as for which they were optimized. Adjustments in the beam position relative to density heterogeneities will therefore inevitably lead to sub-optimal treatment outcomes. As perfect tracking for heterogeneous targets did not result in treatment results equal to static irradiation, imperfect tracking was not further investigated.

The results shown in table 8.3 show that treatment results for perfect retracking are similar to those found for perfect tracking. Although the RMS position error was equal to o for all retracking simulations, treatment outcomes were not equal to irradiation of a static target. These results are in agreement with the suggestion that beam repositioning with respect to density heterogeneities gives rise to dose inhomogeneities as beam weights are not optimal for the new pencil beam positions and shapes. The use of simultaneous rescanning during retracking does not seem to have a considerable effect

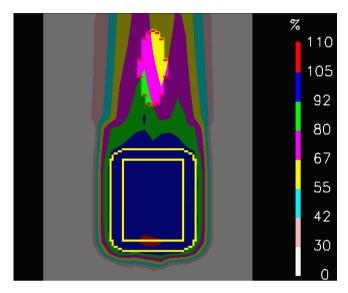


Figure 8.7: Dose distribution in a moving heterogeneous target for perfect retracking with a rescan factor of 4. The displaced target is displayed in its original position. A hotspot can clearly be observed behind the bony region.

on this phenomenon and neither does the rescan factor as RMS position errors are o by definition. These results give the order of magnitude of the best achievable treatment outcomes for tracking and retracking of these heterogeneous targets. The presence of target position uncertainties will only worsen treatment outcomes to an extent determined by the rescan factor. The motion correction, by adjustment of pencil beam positions, is thus by itself a substantial cause of dose inhomogeneity. The dose distribution for retracking of the heterogeneous target moving in the u-direction with an amplitude of 15 mm is given in figure 8.7. The presence of a hot spot behind the bony region is clearly visible.

It should be noted that a rather large 'bone' has been modelled in this work in order to ensure that any effects of proximal density heterogeneities on the tracking algorithm would be clearly observed. Certainly, the 40 mm maximum thickness introduced would be large for ribs, but not necessarily an exaggeration of the thickness of bone when passing through the femoral head. Whether the detrimental effect of the bone on dose homogeneity is a correct representation of reality however, or a side effect of the equivalent path length algorithm, which scales both primary and secondary radiation in depth, needs to be investigated further. This algorithm leads to relatively large beam deformations and the loss in dose homogeneity found in this study might be an exaggeration of real treatments. The use of real-life heterogeneous patient data is therefore recommended to be used to study the effect of density heterogeneities on the tracking performance, plus the use of other dose calculation algorithms such as Monte Carlo techniques. Nevertheless, our preliminary results here indicate a potential problem for tracking algorithms which needs to be investigated further.

8.5 Conclusions

This study shows that spot-scanned proton therapy is very sensitive to position errors in general and especially to position errors parallel to the secondary scanning direction. It is advised to take this into account during patient positioning by aligning the greatest motion amplitude with the primary scanning direction. Tracking of homogeneous targets requires position errors to be less than 1 mm in order to obtain acceptable treatment outcomes. As the use of rescanning significantly reduces the sensitivity to position errors, retracking is considered to be the best and most robust technique for the treatment of moving homogeneous targets. For a rescan factor of 4, retracking has been found to require position errors less than 3 mm to guarantee acceptable treatment outcomes.

In heterogeneous targets, tracking itself has been found to have a negative influence on the dose homogeneity when density gradients are present within the field. The adjustment of the pencil beam position can cause beam shapes to change if differences in density are encountered by the displaced beams. Since beam weights from the original treatment plan no longer correspond to the (deformed) situation for which they are used, hot and cold spots arise. The extent to which this effect can be observed in real treatment situations remains unclear. Rescanning has been found to be an effective method to cope with motion in heterogeneous targets if adequate rescanning factors are used. Even improvements of treatment outcomes with respect to static treatments were found. However, these favorable results come at the cost of increased treatment time, require increased PTV margins and will inevitably result in some smoothing of the resultant delivered dose distribution.

chapter 9

The price of robustness; impact of worst-case optimization on organ-at-risk dose in intensity-modulated proton therapy for oropharyngeal cancer patients

Submitted to Radiotherapy and Oncology

Iris van Dam^{1,2}, Steven van de Water¹, Dennis Schaart², Abrahim Al-Mamgani¹, Ben Heijmen¹, Mischa S. Hoogeman¹

 Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands
 Faculty of Applied Sciences, Section Radiation Detection and Medical Imaging, Delft University of Technology, Delft, the Netherlands

Abstract

Purpose: To investigate the effect of the degree of robustness against setup and range errors on organ-at-risk (OAR) dose and normal tissue complication probabilities (NTCPs) in intensity-modulated proton therapy for oropharyngeal cancer patients.

Material and Methods: For 10 oropharyngeal cases (4 unilateral and 6 bilateral), robust treatment plans were generated using 'minimax' worst-case optimization. We varied the robustness against setup errors ('setup robustness') from 1 to 7 mm and the robustness against range errors ('range robustness') from 1% to 9% (+1 mm). We evaluated OAR doses and NTCP-values for xerostomia, dysphagia and larynx edema.

Results: OAR doses and NTCP-values typically increased with higher degrees of robustness, especially for setup robustness and bilateral cases. Increasing setup robustness from 1 mm to 3, 5, and 7 mm resulted in average NTCP-values increased by 2.2, 5.2 and 8.1 percentage point, whereas they increased by 0.5, 0.9, 1.4, and 1.8 percentage point when increasing range robustness from 1% to 3%, 5%, 7%, and 9%.

Conclusions: A higher degree of setup robustness resulted in considerably higher OAR doses and NTCP-values, especially in bilateral cases, while the impact of range robustness was limited. Prioritizing setup error reduction is therefore recommended, e.g. by implementing online setup verification/correction.

9.1 Introduction

Intensity-modulated proton therapy (IMPT) uses proton pencil beams with individually optimized weights to deliver curative doses to the target with improved sparing of healthy tissues compared with intensity-modulated radiotherapy (75). However, the dose delivery can easily be perturbed by inaccuracies in the preparation and execution phase of the treatment, such as a misalignment of the target with respect to the proton beams and inaccuracies in the calculated proton range. Such inaccuracies may result in underdosing or overdosing the clinical target volume (CTV) and in higher doses delivered to organs-at-risk (OARs) (83, 84, 135). These dose perturbations can be reduced by improving the accuracy of the treatment, for example by implementing image-guided patient positioning and, in the future, in-vivo range verification techniques. Still, treatment uncertainties will always remain.

In IMPT, residual treatment uncertainties are preferably accounted for by performing 'robust' treatment planning, since the use of traditional safety margins around the tumor (and OARs) was found to be inadequate due to additional uncertainties in proton range (99, 101). During robust optimization, certain errors in patient positioning (setup errors) and proton range (range errors) are explicitly included in the plan optimization, improving the expected value or worst-case value of the objective function or individual objectives (82, 88, 89, 136). Robust optimization thereby minimizes the impact of residual treatment uncertainties on the dose delivered to the CTV and OARs. Still, to achieve adequate CTV coverage in all robustly optimized error scenarios, it is expected that robust treatment planning will result in higher doses to surrounding normal tissues, depending on the size of the uncertainties accounted for (82, 89).

The impact of the degree of robustness on the dose received by OARs is of clinical importance. It not only quantifies the price to pay for not being accurate, but it also assists in prioritizing and justifying measures to improve treatment accuracy, e.g. image guidance or in-vivo range verification. The aim of this study is to quantify the clinical impact of the magnitude of setup errors and range errors accounted for by robust optimization in oropharyngeal cancer patients, a treatment site with large density heterogeneities and many OARs close to the CTV. To this purpose, we generated treatment plans with varying robustness against setup and range errors, and evaluated OAR doses and normal tissue complication probabilities (NTCP) for xerostomia, dysphagia and larynx edema.

9 9 Methods and materials

Patient data and dose prescription

CT data of 6 bilateral and 4 unilateral oropharyngeal cancer patients were used in this study. Dose was prescribed according to a simultaneously integrated boost scheme: 66 Gy to the high-dose CTV and 54 Gy to the low-dose CTV, delivered in 30 fractions (74). The median high-dose CTV was 55.0 ml and the median low-dose CTV (including the high-dose CTV) was 186.5 ml. The considered OARs were the parotid glands, submandibular glands, spinal cord, brainstem, larynx, swallowing muscles and oral cavity. Table 9.1 shows the 'wish-list' containing the dose prescriptions for the CTVs and OARs. CTV prescription was chosen such that more than 98% of the CTV received more than 95% of the prescribed dose ($V_{95\%} > 98\%$), in all error scenarios included in the robust optimization. In addition, the CTV receiving more than 107% of the prescribed dose was aimed at to be lower than 2% and the CTV receiving more than 110% of the prescribed dose to be 0%, for all robustly optimized error scenarios. We used a 3-beam arrangement with gantry angles of 60°, 180° and 300°, as proposed in literature (76).

Treatment planning system

Robust IMPT treatment plans were generated using 'Erasmus-iCycle', our in-house developed treatment planning system for fully automated plan generation (66, 70, 86, 87). The algorithm uses 'prioritized' multi-criteria optimization. It does not condense the optimization problem into a single weighted-sum objective function, but it optimizes the different objectives one-by-one according to their priorities as defined by the user in the so-called 'wish-list'. The user can also define constraints, which always have to be met during treatment plan generation. For every patient group a single wish-list can be used, which is fine-tuned in close collaboration with radiation oncologists (70, 87). Table 9.1 shows the wish-list that was used in this study.

Pencil beams were selected and optimized using the resampling method as described by Van de Water et al. (86), an alternative to the use of pencil beams that are distributed over the CTV according to a regular grid. The resampling method iteratively performs: (1) random sampling of candidate pencil beams from a very fine grid, (2) inverse optimization and (3) exclusion of pencil beams with a low contribution. In the current study, resampling was performed using a sample size of 3000 randomly selected candidate pencil beams per iteration. Plan optimization was terminated after 10 resampling iterations, as solutions were then found to have converged (86).

Table 9.1: The wish-list describing the dose prescriptions as used in this study. The priority numbers indicate the order in which objectives are optimized, a low number corresponds to a high priority. The CTVintermediate is a 10 mm transition region between the high-dose and low-dose CTV. The CTV-low' consists of the low-dose CTV excluding the transition region.

Constraints				
	Structure	Туре	Limit	Robust
	CTV-high	minimum	0.97 × 66 Gy	yes
	CTV-intermediate	minimum	0.97 × 54 Gy	yes
	CTV-low'	minimum	0.97 × 54 Gy	yes
Objectives				
Priority	Structure	Type	Goal	Robust
1	CTV-high	maximum	1.08 × 66 Gy	yes
1	CTV-intermediate	maximum	1.08 × 66 Gy	yes
1	CTV-low'	maximum	1.08 × 54 Gy	yes
2	CTV-rings (high-dose conformality)	maximum	0 Gy	no
2	CTV-rings (high-dose conformality)	mean	0 Gy	no
3	Parotid glands	mean	0 Gy	yes
4	Submandibular glands	mean	0 Gy	yes
5	Spinal cord	maximum	20 Gy	yes
5	Brainstem	maximum	20 Gy	yes
6	Larynx	mean	0 Gy	yes
7	Swallowing muscles	mean	0 Gy	yes
7	Oral cavity	mean	0 Gy	yes
8	CTV-rings (low-dose conformality)	maximum	0 Gy	no
8	CTV-rings (low-dose conformality)	mean	0 Gy	no
9	Total spot weight	sum	0 Gp	no

Abbreviations: CTV = clinical target volume; Gp = Giga-protons

The dose calculation algorithm implemented in Erasmus-iCycle was developed at the Massachusetts General Hospital - Harvard Medical School where it is implemented in their in-house developed treatment planning system 'ASTROID' (23). To account for density heterogeneities, the algorithm uses a superposition-convolution method. We used a dose grid resolution of $3 \times 3 \times 3$ mm³. Available proton energies ranged from 70 to 230 MeV and corresponding pencil beam widths ranged from 7 to 3 mm sigma (in-air at the isocenter), respectively. To irradiate superficially located target regions, a range shifter of 75 mm water-equivalent thickness was inserted during the delivery of a field if needed.

Minimax robust optimization

A 'minimax' worst-case approach was used to ensure robustness against setup errors and range errors (88, 89, 136). The method simultaneously included several (error) scenarios and optimized the worst-case value for the constraints and objectives in the wish-list. Nine scenarios were included in the robust optimization: one nominal scenario, setup errors in positive and negative directions along three axes (six scenarios) and positive and negative range errors (two scenarios). Setup errors were modelled by laterally shifting the pencil beams, while range errors were modelled by adjusting the proton energy.

Study design

We generated treatment plans with varying degrees of robustness against setup errors (denoted as 'setup robustness') and robustness against range errors (denoted as 'range robustness'). We used a setup robustness of 3 mm and a range robustness of 3% +1 mm as our 'standard' robustness settings (101, 99, 137). Setup robustness and range robustness were varied separately around these settings. We tested setup robustness of 1, 3, 5 and 7 mm and range robustness of 1%, 3%, 5%, 7% and 9%, the latter always in combination with an absolute range offset of 1 mm. These degrees of robustness were assumed to cover the entire spectrum of realistic treatment uncertainties.

Minimax robust optimization was described as an adequate method when the errors included in the optimization were not larger than the pencil beam width (138). However, these conditions were not strictly satisfied in this study when aiming for high degrees of robustness, which could result in compromised plan robustness for intermediate errors (i.e. errors smaller than the errors included in the optimization). To verify the adequacy of the robust optimization, we performed dose recalculations for all treatment plans while applying variable setup and range errors (see Appendix 9.A).

Plan evaluation

Firstly, we checked for all treatment plans whether the clinical target constraints (V_{95%} > 98% for high-dose and low-dose CTV) and objectives ($V_{107\%} \approx 2\%$ and $V_{110\%} \approx 0\%$ for high-dose CTV) were fulfilled for all error scenarios included in the optimization. Secondly, plan quality was evaluated based on doses to OARs. For OARs with a mean-dose objective, we evaluated the average of the mean-dose values in the robustly optimized error scenarios. For OARs with a maximum-dose objective, we evaluated the worst-case dose.

NTCP-model	Endpoint	Parameters
Jellema <i>et al.</i>	Severe xerostomia after 6 months using quality-of-life questionnaires	Mean dose in both parotid glands and both submandibular glands
Dijkema et al.	<25% salivary flow for individual parotid gland after 1 year	Mean dose in parotid gland
Levendag et al.	Late dysphagia using quality-of-life questionnaires	Mean dose in superior constrictor muscle
Christianen et al.	Grade ≥2 swallowing dysfunction after 6 months using questionnaires	Mean dose in larynx and superior constrictor muscle
Rancati et al.	Grade ≥2 edema after 15 months observed by fiber-optic examination	Mean dose in larynx

Table 9.2: Overview of the normal tissue complication probability (NTCP) models used in the current study.

To estimate the impact of increased OAR doses on quality-of-life, we used published NTCP-models. The models, their endpoints and the model parameters are listed in table 9.2. The probability of xerostomia was evaluated using NTCP-models by Jellema *et al.* (139) and Dijkema *et al.* (140). The probability of dysphagia was evaluated using NTCP-models by Levendag *et al.* (141) and Christianen *et al.* (142). Finally, we evaluated the probability of larynx edema using the NTCP-model by Rancati *et al.* (143).

9.3 Results

All generated treatment plans were clinically acceptable. The worst-case coverage ($V_{95\%}$) was higher than 98% for both CTVs in all treatment plans except one, for which the high-dose CTV coverage was 97.8%. The worst-case $V_{107\%}$ of the high-dose CTV was on average 3.0% over all generated treatment plans, while the $V_{110\%}$ was never higher than 0%. Also for intermediate errors (i.e. errors smaller than the errors included in the optimization), plan robustness was largely maintained as described in the Appendix 9.A.

The NTCP-values as a function of setup robustness are shown in figure 9.1, and as a function of range robustness in figure 9.2. The impact of setup robustness could be substantial, resulting in maximum increases in NTCP-values of 8.8, 19.0 and 29.3 percentage point (pp) when setup robustness increased from 1 mm to 3, 5, and 7 mm, respectively. The impact of range robustness was considerably smaller, resulting in NTCP-values to increase by 3.8, 5.0, 6.2, and 7.8 pp at maximum when range robustness increased from 1% to 3%, 5%, 7%, and 9%, respectively. Especially the impact of setup robustness was patient group-specific. In unilateral cases, NTCP-values were generally low and showed a relatively small increase for higher degrees of robustness. In bilateral cases on the con-

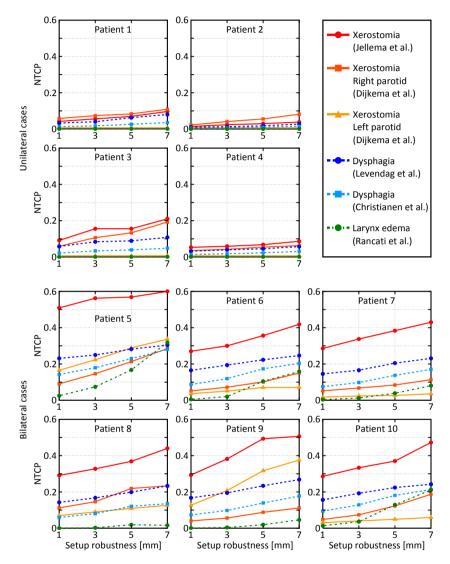


Figure 9.1: Normal tissue complication probabilities (NTCP) for different NTCP-models as a function of the setup robustness for unilateral cases (patients 1-4) and bilateral cases (patients 5-10). Range robustness was set to 3% + 1 mm for all data points in this figure.

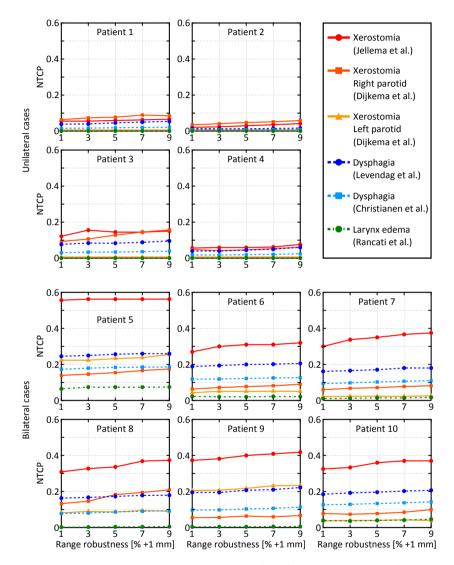


Figure 9.2: Normal tissue complication probabilities (NTCP) for different NTCP-models as a function of the range robustness for unilateral cases (patients 1–4) and bilateral cases (patients 5–10). Setup robustness was set to 3 mm for all data points in this figure.

trary, NTCP-values were relatively high under standard robustness settings and could increase considerably when increasing setup robustness.

Table 9.3 lists the evaluated OAR doses and NTCP-values for the different robustness settings, averaged over the 4 unilateral and the 6 bilateral patients. The table confirms that an increase in setup robustness from 1 to 7 mm had a larger impact on OAR doses (average increase 7.9 Gy over all patients and all OARs) than an increase in range robustness from 1% to 9% (+1 mm) (average increase 1.9 Gy over all patients and all OARs). Next to that, table 9.3 shows that the impact of setup robustness on OAR doses was generally somewhat higher in the bilateral cases than in the unilateral cases, when increasing setup robustness from 1 to 7 mm. However, for some OARs located very close to the CTVs (e.g. ipsilateral salivary glands, spinal cord, SCM, MCM and oral cavity) the impact was observed to be comparable or even higher in the unilateral cases. Nevertheless, the increase in NTCP-values was always higher for the bilateral cases, which was also observed in figure 9.1, as the overall OAR dose level was higher in these cases, which corresponded to a steeper part of the NTCP-curves.

Figure 9.3 shows dose distributions of patient 1 for 1 mm and 7 mm setup robustness and 1% and 9% (+1 mm) range robustness, for each field separately and all three combined. The additionally treated volume around the high-dose CTV is clearly visible for a setup robustness of 7 mm. A higher degree of robustness resulted in a higher contribution of the 300° beam, whereas it caused dose fall-offs to be shallower in the 60° and 180° beams (indicated by arrows).

9.4 Discussion

In this study, we systematically investigated for 10 oropharyngeal cancer patients the impact of the degree of robustness against setup and range errors on doses to OARs. This has, to the best of our knowledge, not been studied before. An increase in the degree of robustness typically resulted in higher, and sometimes in considerably higher, OAR doses and NTCP-values. This effect was especially pronounced for setup robustness and bilateral patients.

Robustness was achieved using the 'minimax' optimization method (88, 89, 136). This is a conservative approach, as only the worst-case value of all included scenarios is optimized. The use of an alternative robust method, such as probabilistic planning in which the expected value is optimized (82), may result in a lower price to be paid for robustness, in terms of OAR doses. Another limitation of this study is the fact that the NTCP-models were constructed using data of (conventional) photon treatments and

Table 9.3: Organ-at-risk doses and NTCP-values [%] for different robustness settings, averaged over the 4 unilateral and the 6 bilateral patients (maximum-dose for spinal cord and brainstem, mean-dose for other organs-at-risk).

			Unilater	al cases						
Setup robustness	[mm]	1	3	5	7	3	3	3	3	3
Range Robustness	[% +1 mm]	3	3	3	3	1	3	5	7	9
Parotid right	[Gy]	12.1	15.3	17.2	19.8	14.6	15.3	16.3	17.3	17.9
Parotid left	[Gy]	0.2	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.2
SMG right	[Gy]	31.4	36.0	39.9	43.5	34.7	36.0	38.0	39.5	40.5
SMG left	[Gy]	4.2	5.0	5.4	6.0	4.7	5.0	4.1	4.0	4.6
Spinal cord	[Gy]	17.0	19.4	21.5	25.3	19.1	19.4	19.2	19.2	19.5
Brainstem	[Gy]	8.4	9.4	9.8	11.5	8.5	9.4	11.9	12.5	15.
Larynx	[Gy]	4.6	7.1	10.5	13.1	6.8	7.1	7.4	7.7	8.0
SCM	[Gy]	26.4	29.8	33.2	37.2	29.8	29.8	30.7	32.2	33.8
MCM	[Gy]	13.9	16.5	20.3	23.1	16.7	16.5	17.3	18.4	19.4
MCI	[Gy]	4.1	5.1	6.3	7.4	4.7	5.1	5.3	5.9	6.1
MCrico	[Gy]	0.8	1.4	2.2	3.1	1.3	1.4	1.5	1.6	1.9
Oesophagus	[Gy]	0.9	1.4	2.4	3.4	1.4	1.4	1.7	1.8	1.9
Oral cavity	[Gy]	12.4	14.0	16.5	18.8	14.3	14.0	13.1	13.2	13.9
Xerostomia (Jellema et al.)	[%]	5.0	7.3	8.0	10.7	6.3	7.3	7.3	7.6	8.3
Xerostomia (R-Dijkema et al.)	[%]	4.3	6.6	8.1	11.1	5.9	6.6	7.5	8.5	9.0
Xerostomia (L-Dijkema <i>et al</i> .)	[%]	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Dysphagia (Levendag et al.)	[%]	3.3	4.4	5.3	6.8	4.2	4.4	4.6	5.0	5.6
Dysphagia (Christianen et al.)	[%]	1.3	1.8	2.4	3.2	1.8	1.8	1.9	2.1	2.3
Larynx edema (Rancati et al.)	[%]	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
			Bilatera	l cases						
Setup robustness	[mm]	1	3	5	7	3	3	3	3	3
Range Robustness	[% +1 mm]	3	3	3	3	1	3	5	7	9
Parotid right	[Gy]	15.3	18.3	22.0	24.9	17.8	18.3	19.3	19.7	20.
Parotid left	[Gy]	14.9	17.8	20.4	22.0	17.2	17.8	17.9	18.4	18.
SMG right	[Gy]	53.9	54.7	56.4	58.0	53.8	54.7	55.5	55.9	56.
SMG left	[Gy]	40.4	42.9	46.2	47.4	40.2	42.9	43.8	45.1	45.
Spinal cord	[Gy]	15.6	16.6	22.0	23.9	19.7	16.6	17.0	16.2	17.
Brainstem	[Gy]	3.2	4.6	10.8	17.0	6.2	4.6	5.1	5.7	6.2
Larynx	[Gy]	19.8	24.7	31.2	34.2	24.5	24.7	25.1	25.4	25.8
SCM	[Gy]	52.4	54.8	57.6	59.5	54.4	54.8	55.4	55.8	56.
MCM	[Gy]	46.6	50.4	54.0	56.1	49.8	50.4	51.1	51.7	52.0
MCI	[Gy]	15.5	19.9	26.2	28.9	19.5	19.9	20.3	20.9	21.
MCrico	[Gy]	5.2	8.2	11.3	13.5	7.4	8.2	8.6	9.0	9.4
Oesophagus	[Gy]	2.4	4.5	7.5	9.7	4.1	4.5	4.5	5.0	5.1
Oral cavity	[Gy]	21.6	23.8	26.5	28.7	23.1	23.8	24.1	24.5	25.
Xerostomia (Jellema et al.)	[%]	32.2	37.4	42.3	47.8	35.6	37.4	38.6	39.8	40.3
Xerostomia (R-Dijkema <i>et al</i> .)	[%]	6.6	9.4	13.8	18.0	8.8	9.4	10.4	11.1	12.0
Xerostomia (L-Dijkema <i>et al</i> .)	[%]	7.4	10.6	14.3	16.7	10.2	10.6	10.8	11.4	11.5
Dysphagia (Levendag et al.)	[%]	16.8	19.4	22.8	25.4	19.0	19.4	20.0	20.5	20.9
Dysphagia (Christianen et al.)	[%]	8.8	11.7	16.3	19.7	11.4	11.7	12.2	12.5	12.8
Larynx edema (Rancati et al.)	[%]	0.8	2.5	7.9	13.8	2.4	2.5	2.6	2.7	2.8

Abbreviations: SMG = submandibular gland; SCM = superior constrictor muscle; MCM = middle constrictor muscle; MCI = inferior constrictor muscle; Mcrico = cricopharyngeal muscle.

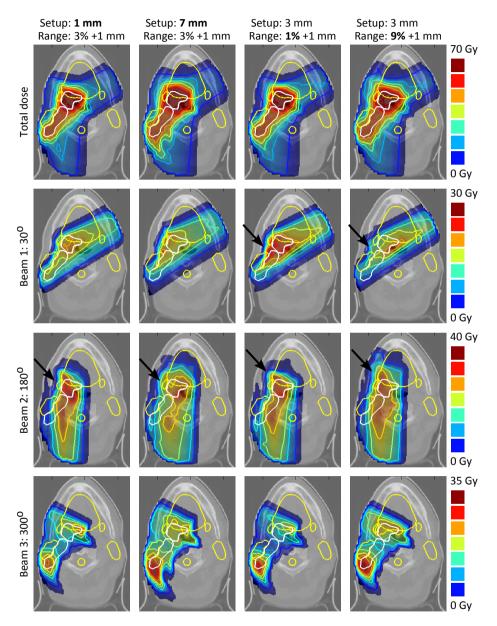


Figure 9.3: Dose distributions (total and per field) of patient 1 for different robustness settings. Target volumes are indicated in white and organs-at-risk in yellow. Arrows indicate differences in dose gradients between different robustness settings.

corresponding dose levels. As a result, the models may have reduced accuracy in predicting side-effects for IMPT treatments.

The larger impact of setup robustness compared with range robustness could be explained by the fact that setup errors occur in three directions and result both in a beam-to-target misalignment and a subsequent range error when different densities are encountered. In contrast, pure range errors occur only in the direction along the beam path. Next to that, range robustness can be achieved to a certain extent by increasing doses within the target volume itself without enlarging the high-dose region around the target (144). The target dose prescriptions used in this study may have provided enough room to achieve range robustness in this way.

To illustrate the clinical consequences of the results found in this study, the NTCPmodels by Jellema et al. (xerostomia) and Levendag et al. (dysphagia) are discussed in more detail below (139, 141). These models describe the most prevalent and harmful toxicities for this patient group and the model endpoints were patient-reported using quality-of-life questionnaires. In clinical practice, the 'standard' setup robustness of 3 mm would require online position verification and correction using in-room imaging techniques. Even when online position verification is performed inclusion of setup robustness is required, as local residual setup errors will always arise in anatomical subregions due to the flexibility of the neck and the resulting daily changes in the patient's pose (97). Offline position verification, still regularly used in photon therapy in combination with a 5 mm safety margin, is likely to require a higher degree of setup robustness of at least 5 mm (99). Reducing setup robustness from 5 mm (offline protocol) to 3 mm (online protocol) would result in average NTCP-values for xerostomia (Jellema et al.) to be reduced by 5.0 percentage point (pp, range: 0.6-11.2 pp) from 42.3% to 37.4% in bilateral cases and by 0.7 pp (range: 0.0-1.4 pp) from 8.0% to 7.3% in unilateral cases. The average probability of developing dysphagia (Levendag et al.) would be reduced by 3.4 pp (range: 2.9–3.9 pp) from 22.8% to 19.4% in bilateral cases and by 1.0 pp (range: 0.5-2.2 pp) from 5.3% to 4.4% in unilateral cases. Reducing range uncertainties from the standard settings of 3% +1 mm to 1% +1 mm would result in average NTCP-values for xerostomia to be lowered by 1.8 pp (range: 0.6-3.8 pp) in bilateral cases and by 1.0 pp (range: 0.0-3.4 pp) in unilateral cases. Average NTCP-values for dysphagia would be reduced by 0.4 pp (range: 0.0–0.8 pp) in bilateral cases and by 0.2 pp (range: -0.1–0.7 pp) in unilateral cases. In clinical practice, this reduction in range uncertainty is likely to require advanced dose calculation methods (137), improved stopping-power estimates using dual-energy CT-scanning (145) or proton radiography (146), and in-vivo range verification (147-150).

The required degree of robustness and corresponding complication probabilities should be weighed against the costs of reducing treatment uncertainties. For example, offline protocols are associated with lower workload and shorter occupancy time of the

treatment room compared with online protocols, and thus with lower costs. NTCPvalues can be lowered substantially by reducing setup uncertainties in bilateral cases, which might justify the use of advanced image guidance techniques for these patients. For unilateral cases, the benefit of reducing setup uncertainties was limited and an offline protocol might be equally effective in terms of OAR sparing. The effect of reducing range robustness was relatively small in both unilateral and bilateral cases, which makes it questionable whether expensive techniques to reduce range uncertainties will be costeffective for oropharyngeal cancer patients. However, we would like to stress that these recommendations only hold for the treatment site investigated in the current study. Other patients, with for example deep-seated tumors or fewer density heterogeneities (e.g. prostate cancer patients), might benefit differently from reduced treatment uncertainties. Next to that, anatomical changes were not accounted for by the robust optimization. The impact of these changes needs to be verified regularly, for example by the acquisition of repeat CT-scans, which could be used for replanning if necessary (135). Finally, we would like to underline the potential role of range verification methods for quality assurance in IMPT.

9.5 Conclusions

Increasing robustness against setup and range errors in IMPT for oropharyngeal cancer patients resulted in higher and potentially in much higher OAR doses and NTCP-values. Setup robustness was found to have a substantially larger impact than range robustness and this impact was considerably larger in bilateral cases compared with unilateral cases. We therefore recommend minimizing residual setup errors for bilateral oropharyngeal cancer patients in particular, e.g. by applying online position verification using in-room imaging and correction strategies.

Acknowledgments

The authors would like to thank Gerda Verduijn MD for sharing her clinical expertise.

chapter 10

Discussion

In this thesis, we have aimed at improving treatment planning and delivery for high-precision robotic radiotherapy using the CyberKnife and intensity-modulated proton therapy (IMPT). We have developed and evaluated several methods to reduce the treatment delivery time and the treatment planning time. This resulted in average delivery time reductions of 16%–38% and optimization times to be shortened by a factor of 2.8–5.6 on average. Next to that, we have investigated for both treatment modalities the sensitivity to treatment uncertainties and the effectiveness of techniques to reduce treatment errors or their impact on the delivered dose. This allowed us to establish favorable treatment strategies to achieve adequate dose delivery in the presence of intrafraction tumor motion and uncertainties in patient position, patient anatomy and proton range (for IMPT). These efforts will contribute to a more efficient, effective and safe delivery of a highly localized dose to the tumor using the CyberKnife and IMPT.

10.1 Efficiency of treatment delivery

In chapter 2, we investigated whether duration of CyberKnife treatments could be reduced by the introduction of more delivery-efficient beam collimation strategies. We generated treatment plans for stereotactic lung treatments either using a mini-multileaf collimator (mMLC) or the 'Iris' variable circular aperture collimator. A mMLC was expected to result in shorter treatment times, since the fields can be shaped more freely compared with the circular fields of the Iris collimator. The gain in delivery time depended on the tumor size and was found to be relatively small for small tumor volumes (<80 cm³), while being substantial for larger tumors. In accordance with the findings of this study, a mMLC has been introduced for the CyberKnife very recently.

An alternative approach to reduce delivery times was presented in chapters 3 and 6, in which we shortened treatment times for high-precision robotic radiotherapy and IMPT by incorporating delivery parameters in the treatment planning process. We especially

focused on the delivery parameters with the highest impact on the delivery time. The number of node positions was reduced in CyberKnife treatment plans, whereas we reduced the number of energy layers in treatment plans for IMPT. As a result, the delivery time per fraction was reduced by 18%-40% for the CyberKnife and by 17%-39% for IMPT treatments (when assuming 2 s energy switching time). The commercial planning system for the CyberKnife 'Multiplan' currently also features a node reduction technique, which enables the delivery time of clinical treatment plans to be reduced.

In the ideal situation, it would be possible to directly optimize the treatment time together with the dose parameters. This would assist the physician to properly balance delivery time and dosimetric plan quality. However, directly including treatment time in the plan optimization is not particularly straightforward. For a parameter to be included in the inverse optimization, it should be described by a continuous, preferably convex, function of the intensity of each individual beam or beamlet, which is typically not the case for the delivery time. For the CyberKnife and IMPT, delivery time depends to a large extent on the number of beams and their spatial distribution. The problem can be simplified by approximating the delivery time as a linear combination of the number of beams, beam directions, energy layers, etc. However, these quantities are discrete numbers that cannot be described using a continuous convex function depending on the individual beam weights. As a result, it is currently not possible to directly optimize the delivery time using our treatment planning system 'Erasmus-iCycle'. Because of the current limitations mentioned above, all algorithms to reduce treatment times as presented in this thesis are heuristics. Surrogate functions and intuitive decision rules are applied in a sequential approach to achieve a reduction in certain delivery parameters, and with that a reduction in treatment time. This means that the solutions are probably suboptimal in terms of the delivery parameters and treatment times. However, it does show that there is a lot of room to improve the time-efficiency of CyberKnife and IMPT treatment plans without affecting the plan quality, by including delivery parameters in the planning process. It is possible to optimize discrete parameters by using 'mixed integer programming' (MIP) and the implementation of a MIP-solver might allow for the treatment time to be included directly in the treatment plan optimization (94, 151). This would therefore be an interesting direction for future research.

Reduction of the treatment time per fraction has recently received great attention in conventional radiotherapy as well. The fraction duration of photon treatments has especially been shortened by the introduction of rotational techniques such as volumetricmodulated arc therapy (VMAT) (8, 9). During VMAT, the patient is irradiated continuously while the gantry is rotated, which eliminates dead time in between the delivery of individual fields. The methods presented in this thesis for robotic radiotherapy and IMPT also aim at reducing the dead time during treatment delivery, but this could be reduced further by introducing continuous irradiation techniques. For the CyberKnife,

this would involve continuous irradiation while adjusting the orientation of the linachead and the position of the robotic manipulator. Robot motion will probably still require more time than gantry rotation in VMAT treatments, but the non-coplanar irradiation in CyberKnife treatments has been shown to provide favorable dose distributions compared with coplanar VMAT irradiation (69, 152). For IMPT, this would involve the use of continuous scanning instead of delivering discrete pencil beams, possibly in combination with continuous gantry rotation (78, 153, 154). However, rotational proton delivery will result in large healthy tissue volumes to receive a low dose, which could have been completely spared (no dose deposition) when applying IMPT from a limited number of beam directions. Switching between proton energy layers is likely to always require a certain amount of dead time, but this dead time is minimized by vendors' efforts to shorten energy switching times. Switching between energy layers can be avoided completely by using the plateau dose region of high-energy proton beams, instead of the Bragg peak (78). The physical advantage of the Bragg peak is then no longer exploited, but the plateau region still has more favorable depth-dose characteristics compared with photon beams. The use of the proton plateau will also reduce the sensitivity of the treatment plan against treatment errors, as the delivered dose is not affected by uncertainties in proton range. However, the use of very high energy protons (e.g. 350 MeV) and beam collimation is recommended in order to achieve a lateral penumbra comparable to the penumbra of high-precision robotic radiotherapy, especially at larger depths (78).

10.2 Efficiency of treatment planning

Besides improving the efficiency of treatment delivery, the efficiency of treatment planning was improved for IMPT in chapter 5 of this thesis. For treatment modalities that can achieve a high degree of modulation by using many individual beams, the optimization time can become very long as it typically scales polynomially to the power 2.5–3.0 with the number of variables in the optimization. Although the total planning time will increase not as strongly due to overhead-processes, a large number of variables can still result in a considerable elongation. The use of the iterative 'resampling' approach, in which a relatively small sample of beams is optimized in each iteration, was shown to result in an improved planning efficiency compared with optimizing a large number of beams at once. The method was initially developed for the CyberKnife (33), but in this thesis it was extended to enable multi-criteria treatment planning for proton therapy as well. As a result, the planning time required to generate an IMPT treatment plan with a certain plan quality was shortened by a factor of 2.8–5.6. Inversely, the plan quality

achievable in a certain planning time was improved. Resampling additionally allows for the plan quality to be monitored during treatment planning and the planning process to be terminated when no further improvements are expected. In this way, the planning efficiency can be improved even more. The resampling technique is a method that allows for many degrees of freedom to be dealt with in an efficient way. This could be of particular interest for future proton systems with very fast energy switching times, which enable the use of many energy layers during treatment delivery and require them to be included in the plan optimization. Also in 4D dose optimization, in which the breathing phase or the time of delivery of each beam is added as an additional dimension (155-157), resampling can play a role to efficiently deal with the strong increase in the degrees of freedom.

Very fast treatment planning in combination with fully automatic treatment planning using our in-house developed treatment planning system Erasmus-iCycle, might allow for so-called 'online adaptive treatment planning'. In online adaptive treatment planning, the treatment plan is updated just before the start of a treatment fraction to the patient anatomy of that moment, for example acquired using an in-room CT scanner. The latest CyberKnife system at our institute is already equipped with a diagnostic in-room CTscanner, as well as the latest IMPT gantry at the Paul Scherrer Institute in Villigen, Switzerland (91). Online adaptive treatment planning will allow patients to receive a truly personalized treatment which is tailored to his/her daily anatomy. This involves adapting the plan to changes in organ shape, tissue density and in the future potentially to changes in biological characteristics. In this way, the best possible dose distribution with the lowest dose in healthy tissues is achieved for each treatment session, thereby reducing the chance of complications.

Online adaptive treatment planning is challenging in terms of calculation time, workflow and logistics. To preserve treatment efficiency and to minimize residual treatment errors, the entire plan adaptation process needs to be completed within 1-2 minutes. Online adaptive treatment planning therefore requires anatomical information to be acquired just before the start of the treatment fraction and subsequent treatment plan generation to be very fast. Current treatment planning algorithms are typically not fast enough to facilitate online plan adaptation. The iterative resampling method can provide part of the required speed-up in treatment planning times. Next to that, its iterative planning approach might be especially suited for treatment plan adaptation. The method intrinsically tries to improve the plan by adapting the previous solution (by iteratively adding and reoptimizing a new random sample of beams). This existing solution can be a previous planning iteration (as implemented in this thesis), but also a previous treatment fraction. When performing adaptive treatment planning using resampling, the treatment plan does not have to be generated from scratch. A small sample of beams can be added to an existing treatment plan, after which the plan is reoptimized. This will allow the

mathematical optimization to start from a solution that is relatively close to the optimal treatment plan, instead of optimizing from scratch. In this way, it might be possible to adjust an existing treatment plan very fast to a new patient anatomy. The effectiveness of this approach will be established in a current research project at the Erasmus MC Cancer Institute within the HollandPTC-consortium ('ADAPTNOW: High-Precision Cancer Treatment by Online Adaptive Proton Therapy')

10.3 Treating moving tumors

The highly localized dose delivery in high-precision robotic radiotherapy and IMPT make these treatments relatively susceptible to treatment errors. This can result in a severe deterioration of the delivered dose, for example when the tumor moves during a treatment fraction (intrafraction motion). Tumors in or close to the thorax move due to respiration (and to a lesser extent due to the heart beat), while tumors in the abdomen generally move due to peristaltic motion and organ filling.

In chapter 4, the impact of intrafraction prostate motion was investigated for robotic radiotherapy using the CyberKnife. During a treatment fraction, the system acquires images with a certain frequency using the integrated stereoscopic X-ray imaging system and it subsequently corrects for prostate displacements by adjusting the position of the robotic manipulator. However, it was shown that a prostate can display very rapid excursions of considerable size and these can remain undetected in between the acquisition of consecutive images (20, 21). Moreover, the prostate was also observed to rotate considerably, whereas rotational corrections can only be performed up to a certain extent (56). We found that a time interval between imaging/corrections of 60 to 180 s and rotational corrections (in combination with translational corrections) up to 5° or 10° were very effective to maintain plan quality. Interestingly, it was also found that very high imaging and correction frequencies could result in a worsening of the delivered dose when treatment errors were not perfectly corrected for. In that case, additional imaging resulted in an elongation of the fraction duration and consequently in larger prostate displacements, which were then not fully accounted for by robot corrections. Very short time intervals might be beneficial when additional imaging would not result in longer treatment times, for example when using MR-guidance, ultra-sound or transponders implanted in the prostate (20, 21, 158, 159). However, the results of this study also show that fairly long time intervals between imaging/corrections already provide sufficient geometrical information to achieve adequate tumor irradiation in CyberKnife prostate treatments.

Tumors that move with respiration display periodic motion instead of the gradual displacements that are typically observed for tumors in the abdomen. This can make the treatment of these tumors even more challenging, especially for IMPT. The timing of spot delivery in IMPT treatments can result in the respiratory motion to interfere with the delivery. This so-called 'interplay' effect causes the relative spot positions to change, giving rise to dose hot-spots and cold-spots within the target volume. Next to that, tumors that move with respiration are often located in regions with large density heterogeneities (e.g. ribs and lungs), which enhance the dose-deteriorating effect of respiratory motion in IMPT treatments.

There are basically three ways to account for respiratory motion in IMPT. The first approach is to reduce impact of respiratory motion by applying a statistical averaging out of the position error at each spot position. An example of this approach is the 'rescanning' technique (80, 160). During rescanning, the dose is delivered in multiple rescans, each time delivering a fraction of the protons at every spot position. The average position error, and the impact on the dose distribution, tends to decrease with an increasing number of rescans. The treatment plan and delivery are both time-independent and information about the tumor position in time is therefore not required. The motion amplitude or a 4D CT-scan should be used during treatment planning to ensure that the moving tumor is fully covered in all breathing phases. The second strategy is to minimize position errors by adjusting or accounting for respiratory motion during treatment delivery, thereby aiming at preserving the planned 3D dose distribution. Examples of this approach are 'real-time tumor tracking' (127, 161), in which the moving tumor is followed by adjusting the spot positions in real-time, or 'gating' (121), in which spots are only delivered in a certain breathing phase or when the tumor is located close to a certain position. The treatment plan is time-invariant whereas the delivery is time-dependent and knowledge of the tumor position during treatment delivery is thus required. The final strategy is to perform '4D plan optimization' by including a 4D CT-scan (or 4D MRI data) in the treatment planning process and adding a time-component to each spot (155-157). Intensity modulation can then be performed in an additional fourth dimension (time) and this allows for the plan quality to be improved by exploiting the changes in anatomy during respiration. For example, the dose received by the surrounding organsat-risk (OARs) can be reduced when certain regions of the tumor are only irradiated in breathing phases in which the distance to an OAR is large. As both the treatment plan and delivery are time-dependent, detailed knowledge about the tumor position during delivery is required.

In chapter 8 we investigated the effectiveness of real-time tumor tracking in IMPT by performing dose simulations for mathematical phantoms. The target was assumed to translate uniformly according to a cos⁴-function and it was irradiated from a single beam direction. We found that tracking errors smaller than 1 mm (root mean square) were required for the homogenous phantom in order to obtain acceptable target dose homogeneity. The sensitivity of tracking to residual position errors could be reduced considerably by combining tracking with rescanning, which was designated as 'retracking'. The use of rescanning alone was most favorable for the heterogeneous phantom, as tracking could cause the pencil beam shape to change in the presence of tissue heterogeneities. Similar observations were recently reported by Zhang et al. (161) in extensive simulations for liver cases using realistic 4D CT-MRI data (CT data deformed according to 4D MRI). Improved sophistication of the tracking mode (2D to 3D tracking) resulted in better target dose homogeneity and conformality, but retracking with a rescan factor of 3 was required to achieve acceptable treatment outcomes in realistic patient cases. However, both studies were performed using a single beam direction and the use of multiple beam directions will allow for a reduction of the rescan factor or for additional rescanning to become superfluous (160). The use of larger spot sizes and multiple treatment fractions will similarly reduce the impact of respiratory motion (162). Tracking and retracking were observed to result in increased normal tissue doses in the beam-entrance region, which is described as the so-called 'inverse interplay' effect. This effect can be avoided by using 4D plan optimization, which allows for OARs and normal tissue to be better spared while adequately irradiating the moving tumor. Both real-time tracking and 4D plan optimization require accurate information about the tumor position during treatment delivery, but uncertainties in time or position will always remain. The impact of inaccuracies can be reduced by additionally applying rescanning, as discussed above, but these uncertainties might also be accounted by using robust treatment planning techniques. Instead of including errors in patient setup or proton range in the plan optimization, as typically done, errors in time or anticipated motion amplitude might also be included as scenarios during robust planning (155, 156). This could be an interesting topic for future research.

10.4 Robust treatment planning

Besides intrafraction tumor motion, the delivered dose in IMPT can also be affected substantially by treatment uncertainties that occur on a longer time-scale, such as uncertainties in patient anatomy, patient setup (i.e. the alignment of the patient with respect to the radiation beams) and proton range. Anatomical changes and setup errors can vary between treatment fractions (interfraction uncertainties), whereas errors in the anticipated proton range are typically constant throughout the entire treatment (systematic uncertainties). Treatment uncertainties are in conventional radiotherapy generally

accounted for by applying a safety margin around the target volume and possibly around OARs as well. This assures adequate dose delivery when treatment errors up to a certain extent occur (within the safety margin). However, safety margins are considered to be inadequate for IMPT, as they do not account for the uncertainties in proton range. This causes the dose to change within the target volume, which cannot be compensated by a mere expansion of the target volume. This was confirmed in chapter 7, in which we investigated the sensitivity of IMPT against uncertainties in anatomy, setup and range, for treatment plans that were generated using a 5 mm safety margin around the target volumes. For 10 head-and-neck cancer patients, we found that a margin could not fully prevent the dose to deteriorate, even when errors were smaller than the applied margin. The dose distribution was mainly affected by setup errors and to a lesser extent by anatomical errors, for which the impact was strongly patient-dependent. The use of margins in proton therapy can help to reduce the impact of treatment uncertainties, but only in a beam-specific fashion and in combination with a uniform dose for each beam direction (single-field uniform dose, SFUD) (163). However, this will result in the beneficial characteristics of IMPT to be sacrificed as it reduces the degrees of freedom during plan optimization.

To account for treatment uncertainties in an effective way, the use of 'robust' treatment planning techniques is proposed for IMPT (99). During robust planning, possible errors in patient setup and proton range are explicitly included in the mathematical optimization, thereby ensuring adequate irradiation when these error scenarios occur. Different approaches to achieve robustness have been described in literature, of which two can be considered most suitable: (1) probabilistic treatment planning (82) and (2) minimax or worst-case optimization (88, 89). In probabilistic planning, the expected value of the objective function or of individual objectives is optimized assuming a certain distribution of treatment errors. The error scenarios included in the optimization are either randomly sampled from the error distribution or evenly distributed and weighted according to their probability. The result is a treatment plan that gives on average the best quality for a population. However, it does not provide robustness against extreme errors. In minimax or worst-case optimization, on the contrary, the included error scenarios are presumed to describe extreme treatment errors or a certain level of extremity (e.g. 85% confidence interval). The worst value of the objective function or of individual objectives is subsequently optimized, instead of weighing the error scenarios according to their probability. Minimax optimization provides a lower bound on the plan quality, whereas probabilistic treatment planning optimizes the most likely outcome. Both strategies only optimize physically feasible dose distributions and the optimization time increases with the number included error scenarios. A relatively high number of scenarios is required in probabilistic treatment planning in order to approximate the probability distribution

with sufficient accuracy. This could explain why minimax optimization is mostly used in (commercial) IMPT treatment planning software.

The advantages, limitations and applicability of both robust methods are not clearly established yet. Firstly, it is unclear whether one of the methods should be preferred in terms of dosimetric plan quality and robustness. Probabilistic treatment planning does not provide robustness against extreme treatment errors, whereas minimax optimization can be regarded as conservative since it accounts for error scenarios that are not very likely to occur. One could also think of a complementary approach, in which the robust methods are applied in an objective-specific or organ-specific fashion. Serial organs and the tumor are typically associated with (near-)minimum dose, (near-)maximum dose or coverage constraints, which should strictly respect a certain value. It might be beneficial to optimize these using the minimax method. Parallel organs, generally associated with mean-dose or dose-volume objectives, could then ideally be optimized using probabilistic planning. A complication-specific approach could also be thought of, optimizing severe complications with the minimax method and less severe complications using probabilistic planning. Secondly, there is no clear understanding of how systematic and random setup uncertainty distributions should be translated into a certain amount of robustness. A traditional margin used in photon therapy can therefore not be converted into equivalent robustness settings and corresponding errors scenarios. In addition, it is unclear whether the included (extreme) error scenarios also provide sufficient robustness against intermediate errors, i.e. errors smaller than the errors accounted for in the robust optimization, which is implicitly the case when using traditional margins in photon therapy. We investigated this by performing dose recalculations for robustly optimized treatment plans while introducing intermediate errors (see Appendix 9.A). On average, adequate target irradiation was preserved for intermediate setup errors, whereas intermediate range errors resulted in a slight increase of the maximum dose in the target volume. However, this should be studied more thoroughly in realistic treatment simulations, as we only included a limited number of error scenarios and ignored treatment fractionation. Finally, since robust methods currently only incorporate setup errors and range errors into the optimization, it is still unknown how interfraction variation in patient anatomy can be accounted for during IMPT treatment planning. Anatomical changes introduce an additional uncertainty in proton range and this variation should therefore ideally be included in the robust plan optimization. However, a method to include anatomical uncertainties is still lacking. Research performed at our institute indicated that a combination of internal margins and robustness against setup errors can be used as a practical solution to account for interfraction motion in locally-advanced prostate cancer patients. An alternative and potentially more effective approach would be to use principle component analysis (PCA) to generate representative anatomical scenarios that can be included in the robust optimization. However, the PCA-model

should then be applicable for a population of patients and provide information about the water-equivalent depths of the organs, while current PCA-models typically only provide geometrical information (164, 165). The introduction of online adaptive treatment planning for IMPT would of course solve this problem of accounting for interfraction organ motion and deformation.

These issues regarding robust treatment planning should be addressed in the very near future, as they hinder the clinical introduction of IMPT for a number of treatment sites. In addition, these issues currently obstruct a fair comparison between IMPT and state-of-the-art photon techniques, in which treatment plans for both modalities are equally robust.

10.5 Cost-effectiveness of intensity-modulation proton therapy

IMPT enables the tumor to be accurately targeted while reducing the amount of dose deposited in surrounding healthy tissue, thereby reducing the risk of complications. However, IMPT treatments are currently also associated with higher costs compared with conventional radiotherapy treatments. Costs of investment and operation are higher than in conventional radiotherapy and are generally not compensated by an increased number of treated patients. As a result, an IMPT treatment fraction is on average approximately 3 times more expensive than a conventional treatment fraction (743 vs. 233 euro), but this can vary from a factor of 0.9 to 3.3 when comparing specific treatments with intensity-modulated radiotherapy (16000 vs. 18000 euro for a prostate cancer treatment and 40000 vs. 12000 euro for a head-and-neck cancer treatment) (24). This has initiated fierce debates on whether the higher treatment costs of IMPT can be justified by an increase in treatment quality and a reduction in side-effects, and on how the evidence for this should be provided (166, 167). This debate is further fueled by the relatively high sensitivity of IMPT to uncertainties in patient position, patient anatomy and proton range. This sensitivity affects the effective and safe delivery of the tumor dose, and requires additional measures to preserve the intended dose distribution, for example by including robustness in the generation of the treatment plans.

In chapter 9, we investigated the price of robust plan optimization as a measure to reduce the impact of treatment uncertainties in oropharyngeal cancer patients, in terms of the dose received by OARs and the consequent normal tissue complication probabilities (NTCPs). Robust optimization was verified to indeed result in a reduction of dose uncertainties for treatment errors up to a certain extent, depending on the degree of robustness. We found that a higher degree of setup robustness could result in considerably higher NTCP-values, especially in bilateral cases. On the contrary, the increase in OAR doses and NTCP-values as a result of a higher degree of range robustness was very small. These results suggest that the use of advanced in-room imaging techniques (which can be expensive) is beneficial to reduce the impact of setup uncertainties, but that it might not be cost-effective in dealing with range errors for this particular patient group. It shows that it can sometimes be more cost-effective to reduce the impact of errors, than to reduce the errors themselves, when the latter involves high costs. These considerations are of importance when improving the cost-effectiveness of a certain radiotherapy treatment modality.

Besides improving the cost-effectiveness of an individual radiotherapy treatment modality, one should also compare the cost-effectiveness of different modalities. The relatively high treatment costs of IMPT should be justified by a sufficiently high gain in plan quality or expected treatment outcome. The treatment modality with the highest cost-effectiveness should then ideally be applied. It can be desirable to make a patientspecific comparison between modalities, when substantial differences between patients in a patient-group exist or when the difference in treatment costs is large. In the Netherlands, a patient-specific comparison between modalities will be mandatory when proton therapy is considered for certain tumor sites. Proton therapy for these so-called 'modelbased indications' will only be reimbursed by insurance companies when NTCP-models predict a substantial gain in expected outcome compared with state-of-the-art photon therapy (168). Automated treatment planning using Erasmus-iCycle, in combination with the efficient planning methods presented in this thesis, is an ideal tool to make these comparisons. It allows for high-quality treatment plans to be generated very efficiently for multiple treatment modalities and enables an objective comparison between plans on a patient-specific basis. Erasmus-iCycle is currently capable of generating treatment plans for IMRT, VMAT, CyberKnife (using cones or mMLC) and IMPT, and has already been used for oropharyngeal cancer patients to automatically generate an additional IMPT treatment plan for comparison purposes. It is of importance for such a comparison to be fair. This means that all plans should have an equal clinical endpoint in terms of tumor dose or tumor control probability (TCP) and that for all plans residual treatmentrelated uncertainties should be accounted for consistently. For example, the expected TCP should be equal for the treatment plans that are being compared on OARs doses.

The development and implementation of efficient methods to perform these treatment plan comparisons between modalities allow for further personalization of radiotherapy treatments. This will also require reliable toxicity prediction models and decision support systems (169). In addition, clinical studies should be conducted in order to verify the modelled gain in clinical practice. Ultimately, it will enable the best treatment to be delivered for a specific patient, while minimizing treatment costs. This is of utmost importance for a society in which the costs of health care are under pressure. In this the-

sis, we have aimed at improving both efficiency and accuracy of high-precision robotic radiotherapy and IMPT. In this way, this work contributes to an improvement of the cost-effectiveness of high-precision radiotherapy.

Appendices

3.A Beam list generation

The beam list contained for a specific patient all available candidate beams that could be used for treatment planning. The direction of a candidate beam is defined by the coordinates of the node position and the coordinates of a target point. The field size of the candidate beam is defined by the selected collimator size. To compute the target points for each node position, a projection of the PTV was made onto a reference plane at 800 mm from each node position. Target points were rectangularly distributed over the reference plane with a spacing corresponding to a 1 mm spacing at the isocenter. Per target point, 12 candidate beams were added to the beam list; one beam for each collimator size. For each possible candidate beam, the area of the beam projection (50% isodose line) located outside the PTV projection was stored as well in the beam list, which is used to improve the efficiency of resampling. Node positions were disabled in advance from the beam list when having beams running through the first or last CT slice or through some critical organs for which the passing of the beams was not allowed, such as the eyes (head-andneck case) and the scrotum (prostate case). Candidate beams were selected by randomly sampling a user-defined number of beams from the beam list. In this study, candidate beam sets of 5000 beams per iteration were used, as this number of beams provided a practical balance between plan quality and optimization time.

3 B Dose calculation

Dose calculations were performed to construct individual dose matrices for each structure involved in the computer optimization, containing the dose deposited (in Gy/MU) in each voxel by every selected candidate beam. The dose was calculated using the tissue

maximum ratio/off axis ratio formalism for data measured in a water phantom and density heterogeneities were accounted for by calculating equivalent path lengths. The use of separate dose matrices for each structure allowed the grid spacing to be adjusted to the size of the structure. Default grid spacing was set to $3 \times 3 \times 3$ mm. Smaller grid spacing was used to ensure sufficient coverage or sparing of small structures. Larger grid spacing was used for large structures such as the lung. In structures outside the PTV on which a maximum-dose constraint was imposed, only voxels at the periphery facing the PTV were considered, as the maximum dose is very likely to be located in these voxels. To generate the skin structure, 3000 voxels were randomly sampled from a large structure, which was segmented using $5 \times 5 \times 3$ mm voxels at 10 mm from the skin in regions where beams could enter the patient. A large area of skin could be taken into account in this way, without having too much skin voxels in the dose calculation and inverse optimization. The high-resolution recalculation of the dose in the entire patient was performed for the final treatment plan using a dose grid spacing of $1 \times 1 \times 1.5$ mm.

3 C Treatment time calculation

Each time a feasible treatment plan had been obtained, the treatment time per fraction was estimated using a dedicated algorithm. The treatment time estimation included the beam-on time, the robot motion time in between node positions and beams, and the time needed for image acquisition. It did not include the time needed for patient setup and time needed to compensate for tumor motion during the treatment. The algorithm was capable of determining the robot path length connecting all the node positions in a treatment plan while skipping those node positions that could be skipped. Moreover, the most time-efficient order of beam delivery was calculated for each node position, requiring minimal linac orientation adjustments. The model parameters (robot speed, linac angular speed, and time per image acquisition) were obtained by fitting the model to treatment times obtained by the vendor's time estimator for six treatment plans. The treatment time estimator developed by the vendor is based on the exact known robotics (acceleration and deceleration parameters), linac output and imaging procedures. Robot speed of the CyberKnife robotic radiosurgery system version G4 (Accuray Inc.), linac output of 800 MU/min, and image pairs acquired every three beams were assumed in the treatment time calculations. A measure of the accuracy of the fitted model was obtained by performing a 'leave-one-out' test. For all six treatment plans, the fraction duration was estimated based on a model fitted using the remaining five treatment times from the vendor software. The obtained fraction durations were then compared to the fraction

Table 4.A.1: Mean values (and standard deviations) of the simulated prostate motion and the residual errors during beam-on time for different time intervals and corrections strategies. The mean values and standard deviations were calculated across all data points, which included all prostate cases, motion data patients and margins. The reported treatment times per fraction are averaged over all prostate cases and both margins.

Standa	rd plans	3D Trans	lations [mm]		Rotation	s [deg]	
Time interval [s]	Treatment time [min]	Prostate motion	Residual error T(full)	Prostate motion	R(≤5°)	Residual erro R(≤10°)	r R(full)
No imaging	28.2	3.2 (2.6)	-	-2.1 (5.9)	-	-	-
15	45.4	4.5 (3.6)	0.5 (0.8)	-3.5 (8.2)	-1.8 (5.9)	-0.9 (4.2)	0.0 (1.6)
60	30.9	3.5 (2.8)	0.7 (1.1)	-2.2 (6.5)	-0.8 (4.3)	-0.3 (3.1)	0.0 (2.3)
180	29.1	3.4 (2.7)	1.0 (1.3)	-2.0 (6.4)	-0.8 (4.4)	-0.3 (3.3)	-0.2 (2.8)
360	28.7	3.5 (2.9)	1.4 (1.4)	-1.8 (7.0)	-0.7 (5.1)	-0.4 (4.0)	-0.3 (3.2)

Adapti	ve plans	3D Trans	ations [mm]		Rotation	s [deg]	
Time interval	Treatment time	Prostate	Residual error	Prostate		Residual erro	
[s]	[min]	motion	T(full)	motion	R(≤5°)	R(≤10°)	R(full)
No imaging	27.0	3.1 (2.5)	-	-2.1 (5.8)	-	-	-
60-On	29.0	3.3 (2.6)	0.6 (0.9)	-2.2 (6.0)	-0.8 (3.9)	-0.3 (2.7)	0.0 (1.8)
60-Off	28.5	3.3 (2.6)	0.7 (1.1)	-2.1 (6.1)	-0.7 (4.0)	-0.2 (2.9)	0.0 (2.3)

Abbreviations: T = translational corrections, R = rotational corrections.

durations obtained by the vendor's time estimator. The average deviation was found to be -0.6%, ranging from -1.3% to 0.7%.

4 A Prostate motion characteristics

The characteristics of the simulated prostate motion are given in table 4.A.1. The table lists the mean values and standard deviations of the translations in 3D and the rotations around the left-right axis, depending on the simulated time interval. The table also contains the residual translational and rotational errors during beam-on time, depending on the time interval and correction strategy. The table shows that the simulated prostate motion was largest for a time interval of 15 s, due to the considerably longer treatment time per fraction compared with the other time intervals. The residual translational errors always decreased with a decrease in time interval, as prostate translations were simulated to be fully corrected after image acquisition. This pattern was also observed for the residual rotational errors, when prostate rotations were perfectly corrected. However, when prostate rotations were not fully corrected, residual rotational errors could increase with

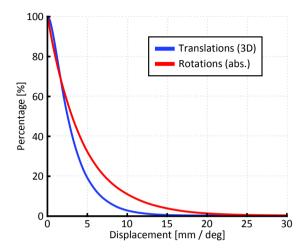


Figure 4.A.1: Cumulative histogram of the simulated prostate translations (in 3D) and rotations (absolute values) for standard treatment plans without robot corrections (and consequently without image acquisition). The average treatment time per fraction was 28.2 min. For clarity, the horizontal axis was limited to 30 mm / 30°.

a decreasing time interval (15 s), due to the longer treatment time per fraction for shorter time intervals. Enabling the adaptive timing mode resulted in smaller residual errors, but the difference was very small compared with the adaptive mode being switched off.

Figure 4.A.1 depicts a cumulative histogram of the translations in 3D and the absolute values of the rotations around the left-right axis, for treatments simulated without robot corrections (and thus without image acquisition). The figure shows that rotations of more than 5° were observed in 32% of the data points and rotations of more than 10° in 11% of the data points.

The prostate rotations around the left-right axis were obtained by converting the anterior-posterior translations into rotations around the prostate apex (see figure 4.1), which might have introduced errors compared with actual prostate rotations. Intrafraction prostate rotations were described in literature by Deutschmann *et al.* for conventional radiotherapy (56). They found absolute rotations around the left-right axis of $2.5^{\circ} \pm 2.3^{\circ}$ (maximum 26.9°) after on average 7.7 min (maximum 15.1 min). When evaluating the prostate motion tracks used in the current study at 7.7 min (and 15.1 min), we observed absolute intrafraction rotations of $2.9^{\circ} \pm 2.6^{\circ}$ (and a maximum of 45.9°). This indicates that our method to derive prostate rotations resulted in outcomes that were similarly distributed as rotations observed during actual treatments. The maximum ro-

tation was substantially larger, but this might be an unreliable parameter for comparison as it only concerns a single data point. Also Aubry et al. reported intrafraction prostate rotations around the LR-axis, which they observed to be -0.5° ± 5.8° for treatment fraction times shorter than 5 min (53). However, as the time between measurements was not clearly mentioned in their paper, a direct comparison with our prostate motion data is not possible.

4 R Treatment plan characteristics

The characteristics of the treatment plans generated in the current study are listed in table 4.B.1. The table shows that the desired dose constraints for the targets were met in all treatment plans (i.e. PTV $V_{100\%} \ge 95\%$, CTV $V_{100\%} \ge 99\%$, PZ $V_{120\%} \ge 95\%$). The PTV coverage was not reported for the o-mm margin plans, since the PTV coincided with the CTV in these treatment plans. Dose distributions in the transversal plane of the 3-mm standard treatment plans are depicted in figure 4.B.1 for the three prostate cases. The figure clearly shows the simultaneously integrated boost that is delivered to the peripheral zone and the urethra dose that is constrained in the target volume.

For the same treatment plans, examples of simulated dose distributions are depicted in figure 4.B.2, for uncorrected intrafraction prostate motion. The 4D dose calculation was performed within a rectangular region encompassing all possible locations of the CTV and high-dose regions of the rectum, bladder and urethra. In the examples, the prostate tended to move posteriorly, which caused the dose to shift in the anterior direction accordingly. The interplay between prostate motion and dose delivery resulted in a degradation of the original dose distributions as well.

Table 4.B.1: Characteristics of the generated treatment plans. The V_{100%} of the PTV was not reported for the 0-mm margin plans, as the PTV coincided with the CTV. The conformity index (Cl) was defined as the ratio of the volume receiving 38 Gy and the volume of the CTV (0-mm margin) or PTV

(3-mm margin).

Structure Type PTV max [Gy]												
· -	1-0	0-mm	3-mm	ш	mm-0	ш	3-mm	μį	0-mm	E.	3-mm	Æ
IV max [G	Standard	Adaptive										
	y] 58.0	58.1	58.0	57.8	58.0	57.8	57.3	58.3	58.1	57.8	57.5	58.6
PTV V100% [%]	- [%]	,	95.5	92.8			8.96	8.96			96.3	96.4
CTV V100% [%]	[%] 99.4	99.5	100	100	9.66	9.66	100	100	99.4	99.4	100	100
Z V120% [%]	6.76 [%]	6.76	98.8	98.9	97.5	97.5	98.0	98.1	98.3	98.4	8.86	99.1
Rectum D1cc [Gy]	iy] 19.4	20.8	28.4	28.0	26.9	28.0	30.9	31.3	22.6	21.9	29.9	30.2
Rectal mucosa max [Gy]	y] 16.5	18.3	24.3	24.6	23.8	25.7	27.0	27.2	23.8	23.5	30.6	30.8
Bladder D1cc [Gy]	5y] 28.1	28.6	33.8	33.6	32.1	32.5	36.2	35.9	28.0	27.9	33.9	34.3
Urethra D5% [Gy]	iy] 40.6	40.5	40.6	40.7	40.4	40.5	40.3	40.5	40.6	40.6	40.4	40.3
Urethra D10% [Gy	Gy] 40.4	40.3	40.4	40.4	40.1	40.1	39.9	40.2	40.4	40.4	40.2	40.0
Urethra D50% [Gy	Gy] 39.2	39.1	39.2	39.2	35.3	35.9	37.9	38.1	39.4	39.2	39.1	39.0
Femur head (left) max [Gy]	y] 12.0	11.0	12.7	12.4	12.6	11.4	17.3	14.2	18.0	19.8	19.0	19.6
Femur head max [Gy] (right)	у] 19.2	18.1	17.1	16.4	15.6	19.1	16.4	17.2	10.7	13.6	15.2	16.3
	1.22	1.22	1.02	1.02	1.31	1.31	1.10	1.08	1.25	1.26	1.05	1.06
Nodes	20	15	19	18	27	16	25	20	18	21	19	17
Beams	257	220	251	250	500	224	166	207	280	278	279	250
MU	85118	79902	75977	77706	71301	68517	64655	62010	91785	103461	96712	87169

Abbreviations: PTV = planning target volume, CTV = clinical target volume, PZ = peripheral zone, CI = conformity index, MU = monitor units.

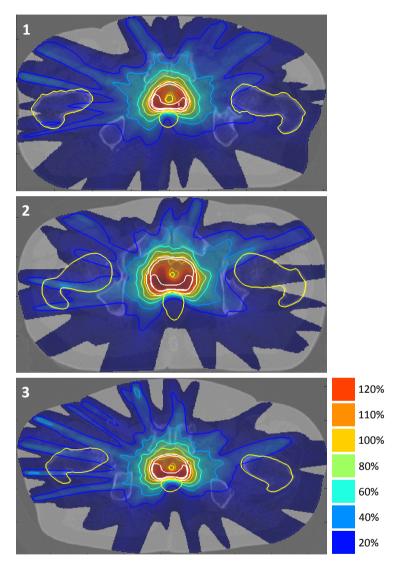


Figure 4.B.1: Dose distributions of the 3-mm standard treatment plans generated for the three patient cases. PTV, CTV and peripheral zone are depicted in white and OARs in yellow. Isodose lines indicate dose levels as a percentage of the prescribed dose (38 Gy).

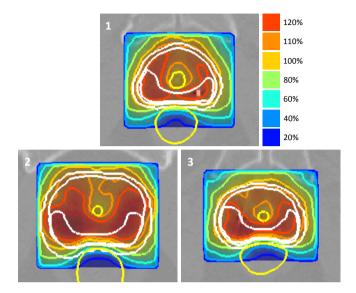


Figure 4.B.2: Close-up view of simulated doses within the region-of-interest for the three prostate cases in the presence of uncorrected intrafraction prostate motion, for the treatment plans shown in figure 4.B.1. The PTV, CTV and peripheral zone are depicted in white and the rectum and urethra in yellow. Isodose lines indicate dose levels as a percentage of the prescribed dose (38 Gy).

4. C Near-minimum CTV dose

The near-minimum dose ($D_{98\%}$) received by the CTV is depicted in figure 4.C.1. The outcomes for the CTV $D_{98\%}$ display the same patterns as observed for the CTV coverage (see figure 4.2). When no motion correction was applied, the CTV $D_{98\%}$ decreased from the average planned value of 38.8 Gy to 33.9 Gy (SD: 4.5 Gy, range: 14.9–39.0 Gy) for the standard o-mm margin plans and from the average planned value of 39.4 Gy to 37.1 Gy (SD: 3.3 Gy, range: 19.3–40.2 Gy) for the standard 3-mm margin plans.

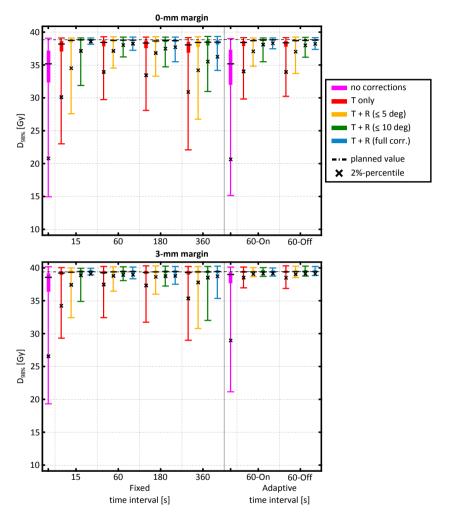


Figure 4.C.1: Near-minimum dose of the CTV (CTV $D_{98\%}$) of the 0-mm (upper graph) and 3-mm (lower graph) margin plans for different correction strategies. Each box plot indicates quartiles. Abbreviations: T = translational corrections, R = rotational corrections.

5.A Convergence

Pencil beams (PBs) were sampled randomly during resampling, which raises the question to what extent the planning outcomes depend on the random seed used. To answer this question, we repeated plan generation five times for a single patient (patient 3) using different random seeds. The results are presented in figure 5.A.1, for thin PBs and a sample size of 5000 PBs. The graph shows that solutions converge well, with the planscore being within 5% (of the maximum value) after two iterations, within 3% after three iterations and within 1% after six iterations.

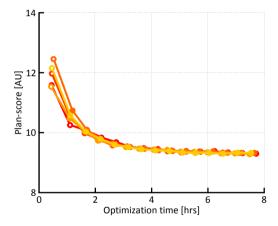


Figure 5.A.1: Plan-score as a function of the optimization time for patient 3. Five resampling runs were performed using thin PBs, a sample size of 5000 beams per iteration and different random seeds.

6. A Wish-lists for oropharyngeal and prostate cases

Table 6.A.1: Wish-lists containing the dose prescriptions for the oropharyngeal and prostate cases. The priority numbers indicate the order in which the objectives are optimized, a low number corresponds to a high-priority objective. A constraint has always to be fulfilled during treatment planning. CTV/ITVintermediate is a 10/15 mm transition region (within CTV/ITV-low) at the CTV/ITV-high and CTV/ITV-low interface. CTV/ITV-low' was constructed by subtracting CTV/ITV-intermediate from CTV/ITV-low.

	Or	opharyngeal cases		
Constrai	nts			
	Structure	Туре	Limit	Robust
	CTV-high	minimum	0.98 × 66 Gy	yes
	CTV-intermediate	minimum	0.98 × 54 Gy	yes
	CTV-low'	minimum	0.98 × 54 Gy	yes
Objectiv	es			
Priority	Structure	Туре	Goal	Robust
1	CTV-high	maximum	1.07 × 66 Gy	yes
1	CTV-intermediate	maximum	1.07 × 66 Gy	yes
1	CTV-low'	maximum	1.07 × 54 Gy	yes
2	CTV-rings (high-dose conformality)	maximum	0.90 × 54 Gy	no
3	Parotid glands	mean	0 Gy	yes
4	Submandibular glands	mean	0 Gy	yes
5	Spinal cord	maximum	20 Gy	yes
5	Brainstem	maximum	20 Gy	yes
6	Larynx	mean	0 Gy	yes
6	Oral cavity	mean	0 Gy	yes
7	Swallowing muscles	mean	0 Gy	yes
8	CTV-rings (low-dose conformality)	maximum	0 Gy	no
9	Logarithm energy-layer weight	sum	0	no
10	Total spot weight	sum	0 Gp	no

 $Abbreviations: CTV = clinical\ target\ volume; ITV = internal\ target\ volume; Gp = Giga-protons.$

Table 6.A.1: Wish-lists containing the dose prescriptions for the oropharyngeal and prostate cases. The priority numbers indicate the order in which the objectives are optimized, a low number corresponds to a high-priority objective. A constraint has always to be fulfilled during treatment planning. CTV/ITV-intermediate is a 10/15 mm transition region (within CTV/ITV-low) at the CTV/ITV-high and CTV/ITV-low interface. CTV/ITV-low' was constructed by subtracting CTV/ITV-intermediate from CTV/ITV-low. (continued)

	Prostate cases									
Constrai	nts									
	Structure	Туре	Limit	Robust						
	ITV-high	minimum	0.97 × 74 Gy	yes						
	ITV-intermediate	minimum	0.99 × 55 Gy	yes						
	ITV-low'	minimum	0.99 × 55 Gy	yes						
Objectiv	es									
Priority	Structure	Туре	Goal	Robust						
1	ITV-high	maximum	1.07 × 74 Gy	yes						
1	ITV-intermediate	maximum	1.07 × 74 Gy	yes						
1	ITV-low'	maximum	1.07 × 55 Gy	yes						
2	ITV-rings (high-dose conformality)	maximum	0.90 × 55 Gy	no						
3	Femur heads	maximum	50 Gy	yes						
4	Rectum	mean	0 Gy	yes						
5	Bladder	mean	0 Gy	yes						
6	Femur heads	mean	0 Gy	yes						
7	ITV-rings (low-dose conformality)	maximum	0 Gy	no						
8	Logarithm energy-layer weight	sum	0	no						
9	Total spot weight	sum	0 Gp	no						

 $Abbreviations: CTV = clinical\ target\ volume; ITV = internal\ target\ volume; Gp = Giga-protons.$

6.B Plan parameters of individual patients

Table 6.B.1: Worst-case target dose parameters and delivery parameters for all patients, treatment plans and fields.

			Oroph	aryngeal cas	e 1				
			Standard o	linical plan		Time-efficient plan			
Parameter		Beam 1	Beam 2	Beam 3	Total	Beam 1	Beam 2	Beam 3	Total
CTV-low	min [Gy]	0.0	3.9	0.0	44.8	0.0	4.8	0.0	45.8
CTV-low	mean [Gy]	21.4	21.4	21.9	64.3	16.5	24.2	23.8	64.4
CTV-low	max [Gy]	45.7	41.7	41.7	72.2	49.3	43.5	47.4	72.0
CTV-low	V _{95%} [%]	0.0	0.0	0.0	99.8	0.0	0.0	0.0	99.8
CTV-low	V _{107%} [%]	0.0	0.0	0.0	77.2	0.0	0.0	0.0	77.2
CTV-low	CI	0.00	0.00	0.00	1.59	0.00	0.00	0.00	1.57
CTV-low	HI-1	1.64	1.37	1.47	0.24	2.04	1.05	1.34	0.24
CTV-low	HI-2 [Gy]	29.3	23.1	27.2	14.7	25.8	20.9	27.8	14.7
CTV-high	min [Gy]	1.1	4.0	0.8	60.9	1.2	6.4	0.5	60.5
CTV-high	mean [Gy]	24.2	20.4	23.5	67.7	18.6	23.7	25.8	67.8
CTV-high	max [Gy]	45.7	38.7	40.0	72.2	49.3	40.7	47.0	72.0
CTV-high	V _{95%} [%]	0.0	0.0	0.0	99.5	0.0	0.0	0.0	99.6
CTV-high	V _{107%} [%]	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.3
CTV-high	CI	0.00	0.00	0.00	1.98	0.00	0.00	0.00	1.98
CTV-high	HI-1	1.26	1.22	1.20	0.08	1.72	0.98	1.13	0.08
CTV-high	HI-2 [Gy]	22.2	20.9	23.3	4.3	22.8	20.2	24.6	4.2
Energy layers		49	44	49	142	34	30	36	100
Spots		819	1474	850	3143	557	1217	792	2566
Total spot weight		1652	2051	1663	5367	1315	2165	1750	5230

Abbreviations: CTV = clinical target volume; ITV = internal target volume; CTV/ITV-low = low-dose target volume; CTV/ITV-high = high-dose target volume; CI = conformity index (volume receiving 95% of the prescription dose divided by the target volume); HI-1= heterogeneity index (($D_{2\%}$ - $D_{98\%}$)/ $D_{50\%}$); HI-2= heterogeneity index ($D_{5\%}$ - $D_{95\%}$);

Table 6.B.1: Worst-case target dose parameters and delivery parameters for all patients, treatment plans and fields. (continued)

			Oroph	naryngeal cas	e 2				
			Standard o	linical plan			Time-effi	cient plan	
Parameter		Beam 1	Beam 2	Beam 3	Total	Beam 1	Beam 2	Beam 3	Tota
CTV-low	min [Gy]	0.0	3.7	11.4	48.0	0.0	3.1	14.3	46.9
CTV-low	mean [Gy]	10.7	22.8	26.6	59.5	6.7	20.6	32.6	59.7
CTV-low	max [Gy]	36.8	41.8	44.4	71.2	33.1	36.6	51.3	71.8
CTV-low	V _{95%} [%]	0.0	0.0	0.0	99.4	0.0	0.0	0.0	99.6
CTV-low	V _{107%} [%]	0.0	0.0	0.0	32.5	0.0	0.0	0.0	32.5
CTV-low	CI	0.00	0.00	0.00	1.80	0.00	0.00	0.00	1.79
CTV-low	HI-1	5.71	1.49	1.05	0.30	9.54	1.00	0.98	0.29
CTV-low	HI-2 [Gy]	31.9	28.3	20.1	15.5	23.1	17.7	22.8	15.5
CTV-high	min [Gy]	4.2	9.9	17.1	64.9	0.5	9.4	19.4	64.8
CTV-high	mean [Gy]	21.7	21.9	27.1	68.1	13.5	20.7	37.1	68.
CTV-high	max [Gy]	36.8	41.8	39.9	71.2	28.3	36.6	51.3	71.8
CTV-high	V _{95%} [%]	0.0	0.0	0.0	99.4	0.0	0.0	0.0	99.3
CTV-high	V _{107%} [%]	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.3
CTV-high	CI	0.00	0.00	0.00	2.08	0.00	0.00	0.00	2.03
CTV-high	HI-1	1.51	1.33	0.89	0.08	2.10	1.07	0.77	0.09
CTV-high	HI-2 [Gy]	29.5	21.0	15.3	4.4	22.6	17.6	20.0	4.5
Energy layers		43	48	27	118	11	23	23	57
Spots		252	661	446	1359	113	529	603	124
Total spot weight		322	965	1242	2528	180	854	1522	255
			Oroph	naryngeal cas	e 3				
			Standard o	linical plan			Time-effi	cient plan	
Parame	eter	Beam 1	Beam 2	Beam 3	Total	Beam 1	Beam 2	Beam 3	Tota
CTV-low	min [Gy]	0.0	1.1	0.0	47.6	0.0	2.9	0.1	47.4
CTV-low	mean [Gy]	15.4	22.0	24.6	61.5	14.4	22.3	25.5	61.6
CTV-low	max [Gy]	47.8	43.4	56.3	71.6	50.0	43.8	57.9	71.9
CTV-low	V _{95%} [%]	0.0	0.0	0.3	99.6	0.0	0.0	1.0	99.6
CTV-low	V _{107%} [%]	0.0	0.0	0.0	53.9	0.0	0.0	0.0	53.6
CTV-low	CI	0.00	0.00	0.01	1.68	0.00	0.00	0.03	1.66
CTV-low	HI-1	2.91	1.36	1.89	0.27	3.19	1.35	1.81	0.27
CTV-low	HI-2 [Gy]	34.8	27.0	43.2	15.7	32.5	24.5	44.5	15.2
CTV-high	min [Gy]	0.2	9.5	2.1	62.4	1.1	9.8	2.1	62.5
CTV-high	mean [Gy]	14.2	25.4	30.0	67.7	13.3	24.8	31.5	67.
CTV-high	max [Gy]	43.0	42.5	56.3	71.6	37.0	43.2	57.9	71.9
CTV-high	V _{95%} [%]	0.0	0.0	0.0	99.8	0.0	0.0	0.0	99.8
CTV-high	V _{107%} [%]	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.3
CTV-high	CI	0.00	0.00	0.00	2.14	0.00	0.00	0.00	2.11
CTV-high	HI-1	2.85	0.87	1.60	0.08	2.53	1.02	1.57	0.07
CTV-high	HI-2 [Gy]	30.6	18.0	40.8	4.4	26.0	20.9	42.1	4.1
Energy layers		49	42	53	144	31	26	36	93
Spots		827	1524	963	3314	677	1247	917	284

Abbreviations: CTV = clinical target volume; ITV = internal target volume; CTV/ITV-low = low-dose target volume; CTV/ITV-high = high-dose target volume; CI = conformity index (volume receiving 95% of the prescription dose divided by the target volume); HI-1= heterogeneity index ($(D_{2\%}-D_{98\%})/D_{50\%}$); HI-2= heterogeneity index ($D_{5\%}-D_{95\%}$);

Total spot weight

Table 6.B.1: Worst-case target dose parameters and delivery parameters for all patients, treatment plans and fields. (continued)

			Oroph	naryngeal cas	e 4				
			Standard o	linical plan			Time-effi	cient plan	
Parame	ter	Beam 1	Beam 2	Beam 3	Total	Beam 1	Beam 2	Beam 3	Total
CTV-low	min [Gy]	0.0	7.9	0.6	50.5	0.0	7.1	0.3	50.4
CTV-low	mean [Gy]	11.3	26.3	20.8	57.4	10.2	20.8	27.6	57.4
CTV-low	max [Gy]	45.9	43.3	44.1	70.9	41.2	35.1	45.8	71.0
CTV-low	V _{95%} [%]	0.0	0.0	0.0	99.5	0.0	0.0	0.0	99.5
CTV-low	V _{107%} [%]	0.0	0.0	0.0	16.2	0.0	0.0	0.0	16.5
CTV-low	CI	0.00	0.00	0.00	1.85	0.00	0.00	0.00	1.79
CTV-low	HI-1	22.18	1.12	1.72	0.29	16.83	1.24	1.38	0.29
CTV-low	HI-2 [Gy]	41.6	24.9	34.2	15.2	36.5	19.6	40.4	15.1
CTV-high	min [Gy]	26.3	17.4	1.7	64.8	27.2	24.6	0.7	65.1
CTV-high	mean [Gy]	37.7	22.4	11.4	68.4	35.4	29.1	5.6	68.2
CTV-high	max [Gy]	45.9	35.2	26.0	70.9	41.2	35.0	15.1	71.0
CTV-high	V _{95%} [%]	0.0	0.0	0.0	99.0	0.0	0.0	0.0	99.3
CTV-high	V _{107%} [%]	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.1
CTV-high	CI	0.00	0.00	0.00	2.41	0.00	0.00	0.00	2.31
CTV-high	HI-1	0.77	0.74	2.64	0.09	0.32	0.45	4.05	0.08
CTV-high	HI-2 [Gy]	23.8	13.0	20.4	5.2	9.7	9.9	11.9	4.8
Energy layers		42	40	25	107	15	22	16	53
Spots		205	591	327	1123	166	459	425	1050
Total spot weight		324	835	887	2046	298	628	1143	2069
			Standard o	linical plan			Time-effi	cient plan	
Parameter		Beam 1	Beam 2	Beam 3	Total	Beam 1	Beam 2	Beam 3	Tota
CTV-low	min [Gy]	0.0	6.2	0.1	50.3	0.0	7.7	0.0	49.6
CTV-low	mean [Gy]	14.8	27.2	23.2	63.5	13.3	29.9	22.3	63.9
CTV-low	max [Gy]	44.4	43.4	52.8	73.0	43.5	51.5	51.9	72.1
CTV-low	V _{95%} [%]	0.0	0.0	0.0	99.6	0.0	0.0	0.0	99.7
CTV-low	V _{107%} [%]	0.0	0.0	0.0	69.3	0.0	0.0	0.0	68.8
CTV-low	CI	0.00	0.00	0.00	1.74	0.00	0.00	0.00	1.70
CTV-low	HI-1	5.43	1.10	1.67	0.25	5.55	1.18	1.72	0.25
CTV-low	HI-2 [Gy]	39.1	25.2	32.0	15.4	36.2	26.5	34.3	15.3
CTV-high	min [Gy]	0.3	7.8	0.4	62.3	0.1	7.9	0.3	62.9
CTV-high	mean [Gy]	17.8	27.4	25.1	68.0	16.5	32.7	21.7	68.4
CTV-high	max [Gy]	44.4	40.9	52.8	73.0	42.3	51.5	51.9	72.1
CTV-high	V _{95%} [%]	0.0	0.0	0.0	99.6	0.0	0.0	0.0	99.7
CTV-high	V _{107%} [%]	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.7
CTV-high	CI	0.00	0.00	0.00	2.33	0.00	0.00	0.00	2.25
CTV-high	HI-1	4.88	1.09	1.57	0.08	5.03	1.16	1.85	0.07
CTV-high	HI-2 [Gy]	39.0	23.9	34.5	4.2	36.5	27.0	37.2	3.8
Energy layers	38	51	29	118	13	20	19	52	
Spots		352	761	443	1556	278	631	433	1342

 $Abbreviations: CTV = clinical\ target\ volume; ITV = internal\ target\ volume; CTV/ITV-low = low-dose\ target\ volume; CTV/ITV-high = low-dose\ target\ volume$ high-dose target volume; CI = conformity index (volume receiving 95% of the prescription dose divided by the target volume); HI-1= heterogeneity index ($(D_{2\%}-D_{95\%})/D_{50\%}$); HI-2= heterogeneity index ($D_{5\%}-D_{95\%}$);

Table 6.B.1: Worst-case target dose parameters and delivery parameters for all patients, treatment plans and fields. (continued)

			Prostate ca:	se 1			
		St	andard clinical pl	an	Time-efficient plan		
Parameter		Beam 1	Beam 2	Total	Beam 1	Beam 2	Tota
ITV-low	min [Gy]	0.2	0.4	52.1	0.4	0.7	51.7
ITV-low	mean [Gy]	30.7	28.0	58.7	32.2	26.6	58.6
ITV-low	max [Gy]	58.9	62.1	79.6	59.5	59.3	80.9
ITV-low	V _{95%} [%]	1.4	2.1	99.4	0.9	1.3	99.4
ITV-low	V _{107%} [%]	0.0	0.0	12.4	0.0	0.0	12.3
ITV-low	CI	0.05	0.08	1.47	0.04	0.07	1.48
ITV-low	HI-1	1.64	1.80	0.41	1.44	2.06	0.41
ITV-low	HI-2 [Gy]	45.8	46.4	20.5	44.3	45.0	20.4
ITV-high	min [Gy]	18.2	23.4	71.4	19.3	23.1	71.7
ITV-high	mean [Gy]	31.2	45.2	76.4	30.8	45.7	76.3
ITV-high	max [Gy]	58.9	62.1	79.6	58.4	58.4	80.9
ITV-high	V _{95%} [%]	0.0	0.0	99.7	0.0	0.0	99.8
ITV-high	V _{107%} [%]	0.0	0.0	0.1	0.0	0.0	0.3
ITV-high	CI	0.00	0.00	1.51	0.00	0.00	1.56
ITV-high	HI-1	1.07	0.69	0.10	0.98	0.60	0.09
ITV-high	HI-2 [Gy]	22.6	22.7	6.4	19.8	20.4	6.0
Energy layers		50	54	104	33	39	72
Spots		1813	1780	3593	1608	1461	3069
Total spot weight		4070	4158	8228	4205	3904	8109

			Prostate cas	se 2			
		Sta	andard clinical pla	an	Т	ime-efficient pla	n
Parameter		Beam 1	Beam 2	Total	Beam 1	Beam 2	Tota
ITV-low	min [Gy]	0.3	0.1	50.6	0.8	0.6	51.2
ITV-low	mean [Gy]	29.7	29.0	58.7	35.4	23.5	58.7
ITV-low	max [Gy]	61.0	59.7	79.5	59.9	59.4	79.5
ITV-low	V _{95%} [%]	2.5	2.6	99.4	1.4	1.7	99.4
ITV-low	V _{107%} [%]	0.0	0.0	12.3	0.0	0.0	12.6
ITV-low	CI	0.09	0.09	1.49	0.08	0.06	1.50
ITV-low	HI-1	1.67	1.89	0.41	1.28	2.66	0.41
ITV-low	HI-2 [Gy]	48.6	48.5	20.4	47.0	46.8	20.6
ITV-high	min [Gy]	22.7	30.6	71.6	27.8	21.3	71.5
ITV-high	mean [Gy]	35.8	40.3	76.0	41.4	34.9	76.2
ITV-high	max [Gy]	48.9	58.3	79.5	55.6	50.1	79.5
ITV-high	V _{95%} [%]	0.0	0.0	99.5	0.0	0.0	99.5
ITV-high	V _{107%} [%]	0.0	0.0	0.0	0.0	0.0	0.0
ITV-high	CI	0.00	0.00	1.53	0.00	0.00	1.56
ITV-high	HI-1	0.53	0.49	0.09	0.56	0.68	0.09
ITV-high	HI-2 [Gy]	14.9	15.1	5.8	19.7	20.8	5.5
Energy layers		50	51	101	38	38	76
Spots		1735	1710	3445	1832	1475	3307
Total spot weight		3985	3882	7867	4487	3290	7777

Abbreviations: CTV = clinical target volume; ITV = internal target volume; CTV/ITV-low = low-dose target volume; CTV/ITV-high = high-dose target volume; CI = conformity index (volume receiving 95% of the prescription dose divided by the target volume); HI-1= heterogeneity index ($(D_{2\%}-D_{98\%})/D_{95\%}$); HI-2= heterogeneity index ($D_{5\%}-D_{95\%}$);

Table 6.B.1: Worst-case target dose parameters and delivery parameters for all patients, treatment plans and fields. (continued)

			Prostate cas	se s			
			andard clinical pla			Time-efficient pla	
Parameter		Beam 1	Beam 2	Total	Beam 1	Beam 2	Total
ITV-low	min [Gy]	0.6	0.1	51.5	1.1	0.2	51.5
ITV-low	mean [Gy]	29.0	29.4	58.4	29.3	29.1	58.4
ITV-low	max [Gy]	59.9	59.5	79.5	60.4	58.8	79.7
ITV-low	V _{95%} [%]	1.2	0.9	99.3	0.8	0.6	99.3
ITV-low	V _{107%} [%]	0.0	0.0	10.7	0.0	0.0	10.7
ITV-low	CI	0.05	0.05	1.46	0.04	0.04	1.47
ITV-low	HI-1	1.69	1.67	0.40	1.66	1.62	0.39
ITV-low	HI-2 [Gy]	46.4	46.3	19.0	45.8	45.8	19.3
ITV-high	min [Gy]	22.2	22.3	71.3	24.2	23.3	70.5
ITV-high	mean [Gy]	39.6	37.4	75.5	41.6	35.6	75.8
ITV-high	max [Gy]	55.8	55.0	79.5	56.1	53.4	79.7
ITV-high	V _{95%} [%]	0.0	0.0	98.7	0.0	0.0	98.8
ITV-high	V _{107%} [%]	0.0	0.0	0.0	0.0	0.0	0.0
ITV-high	CI	0.00	0.00	1.53	0.00	0.00	1.51
ITV-high	HI-1	0.69	0.75	0.10	0.64	0.71	0.09
ITV-high	HI-2 [Gy]	22.9	22.8	5.8	22.7	20.9	5.3
Energy layers		56	56	112	41	42	83
Spots		1989	1981	3970	2164	2136	4300
Total spot weight		4237	4295	8533	4235	4099	8334
			Prostate cas	se 4			
			andard clinical pla			Time-efficient pla	
Parameter		Beam 1	Beam 2	Total	Beam 1	Beam 2	Total
ITV-low	min [Gy]	1.2	0.4	51.7	1.5	0.8	52.1
ITV-low	mean [Gy]	28.9	29.9	58.5	29.4	29.4	58.6
ITV-low	max [Gy]	59.7	58.8	80.2	59.3	59.2	79.5
ITV-low	V _{95%} [%]	1.6	0.9	99.4	1.9	1.5	99.3
ITV-low	V _{107%} [%]	0.0	0.0	11.4	0.0	0.0	11.7
ITV-low	CI	0.10	0.09	1.57	0.10	0.09	1.58
ITV-low	HI-1	1.73	1.56	0.39	1.70	1.64	0.39
ITV-low	HI-2 [Gy]	48.1	48.2	19.6	48.3	48.6	19.9
ITV-high	min [Gy]	22.1	27.9	71.8	21.1	22.4	72.2
ITV-high	mean [Gy]	35.2	40.2	75.3	37.9	38.5	75.5
ITV-high	max [Gy]	49.6	56.8	80.2	56.7	59.2	79.5
ITV-high	V _{95%} [%]	0.0	0.0	99.4	0.0	0.0	99.5
ITV-high	V _{107%} [%]	0.0	0.0	0.1	0.0	0.0	0.0
ITV-high	CI	0.00	0.00	1.59	0.00	0.00	1.59
ITV-high	HI-1	0.60	0.51	0.09	0.76	0.74	0.09
ITV-high	HI-2 [Gy]	16.8	16.5	5.7	24.1	23.4	5.3
Energy layers		56	55	111	40	36	76
,							

Abbreviations: CTV = clinical target volume; ΠV = internal target volume; $CTV/\Pi V$ -low = low-dose target volume; $CTV/\Pi V$ -ligh = high-dose target volume; $CTV/\Pi V$ -low = low-dose target volume; $CTV/\Pi V$ -low-dose target volume; C

Spots

Total spot weight

Table 6.B.1: Worst-case target dose parameters and delivery parameters for all patients, treatment plans and fields. (continued)

			Prostate cas	ie 5			
		Sta	andard clinical pla	an	Time-efficient plan		
Parameter		Beam 1	Beam 2	Total	Beam 1	Beam 2	Total
ITV-low	min [Gy]	0.1	0.1	50.2	0.3	0.4	52.5
ITV-low	mean [Gy]	30.0	28.4	58.4	31.1	27.5	58.5
ITV-low	max [Gy]	59.5	59.5	79.6	59.6	59.3	80.0
ITV-low	V _{95%} [%]	8.6	11.3	99.1	8.6	10.1	99.1
ITV-low	V _{107%} [%]	0.0	0.0	11.0	0.0	0.0	11.4
ITV-low	CI	0.20	0.22	1.50	0.20	0.20	1.51
ITV-low	HI-1	1.72	2.01	0.40	1.65	2.09	0.41
ITV-low	HI-2 [Gy]	53.7	53.6	19.7	53.3	53.1	20.3
ITV-high	min [Gy]	26.1	22.9	71.6	27.4	19.9	71.4
ITV-high	mean [Gy]	38.2	38.9	75.5	41.4	35.7	76.0
ITV-high	max [Gy]	54.8	52.1	79.6	58.1	51.4	80.0
ITV-high	V _{95%} [%]	0.0	0.0	99.4	0.0	0.0	99.6
ITV-high	V _{107%} [%]	0.0	0.0	0.1	0.0	0.0	0.1
ITV-high	CI	0.00	0.00	1.57	0.00	0.00	1.57
ITV-high	HI-1	0.70	0.63	0.09	0.71	0.73	0.09
ITV-high	HI-2 [Gy]	21.6	21.5	5.6	24.4	23.0	5.7
Energy layers		58	51	109	39	38	77
Spots		2223	2137	4360	1996	1851	3847
Total spot weight		4608	4448	9056	4595	4244	8839

Abbreviations: CTV = clinical target volume; ITV = internal target volume; CTV/ITV-low = low-dose target volume; CTV/ITV-high = high-dose target volume; CI = conformity index (volume receiving 95% of the prescription dose divided by the target volume); HI-1= heterogeneity index ($(D_{2\%}-D_{95\%})/D_{50\%}$); HI-2= heterogeneity index ($D_{5\%}-D_{95\%}$);

6.C Performance of individual components

The generation of time-efficient treatment plans differed from the standard clinical plans on two aspects: (1) energy layer reduction and (2) pencil beam resampling. This raises the question to what extent these two components contribute individually.

To this purpose, we compared treatment plans that were generated using: (1) regular grid planning without energy layer reduction (standard clinical plan), (2) regular grid planning with energy layer reduction, (3) pencil beam resampling without energy layer reduction, and (4) pencil beam resampling with energy layer reduction (time-efficient plan). The results are listed in table 6.C.1. The use of energy layer reduction only (in combination with traditional regular grid planning) resulted in the number of energy layers lowered by 8% and 9% on average for the oropharyngeal and prostate cases, respectively. The use of pencil beam resampling only (without energy layer reduction) caused the number of energy layers to increase slightly for both the oropharyngeal cases and the prostate cases, compared with the standard clinical plan.

The substantial reduction in energy layers observed in this study (45% and 28%) can thus only be achieved by applying energy layer reduction in addition to pencil beam resampling. This is probably due to the iterative approach of pencil beam resampling, which allows for new spots to be added to the treatment plan, thereby giving room for a redistribution of spots within the remaining energy layers.

Table 6.C.1: Average worst-case dose parameters and delivery parameters for plans generated using different planning methods and with/without energy layer reduction.

Oropharyngeal case Parameter	s (n=5)	Standard clinical plan Regular grid without E-layer reduction	Regular grid with E-layer reduction	Resampling without E-layer reduction	Time-efficient plar Resampling with E-layer reduction	
rarameter		E-layer reduction	E-layer reduction	E-layer reduction	E-layer reduction	
CTV-low	V95% [%]	99.5	99.5	99.6	99.6	
CTV-high	V95% [%]	99.4	99.5	99.5	99.5	
CTV-high	V107% [%]	0.3	0.3	0.3	0.3	
Parotid glands	mean [Gy]	14.1	14.1	14.1	14.1	
Submandibular glands	mean [Gy]	34.7	34.7	34.7	34.7	
Spinal cord	max [Gy]	20.7	20.7	20.4	20.3	
Brainstem	max [Gy]	19.4	20.1	15.9	19.4	
Larynx	mean [Gy]	18.0	18.0	18.0	18.0	
Oral cavity	mean [Gy]	19.5	19.5	19.5	19.5	
Swallowing muscles	mean [Gy]	23.9	23.9	23.9	23.8	
CI-low		2.23	2.23	2.20	2.20	
CI-high		2.19	2.19	2.13	2.13	
Beam directions		3	3	3	3	
Energy layers		126	115	128	71	
Spots		2099	2281	1880	1809	
Total spot weight	[Gp]	3899	3907	3835	3855	
Delivery time (1 s)	[min]	3.9	3.8	3.9	3.0	
Delivery time (2 s)	[min]	6.0	5.7	6.1	4.2	
Delivery time (5 s)	[min]	12.3	11.5	12.5	7.7	

Prostate cases (n=5)		Standard clinical plan Regular grid without	Regular grid with	Resampling without	Time-efficient plan Resampling with
Parameter		E-layer reduction	E-layer reduction	E-layer reduction	E-layer reduction
ITV-low	V95% [%]	99.3	99.3	99.4	99.3
ITV-high	V95% [%]	99.3	99.3	99.4	99.5
ITV-high	V107% [%]	0.1	0.1	0.0	0.1
Femur left	max [Gy]	50.9	50.9	50.8	50.8
Femur right	max [Gy]	50.3	50.3	50.4	50.2
Rectum	mean [Gy]	26.3	26.2	26.2	26.2
Bladder	mean [Gy]	39.2	39.2	39.2	39.2
Femur left	mean [Gy]	27.9	27.9	27.9	27.9
Femur right	mean [Gy]	30.9	30.9	30.9	30.9
CI-low		1.61	1.61	1.61	1.62
CI-high		1.54	1.54	1.56	1.56
Beam directions		2	2	2	2
Energy layers		107	98	108	77
Spots		3835	3780	3646	3649
Total spot weight	[Gp]	8326	8312	8288	8169
Delivery time (1 s)	[min]	3.4	3.3	3.4	2.9
Delivery time (2 s)	[min]	5.2	4.9	5.2	4.2
Delivery time (5 s)	[min]	10.6	9.8	10.6	8.0

 $Abbreviations: \ CTV = clinical \ target \ volume; \ ITV = internal \ target \ volume; \ CTV/ITV-low = low-dose \ target \ volume; \ CTV/ITV-high = high-dose \ target \ volume; \ CI-low = low-dose \ conformity \ index; \ CI-high = high-dose \ conformity \ index; \ Gp = Giga-protons.$

7.A Optimization method

We have generated IMPT treatment plans using Erasmus-iCycle, an in-house developed multi-criteria treatment planning system (66, 110), which was extended with proton pencil beam scanning. The system performs 'prioritized' or 'lexicographic' optimization, i.e. treatment objectives are optimized successively according to their pre-selected priorities. The platform switches to optimizing the next objective with a lower priority when the desired or optimal value for the current objective is obtained and constrained. In this way, pareto-optimal treatment plans are produced. The user defines the hard constraints and the objectives with their priorities in the so-called 'wish-list'. For spinal cord and brainstem, we have done a step-wise optimization (66, 110). We have used the same wish-list for each patient in our patient group, resulting in fully automated treatment planning. Naturally, all treatment plans were also inspected visually. Our patient wishlist is displayed in table 7.A.1. The result of the optimization was a set of pencil beams with an energy and weight in Giga protons.

Table 7.A.1: Wish-list with planning aims for our head-and-neck cancer patient group. For the target and OARs, we have indicated (1) the direction and type of optimization problem (†to indicate maximize, \dot to indicate minimize), (2) whether minimum, maximum, mean or sum is optimized, (3) the dose objective, and (4) the priority numbers that indicate the order in which the objectives are optimized (a low number corresponds to a high-priority). A constraint must always be fulfilled during treatment planning.

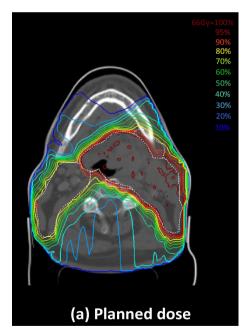
Objective	Direction	Туре	Value	Priority
PTV-66 Gy	↑	minimum	0.99 × 66	Constraint
PTV-54 Gy	↑	minimum	0.99 × 54	Constraint
PTV-66 Gy	↓	maximum	1.07 × 66	1
PTV-54 Gy	↓	maximum	1.07 × 54	2
PTV-rings (conformality)	↓	maximum	0	3
Spinal cord	↓	maximum	15	4
Brain stem	↓	maximum	15	4
Parotid glands	↓	mean	0	5
Submandibular glands	↓	mean	0	6
Swallowing muscles	↓	mean	0	7
Larynx	↓	mean	0	7
Oral cavity	↓	mean	0	8
Esophagus	↓	mean	0	9
Spinal cord	↓	maximum	0	10
Brain stem	↓	maximum	0	10
Unspecified tissue	↓	maximum	0	11
# Giga Protons	↓	sum	0	12

7.8 Example of the impact of anatomical changes

For patient 9, a large increase in hotspots in CTV-66 Gy was found. The reason for the large value of the $V_{107\%}$ is that the anatomy of patient 9 has changed substantially during treatment, leading to incorrect matching of the pencil beams in the target volume. There are two effects which contribute to the anatomical change:

- 1. The shrinkage of the primary tumor and nodal mass: the volume of the CTV-66 Gy decreases from 178 cm³ to 132 cm³, as can be seen from table 7.1. In figure 7.B.1 we pointed out the tumor shrinkage.
- 2. The patient has lost weight. This can be seen from figure 7.B.1, where we pointed out where the patient has lost weight.

In figure 7.B.1 (a) the planned dose distribution is shown, while in (b) the dose distribution recalculated on the bone aligned repeat CT is shown in the transverse plane. The anatomical changes are indicated.



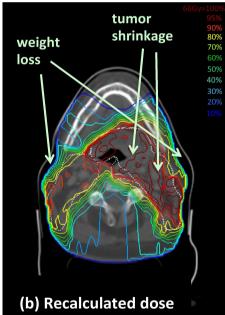


Figure 7.B.1: An example of a dose deviation caused by anatomical changes for patient 9. Planned (a) and recalculated (b) dose distribution, the latter displayed on the repeat CT scan. A large number of hotspots is seen in addition to underdosed areas (b).

7 C Example of the impact of anatomical changes

For patient 6, a large increase (9.2 Gy) in brainstem dose was observed. In figure 7.C.1 (a) the planned dose distribution is displayed, while in (b) the recalculated dose distribution is shown displayed on the repeat CT scan. The dose increase in the brainstem is evident from figure 7.C.1 (b). To understand what caused the increase, we displayed the planned and recalculated dose of the individual beam directions in (c) (d), (e), (f), (g), and (h). It becomes apparent from (e) and (f) that it is field 2, which contributed most to the dose increase in the brainstem. Changes in air-cavities and a change in the pose of the patient that occurred despite the bone alignment caused tissue changes in the beam path and a subsequent dose increase to the brainstem.

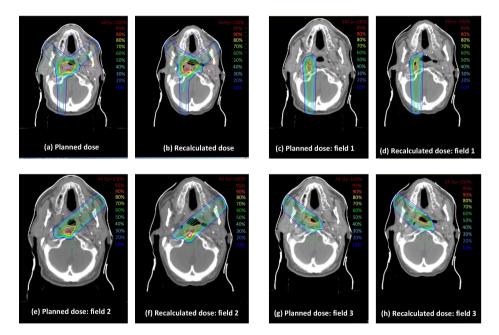
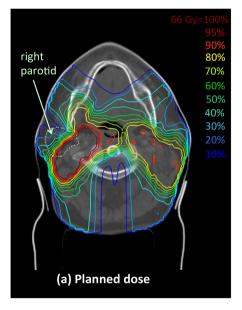


Figure 7.C.1: Dose deviation caused by anatomical changes for patient 6. Planned (a) and recalculated (b) dose distribution, the latter displayed on the repeat CT scan. Furthermore, planned (c) and recalculated (d) dose for field 1, planned (e) and recalculated (f) dose for field 2, and planned (g) and recalculated (h) dose for field 3.

7.D Example of the impact of combined errors

In patient 10, we found the largest increase in dose (11.3 Gy) to the parotid gland (right) for the scenario of combined errors. In figure 7.D.1 (a), the planned dose distribution is displayed together with the parotid gland. In figure 7.D.1 (b) is illustrated what happens when applying the planned dose on the repeat CT scan. Some weight loss is seen when comparing the planning and the repeat CT scan. The high dose gradient that spared the parotid gland in the planning CT scan was partly lost in the recalculated dose, resulting in a dose increase in the right parotid. In addition, inspection of the CT slices showed that the parotid gland shrank such that a larger part of the gland ended up in the high dose region. Finally, in this simulation, the dose was further increased by the simulated setup-error that moved the parotid gland into the high dose region.



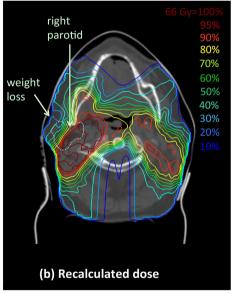


Figure 7.D.1: Dose deviation example for patient 10. Planned (a) and recalculated (b) dose distribution, the latter displayed on the repeat CT scan.

7.E Example of the impact of combined errors

In patient 8, we observed a large dose increase in the spinal cord in the scenario with all error sources combined. The reason for the large increase in dose in patient 8 is a beam overshoot, causing the proton pencil beams to reach the spinal cord. In figure 7.E.1 (a) the planned dose distribution is shown for 1 beam direction. In figure 7.E.1 (b) the delivered dose is recalculated in the repeat CT scan, where the reason for the beam overshoot becomes evident. The beam overshoot is caused by a decrease of the amount of material the proton beam traverses caused by weight loss. In our simulation the dose impact was amplified by a simulated range error of -3%. If the anatomical change would not be present, the overshoot caused by a range error of -3% would hardly cause a dose increase to the cord.

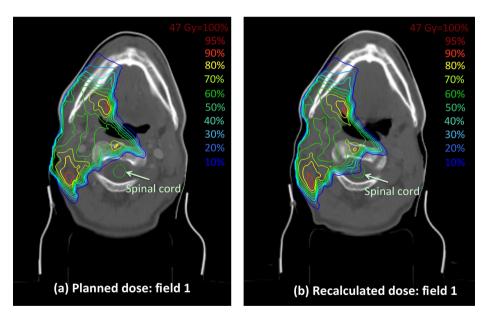


Figure 7.E.1: Dose deviation caused by combined errors for patient 8. Planned (a) and recalculated (b) dose distribution from one of the beam directions. The recalculated dose is displayed on the repeat CT scan.

7 F Comparison of plans with 3, 5 and 7 beam directions

Figures 7.F.1 and 7.F.2 summarize the dose impacts for the scenario of combined errors for the 3, 5 and 7 beam plans. To analyze the differences in robustness between the 3, 5 and 7 beam direction plans, we evaluated for the target region the standard deviation of delivered minus planned $V_{95\%}$, $V_{107\%}$, and $D_{98\%}$, while for the OARs we evaluated the standard deviation of delivered minus planned max dose (spinal cord, brainstem) and the standard deviation of delivered minus planned mean dose (other OARs). A significantly smaller value of the standard deviation would indicate an increased robustness. Since we have one patient group and 3 independent measurements, we used the Anova repeated measurements (AnovaRM) test. For the target region, the differences were in all cases insignificant (p > 0.05). Looking at the $V_{107\%}$ in figure 7.F.1, it may seem that the robustness of the 5 beam plans of the $V_{107\%}$ is worse than that of the 3 and 7 beam direction plans. However, the p-value resulting from the AnovaRM test was not significant: p = 0.104. For the OARs, in a few cases the AnovaRM test yielded a significant difference (p < 0.05), however, in all cases the differences in standard deviations were below about 1 Gy, which is usually not considered as clinically significant.

We suggest that including more beam directions does not necessarily solve dose deteriorations resulting from pencil beam mismatching or anatomical changes. A different pencil beam placement or a different choice of the beam directions, for example obtained with beam-angle optimization, could alter our observation.

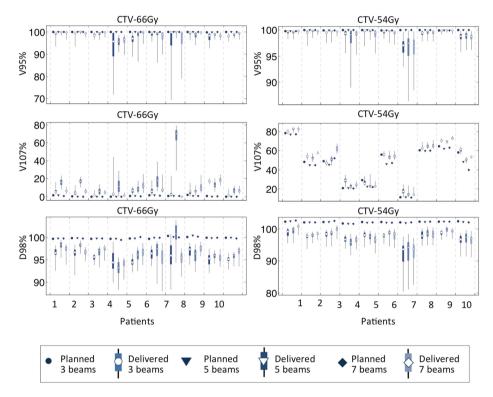


Figure 7.F.1: Comparison of plans with 3, 5 and 7 beam directions. Summary of dose effects with all uncertainties included for the 10 patients: $V_{95\%}$ and $V_{107\%}$ (for CTV-66 Gy and CTV-54 Gy). The solid markers represent the median, boxes spread between the 25th and 75th percentile, and lines extend to the minimum and maximum of the 40 simulations.

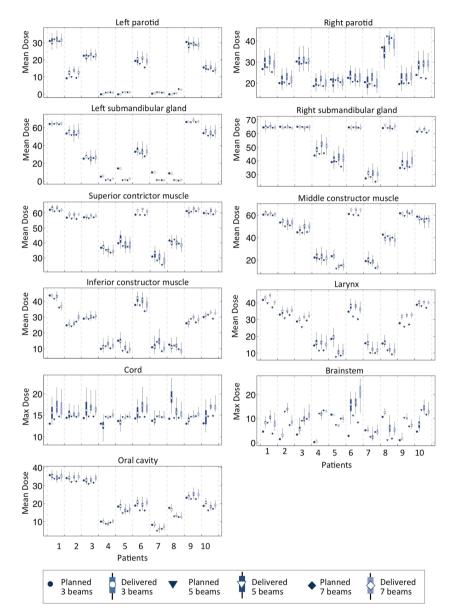


Figure 7.F.2: As in figure 7.F.1, but summary of dose effects for OARs. Maximum dose ($=D_{2\%}$, cord and brainstem) and mean dose (other OARs).

9 A Evaluation of treatment plan robustness

To evaluate the behavior of robustly optimized treatment plans when applying treatment errors other than the errors included in the plan optimization, we performed plan recalculations. For all patients, dose distributions were recalculated for treatment plans generated with a setup robustness ranging from 1 to 7 mm, while applying setup errors ranging from 1 to 7 mm. Similarly, the dose was recalculated for treatment plans generated with a range robustness ranging from 1% to 9% (+1 mm), while applying setup errors ranging from 1% to 9% (+1 mm). In the recalculations, setup errors and range errors were not combined and setup errors were applied only along one of the three axes. We evaluated the coverage ($V_{95\%} > 98\%$) of the low-dose and high-dose CTVs and the $V_{107\%}$ ($\approx 2\%$) and $V_{110\%}$ (\approx 0%) of the high-dose CTV.

The results of the recalculations are depicted in figure 9.A.1. The graphs display the outcomes averaged over all recalculations of all 10 patients. In general, the treatment plans were found to be robust against treatment errors smaller than or equal to the errors included in the plan optimization. The coverage ($V_{95\%}$) of the low-dose and high-dose CTVs was higher than 98% in nearly all recalculations for intermediate errors (i.e. errors within the range of errors included in the robust optimization). Beyond the errors optimized for, the coverage decreased rapidly. For setup errors, intermediate errors always resulted in lower (better) V_{107%} and V_{110%} values compared with the planned values, also for high degrees of robustness. However, the $V_{107\%}$ and $V_{110\%}$ were found to worsen slightly for intermediate range errors, with the worsening depending on the distance to the errors included in the optimization. The mean $V_{107\%}$ could be higher than 2% for treatment plans with a range robustness of 7% and 9% (+1 mm). Similar behavior was observed for the $V_{110\%}$, although the mean $V_{110\%}$ never exceeded 0.1%. In clinical practice, one could consider including additional intermediate error scenarios for very high degrees of range robustness.

These results also show that robustness was not maintained when the actual errors were larger than the errors for which robustness was assured. This highlights that the required robustness should not be underestimated, because target coverage could be compromised.

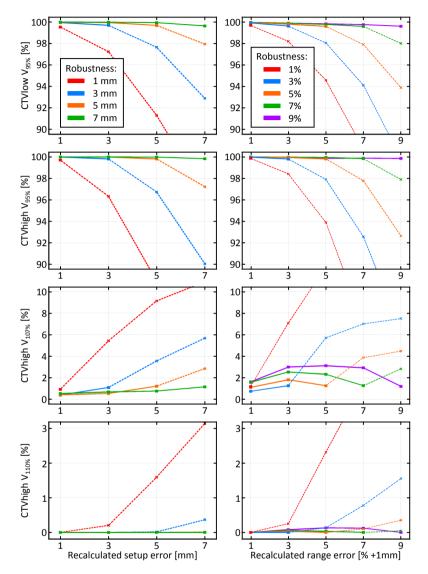


Figure 9.A.1: Results of dose recalculations with varying setup errors (left column) and range errors (right column) for treatment plans with varying setup robustness (left column) and range robustness (right column), respectively. Colors indicate the different degrees of setup or range robustness. Square markers and solid lines represent results for errors smaller or equal to the error included in the optimization. Cross markers and dotted lines indicate treatment errors exceeding the errors optimized for.

References

- 1. Rosenblatt E, Izewska J, Anacak Y, *et al.* Radiotherapy capacity in European countries: an analysis of the Directory of Radiotherapy Centres (DIRAC) database. *Lancet Oncol.* 2013;14:e79–e86.
- 2. Convery DJ, Rosenbloom ME. The generation of intensity-modulated fields for conformal radiotherapy by dynamic collimation. *Phys. Med. Biol.* 1992;37:1359–1374.
- 3. Webb S. Optimizing the planning of intensity-modulated radiotherapy. *Phys. Med. Biol.* 1994;39:2229–2246.
- 4. Intensity Modulated Radiation Therapy Collaborative Working Group. Intensity-modulated radiotherapy: current status and issues of interest. *Int. J. Radiat. Oncol. Biol. Phys.* 2001;51:880–914.
- 5. Van Herk M. Different Styles of Image-Guided Radiotherapy. *Semin. Radiat. Oncol.* 2007;17:258–267.
- 6. Verellen D, Ridder MD, Storme G. A (short) history of image-guided radiotherapy. *Radiother. Oncol.* 2008;86:4–13.
- 7. Mackie TR, Holmes T, Swerdloff S, *et al.* Tomotherapy: A new concept for the delivery of dynamic conformal radiotherapy. *Med. Phys.* 1993;20:1709–1719.
- 8. Yu CX. Intensity-modulated arc therapy with dynamic multileaf collimation: an alternative to tomotherapy. *Phys. Med. Biol.* 1995;40:1435–1449.
- 9. Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med. Phys.* 2008;35:310–317.
- 10. Adler JR, Murphy MJ, Chang DS, et al. Image-guided robotic radiosurgery. *Neuro-surgery*. 1999;44:1299–1306.
- 11. Kilby W, Dooley JR, Kuduvalli G, et al. The CyberKnife® Robotic Radiosurgery System in 2010. *Technol. Cancer Res. Treat.* 9:433–452.
- 12. Lomax A. Intensity modulation methods for proton radiotherapy. *Phys. Med. Biol.* 1999;44:185–205.
- 13. Lomax AJ, Boehringer T, Coray A, *et al.* Intensity modulated proton therapy: A clinical example. *Med. Phys.* 2001;28:317–324.
- 14. Van der Voort van Zyp NC, Prévost J-B, Hoogeman MS, *et al.* Stereotactic radiotherapy with real-time tumor tracking for non-small cell lung cancer: clinical outcome. *Radiother. Oncol.* 2009;91:296–300.

- 15. Aluwini S, Rooij P van, Hoogeman M, *et al.* Stereotactic body radiotherapy with a focal boost to the MRI-visible tumor as monotherapy for low- and intermediate-risk prostate cancer: early results. *Radiat. Oncol.* 2013;8:1–7.
- 16. Ares C, Hug EB, Lomax AJ, et al. Effectiveness and Safety of Spot Scanning Proton Radiation Therapy for Chordomas and Chondrosarcomas of the Skull Base: First Long-Term Report. Int. J. Radiat. Oncol. Biol. Phys. 2009;75:1111–1118.
- 17. Rutz HP, Weber DC, Goitein G, et al. Postoperative Spot-Scanning Proton Radiation Therapy for Chordoma and Chondrosarcoma in Children and Adolescents: Initial Experience at Paul Scherrer Institute. Int. J. Radiat. Oncol. Biol. Phys. 2008;71:220–225.
- 18. Frank SJ, Cox JD, Gillin M, *et al.* Multifield Optimization Intensity Modulated Proton Therapy for Head and Neck Tumors: A Translation to Practice. *Int. J. Radiat. Oncol. Biol. Phys.* 2014;89:846–853.
- 19. Pöll JJ, Hoogeman MS, Prévost J-B, *et al.* Reducing monitor units for robotic radio-surgery by optimized use of multiple collimators. *Med Phys.* 2008;35:2294–2299.
- 20. Kupelian P, Willoughby T, Mahadevan A, *et al.* Multi-institutional clinical experience with the Calypso System in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2007;67:1088–1098.
- 21. Langen KM, Willoughby TR, Meeks SL, *et al.* Observations on Real-Time Prostate Gland Motion Using Electromagnetic Tracking. *Int. J. Radiat. Oncol. Biol. Phys.* 2008;71:1084–1090.
- 22. Pedroni E, Bacher R, Blattmann H, *et al.* The 200-MeV proton therapy project at the Paul Scherrer Institute: Conceptual design and practical realization. *Med. Phys.* 1995;22:37–53.
- 23. Kooy HM, Clasie BM, Lu H-M, *et al.* A Case Study in Proton Pencil-Beam Scanning Delivery. *Int. J. Radiat. Oncol. Biol. Phys.* 2010;76:624–630.
- 24. Peeters A, Grutters JPC, Pijls-Johannesma M, *et al.* How costly is particle therapy? Cost analysis of external beam radiotherapy with carbon-ions, protons and photons. *Radiother. Oncol.* 2010;95:45–53.
- 25. Echner GG, Kilby W, Lee M, *et al.* The design, physical properties and clinical utility of an iris collimator for robotic radiosurgery. *Phys. Med. Biol.* 2009;54:5359–5380.
- 26. Liu R, Buatti JM, Howes TL, *et al.* Optimal number of beams for stereotactic body radiotherapy of lung and liver lesions. *Int. J. Radiat. Oncol. Biol. Phys.* 2006;66:906–912.
- 27. De Pooter JA, Romero AM, Wunderink W, *et al.* Automated non-coplanar beam direction optimization improves IMRT in SBRT of liver metastasis. *Radiother Oncol.* 2008;88:376–381.
- 28. Hoogeman M, Prevost J-B, Nuyttens J, et al. Clinical Accuracy of the Respiratory

- Tumor Tracking System of the Cyberknife: Assessment by Analysis of Log Files. Int. J. Radiat. Oncol. Biol. Phys. 2009;74:297-303.
- 29. Timmerman R, Galvin J, Michalski J, et al. Accreditation and quality assurance for Radiation Therapy Oncology Group: Multicenter clinical trials using Stereotactic Body Radiation Therapy in lung cancer. *Acta Oncol.* 2006;45:779–786.
- 30. Woudstra E, Heijmen BJM, Storchi PRM. A comparison of an algorithm for automated sequential beam orientation selection (Cycle) with simulated annealing. Phys Med Biol. 2008;53:2003-2018.
- 31. De Pooter JA, Méndez Romero A, Jansen WPA, et al. Computer optimization of noncoplanar beam setups improves stereotactic treatment of liver tumors. Int. J. Radiat. Oncol. Biol. Phys. 2006;66:913-922.
- 32. Breedveld S, Storchi PRM, Heijmen BJM. The equivalence of multi-criteria methods for radiotherapy plan optimization. *Phys. Med. Biol.* 2009;54:7199–7209.
- 33. Schweikard A, Schlaefer A, Adler JR. Resampling: an optimization method for inverse planning in robotic radiosurgery. Med Phys. 2006;33:4005-4011.
- 34. Borst GR, Ishikawa M, Nijkamp J, et al. Radiation pneumonitis after hypofractionated radiotherapy: evaluation of the LQ(L) model and different dose parameters. *Int.* J. Radiat. Oncol. Biol. Phys. 2010;77:1596-1603.
- 35. Meedt G, Alber M, Nüsslin F. Non-coplanar beam direction optimization for intensitymodulated radiotherapy. Phys. Med. Biol. 2003;48:2999-3019.
- 36. Alber M, Meedt G, Nüsslin F, et al. On the degeneracy of the IMRT optimization problem. Med. Phys. 2002;29:2584-2589.
- 37. Heijmen B, van de Water S, Breedveld S, et al. A Variable Circular Collimator: A Time-efficient Alternative for a Multi-leaf Collimator in Robotic Radiosurgery? Int. J. Radiat. Oncol. Biol. Phys. 2009;75:S674-S675.
- 38. Woudstra E, Storchi PR. Constrained treatment planning using sequential beam selection. *Phys Med Biol.* 2000;45:2133-2149.
- 39. Woudstra E, Heijmen BJM, Storchi PRM. Automated selection of beam orientations and segmented intensity-modulated radiotherapy (IMRT) for treatment of oesophagus tumors. Radiother Oncol. 2005;77:254-61.
- 40. Fuller DB, Naitoh J, Lee C, et al. Virtual HDRSM CyberKnife Treatment for Localized Prostatic Carcinoma: Dosimetry Comparison With HDR Brachytherapy and Preliminary Clinical Observations. Int. J. Radiat. Oncol. Biol. Phys. 2008;70:1588-1597.
- 41. Breedveld S, Storchi PRM, Keijzer M, et al. A novel approach to multi-criteria inverse planning for IMRT. Phys. Med. Biol. 2007;52:6339-6353.
- 42. Murphy MJ, Lin P-S, Ozhasoglu C. Intra-fraction dose delivery timing during stereotactic radiotherapy can influence the radiobiological effect. Med Phys. 2007;34:481-484.

- 43. Paganetti H. Changes in tumor cell response due to prolonged dose delivery times in fractionated radiation therapy. *Int J Radiat Oncol Biol Phys.* 2005;63:892–900.
- 44. Dooley JR, Sims C, Orr GL, et al. SU-GG-T-544: Delivery Time Minimization for Robotic Radiosurgery and IMRT Treatments. Med. Phys. 2010;37:3312.
- 45. Schlaefer A, Schweikard A. Stepwise multi-criteria optimization for robotic radiosurgery. Med. Phys. 2008;35:2094-2103.
- 46. Llacer J, Li S, Agazaryan N, et al. Non-coplanar automatic beam orientation selection in cranial IMRT: a practical methodology. Phys. Med. Biol. 2009;54:1337-1368.
- 47. Das S, Cullip T, Tracton G, et al. Beam orientation selection for intensity-modulated radiation therapy based on target equivalent uniform dose maximization. Int. J. Radiat. Oncol. Biol. Phys. 2003;55:215-224.
- 48. Rowbottom CG, Nutting CM, Webb S. Beam-orientation optimization of intensitymodulated radiotherapy: clinical application to parotid gland tumours. Radiother. Oncol. 2001;59:169-177.
- 49. Breedveld S, Storchi P, Heijmen B. Multi-criteria beam angle IMRT optimization. Radiother. Oncol. 2009;92:S204.
- 50. Van Herk M, Bruce A, Guus Kroes AP, et al. Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. Int. J. Radiat. Oncol. Biol. Phys. 1995;33:1311-1320.
- 51. Hoogeman MS, van Herk M, de Bois J, et al. Strategies to reduce the systematic error due to tumor and rectum motion in radiotherapy of prostate cancer. Radiother. Oncol. 2005;74:177-185.
- 52. De Boer HC., van Os MJ., Jansen PP, et al. Application of the No Action Level (NAL) protocol to correct for prostate motion based on electronic portal imaging of implanted markers. Int. J. Radiat. Oncol. Biol. Phys. 2005;61:969–983.
- 53. Aubry J-F, Beaulieu L, Girouard L-M, et al. Measurements of intrafraction motion and interfraction and intrafraction rotation of prostate by three-dimensional analysis of daily portal imaging with radiopaque markers. Int. J. Radiat. Oncol. Biol. Phys. 2004;60:30-39.
- 54. Xie Y, Djajaputra D, King CR, et al. Intrafractional Motion of the Prostate During Hypofractionated Radiotherapy. Int. J. Radiat. Oncol. Biol. Phys. 2008;72:236–246.
- 55. Mutanga TF, de Boer HCJ, Rajan V, et al. Day-to-Day Reproducibility of Prostate Intrafraction Motion Assessed by Multiple kV and MV Imaging of Implanted Markers During Treatment. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;83:400–407.
- 56. Deutschmann H, Kametriser G, Steininger P, et al. First Clinical Release of an Online, Adaptive, Aperture-Based Image-Guided Radiotherapy Strategy in Intensity-Modulated Radiotherapy to Correct for Inter- and Intrafractional Rotations of the Prostate. Int. J. Radiat. Oncol. Biol. Phys. 2012;83:1624–1632.

- 57. Lin Y, Liu T, Yang W, et al. The Non-Gaussian Nature of Prostate Motion Based on Real-Time Intrafraction Tracking. Int. J. Radiat. Oncol. Biol. Phys. 2013;87:363-369.
- 58. Langen KM, Lu W, Willoughby TR, et al. Dosimetric Effect of Prostate Motion During Helical Tomotherapy. Int. J. Radiat. Oncol. Biol. Phys. 2009;74:1134-1142.
- 59. Langen KM, Chauhan B, Siebers JV, et al. The Dosimetric Effect of Intrafraction Prostate Motion on Step-and-Shoot Intensity-Modulated Radiation Therapy Plans: Magnitude, Correlation With Motion Parameters, and Comparison With Helical Tomotherapy Plans. Int. J. Radiat. Oncol. Biol. Phys. 2012;84:1220–1225.
- 60. Mutanga TF, Boer HCJ de, Rajan V, et al. Software-controlled, highly automated intrafraction prostate motion correction with intrafraction stereographic targeting: System description and clinical results. *Med. Phys.* 2012;39:1314–1321.
- 61. Aluwini S, van Rooij P, Hoogeman M, et al. CyberKnife Stereotactic Radiotherapy as Monotherapy for Low- to Intermediate-Stage Prostate Cancer: Early Experience, Feasibility, and Tolerance. J. Endourol. 2010;24:865-869.
- 62. Hoogeman MS, Nuyttens JJ, Levendag PC, et al. Time Dependence of Intrafraction Patient Motion Assessed by Repeat Stereoscopic Imaging. Int. J. Radiat. Oncol. Biol. Phys. 2008;70:609-618.
- 63. Murphy MJ. Intrafraction Geometric Uncertainties in Frameless Image-Guided Radiosurgery. Int. I. Radiat. Oncol. Biol. Phys. 2009;73:1364-1368.
- 64. Gevaert T, Boussaer M, Engels B, et al. Evaluation of the clinical usefulness for using verification images during frameless radiosurgery. Radiother. Oncol. 2013;108:114-117.
- 65. Kang KM, Chai GY, Jeong BK, et al. Estimation of optimal margin for intrafraction movements during frameless brain radiosurgery. Med. Phys. 2013;40:051716.
- 66. Breedveld S, Storchi PRM, Voet PWJ, et al. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. Med. Phys. 2012;39:951-963.
- 67. Van de Water S, Hoogeman MS, Breedveld S, et al. Shortening treatment time in robotic radiosurgery using a novel node reduction technique. Med. Phys. 2011;38: 1397-1405.
- 68. Van de Water S, Hoogeman MS, Breedveld S, et al. Variable circular collimator in robotic radiosurgery: a time-efficient alternative to a mini-multileaf collimator? Int. J. Radiat. Oncol. Biol. Phys. 2011;81:863-870.
- 69. Rossi L, Breedveld S, Heijmen BJM, et al. On the beam direction search space in computerized non-coplanar beam angle optimization for IMRT—prostate SBRT. Phys. Med. Biol. 2012;57:5441-5458.
- 70. Voet PWJ, Dirkx MLP, Breedveld S, et al. Toward Fully Automated Multicriterial Plan Generation: A Prospective Clinical Study. Int. J. Radiat. Oncol. Biol. Phys. 2013;85:866-872.

- 71. Demanes DJ, Martinez AA, Ghilezan M, et al. High-Dose-Rate Monotherapy: Safe and Effective Brachytherapy for Patients With Localized Prostate Cancer. Int. J. Radiat. Oncol. Biol. Phys. 2011;81:1286-1292.
- 72. Kang JH, Wilkens JJ, Oelfke U. Non-uniform depth scanning for proton therapy systems employing active energy variation. *Phys. Med. Biol.* 2008;53:N149-N155.
- 73. Hillbrand M, Georg D. Assessing a set of optimal user interface parameters for intensity-modulated proton therapy planning. J. Appl. Clin. Med. Phys. 2010;11:93– 104.
- 74. Widesott L, Pierelli A, Fiorino C, et al. Intensity-Modulated Proton Therapy Versus Helical Tomotherapy in Nasopharynx Cancer: Planning Comparison and NTCP Evaluation. Int. J. Radiat. Oncol. Biol. Phys. 2008;72:589-596.
- 75. Van de Water TA, Lomax AJ, Bijl HP, et al. Potential Benefits of Scanned Intensity-Modulated Proton Therapy Versus Advanced Photon Therapy With Regard to Sparing of the Salivary Glands in Oropharyngeal Cancer. Int. J. Radiat. Oncol. Biol. Phys. 2011;79:1216-1224.
- 76. Van de Water TA, Lomax AJ, Bijl HP, et al. Using a Reduced Spot Size for Intensity-Modulated Proton Therapy Potentially Improves Salivary Gland-Sparing in Oropharyngeal Cancer. Int. J. Radiat. Oncol. Biol. Phys. 2012;82:e313-e319.
- 77. Widesott L, Lomax AJ, Schwarz M. Is there a single spot size and grid for intensity modulated proton therapy? Simulation of head and neck, prostate and mesothelioma cases. Med. Phys. 2012;39:1298-1308.
- 78. Schippers JM, Lomax AJ. Emerging technologies in proton therapy. Acta Oncol. 2011;50:838-850.
- 79. Pedroni E, Bearpark R, Böhringer T, et al. The PSI Gantry 2: a second generation proton scanning gantry. Z. Für Med. Phys. 2004;14:25–34.
- 80. Zenklusen SM, Pedroni E, Meer D. A study on repainting strategies for treating moderately moving targets with proton pencil beam scanning at the new Gantry 2 at PSI. Phys. Med. Biol. 2010;55:5103-5121.
- 81. Pflugfelder D, Wilkens JJ, Oelfke U. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. Phys. Med. Biol. 2008;53:1689-1700.
- 82. Unkelbach J, Bortfeld T, Martin BC, et al. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. Med. Phys. 2009;36:149-163.
- 83. Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. Phys. Med. Biol. 2008;53:1027-1042.
- 84. Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncer-

- tainties 2: the potential effects of inter-fraction and inter-field motions. Phys. Med. Biol. 2008;53:1043-1056.
- 85. Goitein M, Jermann M. The Relative Costs of Proton and X-ray Radiation Therapy. Clin. Oncol. 2003;15:S37-S50.
- 86. Van de Water S, Kraan AC, Breedveld S, et al. Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams. Phys. Med. Biol. 2013;58:6969-6983.
- 87. Voet PWJ, Dirkx MLP, Breedveld S, et al. Fully automated volumetric modulated arc therapy plan generation for prostate cancer patients. Int. J. Radiat. Oncol. Biol. Phys. 2014;88:1175-1179.
- 88. Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. Med. Phys. 2011;38:1672–1684.
- 89. Chen W, Unkelbach J, Trofimov A, et al. Including robustness in multi-criteria optimization for intensity-modulated proton therapy. Phys. Med. Biol. 2012;57:591-608.
- 90. Thörnqvist S, Muren LP, Bentzen L, et al. Degradation of target coverage due to interfraction motion during intensity-modulated proton therapy of prostate and elective targets. Acta Oncol. 2013;52:521-527.
- 91. Safai S, Bula C, Meer D, et al. Improving the precision and performance of proton pencil beam scanning. Transl. Cancer Res. 2012;1:196-206.
- 92. Flanz J, Bortfeld T. Evolution of Technology to Optimize the Delivery of Proton Therapy: The Third Generation. *Semin. Radiat. Oncol.* 2013;23:142–148.
- 93. Siebenthal M von, Székely G, Lomax AJ, et al. Systematic errors in respiratory gating due to intrafraction deformations of the liver. Med. Phys. 2007;34:3620-3629.
- 94. Cao W, Lim G, Liao L, et al. Proton energy optimization and reduction for intensitymodulated proton therapy. Phys. Med. Biol. 2014;59:6341-6354.
- 95. Simone II CB, Ly D, Dan TD, et al. Comparison of intensity-modulated radiotherapy, adaptive radiotherapy, proton radiotherapy, and adaptive proton radiotherapy for treatment of locally advanced head and neck cancer. Radiother. Oncol. 2011;101:376-
- 96. Barker Jr JL, Garden AS, Ang KK, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int. J. Radiat. Oncol. Biol. Phys. 2004;59:960-970.
- 97. Van Kranen S, van Beek S, Rasch C, et al. Setup Uncertainties of Anatomical Sub-Regions in Head-and-Neck Cancer Patients After Offline CBCT Guidance. Int. J. Radiat. Oncol. Biol. Phys. 2009;73:1566-1573.
- 98. Albertini F, Bolsi A, Lomax AJ, et al. Sensitivity of intensity modulated proton therapy plans to changes in patient weight. Radiother. Oncol. 2008;86:187-194.
- 99. Liu W, Frank SJ, Li X, et al. Effectiveness of robust optimization in intensity-

- modulated proton therapy planning for head and neck cancers. *Med. Phys.* 2013;40: 051711.
- 100. Liu W, Zhang X, Li Y, *et al.* Robust optimization of intensity modulated proton therapy. *Med. Phys.* 2012;39:1079–1091.
- 101. Liu W, Frank SJ, Li X, *et al.* PTV-based IMPT optimization incorporating planning risk volumes vs robust optimization. *Med. Phys.* 2013;40:021709.
- 102. Inaniwa T, Kanematsu N, Furukawa T, *et al.* A robust algorithm of intensity modulated proton therapy for critical tissue sparing and target coverage. *Phys. Med. Biol.* 2011;56:4749–4770.
- 103. Fredriksson A. A characterization of robust radiation therapy treatment planning methods—from expected value to worst case optimization. *Med. Phys.* 2012;39:5169–5181.
- 104. Castadot P, Lee JA, Geets X, *et al.* Adaptive Radiotherapy of Head and Neck Cancer. *Semin. Radiat. Oncol.* 2010;20:84–93.
- 105. Vásquez Osorio EM, Hoogeman MS, Al-Mamgani A, et al. Local Anatomic Changes in Parotid and Submandibular Glands During Radiotherapy for Oropharynx Cancer and Correlation With Dose, Studied in Detail With Nonrigid Registration. Int. J. Radiat. Oncol. Biol. Phys. 2008;70:875–882.
- 106. Teguh DN, Levendag PC, Voet PWJ, *et al.* Clinical Validation of Atlas-Based Auto-Segmentation of Multiple Target Volumes and Normal Tissue (Swallowing/Mastication) Structures in the Head and Neck. *Int. J. Radiat. Oncol. Biol. Phys.* 2011;81:950–957.
- 107. Revision Radiology. *CT metal artifact reduction. Available at: http://www.revisionrads.com/.*
- 108. Van Herk M, Remeijer P, Rasch C, *et al.* The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 2000;47:1121–1135.
- 109. Hong L, Goitein M, Bucciolini M, et al. A pencil beam algorithm for proton dose calculations. *Phys. Med. Biol.* 1996;41:1305–1330.
- 110. Voet PWJ, Breedveld S, Dirkx MLP, *et al.* Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head and neck IMRT and implications for VMAT. *Med. Phys.* 2012;39:4858–4865.
- 111. ICRU Report 83. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT), vol. 10. Oxford, UK: Oxford University Press; 2010.
- 112. Cao W, Lim GJ, Lee A, *et al.* Uncertainty incorporated beam angle optimization for IMPT treatment planning. *Med. Phys.* 2012;39:5248–5256.
- 113. Eickhoff H, Haberer T, Kraft G, *et al.* The GSI Cancer Therapy Project. *Strahlenther. Onkol.* 1999;175:21–24.

- 114. Schulz-Ertner D, Nikoghosyan A, Thilmann C, *et al.* Results of carbon ion radiotherapy in 152 patients. *Int. J. Radiat. Oncol. Biol. Phys.* 2004;58:631–640.
- 115. Boyer AL, Goitein M, Lomax AJ, *et al.* Radiation in the treatment of cancer. *Phys. Today.* 2002;55:34–36.
- 116. Wilson RR. Radiological Use of Fast Protons. Radiology. 1946;47:487-491.
- 117. Phillips MH, Pedroni E, Blattmann H, *et al.* Effects of respiratory motion on dose uniformity with a charged particle scanning method. *Phys. Med. Biol.* 1992;37:223–234.
- 118. Seiler PG, Blattmann H, Kirsch S, *et al.* A novel tracking technique for the continuous precise measurement of tumour positions in conformal radiotherapy. *Phys. Med. Biol.* 2000;45:N103–N110.
- 119. Bert C, Grözinger SO, Rietzel E. Quantification of interplay effects of scanned particle beams and moving targets. *Phys. Med. Biol.* 2008;53:2253–2265.
- 120. Bortfeld T, Jokivarsi K, Goitein M, *et al.* Effects of intra-fraction motion on IMRT dose delivery: statistical analysis and simulation. *Phys. Med. Biol.* 2002;47:2203–2220.
- 121. Keall P, Vedam S, George R, *et al.* The clinical implementation of respiratory-gated intensity-modulated radiotherapy. *Med. Dosim.* 2006;31:152–162.
- 122. Papiez L, Rangaraj D, Keall P. Real-time DMLC IMRT delivery for mobile and deforming targets. *Med. Phys.* 2005;32:3037–3048.
- 123. Keall PJ, Cattell H, Pokhrel D, *et al.* Geometric accuracy of a real-time target tracking system with dynamic multileaf collimator tracking system. *Int. J. Radiat. Oncol. Biol. Phys.* 2006;65:1579–1584.
- 124. Grözinger SO, Rietzel E, Li Q, *et al.* Simulations to design an online motion compensation system for scanned particle beams. *Phys. Med. Biol.* 2006;51:3517–3531.
- 125. Grözinger SO, Bert C, Haberer T, *et al.* Motion compensation with a scanned ion beam: a technical feasibility study. *Radiat. Oncol.* 2008;3:34.
- 126. Bert C, Rietzel E. 4D treatment planning for scanned ion beams. *Radiat. Oncol.* 2007;2:24.
- 127. Bert C, Saito N, Schmidt A, *et al.* Target motion tracking with a scanned particle beam. *Med. Phys.* 2007;34:4768–4771.
- 128. Siebenthal M von. Analysis and modelling of respiratory livermotion using 4DMRI. ETH Dissertation No. 17613. 2008.
- 129. Lomax AJ, Böhringer T, Bolsi A, *et al.* Treatment planning and verification of proton therapy using spot scanning: Initial experiences. *Med. Phys.* 2004;31:3150–3157.
- 130. Lujan AE, Larsen EW, Balter JM, *et al.* A method for incorporating organ motion due to breathing into 3D dose calculations. *Med. Phys.* 1999;26:715–720.
- 131. George R, Vedam SS, Chung TD, *et al.* The application of the sinusoidal model to lung cancer patient respiratory motion. *Med. Phys.* 2005;32:2850–2861.

- 132. Siebenthal M von, Székely G, Gamper U, *et al.* 4D MR imaging of respiratory organ motion and its variability. *Phys. Med. Biol.* 2007;52:1547–1564.
- 133. Schaffner B, Pedroni E, Lomax A. Dose calculation models for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity effects in the analytical dose calculation. *Phys. Med. Biol.* 1999;44:27–41.
- 134. Vedam SS, Keall PJ, Docef A, *et al.* Predicting respiratory motion for four-dimensional radiotherapy. *Med. Phys.* 2004;31:2274–2283.
- 135. Kraan AC, van de Water S, Teguh DN, *et al.* Dose Uncertainties in IMPT for Oropharyngeal Cancer in the Presence of Anatomical, Range, and Setup Errors. *Int. J. Radiat. Oncol. Biol. Phys.* 2013;87:888–896.
- 136. Fredriksson A, Bokrantz R. A critical evaluation of worst case optimization methods for robust intensity-modulated proton therapy planning. *Med. Phys.* 2014;41:081701.
- 137. Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys. Med. Biol.* 2012;57:R99–R117.
- 138. Casiraghi M, Albertini F, Lomax AJ. Advantages and limitations of the "worst case scenario" approach in IMPT treatment planning. *Phys. Med. Biol.* 2013;58:1323–1339.
- 139. Jellema AP, Doornaert P, Slotman BJ, et al. Does radiation dose to the salivary glands and oral cavity predict patient-rated xerostomia and sticky saliva in head and neck cancer patients treated with curative radiotherapy? *Radiother. Oncol.* 2005;77:164–171.
- 140. Dijkema T, Raaijmakers CPJ, Ten Haken RK, *et al.* Parotid Gland Function After Radiotherapy: The Combined Michigan and Utrecht Experience. *Int. J. Radiat. Oncol. Biol. Phys.* 2010;78:449–453.
- 141. Levendag PC, Teguh DN, Voet P, *et al.* Dysphagia disorders in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscle: A dose-effect relationship. *Radiother. Oncol.* 2007;85:64–73.
- 142. Christianen MEMC, Schilstra C, Beetz I, *et al.* Predictive modelling for swallowing dysfunction after primary (chemo)radiation: Results of a prospective observational study. *Radiother. Oncol.* 2012;105:107–114.
- 143. Rancati T, Schwarz M, Allen AM, *et al.* Radiation Dose–Volume Effects in the Larynx and Pharynx. *Int. J. Radiat. Oncol. Biol. Phys.* 2010;76:S64–S69.
- 144. Petit S, Seco J, Kooy H. Increasing maximum tumor dose to manage range uncertainties in IMPT treatment planning. *Phys. Med. Biol.* 2013;58:7329–7341.
- 145. Yang M, Virshup G, Clayton J, *et al.* Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Phys. Med. Biol.* 2010;55:1343–1362.
- 146. Testa M, Verburg JM, Rose M, et al. Proton radiography and proton computed to-

- mography based on time-resolved dose measurements. *Phys. Med. Biol.* 2013;58: 8215–8233.
- 147. Testa M, Bajard M, Chevallier M, *et al.* Real-time monitoring of the Bragg-peak position in ion therapy by means of single photon detection. *Radiat. Environ. Biophys.* 2010;49:337–343.
- 148. Biegun AK, Seravalli E, Lopes PC, *et al.* Time-of-flight neutron rejection to improve prompt gamma imaging for proton range verification: a simulation study. *Phys. Med. Biol.* 2012;57:6429–6444.
- 149. Min CH, Lee HR, Kim CH, *et al.* Development of array-type prompt gamma measurement system for in vivo range verification in proton therapy. *Med. Phys.* 2012;39:2100–2107.
- 150. Smeets J, Roellinghoff F, Prieels D, *et al.* Prompt gamma imaging with a slit camera for real-time range control in proton therapy. *Phys. Med. Biol.* 2012;57:3371–3405.
- 151. Wake GMGH, Boland N, Jennings LS. Mixed integer programming approaches to exact minimization of total treatment time in cancer radiotherapy using multileaf collimators. *Comput. Oper. Res.* 2009;36:795–810.
- 152. Dong P, Lee P, Ruan D, et al. 4π Non-Coplanar Liver SBRT: A Novel Delivery Technique. *Int. J. Radiat. Oncol. Biol. Phys.* 2013;85:1360–1366.
- 153. Schätti A, Meer D, Lomax AJ. First experimental results of motion mitigation by continuous line scanning of protons. *Phys. Med. Biol.* 2014;59:5707–5723.
- 154. Seco J, Gu G, Marcelos T, *et al.* Proton Arc Reduces Range Uncertainty Effects and Improves Conformality Compared With Photon Volumetric Modulated Arc Therapy in Stereotactic Body Radiation Therapy for Non-Small Cell Lung Cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2013;87:188–194.
- 155. Graeff C, Lüchtenborg R, Eley JG, *et al.* A 4D-optimization concept for scanned ion beam therapy. *Radiother. Oncol.* 2013;109:419–424.
- 156. Graeff C. Motion mitigation in scanned ion beam therapy through 4D-optimization. *Phys. Med.* 2014;30:570–577.
- 157. Eley JG, Newhauser WD, Lüchtenborg R, *et al.* 4D optimization of scanned ion beam tracking therapy for moving tumors. *Phys. Med. Biol.* 2014;59:3431–3452.
- 158. Lagendijk JJW, Raaymakers BW, Raaijmakers AJE, et al. MRI/linac integration. Radiother. Oncol. 2008;86:25–29.
- 159. Fontanarosa D, Meer S van der, Bamber J, *et al.* Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management. *Phys. Med. Biol.* 2015;60:R77–R114.
- 160. Knopf A-C, Hong TS, Lomax A. Scanned proton radiotherapy for mobile targets—the effectiveness of re-scanning in the context of different treatment planning approaches and for different motion characteristics. *Phys. Med. Biol.* 2011;56:7257–7271.

- 161. Zhang Y, Knopf A, Tanner C, *et al.* Online image guided tumour tracking with scanned proton beams: a comprehensive simulation study. *Phys. Med. Biol.* 2014;59: 7793–7817.
- 162. Grassberger C, Dowdell S, Lomax A, *et al.* Motion Interplay as a Function of Patient Parameters and Spot Size in Spot Scanning Proton Therapy for Lung Cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2013;86:380–386.
- 163. Park PC, Zhu XR, Lee AK, et al. A Beam-Specific Planning Target Volume (PTV) Design for Proton Therapy to Account for Setup and Range Uncertainties. *Int. J. Radiat. Oncol. Biol. Phys.* 2012;82:e329–e336.
- 164. Budiarto E, Keijzer M, Storchi PR, *et al.* A population-based model to describe geometrical uncertainties in radiotherapy: applied to prostate cases. *Phys. Med. Biol.* 2011;56:1045–1061.
- 165. Thörnqvist S, Hysing LB, Zolnay AG, *et al.* Adaptive radiotherapy in locally advanced prostate cancer using a statistical deformable motion model. *Acta Oncol.* 2013;52:1423–1429.
- 166. Goitein M, Cox JD. Should Randomized Clinical Trials Be Required for Proton Radiotherapy? *J. Clin. Oncol.* 2008;26:175–176.
- 167. Suit H, Kooy H, Trofimov A, *et al.* Should positive phase III clinical trial data be required before proton beam therapy is more widely adopted? No. *Radiother. Oncol.* 2008;86:148–153.
- 168. Langendijk JA, Lambin P, De Ruysscher D, *et al.* Selection of patients for radiotherapy with protons aiming at reduction of side effects: The model-based approach. *Radiother. Oncol.* 2013;107:267–273.
- 169. Lambin P, Roelofs E, Reymen B, *et al.* "Rapid Learning health care in oncology" An approach towards decision support systems enabling customised radiotherapy." *Radiother. Oncol.* 2013;109:159–164.

List of publications

- Van de Water S, Kooy HM, Heijmen BJM, Hoogeman MS. Shortening Delivery Times of Intensity Modulated Proton Therapy by Reducing Proton Energy Layers During Treatment Plan Optimization. Int. J. Radiat. Oncol. Biol. Phys. 2015. Accepted for publication.
- Van de Water S, Valli L, Aluwini S, Lanconelli N, Heijmen B, Hoogeman M. Intrafraction Prostate Translations and Rotations During Hypofractionated Robotic Radiation Surgery: Dosimetric Impact of Correction Strategies and Margins. Int. J. Radiat. Oncol. Biol. Phys. 2014. 88, 1154–1160.
- 3. Kraan AC, Van de Water S, Teguh DN, Al-Mamgani A, Madden T, Kooy HM, Heijmen BJM, Hoogeman MS. Dose Uncertainties in IMPT for Oropharyngeal Cancer in the Presence of Anatomical, Range, and Setup Errors. Int. J. Radiat. Oncol. Biol. Phys. 2013. 87, 888–896.
- 4. Van de Water S, Kraan AC, Breedveld S, Schillemans W, Teguh DN, Kooy HM, Madden TM, Heijmen BJM, Hoogeman MS. Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams. Phys. Med. Biol. 2013. 58, 6969–6983.
- 5. Van de Water S, Hoogeman MS, Breedveld S, Nuyttens JJME, Schaart DR, Heijmen, BJM. Variable circular collimator in robotic radiosurgery: a time-efficient alternative to a mini-multileaf collimator? Int. J. Radiat. Oncol. Biol. Phys. 2011. 81, 863–870.
- 6. Van de Water S, Hoogeman MS, Breedveld S, Heijmen BJM. Shortening treatment time in robotic radiosurgery using a novel node reduction technique. Med. Phys. 2011. 38, 1397–1405.
- Van der Voort van Zyp NC, Hoogeman MS, van de Water S, Levendag PC, van der Holt B, Heijmen BJM, Nuyttens JJ. Stability of markers used for real-time tumor tracking after percutaneous intrapulmonary placement. Int. J. Radiat. Oncol. Biol. Phys. 2011. 81, e75–e81.

- 8. Van der Voort van Zyp NC, Hoogeman MS, van de Water S, Levendag PC, van der Holt B, Heijmen BJM, Nuyttens JJ. Clinical introduction of Monte Carlo treatment planning: a different prescription dose for non-small cell lung cancer according to tumor location and size. Radiother. Oncol. 2010. 96, 55–60.
- 9. Van de Water S, Kreuger R, Zenklusen S, Hug E, Lomax AJ. Tumour tracking with scanned proton beams: assessing the accuracy and practicalities. Phys. Med. Biol. 2009. 54, 6549–6563.

Summary

High-precision robotic radiotherapy and intensity-modulated proton therapy (IMPT) are two relatively new radiotherapy techniques that particularly aim at a highly localized delivery of a curative dose to the tumor, while achieving excellent sparing of the surrounding healthy tissues and critical organs. However, the use of hundreds (in high-precision robotic radiotherapy) or thousands (in IMPT) of small beams to achieve a highly localized dose delivery is also associated with a number of issues regarding the efficiency and accuracy of treatment planning and delivery. These issues currently obstruct a more widespread use of high-precision robotic radiotherapy and IMPT. In this thesis, we have aimed at improving treatment planning and treatment delivery for both treatment modalities in order to overcome these obstacles.

Part I: High-precision robotic radiotherapy using the CyberKnife

The CyberKnife robotic radiosurgery system consists of a compact 6-MV linear accelerator mounted on a robotic manipulator. The robot can be positioned at approximately 120 so-called 'node positions'. The node positions are distributed semi-spherically around the patient, allowing the patient to be irradiated from many non-coplanar directions. From each node position, multiple beams can be delivered by adjusting the orientation of the linear accelerator. The integrated stereoscopic imaging system allows for 'tumor tracking' to be performed. Tumor displacements can be compensated during a treatment fraction by acquiring images with a certain time interval and applying subsequent robot corrections. Although very promising clinical outcomes have been achieved in hypofractionated treatments (typically 3–4 fractions) for a selection of treatment sites, CyberKnife treatments are also associated with a number of issues. Firstly, delivery times per fraction are relatively long and this hinders a more widespread application of the CyberKnife in other (conventionally fractionated) treatments. Secondly, tumor tracking using the integrated imaging system has a limited temporal and geometric accuracy, whereas prostate tumors can display very rapid excursions of considerable size. However,

the resulting dosimetric accuracy of tumor tracking and the optimal tracking settings have not been established yet for CyberKnife prostate treatments. We have addressed these issues regarding the efficiency and accuracy of CyberKnife treatments in Part I of this thesis.

In chapter 2, we investigated the effect of beam collimation on the efficiency of treatment delivery. CyberKnife systems are traditionally equipped with circular collimators. However, a mini-multileaf collimator (mMLC) does not restrict the fields to a circular shape and is therefore expected to enable more efficient dose delivery. We compared a mMLC with the 'Iris' variable circular aperture in terms of the expected delivery time per fraction and the achievable plan quality. For 10 early-stage non-small-cell lung cancer patients, mMLC plans were generated first using the beam-angle optimization algorithm 'Cycle' while applying beams' eye view shaped fields with uniform fluence. These mMLC plans typically contained 10-16 node positions, selected from the 117 available node positions. Second, intensity-modulated Iris plans were generated using the same node positions selected in the mMLC plans. In this way, time-efficient treatment plans were generated for both collimators. The use of a mMLC was found to result in a considerable reduction of the delivery time per fraction. For a comparable mean lung dose, the average delivery time per fraction was 12.2 min (range: 10.8-13.5 min) for the mMLC plans and 18.4 min (range: 12.9-28.5 min) for the Iris plans. The gain in delivery time was especially large for large tumor volumes. The Iris collimator outperformed the mMLC in terms of achievable plan quality, as it allowed for intensity-modulation to be performed within the field from each node position whereas uniform fields were used in the mMLC plans. In accordance with the findings of this study, a mMLC has been introduced for the CyberKnife very recently.

A reduction of the number of node positions in a CyberKnife treatment plan is likely to improve treatment efficiency, as a large proportion of the delivery time is spent on robot travel between node positions. A method to incorporate the reduction of node position into the treatment planning process is presented in **chapter 3**. The node reduction technique involved repeated inverse optimization, while excluding low-contribution node positions from the treatment plan and sampling new candidate beams from the remaining node positions in each iteration. Treatment plans were generated with and without node reduction for two lung cases of different complexity, one oropharyngeal case and one prostate case. Node positions were reduced by 77% on average at the cost of an increase in the number of beams and total monitor units of 28% and 9%, respectively. The delivery time per fraction was shortened by 25% (range: 18%–40%) on average. Treatment times were even reduced by 54% on average when compared with the clinical planning approach (no repeated optimization). Node reduction did not affect the achievable plan quality, but the use of additional constraints to guarantee low-dose conformality and to avoid unacceptable skin dose is recommended.

In **chapter** 4, we investigated the dosimetric impact of intrafraction prostate motion and we assessed the effectiveness of tumor tracking and the applied safety margin. Using treatment plans of 3 patients and prostate motion tracks of 17 patients, we performed extensive dosimetric simulations while varying (1) the time interval between robot corrections (15, 60, 180, or 360 s), (2) the extent of robot corrections (no corrections, translational corrections only, or translational corrections combined with rotational corrections up to 5°, 10°, or perfect rotational correction) and (3) the applied safety margin (0 mm or 3 mm). Intrafraction prostate motion was found to substantially affect the delivered dose but this could be compensated effectively by applying robot corrections with a time interval of 60 to 180 s. To achieve adequate irradiation of the tumor in 98% of the treatments, translational and rotational corrections up to 10° were required when using a o-mm safety margin, whereas translational and rotational corrections up to 5° were required when using 3-mm margin plans. Applying a o-mm safety margin resulted in considerably better sparing of the rectum and bladder.

Part II: Intensity-modulated proton therapy

In IMPT, patients are irradiated using thousands of small proton pencil beams (or 'spots'). Since protons are positively charged particles, they interact differently with tissue than photons. Protons have a finite range in tissue, depending on the initial energy of the proton beam, and deposit a high peak-dose just before the end-of-range. This socalled 'Bragg peak' is very suitable to deliver a highly localized dose to the tumor. The dose in front of the Bragg peak is relatively low and beyond the Bragg peak no dose is deposited at all, which enables excellent sparing of surrounding healthy tissues. Spots are delivered one-by-one, each time adjusting the lateral spot position, proton energy (to adjust the proton range), and intensity. In this way, the entire tumor volume is treated. Although promising treatment results have been achieved, IMPT is also associated with a number of challenges. Firstly, the use of thousands of proton beams affects the efficiency of planning and delivery. Improved efficiency is likely to result in improved plan quality, patient comfort and treatment accuracy, and maybe most importantly, can help to reduce treatment costs. IMPT treatment costs are currently high compared with photon techniques. Secondly, the highly localized dose deposition results in a high sensitivity of IMPT treatment plans to treatment uncertainties. This can result in severe degradations of the actual delivered dose and might therefore affect the efficacy and safety of IMPT.

In Part II of this thesis, we have addressed these issues concerning the efficiency and sensitivity of IMPT treatments.

In chapter 5, we developed a novel time-efficient treatment planning approach for IMPT, called 'pencil beam resampling'. This planning technique was implemented into our in-house developed multi-criteria treatment planning system 'Erasmus-iCycle'. Resampling consisted of repeatedly performing: (1) random selection of candidate spots from a very fine grid, (2) inverse multi-criteria optimization, and (3) exclusion of low-weighted spots. The newly selected candidate spots were added to the spots in the existing solution, causing the solution to improve with each iteration. For five head-and-neck cancer patients, we compared pencil beam resampling with the traditional planning method in which spots are distributed over a regular grid. For a comparable plan quality, resampling reduced optimization time by a factor of 2.8 and 5.6 on average (7.8 and 17.0 at maximum) compared with the use of anisotropic and isotropic grids, respectively. Doses to organs-at-risk were generally reduced when using resampling, with median dose reductions ranging from -4% to 42%. Resampling was found to be especially effective when using thin proton pencil beams.

For current commercial IMPT equipment, the time required to adjust the energy of the proton beam is the major component of the treatment time. A reduction of the number of energy layers in IMPT treatment plans is therefore likely to result in a shortening of delivery times. A method to reduce the number of energy layers while preserving robust plan quality is presented in chapter 6. The method was implemented into our inhouse developed planning system 'Erasmus-iCycle' and is similar to the node reduction technique presented in chapter 3. It consisted of two components: (1) minimizing the logarithm of the total spot weight per energy layer, and (2) iteratively excluding lowweighted energy layers. We benchmarked the method by comparing a robust 'timeefficient plan' (with energy layer reduction) with a robust 'standard clinical plan' (without energy layer reduction) for five oropharyngeal cases and five prostate cases. Worst-case robust optimization was performed, accounting for setup errors of 3 mm and range errors of 3% +1 mm. The method resulted in the average number of energy layers to be reduced by 45% for the oropharyngeal cases and by 28% for the prostate cases. When assuming 1, 2 or 5 s energy switching time, the average delivery time was shortened by 25%, 32% or 38% for the oropharyngeal cases, and by 16%, 20% or 24% for the prostate cases.

The impact of uncertainties in patient setup, patient anatomy and proton range on the delivered dose in IMPT treatments was investigated in **chapter 7**. For 10 head-and-neck cancer patients, IMPT treatment plans were generated using 3, 5 or 7 beam directions and applying safety margins of 5 mm around the high-dose and low-dose target volumes. The actual delivered dose was simulated by performing dose recalculations (>3700) in the presence of setup, range and anatomical errors. Uncertainties in patient anatomy

were simulated using a repeat CT-scan, acquired halfway the treatment course, and nonrigid registration to perform dose accumulation. The repeat CT-scans were also used to simulate adaptive planning by generating a new treatment plan for the second half of the treatment fractions. For 3-beam plans, we found that adequate tumor irradiation was achieved in only 69% and 88% of the simulations for the high-dose and low-dose targets, respectively, when including all error sources. The dose in organs-at-risk (OARs) showed a considerable spread around their planned values. The error source causing major dose deviations varied between patients. Plan adaptation based on repeat imaging resulted in a considerable improvement of the tumor dose. Adequate tumor doses were achieved in 96% (high-dose target) and 100% (low-dose target) of the simulations when a new treatment plan was generated halfway the treatment course. Increasing the number of beam directions was not found to improve plan robustness.

In chapter 8, we investigated the impact of respiratory motion on IMPT treatment delivery and we assessed the effectiveness of real-time tumor (re)tracking. Using a 4D dose calculation algorithm and simulated respiratory motion, we implemented a realtime tumor tracking algorithm and applied it to a simple target volume in both homogenous and heterogeneous mathematical phantoms. For tracking and retracking (a hybrid solution combining tumor tracking and rescanning), we assessed different degrees of tracking-accuracy: (1) no tracking (uncorrected irradiation), (2) perfect tracking (no time delays and exact knowledge of target position) and (3) imperfect tracking (simulated time delays or position prediction errors). For the homogeneous phantom, we found that the intended target dose was maintained for all motion amplitudes and phases when performing perfect tracking. Time delays and position errors (imperfect tracking) could result in a substantial degradation of the delivered dose, but this degradation was reduced considerably by applying retracking. In the heterogeneous phantom, the planned target dose could not be preserved using tracking, as adjustments of the pencil beam positions could cause the pencil beam shape to change in the presence of density heterogeneities. The use of rescanning with a high rescan-factor was found to give the most favorable dose distributions for the heterogeneous phantom.

In IMPT, robust treatment planning is preferred over the use of traditional safety margins to deal with uncertainties in patient setup and proton range. However, the magnitude of errors accounted for during robust treatment planning is likely to have an effect on the dose received by OARs. In chapter 9, we quantified the clinical impact of the degree of robustness against setup and range errors on OAR doses and normal tissue complication probabilities (NTCPs). For 10 oropharyngeal cancer patients, we generated robust treatment plans using 'minimax' worst-case optimization. For each patient, the setup robustness was varied from 1 to 7 mm and the range robustness was varied from 1% to 9% (+1 mm). The treatment plans were evaluated on the dose received by OARs and NTCP-values for xerostomia, dysphagia and larynx edema. A higher degree of setup

212 | Summary

robustness was found to result in considerably higher OAR doses and NTCP-values, especially in bilateral cases. In contrast, the impact of range robustness was very limited. We therefore recommend giving a high priority to setup error reduction, for example by implementing online setup verification and correction.

Samenvatting

Hoge-precisie robotbestraling en intensiteitsgemoduleerde protonentherapie ('IMPT') zijn twee relatief nieuwe radiotherapietechnieken die als doel hebben om een zeer gelokaliseerde dosis af te geven in de tumor, en tegelijkertijd het omliggende gezonde weefsel en de kritieke organen te sparen. Echter, het gebruik van honderden (in hoge-precisie robotbestraling) of duizenden (in IMPT) kleine bundels om die gelokaliseerde dosisafgifte te bewerkstelligen, zorgt ook voor een aantal uitdagingen op het gebied van de efficiëntie en nauwkeurigheid van de behandeling. Deze uitdagingen belemmeren momenteel een bredere toepassing van hoge-precisie robotbestraling en IMPT. In dit proefschrift wordt ingegaan op de verbetering van de planning en afgifte van deze twee behandelvormen teneinde deze obstakels te overwinnen.

Deel I: Hoge-precisie robotbestraling met de CyberKnife

Het CyberKnife robotbestralingssysteem bestaat uit een compacte 6-MV lineaire versneller die gemonteerd is op een robotarm. De robotarm kan gepositioneerd worden op ongeveer 120 verschillende posities die verdeeld zijn rondom de patiënt, de zogenaamde 'node-posities'. Daardoor kan de patiënt vanuit veel (niet-coplanaire) richtingen bestraald worden. Vanaf iedere node-positie worden meerdere bundels afgegeven door de richting van de lineaire versneller aan te passen. Het geïntegreerde stereoscopische beeldvormingssysteem maakt het mogelijk om zogenaamde 'tumor tracking' toe te passen. Tumorbewegingen worden hierbij gedurende een behandelfractie gevolgd door regelmatig röntgenbeelden te maken en met de robot voor verplaatsingen van de tumor te corrigeren. Hoewel veelbelovende klinische resultaten zijn behaald voor gehypofractioneerde behandelingen (meestal 3-4 fracties), hebben CyberKnife behandelingen ook een aantal beperkingen. Ten eerste is de behandeltijd per fractie relatief lang wat een bredere toepassing van de CyberKnife voor andere (conventioneel gefractioneerde) behandelingen in de weg staat. Daarnaast heeft 'tumor tracking' met behulp van het geïntegreerde röntgensysteem een beperkte temporele en geometrische nauwkeurigheid, terwijl prostaattumoren zeer snelle en grote bewegingen kunnen vertonen. De resulterende dosimetrische nauwkeurigheid van 'tumor tracking' en de optimale volgstrategie zijn tot op heden nog niet vastgesteld voor prostaatbehandelingen met de CyberKnife . Verbeteringen aangaande de efficiëntie en nauwkeurigheid van CyberKnife behandelingen zullen in Deel I van dit proefschrift aan bod komen.

In hoofdstuk 2 wordt ingegaan op het effect van bundelcollimatie (het definiëren van de vorm van het stralingsveld) op de efficiëntie van de behandeling. De CyberKnife is van oorsprong uitgerust met circelvormige collimatoren. Een mini-multileaf collimator (mMLC) maakt het daarentegen mogelijk om het bestralingsveld iedere gewenste vorm te geven wat naar verwachting tot een efficiëntere dosisafgifte leidt. Om dit te onderzoeken hebben we de behandeltijd per fractie en de plankwaliteit van de CyberKnife vergeleken voor een mMLC en de 'Iris' variable circelvormige collimator. Voor 10 patiënten met stadium I niet-klein-cellig longcarcinoom zijn allereerst mMLC plannen gemaakt met behulp van het bundelhoek optimalisatiealgoritme 'Cycle'. De bestralingsbundels werden gevormd naar de projectie van de tumor en hadden een uniforme intensiteit. In deze mMLC plannen werden 10-16 node-posities gebruikt, geselecteerd uit de 117 beschikbare node-posities. Vervolgens werden intensiteitsgemoduleerde Iris plannen gegenereerd, waarbij gebruik gemaakt werd van de node-posities die geselecteerd waren in de mMLC plannen. Op deze manier werden tijdsefficiënte behandelingen gemaakt voor beide collimatoren. Het gebruik van een mMLC resulteerde in een aanzienlijke verkorting van de behandeltijd per fractie. De gemiddelde fractieduur was 12,2 min (10,8-13.5 min) voor de mMLC plannen en 18,4 min (12,9-28,5 min) voor de Iris plannen, bij een gelijke gemiddelde longdosis. De winst in behandeltijd was vooral groot voor grote tumoren. De Iris collimator overtrof de mMLC op het gebied van haalbare plankwaliteit door het gebruik van intensteitsgemoduleerde velden in plaats van de uniforme velden in mMLC plannen. In overeenstemming met de bevindingen van deze studie is een mMLC recentelijk geïntroduceerd voor de CyberKnife.

De behandeltijd kan ook gereduceerd worden door een vermindering van het aantal node-posities in een CyberKnife behandelplan, aangezien een groot deel van de behandeltijd wordt besteed aan het bewegen van de robot tussen de node-posities. In hoofdstuk 3 wordt een methode beschreven om de node-posities te verminderen gedurende het planningsproces. Gedurende de 'node-reductie' wordt de inverse optimalisatie herhaaldelijk uitgevoerd, waarbij iedere keer node-posities met een lage bijdrage verwijderd worden uit het behandelplan. De inverse optimalisatie wordt vervolgens opnieuw uitgevoerd met nieuwe kandidaat-bundels geselecteerd van de overgebleven node-posities. We hebben deze methode geëvalueerd door behandelplannen te genereren mét en zonder deze node-reductie voor een orofarynxcarcinoom, een prostaattumor en twee longtumoren van verschillende complexiteit. Het aantal node-posities werd gemiddeld met 77% gereduceerd terwijl het aantal bundels en monitoreenheden met respectievelijk 28% en 9% toenamen. De fractieduur werd hiermee gemiddeld met 25% verkort (18%–40%).

Ten opzichte van de klinische planningsaanpak (geen herhaalde optimalisatie) werd zelfs een gemiddelde tijdswinst van 54% gerealiseerd. Het reduceren van node-posities had geen effect op de haalbare plankwaliteit, al wordt het gebruik van extra planningsrestricties aangeraden om dosisconformaliteit te garanderen en ongewenste huiddosis te vermijden.

Hoofdstuk 4 beschrijft een studie naar de dosimetrische impact van intrafractie prostaatbeweging en de effectiviteit van 'tumor tracking' en de gehanteerde veiligheidsmarges. Gebruikmakend van de behandelplannen van 3 patiënten en prostaatbewegingsdata van 17 patiënten, zijn uitgebreide simulaties uitgevoerd waarbij de volgende parameters gevarieerd werden: (1) het tijdsinterval tussen robotcorrecties (15, 60, 180, of 360 s), (2) de grootte van robotcorrecties (geen correcties, enkel translatiecorrecties, of translatiecorrecties gecombineerd met rotatiecorrecties tot 5°, 10°, of volledige correctie) en (3) de toegepaste veiligheidsmarge (o mm of 3 mm). Intrafractie prostaatbeweging beïnvloedde de afgegeven dosis aanzienlijk, maar dat kon effectief gecompenseerd worden met behulp van robotcorrecties met een tijdsinterval van 60-180 s. Om de tumor adequaat te bestralen in 98% van de behandelingen zijn translatie- en rotatiecorrecties tot 10° vereist voor o mm veiligheidsmarge, terwijl translatie- en rotatiecorrecties tot 5° benodigd zijn voor 3 mm veiligheidsmarge. Het gebruik van 0 mm marge resulteerde in aanzienlijk betere sparing van het rectum en de blaas.

Deel II: intensiteitsgemoduleerde protonentherapie

Bij IMPT worden patiënten bestraald met duizenden zeer kleine protonenbundels (of 'spots'). Aangezien protonen positief geladen deeltjes zijn, hebben ze een andere interactie met weefsel dan fotonen (die bijvoorbeeld gebruikt worden in de CyberKnife behandelingen). Protonen hebben een beperkte dracht in het weefsel, afhankelijk van de betreffende protonenenergie. De dosisafgifte van de protonen neemt eerst heel langaam toe met toenemende diepte en stijgt tot een hoge dosispiek vlak voor de maximale indringdiepte. Deze zogenaamde 'Bragg-piek' is zeer geschikt om een gelokaliseerde dosis af te geven in de tumor. De dosis vóór de Bragg-piek is namelijk relatief laag en na de Bragg-piek wordt helemaal geen dosis meer afgegeven, met het gevolg dat het omliggende gezonde weefsel goed gespaard kan worden. Tijdens de behandeling worden de protonenspots één voor één afgegeven, waarbij steeds de laterale positie, de protonenenergie en de intensiteit aangepast worden. Op deze wijze kan het gehele tumorvolume bestraald worden. Met IMPT zijn veelbelovende resultaten verkregen, maar tegelijkertijd wordt

IMPT ook gekenmerkt door een aantal uitdagingen. Allereerst beperkt het gebruik van het grote aantal (duizenden) protonenbundels de efficiëntie van de planning en de afgifte. Een verbeterde efficiëntie zal leiden tot een verbetering van plankwaliteit, patiëntcomfort en behandelnauwkeurigheid, en misschien nog wel belangrijker, het kan bijdragen aan een vermindering van de behandelkosten. De kosten van IMPT behandelingen zijn momenteel hoog in vergelijking met die van fotonenbestraling. Daarnaast leidt de sterk gelokaliseerde dosisafgifte van IMPT tot een hoge gevoeligheid voor behandelonzekerheden. Dit kan resulteren in ernstige verslechteringen van de afgeven dosisverdeling wat weer impact heeft op de effectiviteit en veiligheid van IMPT behandelingen. In Deel II van dit proefschrift wordt nader ingegaan op mogelijke verbeteringen aangaande de efficiëntie en de gevoeligheid van IMPT.

Een nieuwe en tijdsefficiënte planningsmethode voor IMPT, 'resampling' genaamd, wordt gepresenteerd in hoofdstuk 5. Deze methode is geïmplementeerd in het planningssysteem 'Erasmus-iCycle', en onderscheidt zich van traditionele planningsmethoden door zijn iteratieve aanpak waarbij herhaaldelijk de volgende stappen worden uitgevoerd: (1) random selectie van kandidaat-bundels van een zeer fijn raster, (2) inverse optimalisatie, en (3) het verwijderen van bundels met een lage bijdrage. De nieuw geselecteerde kandidaat-bundels worden steeds toegevoegd aan de bundels in de bestaande oplossing, waardoor het behandelplan enkel kan verbeteren in iedere iteratie. Voor vijf patiënten met een orofarynxcarcinoom is de optimalisatietijd en de plankwaliteit van deze resampling methode vergeleken met die van de traditionele planningsmethode waarbij bundels verdeeld zijn over een regelmatig raster. Het gebruik van resampling resulteerde in een reductie van de optimalisatietijd met een factor van 2,8 en 5,6 (max. 7,8 en 17,0) ten opzichte van anisotrope en isotrope reguliere rasters, voor een gelijke plankwaliteit. Resampling leidde tevens tot lagere doses in kritieke organen, met een mediane verlaging variërend van -4% tot 42%. Resampling bleek vooral effectief te zijn bij het gebruik van smalle protonenbundels.

Voor huidige commerciële IMPT apparatuur wordt de behandeltijd grotendeels bepaald door de tijd die nodig is om de energie van de protonenbundel aan te passen. Een vermindering van het aantal energielagen in een behandelplan zal dan ook zorgen voor een kortere fractieduur. In **hoofdstuk** 6 wordt een methode gepresenteerd waarmee het aantal energielagen verminderd kan worden zonder dat dit ten koste gaat van de robuuste plankwaliteit. De methode is geïmplementeerd in het planningssysteem 'Erasmus-iCycle' en heeft overeenkomsten met de methode beschreven in hoofdstuk 3. De methode bestaat uit twee componenten: (1) minimalisatie van de logaritme van het totale bundelgewicht per energielaag, en (2) iteratief excluderen van energielagen met een lage bijdrage. De methode is gebenchmarkt door robuuste 'tijdsefficiënte plannen' (mét energielaagreductie) met robuuste 'standaard klinische plannen' (zonder energielaag-reductie) te vergelijken voor vijf patiënten met een orofarynxcarcinoom en vijf patiënten met een

prostaattumor. 'Worst-case' robuuste optimalisatie werd toegepast, waarbij mogelijke positioneringsfouten van 3 mm en onzekerheden in de protonendracht van 3% + 1 mm meegenomen werden in de optimalisatie. De methode resulteerde in een gemiddelde reductie van het aantal energielagen van 45% (orofarynxcarcinoom) en 28% (prostaattumor). De gemiddelde behandeltijd werd verkort met 25%, 32% of 38% (orofarynxcarcinoom) en met 16%, 20% of 24% (prostaattumor), voor schakeltijden tussen energielagen van 1, 2 of 5 s.

In hoofdstuk 7 is de impact van onzekerheden in patiëntpositionering, anatomie en protonendracht op de afgegeven dosis in IMPT behandelingen onderzocht. Voor 10 hoofd-hals kankerpatiënten zijn behandelplannen gemaakt met 3, 5 of 7 bundelrichtingen. Een veiligheidsmarge van 5 mm werd aangebracht rondom het hoge-dosis doelgebied en het lage-dosis doelgebied. De daadwerkelijk afgegeven dosis werd gesimuleerd door (>3700) herberekeningen van de dosis uit te voeren met verschillende fouten in patiëntpositionering, protonendracht en anatomie. Onzekerheden in de anatomie van een patiënt werden gesimuleerd met behulp van een herhaal-CT-scan (scan halverwege de behandeling) en met niet-rigide registratie om dosisoptelling te kunnen uitvoeren. De herhaal-CT-scans werden ook gebruikt om adaptieve planning te simuleren door een nieuw behandelplan te maken voor de tweede helft van de behandelfracties. Uit de resultaten bleek dat een verhoging van het aantal bundelrichtingen niet leidde tot een verbeterde robuustheid van de behandelplannen. Voor behandelplannen met 3 bundelrichtingen, bleek dat wanneer alle bovengenoemde foutbronnen (patiëntpositionering, anatomie en protonendracht) werden meegenomen, de adequate tumorbestraling slechts in 69% en 88% van de simulaties gehaald werd, voor respectievelijk het hoge-dosis en lage-dosis doelgebied. De dosis in kritieke organen vertoonde aanzienlijke spreiding rondom de geplande waarden. De foutbron met de grootste impact op de dosisverdeling verschilde tussen patiënten. Adaptieve planning met behulp van de herhaal-CT-scan leidde tot een forse verbetering van de tumordosis. Adequate bestraling werd in 96% (hoge-dosis doelvolume) en 100% (lage-dosis doelvolume) van de simulaties behaald wanneer het plan aangepast werd halverwege de behandeling.

De invloed van ademhalingsbeweging en de effectiviteit van 'real-time tumor (re)tracking' (het continue volgen van een bewegende tumor met de protonenbundel) zijn onderzocht voor IMPT in hoofdstuk 8. Gebruik makend van een 4D dosisberekeningsalgoritme en gesimuleerde ademhalingsbeweging, is een algoritme voor 'real-time tumor tracking' toegepast op eenvoudige homogene en heterogene in-silico fantomen. Voor 'tracking' en 'retracking' (een combinatie van 'tracking' en 'rescanning') zijn verschillende theoretische behandelscenario's onderzocht: (1) geen 'tracking' (ongecorrigeerde beweging), (2) perfecte 'tracking' (geen vertragingen en exacte kennis van de tumorpositie), en (3) imperfecte 'tracking' (gesimuleerde tijdsvertragingen of voorspellingsfouten in de tumorpositie). In het homogene fantoom werd de beoogde tumordosis met perfecte 'tracking' gehandhaafd voor alle bewegingsamplitudes en fases. Tijdsvertragingen en positie-onnauwkeurigheden (geen of imperfecte 'tracking') kunnen echter leiden tot een aanzienlijke verslechtering van de afgegeven dosis in de tumor, maar dit kan effectief worden tegengegaan door het toepassen van 'retracking'. In het heterogene fantoom kon de geplande dosis niet behouden worden, zelfs niet met perfecte 'tracking'. Dit kwam doordat het aanpassen van de bundelposities veranderingen teweeg bracht in de bundelvorm in de aanwezigheid van verschillende dichtheden. Het gebruik van 'rescanning' met een hoge rescan-factor bleek de meest gunstig strategie voor het heterogene fantoom.

Om met onzekerheden in patiëntpositionering en protonendracht om te gaan wordt bij IMPT de voorkeur gegeven aan robuuste planningsmethodes boven het gebruik van traditionele veiligheidsmarges. Echter, de grootte van de fouten waarmee tijdens robuuste planning rekenig gehouden wordt heeft een effect op de dosis die de kritieke organen ontvangen. In hoofdstuk 9 is de klinische impact van de mate van robuustheid tegen onzekerheden in positionering en protonendracht gekwantificeerd in termen van de dosis in kritieke organen en de kans op bijwerkingen. Voor 10 patiënten met een orofarynxcarcinoom hebben we robuuste plannen gegenereerd met behulp van 'worst-case' optimalisatie. Voor elke patiënt is de robuustheid tegen positioneringsfouten gevarieerd van 1 tot 7 mm en de robuustheid tegen fouten in de protonendracht van 1% tot 9% (+1 mm). De behandelplannen werden vervolgens beoordeeld op de dosis in kritieke organen en de kans op bijwerkingen (voor xerostomie, dysfagie en lanrynxoedeem). Een hogere mate van robuustheid tegen positioneringsfouten resulteerde in aanzienlijk hogere dosis in kritieke organen en een verhoogde kans op bijwerkingen, met name voor de bilaterale patiëntengroep. De robuustheid tegen onzekerheden in protonendracht had daarentegen relatief weinig invloed op de dosis en het complicatierisico. Geadviseerd wordt dan ook om een hoge prioriteit te geven aan het minimaliseren van fouten in de patiëntpositionering, bijvoorbeeld door het implementeren van online positieverificatie en correctie.

Dankwoord

De afgelopen jaren heb ik met veel plezier aan mijn onderzoek gewerkt. Graag wil ik iedereen bedanken die daar op één of andere manier aan bijgedragen heeft. Veel waardering gaat uit naar mijn collega's van de afdeling radiotherapie –collega's van de klinische fysica, laboranten en radiotherapeuten– voor de inspirerende, motiverende en gezellige samenwerking. Daarnaast ben ik mijn vrienden en (schoon-)familie erg dankbaar voor hun steun en interesse in mijn onderzoek. In het bijzonder wil ik de volgende mensen bedanken voor hun bijdrage aan dit proefschrift.

Mijn promotor, Ben Heijmen. Beste Ben, het waren jouw colleges tijdens mijn studie die me inspireerden om voor dit vakgebied te kiezen en dit heeft, via een afstudeeropdracht, uiteindelijk geleid tot dit proefschrift. Je voorzag mijn artikelen altijd van grondig en kritisch commentaar en ook onze discussies vond ik erg waardevol. Tevens wil ik je bedanken voor de mogelijkeid om mijn onderzoek de komende jaren als postdoc op jouw afdeling te continueren.

Mijn copromotor, Mischa Hoogeman. Beste Mischa, ik heb voornamelijk met jou samengewerkt en dat was me een groot genoegen. Je liet me vrij om een eigen richting aan mijn onderzoek te geven en tegelijkertijd kon ik altijd bij je aankloppen voor overleg, interessante ideeën en commentaar op mijn artikelen. Ik kijk er erg naar uit om onze samenwerking de komende jaren voort te zetten.

Hanne Kooy, jij bent zo genereus geweest om jullie dosisberekeningsalgoritme voor protonen met ons te delen en daarmee heb je een essentiële bijdrage geleverd aan het protonenonderzoek op onze afdeling. Daarnaast waardeer ik het enorm dat we regelmatig een beroep hebben mogen doen op jouw kennis en ervaring op het gebied van protonentherapie.

Tony Lomax, you offered me the possibility to perform a research project in the exciting field of proton therapy. During my internship at PSI I experienced how inspiring and rewarding research can be and it marked the beginning of my scientific career. I would like to thank you for giving me this opportunity and for supporting me with my very first article.

Sebastiaan Breedveld, jouw werk op het gebied van plan optimalisatie is de basis geweest voor al mijn onderzoek. Je bent daarmee een onmisbare schakel geweest in het werk wat ik de afgelopen jaren heb gedaan. Heel veel dank daarvoor. We hadden regelmatig een andere kijk op treatment planning maar dat zorgde juist voor inspirerende discussies en nieuwe ideeën.

Inger-Karine Kolkman-Deurloo, inhoudelijk hebben we weinig samengewerkt, maar het was fijn om jou als collega te hebben. Zo kon ik in voor mij moeilijkere tijden altijd bij je binnen lopen om even te praten.

Mijn afstudeerstudenten Lorella, Iris, Charlotte, Christiana en Sebastian. Ik kijk met veel plezier terug op de tijd dat ik jullie (gedeeltelijk) heb mogen begeleiden. We hebben aan uitdagende projecten gewerkt en jullie hebben daarbij mooie resultaten behaald. Ik heb veel van jullie geleerd; inhoudelijk door jullie vernieuwende ideeën, maar ook op het gebied van begeleiding.

Mijn kamergenotes Sabrina, Aafke en Noëlle. Met jullie op de kamer heb ik een fijne tijd gehad. Naast dat we elkaar hielpen als dat nodig was, werden onze werkdagen ook gekenmerkt door koffie, muziek, gezelligheid en geluidsoverlast.

Mijn paranimfen, Noëlle en Dennis. Lieve Noëlle, mijn promotietijd begon bij jou op de kamer. Jij hebt mij geholpen met het schrijven van mijn eerste artikelen en ik kon jou weer ondersteunen met rekenwerk voor jouw onderzoek. Dit groeide uiteindelijk uit tot een waardevolle vriendschap waarin je een grote steun was in verdrietige tijden. Dennis, lieve broer, ik ben heel blij dat jij tijdens mijn promotie naast me staat. De afgelopen jaren hebben we samen veel meegemaakt en daardoor is onze band alleen maar hechter geworden. Ik weet dat ik altijd op je kan bouwen en ben trots dat ik jou als broer heb.

Lieve Roel en Carol, liefdevollere ouders en grotere steun had ik me niet kunnen wensen. Ik ben jullie heel erg dankbaar voor alles wat jullie me gegeven hebben. Roel, je had mijn promotie zo graag mee willen maken; ik mis je op dit soort momenten.

Lieve Rosa, wat is het toch geweldig om jou in mijn leven te hebben!

Curriculum vitae

Steven van de Water was born on January 24, 1983 in Veghel, the Netherlands, where he finished high school at the Zwijsen College in 2000. In the same year he started the Bachelor Aerospace Engineering at the Delft University of Technology. During his study he was a board member of several committees related to student societies in Delft and took one year off to be the president of a student rowing club. He obtained his Bachelor's degree in 2006 and continued with the Master Biomedical Engineering, specialization Clinical Physics. In 2007, he spent 4 months in Switzerland at the Paul Scherrer Institute for a research internship on proton therapy. Upon returning in the Netherlands, he started with his Master thesis project on CyberKnife treatment planning at the Department of Radiation Oncology of the Erasmus MC Cancer Institute in Rotterdam. This resulted in obtaining his Master's degree in 2009 with 'Cum Laude' distinction. In March 2009, he started his PhD at the same department, investigating different aspects of high-precision radiotherapy, which resulted in this thesis. Steven now continues his research on improving treatment planning and delivery of proton therapy at the Erasmus MC Cancer Institute.

PhD portfolio

Name PhD student: Steven van de Water Erasmus MC Department: Radiation Oncology Research School: Molecular Medicine PhD period: 2009-2015 Promotor: Prof. dr. B.J.M. Heijmen Copromotor: dr. M.S. Hoogeman

1. PhD Training	
General courses	
Biomedical English Writing and Communication	2009
Research management	2010
International conferences	
ESTRO 29, the Netherlands (oral presentation)	2009
CyberKnife Society Scientific Meeting, USA (oral presentation)	2010
AAPM 52, USA (oral presentation)	2010
AAPM 54, USA (2 oral presentations)	2012
ESTRO 32, Switzerland (oral presentation)	2013
PTCOG 52, Germany (oral presentation)	2013
ESTRO 33, Vienna (oral presentation and poster presentation)	2014
National conferences	
RKF meeting, Leiden (oral presentation)	2009
RKF meeting, Amsterdam (VUmc) (oral presentation)	2012
RKF meeting, Amsterdam (AMC) (oral presentation)	2012
NVKF conferentie, Zeist (oral presentation)	2014
RKF meeting, Utrecht (oral presentation)	2014
Wetenschappelijke vergadering NVRO, Amersfoort (oral presentation)	2014
Presentations as invited speaker	
AERO Physics meeting (CyberKnife), Vienna	2014
Erasmus MC SURE symposium, Rotterdam	2014
In-house presentations	
Journal club	2010(2x), 2011, 2012, 2014
Werkbespreking fysica	2009(2x), 2011(2x), 2012, 201-
Refereeravond	2009, 2010, 2012, 2013
2. Teaching	
Lecturing	
Guest lecturer TU Delft: 'Planning for Particle Therapy'	2009, 2010, 2012, 2013, 2014
Guest lecturer RTT Skillslab: 'Planning for Particle Therapy'	2011, 2013, 2014
Lecture statistical methods RTT Skillslab	2010 (2x)
Supervising Master students	
Lorella Valli, Alma Mater Studiorum, Bologna	2011
Iris van Dam, TU Delft	2012-2013
Charlotte Hartman, TU Delft	2013–2014
Christiana Balta, TU Delft	2014
Sebastian van der Voort, TU Delft	2014–2015

