Evaluation and Application of Automated Treatment Planning in
Radiation Therapy

Abdul Wahab Sharfo

Evaluation and Application of Automated Treatment Planning in Radiation Therapy

Evaluation and Application of Automated Treatment Planning in Radiation Therapy

Evaluatie en applicatie van het automatisch genereren van behandelplannen in de radiotherapie

Proefschrift

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

Prof. dr. F.A. van der Duijn Schouten

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

dinsdag 23 maart 2021 om 15.30 uur door

Abdul Wahab M. Sharfo

geboren te Aleppo, Syrië

ERASMUS UNIVERSITY ROTTERDAM

Promotiecommissie:

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

Overige leden: Prof. dr. R.A. Nout
                       Prof. dr. J.J. Sonke
                       Prof. dr. D. Verellen

Copromotor: Dr. ir. M.L.P Dirkx

وقل ربِ زدني علماً

Contents

Introduction

1.1 Radiation therapy

Radiation therapy (RT) is frequently used to treat cancer by exposing the patient to ionizing photons, electrons or protons to kill malignant tumor cells. In Europe, about half of the cancer patients receive radiation therapy during their course of disease. Often, it is applied in conjunction with other treatment modalities, mainly surgery or chemotherapy. In *external beam radiation therapy* (EBRT), the patient is treated with an ionizing radiation source that is positioned outside the body. EBRT with a linear accelerator is the most commonly applied technique.

The goal of radiation therapy is to eradicate cancer cells through high doses of ionizing radiation while minimizing the dose in surrounding normal tissues in order to avoid or minimize negative side effects of the treatment. Radiation therapy can be prescribed with curative intent, aiming at elimination of all malignant tumor cells and thereby at prevention of cancer recurrence, or with palliative intent to relieve symptoms like pain caused by cancer.

Generally, the total radiation dose is delivered in multiple fractions over a period of several days or weeks to reduce the damage to healthy tissues, especially the sensitive *organs-at-risk* (OAR). Advantages of radiation therapy compared to other treatment modalities include that the treatment is *i)* non-invasive with the potential of organ sparing, *ii)* largely selective for the patient's individual tumor, and *iii)* relatively cheap. However, radiation therapy can cause both acute (i.e., during and shortly after treatment)

1

and late (i.e., months or years after treatment) side effects. Excessive toxicity may result from a *sub-optimal* treatment plan.

1.2   External beam radiation therapy treatment planning

Prior to treatment delivery, a *computed tomography* (CT) scan is acquired and used to contour the 3-dimensional patient anatomy with the exact positions and shapes of the target/tumor volume and OARs. The radiation oncologist prescribes the radiation dose to be delivered to the target and the maximum tolerable dose constraints and objectives for OARs, such as the spinal cord, heart or lungs. Based on this information, a dosimetrist will then generate a treatment plan using a software application called *treatment planning system* (TPS).

The aim of treatment planning is to establish treatment machine parameters (beam directions, field sizes, etc.) that will result in delivery of the requested target dose with minimal doses in the OARs, in line with the defined constraints and objectives. Treatment planning involves solving a large multi-criterial problem with conflicting and prioritized objectives. Optimizing both beam angles and beam intensity profiles is a non-convex combinatorial multi-criterial optimization problem. With the current TPSs, treatment plans are generated in an iterative *trial-and-error* process. With this so-called manual planning the treatment plan quality may highly depend on the available planning time, skills and experience of the dosimetrist. For individual patients, this can result in plans that are clinically acceptable but sub-optimal, which can increase the probability of developing radiation-associated complications.

1.2.1   Automated radiotherapy treatment plan optimization

Over the past decades, the dramatic increase of computer power, improvement of dose calculation algorithms, and the availability of *multi-leaf collimators* (MLCs) have led to more precise and conformal radiation treatments, e.g., by using *intensity-modulated radiation therapy* (IMRT) or *volumetric-modulated arc therapy* (VMAT). More recently, much effort has been put in the development of algorithms to automate the labor-intensive, trial-and-error process of treatment planning. In Rotterdam, Erasmus-iCycle has been developed and clinically implemented for automated and integrated multi-criterial fluence map and beam angle optimization. For an individual patient, Erasmus-iCycle automatically generates a *Pareto-optimal* plan that is clinically favorable. In *Pareto-optimal* plans, no objective can be improved without a sacrifice in another.

1.3 Thesis overview

This thesis focuses on the evaluation of automated treatment planning with Erasmus-iCycle and on its use in treatment planning studies for development and comparison of different treatment techniques. By automating the planning based on exactly the same decision-making strategy '*wish-list*' for all involved techniques, often observed bias in published treatment planning studies can significantly be reduced. Due to the automation of treatment planning, the number of involved treatment plans can also be significantly enhanced compared to literature.

In Chapter 2, we developed and validated a method for fully automated generation of VMAT plan-libraries for plan-of-the-day adaptive radiotherapy for advanced cervical cancer. The quality of the automatically generated VMAT and IMRT plans was evaluated by comparison with plans that were manually generated by an expert dosimetrist in the absence of time pressure. In addition, treatment delivery times with different multi-leaf collimators were compared.

In Chapters 3-5, we compared fully automated multi-criteria VMAT plan generation using Erasmus-iCycle with manual treatment planning in absence of time pressure in collaborations with international partners. Chapter 3 describes automated VMAT planning for spinal metastases. In Chapter 4, the clinical advantages of automated VMAT plan generation were investigated for the treatment of gastric cancer patients by comparing dosimetric parameters and expected *normal tissue complication probabilities* (NTCPs). Chapter 5 presents the evaluation of automated VMAT planning for sequential-boost whole pelvic prostate cancer treatment using a dosimetric and a blinded physician's plan scoring.

In Chapters 6-9, Erasmus-iCycle was used to systematically compare different treatment delivery techniques. In Chapter 6, for locally-advanced cervical cancer patients, single- and dual-arc VMAT plans and IMRT plans with 9, 12, and 20 equi-angular beams were mutually compared. In Chapters 7 and 8, we evaluated the benefit of a novel treatment approach, designated VMAT+, i.e., the addition of a single or a few IMRT beams with optimized non-coplanar beam angles to a VMAT delivery, with simultaneous optimization of the VMAT parameters, the non-coplanar beam configuration, and the intensity profiles of these beams. Chapter 7 investigates the application of the VMAT+ technique for *stereotactic body radiation therapy* (SBRT) of liver metastases. Chapter 8 extends the VMAT+ approach to prostate SBRT patients, and demonstrates the use of a simple non-coplanar two-beam angle class solution to improve VMAT plans without the need of a lengthy, individualized non-coplanar beam angle optimization for each patient. In Chapter 9, the use of Erasmus-iCycle for fully automated non-coplanar CyberKnife®

1

(CK) treatment planning is described. For prostate SBRT we used this system to compare automatically generated CK plans to automatically generated coplanar VMAT plans.

In Chapter 10, we applied automated treatment planning to retrospectively analyze dose delivery in the HYPRO multi-institutional randomized trial comparing conventional and hypo-fractionated EBRT for prostate cancer (Aluwini *et al.* 2015, 2016; Incrocci *et al.* 2016). In total, plans of 725 prostate cancer patients were analyzed in this planning study. Differences in plan quality between the manually generated, clinically delivered plans and the automatically generated plans were quantified using dose-volume histogram parameters and NTCP models for late gastrointestinal toxicity, derived from reported toxicity in the HYPRO trial. The HYPRO trial showed some enhanced toxicity for the hypo-fractionation arm. In this study, we also compared the expected reduced toxicity for the investigated hypo-fractionation schedule in case treatment plans would have been generated with automated planning, with the (accepted) observed toxicity for the conventional-fractionation schedule.

In Chapter 11, the use of Erasmus-iCycle for prospective feedback on plan quality in a multi-institutional randomized trial on hepatocellular carcinoma was investigated. This plan feedback was part of the *quality assurance* (QA) program of the trial. Based on feedback, institutes could adapt their plan prior to start of treatment.

In Chapter 12, we conclude with a discussion on the findings of this work, followed by a future outlook.

Validation of Fully Automated VMAT Plan Generation for Library-Based Plan-of-the-Day Cervical Cancer Radiotherapy

Abdul Wahab M. Sharfo, Sebastiaan Breedveld, Peter W.J. Voet, Sabrina T. Heijkoop, Jan-Willem M. Mens, Mischa S. Hoogeman and Ben J.M. Heijmen

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

Abstract

Purpose: To develop and validate fully automated generation of VMAT plan-libraries for plan-of-the-day adaptive radiotherapy in locally-advanced cervical cancer.

Material and Methods: Our framework for fully automated treatment plan generation (Erasmus-iCycle) was adapted to create dual-arc VMAT treatment plan-libraries for cervical cancer patients. For each of 34 patients, automatically generated VMAT plans (autoVMAT) were compared to manually generated, clinically delivered 9-beam IMRT plans (CLINICAL), and to dual-arc VMAT plans generated manually by an expert planner (manVMAT). Furthermore, all plans were benchmarked against 20-beam equi-angular IMRT plans (autoIMRT). For all plans, a PTV coverage of 99.5% by at least 95% of the prescribed dose (46 Gy) had the highest planning priority, followed by minimization of V_{45Gy} for small bowel (SB). Other organs-at-risk (OARs) considered were bladder, rectum, and sigmoid.

Results: All plans had a highly similar PTV coverage, within the clinical constraints (above). After plan normalizations for exactly equal median PTV doses in corresponding plans, all evaluated OAR parameters in autoVMAT plans were on average lower than in the CLINICAL plans with an average reduction in SB V_{45Gy} of 34.6% ($p < .001$). For 41/44 autoVMAT plans, SB V_{45Gy} was lower than for manVMAT ($p < .001$, average reduction 30.3%), while SB V_{15Gy} increased by 2.3% ($p = .011$). AutoIMRT reduced SB V_{45Gy} by another 2.7% compared to autoVMAT, while also resulting in a 9.0% reduction in SB V_{15Gy} ($p < .001$), but with a prolonged delivery time. Differences between manVMAT and autoVMAT in bladder, rectal and sigmoid doses were ≤ 1%. Improvements in SB dose delivery with autoVMAT instead of manVMAT were higher for empty bladder PTVs compared to full bladder PTVs, due to differences in concavity of the PTVs.

Conclusions: Quality of automatically generated VMAT plans was superior to manually generated plans. Automatic VMAT plan generation for cervical cancer has been implemented in our clinical routine. Due to the achieved workload reduction, extension of plan-libraries has become feasible.

2.1 Introduction

There is a growing evidence that more conformal dose distributions can limit frequently observed side effects in external beam radiotherapy (EBRT) of cervical cancer patients (Chen *et al.* 2007; Fernandez-Ots & Crook 2013). The concavely shaped planning target volume (PTV) of locally-advanced cervical cancer patients requires intensity-modulated radiotherapy (IMRT) or volumetric-modulated arc therapy (VMAT) techniques in order to improve conformality and subsequent sparing of surrounding healthy tissues compared to 3-dimension conformal radiotherapy (3D-CRT) (Mundt *et al.* 2002; van de Bunt *et al.* 2006).

The use of highly conformal dose distributions for cervical cancer is challenged by the large inter-fraction variability in the shape and position of the target volume (Ahmad *et al.* 2011). To avoid the need for large margins to compensate for the geometrical uncertainties (Bondar *et al.* 2012), our institute has implemented a plan-of-the-day (PotD) strategy using individualized libraries of treatment plans created in the treatment preparation phase (Bondar *et al.* 2012; Heijkoop *et al.* 2014). During the daily treatment fractions, a CBCT-scan is acquired prior to dose delivery to select the plan from the library that best fits the patient's internal anatomy. The library is constructed using two planning CTs made with a full and an empty bladder to construct internal target volumes (ITVs) encompassing the cervix-uterus motion for sub-ranges in bladder volume. Up to recently, the library contained two 9-beam IMRT plans (one for empty-to-half-full and one for half-full-to-full bladders) for patients with a uterine fundus displacement between the full and empty bladder CT-scan larger than 2.5 cm (large movers). For patients with smaller fundus mobility (small movers), the library contained one 9-beam IMRT plan covering all bladder-induced cervix-uterus motion. Apart from the IMRT plans, a 4-field motion-robust 3D-CRT plan was always added, to be selected in fractions with inadequate target coverage by the IMRT plans, or in case of poor CBCT image quality prohibiting proper evaluation of the patient's anatomy (see for full details Heijkoop *et al.* (2014)). All plans were generated manually.

A library-based PotD approach with multiple pre-treatment generated plans per patient requires an enhanced planning effort. The generally applied, manual trial-and-error process for treatment plan generation is time consuming, and the outcome does largely depend on the efforts and skills of the planner and on allotted planning time. This limits the widespread clinical implementation of library-based plan-of-the-day strategies or puts constraints on the number of plans in a library. In previous studies, we have demonstrated for head-and-neck and prostate cancer that automated treatment plan generation could substantially reduce the workload compared to conventional planning with equal or superior plan quality (Voet *et al.* 2013b, 2014).

The purpose of this work was to develop and validate automatic dual-arc VMAT treatment planning for cervical cancer patients, based on our in-house platform for automated plan generation. For each automated plan generation, a manually contoured planning CT-scan was used as input. For validation of plan quality, fully automatically generated dual-arc VMAT plans (autoVMAT) were compared with *i)* manually generated, clinical 9-beam IMRT plans (CLINICAL), *ii)* dual-arc VMAT plans, manually generated by an expert planner (manVMAT), and *iii)* automatically generated 20-beam equi-angular IMRT plans (autoIMRT). For the large movers, separate analyses were performed for the plans generated for the empty-to-half-full-bladder PTVs and the half-full-to-full-bladder PTVs.

All final plans were generated with our clinical treatment planning system (TPS), with an implemented Monte Carlo dose calculation engine. For a sub-group of patients, automatically generated VMAT and IMRT plans were executed at a treatment machine equipped with a novel, fast multi-leaf collimator for treatment time assessment and comparison.

2.2 Material and methods

2.2.1 Patients

In this study, the CT- and treatment plan data of 34 patients without pelvic or para-aortic nodal involvement was used to validate multi-criterial autoVMAT planning. All patient-related information was fully anonymized prior conducting the research. According to the regulations of the Ethics Committee of Erasmus MC, no ethical approval for this retrospective study was needed as there was no impact on treatment and the applied patient data. EBRT was delivered to a dose of 46 Gy in 2 Gy daily fractions, followed by brachytherapy. Thirty-one patients were treated in prone position, lying on a small bowel displacement system (belly board), and three patients were treated supine. To limit the dose to healthy tissues, all patients were treated according to the plan-of-the-day protocol with 9-beam IMRT plans, as briefly described in the Introduction (Section 2.1) and extensively explained in Bondar *et al.* (2012), and Heijkoop *et al.* (2014). For 24 patients, the library contained a single IMRT plan (small movers, see Section 2.1), while the other 10 patients had two IMRT plans (for empty-to-half-full and half-full-to-full bladder filling, respectively; large movers), resulting in a total of forty-four 9-beam IMRT plans. Apart from the 34 patients used for validation of automatic planning, 5 different training patients were used for the '*wish-list*' generation (below).

All 44 clinical plans were produced with the clinical TPS (Monaco® version 3.3, Elekta AB, Sweden) for a photon beam energy of 10 MV, using an Elekta Synergy® linear accelerator equipped with the MLCi2 multi-leaf collimator (MLC).

2.2.2 Plan generations

For each of the 44 clinically applied structure sets in the 34 planning CT-scans, a dual-arc VMAT plan was automatically generated (autoVMAT) and compared with the clinically delivered 9-beam IMRT plan (CLINICAL), and a manually generated dual-arc VMAT plan (manVMAT). Finally, the autoVMAT plan was also benchmarked against an automatically generated 20-beam equi-angular IMRT plan, delivered with dynamic multi-leaf collimation (autoIMRT) (technical details below).

All clinical and study plans were generated in line with our clinical PotD protocol. To establish the PTV, the cervix-uterus ITV (see Section 2.1), nodal clinical target volume (CTV) and parametria were combined and expanded by a 1-cm margin in all directions. At least 99.5% of the PTV should receive ≥ 95% of the prescribed dose (D^p), and no more than 0.2% of the PTV may receive ≥ 110% of the D^p. Dose distributions should be conformal while maximally sparing the organs-at-risk (small bowel (SB), rectum, bladder, and sigmoid), giving highest priority to SB sparing (Roeske *et al.* 2003). As in clinical practice, for all 44 structure sets, autoVMAT, manVMAT, and autoIMRT plans were produced for an Elekta Synergy® linear accelerator equipped with the MLCi2 multi-leaf collimator (MLC) and a photon beam energy of 10 MV. For treatment time studies, for a sub-group of 5 patients, automatic plans were also created for the much faster Agility™ MLC (Elekta AB, Sweden) with 160 leaves, with effective leaf speeds of up to 6.5 cm/sec (Cosgrove *et al.* 2009; Bedford *et al.* 2013).

2.2.3 Clinical IMRT plan generation

The clinically applied 9-beam, step-and-shoot IMRT plans were manually generated by dosimetrists in Monaco®, using pre-optimized beam angles established with Erasmus-iCycle (Breedveld *et al.* 2012).

2.2.4 Manual generation of VMAT plans (manVMAT)

For each planning CT-scan/structure set, a dual-arc VMAT plan (manVMAT) was generated manually with the clinical TPS in a trial-and-error process by an expert planner (A.W.M.S.), in the absence of time pressure and without prior knowledge of results of other planning strategies.

2.2.5 Automated generation of VMAT and 20-beam IMRT plans (autoVMAT, autoIMRT)

In a previous study (Sharfo *et al.* 2015), our in-house Erasmus-iCycle optimizer was prepared for fully automated multi-criteria generation of treatment plans for cervical cancer patients. These were non-deliverable plans as only fluence optimization was performed (no segmentation). In that study, Erasmus-iCycle plans were used by a human planner to assist in the manual generation of deliverable plans with the clinical TPS. We observed that IMRT with 12 or more beams outperformed dual-arc VMAT in overall plan quality.

In this study, we built a system for fully automated generation of deliverable IMRT and VMAT plans for cervical patients. In this system, Erasmus-iCycle is used as a pre-optimizer for automated multi-criterial generation of a (non-deliverable) plan. The dose distribution of this plan is then automatically converted into a patient-specific template for automatic generation of a deliverable plan by Monaco®. The building of this system for fully automated, multi-criterial generation of deliverable plans consisted of several steps and was done with plans of so-called tuning patients. In the first step, we further enhanced the Erasmus-iCycle plan quality for cervical cancer patients by optimizing the so-called wish-list (below) to better stress minimization of the small bowel V_{45Gy}, as the most important OAR objective. Secondly, the automated building of patient-specific templates for automated plan generation by Monaco® was tuned such that Erasmus-iCycle plans were properly re-constructed in Monaco®. The generated patient-specific templates have a layering (Monaco® terminology) determined by the Erasmus-iCycle priorities, in addition to parallel and serial cost-functions with volume-effect parameters for OARs (details in Voet *et al.* (2014)).

Although Erasmus-iCycle is an algorithm for automated optimization of IMRT fluence profiles and selection of beam configurations (Breedveld *et al.* 2012), in this study, only beam profiles were optimized. To this purpose, Erasmus-iCycle uses the *2-phase-ε-constraint* (2*pεc*) lexicographic optimization algorithm for multi-criterial generation of a *Pareto-optimal* IMRT plan. Plan generation with Erasmus-iCycle is based on a wish-list, which contains hard constraints that are strictly met, and prioritized objectives to steer the multi-criterial optimization (Breedveld *et al.* 2007, 2009). The higher the priority of an objective, the higher the chance that the planning aim will be achieved, or even superseded. For each treatment site, the wish-list is generated in an iterative process where initial results are presented to the physician, and their feedback is used to adjust the wish-list, after which new plans are generated and presented to the physician. The result at the end of the process is a single wish-list, applicable to all patients of a certain treatment group.

Table 2.1: Wish-list applied for all automatic plan generations (autoVMAT and autoIMRT).

Constraints				
	Structure	**Type**	**Limit**	
	PTV	Max	105% of D^p	
	PTV Shell 40 mm	Max	50% of D^p	
	Unspecified Tissues	Max	105% of D^p	

Objectives				
Priority	**Structure**	**Type**	**Goal**	**Parameters**
1	PTV	↓ LTCP	0.5	$D^p = 46$ Gy, $\alpha = 0.8$
2	PTV Shell 3 mm	↓ Max	95% of D^p	
3	PTV Shell 15 mm	↓ Max	75% of D^p	
4	Small Bowel	↓ EUD	45 Gy	$k = 4$
5	Small Bowel	↓ EUD	15 Gy	$k = 3$
6	Small Bowel	↓ Mean	10 Gy	
7	Sigmoid	↓ EUD	30 Gy	$k = 4$
	Colon	↓ EUD	30 Gy	$k = 4$
8	Small Bowel	↓ Max	D^p	
9	PTV Shell 25 mm	↓ Max	60% of D^p	
	Skin Ring 20 mm	↓ Max	35% of D^p	
10	Bladder	↓ EUD	30 Gy	$k = 4$
11	Rectum	↓ EUD	30 Gy	$k = 4$
12	Sigmoid	↓ Mean	25 Gy	
	Colon	↓ Mean	25 Gy	
13	Bladder	↓ Mean	25 Gy	
14	Rectum	↓ Mean	25 Gy	

Abbreviations: D^p = prescribed dose, LTCP = *Logarithmic Tumor Control Probability*, α = cell sensitivity parameter, EUD = *Equivalent Uniform Dose*, k = volume-effect parameter.

The wish-list used in this study is presented in Table 2.1. The highest priority objective was given to the PTV, using the *logarithmic tumor control probability* (LTCP) function to obtain adequate PTV coverage for all plans (see Breedveld *et al.* (2012), and Voet *et al.* (2014) for details on the LTCP). The applied cell sensitivity parameter (α) was set to 0.8. To ensure steep dose fall-offs outside the PTV, as well as conformal dose distributions, one constraint and three objectives (priorities 2, 3 and 9) for shells around the PTV were defined. In line with clinical practice, SB had the highest priority objectives for OAR sparing (priorities 4, 5, 6 and 8 in Table 2.1). Priority 4, reduction of the SB *equivalent uniform dose* (EUD) with volume-effect parameter ($k = 4$) focuses on the reduction of relatively high doses. A skin ring of 2 cm wide, from the body contour towards the patient's internal, was defined to control entrance doses (priority 9). For sigmoid, colon, bladder and rectum, EUD and mean dose objectives were used as shown in Table 2.1.

2.2.6 Plan quality evaluations

Generated plans were visually checked and verified for clinical use by a clinician (S.T.H. or J.W.M.M.) and a medical physicist (M.S.H.). For quantitative comparisons, autoVMAT, CLINICAL, manVMAT and autoIMRT plans were first normalized to obtain exactly equal median PTV doses and were then mutually compared regarding OAR sparing. For SB,

the volume percentages receiving more than 45 Gy (V_{45Gy}, the most relevant OAR plan parameter for clinical plan evaluation), and more than 15 Gy (V_{15Gy}) (correlated to grade ≥ 3 acute SB toxicity (Kavanagh *et al.* 2010)), and the mean dose, D_{mean}, were analyzed. For the other OARs, only D_{mean} was evaluated. Quantitative evaluation of plans also included the conformity index $CI_{95\%}$ (ratio of the total volume receiving 95% of the prescribed dose to the PTV) (Feuvret *et al.* 2006). Paired two-sided Wilcoxon signed-rank tests were performed to compare the different planning strategies. Differences were considered statistically significant for $p < .05$.

2.2.7 Delivery time comparisons

For a sub-group of 5 patients, autoVMAT and autoIMRT plans were generated for both MLCi2 and the fast Agility™ MLCs, and delivered at a treatment machine. The experiments for the Agility™ MLC were performed at the Dutch Cancer Institute, as this MLC was not available at Erasmus MC. A stopwatch was used to measure total delivery times (time elapsed from pushing the beam-on button till delivery of the last MU).

2.3   Results

All generated plans were clinically acceptable and achieved adequate target coverage (i.e., at least 99.5% of the PTV received 95% of the prescribed dose), and PTV doses of more than 110% did not occur. After plan normalizations to obtain equal median PTV doses (Section 2.2.2), differences between corresponding autoVMAT, CLINICAL, manVMAT and autoIMRT plans in PTV $V_{95\%}$ were within 0.3%. Table 2.2 reports pair-wise comparisons of investigated planning strategies regarding OAR dose delivery. Figure 2.1 shows for all planning strategies, mean differences in OAR plan evaluation parameters compared to autoVMAT.

2.3.1   autoVMAT vs. Clinical plan quality

Automatically generated dual-arc VMAT plans showed superior plan quality compared to the CLINICAL plans (Table 2.2, Figure 2.1). For all OAR evaluation parameters, the mean values were on average lowest for autoVMAT, for 4/7 parameters the difference was statistically significant ($p < .001$, Table 2.2). For 43/44 plans, SB V_{45Gy} was lowest with autoVMAT, and for 35/44, SB D_{mean} was lowest (Table 2.2).

2.3.2   autoVMAT vs. manVMAT plan quality

In 41/44 cases, autoVMAT was lowest with respect to the most important OAR plan parameter, i.e. SB V_{45Gy} ($p < .001$, average reduction 30.3%), while SB V_{15Gy} increased by

(a) all plans
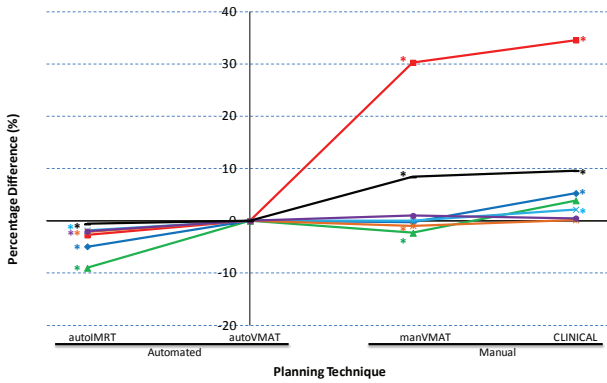
(b) half-full-to-full-bladder PTVs
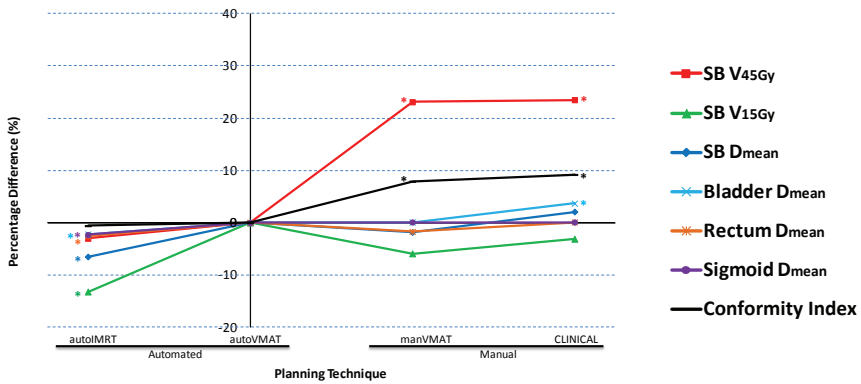
(c) empty-to-half-full-bladder PTVs
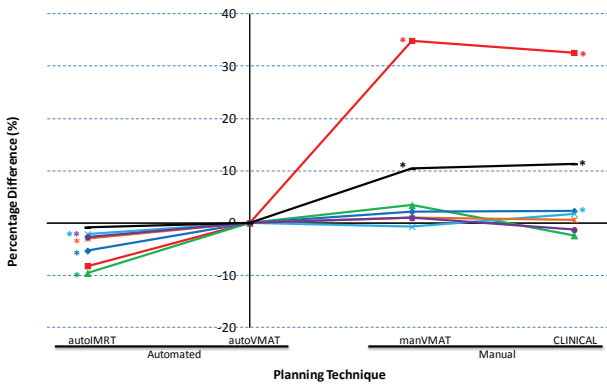
Figure 2.1: Percentage differences in mean OAR plan parameter values for all evaluated planning strategies, compared to autoVMAT. **(a)** for all 44 plans, **(b)** and **(c)** for the 10 patients with large bladder filling induced PTV displacements. Positive values point at an advantage for autoVMAT. Average absolute values for autoVMAT: SB V_{45Gy} = 15.4%, SB V_{15Gy} = 68.7%, SB D_{mean} = 25.5 Gy, Bladder D_{mean} = 43.5 Gy, Rectum D_{mean} = 42.9 Gy, Sigmoid D_{mean} = 42.3 Gy and $CI_{95\%}$ = 1.174. ∗: difference with autoVMAT is statistically significant.

Table 2.2: Pair-wise comparisons of planning strategies.

		autoIMRT		autoVMAT		manVMAT		CLINICAL	
		p-value	*#Plans*	*p-value*	*#Plans*	*p-value*	*#Plans*	*p-value*	*#Plans*
	a			NS	(26 | 18)	< .001	(44 | 00)	< .001	(42 | 02)
	b			< .001	(37 | 07)	NS	(22 | 22)	< .001	(31 | 12)
	c			< .001	(40 | 04)	NS	(26 | 18)	< .001	(41 | 03)
autoIMRT	**d**			< .001	(42 | 02)	.001	(33 | 11)	< .001	(44 | 00)
	e			< .001	(41 | 03)	.026	(29 | 15)	< .001	(37 | 07)
	f			< .001	(29 | 02)	< .001	(26 | 05)	.002	(24 | 07)
	g			.018	(28 | 16)	< .001	(44 | 00)	< .001	(39 | 05)
	a	0.3 ± 1.5 (2.7%)				< .001	(41 | 03)	< .001	(43 | 01)
	b	4.5 ± 4.7 (9.0%)				.011	(14 | 30)	NS	(26 | 18)
	c	1.0 ± 0.7 (4.9%)				NS	(16 | 28)	< .001	(35 | 09)
autoVMAT	**d**	0.8 ± 0.5 (1.8%)				NS	(21 | 23)	< .001	(39 | 05)
	e	0.9 ± 0.6 (2.1%)				.041	(15 | 29)	NS	(26 | 18)
	f	0.8 ± 0.6 (2.0%)				NS	(18 | 13)	NS	(19 | 12)
	g	0.01 ± 0.02 (0.6%)				< .001	(44 | 00)	< .001	(39 | 05)
	a	3.8 ± 2.7 (32.7%)		3.5 ± 3.0 (30.3%)				NS	(26 | 18)
	b	1.5 ± 11.6 (6.4%)		-2.9 ± 10.4 (-2.3%)				.037	(34 | 10)
	c	0.8 ± 2.3 (5.0%)		-0.3 ± 2.1 (-0.2%)				.003	(34 | 10)
manVMAT	**d**	0.8 ± 1.4 (1.8%)		0.0 ± 1.3 (0.0%)				< .001	(36 | 08)
	e	0.5 ± 1.2 (1.2%)		-0.4 ± 1.3 (-0.9%)				.006	(32 | 12)
	f	1.1 ± 0.9 (3.0%)		0.3 ± 1.0 (1.0%)				NS	(14 | 17)
	g	0.11 ± 0.05 (8.9%)		0.10 ± 0.05 (8.4%)				NS	(25 | 17)
	a	4.7 ± 4.2 (37.2%)		4.4 ± 4.1 (34.6%)		0.9 ± 3.7 (6.0%)			
	b	7.3 ± 11.5 (12.9%)		2.8 ± 12.1 (3.9%)		5.8 ± 15.9 (6.3%)			
	c	2.4 ± 1.7 (10.2%)		1.4 ± 1.8 (5.3%)		1.7 ± 3.0 (5.1%)			
CLINICAL	**d**	1.7 ± 1.0 (4.0%)		1.0 ± 1.0 (2.2%)		1.0 ± 1.4 (2.2%)			
	e	1.0 ± 1.2 (2.3%)		0.1 ± 1.3 (0.2%)		0.5 ± 1.3 (1.1%)			
	f	0.8 ± 1.3 (1.5%)		0.0 ± 1.4 (0.5%)		-0.3 ± 1.9 (-1.5%)			
	g	0.13 ± 0.11 (10.2%)		0.12 ± 0.10 (9.7%)		0.02 ± 0.10 (1.3%)			
		Δ*mean ± 1SD (PD)*		Δ*mean ± 1SD (PD)*		Δ*mean ± 1SD (PD)*			

Abbreviations: NS = no statistically significant difference i.e. $p > .05$, SD = Standard Deviation, PD = percentage difference. In each comparison (table cell), going from top to bottom, data refer to (**a**) SB V_{45Gy}, (**b**) SB V_{15Gy}, (**c**) SB D_{mean}, (**d**) Bladder D_{mean}, (**e**) Rectum D_{mean} (**f**) Sigmoid D_{mean}, and (**g**) $CI_{95\%}$. Below the table diagonal, mean point differences with 1 SD, and mean relative percentage differences (between brackets) are shown, while *p*-values and the numbers of plans with lowest OAR dose for one or the other strategy are presented above the diagonal; (*n*|*m*): for *n* plans, the strategy indicated at the vertical axis has lowest OAR dose, while for *m* plans the strategy mentioned at the horizontal axis is superior. Positive Δmean indicates that the strategy along the horizontal axis is superior (reduced OAR dose delivery).

2
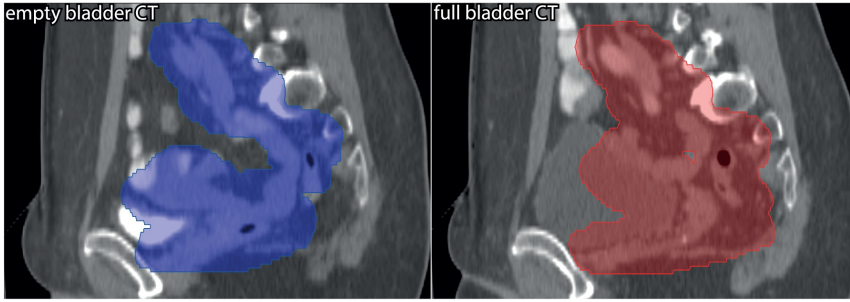
Figure 2.2: For a study patient, sagittal views of the empty-bladder CT-scan with the empty-to-half-full-bladder PTV (blue) projected on it (left), and of the full-bladder CT-scan with the half-full-to-full-bladder PTV (red) on the (right). The higher degree of concavity of the empty-to-half-full-bladder PTV increases the complexity of small bowel sparing in manual planning.

on average 2.3% ($p = .011$), and SB D_{mean} by 0.2% (NS) (Table 2.2, Figure 2.1a). For the other OARs, differences were small ($\leq 1\%$). All 44 autoVMAT plans had lower conformity index, $CI_{95\%}$, than the corresponding manVMAT plan ($p < .001$).

As illustrated in Figure 2.1b and Figure 2.1c, for the large movers, the increase in plan quality with autoVMAT compared to manVMAT was most pronounced for PTVs for empty-to-half-full bladder. With autoVMAT, SB V_{45Gy} reduced on average by 34.9% (SD = 25.9%, range: [-0.5%, 92.7%]) for these PTVs, compared to 23.1% (SD = 15.7%, range: [-0.9%, 55.3%]) for the PTVs belonging to half-full-to-full bladder. For empty-to-half-full bladder PTVs, SB V_{15Gy} was on average reduced by 3.4% using autoVMAT instead of manVMAT, while there was an increase by 5.9% for half-full-to-full bladder PTVs. With autoVMAT, SB D_{mean} was 2.2% lower for empty-to-half-full bladder PTVs, while for half-full-to-full bladder PTVs an average increase of 1.8% was observed. The enhanced positive impact of autoVMAT for empty-to-half-full bladder PTVs is attributed to the higher degree of concavity of these PTVs (see Figure 2.2 for example), making it relatively difficult for the human planner to optimally spare the SB in the manVMAT plan.

2.3.3   autoVMAT vs. autoIMRT plan quality

Compared to autoVMAT, in automatically generated 20-beam equi-angular IMRT plans (autoIMRT) SB V_{45Gy} reduced by 2.7%, while also resulting in a 9.0% reduction in SB V_{15Gy}, and a reduction of 4.9% in SB D_{mean} ($p < .001$, Table 2.2). Differences in bladder, rectal and sigmoid doses were small, but all in favor of autoIMRT (Figure 2.1).

Table 2.3: Measured treatment delivery times for autoVMAT and autoIMRT using the MLCi2 and Agility™ MLCs.

Planning Technique	Patient	MLCi2 Treatment Delivery Time (min)	Agility™ Treatment Delivery Time (min)	Difference (min/%)
autoVMAT	Pt 1	4.5	2.9	1.6 / 35.6
	Pt 2	3.8	2.7	1.1 / 28.9
	Pt 3	4.1	3.2	0.9 / 22.0
	Pt 4	4.4	3.1	1.3 / 29.5
	Pt 5	3.9	2.6	1.3 / 33.3
	Mean	**4.1**	**2.9**	**1.2 / 29.9**
autoIMRT	Pt 1	15.0	9.8	5.2 / 34.7
	Pt 2	17.6	9.5	8.1 / 46.0
	Pt 3	14.0	7.9	6.1 / 43.6
	Pt 4	17.5	10.4	7.1 / 40.6
	Pt 5	14.1	7.8	6.3 / 44.7
	Mean	**15.6**	**9.1**	**6.6 / 41.9**

2.3.4 Planning times

To maximally challenge the developed procedure for automated plan generation (autoVMAT and autoIMRT) regarding plan quality, the expert planner was allowed as much time as needed for the manual generation of high-quality manVMAT plans. Generation of the latter took on average 6 h, including 3 h of hands-on planning time and Monte Carlo dose calculations. For the manual generation of the CLINICAL plans, planning times were not recorded. We estimate that these times were around 2.5 h on average.

For fully automatic plan generation, Erasmus-iCycle optimizations took on average 3 h calculation time, and the automatic re-construction of a clinically deliverable plan in Monaco® took on average 4.5 h for autoVMAT and 2.5 h for autoIMRT. Optimizations for automated plan generation were often performed in evenings or during the night.

2.3.5 Treatment time comparisons

Table 2.3 shows for a sub-group of 5 patients measured treatment times for the generated autoVMAT and autoIMRT plans. Compared to MLCi2, the Agility™ MLC exhibits treatment time reductions of ~30% and 42% for autoVMAT, and autoIMRT, respectively, while preserving the same plan quality. With Agility™, mean treatment times for delivery of autoVMAT and autoIMRT plans were 2.9 and 9.1 min.

2.4 Discussion

We have developed and validated a fully automated VMAT treatment planning solution for plan-library filling in adaptive radiotherapy for cervical cancer patients. Based on the

positive findings presented here, the method is now in routine clinical use. Compared to the previous use of manually generated, 9-beam IMRT plans, the automation rendered hands-on planning obsolete and resulted in higher quality dose distributions. Due to the switch to VMAT, the daily treatment time was also reduced, reducing the negative impact of time-dependent intra-fraction bladder filling and changes in patient setup (Hoogeman *et al.* 2008). We demonstrated that autoVMAT plans were also of higher quality than VMAT plans manually generated by an expert planner in the absence of time pressure. For autoVMAT, SB V_{45Gy} was reduced by 30.3% on average while SB V_{15Gy} increased by 2.3%. This is related to the fact that, according to the applied wish-list in this study, minimization of the dose in the small bowel at the 45 Gy level is given higher planning priority than at the 15 Gy level, and this likely reduces grade $\geq$ 3 acute SB toxicity (Kavanagh *et al.* 2010). The benchmark with automatically generated 20-beam equi-angular IMRT plans (autoIMRT) demonstrated a superior plan quality compared to autoVMAT. However, with the currently available MLCi2 multi-leaf collimators in our department, the treatment delivery time would increase from 4.1 min for autoVMAT to 15.6 min for autoIMRT. This increase led to the departmental decision to always treat with autoVMAT in order to limit the impact of intra-fraction bladder filling and patient motion, and to increase patient throughput. However, the experiments with a much faster Agility™ MLC, performed in a different institution, showed delivery times of 2.9 and 9.1 min for autoVMAT and autoIMRT, respectively. With 9.1 min, autoIMRT might be the treatment modality of choice for selected patients with a high plan quality gain compared to autoVMAT, in case of availability of an Agility™ MLC.

Until recently, plan-libraries for cervical cancer patients always contained a 3D-CRT back-up plan, to be applied in case of an occasional daily dosimetric mismatch between the VMAT plan(s) in the library and the anatomy-of-the-day, or in case of poor CBCT image quality (see Section 2.1). To further exploit the possibility of automated planning, we have now replaced the manually generated 3D-CRT plan by an automatically generated VMAT plan with a generous CTV-PTV margin, to be used as back-up plan. The reduction of workload related to automated planning does also allow extension of the pre-treatment established, patient-specific libraries with more plans, each with a smaller CTV-PTV margin. This extension is currently being investigated, as well as extension based on daily acquired CBCTs.

In this paper, we investigated automated planning for patients that did not have involvement of pelvic or para-aortic lymph nodes. For these patients, only the pelvic nodes (superior field border at the level of the aortic bifurcation) are irradiated to the prescribed dose of 46 Gy, delivered in 23 fractions. For locally-advanced cervical cancer patients with positive pelvic or para-aortic lymph nodes, toxicity can be substantial due to enlarged fields in superior direction and increased (boost) dose, while the dose that can be given is restricted by tolerances of the surrounding OARs such as

kidneys, spinal cord, and small bowel. This can increase the complexity of treatment planning, requiring an enhanced workload effort, as well as skills and experience of the planner. We believe that also for these challenging patients, automated planning holds a strong promise for treatment plan improvement, compared to the current manual planning. However, involved investigations were outside the scope of this paper. To the best of our knowledge, this is the first paper on fully automated, multi-criterial generation of clinically deliverable plans for cervical cancer. It is also the first paper on filling of libraries with automatically generated, clinically deliverable plans, used in plan-of-the-day radiotherapy. In a previous study, Voet *et al.* (2014) developed a system for fully automatic VMAT plan generation for prostate cancer patients, based on an Erasmus-iCycle pre-optimization and final plan generation in Monaco®. In the current study we adapted this method for fully automatic, multi-criterial generation of deliverable plans for cervical cancer patients, both for VMAT and for static gantry IMRT. Recently, Quan *et al.* (2012a,b) developed an automatic inverse planning algorithm for dual-arc VMAT and IMRT for prostate and lung. The initial objectives in their work were defined on the basis of previous planning experience. Their 'autoplans' were consistently better, or no worse, than the manual plans in terms of tumor coverage and normal tissue sparing (Zhang *et al.* 2011). In another study, Wu *et al.* (2013) developed a method using overlap volume histograms and IMRT data to guide and automate head-and-neck VMAT planning. In terms of PTV coverage and OAR sparing, their automatically IMRT-data driven VMAT plans were comparable to clinical IMRT plans. A large and rather mature IMRT plan dataset was required to effectively guide VMAT planning. In most recent studies (Shiraishi *et al.* 2015; Tol *et al.* 2015) a comprehensive knowledge-based method was used for predicting achievable dose-volume histograms (DVHs) to standardize and improve treatment planning. Tol *et al.* (2015) evaluated the method for head-and-neck cancer using a library of different patient plans to make a model that can predict achievable DVHs for defining optimization objectives for new patients. The performance of the proposed method depended highly on the geometrical characteristics of the model-library and the quality of the IMRT plan database. Automated plan generation with Erasmus-iCycle, as described in Breedveld *et al.* (2007, 2009, 2012) and briefly summarized above, does not depend on a database with DVHs of previously treated patients. Instead, the generation of the tumor site specific wish-lists, containing hard constraints and planning objectives with priorities, results in a generalized approach of knowledge-based planning, using previous planning experience as well as new insights for example obtained from clinical studies or literature. The quality of automatically generated new plans does not depend on the quality of the plans in a database. Moreover, in case of an adaptation in the clinical protocol, automatic plan generation does only require an adaptation of the wish-list, and does not have to wait until a new database with manually generated, high-quality plans has been produced first.

2.5 Conclusion

A method for fully automated VMAT treatment plan generation for filling of plan-libraries in adaptive radiotherapy for advanced cervical cancer has been developed and validated. Quality of automatically generated VMAT plans was superior to manually generated plans, while planning workload was fully avoided. Based on the study results, automatic VMAT plan generation has been implemented in clinical routine. Due to the absence of planning workload, extension of plan-libraries to enhance OAR sparing has become feasible.

2.6 Acknowledgments

Fully Automated Treatment Planning of Spinal Metastases – A Comparison to Manual Planning of Volumetric Modulated Arc Therapy for Conventionally Fractionated Irradiation

Daniel Buergy[1], Abdul Wahab M. Sharfo[2], Ben J.M. Heijmen[2], Peter W.J. Voet[3], Sebastiaan Breedveld[2], Frederik Wenz[1], Frank Lohr[1] and Florian Stieler[1]

[1] Department of Radiation Oncology, Universitätsmedizin Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

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

[3] Elekta B.V., De Maas 26, 5684 PL Best, The Netherlands

Abstract

Background: Planning for volumetric-modulated arc therapy (VMAT) may be time consuming and its use is limited by available staff resources. Automated multi-criterial treatment planning can eliminate this bottleneck. We compared automatically created (auto) VMAT plans generated by Erasmus-iCycle to manually created VMAT plans for treatment of spinal metastases.

Methods: Forty-two targets in 32 patients were analyzed. Lungs and kidneys were defined as organs-at-risk (OARs). Twenty-two patients received radiotherapy on kidneys level, 17 on lungs level, and 3 on both levels.

Results: All Erasmus-iCycle plans were clinically acceptable. When compared to manual plans, planning target volume (PTV) coverage of auto plans was significantly better. The homogeneity index did not differ significantly between the groups. Mean dose to OARs was lower in auto plans concerning both kidneys and the left lung. One hot-spot (> 110% of $D_{50\%}$) occurred in the spinal cord of one auto plan (33.2 Gy, $D_{50\%}$: 30 Gy). Treatment time was 7% longer in auto plans.

Conclusions: Erasmus-iCycle plans showed better target coverage and sparing of OARs at the expense of minimally longer treatment times (for which no constraint was set).

3.1 Background

Fractionated or single-dose radiotherapy to spinal metastases of solid tumors is one of the most frequently performed radiation treatments (Tiwana *et al.* 2016). Highly conformal radiation techniques such as stereotactic body radiation therapy (SBRT), intensity-modulated radiation therapy (IMRT) or VMAT are increasingly being used as their application becomes easier and dose distributions confer theoretical advantages (Stieler *et al.* 2011), albeit so far without proof of superior clinical outcome. On the other hand, generation of an optimal IMRT or VMAT plan is often time consuming and it has been shown that staff limitations are correlated with restricted use of new techniques such as IMRT even in the developed world (Mayles 2010; Shikama *et al.* 2014). Automated planning of IMRT and VMAT may reduce the workload which is associated with manual "*trial-and-error*" approaches by around 50% (Voet *et al.* 2014).

Erasmus-iCycle, developed at the Erasmus MC Cancer Institute, is an optimizer for multi-criterial beam profile optimization and beam angle selection for coplanar and non-coplanar IMRT (Breedveld *et al.* 2007, 2009, 2012; Rossi *et al.* 2012; Voet *et al.* 2012, 2014). Other solutions for multi-criterial beam angle optimization have been proposed by Schreibmann *et al.* (2004), and Craft & Monz (2010). Craft & Monz (2010) presented a full multi-beam space Pareto navigation tool. However, both are *a posteriori* methods, i.e., the algorithm generates different sets of beam angles and intensity profiles from which the user selects the plan afterward. The Erasmus-iCycle solution is an *a priori* approach which enables the user to define a site-specific set of criteria '*wish-list*' which may not be violated (constraints) or have to be met as well as possible, or better (objectives). Objectives have assigned priorities to steer the multi-criterial planning towards favorable trade-offs between the various treatment goals. If the priority of an objective is higher, the probability that the corresponding objective is met increases. Hard constraints are always respected in Erasmus-iCycle plans.

Beam directions are selected from candidate directions which can be restricted, e.g., in case of coplanar treatments (Breedveld *et al.* 2007, 2009; Sharfo *et al.* 2015). For fully automated generation of plans that are clinically delivered, Erasmus-iCycle auto plans are automatically re-constructed and segmented in the clinical treatment planning system (TPS) (Sharfo *et al.* 2015). This study intends to validate VMAT plans for treatment of spinal metastases, generated with this approach, and to compare the quality of automatically generated plans with plans that were manually created by experienced treatment planners. Different spine regions pose different optimization problems as a consequence of different OARs being relevant for the treatment. Therefore cervical, thoracic, and lumbar targets were included in the study design.

3.2 Methods

3.2.1 Manual plan generation

Forty-two clinical target volumes (CTVs) in the spinal column of 32 patients were manually delineated for clinical routine treatments. CTVs were anisotropically expanded to PTVs that were the basis for all further analyses. All PTVs were reviewed by expert radiation oncologists (F.W., F.L.).

Manual plans were created by expert treatment planners and also reviewed by radiation oncologists. All manually generated plans were calculated using the Monaco® treatment planning system (Elekta Ltd., Crawly, UK) version 3.2 or later which supports static IMRT, dynamic IMRT, and VMAT. Informed consent was obtained from all patients for anonymized processing of their clinical data. The study was approved by the ethics committee of Heidelberg University, Medical Faculty Mannheim (2016-806R-MA). For this study we assumed lungs or kidneys to be (potentially pre-irradiated) OARs, therefore we included only spinal regions which were at the level of kidneys ($n = 22$), lungs ($n = 17$), or both ($n = 3$), resulting in 90 (45×2) OARs and 42 PTVs in total. Spinal re-irradiation plans were not tested in this study; nevertheless, we contoured the spinal cord to identify any hot-spots in this area within the target volume. Most plans were at the thoracic, thoracolumbar or lumbar level. In 4 cases, soft tissue metastases or rib metastases were included in the PTV. Prescribed doses ranged from 30 to 40 Gy in 10–20 fractions which are commonly applied treatment regimens in patients with sufficient life expectancy and good general condition (Eastern Cooperative Oncology Group, ECOG 0/1; Karnofsky Performance Status, KPS > 80%). Generally, 40 Gy regimens were applied to patients who received postoperative radiotherapy. We did not include patients who had to be treated with spinal cord sparing plans after prior full dose irradiation of the spinal cord. While some patients had in fact received prior radiotherapy, this previous therapy did not require sparing of spinal cord to meet dose constraints (Marks *et al.* 2010b). Details on irradiation sites are summarized in Table 3.1. Further information such as primary tumor site of each patient is provided in Table 3.2.

3.2.2 Automated VMAT plan generation with Erasmus-iCycle (auto)

Target volumes were identical to those used in the manual planning approach. General principles of iCycle plan generation are described above. Further details on auto-planning with the Erasmus-iCycle/Monaco system have been previously provided by Voet *et al.* (2014). A site-specific wish-list for our patient cohort is shown in Table 3.3. Apart from objectives for the PTV, the kidneys and the lungs, shells around the PTV are

3

Table 3.1: Characteristics of spinal irradiation plans ($n = 42$).

		Number of sites
Irradiation site	Cervical spine	None
	Cervical and thoracic spine	2
	Thoracic spine*	15
	Thoracic and lumbar spine	17
	Thoracic and lumbar spine, Sacrum	3
	Lumbar spine	4
	Lumbar spine and Sacrum	1
Primary tumor site	Breast Cancer	15
	Prostate Cancer	11
	Lung Cancer	5
	Non-Small Cell Lung Cancer	3
	Small Cell Lung Cancer	2
	Gastric and Oesophageal Cancer	5
	Multiple Myeloma	3
	Head-and-Neck	1
	Unknown Primary	1
	Urothelial Cell Carcinoma	1
Organs-at-risk	Kidneys	22
	Lungs	17
	Lungs and kidneys in one plan[†]	3
Median dose	40 Gy	9
	30 Gy	33

* Including 3 plans in which a rib metastasis was included into the irradiation field. In one plan a rib metastasis and a soft tissue metastatic site were included into the irradiation field.

[†] In 3 plans, lungs and kidneys were both considered OARs. For statistical considerations, these plans were evaluated in both groups concerning calculations for organs-at-risk. Concerning PTV coverage, keeping the patients in both groups would have weighted these 3 plans double, therefore all statistics were re-calculated under inclusion/exclusion of these patients. This did not change statistical significance in any case (i.e., only numerical changes occurred).

Table 3.2: Characteristics of spinal irradiation plans and primary tumor site.

Analyzed OAR	Plan number	Patient number	PTV location	Primary
Kidneys	1	1	T11-L5	Prostate
	2	2	T12-L5	Prostate
	3	3	T11-L2	Breast
	5	4	T10-L6	Multiple Myeloma
	6	5	T10-S3	Breast
	7	6	T2-L5	Breast
	9	7	L1-S5	Prostate
	11	9	T11-L1	Prostate
	12	10	T12-L2	Breast
	14	11	T11-L2	NSCLC
	15	12	T10-L4	Multiple Myeloma
	16	13	T12-L2	Breast
	17	14	L1-L5	Oesophageal
	19	15	T12-L5	Prostate
	21	16	T12-L3	H&N SCC
	22	17	L1	Breast
	23	18	T12-S1	Prostate
	26	20	T8-S5 / Pelvis	Breast
	28	21	T9-L1	Adenocarcinoma CUP
	30	23	T12-L1	Breast
	31	24	T8-L3	SCLC
	33	25	L3	Gastric Cancer
Kidneys & **Lungs**	10	8	T10-L3	Urothelial
	25	19	T11-L1	NSCLC
	29	22	T7-L1	Breast
Lungs	34	25	T9-T11	Gastric
	35	25	T3-T6	Gastric
	36	26	T5-T10	Breast
	37	27	T3-T6	Prostate
	38	28	T6-T11	Prostate
	39	29	T7-T9	Prostate
	40	30	T3-T8	NSCLC (Adeno)
	41	31	C6-T4	Breast
	42	32	T2-T4	Multiple Myeloma
	4	3	T6-T8	Breast
	8	6	T6-T11	Breast
	13	10	C7-T5	Breast
	18	14	T7-T10	Oesophageal
	20	15	T3-T8	Prostate
	24	18	T2-T9	Prostate
	27	20	T3	Breast
	32	24	T4-T5	SCLC

Table 3.3: Wish-list for automatic plan generation.

Constraints			
Structure	**Type**	**Limit**	
PTV	Max	105% of D^p	
Unspecified Tissues	Max	105% of D^p	

Objectives				
Priority	**Structure**	**Type**	**Goal**	**Parameters**
1	PTV	↓ LTCP	0.5	$D^p = 30(40)$ Gy, $\alpha = 0.75$
2	PTV Shell 5 mm	↓ Max	80% of D^p	
	PTV Shell 20 mm	↓ Max	50% of D^p	
	Skin Ring 30 mm	↓ Max	25% of D^p	
	Right Kidney	↓ Mean	15% of D^p	
	Left Kidney	↓ Mean	15% of D^p	
	Right Lung	↓ Mean	15% of D^p	
	Left Lung	↓ Mean	15% of D^p	
	PTV Shell 40 mm	↓ Max	25% of D^p	
3	PTV Shell 50 mm	↓ Max	15% of D^p	

Abbreviations: D^p = prescribed dose, LTCP = *Logarithmic Tumor Control Probability*, α = cell sensitivity.

used to steer on conformality. Two cases were included which required beam restrictions, both because patients were unable to lift their arms.

3.2.3 Plan quality and statistical analysis

Plan quality was estimated by calculating the dose to 98% ($D_{98\%}$) and to 2% ($D_{2\%}$) of the PTV. Both parameters were used to compute the homogeneity index (HI, (Ohtakara *et al.* 2012)) which was defined as the ratio between the difference of $D_{2\%}$ and $D_{98\%}$ and the median dose ($D_{50\%}$), i.e.,:

$$HI = \frac{D_{2\%} - D_{98\%}}{D_{50\%}}$$

In addition, we calculated the volume receiving 95% of the prescribed dose ($V_{95\%}$ i.e., $V_{28.5Gy}$, or V_{38Gy}, depending if the prescribed dose was 30 Gy or 40 Gy), and based on $V_{95\%}$ a conformity index (CI, reviewed in Ohtakara *et al.* (2012)) was calculated as follows:

$$CI = \frac{V_{95\%}}{PTV}$$

An optimal CI would be considered to be 1; however, values can be above or below 1. To consider both over- and under-dosage for statistical comparison of CI, we calculated the difference to 1 as follows:

$$CI_{diff} = |CI - 1|$$
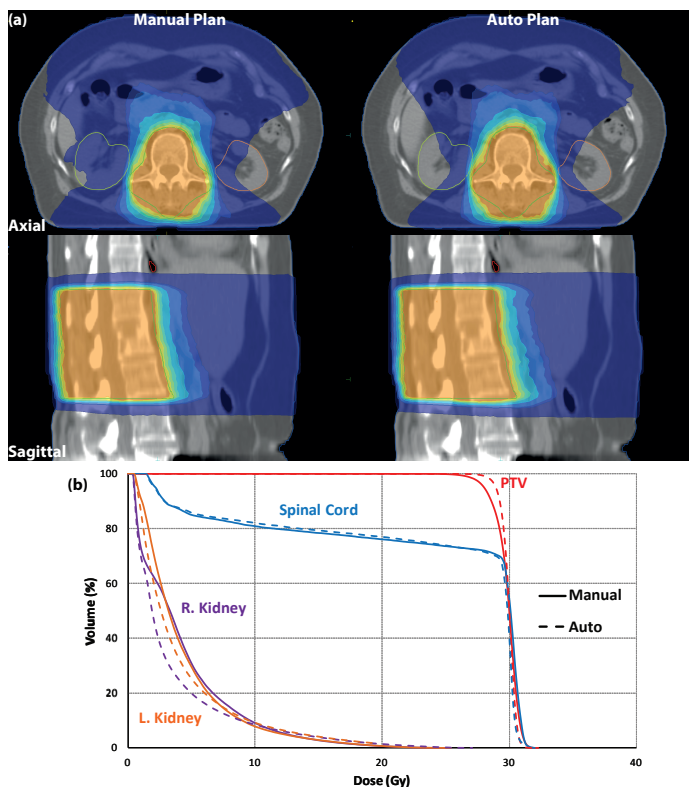
Figure 3.1: a) Dose distribution of an irradiation plan on thoracolumbar level (T12-L2). The patient had metastatic breast cancer, irradiation dose applied with this plan was 30 Gy in 10 fractions. Manual plan shown on the left, and auto plan on the right.
b) Dose-volume histogram, continuous line represents manual plan, and dashed line represents auto plan.

For OARs, we calculated the mean dose (D_{mean}), and in case of the spinal cord, the maximum dose to any hot-spot (D_{max}). All parameters were compared directly between manually generated plans and auto plans. Differences between variables were computed using the two-sample paired Wilcoxon test. Statistical significance was defined as $p <$ .05 (two-sided testing). Clinical acceptability of the plans was evaluated by an expert radiation oncologist (F.L.).

3.3   Results

All automatically created plans were clinically deliverable and acceptable. Figure 3.1 and Figure 3.2 show dose distributions, and dose-volume histogram comparisons for a kidney-level target, and a lung-level target, respectively.

3
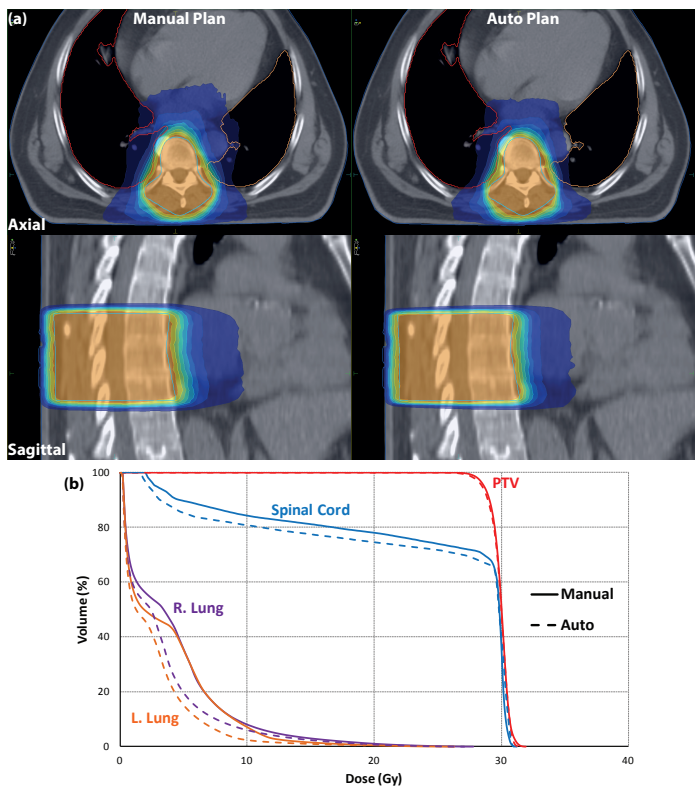
Figure 3.2: a) Dose distribution of an irradiation plan on thoracic level (T7-T9). The patient had metastatic prostate cancer, irradiation dose applied with this plan was 30 Gy in 10 fractions. Manual plan shown on the left, and auto plan on the right.
b) Dose-volume histogram, continuous line represents manual plan, and dashed line represents auto plan.
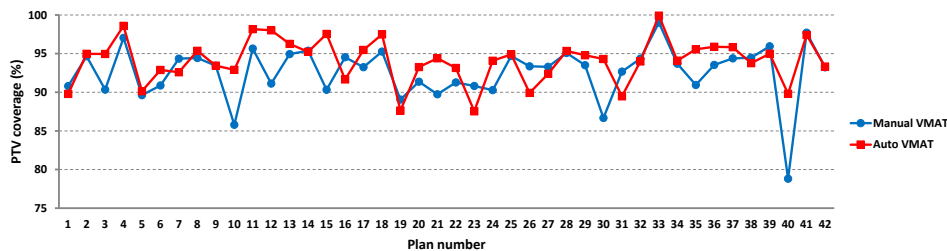
Figure 3.3: Volume receiving 95% of the prescribed dose, i.e., $V_{28.5Gy}$ if the prescribed dose was 30 Gy or V_{38Gy} if it was 40 Gy.

3.3.1 Target volumes

PTV coverage, defined as V_{38Gy} or $V_{28.5Gy}$ was higher in auto plans when compared to manual plans when the whole patient population was analyzed, with an average difference of 1.47% (SD = 3.17%, p = .008, n = 42), see Figure 3.3. When sub-groups were analyzed, the difference was only significant in lung-level plans (p = .004, n = 20) but not at kidney levels (p = .110, n = 25). This pattern occurred irrespective whether 3 patients with both kidneys and lungs as OARs in one plan were added to the lung or to the kidney group or to both. Dose parameters were as follows: $D_{98\%}$ was numerically higher in auto plans when compared to manual plans but this was not statistically significant in the whole population (p = .479), nor in any sub-group (p > .05 for both lung and kidney levels). $D_{2\%}$ was significantly higher in manual plans over the whole patient population (p < .001), however the difference was clinically irrelevant (mean: 33.4 Gy vs. 33.2 Gy; median 31.1 Gy vs. 31.0 Gy). Sub-group analysis showed that $D_{2\%}$ was significantly lower in auto plans at kidney level (p = .003), however the difference was marginal and not clinically relevant. In lung-level plans, $D_{2\%}$ did not differ significantly between auto and manual plans.

HI as defined above was not significantly different between auto and manual plans in the whole group or in any sub-group (p > .05 for all comparisons). Considering the whole population, CI was slightly but statistically significant higher in auto plans than in manual plans: 1.014 ± 0.12 (1 SD) vs. 1.044 ± 0.11. For targets at the kidney level, the mean CIs for manual plans and auto plans were 0.99 ± 0.071 and 1.022 ± 0.067, respectively, and for lung-level targets 1.036 ± 0.157 and 1.065 ± 0.144. The differences between auto and manual plans can be explained by a slightly more pronounced underdosage in manual plans (minimum CI in manual plans: 0.88 [kidney level] and 0.81 [lung level]; minimum CI in auto plans: 0.92, and 0.95, respectively). CI_{diff} did not differ significantly between auto and manual plans.

3.3.2 Organs-at-risk

In the pooled analysis of all plans, mean dose to OARs (kidneys and lungs pooled) was significantly lower in auto plans when compared to manual plans on both the right side (p = .001, n = 45, relative difference -10.9% [mean value auto vs. manual]) and on the left side (p < .001, n = 45, -19.5%). Sub-group analyses in patients with kidneys as OARs, including 3 patients with lungs and kidneys, showed that auto plans were associated with lower dose to both kidneys (p < .001, n = 25, -18.6% right kidney, and p < .001, n = 25, -23.1%, left kidney), see Figure 3.4a for details. In lung-level plans, again including 3 plans in which both lungs and kidneys were included as OARs, only the left lung showed a significantly lower mean dose when auto-planning was used (p = .004, n = 20, -13.7%).
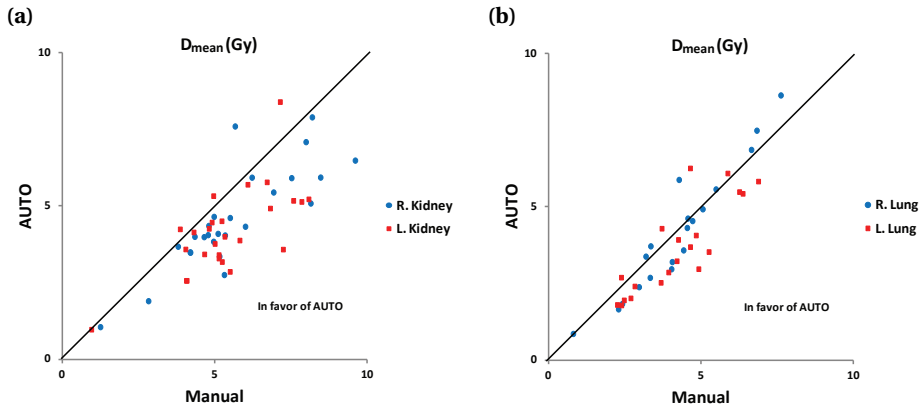
Figure 3.4: Comparison of mean dose to OARs. Each marker represents the mean dose in the manual plan vs. the auto plan to the kidneys (**a**) and to the lungs (**b**). For data points right of the unity line, auto-planning yielded better sparing of OARs.

Mean dose to the right lung did not differ significantly between manual and auto plans ($p = .827$, $n = 20$, -1.3%), see Figure 3.4b for details. There was no significant difference between the maximum doses (D_{max}) to the spinal cord between auto and manual groups in the whole population or in any sub-group. However, one patient in the auto group had a hot-spot in the spinal cord (> 110% of prescribed dose: 33.2 Gy, $D_{50\%}$: 30 Gy).

3.3.3 Estimated treatment time and required monitor units

Auto plans required slightly more (estimated) treatment time in the whole population ($p < .001$, relative difference 7%) and in the sub-groups of kidney-level plans ($p < .001$, relative difference 9.7%), and lung-level plans ($p = .046$, relative difference 3.6%) as no hard constraint was placed on treatment time. This was a consequence of auto plans requiring more monitor units in the whole population and in any sub-group (p-values and relative differences as follows: whole population, $p < .001$, 17.9%; kidneys-level: $p < .001$, 15.4%; lungs-level: $p = .001$, 20.6%).

3.4 Discussion

Though highly conformal radiotherapy is not mandatory in the treatment of painful bone metastases (Lutz *et al.* 2011), intensity modulation with static beams or VMAT allows to deliver a highly focused dose distribution in every clinical situation and may be beneficial for patients with oligometastatic disease (Corbin *et al.* 2013) who require dose escalation or patients who had received prior courses of radiotherapy limiting

doses to OARs in case of re-irradiation (Kirkpatrick *et al.* 2010; Damast *et al.* 2011). Highly conformal radiotherapy is, however, not always applied when appropriate for its perceived resource intensity, both in personnel and planning time (Mayles 2010). Unexpectedly, the reason for not performing advanced radiotherapy techniques despite a clear indication seems to be a shortage of sufficiently trained personnel rather than machine shortage (Mayles 2010; AlDuhaiby *et al.* 2012; Shikama *et al.* 2014). AlDuhaiby *et al.* (2012) described in a Canadian survey that limitations in CT simulator or linear accelerator configuration impaired IMRT implementation in only 10% of cases. The most relevant factors hindering IMRT implementation was the need to train existent treatment planners (50%) as well as the need to hire more planners (30%). Personal shortage is even more relevant in VMAT settings as these are more time consuming as compared to fixed field IMRT (Oliver *et al.* 2009); "increased planning time required for generating the VMAT plans" has been described by Rao *et al.* (2010) as the main disadvantage of VMAT. Therefore, there is a need for tools that reduce planning time and ease the process of developing a radiotherapy plan without compromising plan quality.

Various auto-planning strategies are currently being developed (Breedveld *et al.* 2012), these include approaches for beam angle selection (e.g., (Potrebko *et al.* 2008)), and integrated beam weight optimization algorithms. The latter can be based on global optimization (Lee *et al.* 2006), or on sequential beam selection (de Pooter *et al.* 2008). As discussed above, multi-criterial optimization systems have been proposed (Schreibmann *et al.* 2004; Craft & Monz 2010), that enable the user to select the plan out of different sets of beam angles and intensity profiles after optimization (*a posteriori* setting). Template based auto-planning solutions are already commercially available and have been applied to various target paradigms such as head-and-neck (Hazell *et al.* 2016), esophageal targets (Fogliata *et al.* 2015a) and breast cancer (Fogliata *et al.* 2015b).

Erasmus-iCycle plans, which provide a template-free approach have been successfully applied to head-and-neck (Voet *et al.* 2013b), prostate (Voet *et al.* 2014), and cervical cancer (Sharfo *et al.* 2015) target volumes to create clinically deliverable plans.

Our results show that fully automated VMAT treatment planning with the Erasmus-iCycle/Monaco system for spinal metastases was non-inferior to conventional treatment planning by expert dosimetrists or medical physicists. Automatically generated plans with this system outperformed manual plans in terms of sparing of OARs in 3 of 4 predefined organs and also PTV coverage was favorable when compared to manual plans. We observed one hot-spot in an automatically generated plan in the spinal cord (33.2 Gy, $D_{50\%}$: 30 Gy). This problem can be addressed by modifying the wish-list to include a maximum dose cost-function on the spinal cord. In our analysis, however, we applied a quadratic overdose cost-function which lead to maximum doses that were comparable to maximum doses observed in manual plans in all but this patient

and allowed for more degrees of freedom regarding dose reduction to other OAR. If the maximum dose has to be strictly controlled, e.g., in case of re-irradiation, a maximum dose cost-function should be applied. Otherwise, the situation would be evaluated during the manual approval process which might prompt occasional re-planning with a maximum dose cost-function. Monitor units and estimated treatment time were higher in auto plans; however, this difference was not considered clinically relevant in the treated patient population and resulted in a negligible prolongation of treatment times. The difference in treatment times was expected as the aim of this study was to generate high-quality plans with clinically acceptable delivery times. Monitor units and delivery times could be reduced in the auto-plans if required. In contrast to a study in a more complex setting such as head-and-neck cancer (Voet *et al.* 2013b), the observed dosimetric advantages of Erasmus-iCycle for treatment of spinal metastases were on average of low clinical relevance although advantages across the population were also observed.

External audits have shown that experienced treatment centers may yield superior IMRT plans (Nelms *et al.* 2012). Therefore, as postulated most recently by Fogliata *et al.* (2015a), quality improvements of automated planning vs. manual plan generation may even be more pronounced in target paradigms that are not frequently treated in a particular center.

3.5 Conclusion

Our data add to the growing evidence (Voet *et al.* 2013b, 2014; Sharfo *et al.* 2015) that automated treatment planning might be an alternative to manual planning, reducing the workload of medical physicists and dosimetrists while maintaining or improving plan quality.

Automated VMAT Planning for Postoperative Adjuvant Treatment of Advanced Gastric Cancer

Abdul Wahab M. Sharfo[1]**, Florian Stieler**[2]**, Oskar Kupfer**[2]**, Ben J.M. Heijmen**[1]**, Maarten L.P. Dirkx**[1]**, Sebastiaan Breedveld**[1]**, Frederik Wenz**[2]**, Frank Lohr**[3]**, Judit Boda-Heggemann**[2] **and Daniel Buergy**[2]

[1]Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands
[2]Department of Radiation Oncology, Universitätsmedizin Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
[3]Unita Operativa di Radioterapia, Dipartimento di Oncologia, Az. Ospedaliero-Universitaria di Modena, Modena, Italy

Abstract

Background: Postoperative/adjuvant radiotherapy of advanced gastric cancer involves a large planning target volume (PTV) with multi-concave shapes which presents a challenge for volumetric-modulated arc therapy (VMAT) planning. This study investigates the advantages of automated VMAT planning for this site compared to manual VMAT planning by expert planners.

Methods: For 20 gastric cancer patients in the postoperative/adjuvant setting, dual-arc VMAT plans were generated using fully automated multi-criterial treatment planning (autoVMAT), and compared to manually generated VMAT plans (manVMAT). Both automated and manual plans were created to deliver a median dose of 45 Gy to the PTV using identical planning and segmentation parameters. Plans were evaluated by two expert radiation oncologists for clinical acceptability. AutoVMAT and manVMAT plans were also compared based on dose-volume histogram (DVH) and predicted normal tissue complication probability (NTCP) analysis.

Results: Both manVMAT and autoVMAT plans were considered clinically acceptable. Target coverage was similar (manVMAT: 96.6 ± 1.6%, autoVMAT: 97.4 ± 1.0%, $p = .085$). With autoVMAT, median kidney dose was reduced on average by > 25%; (for left kidney from 11.3 ± 2.1 Gy to 8.9 ± 3.5 Gy ($p = .002$); for right kidney from 9.2 ± 2.2 Gy to 6.1 ± 1.3 Gy ($p < .001$)). Median dose to the liver was lower as well (18.8 ± 2.3 Gy vs. 17.1 ± 3.6 Gy, $p = .048$). In addition, D_{max} of the spinal cord was significantly reduced (38.3 ± 3.7 Gy vs. 31.6 ± 2.6 Gy, $p < .001$). Substantial improvements in dose conformity and integral-dose were achieved with autoVMAT plans (4.2% and 9.1%, respectively; $p < .001$). Due to the better OAR sparing in the autoVMAT plans compared to manVMAT plans, the predicted NTCPs for the left and right kidney and the liver-PTV were significantly reduced by 11.3%, 12.8%, 7%, respectively ($p \leq .001$). Delivery time and total number of monitor units were increased in autoVMAT plans (from 168 ± 19 s to 207 ± 26 s, $p = .006$) and (from 781 ± 168 MU to 1001 ± 134 MU, $p = .003$), respectively.

Conclusions: For postoperative/adjuvant radiotherapy of advanced gastric cancer, involving a complex target shape, automated VMAT planning is feasible and can

substantially reduce the dose to the kidneys and the liver without compromising the target dose delivery.

4.1   Background

In the Intergroup 0116 (INT-0116) study, adjuvant chemoradiotherapy (CRT) improved overall survival and progression-free survival in patients with gastric cancer compared to surgery alone (Macdonald *et al.* 2001; Smalley *et al.* 2012). The benefit of combined modality treatment was later confirmed in Asian studies with D2-resected patients (Lee *et al.* 2012; Zhu *et al.* 2012; Park *et al.* 2015). European trials in gastroesophageal junction cancers showed survival benefits compared to surgery with neoadjuvant CRT (van Hagen *et al.* 2012) or perioperative chemotherapy without radiotherapy (Cunningham *et al.* 2006; Ychou *et al.* 2011). Randomized trials comparing adjuvant CRT to (perioperative) chemotherapy indicated clinical benefits for CRT in sub-groups, but failed in showing a consistent benefit in European and Asian trials (Buergy *et al.* 2012; Zhu *et al.* 2012; Park *et al.* 2015; Verheij *et al.* 2016). Therefore, the role of adjuvant CRT in gastric carcinoma compared to other approaches, such as perioperative chemotherapy, is not yet clearly defined (Buergy *et al.* 2012). The use of modern radiotherapy approaches might improve the risk-/benefit ratio in favor of radiotherapy in gastric carcinoma. Recent data showed the feasibility of using intensity-modulated radiation therapy (IMRT) or volumetric-modulated arc therapy (VMAT) for this purpose (Dikken *et al.* 2011; Zhu *et al.* 2012; Boda-Heggemann *et al.* 2013; Verheij *et al.* 2016). Postoperative radiotherapy of advanced gastric cancer involves a large target volume with multi-concave shapes which presents a challenge for VMAT planning with a risk of protocol deviations. In a similar setting, radiotherapy for esophageal carcinoma indicated that the experience of the treatment planner may largely affect plan quality, as shown by large differences in plan quality among planners within one institute (Berry *et al.* 2016b). In such a setting, automated treatment planning might have several advantages, including facilitated central review in clinical trials as well as improvement of less experienced treatment planner's performance.

Over the years, in-house developed as well as commercial algorithms have attempted to automate the trial-and-error process in order to create optimal plans, reduce user variability and improve the quality and efficiency of the resulting plans. Knowledge-based planning uses a model-library of previously generated plans to predict new treatment plan parameters (Appenzoller *et al.* 2012; Yuan *et al.* 2012), while multi-criterial optimization generates a set of *Pareto-optimal* plans (Breedveld *et al.* 2007; Monz *et al.* 2008). Erasmus-iCycle is an optimizer for fully automated multi-criterial beam profile optimization and beam angle selection for coplanar and non-coplanar IMRT, developed at the Erasmus MC Cancer Institute (Breedveld *et al.* 2012; Voet *et al.* 2013b, 2014; Sharfo *et al.* 2016; Della Gala *et al.* 2017). In combination with the Monaco® treatment planning system (TPS), Erasmus-iCycle is currently used

in clinical practice for IMRT and VMAT plan generation for prostate, head-and-neck, cervical and lung cancer patients (Voet *et al.* 2013b, 2014; Sharfo *et al.* 2016; Della Gala *et al.* 2017). Several studies on postoperative gastric cancer patients have dosimetrically evaluated different radiotherapy techniques (Li *et al.* 2014; Zhang *et al.* 2015; Mondlane *et al.* 2017; Onal *et al.* 2018). To our knowledge, no study has been published considering the feasibility and advantages of automated treatment planning for this treatment site. In this study, we investigated to what extent automated treatment planning using Erasmus-iCycle results in improved VMAT plan quality for advanced gastric cancer patients compared to plan generation by an expert planner.

4.2 Methods

4.2.1 Patients

A total of 20 patients with advanced gastric cancer who received radiotherapy treatment were included in the study. Clinical details are shown in Table 4.1. Written informed consent was obtained from all patients for anonymized usage of treatment planning data. Our study protocol for retrospective evaluation of automated planning using Erasmus-iCycle was approved by the ethics committee of Heidelberg University, Medical Faculty Mannheim (2016-806R-MA).

4.2.2 Treatment plan generation

All manually and automatically generated dual-arc VMAT plans (manVMAT and autoVMAT, respectively) were prescribed to deliver a median dose of 45 Gy to the target in 25 fractions. The primary planning objective was achieving adequate PTV coverage while maximally sparing the organs-at-risk (OARs). All plans were generated for delivery at a Versa HD™ linear accelerator (Elekta AB, Stockholm, Sweden), equipped with an Agility™ multi-leaf collimator and a photon beam energy of 10 MV.

All manVMAT plans were generated by expert treatment planners with the Monaco® TPS, version 5.11 (Elekta AB, Stockholm, Sweden) using template-based optimization cost-functions. The employed optimization template is set according to the current clinical guidelines used at the Medical University of Mannheim. For each patient, the optimization cost-functions' parameters were iteratively tweaked to improve the dose distribution and achieve optimal target coverage while respecting OAR constraints. As explained below, also the final autoVMAT plans were generated with Monaco®, for which we used the same software version. All Monte Carlo dose calculations in Monaco® were performed using a 1% dose variance, and a dose grid resolution of 3 mm. A maximum of

Table 4.1: Clinical details of patients with gastric carcinoma.

		Number of patients
Sex	Female	9
	Male	11
Age at diagnosis	< 50	4
	50-55	5
	56-60	5
	61-65	2
	> 65	4
Staging		
T	cT1	2
	cT2a	5
	cT2b	4
	cT3	8
	cT4	1
N	N0	2
	N1	13
	$\geqslant$ N2	5
M	M0	19
	M1	0
	Mx	1
Resection margins	R0	16
	R1	1
	Unknown	3
Tumor site	Cardia	1
	Antrum	2
	Corpus	3
	Lesser curvature	2
	Greater curvature	1
	Diffuse	5
	Unknown	6

Table 4.2: Applied wish-list for automatic VMAT plan generation for gastric cancer patients.

Constraints

	Structure	Type	Limit
	PTV	Max	105% of D^p
	PTV Shell 39 mm	Max	50% of D^p
	Patient	Max	105% of D^p

Objectives

Priority	Structure	Type	Goal	Parameters
1	PTV	↓ LTCP	0.4	$D^p = 45$ Gy, $\alpha = 4$
2	PTV Shell 3 mm	↓ Max	90% of D^p	
3	Left and Right Kidney	↓ Mean	8 Gy	
4	Liver	↓ Mean	15 Gy	
5	Heart	↓ EUD	15 Gy	$k = 6$
6	Spinal Cord	↓ EUD	25 Gy	$k = 12$
7	PTV Shell 18 mm	↓ Max	40% of D^p	
	Skin Ring 21 mm	↓ Max	25% of D^p	
8	Left and Right Kidney	↓ Dose-Volume	25%	12 Gy
	Liver	↓ Dose-Volume	30%	24 Gy
9	Left and Right Lung	↓ Dose-Volume	50%	20 Gy

Abbreviations: D^p = prescribed dose, LTCP = *Logarithmic Tumor Control Probability*,
α = cell sensitivity parameter to achieve adequate target coverage, EUD = *Equivalent Uniform Dose, k*
= volume-effect parameter.

140 control points per arc was allowed. Also the other segmentation settings were kept identical for manual and automated plan generation.

4.2.3 Automated multi-criterial VMAT plan generation with Erasmus-iCycle/Monaco

The in-house developed Erasmus-iCycle/Monaco platform for fully automated multi-criteria plan generation (Breedveld *et al.* 2012; Voet *et al.* 2013b, 2014; Sharfo *et al.* 2016; Della Gala *et al.* 2017) was configured to generate clinically deliverable VMAT plans for gastric cancer. For each of the study patients, Erasmus-iCycle was used to first automatically generate a *Pareto-optimal* plan with clinically favorable trade-offs between treatment objectives using a '*wish-list*' developed for gastric cancer (as detailed in Table 4.2). Based on this Erasmus-iCycle plan, a patient-specific Monaco® template was then created fully automatically, to be used in the Monaco® TPS for automated generation of a deliverable dual-arc VMAT plan that mimicked the initial Erasmus-iCycle plan. The applied wish-list contains constraints to be strictly fulfilled, and clinical plan objectives with ascribed priorities to be met as closely as possible or superseded (see Table 4.2 for details). Three constraints were used to control the maximum dose in the PTV and the patient (i.e., the outline of the patient's external surface including the PTV, delineated OARs and the unspecified tissues), as well as the dose conformity outside the PTV. In order to achieve a homogeneous and adequate PTV dose coverage, the highest priority objective was given to the PTV using the logarithmic tumor control probability function (LTCP) (Alber & Reemtsen 2007), followed by a shell around the

PTV to ensure a steep gradient outside the PTV (priority 2). In line with the clinical practice, minimizing the mean dose in the kidneys was the OAR objective with highest priority (priority 3), followed by the mean dose in the liver (priority 4). Subsequently, equivalent uniform dose (EUD) objectives with volume-effect parameters ($k = 6$ and 12) (Niemierko 1997), focusing on reduction of the midrange and high dose in the heart and spinal cord were used (priorities 5 and 6), respectively. To control the dose conformity and entrance dose, a shell at 18 mm from the PTV, as well as a skin ring of 21 mm wide from the body contour towards the patient's internal were defined (priority 7). Additionally, dose-volume objectives for the kidneys, liver, and the lungs were used with lower priorities.

4.2.4   Plan evaluation and comparison

All plans were evaluated by expert radiation oncologists (F.L., J.B.H., and D.B.) for clinical acceptability. For fair dosimetric plan comparisons, all manVMAT and autoVMAT plans were first normalized to obtain equal median PTV dose (i.e., ($D_{50\%}$) = 45 Gy). In accordance with the International Commission on Radiation Units and Measurements Report No. 83, near-minimum and near-maximum doses ($D_{98\%}$ and $D_{2\%}$, respectively) in the PTV were evaluated, from which the homogeneity index ($HI = (D_{2\%} - D_{98\%})/D_{50\%}$) was computed. Additionally, dose conformity was estimated by calculating the conformity index ($CI = (TV_{RI})^2/(TV * V_{RI})$), i.e., ratio of the target volume covered by the reference isodose level (TV_{RI}) to the target volume (TV) and volume of the reference isodose (V_{RI}). Quantitative analyses of OAR doses included D_{mean}, $D_{30\%}$, $D_{50\%}$, $D_{60\%}$ in the kidneys and the liver-PTV, mean and maximum doses (D_{mean}, and D_{max}) in the spinal cord and the heart, volumes of the kidneys receiving more than 12 Gy (V_{12Gy}) and 20 Gy (V_{20Gy}) (Dawson *et al.* 2010), volumes of the liver-PTV receiving more than 24 Gy (V_{24Gy}) and 30 Gy (V_{30Gy}), and D_{mean} in the lungs and volumes of the lungs receiving more than 5 Gy (V_{5Gy}) and 20 Gy (V_{20Gy}) (Marks *et al.* 2010a). In addition, integral-patient-doses were evaluated by assessing patient D_{mean} and volumes receiving V_{5Gy}, $V_{11.25Gy}$ and $V_{22.5Gy}$. Estimated treatment delivery time and total number of monitor units (MUs) for each plan were also quantified.

From the DVHs, normal tissue complication probabilities (NTCPs) for the kidneys and the liver were estimated using the Lyman-Burman-Kutcher (LKB) NTCP model for late effects, considering tolerance dose $TD_{50/5}$ = 12 Gy, n = 0.70, m = 0.26 for the kidneys (Dawson *et al.* 2010), and $TD_{50/5}$ = 30 Gy, n = 0.32, m = 0.15 for liver failure (Dawson *et al.* 2002), where $TD_{50/5}$ refers to the dose to the whole organ which lead to complication in 50% of the population at 5 years, m relates to the steepness of the dose-response curve, and n represents the volume-effect in the LKB model. For this purpose, all plans were first normalized to 1.5 Gy per fraction using α/β = 2.5 Gy.

Differences in dosimetric parameters between manVMAT and autoVMAT plans were analyzed using SPSS® software version 21 (SPSS®, Inc., Chicago, USA) and presented as the mean ± 1 standard deviation. Paired two-sided Wilcoxon signed-rank tests were performed to assess statistical significance of observed differences, considering $p < .05$ statistically significant.

4.3 Results

4.3.1 Target volume dosimetric evaluations

All automatically and manually generated VMAT plans were clinically acceptable and achieved adequate target coverage. Clinical acceptability of the plans was evaluated by radiation oncologists with experience in gastric cancer treatment (J.B.H., F.L. and D.B.). Differences between the autoVMAT and manVMAT plans in $V_{95\%}$ or $D_{98\%}$ were not statistically significant (Table 4.3). However, autoVMAT plans exhibited a significantly lower near-maximum dose in the PTV, resulting in a better target dose homogeneity (0.09 ± 0.01 vs. 0.10 ± 0.02 ($p = .003$)). For autoVMAT, the dose conformity was significantly improved as well (0.91 ± 0.02 vs. 0.88 ± 0.03 ($p < .001$)). As an example, Figure 4.1a shows the resulting dose distributions from the autoVMAT and manVMAT plans for patient 8. As is evident from this figure and the corresponding dose-volume histograms in Figure 4.1b, autoVMAT resulted in favorable dose conformity and better OAR sparing.

4.3.2 Organs-at-risk dosimetric evaluations

Figure 4.2 shows the observed absolute differences in dosimetric parameters for each of the study patients. Overall, autoVMAT plans had more favorable dose distributions, resulting in reduced dose delivery to the kidneys, liver, spinal cord, heart and lungs, without deteriorating the PTV dose coverage (Table 4.3 and Figure 4.2). Median doses ($D_{50\%}$) to the left kidney, right kidney, and liver-PTV were significantly reduced by 28% (from 11.3 ± 2.1 Gy to 8.9 ± 3.5 Gy, $p = .002$), 39% (from 9.2 ± 2.2 Gy to 6.1 ± 1.3 Gy, $p < .001$), and 11% (from 18.8 ± 2.3 Gy to 17.1 ± 3.6 Gy, $p = .048$), respectively. The V_{20Gy} for the left and right kidney with autoVMAT were lower than those for the manVMAT plans (Table 4.3). In addition, the V_{30Gy} for the liver-PTV was significantly decreased with autoVMAT plans by 36% (from 17.7 ± 6.2% to 12.9 ± 6.0%, $p = .001$). The maximum dose in the spinal cord was on average reduced by 6.7 ± 3.2 Gy ($p < .001$) with autoVMAT. Furthermore, the integral-dose in the patient and the dose conformity were also significantly better with autoVMAT (Table 4.3).

(a)

(b)

4
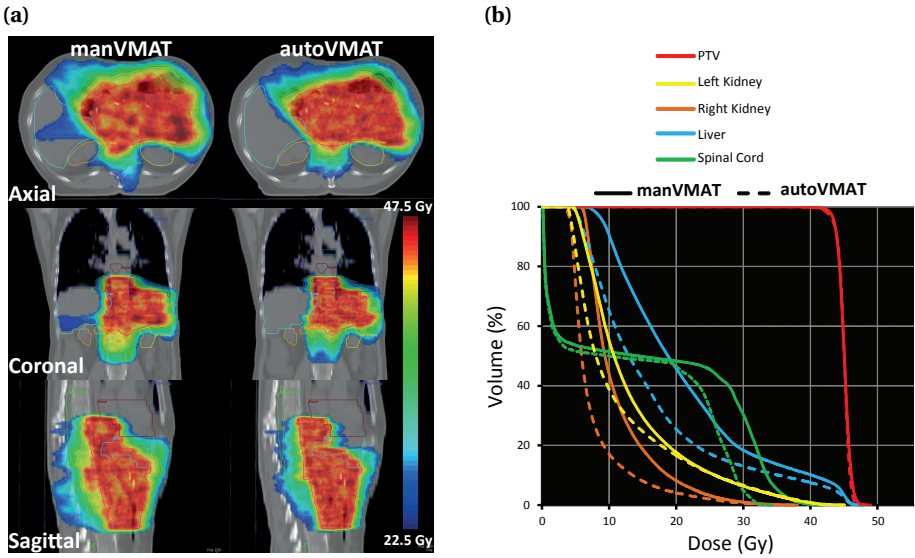
Figure 4.1: a) Comparison of dose distributions for the manVMAT (left) and autoVMAT plans (right) for patient 8 on the axial, coronal and sagittal planes, **b)** Dose-volume histograms for the manVMAT (solid lines) and the autoVMAT (dashed lines) plans of this patient.
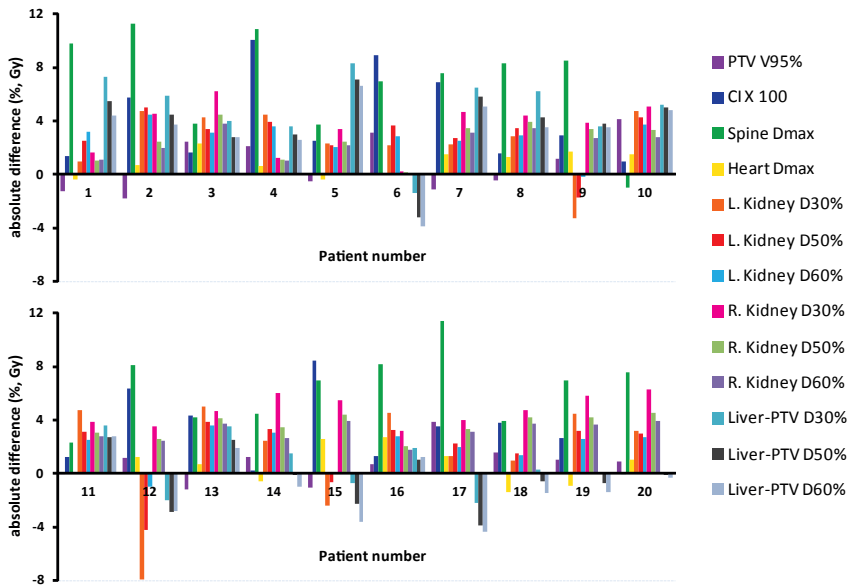
Figure 4.2: Differences in dosimetric plan parameters between autoVMAT and manVMAT plans for each of the 20 study patients. Positive values are in favor of the autoVMAT plans.

Table 4.3: Comparison of dosimetric parameters between autoVMAT and manVMAT plans. Population mean values for the 20 study patients and corresponding standard deviations are reported.

Structure	Parameter	autoVMAT	manVMAT – autoVMAT	
		Mean ± SD	Mean diff. ± SD	*p*-value
PTV	$V_{95\%}$ (%)	97.4 ± 1.0	-0.8 ± 1.7	.085
	$D_{98\%}$ (Gy)	42.5 ± 0.5	-0.2 ± 0.7	.14
	$D_{2\%}$ (Gy)	46.3 ± 0.1	0.5 ± 0.2	**< .001**
	HI	0.09 ± 0.01	0.02 ± 0.02	**.003**
	CI	0.91 ± 0.02	-0.04 ± 0.03	**< .001**
Left Kidney	D_{mean} (Gy)	12.7 ± 4.3	1.8 ± 1.6	**.001**
	$D_{30\%}$ (Gy)	14.9 ± 8.1	2.1 ± 3.3	**.013**
	$D_{50\%}$ (Gy)	8.9 ± 3.5	2.4 ± 2.2	**.002**
	$D_{60\%}$ (Gy)	7.3 ± 1.9	2.4 ± 1.4	**< .001**
	V_{12Gy} (%)	32.2 ± 14.0	11.7 ± 7.7	**< .001**
	V_{20Gy} (%)	19.2 ± 13.2	1.0 ± 3.9	.117
	$NTCP$ (%)	45.1 ± 35.8	11.3 ± 10.9	**.001**
Right Kidney	D_{mean} (Gy)	7.5 ± 1.7	3.0 ± 1.2	**< .001**
	$D_{30\%}$ (Gy)	7.8 ± 1.7	4.1 ± 1.6	**< .001**
	$D_{50\%}$ (Gy)	6.1 ± 1.3	3.1 ± 1.3	**< .001**
	$D_{60\%}$ (Gy)	5.5 ± 1.2	2.7 ± 1.1	**< .001**
	V_{12Gy} (%)	12.4 ± 4.8	18.9 ± 8.4	**< .001**
	V_{20Gy} (%)	3.8 ± 2.8	3.0 ± 2.8	**.001**
	$NTCP$ (%)	4.2 ± 3.8	12.8 ± 7.2	**< .001**
Liver-PTV	D_{mean} (Gy)	18.6 ± 3.3	1.8 ± 2.8	**.02**
	$D_{30\%}$ (Gy)	21.9 ± 4.1	2.8 ± 3.2	**.003**
	$D_{50\%}$ (Gy)	17.1 ± 3.6	1.7 ± 3.3	**.048**
	$D_{60\%}$ (Gy)	15.1 ± 3.4	1.2 ± 3.3	.14
	V_{24Gy} (%)	25.4 ± 11.6	7.1 ± 8.8	**.004**
	V_{30Gy} (%)	12.9 ± 6.0	4.7 ± 4.6	**.001**
	$NTCP$ (%)	31.1 ± 23.2	7.0 ± 7.0	**.001**
Heart	D_{mean} (Gy)	12.7 ± 4.7	0.7 ± 1.8	.117
	D_{max} (Gy)	46.8 ± 0.6	0.8 ± 1.2	**.009**
Spinal cord	D_{mean} (Gy)	13.9 ± 2.6	2.9 ± 1.5	**< .001**
	D_{max} (Gy)	31.6 ± 2.6	6.7 ± 3.2	**< .001**
Left Lung	D_{mean} (Gy)	7.3 ± 3.2	0.6 ± 0.3	**< .001**
	V_{5Gy} (%)	32.7 ± 15.6	1.4 ± 1.0	**< .001**
	V_{20Gy} (%)	14.0 ± 7.2	2.2 ± 1.8	**< .001**
Right Lung	D_{mean} (Gy)	5.4 ± 3.3	0.3 ± 0.5	**.013**
	V_{5Gy} (%)	31.2 ± 16.8	1.1 ± 2.1	**.01**
	V_{20Gy} (%)	7.0 ± 7.6	0.8 ± 1.8	.078
Patient	D_{mean} (Gy)	9.4 ± 1.9	0.5 ± 0.3	**< .001**
	V_{5Gy} (%)	40.2 ± 8.9	1.6 ± 0.8	**< .001**
	$V_{11.25Gy}$ (%)	30.0 ± 6.6	1.8 ± 1.7	**.001**
	$V_{22.5Gy}$ (%)	16.2 ± 3.3	1.5 ± 1.1	**< .001**

Abbreviations: HI = *Homogeneity Index*, CI = *Conformity Index*, NTCP = *Normal Tissue Complication Probability*.
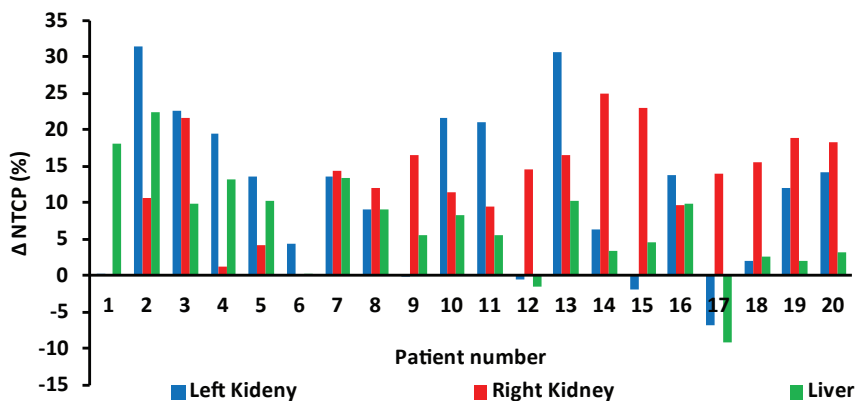
Figure 4.3: Differences in the predicted normal tissue complication probabilities (NTCP) between autoVMAT and manVMAT plans for the 20 study patients. Positive values are in favor of the autoVMAT plans.

4.3.3 NTCPs evaluations

The published sets of parameters used for calculations of NTCPs for the kidneys and liver-PTV resulted in overall lower probability of late complications with autoVMAT compared to manVMAT plans. In 16/20 patients, the predicted NTCPs were lowest with autoVMAT for the left kidney, whereas all 20 patients had lower predicted NTCPs with autoVMAT for the right kidney (Figure 4.3). For 17/20 patients, the NTCP predictions for liver-PTV were lower for the autoVMAT plans. The average NTCP values resulting from autoVMAT, given in Table 4.3, were significantly reduced by 11.3%, 12.8% and 7% for the left and right kidney and the liver-PTV, respectively.

4.3.4 Planning and treatment delivery times

All manVMAT plans in this study were generated using template-based manual planning. Hands-on tweaking time of the optimization cost-functions was on average 30 min (range 5-60 min). The optimization and dose calculation of autoVMAT plans in Erasmus-iCycle/Monaco were fully automated and therefore did not require any hands-on time.

Compared to manVMAT plans a significantly more MUs were required for delivery of the autoVMAT plans (1001 ± 134 MU vs. 781 ± 168 MU, p = .003), resulting in a prolonged (estimated) treatment delivery time (207 ± 26 s vs. 168 ± 19 s, p = .006).

4.4   Discussion

Gastric cancer is the third leading cause of cancer death in both sexes with yearly 723,000 deaths worldwide (Ferlay *et al.* 2015). Protocol deviations can reduce clinical efficacy of radiotherapy in complex treatment geometries, like gastric carcinoma. In the centralized review of the INT-0116 study, minor or major protocol deviations were observed in 35% of patients (Smalley *et al.* 2012). Even after correction of errors, major protocol deviations were still found in 6.5% of the plans that were actually irradiated (Macdonald *et al.* 2001; Smalley *et al.* 2012). The clinical impact of protocol deviations in INT-0116 has not been reported. However, in other complex geometries it was shown that protocol violations during radiation therapy were correlated with reduced overall survival (Wong *et al.* 2017). Modern radiotherapy approaches such as IMRT and VMAT have several dosimetric advantages but they are associated with increased complexity (Lohr *et al.* 2003; Wieland *et al.* 2004; Buergy *et al.* 2012; Haneder *et al.* 2012). Very high protocol deviations are therefore possible; for example, RTOG 0529 showed 81% protocol deviations at first plan review (Kachnic *et al.* 2013). Albeit in small numbers, RTOG 0022 showed local failure rates of 50% (2/4) vs. 6% (3/49) in oropharyngeal cancer patients treated with and without protocol deviations, respectively (Eisbruch *et al.* 2010). Protocol deviations may, amongst other things, be caused by inaccurate delineation or sub-optimal treatment planning. For gastric cancer, the percentage of these deviations is unknown. In a recent report on RTOG 0933 by Gondi *et al.* (2015), unacceptable radiotherapy protocol deviations were observed in 25% of cases at a rapid review process: 11/21 cases had contouring deviations, 5/21 cases had unacceptable planning deviations and 5/21 had unacceptable deviations of contouring and planning. Also in a recent study of Habraken *et al.* (2017) on hepatocellular carcinoma patients, it was demonstrated that pre-treatment plan review was important to reduce unacceptable protocol deviations. Additionally, they demonstrated that automated treatment planning could be used to identify sub-optimal treatment plans. This might be especially relevant in trials in which low numbers of patients are eligible or small numbers of patients are treated in a participating hospital.

Over the past years, several studies have investigated the advantages of automated treatment planning compared to manual planning using in-house developed or commercial algorithms (Breedveld *et al.* 2012; Voet *et al.* 2013b, 2014; Fogliata *et al.* 2014, 2015a; Sharfo *et al.* 2016; Hussein *et al.* 2016; Della Gala *et al.* 2017). Treatment planning with Erasmus-iCycle/Monaco is fully automated and has been successfully validated for clinical use in several treatment sites including prostate (Voet *et al.* 2014), head-and-neck (Voet *et al.* 2013b), cervix (Sharfo *et al.* 2016), and lung (Della Gala *et al.* 2017). Knowledge-based planning using a model-library of previously generated

plans to predict treatment plan parameters for a new patient was configured and tested in pelvic anatomy (Hussein *et al.* 2016), lung (Fogliata *et al.* 2014), and esophageal cancer (Fogliata *et al.* 2015a).

Numerous dosimetric studies have evaluated different radiotherapy techniques for gastric cancer by comparing IMRT, VMAT, helical tomotherapy, 3D-CRT and proton therapy (Li *et al.* 2014; Zhang *et al.* 2015; Mondlane *et al.* 2017; Onal *et al.* 2018). Until now, no study has been published showing the possibility/advantages of automated treatment planning for postoperative gastric cancer patients. In this study, Erasmus-iCycle was used to automatically generate VMAT plans for gastric cancer patients. For this site, manual treatment optimization is a large challenge, even for an expert planner, due to the multi-concave shape, the extent of the target volume, and the close proximity to many radio-sensitive organs (i.e., kidneys, liver, heart, and spinal cord). Another difficulty in postoperative radiotherapy in gastric cancer patients is that radiation tolerance doses of OARs are relatively low. Compared to the manVMAT plans, generated by an expert planner, plan quality was significantly better for the autoVMAT plans, only requiring slightly longer treatment delivery times. Specifically, dose conformity and sparing of organs-at-risk were improved, while the clinical importance of the observed longer treatment delivery time and the increased MUs is considered to be low. As a result of higher modulation, the number of MUs was significantly higher in the autoVMAT than in manVMAT plans, this could lead to challenges in radiation delivery. The integral-dose was significantly reduced in autoVMAT plans, this may be beneficial especially for young patients to avoid secondary tumors, and most notably showed significantly reduced NTCP for liver and both kidneys, although it is not yet clear to what extent the observed dosimetric advantages of autoVMAT versus an experienced planner will really translate into a clinical benefit. Automated treatment planning can improve the efficiency of the treatment planning process and reduce its user dependency, this in turns might lead to more consistent and uniform outcomes in treatment planning studies and clinical trials. Apart from the improved plan quality and reduced predicted complications, automated planning also eliminated the planning hands-on time required for planning.

4.5   Conclusion

Automated treatment planning is of great value for complex treatment sites like for gastric cancer. Compared to manual planning by an expert planner, plan quality could be largely improved, while drastically reducing treatment planning workload.

Automated Volumetric Modulated Arc Therapy Planning for Whole Pelvic Prostate Radiotherapy

Martin Buschmann[1,2], **Abdul Wahab M. Sharfo**[3], **Joan Penninkhof**[3], **Yvette Seppenwoolde**[1,2], **Gregor Goldner**[1], **Dietmar Georg**[1,2], **Sebastiaan Breedveld**[3] **and Ben J.M. Heijmen**[3]

[1] Department of Radiation Oncology, Medical University of Vienna/AKH Wien, Vienna, Austria

[2] Christian Doppler Laboratory for Medical Radiation Research for Radiation Oncology, Medical University of Vienna, Vienna, Austria

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

Abstract

Background: For several tumor entities, automated treatment planning has improved plan quality and planning efficiency, and may enable adaptive treatment approaches. Whole-pelvic prostate radiotherapy (WPRT) involves large concave target volumes, which present a challenge for volumetric-modulated arc therapy (VMAT) optimization. This study evaluates automated VMAT planning for WPRT-VMAT and compares the results with manual expert planning.

Methods: A system for fully automated multi-criterial plan generation was configured for each step of sequential-boost WPRT-VMAT, with final (autoVMAT) plans being automatically calculated by the Monaco® treatment planning system (TPS; Elekta AB, Stockholm, Sweden). Configuration was based on manually generated VMAT plans (manualVMAT) of 5 test patients, the planning protocol, and discussions with the treating physician on wishes for plan improvements. AutoVMAT plans were then generated for another 30 evaluation patients and compared to manualVMAT plans. For all 35 patients, manualVMAT plans were optimized by expert planners using the Monaco® TPS.

Results: AutoVMAT plans exhibited strongly improved organ sparing and higher conformity compared to manualVMAT. On average, mean doses (D_{mean}) of bladder and rectum were reduced by 10.7 Gy and 4.5 Gy, respectively, by autoVMAT. Prostate target coverage ($V_{95\%}$) was slightly higher (+0.6%) with manualVMAT. In a blinded scoring session, the radiation oncologist preferred autoVMAT plans to manualVMAT plans for 27/30 patients. All treatment plans were considered clinically acceptable. The workload per patient was reduced by > 70 min.

Conclusion: Automated VMAT planning for complex WPRT dose distributions is feasible and creates treatment plans that are generally dosimetrically superior to manually optimized plans.

5.1   Background

Modern linear accelerator (linac)-based prostate radiotherapy (RT) is commonly delivered by intensity-modulated radiation therapy (IMRT) and, more recently, by volumetric-modulated arc therapy (VMAT). However, treatment planning is often an iterative trial-and-error process, and the resulting plan quality is strongly dependent on the planner's experience.

The clinical target volume (CTV) usually comprises the prostate gland, but in intermediate- and high-risk disease, the seminal vesicles and pelvic lymph nodes are also commonly irradiated. This technique of whole-pelvic prostate RT (WPRT) has shown clinical benefit compared to local prostate-only RT in several studies (Roach *et al.* 2003; Mantini *et al.* 2011), but its role remains controversial (Morikawa & Roach 2011; Amini *et al.* 2015). The optimization of WPRT-IMRT/VMAT plans is more challenging and time consuming than for the local prostate treatments, due to the larger and more complex target volumes.

In recent years, multi-criterial optimization (MCO) has been applied in the field of treatment plan optimization with the aim of generating *Pareto-optimal* plans which cannot be improved further in one aspect without worsening another. Another motivation for advanced optimization is avoidance of time-consuming iterative treatment planning, which may still lead to treatments lacking Pareto optimality. An MCO approach developed by Craft *et al.* (2007); Craft & Monz (2010) uses manual navigation of a plan-library that spans the Pareto space. This strategy can be classified as an *a posteriori* technique, where manual interaction by the user is necessary to select the best-fitting treatment plan. Erasmus-iCycle is a multi-criterial optimizer for beam profiles and beam angle selection (Breedveld *et al.* 2012) which employs a *a priori* strategy. The user defines a treatment site-specific optimization protocol, a so-called '*wish-list*', containing the goal functions that are optimized in a specific order, defined by assigned priorities. The priorities steer the fully automated multi-criterial plan generation. In addition, the wish-list can contain hard constraints that must not be violated. Generated plans are *Pareto-optimal* and clinically favorable (Voet *et al.* 2013b, 2014; Della Gala *et al.* 2017). Currently, Erasmus-iCycle is used as a pre-optimizer; optimized iCycle plans are subsequently re-constructed with Monaco® TPS (Elekta AB, Stockholm, Sweden) to generate deliverable IMRT or VMAT plans (Voet *et al.* 2014; Sharfo *et al.* 2015). After the development of a wish-list and a translation strategy to the TPS in a multi-disciplinary team, treatment planning with Erasmus-iCycle/Monaco is fully automatic. The algorithm is successfully applied in clinical routine for several tumor sites, including head-and-neck (Voet *et al.* 2013b), local prostate cancer (Voet *et al.* 2014), cervix (Sharfo *et al.* 2016), lung (Della Gala *et al.* 2017), and spinal metastases (Buergy

et al. 2017). Up until now, the algorithm has only been validated against manual planning at a single institution, and the previously demonstrated benefit of auto-planning may be explainable by institution-specific protocols.

This study investigates the feasibility of automated generation of VMAT plans for WPRT with Erasmus-iCycle/Monaco and compares the resulting dose distributions with manually optimized VMAT plans that were generated by expert planners in a different center, also using Monaco®.

5.2   Materials and methods

5.2.1   Patients

From the clinical database of the department of Radiation Oncology in Vienna, 35 prostate cancer patients who were treated with WPRT-VMAT between May 2014 and July 2016 were randomly selected for this study. Patients with hip implants and bladder filling less than 100 ml on the planning CT-scan were not included. Each patient had gold fiducials implanted in the prostate for improved target localization during image guidance. A CT-scan with 2 mm slice thickness was used for treatment planning. Patients were scanned and treated with a rectal balloon. The prostate was delineated as the primary CTV (CTV_P), and the pelvic lymph nodes including seminal vesicles as well as the CTV_P were defined as the secondary CTV (CTV_{LN}). Rectum, bladder, bowel bag, and femoral heads were manually contoured as organs-at-risk (OAR). The bowel was delineated up to a 2 cm extension of the target in the cranial direction. Considering the use of daily image guidance, setup margins of 5 mm were applied to create two planning target volumes: PTV_P and PTV_{LN}.

The total radiation dose was prescribed in a two-phase Vienna-specific protocol: the first phase delivered 60 Gy and 50 Gy in 25 fractions to PTV_P and PTV_{LN}, respectively, using a simultaneous-integrated-boost (SIB) technique. The second phase was a sequential-boost of 13 Gy in 5 fractions to PTV_P. Clinical acceptability of the treatment plans was assessed through a list of clinically applied organ dose constraints which is presented in Table 5.1, and by clinical judgment of the treating physician.

5.2.2   Treatment planning and delivery systems

Prior to initiation of this study, discussions between the radiation oncologist (G.G.), the dosimetrist, and the physicists (M.B., A.S., Y.S., J.P., B.H.) were held to define the guidelines and preferences for dose distributions. A target dose criterion was defined as a minimum target coverage corresponding to total volume irradiated with 95% of the prescribed dose ($V_{95\%}$) = 96% for PTV_P and PTV_{LN} in the total summed dose (phase

Table 5.1: Clinical dose constraints for organs-at-risk for the total summed dose.

Rectum	D_{max}	78 Gy
	V_{65Gy}	20%
	V_{60Gy}	40%
	V_{55Gy}	45%
	V_{50Gy}	50%
Bowel bag	D_{max}	56 Gy
	V_{50Gy}	10%
	V_{45Gy}	15%
	V_{40Gy}	20%
Bladder	D_{max}	78 Gy
	V_{65Gy}	20%
	V_{55Gy}	40%
	V_{50Gy}	50%
	V_{35Gy}	80%
Femoral Heads	D_{max}	55 Gy
	V_{45Gy}	5%

D_{max}: maximum dose, V_{xGy}: relative volume of organ receiving at least x Gy.

1 plus phase 2). The physician expressed his preference for high target coverage and therefore the aim was to reach approximately $V_{95\%} = 98\%$ during optimization of the two single treatment plans.

All final autoVMAT and manualVMAT plans were generated with the Monaco® TPS version 5.11 for a Versa HD™ linac (Elekta) with 10 MV photons. All treatments were planned to be delivered by full 360° coplanar arcs. Phase 1 was planned as dual-arc and the boost phase 2 as single-arc, with the iso-center being placed in the center of PTV_{LN} and PTV_P, respectively.

5.2.3   Manual VMAT planning

Two experienced planners from the Vienna team performed VMAT treatment planning (manualVMAT) in Monaco® for all 35 patients according to current clinical guidelines, with commonly used templates of optimization cost-functions. The employed templates are described in Section 5.7. For each patient, the optimization parameters were iteratively tweaked to improve the dose distributions. This process was performed in the absence of time constraints and the planner only stopped when no further improvement in plan quality could be achieved. The two-plan phases were optimized and finalized independently before the summed dose was inspected to check the OAR constraints. Each patient was planned by just one planner.

Table 5.2: Applied wish-lists for automatic volumetric-modulated arc therapy plan generation for **(A)** phase 1 (i.e., whole pelvis) and **(B)** phase 2 (i.e., boost) plans. A detailed description is included in Section 5.6. The parameters k and α are specific to the iCycle algorithm (Breedveld *et al.* 2012).

(A)

Constraints

Structure	Type	Limit
PTV_P	Max	105% of D_{high}
PTV_P	Mean	101% of D_{high}
PTV_{LN} - PTV_{P+5mm}	Max	105% of D_{low}
PTV_{LN} - PTV_{P+5mm}	Mean	101% of D_{low}
Left & Right Femoral Head	Max	70% of D_{high}
PTV_{LN} Shell 50 mm	Max	50% of D_{high}
Unspecified Tissues	Max	105% of D_{high}

Objectives

Priority	Structure	Type	Goal	Parameters
1	PTV_P	↓ LTCP	0.65	D_{high} = 60 Gy, α = 0.7
2	PTV_{LN}	↓ LTCP	0.5	D_{low} = 50 Gy, α = 0.9
3	PTV_{LN} Shell 5 mm	↓ Max	90% of D_{low}	
4	Rectum	↓ EUD	70% of D_{high}	k = 12
		↓ EUD	30% of D_{high}	k = 4
5	Bladder	↓ EUD	30% of D_{high}	k = 4
6	PTV_{LN} Shell 15 mm	↓ Max	60% of D_{low}	
	PTV_{LN} Shell 25 mm	↓ Max	40% of D_{low}	
7	Bowel Bag	↓ EUD	30% of D_{high}	k = 8
8	Skin Ring 20 mm	↓ Max	30% of D_{high}	
9	Rectum	↓ Mean	20% of D_{low}	
10	Bladder	↓ Mean	20% of D_{low}	
11	Bowel Bag	↓ Mean	20% of D_{low}	
12	Left & Right Femoral Head	↓ EUD	10% of D_{high}	k = 5

(B)

Constraints

Structure	Type	Limit
PTV_P	Max	105% of D^p
PTV_P	Mean	101% of D^p
Left & Right Femoral Head	Max	70% of D^p
PTV_P Shell 50 mm	Max	50% of D^p
Unspecified Tissues	Max	105% of D^p

Objectives

Priority	Structure	Type	Goal	Parameters
1	PTV_P	↓ LTCP	0.5	D^p = 13 Gy, α = 0.95
2	PTV_P Shell 5 mm	↓ Max	90% of D^p	
3	Rectum	↓ EUD	70% of D^p	k = 16
4	Bladder	↓ EUD	60% of D^p	k = 12
5	PTV_P Shell 15 mm	↓ Max	50% of D^p	
	PTV_P Shell 25 mm	↓ Max	30% of D^p	
	Skin Ring 20 mm	↓ Max	30% of D^p	
6	Rectum	↓ Mean	10% of D^p	
7	Bladder	↓ Mean	10% of D^p	
8	Left & Right Femoral Head	↓ EUD	10% of D^p	k = 5

Abbreviations: PTV_P = planning target volume of the prostate, PTV_{LN} = planning target volume of the lymph nodes, PTV_{P+5mm} = a 5 mm transition region within PTV_{LN}, D_{high} = high (= 60 Gy) prescribed dose, D_{low} = low (= 50 Gy) prescribed dose, D^p = prescribed dose, LTCP = *Logarithmic Tumor Control Probability*, α = cell sensitivity parameter to achieve adequate target coverage, EUD = *Equivalent Uniform Dose*, k = volume-effect parameter.

5

5.2.4   Automated VMAT planning with Erasmus-iCycle/Monaco

For automated VMAT planning (autoVMAT), the Erasmus-iCycle algorithm (Breedveld *et al.* 2012) was used as a pre-optimizer. To this purpose, a VMAT plan was approached by using 20 equi-angularly positioned IMRT beams. The two wish-lists for the two treatment phases used in this study are presented in Table 5.2 and a detailed description is provided in Section 5.6. The segmentation settings in Monaco® were identical for manual and automated plan generation, and are described in Section 5.6.

The autoVMAT strategy was developed using a random sub-set of 5 of the total of 35 study patients. This training cohort was used to configure the Erasmus-iCycle/Monaco system for high-quality automated VMAT plan generation, i.e., development of the site-specific wish-lists and automatic translations of pre-optimization results into patient-specific Monaco® templates for both treatment phases. A first strategy for autoVMAT was created according to clinical guidelines and applied on the training cohort to create deliverable VMAT plans. A '*flavor session*' with the radiation oncologist and medical physicists was then held, where manualVMAT and autoVMAT plans for the training patients were blindly presented and discussed. Using the results and physicians' preferences from this flavor session, the autoVMAT configurations were refined and then applied to the 30 patients not used for configuration (evaluation patients).

5.2.5   Dosimetric plan evaluations

For the 30 evaluation patients, the dose distributions of the individual treatment phases and the summed total dose as generated by autoVMAT and manualVMAT were analyzed and compared. Normal tissue sparing was assessed by extracting the near-maximum dose ($D_{2\%}$) and mean dose (D_{mean}) for each OAR, and the total volume irradiated at 95% and 50% of the respective prescription dose. Additionally, for the rectum, V_{70Gy} and V_{55Gy} were assessed in the summed dose, as these parameters are correlated to rectal toxicity (Michalski *et al.* 2010) (adjusted for hypofractionation according to the linear quadratic model with α/β = 3 Gy). For target structures, $V_{95\%}$ (target coverage), $D_{2\%}$, and D_{mean} were analyzed. The homogeneity index (HI) and two conformity indices (*CI*, $CI_{50\%}$) were calculated as follows:

$$HI = \frac{D_{2\%} - D_{98\%}}{D_{50\%}}$$

$$CI = \frac{V_{95\%}}{V_{PTV}}$$

$$CI_{50\%} = \frac{V_{50\%}}{V_{PTV}}$$

where $V_{95\%}$ and $V_{50\%}$ describe the total volumes irradiated with 95% and 50% of the prescribed dose, respectively. For the determination of $D_{2\%}$, D_{mean}, and HI in the PTV_{LN}, the SIB volume including a 2 cm margin was subtracted.

Dosimetric differences between autoVMAT and manualVMAT were tested for statistical significance by a two-sided Wilcoxon matched-paired signed-rank test (p < .05). No rescaling of dose distributions was performed, since D_{mean} was equivalent in the total plan sum for autoVMAT and manualVMAT. The number of monitor units (MU) for each of the two treatment phases was noted.

5.2.6   Physician's plan scoring

In a blinded plan scoring session, the radiation oncologist was simultaneously presented with the total summed plans created by autoVMAT and manualVMAT for all 30 study patients. The cumulative DVHs for all structures, the dose distributions in the TPS, and a table with the most important dosimetric parameters were provided to the physician. Based on this information, the physician rated the two distributions relatively to each other according to three categories:

- Plan A is considerably better than plan B

- Plan B is considerably better than plan A

- Plan A and plan B are of similar quality

5.2.7   Dosimetric verification measurements

Deliverability of autoVMAT and manualVMAT plans was checked by performing verification dose measurements for 10 randomly selected patients. Dose distributions were measured with the $Delta^4$ phantom (ScandiDos, Uppsala, Sweden) and compared to planned doses through a γ analysis (3% local dose difference, 3 mm).

5.3   Results

5.3.1   Dosimetric plan evaluations

All generated autoVMAT and manualVMAT dose distributions fulfilled the criteria for clinical acceptability in terms of OAR sparing and target coverage. Population average DVHs for the 30 evaluation patients for the single plan phases and the total treatment are shown in Figure 5.1, pointing toward meaningful advantages for autoVMAT. Dose–volume parameters of autoVMAT and differences with manualVMAT are presented in Table 5.3.
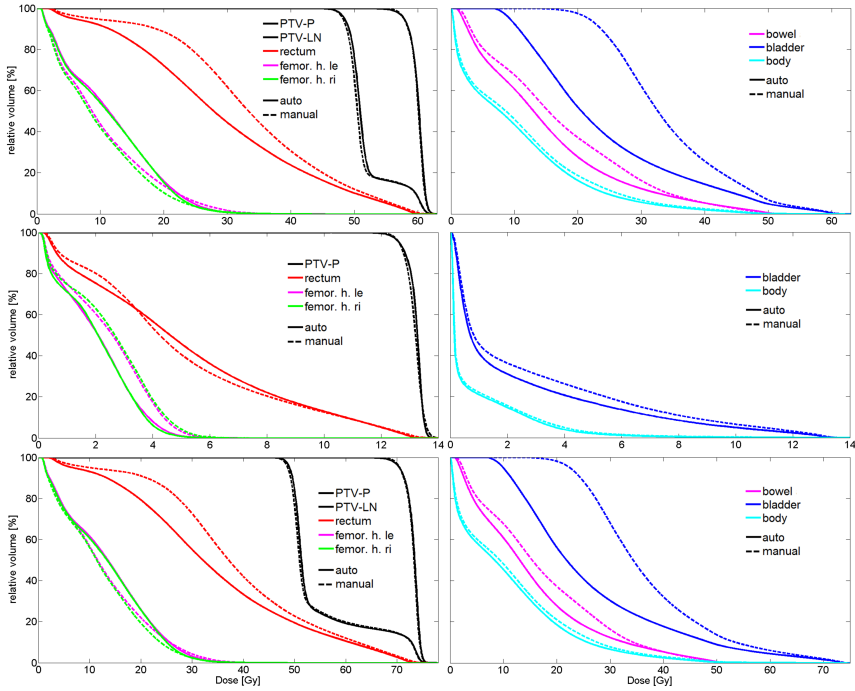
Figure 5.1: Mean dose–volume histograms (DVHs) for automated and manual volumetric-modulated arc therapy (autoVMAT, manualVMAT) plans for the 30 evaluation patients. Top: Phase 1 whole-pelvis plan; Middle: Phase 2 boost plan; Bottom: Sum plan; femor. h. = femoral head, PTV_P = planning target volume prostate, PTV_{LN} = planning target volume lymph nodes.

Figure 5.2 depicts an exemplary dose distribution in an axial slice of the summed treatment for both optimization techniques. Most striking is the improved bladder sparing with autoVMAT, which was also the most important advantage of autoVMAT for the entire study population.
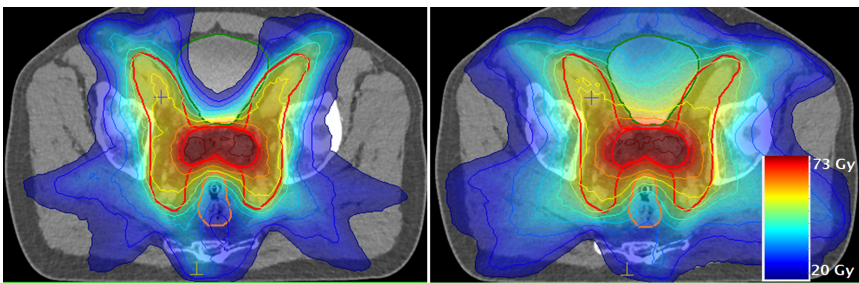
Figure 5.2: Exemplary dose distribution of the total summed plan. Left: Automated volumetric-modulated arc plan (autoVMAT); Right: Manual VMAT plan (manualVMAT).

In general, dose distributions in the targets were similar, with small differences as is shown by the DVHs in Figure 5.1, and Table 5.3. D_{mean} in PTV_P was identical in phase 1 and the summed dose in both optimizations. Target coverage was similar for both optimization methods, but was statistically significantly higher in manualVMAT in the PTV_P by a small amount of 0.6% points on average. In the PTV_{LN}, $V_{95\%}$ was higher in autoVMAT in the summed plan. Significantly lower OAR doses were observed in rectum, bladder, and bowel in autoVMAT compared to manualVMAT, for both phases and for the sum plan. Most prominent was the lower bladder dose in autoVMAT, with average D_{mean} values in the sum plan of 26.3 Gy and 37 Gy for autoVMAT and manualVMAT, respectively. Doses to the femoral heads were generally higher in autoVMAT, but the average difference in D_{mean} was below 1.7 Gy.

AutoVMAT dose distributions were more conformal than manualVMAT doses, which was especially evident in the smaller $CI_{50\%}$ in both planning phases. The volume of the 50% isodose was decreased on average by 852 ml in phase 1 (50% = 25 Gy) and 63 ml in phase 2 (50% = 6.5 Gy) with autoVMAT. ManualVMAT plans delivered slightly more homogeneous target doses. The average number of MUs were 664/1029 for phase 1 and 563/716 for phase 2 in manualVMAT and autoVMAT, respectively.

5.3.2 Physician's plan scoring

During the blinded scoring session, the radiation oncologist considered all plans from both optimizations clinically acceptable. He preferred the autoVMAT plan in 27 out of 30 patients and the manualVMAT in the other 3 patients. He chose the autoVMAT plans mainly for their increased OAR sparing, while the occasional small advantages of manualVMAT in PTV_P coverage were not considered clinically relevant in most cases. However, in the three patients where manualVMAT was considered better, the $V_{95\%}$ and D_{min} were noticeably lower in autoVMAT. After the scoring session, the MUs of the treatment plans in these three patients were rescaled to reach identical values of $V_{95\%}$ in the PTV_P in the autoVMAT and the manualVMAT plans, and another blinded scoring session was held with these three patients. The physician then preferred the autoVMAT plans in the second round in all cases.

Table 5.3: Dosimetric parameters of autoVMAT plans and differences to manualVMAT plans.

	Phase 1 whole pelvis plan			Phase 2 boost plan			Total dose		
	autoVMAT	manualVMAT - autoVMAT	p-value	autoVMAT	manualVMAT - autoVMAT	p-value	autoVMAT	manualVMAT - autoVMAT	p-value
	Mean ± SD	Mean ± SD		Mean ± SD	Mean ± SD		Mean ± SD	Mean ± SD	
CI	1.16 ± 0.04	0.05 ± 0.05	**<.001**	1.10 ± 0.05	0.07 ± 0.05	**<.001**			
$CI_{50\%}$	4.50 ± 0.35	1.17 ± 0.39	**<.001**	3.93 ± 0.35	0.64 ± 0.24	**<.001**			
CI SIB	1.10 ± 0.03	0.02 ± 0.04	**.026**						
HI	0.10 ± 0.01	-0.01 ± 0.01	**<.001**	0.09 ± 0.01	0.00 ± 0.01	NS			
HI SIB	0.08 ± 0.01	-0.01 ± 0.01	**.003**						
PTV_P $V_{95\%}$ (%)	97.56 ± 0.76	0.49 ± 0.93	**.023**	98.27 ± 0.57	0.50 ± 0.48	**<.001**	97.75 ± 0.61	0.60 ± 0.77	**<.001**
PTV_P $D_{2\%}$ (Gy)	61.73 ± 0.11	-0.11 ± 0.28	**.039**	13.60 ± 0.02	0.08 ± 0.08	**<.001**	75.06 ± 0.11	-0.12 ± 0.28	**.020**
PTV_P D_{mean} (Gy)	59.93 ± 0.12	-0.05 ± 0.25	NS	13.16 ± 0.03	-0.04 ± 0.06	**<.001**	73.08 ± 0.13	-0.09 ± 0.27	NS
PTV_{LN} $V_{95\%}$ (%)	98.14 ± 0.70	-0.28 ± 0.84	NS				99.30 ± 0.31	-0.36 ± 0.60	**.003**
PTV_{LN} $D_{2\%}$ (Gy)	52.38 ± 0.20	-0.56 ± 0.28	**<.001**				54.42 ± 1.26	-0.18 ± 0.59	NS
PTV_{LN} D_{mean} (Gy)	50.33 ± 0.18	-0.41 ± 0.20	**<.001**				50.77 ± 0.19	-0.35 ± 0.21	**<.001**
Rectum $D_{2\%}$ (Gy)	57.58 ± 1.34	0.90 ± 0.75	**<.001**	12.66 ± 0.30	0.15 ± 0.12	**<.001**	70.17 ± 1.63	1.05 ± 0.80	**<.001**
Rectum D_{mean} (Gy)	29.18 ± 3.63	4.63 ± 2.08	**<.001**	5.05 ± 0.84	-0.08 ± 0.36	NS	34.23 ± 4.00	4.53 ± 2.16	**<.001**
Rectum V_{70Gy} (%)							2.68 ± 1.58	0.60 ± 0.57	**<.001**
Rectum V_{55Gy} (%)							14.59 ± 4.83	2.22 ± 1.76	**<.001**
Bladder $D_{2\%}$ (Gy)	55.30 ± 4.01	1.40 ± 1.47	**<.001**	11.72 ± 1.33	0.69 ± 0.72	**<.001**	66.67 ± 5.76	2.25 ± 2.18	**<.001**
Bladder D_{mean} (Gy)	24.03 ± 3.73	10.25 ± 3.03	**<.001**	2.20 ± 1.07	0.44 ± 0.31	**<.001**	26.28 ± 4.18	10.68 ± 3.21	**<.001**
Hip Left $D_{2\%}$ (Gy)	26.38 ± 3.28	2.14 ± 3.33	**.002**	4.46 ± 0.55	0.56 ± 0.43	**<.001**	30.00 ± 3.70	2.21 ± 3.23	**<.001**
Hip Left D_{mean} (Gy)	11.53 ± 1.38	-1.22 ± 1.28	**<.001**	1.98 ± 0.36	0.40 ± 0.26	**<.001**	13.52 ± 1.63	-0.83 ± 1.32	**.003**
Hip Right $D_{2\%}$ (Gy)	26.01 ± 3.38	0.18 ± 3.13	NS	4.39 ± 0.47	0.83 ± 0.58	**<.001**	29.54 ± 3.45	0.20 ± 3.11	NS
Hip Right D_{mean} (Gy)	11.32 ± 1.25	-1.64 ± 1.05	**<.001**	1.95 ± 0.30	0.50 ± 0.22	**<.001**	13.29 ± .144	-1.15 ± 1.16	**<.001**
Bowel $D_{2\%}$ (Gy)	45.17 ± 3.29	0.26 ± 1.73	NS				45.30 ± 3.28	0.26 ± 1.73	NS
Bowel D_{mean} (Gy)	15.22 ± 3.29	2.07 ± 1.24	**<.001**				15.27 ± 3.30	2.09 ± 1.24	**<.001**
MU	1029 ± 130	-365 ± 115	**<.001**	716 ± 89	-153 ± 70	**<.001**			

Abbreviations: autoVMAT = automated volumetric-modulated arc plan, manualVMAT = manual volumetric-modulated arc plan, NS = non-significant, PTV_P = planning target volume of the prostate, PTV_{LN} = planning target volume of the lymph nodes, CI = *Conformity Index*, HI = *Homogeneity Index*, SIB = simultaneous-integrated-boost, $V_{95\%}$ = volume receiving at least 95% of the prescribed dose, D_{mean} = mean dose, $D_{2\%}$ = minimum dose delivered to the most exposed 2% of the structure, V_{70Gy} and V_{55Gy} = the volumes receiving 70 Gy, respectively 55 Gy, or more, MU = *Monitor Units*, SD = Standard Deviation.

5

5.3.3 Planning workload

The planning workload in manualVMAT was assessed by recording the time for parameter tweaking and re-optimization after the initial optimization run, as the planner would usually not wait in front of the workstation during the first calculation. The average manual planning time was 54 min for a phase 1 plan and 22 min for a phase 2 plan, which accumulated to 76 min workload per patient. The additional optimization time in Erasmus-iCycle for autoVMAT was in the range of 3-6 h for phase 1 and 20-40 min for phase 2.

5.3.4 Dosimetric verification measurements

All treatment plans, both autoVMAT and manualVMAT, exhibited γ pass rates (GPR) > 95% in the verification measurements, fulfilling the in-house QA requirements. The mean GPR in phase 1/phase 2 were 98.3%/99.3% for autoVMAT and 98.6%/99.6% for manualVMAT. Differences between the verification results of both optimization methods were not statistically significant.

5.4 Discussion

In this study, the potential of fully automated, *a priori* multi-criterial VMAT planning for whole-pelvis prostate treatments was investigated for the first time. This treatment site involves a large concave-shaped target with a SIB, and therefore presents a new and complex optimization problem for the algorithm.

Automated planning significantly improved OAR sparing compared to manual planning. In the total sum plans, the dose to bladder, rectum, and bowel was substantially reduced. Although a slightly higher target coverage in the prostate PTV was observed in manualVMAT, this small difference was not considered clinically relevant. Additionally, the conformity strongly improved with autoVMAT. The dosimetric advantages were also confirmed by the radiation oncologist, who preferred the autoVMAT plan over manualVMAT in 90% of cases in a blinded plan comparison session.

The driving forces of auto-planning with Erasmus-iCycle are the wish-lists, which essentially define planning protocols or recipes to be followed by the algorithm for automatic plan generation. The wish-lists are specific for a patient population. In contrast to generally applied planning templates, the wish-lists do not need patient-specific fine-tuning. Wish-lists are generated in an iterative procedure, using repeated automated plan generations for a small group of test patients.

AutoVMAT planning in this study was indeed fully automatic with no hands-on time, resulting in a reduction of more than 70 min of manual planning time. However, the

creation and refining of wish-lists in a multi-disciplinary team prior to auto-planning for a patient population is time-consuming, which needs to be considered in the departmental logistics.

The number of MUs was significantly higher in autoVMAT than in manualVMAT plans, as a result of higher modulation. The increased plan complexity might lead to challenges in radiation delivery, but no delivery errors could be detected in the verification measurements. More MUs might also lead to more head scatter and higher peripheral doses.

A commercially available *a posteriori* MCO approach (Bokrantz 2012) has also demonstrated its potential in WPRT-VMAT planning (Buschmann *et al.* 2016). Knowledge-based treatment planning strategies have also emerged, which use a model created from a library of approved plans to optimize dose distributions for a new patient (Hussein *et al.* 2016). However, these approaches still require manual interaction or the establishment of a library of high-quality treatment plans. One limitation of this study is that the *Pareto-optimal* pre-optimizer (with sub-sequent re-construction in Monaco®) in autoVMAT was not compared to a manual multi-criterial optimizer (Craft & Monz 2010; Bokrantz 2012), but rather to current standard clinical practice planning.

Precise concurrent targeting of the prostate and the pelvic lymph nodes with radiation can be challenging, as both target volumes can move independently. The lymph nodes remain relatively fixed to the bony anatomy, while the prostate may shift more than 15 mm (Bylund *et al.* 2008; Peng *et al.* 2011) in position with variable bladder and rectum filling. Therefore, adaptive RT (ART) has been proposed for WPRT by employing off-line re-planning or pre-planned plan-libraries (Qi *et al.* 2014; Cantin *et al.* 2015; McVicar *et al.* 2016). These approaches are rarely clinically realized due to the high workload. With the demonstrated efficiency increase achieved through auto-planning, multiple VMAT planning and the implementation of ART for WPRT may be feasible. Another advantage of automated planning is the absence of inter- and intra-planner variations, and a higher homogeneity in plan quality. This may lead to more consistent outcomes in treatment planning studies and clinical trials.

A limitation of the presented (clinical) planning work-flow is that the boost-plan is created in the initial planning step and possible systematic changes in anatomy over the first weeks of treatment are not incorporated.

Hip prostheses are rather common in the prostate cancer patient population, but these patients were not considered in this study. WPRT planning can be very complex in the presence of hip implants. Voet *et al.* (2013a) investigated automated planning for treatment of prostate plus seminal vesicles (no nodes), where automated beam angle optimization was also considered.

Elekta AB is currently working on the integration of Erasmus-iCycle into their commercial Monaco® TPS.

5.5   Conclusion

Fully automated VMAT planning of whole-pelvic prostate treatments with large and complex target volumes created dosimetrically superior treatment plans compared to plans generated manually by experts. The autoVMAT plans were favored by the radiation oncologist in 90% of patients and the planning workload was considerably reduced.

5.6   Automated VMAT planning with Erasmus-iCycle/ Monaco

Automated treatment plan generation with Erasmus-iCycle for an individual patient is based on a site-specific 'wish-list' containing the clinical plan objectives with ascribed priorities and hard constraints to be strictly obeyed. The wish-list used for the whole pelvis contains constraints to control maximum and mean dose in the targets and OARs, and dose fall-off outside the PTV. PTV coverage and its homogeneous dose distribution are optimized using the Logarithmic Tumor Control Probability (LTCP) function (Alber & Reemtsen 2007) with the highest priority (priorities 1 and 2), followed by a shell around the PTV to create a conformal dose distribution (priority 3). Equivalent uniform dose (EUD) (Niemierko 1997) objectives with volume effect parameter ($k = 12$ and 4) focusing on reduction of the high and midrange dose in the rectum and bladder, respectively (priorities 4 and 5). Two more shells around the PTV (priority 6 in Table 5.2) were defined to make a conformal dose distribution. To reduce mid- to high-range dose in the bowel bag another EUD with ($k = 8$) was defined (priority 7). A skin ring of 2 cm wide, from the body contour towards the patient's internal was defined to control entrance doses (priority 8). Mean dose in the rectum, bladder and bowel bag were then reduced with priorities 9,10, and 11. For the left and right femoral heads, EUD objectives were used as shown in Table 5.2. After plan optimization in Erasmus-iCycle, a template containing the achieved results for constraints and objectives is automatically created and fed into the clinical TPS (Monaco® v5.11, Elekta AB, Sweden) to generate, without planner interaction, clinically deliverable dual-arc VMAT plans.

The following segmentation settings in Monaco® were used and kept constant for all automated and manual plan in both planning phases:

- Fluence smoothing = medium

- Maximum number of control points per arc = 230

- Minimum segment width = 1 cm

- Sector increment = 30°

5.7 Manual VMAT planning with Monaco

In Table 5.4 and Table 5.5, the cost-function templates for manual planning in Monaco® are presented. Plan optimizations were initiated with these population-based templates and in the next iterative steps functions and parameters were manually tweaked to obtain the best possible dose distributions.

Table 5.4: Applied Monaco® template for manual VMAT plan generation for phase 1 (i.e., whole pelvis) plan.

Volume	Type	Dose level	Parameters	Shrink margin
PTV_P	Target EUD	60 Gy	Cell sensitivity: 0.5	
	Quadratic overdose	61 Gy	RMS Dose excess: 0.17 Gy	
	Target penalty	58.8 Gy	Min. volume: 99%	
PTV_{LN}	Target EUD	50 Gy	Cell sensitivity: 0.5	
	Quadratic overdose	51 Gy	RMS Dose excess: 0.3 Gy	0.5 cm
	Quadratic overdose	52 Gy	RMS Dose excess: 0.1 Gy	0.5 cm
	Target penalty	49.5 Gy	Min. volume: 99%	
Rectum	Serial	46 Gy	PLE 15	
	Serial	32 Gy	PLE 10	1 cm
Bladder	Serial	41 Gy	PLE 10	
	Serial	32 Gy	PLE 6	1 cm
Bowel	Serial	34 Gy	PLE 10	
	Quadratic overdose	50 Gy	RMS Dose excess: 0.1 Gy	
Left & **Right Femoral Head**	Serial	24 Gy	PLE 6	
Body	Quadratic overdose	53 Gy	RMS Dose excess: 0.08 Gy	
	Quadratic overdose	41 Gy	RMS Dose excess: 0.17 Gy	0.6 cm
	Quadratic overdose	28 Gy	RMS Dose excess: 0.7 Gy	1.5 cm

Abbreviations: PTV_P = planning target volume of the prostate, PTV_{LN} = planning target volume of the lymph nodes, EUD = Equivalent Uniform Dose, PLE = Power Law Exponent, RMS = Root Mean Square.

Table 5.5: Applied Monaco® template for manual VMAT plan generation for phase 2 (i.e., boost) plan.

Volume	Type	Dose level	Parameters	Shrink margin
PTV_P	Target EUD	13 Gy	Cell sensitivity: 0.5	
	Quadratic overdose	13.1 Gy	RMS Dose excess: 0.3 Gy	
	Target penalty	12.9 Gy	Min. volume: 98%	
Rectum	Serial	10 Gy	PLE 15	
	Serial	5 Gy	PLE 10	1 cm
Bladder	Serial	9 Gy	PLE 10	
Left & **Right Femoral Head**	Serial	3 Gy	PLE 6	
Body	Quadratic overdose	10 Gy	RMS Dose excess: 0.16 Gy	
	Quadratic overdose	5 Gy	RMS Dose excess: 0.3 Gy	1.5 cm

Abbreviations: PTV_P = planning target volume of the prostate, EUD = Equivalent Uniform Dose, PLE = Power Law Exponent, RMS = Root Mean Square.

Comparison of VMAT and IMRT Strategies for Cervical Cancer Patients Using Automated Planning

Abdul Wahab M. Sharfo, Peter W.J. Voet, Sebastiaan Breedveld, Jan-Willem M. Mens, Mischa S. Hoogeman and Ben J.M. Heijmen

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

Abstract

Background and Purpose: In a published study on cervical cancer, 5-beam IMRT was inferior to single-arc VMAT. Here we compare 9, 12, and 20 beam IMRT with single- and dual-arc VMAT.

Material and Methods: For each of 10 patients, automated plan generation with the in-house Erasmus-iCycle optimizer was used to assist an expert planner in generating the five plans with the clinical TPS.

Results: For each patient, all plans were clinically acceptable with a high and similar PTV coverage. OAR sparing increased when going from 9 to 12 to 20 IMRT beams, and from single- to dual-arc VMAT. For all patients, 12 and 20 beam IMRT were superior to single- and dual-arc VMAT, with substantial variations in gain among the study patients. As expected, delivery of VMAT plans was significantly faster than delivery of IMRT plans.

Conclusions: Often reported increased plan quality for VMAT compared to IMRT has not been observed for cervical cancer. Twenty- and 12-beam IMRT plans had a higher quality than single- and dual-arc VMAT. For individual patients, the optimal delivery technique depends on a complex trade-off between plan quality and treatment time that may change with introduction of faster delivery systems.

6.1   Introduction

Volumetric-modulated arc therapy (VMAT) and intensity-modulated radiotherapy with static gantry angles (IMRT) have been compared for various tumor sites, both regarding plan quality and treatment time (Teoh *et al.* 2011). In most studies, VMAT was a new technology and plans were compared to previously delivered clinical IMRT plans. Systematic comparisons with variations in the number of IMRT beams and VMAT arcs are scarce (Guckenberger *et al.* 2009; Bortfeld 2010; Quan *et al.* 2012a). Several studies on advanced cervical cancer have demonstrated reduced OAR doses with IMRT compared to 3D-conformal radiotherapy (Portelance *et al.* 2001; Mundt *et al.* 2002; Heron *et al.* 2003; Ahamad *et al.* 2005; van de Bunt *et al.* 2006). To the best of our knowledge, only Cozzi *et al.* (2008) have compared VMAT and IMRT for cervical cancer patients.

Recently, we have developed Erasmus-iCycle, an optimizer for automated, multi-criterial beam profile optimization and beam angle selection for coplanar and non-coplanar IMRT (Breedveld *et al.* 2007, 2009, 2012; Rossi *et al.* 2012; Voet *et al.* 2012, 2013a,b). In Erasmus-iCycle, the common manual, trial-and-error tweaking of plan parameters by dosimetrists is replaced by a fully automated procedure, based on lexicographic optimization. The automation allows objective comparison of treatment strategies, e.g. with variable numbers of IMRT beams, or considering non-coplanar vs. coplanar IMRT (Rossi *et al.* 2012; Voet *et al.* 2012, 2013b). Currently, Erasmus-iCycle optimizes beam fluences. For generation of clinical plans, Erasmus-iCycle plans are re-constructed and segmented in the clinical treatment planning system (TPS). In a recent prospective study, we have demonstrated superior plan quality of Erasmus-iCycle plans that were 'manually' reconstructed in the clinical TPS (semi-automatic plan generation), compared to the clinical routine of trial-and-error treatment planning. In 97% of cases, the treating physician selected the semi-automatically generated plan for treatment rather than the IMRT plan generated manually by a dosimetrist without prior knowledge of the Erasmus-iCycle result (Voet *et al.* 2013b). Apart from the improved plan quality, the semi-automated procedure also reduced the planning hands-on time by 50% compared to manual planning.

Cozzi *et al.* (2008) have retrospectively compared single-arc VMAT plans with previously delivered 5-field equi-angular IMRT plans for treatment of cervical cancer patients. In this study, we have systematically compared single- and dual-arc VMAT plans with IMRT plans with 9, 12 and 20 beams. All these plans were generated semi-automatically (above). Apart from the mutual comparisons, the semi-automatic plans were also compared to clinically delivered 9-beam IMRT plans.

6.2 Materials and methods

6.2.1 Patients and clinical treatment plans

Planning CT-scans of 10 randomly selected, previously treated cervical cancer patients were included in this study. All scans had delineated small bowel (SB), rectum, bladder, and a planning target volume (PTV). To establish the PTV, an internal target volume (ITV) was first constructed using a pre-treatment established motion-model from a full and an empty bladder CT-scan, which was then uniformly expanded by a 1 cm margin (Bondar *et al.* 2012). Patients were treated in prone position lying on a small bowel displacement system (belly board), and using step-and-shoot 9 field IMRT with 10 MV photons at a Synergy® linear accelerator (Elekta AB, Sweden). The linacs were equipped with the MLCi2 multi-leaf collimator (MLC) and the dose rate was 600 MU/min. The prescription dose was 46 Gy delivered in 23 fractions. As defined by the clinical protocol, all plans were created to deliver ≥ 95% of the prescribed dose to a minimum of 99.5% of the PTV, and no more than 0.2% of the PTV should receive ≥ 110%. Cervical cancer target volumes are generally large. For this reason we did not use commonly applied $D_{98\%}$ and $D_{2\%}$ parameters for these patients to evaluate low and high PTV doses, respectively. The main planning goal was to create highly conformal plans providing the best sparing of the delineated OARs with a higher concern for the SB which has the most severe complications (Roeske *et al.* 2003). All plans were generated with our clinical TPS (Monaco® version 3.3, Elekta AB, Sweden).

6.2.2 Systematic comparisons of planning strategies

For each patient, 5 plans were semi-automatically generated in a two-step process. In the first step, Erasmus-iCycle ((Breedveld *et al.* 2012), and Section 6.2.3) was used for fully automatic generation of a high quality 20-beam equi-angular IMRT plan. Using the Erasmus-iCycle results for guidance, an expert planner (Abdul Wahab M. Sharfo) then manually designed 5 plans with the clinical TPS; single-arc VMAT (VMAT), dual-arc VMAT (2VMAT), and IMRT plans with 9 (9DMLC), 12 (12DMLC) and 20 (20DMLC) equi-angular beams, delivered with dynamic MLC. To this purpose, the Erasmus-iCycle output was used to generate a patient specific plan prescription for Monaco® (Voet *et al.* 2014). For all plans, the goal was to maximally approach (or supersede) the corresponding Erasmus-iCycle plan results, considering the established hard constraints and priorities for the various objectives (Section 6.2.3). In order to avoid bias in planning results, plan generations for individual patients were spread out in time. Moreover, for each patient, the five plans were generated in an arbitrary order, and consecutive generation of plans for a single patient was avoided. The expert stopped working on the

plans if plan quality leveled off and no significant further improvement was expected; there was no time limit. As in clinical practice, semi-automatic plans were generated for 10 MV beams, the MLCi2 multi-leaf collimator, and a maximum dose rate of 600 MU/min.

Generated plans were visually inspected by the clinician participating in this study (Jan Willem M. Mens) to assess acceptability for clinical use. For small bowel, the volumes receiving > 15 Gy and > 45 Gy, V_{15Gy} and V_{45Gy} (Kavanagh *et al.* 2010), and the mean dose, D_{mean}, were used for quantitative plan comparisons. For rectum and bladder, D_{mean} was used. Paired two-sided Wilcoxon signed-rank tests were performed to compare the different planning strategies. Differences were considered statistically significant for $p < .05$.

For a sub-set of 5 patients, all 5 semi-automatically generated plans were actually delivered at one of our treatment units in the absence of the patient to record treatment delivery times.

6.2.3 Automatic plan generation with Erasmus-iCycle

Erasmus-iCycle has been described in detail in (Breedveld *et al.* 2012). In this study, the optimizer was used for fluence profile optimization, using a fixed, equi-angular 20-beam setup (Section 6.2.2). Plan optimization with Erasmus-iCycle is based on an a priori defined '*wish-list*' with hard constraints, and prioritized objectives with goal values (Breedveld *et al.* 2007, 2009). Each treatment site has a fixed wish-list. In several studies, we have observed superior plan quality with this approach compared to 'manual' trial-and-error tweaking of TPS parameters by dosimetrists in a clinical setting (Rossi *et al.* 2012; Voet *et al.* 2012, 2013b).

The cervical cancer wish-list as used in this study is presented in Table 6.1. The constraints for shells around the PTV at 1.5 cm, 2.5 cm and 4 cm were defined to ensure a steep dose fall-off outside the PTV, as well as a conformal dose distribution. Among the objectives, PTV coverage had the highest priority. To achieve adequate coverage, the Logarithmic Tumor Control Probability (LTCP) was used as objective function (Alber & Reemtsen 2007). A skin ring of 2 cm wide, extending from the body contour towards the PTV, was defined to control the entrance dose (priority 2). In line with current clinical practice, sparing of SB had a higher priority than reducing delivered dose to the rectum and bladder. Reduction of the mean SB, rectum and bladder doses was performed using a multi-level approach (Breedveld *et al.* 2012), i.e., by repeated use of the objective functions with decreasing priorities and goal values (Table 6.1).

Table 6.1: Applied wish-list for all study patients.

Constraints

Volume	Type	Limit
PTV	Max	105% of D^p
PTV Shell 15 mm	Max	75% of D^p
PTV Shell 25 mm	Max	65% of D^p
PTV Shell 40 mm	Max	50% of D^p
Unspecified Tissues	Max	105% of D^p

Objectives

Level	Priority	Volume	Type	Goal	Parameters
	1	PTV	↓ LTCP	1	D^p = 46 Gy, α = 0.75
	2	Skin Ring 20 mm	↓ Max	50% of D^p	
	3	Small Bowel	↓ Mean	40 Gy	
1	4	Rectum	↓ Mean	40 Gy	
	5	Bladder	↓ Mean	40 Gy	
	6	Small Bowel	↓ Mean	20 Gy	
2	7	Rectum	↓ Mean	20 Gy	
	8	Bladder	↓ Mean	20 Gy	
	9	Small Bowel	↓ Mean	10 Gy	
3	10	Rectum	↓ Mean	10 Gy	
	11	Bladder	↓ Mean	10 Gy	
	12	Unspecified Tissues	↓ Mean	− − −	

Abbreviations: D^p = prescribed dose, LTCP = *Logarithmic Tumor Control Probability*, α = cell sensitivity parameter.

6.3   Results

6.3.1   Semi-automatic plan generation

All generated plans were clinically acceptable. Based on achieved plan parameters of an Erasmus-iCycle 20-beam IMRT plan, the planner succeeded in all cases to generate a clinically deliverable 20DMLC Monaco® plan with highly similar plan quality. As in the clinical plans, in all semi-automatically generated plans, at least 99.5% of the PTV received 95% or more of the prescribed dose, while the volume receiving > 110% was below 0.2%. For the whole population and all plans, the minimum dose in the PTV-1cc was 92.2% ± 1.1% of the prescribed dose. For each patient, differences in PTV coverage and PTV mean and minimum doses among the various plans were insignificant. Because of the high similarities in PTV doses in all plans for each of the patients, in the remainder of this section, plan quality comparisons are based on differences in OAR dose delivery. Figure 6.1 shows for the included patient 8, comparisons between 20DMLC, 2VMAT and clinical dose distributions in axial planes. Conformality is clearly highest for 20DMLC and lowest for the clinical plan.
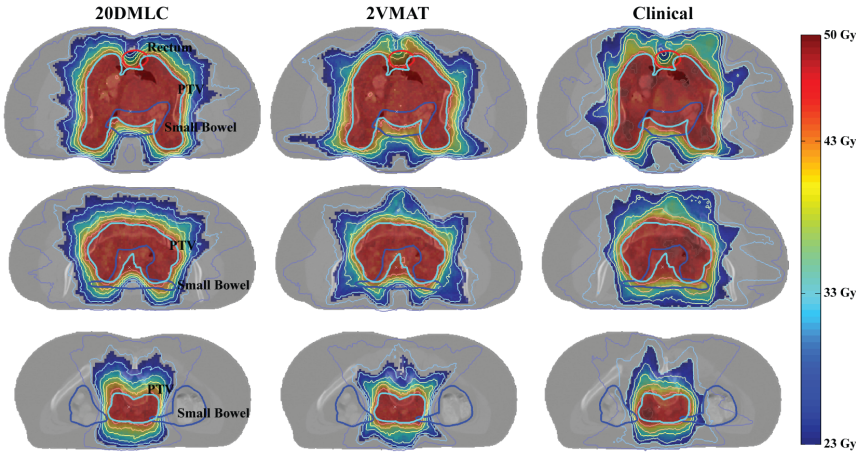
Figure 6.1: Axial slices for patient 8 showing the differences in dose distribution between 20DMLC, 2VMAT, and clinical plans. The shown isodose lines are 30%, 40%, 50%, 60%, 70%, 80%, 95% and 107% of the prescribed dose, respectively.

6.3.2 Semi-automatic vs. clinical plan quality

The last row of Table 6.2 demonstrates that for all semi-automatic IMRT and VMAT procedures (20DMLC, 12DMLC, 9DMLC, 2VMAT, VMAT), the 5 OAR plan parameters were on average lower than the corresponding parameters in the clinical plans. For the 3 SB plan parameters, reductions in the semi-automatic procedures were always statistically significant (last column of Table 6.2). For the (lower priority) rectum and bladder mean doses, reductions were statistically significant for 20DMLC, 12DMLC, 9DMLC, and 2VMAT, but not for VMAT. The last column in Table 6.2 shows that for each semi-automatic procedure, the majority of patients had lower OAR parameter values than in the clinical plan. Superiority of the semi-automatically generated plans over the clinically applied plans is confirmed in Figure 6.2.

6.3.3 Semi-automatic dual-arc vs. single-arc VMAT plan quality

On average, semi-automatic dual-arc VMAT planning resulted in reduced OAR dose delivery compared to single-arc (Figure 6.2 and Table 6.2). Each of the 5 evaluated OAR plan parameters was lowest with 2VMAT for 50% or more of patients (Table 6.2). For SB V_{45Gy} and SB D_{mean}, the differences between 2VMAT and VMAT were statistically significant (Table 6.2). While both for 2VMAT and VMAT, SB V_{45Gy}, SB V_{15Gy}, and SB D_{mean} were statistically significantly lower than in the clinical plan, significant reductions for rectum and bladder D_{mean} were only observed for dual-arc planning (Table 6.2).

Table 6.2: Pair-wise comparisons of planning strategies.

		20DMLC		12DMLC		9DMLC		2VMAT		VMAT		Clinical	
		p-value	*Pts*	*p-value*	*Pts*	*p-value*	*Pts*	*p-value*	*Pts*	*p-value*	*Pts*	*p-value*	*Pts*
	a			NS	(6|4)	.031	(7|3)	NS	(7|3)	.006	(9|1)	.002	(10|0)
	b			.037	(9|1)	.004	(9|1)	.002	(10|0)	.008	(9|1)	.002	(10|0)
20DMLC	**c**			.002	(10|0)	.002	(10|0)	.004	(10|0)	.002	(10|0)	.002	(10|0)
	d			.01	(8|2)	.006	(9|1)	.002	(10|0)	.002	(10|0)	.002	(10|0)
	e			NS	(8|2)	NS	(7|3)	.002	(10|0)	.002	(10|0)	.002	(10|0)
	a	1.1/15.2				NS	(6|4)	NS	(5|5)	NS	(8|2)	.002	(10|0)
	b	2.6/50.0				NS	(8|2)	.006	(9|1)	.02	(8|2)	.002	(10|0)
12DMLC	**c**	0.6/20.9				.002	(10|0)	.006	(9|1)	.004	(9|1)	.002	(10|0)
	d	0.5/43.0				NS	(7|3)	.004	(9|1)	.004	(10|0)	.004	(9|1)
	e	0.3/42.1				NS	(7|3)	.002	(10|0)	.004	(10|0)	.002	(10|0)
	a	1.2/15.3		0.1/15.8				.01	(3|7)	NS	(7|3)	.002	(10|0)
	b	4.8/51.1		2.2/52.3				.027	(8|2)	NS	(7|3)	.004	(9|1)
9DMLC	**c**	1.3/21.2		0.7/21.5				NS	(7|3)	NS	(7|3)	.004	(9|1)
	d	0.8/43.1		0.2/43.4				NS	(8|2)	.049	(8|2)	.006	(9|1)
	e	0.5/42.2		0.2/42.4				.004	(9|1)	.002	(10|0)	.002	(10|0)
	a	0.2/14.8		-0.9/15.3		-1.0/15.4				.012	(9|1)	.002	(10|0)
	b	9.0/53.2		6.4/54.4		4.2/55.5				NS	(5|5)	.01	(8|2)
2VMAT	**c**	1.7/21.4		1.1/21.7		0.5/22.1				.027	(9|1)	.002	(10|0)
	d	1.2/43.3		0.6/43.6		0.4/43.7				NS	(7|3)	.037	(8|2)
	e	1.3/42.6		1.0/42.8		0.8/42.9				NS	(8|2)	.014	(9|1)
	a	1.3/15.4		0.3/15.9		0.2/16.0		1.2/15.5				.004	(9|1)
	b	9.6/53.5		7.1/54.8		4.9/55.9		0.7/58.0				.002	(10|0)
VMAT	**c**	2.1/21.6		1.5/21.9		0.9/22.3		0.4/22.5				.002	(10|0)
	d	1.5/43.5		1.0/43.9		0.8/43.8		0.4/44.1				NS	(7|3)
	e	1.8/42.9		1.5/43.0		1.3/43.1		0.5/43.5				NS	(7|3)
	a	4.1/16.8		3.1/17.3		3.0/17.4		4.0/16.9		2.8/17.4			
	b	17.4/57.4		14.8/58.6		12.6/59.7		8.4/61.8		7.7/62.2			
Clinical	**c**	4.5/22.8		3.9/23.1		3.3/23.5		2.8/23.7		2.4/23.9			
	d	1.9/43.7		1.4/44.0		1.2/44.1		0.8/44.3		0.4/44.5			
	e	2.5/43.2		2.2/43.4		2.0/43.5		1.2/43.9		0.7/44.1			

Abbreviations: NS = no statistically significant difference i.e. $p > .05$.
In each comparison (table cell), going from top to bottom, data refer to (**a**) SB V_{45Gy}, (**b**) SB V_{15Gy}, (**c**) SB D_{mean}, (**d**) Bladder D_{mean}, and (**e**) Rectum D_{mean}. Below the table diagonal, the "A/B" in the cells refer to A: plan parameter value for the strategy along the vertical axis minus the parameter value for the strategy along the horizontal axis, averaged over the 10 patients in the study, B: average of the patient-mean OAR parameter values in the two compared strategies. Cells above the diagonal show *p*-values and "(*n*|*m*)": for *n* patients, the strategy indicated at the horizontal axis has lowest OAR dose, while for *m* patients the strategy mentioned at the vertical axis is superior.

6
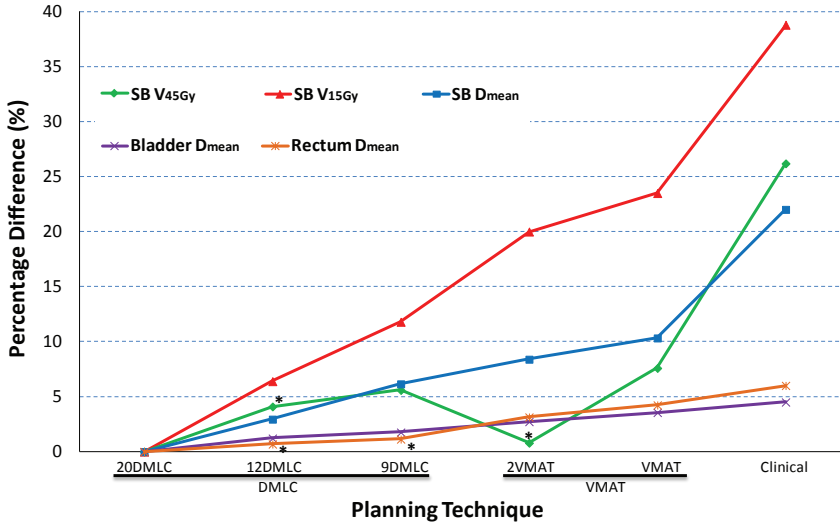
Figure 6.2: For the semi-automatic planning strategies 12DMLC, 9DMLC, 2VMAT and VMAT and the clinical plans, percentage increases of mean OAR plan parameter values relative to the semi-automatic 20DMLC. Average absolute values for 20DMLC: SB V_{45Gy} = 102 cc, SB V_{15Gy} = 345 cc, SB D_{mean} = 20.6 Gy, Rectum D_{mean} = 42 Gy and Bladder D_{mean} = 42.7 Gy. ∗: Difference from 20DMLC not statistically significant.

6.3.4   Semi-automatic 20DMLC vs. 12DMLC vs. 9DMLC plan quality

Figure 6.2 and Table 6.2 show that all mean OAR parameter values reduce when going from 9 to 12 beams and from 12 to 20. Both beam additions result in a benefit for the majority of patients (Table 6.2).

6.3.5   Semi-automatic IMRT vs. VMAT plan quality

Figure 6.2 and Table 6.2 demonstrate that all semi-automatic IMRT strategies (20DMLC, 12DMLC, and 9DMLC) have better OAR sparing than single-arc VMAT plans. The differences are also mostly statistically significant (Table 6.2). For all OAR parameters, the majority of patients also benefited from 20DMLC and 12DMLC compared to dual-arc VMAT (2VMAT) (Table 6.2). Figure 6.3 and Figure 6.4 provide for each patient separately, detailed comparisons between 20DMLC and 2VMAT, and the clinical plan (Figure 6.3 only). Both figures underline the overall superiority of semi-automatic 20DMLC plan quality compared to semi-automatic 2VMAT. There are however substantial variations among patients in the differences between the two strategies. For patients 2, 8 and 9 these differences are relatively small, while patients 3–6 and 10 show higher gains of 20DMLC. Treatment times of 20DMLC plans are clearly enhanced compared to 2VMAT

plans. In Section 6.4, improved plan quality is discussed in the context of treatment time enhancement.

6.3.6   Planning times

Fully automated plan generation with Erasmus-iCycle to generate input for the manual planning in the clinical TPS took on average 4.5 h per patient and was done overnight (no workload). The average hands-on planning time for generating a plan in the clinical TPS, based on Erasmus-iCycle output, was 1.5 h with a minimum of 0.5 h and a maximum of 3 h; VMAT hands-on planning times were almost a factor of 2 larger than that for DMLC. The average hands-on planning time spent on generation of clinical plans was 3 h with a range between 2 and 5 h, depending on the complexity of the case. The mentioned times for plan generation in the clinical TPS include times for Monte Carlo dose calculation.

6.3.7   Treatment delivery times

For the studied sub-set of 5 patients, the recorded mean treatment delivery times for VMAT, 2VMAT, 9DMLC, 12DMLC and 20DMLC were 3.0 [2.9, 3.2], 4.1 [3.8, 4.5], 10.5 [9.1, 12], 12.8 [10.3, 15.2], and 15.6 min. [14, 17.6], respectively.

6.4   Discussion

VMAT and IMRT plans have been compared in many studies and for various tumor sites (Teoh *et al.* 2011). In most studies, VMAT was a novel treatment technique, compared to previously delivered IMRT. Often, VMAT and IMRT plans were generated by different planners, and possibly under different conditions (e.g. available planning time). In this study, we have systematically compared single- and dual-arc VMAT, and IMRT with 9, 12, and 20 beams for treatment of locally-advanced cervical cancer patients. All plans were generated in a semi-automatic procedure. For each patient, the basis for the production of the five semi-automatic plans was a 20-beam equi-angular IMRT plan, automatically generated with the in-house Erasmus-iCycle. The output (i.e. achieved plan parameters) of this Erasmus-iCycle plan was used as a guide for manual construction of the two VMAT and three IMRT plans in the clinical TPS. To the best of our knowledge this is the first comprehensive study presenting for a single treatment site a comparison of clinically deliverable plans (i.e. generated with a clinical TPS) with variation of both the number of IMRT beams and the number of VMAT arcs. Previous studies by Bortfeld (2010), and Quan *et al.* (2012a) were based on plans generated with non-clinical optimization algorithms. Guckenberger *et al.* (2009) investigated VMAT with multiple arcs for complex targets compared to one clinical IMRT plan.

6
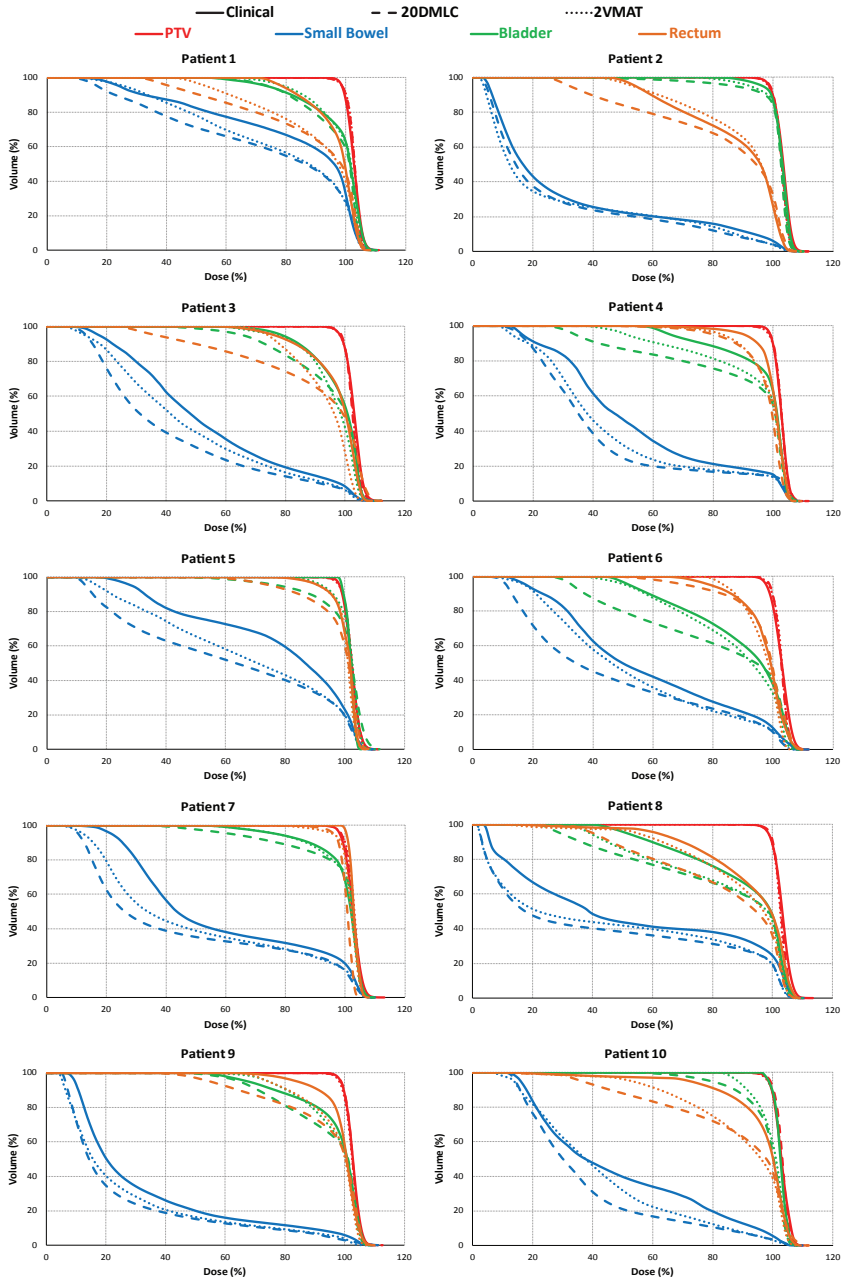
Figure 6.3: DVHs for PTV (red), SB (blue), bladder (green) and rectum (orange) for all study patients for clinical (solid), 20DMLC (dashed), and 2VMAT (dotted) plans.
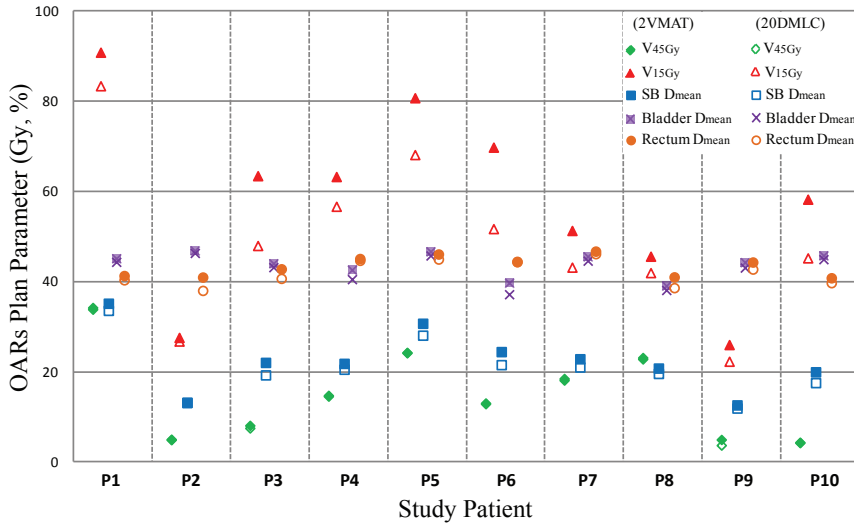
Figure 6.4: For each of the 10 patients in this study (P1-P10 along the horizontal axis), OAR plan parameters are presented for dual-arc VMAT (2VMAT) and 20-beam IMRT (20DMLC).

Instead of generating a 20-beam equi-angular IMRT plan, Erasmus-iCycle could also have been used to generate for each patient a 20-beam plan with an individualized beam setup. However, a study for a sub-set of patients did not show significant plan improvements for individualized beam arrangements, while calculation times went up. The same study showed no significant improvements for plans with more than 20 beams. The presented hands-on planning times in Section 6.3 demonstrate that construction of IMRT and VMAT plans in the clinical TPS, based on an automatically generated 20-beam Erasmus-iCycle plan, is a time-efficient method that sets the workload to minimal by the automatic generation of Monaco® cost-functions.

In a previous prospective clinical study on head-and-neck cancer we have demonstrated clear superiority of the semi-automatic procedure over the general trial-and-error treatment planning by dosimetrists (Voet *et al.* 2013b). In this study, we have confirmed the superiority of semi-automatic planning for cervical cancer patients in a retrospective comparison; for each patient, all five semi-automatically generated plans had better OAR sparing than the clinically applied 9-beam IMRT plan. Plan quality increased when the number of beams (IMRT) or arcs (VMAT) went up. This can be attributed to increases in the degrees of freedom for generating high quality plans. Interestingly, 12DMLC and 20DMLC plans had better OAR sparing than dual- and single-arc VMAT. Detailed analyses demonstrated that there were substantial variations among patients in the OAR sparing advantage of 20DMLC compared to 2VMAT. Possibly, the lower number of directions with IMRT compared to VMAT is compensated by an

increase in beam modulation. To our best knowledge, only Cozzi *et al.* (2008) have previously compared VMAT and IMRT plans for treatment of cervical cancer. The involved clinically delivered IMRT plans had 5 equi-angularly spaced beams. For VMAT, only single-arc dose delivery was considered. They found superior plan quality with VMAT compared to IMRT. VMAT improved dose homogeneity and sparing of OARs in the medium- to high-dose region. For a sub-set of patients, we also generated 5-beam IMRT plans and also observed clearly inferior plan quality compared to both single- and dual-arc VMAT (data not provided), confirming the findings of Cozzi *et al.* (2008).

In a theoretical study, Bortfeld & Webb (2008) have discussed plan quality limitations of single-arc VMAT with very short delivery times. The current study also finds limitations of VMAT, not only for single-arc VMAT, but also for dual-arc. Moreover, the evidence provided in our paper is all obtained from clinically deliverable dose distributions, approved by a physician. In contrast, published planning studies often report better or similar plan quality for single- or dual-arc VMAT compared to IMRT, or advantages for PTV and disadvantages for OARs, or vice versa (Teoh *et al.* 2011).

In line with overwhelming evidence in the literature, we have also observed substantial treatment delivery time reductions with VMAT compared to IMRT. Compared to 20DMLC and 12DMLC, dose delivery times with 2VMAT reduced on average by 11.5 and 8.7 min, respectively. These studies were performed with the MLCi2 multi-leaf collimator. Recently, the Agility™ MLC (Elekta AB) with 160 leaves with very fast motion has shown to reduce VMAT delivery times by 40% when compared to MLCi2 with equivalent plan quality and delivery accuracy (Bedford *et al.* 2013). For a sub-set of 5 patients we observed mean reductions in the predicted treatment delivery times when using the Agility™ MLC instead of MLCi2 of 1.5 min/ 42.9% for single-arc VMAT and 2.1 min/42% for dual-arc VMAT, and 6.5 min/42.5%, 5.7 min/43.9%, and 7.2 min/45%, for IMRT with 9, 12, and 20 beams, respectively. Treatment time reductions have a potential clinical impact in terms of reducing the risk of intra-fraction motion and enhancing patient comfort during treatment. Moreover, short treatment times can increase patient throughput in the clinic and eventually reduce waiting lists and costs. On the other hand, in this paper we have observed for some cervical cancer patients clearly enhanced plan quality for IMRT compared to VMAT, especially when 20 IMRT beams are applied. Based on the results of this study, the clinical decision has been made to start treating cervical cancer patients with dual-arc VMAT. This decision will be re-considered when the fast Agility™ MLCs will become available in our department. We are currently developing a system for fully automatic generation of a set of IMRT and VMAT plans for each individual patient to evaluate on a per patient basis plan quality differences versus delivery time differences. These studies will include modeling of intra-fraction motion.

In this study, we have investigated IMRT delivered with dynamic multi-leaf collimation. From a theoretical point of view, this approach could result in a higher plan

quality than step-and-shoot IMRT because of the increased degree of freedom. It could however also result in somewhat longer treatment times. In the end this all depends on implementation in the (commercial) TPS software. Here we have chosen for the dynamic IMRT technique as we mainly focused on plan quality.

A potential limitation of this study is that plans were generated semi-automatically, possibly resulting in some 'noise' in our results. In a recent study, we have developed a system for fully automatic VMAT plan generation for prostate cancer patients, based on an Erasmus-iCycle pre-optimization (Voet *et al.* 2014). Currently, efforts are on-going to also realize this for cervical cancer patients and to extend the method for static gantry IMRT.

6.5   Conclusion

For locally-advanced cervical cancer patients we have systematically compared single- and dual-arc VMAT with 9, 12, and 20 beam IMRT. IMRT plan quality increased when going from 9 to 12 to 20 beams, and dual-arc VMAT was superior to single-arc. For all patients, the 20-beam IMRT plan had the overall highest quality. On the other hand, the VMAT plans had lowest treatment delivery times. For individual patients, the optimal delivery technique is dependent on a complex trade-off between plan quality and treatment time.

6

VMAT Plus a Few Computer-Optimized Non-Coplanar IMRT Beams 'VMAT+' Tested for Liver SBRT

Abdul Wahab M. Sharfo, Maarten L.P. Dirkx, Sebastiaan Breedveld, Alejandra Méndez Romero and Ben J.M. Heijmen

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

Abstract

Purpose: To propose a novel treatment approach, designated VMAT+, involving addition of < 5 IMRT beams with computer-optimized non-coplanar orientations to VMAT, and evaluate it for liver Stereotactic Body Radiation Therapy (SBRT). VMAT+ is investigated as an alternative for *i)* coplanar VMAT and *ii)* multi-beam non-coplanar treatment.

Methods/Material: For fifteen patients with liver metastases, VMAT+ plans were compared with *i)* dual-arc VMAT and *ii)* 25-beam, non-coplanar treatment with computer-optimized beam orientations (25-NCP). All plans were generated fully automatically for delivery of the highest feasible tumor Biologically Effective Dose (BED). OAR doses, intermediate-dose-spillage, dose-compactness, and measured delivery times were evaluated.

Results: With VMAT the maximum achievable tumor BED was equal to that of 25-NCP. Conversely, VMAT resulted in a lower tumor BED in 5 patients. Compared to VMAT, VMAT+ yielded significant dose reductions in OARs. Intermediate-dose-spillage and dose-compactness were significantly improved by 9.8% and 17.3% ($p \leq .002$), respectively. Treatment times with VMAT+ were only enhanced by 4.1 min on average, compared to VMAT (8.4 min). Improvements in OAR sparing with 25-NCP, compared to VMAT+, were generally modest and/or statistically insignificant, while delivery times were on average 20.5 min longer.

Conclusions: For liver SBRT, VMAT+ is equivalent to time-consuming treatment with 25 non-coplanar beams in terms of achievable tumor BED. Compared to VMAT, OAR sparing and intermediate-dose-spillage are significantly improved, with minor increase in delivery time.

7.1   Introduction

There is growing interest in the use of Stereotactic Body Radiation Therapy (SBRT) as a treatment option for liver tumors. In contrast to conventionally fractionated radiotherapy, SBRT involves delivery of a prescribed high dose to the target in either a single or a small number of treatment fractions, thereby delivering a much higher Biological Effective Dose (BED) (Herfarth *et al.* 2001; Schefter *et al.* 2005; Méndez Romero *et al.* 2006; Rusthoven *et al.* 2009; Goodman *et al.* 2010; van der Pool *et al.* 2010; Chang *et al.* 2011; Liu *et al.* 2013). SBRT provides high local control rates for metastatic liver lesions with small toxicity (Herfarth *et al.* 2001; Schefter *et al.* 2005; Méndez Romero *et al.* 2006; Rusthoven *et al.* 2009; Goodman *et al.* 2010; van der Pool *et al.* 2010; Chang *et al.* 2011; Liu *et al.* 2013). Dose escalation using a smaller number of treatment fractions, resulting in a higher tumor BED, has shown a benefit (Schefter *et al.* 2005; Rusthoven *et al.* 2009; Goodman *et al.* 2010).

To achieve high dose conformality in liver SBRT, non-coplanar beam arrangements have been promoted to minimize the dose in normal tissues compared to coplanar techniques (de Pooter *et al.* 2006, 2008; Dong *et al.* 2013; Woods *et al.* 2016). Coplanar volumetric-modulated arc therapy (VMAT) is also used, with the advantage of a much shorter treatment delivery time (Wang *et al.* 2014).

In this paper we propose a novel treatment approach, designated VMAT+, combining fast coplanar dual-arc VMAT with a few (< 5) IMRT beams with computer-optimized non-coplanar beam angles. Hypothesized advantages of VMAT+ are *i)* higher plan quality than for regular dual-arc VMAT with limited increase in treatment delivery time, and *ii)* faster delivery than a non-coplanar treatment with many (25) IMRT beams with computer-optimized directions (25-NCP) with similar plan quality. These hypotheses were investigated for liver SBRT. To this purpose, deliverable VMAT, VMAT+, and 25-NCP plans were generated fully automatically (Breedveld *et al.* 2012; Voet *et al.* 2012, 2014) for 15 study patients for comparisons of plan quality and treatment time.

7.2   Materials and methods

7.2.1   Patient data

Fifteen patients with a single liver metastasis were included in this simulation study. The planning target volume (PTV) ranged from 34.6 cm^3 to 167.0 cm^3, and the healthy liver volume, defined as total liver minus gross tumor volume (GTV), ranged from 1076 cm^3 to 2306 cm^3 (see Table 7.1). In line with the Dutch consensus guideline for SBRT of liver metastases, if feasible, VMAT, VMAT+ and 25-NCP plans were generated for delivery of 3 fractions of 20 Gy (BED = 180 Gy, assuming a tumor α/β-ratio of 10 Gy), prescribed to

Table 7.1: Patient characteristics and the maximum achieved tumor BEDs for the studied treatment techniques.

Patient	Age	Gender	PTV (cm³)	Healthy liver volume (cm³)	Achieved tumor BED (Gy) 25-NCP/ VMAT+	VMAT
1	86	M	56.3	1361	180	180
2	44	M	132.6	1515	180	180
3	72	M	83.2	1142	180	180
4	56	F	34.6	1334	180	180
5	78	M	90.1	1153	180	180
6	77	M	121.6	1597	180	180
7	67	M	74.6	1559	180	180
8	60	M	141.5	1930	180	180
9	74	M	41.1	1464	105	105
10	37	F	66.2	1209	105	105
11	59	M	157.7	2306	180	132
12	69	F	63.7	1617	180	132
13	66	M	136.7	1718	180	132
14	77	M	97.4	1328	180	132
15	82	F	167.0	1076	132	105

Abbreviations: PTV = *Planning Target Volume*, 25-NCP = 25-beam non-coplanar IMRT plans, VMAT = *Volumetric-Modulated Arc Therapy*, VMAT+ = VMAT supplemented with computer-optimized non-coplanar IMRT beams.

Table 7.2: Applied OAR clinical constraints in each fractionation regime according to the national treatment planning consensus guidelines for liver metastases.

Fractionation regimen	3 × 20 Gy BED = 180 Gy	5 × 12 Gy BED = 132 Gy	8 × 7.5 Gy BED = 105 Gy
Healthy liver	$V_{15Gy} \geq 700\,cm^3$ *	$V_{18Gy} \geq 700\,cm^3$ *	$V_{21.6Gy} \geq 700\,cm^3$ *
Spinal cord	$D_{max} \leq 18$ Gy	$D_{max} \leq 22.5$ Gy	$D_{max} \leq 27.2$ Gy
Esophagus	$D_{max} \leq 27$ Gy	$D_{max} \leq 33$ Gy	$D_{max} \leq 40$ Gy
Kidneys	$V_{15Gy} \leq 33\%$	$V_{18Gy} \leq 33\%$	$V_{21.6Gy} \leq 33\%$
Duodenum, Bowel &	$D_{max} \leq 30$ Gy	$D_{max} \leq 36.5$ Gy	$D_{max} \leq 44$ Gy
Stomach	$V_{22.5Gy} \leq 5\,cm^3$	$V_{26Gy} \leq 5\,cm^3$	$V_{32.8Gy} \leq 5\,cm^3$

Abbreviations: OAR = Organ-at-Risk, BED = *Biologically Effective Dose*.
* Volume to receive the indicated dose or less should be at least 700 cm³.

the 80% isodose line. At least 95% of the PTV should be covered by the prescription dose, while adhering to the imposed constraints for the healthy liver, spinal cord, esophagus, kidneys, duodenum, bowel and stomach (Table 7.2). If sufficient target coverage with 3 × 20 Gy was not feasible, a plan was generated to deliver the total dose of 60 Gy in either 5 fractions of 12 Gy (BED = 132 Gy), or 8 fractions of 7.5 Gy (BED = 105 Gy), allowing higher OAR planning constraints, as calculated with an OAR α/β-ratio of 3 Gy (Table 7.2).

7.2.2 Automated plan generation

All plans in this study were generated in a two-step process. First, a plan was created using our in–house developed multi-criterial optimizer, Erasmus-iCycle (Breedveld *et al.* 2012). Based on this plan, a patient-specific plan-template for the Monaco® treatment planning system version 5.0 (Elekta AB, Stockholm, Sweden) was generated. Next, Monaco® was used to generate a clinically deliverable plan that mimicked the Erasmus-iCycle plan. The whole procedure is fully automated, i.e., no manual

interference is needed. This system is in routine clinical use for prostate cancer VMAT (Voet *et al.* 2014), head-and-neck cancer IMRT and VMAT (Voet *et al.* 2012), cervical cancer VMAT (Sharfo *et al.* 2015, 2016; Heijkoop *et al.* 2016) and advanced-lung-cancer VMAT. In several studies we have demonstrated that plan quality with automated planning is equivalent or superior to manual planning (Voet *et al.* 2012, 2014; Sharfo *et al.* 2015, 2016; Heijkoop *et al.* 2016).

Automated multi-criterial plan generation with Erasmus-iCycle has been described extensively elsewhere (Breedveld *et al.* 2012). Therefore, we only briefly summarize the aspects specific to this study. Treatment plan generation with Erasmus-iCycle is based on a tumor site specific '*wish-list*' containing hard constraints to be strictly obeyed and plan objectives with ascribed priorities, and a set of user defined candidate beam orientations. The objective priorities are used for the multi-criterial plan generation, making trade-offs between all objectives in line with our clinical wishes. For each fractionation regimen, a single wish-list (Table 7.3) was used to generate the VMAT, VMAT+, and 25-NCP plans for all patients, facilitating objective plan comparisons. Apart from beam fluence optimization, Erasmus-iCycle also features automated coplanar and non-coplanar beam angle optimization, based on sequential beam addition (Breedveld *et al.* 2012; Voet *et al.* 2012, 2013a; Rossi *et al.* 2012, 2015), which was used in this study to generate 25-NCP and VMAT+ plans. The non-coplanar beam selection search-space in Erasmus-iCycle consisted of 302 candidate beams separated by about 10 degrees and homogeneously distributed across the part of the sphere where collision between the patient/couch and the gantry were avoided. For VMAT+ plan generation, Erasmus-iCycle first generated a multi-beam coplanar plan to simulate VMAT. Then, optimal non-coplanar beams were sequentially added, every time followed by a full fluence re-optimization, i.e., including the coplanar beams and all selected non-coplanar beams.

All plans were generated for an Elekta Synergy® treatment machine equipped with an MLCi2 collimator with a leaf width of 10 mm. A 10 MV photon beam energy was used. Dose calculations in Monaco® were performed with a dose grid resolution of 3 mm. For VMAT, a maximum of 200 control points per arc were used, while using 30 control points per IMRT beam, delivered using dynamic multi-leaf collimation. A 5 degree collimator angle was used for all arcs/beams, and the maximum dose rate was 600 MU/min.

7.2.3   VMAT+ for enhanced tumor BED

As described in detail in Section 7.3, for some patients, the maximum achievable tumor BED with VMAT was lower than for 25-NCP. For these patients, we investigated whether VMAT+ could be used to enhance the BED to the level of 25-NCP.

Table 7.3: Wish-list for automatic plan generation.

Constraints

Volume	Type	Limit
PTV, GTV	Max	120% of D^p
PTV Shell 5 mm	Max	93% of D^p
Spinal Cord	Max	$15^*, 20^\dagger, 25^\ddagger$ Gy
Esophagus	Max	$25^*, 30^\dagger, 34^\ddagger$ Gy
Stomach, Bowel, Duodenum	Max	$27^*, 34^\dagger, 39^\ddagger$ Gy
Healthy Liver	Dose-Volume	$D_{700cm^3} < 15^*, 18^\dagger, 21.6^\ddagger$ Gy

Objectives

Priority	Volume	Type	Goal	Parameters
1	GTV	↓ LTCP	0.1	$D^p = 60$ Gy, $\alpha = 0.3$
	PTV	↓ LTCP	0.1	$D^p = 60$ Gy, $\alpha = 0.4$
2	Duodenum	↓ EUD	$19.5^*, 24^\dagger, 28^\ddagger$ Gy	$k = 20$
3	Bowel, Stomach	↓ EUD	$19.5^*, 24^\dagger, 28^\ddagger$ Gy	$k = 20$
4	Esophagus	↓ EUD	$22.5^*, 28^\dagger, 30^\ddagger$ Gy	$k = 20$
	Spinal Cord	↓ EUD	10 Gy	$k = 20$
5	PTV Shell 10 mm	↓ Max	75% of D^p	
	Patient – Ring	↓ Max	$15^*, 18^\dagger, 21.6^\ddagger$ Gy	60 mm
6	PTV Shell 20 mm	↓ Max	50% of D^p	
7	PTV Shell 30 mm	↓ Max	40% of D^p	
	Skin Ring 30 mm	↓ Max	35% of D^p	
8	Healthy Liver	↓ Mean	10 Gy	
9	Kidneys	↓ EUD	10 Gy	$k = 15$
10	PTV Shell 40 mm	↓ Max	30% of D^p	
	PTV Shell 50 mm	↓ Max	25% of D^p	
11	Duodenum, Bowel, Stomach	↓ Mean	10 Gy	

Abbreviations: D^p = prescribed dose (i.e., 60 Gy), LTCP = *Logarithmic Tumor Control Probability*, EUD = Equivalent Uniform Dose, α = cell sensitivity parameter to achieve adequate target coverage, k = volume parameter.
$*, \dagger, \ddagger$: Applied values for the 3, 5 and 8 fractionation regimen, respectively.

7.2.4 Comparison of VMAT+3 and VMAT for fixed tumor BED

For all patients, the plan quality of VMAT+ and VMAT was compared for a fixed tumor BED per patient. To this purpose, for each patient, a VMAT+3 plan (VMAT + 3 optimized non-coplanar IMRT beams) was generated for the highest achievable BED with VMAT.

7.2.5 Comparison of VMAT+3 and 25-NCP for fixed tumor BED

To compare VMAT+ with 25-NCP, for each patient, a VMAT+3 plan was generated for the maximum achievable tumor BED for 25-NCP.

7.2.6 Plan comparison, normalization, and evaluation

In the above described plan quality comparisons between VMAT+3 and VMAT, and VMAT+3 and 25-NCP, the aim was to always compare plans with exactly equal PTV coverage. Ideally each plan had a coverage of 95%, i.e. $V_{60Gy} = 95\%$ of PTV. In practice, small coverage differences between plans (0.1% ± 1.1% (1SD)) did occur. In these cases, generated plans were first normalized to achieve identical PTV dose coverage. Only for patients 9 and 10, the plans were normalized to $V_{60Gy} = 88\%$ of the PTV,

because otherwise the maximum dose constraint for respectively the esophagus and stomach would be exceeded for all treatment delivery techniques. Differences in mean PTV dose (D_{mean}), near-minimum PTV dose ($D_{98\%}$), near-maximum PTV dose ($D_{2\%}$), conformation number ($CI = (TV_{RI})^2/(TV * V_{RI})$, i.e., ratio of the target volume covered by the reference isodose (TV_{RI}) to the target volume (TV) and volume of the reference isodose (V_{RI}) (Feuvret *et al.* 2006), and near-maximum doses ($D_{2\%}$) in the OARs were quantified. For the healthy liver we also evaluated D_{mean}, D_{700cm^3} (maximum dose in the 700 cm^3 of healthy liver receiving the lowest doses), and the healthy liver volume receiving a dose less than the corresponding clinical constraint x Gy (V_{xGy}, e.g., V_{15Gy} for the 3-fraction regimen, see Table 7.2). To quantify the dose gradient and compactness outside the PTV, the maximum dose at 2 cm from the PTV surface (D_{2cm}), the intermediate-dose-spillage ($R_{50\%} = V_{50\%}/V_{PTV}$, i.e., ratio of patient volume encompassed by the 50% isodose line and the PTV), and the volume encompassed by the 25% isodose line ($V_{25\%}$) were evaluated. Visual plan inspections were performed to ensure a conformal dose distribution and the absence of any undesirable hot-spots in the unspecified tissues. Paired two-sided Wilcoxon signed-rank tests were performed to compare the different planning strategies. Differences were considered statistically significant for $p < .05$.

7.2.7 Delivery time comparisons, dosimetric QA, and degree of modulation

Generated VMAT, VMAT+3, and 25-NCP plans were delivered at an Elekta Synergy® linac (Elekta AB, Sweden) for delivery time comparisons and dosimetric QA.

Non-coplanar beams were delivered in an order that ensured minimal delivery time. A stopwatch was used to measure *i)* for each beam the beam-on-time (time elapsed from pushing the beam-on button till delivery of the last MU) *ii)* gantry-travel-times (time elapsed to rotate the gantry from a certain fixed angle to another while the beam is off; gantry motion was steered from the control room), and *iii)* couch-travel-times (time required to enter the treatment room and rotate the treatment couch to its next position).

For dosimetric QA, the VMAT, VMAT+3, and 25-NCP plans were delivered while irradiating a PTW 2D-Array seven29™ and Octavius™ phantom (PTW, Freiburg, Germany). The measurements were compared to TPS predictions using a commercial QA software package (PTW VeriSoft® version 6.2) with 3% global maximum dose and 3 mm distance to agreement (3%/3 mm) and 2%/2 mm as reference criteria for γ-analyses. A Gamma passing criterion of 90% was used (Low *et al.* 1998; Ezzell *et al.* 2009).

Apart from the dosimetric QA, VMAT, VMAT+ and 25-NCP plans were also compared regarding the total number of MU and the degree of modulation, defined as the sum

Table 7.4: OAR constraints for patients 11 – 15 prohibiting a VMAT tumor BED equal to that of 25-NCP, the minimum number of NCP beams in VMAT+ to lift the limitation, and their total weights.

Patient	Limiting constraint	# NCP beams to lift limitation	Total weights of NCP beams (%)
11	Bowel $V_{22.55Gy} < 5$ cm^3	3	47.7
12	Esophagus $D_{max} < 27$ Gy	1	14.8
13	Bowel $D_{max} < 30$ Gy	2	39.9
14	Stomach $V_{22.55Gy} < 5$ cm^3	3	45.8
15	Healthy liver $V_{18Gy} > 700$ cm^3	4	46.1

Abbreviations: OAR = Organ-at-Risk, VMAT = *Volumetric-Modulated Arc Therapy*, NCP = non-coplanar.

of the MU of all segments divided by the sum over all segments of ((segment area × segment MU)/total beam area), as reported for each plan by the Monaco® TPS.

Patient group averages of the percentage Gamma passing rate, total MU, and degree of modulation of VMAT, VMAT+3 and 25-NCP plans were compared using paired two-sided Wilcoxon signed-rank tests.

7.3 Results

7.3.1 VMAT+ for enhanced tumor BED

Table 7.1 (last 2 columns) shows for all investigated treatment techniques the maximum achievable tumor BED. For each patient, the maximum feasible tumor BED for 25-NCP could also be realized with VMAT+ (last but one column in Table 7.1). For patients 1–10, this BED could also be obtained with VMAT, while this was not possible for patients 11–15 (last column). By adding only 1, 2, 3 (2 patients) or 4 individualized non-coplanar IMRT beams to VMAT, the tumor BED for these 5 patients could be enhanced to become equal to 25-NCP (Table 7.4). The total weight of the 1–4 non-coplanar beams, defined as the percentage contribution to the mean PTV dose was on average 39% (range: 15–48%).

7.3.2 Comparison of VMAT+3 and VMAT for fixed tumor BED

Figure 7.1, Figure 7.2, and Table 7.5 show superior plan quality for VMAT+3 compared to VMAT. While PTV dose delivery was similar, both OAR sparing, dose compactness and intermediate-dose-spillage were improved. Significant improvements were observed for the healthy liver, esophagus, spinal cord, and stomach. The healthy liver volume receiving a dose less than x Gy was increased by on average (81 cm^3 ± 76 cm^3 (1 SD), p = .001). Intermediate-dose-spillage as calculated by $R_{50\%}$ and $V_{25\%}$ were improved by 9.8% and 17.3% ($p \leq .002$), respectively. The total weight of the 3 non-coplanar beams in the VMAT+3 plans was on average 35.0% ± 7.9% (1 SD). As is evident from Figure 7.1 and
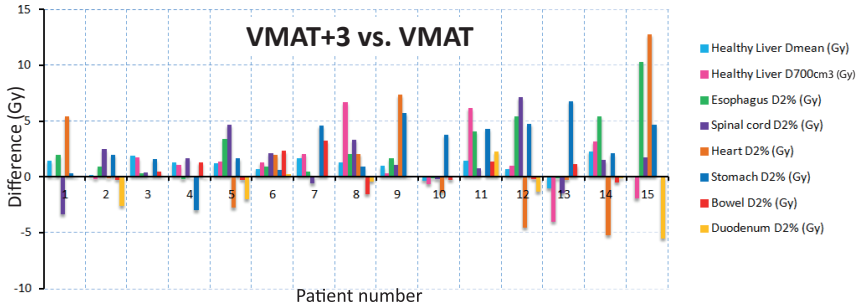
Figure 7.1: Differences in dosimetric plan parameters between VMAT+3 and VMAT plans for individual patients. Positive values are in favor of VMAT+3. D_{700cm^3} = maximum dose in the 700 cm^3 of the healthy liver receiving the lowest doses.
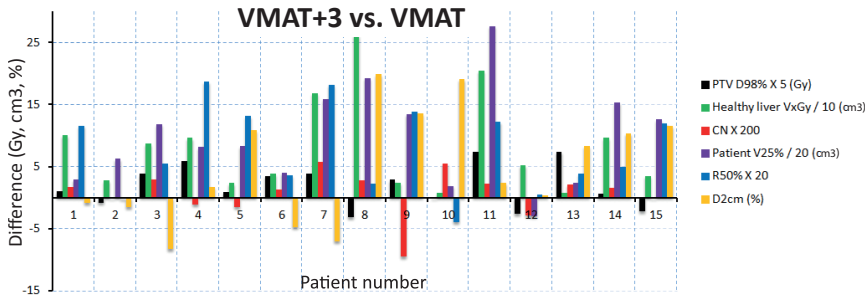
Figure 7.2: Differences in dosimetric plan parameters between VMAT+3 and VMAT plans for individual patients. Positive values are in favor of VMAT+3 plans. V_{xGy} = healthy liver volume receiving a dose less than x Gy (x = 15, 18, or 21.6 for BED = 180, 132, and 105 Gy, respectively, see Table 7.2), CN = *Conformation Number*, $R_{50\%} = V_{50\%}/V_{PTV}$ = intermediate-dose-spillage. D_{2cm} = maximum dose at 2 cm from the PTV.

Figure 7.2, some patients benefited more from VMAT+3 than others. Figure 7.3 shows for patient 11, for the 5 fraction regime, the improvement in dose compactness and dose conformality when using VMAT+3 compared to VMAT.

7.3.3 Comparison of VMAT+3 and 25-NCP for fixed tumor BED

For patients 1–10, the achieved tumor BED of the VMAT+3 plans, used in the above comparisons with VMAT, was equal to the BED for 25-NCP. Therefore, these plans were also used to compare VMAT+3 with 25-NCP. For patients 11–14, a new VMAT+3 plan was generated to match the 25-NCP tumor BED. For patient 15, this BED could only be obtained with a VMAT+4 plan (see also Table 7.4). VMAT+3 (VMAT+4 for patient 15) and 25-NCP are dosimetrically compared in Table 7.6. On average, the total weight of the 3 non-coplanar beams in VMAT+3 was 37% ± 9%. No statistically significant

Table 7.5: Comparison of dosimetric parameters of VMAT+3 and VMAT plans. For each patient, plans were generated for the same tumor BED (i.e., the highest achievable with VMAT, Table 7.1). Positive mean differences hint at an advantage for VMAT+3.

Structure	Parameter	VMAT	VMAT+3 vs. VMAT	
		Mean ± SD	Mean diff. ± SD	*p-value*
PTV	D_{mean} (Gy)	65.9 ± 0.9	0.2 ± 0.9	.09
	$D_{98\%}$ (Gy)	55.0 ± 7.4	0.4 ± 0.7	.07
	$D_{2\%}$ (Gy)	72.0 ± 1.5	0.4 ± 1.5	.06
	CN	0.83 ± 0.07	0.00 ± 0.02	.2
Healthy liver	D_{mean} (Gy)	13.7 ± 2.9	0.9 ± 0.9	**.006**
	D_{700cm^3} (Gy)	8.6 ± 6.0	1.2 ± 2.7	.08
	V_{xGy} (cm^3)	1019 ± 224	81 ± 76	**.001**
Esophagus	$D_{2\%}$ (Gy)	12.6 ± 9.0	2.4 ± 2.9	**.002**
Spinal cord	$D_{2\%}$ (Gy)	8.9 ± 3.7	1.4 ± 2.5	**.04**
Heart	$D_{2\%}$ (Gy)	12.8 ± 12.1	1.1 ± 4.8	.6
Stomach	$D_{2\%}$ (Gy)	13.4 ± 7.1	2.7 ± 2.5	**.003**
Bowel	$D_{2\%}$ (Gy)	9.3 ± 8.6	0.5 ± 1.3	.4
Duodenum	$D_{2\%}$ (Gy)	4.2 ± 6.5	-1.0 ± 2.1	.2
Patient	D_{2cm} (%)	71.3 ± 9.4	5.0 ± 9.0	.06
	$V_{25\%}$ (cm^3)	1026 ± 400	195 ± 159	**.001**
	$R_{50\%}$	3.6 ± 1.1	0.4 ± 0.3	**.002**

Abbreviations: Mean diff. = mean difference, SD = *Standard Deviation*, VMAT+3 = VMAT supplemented with 3 computer-optimized non-coplanar IMRT beams, PTV = *Planning Target Volume*, $D_{98\%}$ = near minimum dose, $D_{2\%}$ = near maximum dose, CN = *Conformity Number*, D_{700cm^3} = maximum dose in the 700 cm^3 of the healthy liver receiving the lowest doses, V_{xGy} = healthy liver volume receiving a dose less than x Gy (x = 15, 18, or 21.6 for BED = 180, 132, and 105 Gy, respectively, see Table 7.2), D_{2cm} = maximum dose at 2 cm from the PTV, $R_{50\%} = V_{50\%}/V_{PTV}$ = intermediate-dose-spillage.
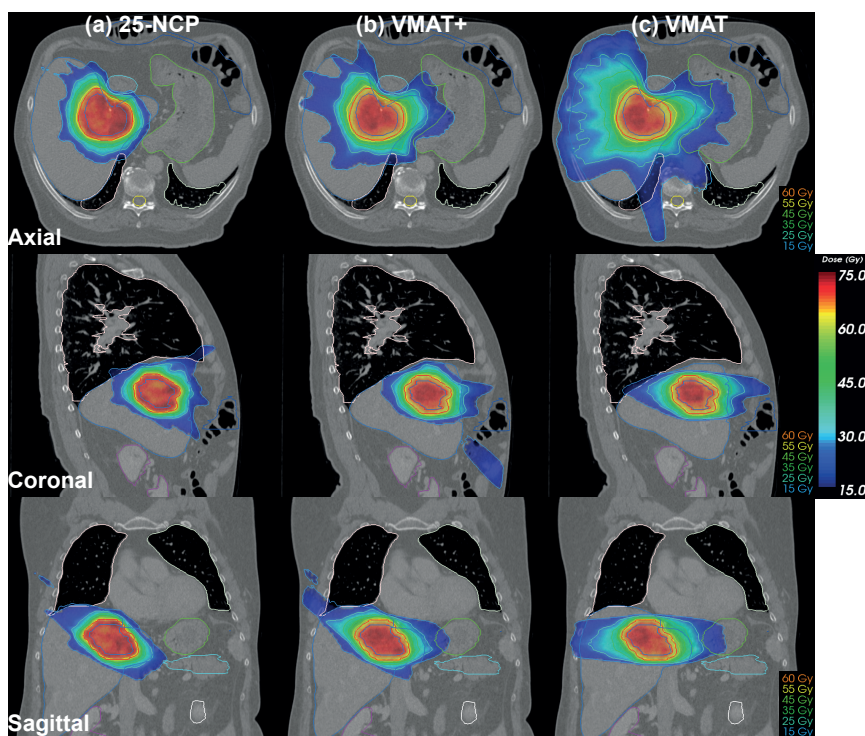
7

7

Figure 7.3: Dose distributions of 25-NCP, VMAT+3, and VMAT plans for patient 11 for the 5-fraction regimen, clearly showing better intermediate-dose conformality for VMAT+3 than for VMAT.

Table 7.6: Dosimetric comparisons of 25-NCP with VMAT+3. For each patient, the tumor BED was the same in both plans, i.e. the highest achievable with 25-NCP (Table 7.1). Positive mean differences hint at an advantage for VMAT+3.

Structure	Parameter	25-NCP	VMAT+3 vs. 25-NCP	
		Mean ± SD	Mean diff. ± SD	*p-value*
PTV	D_{mean} (Gy)	66.1 ± 0.9	0.0 ± 0.8	.8
	$D_{98\%}$ (Gy)	55.5 ± 6.1	-0.7 ± 1.6	.05
	$D_{2\%}$ (Gy)	72.0 ± 1.0	-0.2 ± 1.0	.5
	CN	0.84 ± 0.05	-0.01 ± 0.01	**.01**
Healthy liver	D_{mean} (Gy)	12.0 ± 2.7	-0.8 ± 1.1	**.01**
	D_{700cm^3} (Gy)	7.0 ± 3.6	-0.5 ± 1.6	.3
	V_{xGy} (cm^3)	1119 ± 211	-50 ± 41	**.002**
Esophagus	$D_{2\%}$ (Gy)	9.1 ± 5.9	-0.4 ± 2.2	.5
Spinal cord	$D_{2\%}$ (Gy)	6.4 ± 3.4	-0.8 ± 2.4	.3
Heart	$D_{2\%}$ (Gy)	10.6 ± 8.8	-1.1 ± 5.4	.9
Stomach	$D_{2\%}$ (Gy)	8.7 ± 6.2	-1.9 ± 2.6	**.03**
Bowel	$D_{2\%}$ (Gy)	9.1 ± 5.9	0.5 ± 4.2	0.2
Duodenum	$D_{2\%}$ (Gy)	6.5 ± 4.7	2.1 ± 2.6	**.02**
Patient	D_{2cm} (%)	55.2 ± 7.4	-12.1 ± 7.6	**.001**
	$V_{25\%}$ (cm^3)	673 ± 270	-154 ± 52	**.001**
	$R_{50\%}$	2.8 ± 0.8	-0.5 ± 0.3	**.001**

Abbreviations: Mean diff. = mean difference, SD = *Standard Deviation*, VMAT+3 = VMAT supplemented with 3 computer-optimized non-coplanar IMRT beams, 25-NCP = treatment with 25 computer-optimized non-coplanar IMRT beams, PTV = *Planning Target Volume*, $D_{98\%}$ = near minimum dose, $D_{2\%}$ = near maximum dose, CN = *Conformity Number*, D_{700cm^3} = maximum dose in the 700 cm^3 of the healthy liver receiving the lowest doses, V_{xGy} = healthy liver volume receiving a dose less than x Gy (x = 15, 18, or 21.6 for BED = 180, 132, and 105 Gy, respectively, see Table 7.2), D_{2cm} = maximum dose at 2 cm from the PTV, $R_{50\%}$ = $V_{50\%}/V_{PTV}$ = intermediate-dose-spillage.

differences in PTV parameters were observed, except for a small difference in CN (Table 7.6). Improvements in OAR sparing with 25-NCP compared to VMAT+3 were generally modest and/or not statistically significant. The largest advantages of 25-NCP were observed for intermediate-dose-spillage, showing a reduction in $V_{25\%}$ of 24.4% (p = .001) and in $R_{50\%}$ of 16.7% (p = .001), and the maximum dose at 2 cm from the PTV (D_{2cm}) with a reduction of 23.6% (p = .001). The enhanced dose compactness for 25-NCP compared to VMAT+ is also apparent from the dose distributions in Figure 7.3.

7.3.4 Delivery time comparisons, dosimetric QA, and degree of modulation

For each patient, VMAT, VMAT+3 and 25-NCP plans with the highest possible tumor BED feasible for VMAT (Table 7.1, last column) were delivered on the treatment machine for delivery time comparisons and dosimetric QA.

The number of monitor units, beam-on-times, gantry-travel-times, couch-travel-times and total delivery times are shown for each patient in Figure 7.4. The average total delivery times for VMAT, VMAT+ and 25-NCP plans were 8.4 ± 2.3 min (1SD), 12.5 ± 3.8 min, and 33.1 ± 4.7 min, respectively. While for VMAT+3 a relatively small (4.1 min on average) increase in total delivery time was observed compared to VMAT, treatment times of 25-NCP were on average prolonged by as much as 20.5 min compared to VMAT+3. VMAT+3 plans required on average 5292 ± 2028 MU, compared to 4541 ± 1421 MU for VMAT (p = .005) and 4612 ± 2131 for 25-NCP (p = .011).

The QA measurements showed an average γ-passing rate of 99.9% ± 0.1% [99.5%, 100%] using 3% global maximum dose and 3 mm distance to agreement (VMAT: median = 100% (95% confidence interval (CI), 99.9 – 100%), VMAT+3: median = 100% (95% CI, 99.9 – 100%), and 25-NCP: median = 100% (95% CI, 99.8 – 100%)). When using 2%/2 mm, the γ-passing rate of all plans was on average 99.4% ± 1.0% [95.9%, 100%] (VMAT: median = 99.7% (95% CI, 99.3 – 99.9%), VMAT+3: median = 100% (95% CI, 98.6 – 100%), 25-NCP: median = 99.5% (95% CI, 98.7 – 99.7%)). No statistically significant differences in Gamma passing rates for VMAT, VMAT+ and 25-NCP were observed, (p > .05, Wilcoxon signed-rank).

The degree of modulation of delivered plans was on average significantly higher for VMAT+3 plans (4.0 ± 0.9), compared to 3.4 ± 1.0 (p = .005) and 3.5 ± 0.7 (p = .006) for VMAT and 25-NCP, respectively. No statistically significant correlation was found between the Gamma passing rates for 3%/3 mm and 2%/2 mm and the modulation degree for any of the studied techniques, (p > .05, Spearman's rho).

7.4   Discussion

For liver SBRT, multiple non-coplanar beams are generally used to achieve high dose conformality and minimize the dose in normal tissues (de Pooter *et al.* 2008; Dong *et al.* 2013; Wang *et al.* 2014; Paik *et al.* 2015). Currently, the greatest concern for using a large number of non-coplanar beams in SBRT is treatment delivery time. Speed of treatment delivery is particularly important in SBRT to minimize the impact of intra-fraction motion. VMAT is a time-efficient delivery technique (Woods *et al.* 2016), but may result in lower plan quality. Dong *et al.* (2013) demonstrated the dosimetrical benefit of a fully non-coplanar treatment technique over VMAT, by applying a novel 4π delivery technique to optimize non-coplanar beam orientation and fluences. To achieve superior plan quality, between 14 and 22 non-coplanar beams were required, at the cost of much longer treatment delivery times. Nevertheless, as shown in a recent paper of Woods *et al.* (2016), the 4π technique allowed for a clinically relevant dose escalation with the intention to improve local tumor control in liver SBRT.
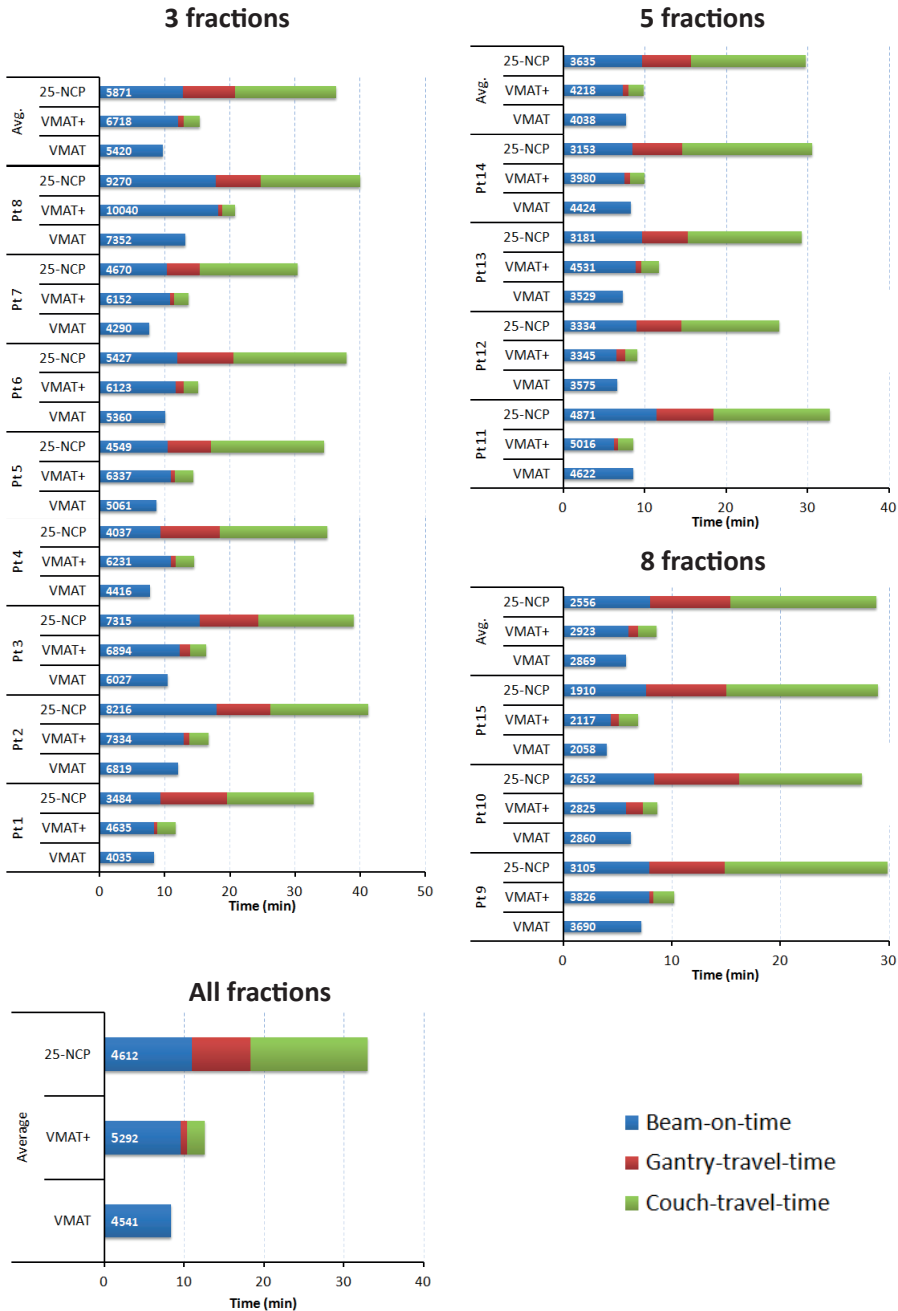
Figure 7.4: For all patients, measured beam-on-times, gantry-travel-times, couch-travel-times, and total delivery times for the VMAT, VMAT+3 and 25-NCP plans. The numbers in the bars represent delivered MUs.

In this study, VMAT+, i.e., dual-arc VMAT combined with 1–5 IMRT beams with computer-optimized, non-coplanar orientations, is proposed as an alternative both for regular dual-arc VMAT and for multi-beam non-coplanar treatment. It was hypothesized that *i)* VMAT+ would result in a higher plan quality than VMAT, with only a small and acceptable increase in delivery time, and *ii)* VMAT+ would have substantially shorter delivery times than multi-beam non-coplanar treatment with no, or only a small and acceptable loss in plan quality. These hypotheses were tested for liver SBRT, using automatically generated treatment plans that were deliverable on a regular treatment unit. VMAT+ indeed resulted in improved healthy tissue sparing compared to VMAT and especially in reduced intermediate-dose-spillage. Due to the improved OAR sparing, the feasible tumor BED went up by 48 Gy for 4 patients and by 27 Gy for 1 patient, in all 5 cases reaching the level of the maximum achievable BED for 25-NCP. The enhanced BEDs with VMAT+ could be obtained by adding only 1–4 non-coplanar beams to VMAT. Compared to VMAT, VMAT+ resulted in a relatively small increase in treatment time (4.1 min). While plan quality improvement of 25-NCP compared to VMAT+3 was modest, the treatment delivery time was 20.5 min longer. Due to the shorter treatment delivery time, VMAT+ has the potential to enhance patient comfort and lower the cost per treatment, as compared to 25-NCP. Currently, efforts are on-going to explore the advantages of VMAT+ for other treatment sites.

The increase in plan quality of VMAT+ compared to VMAT with only small increase in treatment time, and the largely reduced treatment time with only moderately reduced plan quality compared to 25-NCP was related to enhanced numbers of MU and higher degrees of modulation. On the other hand, the dosimetric QA comparisons showed no statistically significant differences between VMAT, VMAT+3 and 25-NCP plans in Gamma passing rates.

In this study, enhanced OAR sparing with the VMAT+ approach has been quantified in a systematic comparison of VMAT with VMAT+3, i.e. for all patients, 3 optimized non-coplanar beams were added. For the same tumor BED, VMAT+3 resulted in a significantly improved healthy tissue sparing, while minimally enhancing delivery time. As described in Section 7.3 and above, for some patients, VMAT+1 or VMAT+2 already allowed for dose escalation compared to regular VMAT, due to enhanced OAR sparing. In this study, we have not systematically investigated potential benefits of VMAT+1 and VMAT+2 in terms of improved healthy tissue sparing compared to VMAT for the same tumor BED with a potentially even more limited increase in treatment time than VMAT+3.

As shown in Figure 7.4, 45% of the total delivery time of 33.1 min for 25-NCP treatments was related to couch travel, largely due to the need of entering the treatment room for treatment couch re-positioning during non-coplanar treatment. With a system to enable re-positioning from the linac control room, delivery times of 25-NCP could be

reduced, but they would unavoidably remain longer than for VMAT and VMAT+ (Note, the latter would also benefit from remote couch re-positioning).

In this paper, the total number of control points in VMAT+ plans was higher than for corresponding VMAT plans, related to the addition of the IMRT beams. Therefore, for VMAT+ the degree of freedom for generating an optimal plan was enhanced. As demonstrated, VMAT+ resulted in clearly improved plan quality with only a small increase in treatment time. Although not explicitly investigated in this study, in our experience, an enhanced number of control points for VMAT would not have resulted in a meaningful improved plan quality.

Several other approaches for combining arc therapy with IMRT have recently been proposed (Robar & Thomas 2012; Matuszak *et al.* 2013; Hoover *et al.* 2015; Zhao *et al.* 2015a,b; MacFarlane *et al.* 2016), but in all of them the IMRT beams were in an arc plane, in contrast with the VMAT+ approach. None of the published studies used fully automated and multi-criterial plan generation, and none investigated the combined approach for liver SBRT. These differences, on top of differences in applied technology and study design hamper direct comparisons with VMAT+. BrainLAB introduced HybridArc™, combining IMRT with dynamic conformal arcs (DCA), instead of VMAT in our approach. In this system, the user should define the relative dose contribution from IMRT beams to the total dose. For complex cranial cases, Robar & Thomas (2012) found that adding 3–15 IMRT beams to dynamic arcs enhanced plan quality, whereas for prostate cancer patients 5 IMRT beams were adequate. For cranial cases, HybridArc™ showed improved dose conformity compared with dynamic conformal arc and IMRT. For prostate cases, dose conformity and homogeneity were improved compared with DCA and IMRT, respectively. Differences in required MU compared to IMRT were site-dependent. Matuszak *et al.* (2013) proposed and validated a strategy for adding in-plane IMRT beams with optimal directions to a single-arc VMAT plan. These in-plane beams were selected based on an analysis of dose gradients in the VMAT plan. Their study demonstrated for a prostate phantom case, a pancreas case and a brain case that additional modulation from 3–5 "*optimal*" beam angles improved plan quality compared with single-arc VMAT. When adding three IMRT beams, the plans required 22.4– 43.7% fewer MU/Gy than a seven beam IMRT plan. Recently, Hoover *et al.* (2015) and MacFarlane *et al.* (2016) reported another method for combining VMAT with in-plane IMRT with optimized directions. Generated plans for head-and-neck cancer, lung cancer, and prostate cancer were superior in quality when compared to stand-alone IMRT or VMAT plans, mainly due to reduced OAR doses. Delivery times were comparable with VMAT. Zhao *et al.* (2015a,b) proposed a combination of VMAT with IMRT, generating a VMAT plan while using an IMRT plan, delivering half of the prescribed dose, as a base plan. Five - nine coplanar IMRT beams were used. Compared to VMAT, the hybrid plans significantly improved target dose homogeneity and conformity as well as OAR sparing

for nasopharyngeal and lung tumors. Delivery times were substantially longer than for VMAT. Compared to IMRT, the difference in delivery time was dependent on treatment site.

This study for liver SBRT showed clear advantages of VMAT+, compared to VMAT and 25-NCP. Important for this study was the availability of our in-house Erasmus-iCycle application for automatic, multi-criterial plan optimization, including non-coplanar beam angle selection. Currently, Erasmus-iCycle is not commercially available. However, Elekta AB is preparing integration in their commercial Monaco® TPS.

7.5   Acknowledgments

7

Complementing Prostate SBRT VMAT with a Two-Beam Non-Coplanar IMRT Class Solution to Enhance Rectum and Bladder Sparing with Minimum Increase in Treatment Time

Abdul Wahab M. Sharfo, Linda Rossi, Shafak Aluwini, Maarten L.P. Dirkx, Sebastiaan Breedveld and Ben J.M. Heijmen

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

Abstract

Background and Purpose: Enhance rectum and bladder sparing in prostate SBRT with minimum increase in treatment time by complementing VMAT with a two-beam non-coplanar IMRT *class solution* (CS).

Materials and Methods: For twenty patients, an optimizer for automated multi-criteria planning with integrated beam angle optimization (BAO) was used to generate VMAT plans, supplemented with five non-coplanar IMRT beams with individualized orientations (VMAT+5). A CS with two most preferred directions in VMAT+5 was then defined for VMAT+CS. VMAT+CS was compared with automatically generated *i)* dual-arc coplanar VMAT plans (VMAT), *ii)* VMAT+5 plans, and *iii)* IMRT plans with 30 patient-specific non-coplanar beam orientations (30-NCP). Differences in PTV doses, OAR sparing, and computation and treatment delivery times were quantified.

Results: Compared to VMAT, VMAT+CS significantly reduced rectum and bladder doses, and the dose bath. Mean relative differences in rectum D_{mean}, D_{1cc}, V_{40GyEq} and V_{60GyEq} were 19.4 ± 10.6%, 4.2 ± 2.7%, 34.9 ± 20.3%, and 39.7 ± 23.2%, respectively (all p < .001). Total delivery times with VMAT+CS only increased by 1.9 ± 0.7 min compared to VMAT (9.1 ± 0.7 min). The dosimetric quality of VMAT+CS plans was equivalent to VMAT+5, while optimization times were reduced by a factor of 25 due to avoidance of BAO. Compared to VMAT+CS, the 30-NCP plans were only favorable in terms of dose bath, at the cost of much enhanced optimization and delivery times.

Conclusions: The proposed two-beam non-coplanar class solution to complement coplanar VMAT resulted in substantial plan quality improvements with minor increases in treatment delivery times. In addition, patient-specific non-coplanar beam angle optimization was superfluous.

8.1 Introduction

Stereotactic body radiation therapy (SBRT) is becoming the standard treatment radiotherapy option for several primary and metastatic tumors (Guckenberger *et al.* 2017; Videtic *et al.* 2017; Schneider *et al.* 2018; Morgan *et al.* 2018; Wilke *et al.* 2019; Glicksman *et al.* 2020; Lee *et al.* 2020; Jang *et al.* 2020). In prostate SBRT, volumetric-modulated arc therapy (VMAT) has been promoted because of its short treatment delivery time (Alongi *et al.* 2013; Lin *et al.* 2014a; Kang *et al.* 2017). On the other hand, several studies have shown that use of non-coplanar beam arrangements minimizes doses in the normal tissues compared to coplanar VMAT, at the cost of enhanced treatment delivery times (Rossi *et al.* 2012; Dong *et al.* 2014; Rossi *et al.* 2018). To overcome the prolonged treatment times of non-coplanar IMRT beam arrangements, recent work has been focusing on increasing the delivery efficiency by employing non-coplanar arcs instead. This showed promise due to a drastic reduction in the treatment time, while obtaining a high plan quality (Burghelea *et al.* 2016; Kearney *et al.* 2017, 2018; Lyu *et al.* 2018). Recently, we proposed a novel treatment approach for liver SBRT, designated VMAT+, complementing VMAT with a few non-coplanar IMRT beams with computer-optimized, patient-specific orientations to enhance plan quality, while keeping delivery time low (Sharfo *et al.* 2017). However, plan optimization times for VMAT+ were long because of the need for individualized beam angle optimization (BAO).

In this study, we used our in-house developed algorithm for automated multi-criterial planning with integrated BAO to explore opportunities for enhancing prostate SBRT dose distributions by complementing dual-arc coplanar VMAT with non-coplanar IMRT beams. To keep the total delivery time limited, the investigated maximum number of complementary non-coplanar beams was five (VMAT+5). Based on the selected beam orientations in the VMAT+5 plans, a *class solution* (CS) for the non-coplanar beams was defined. Final VMAT+CS plans were benchmarked against automatically generated *i)* dual-arc coplanar VMAT plans (VMAT), *ii)* VMAT+5 plans, and *iii)* 30-beam non-coplanar IMRT plans with computer-optimized beam orientations (30-NCP). Differences in dosimetric plan parameters, computation and treatment delivery times were analyzed. Dose measurements were performed to verify deliverability of generated plans at the treatment unit.

8

8.2   Materials and Methods

8.2.1   Patient data

Planning CT-scans of 20 randomly selected prostate SBRT patients were included in this study. In all CT-scans the rectum (outer contour), rectal mucosa (3mm wall), bladder, urethra, femoral heads, scrotum, penis and prostate were delineated. The average planning target volume (PTV) size was 91.2 cc [57.8 – 142.3 cc] (PTV was defined as prostate plus 3 mm isotropic margin). Dose was delivered in 4 fractions of 9.5 Gy, emulating high-dose rate (HDR) brachytherapy with highly heterogeneous dose distributions (Aluwini *et al.* 2010). Plan acceptability was subject to the dosimetric constraints presented in Table 8.1.

Table 8.1: Clinically applied dose constraints for prostate SBRT.

Structure	Parameter	Tolerance Limit
PTV	$V_{100\%}$	95%
Rectum	D_{max}	38 Gy
	D_{1cc}	32.3 Gy
Rectum mucosa	D_{max}	28.5 Gy
Bladder	D_{max}	41.8 Gy
	D_{1cc}	38 Gy
Urethra	$D_{5\%}$	45.5 Gy
	$D_{10\%}$	42 Gy
	$D_{50\%}$	40 Gy
Femoral heads	D_{max}	24 Gy

8.2.2   Automated plan generation

For each patient, the Erasmus-iCycle multi-criterial optimizer was used to automatically generate one treatment plan per investigated technique (VMAT+CS, VMAT, VMAT+5, 30-NCP) that is both *Pareto-optimal* and clinically favorable (Breedveld *et al.* 2012). For practical and legal reasons, Erasmus-iCycle plans are not directly used clinically. Instead, the Erasmus-iCycle plan is automatically converted into a clinically deliverable plan by the Monaco® treatment planning system (TPS) (Elekta AB, Stockholm, Sweden) (Voet *et al.* 2014). For this purpose, a patient-specific Monaco® template is automatically made based on the Erasmus-iCycle dose distribution. In case plan generation includes BAO, the optimal angles are established with Erasmus-iCycle, and are then used as fixed angles in the subsequent Monaco® plan generation. Many studies have demonstrated superiority of these automatically generated plans compared to manually generated

plans (Rossi *et al.* 2018; Sharfo *et al.* 2016, 2018; Buschmann *et al.* 2016; Heijmen *et al.* 2018).

Multi-criterial plan generation with Erasmus-iCycle is based on a tumor site-specific '*wish-list*' containing hard constraints to be strictly obeyed and plan objectives with ascribed priorities. For BAO, a set of candidate beam orientations has to be defined as well (Breedveld *et al.* 2012). In this study, we used a published wish-list for prostate SBRT (Rossi *et al.* 2018). For the optimizations with integrated BAO (VMAT+5 and 30-NCP), the non-coplanar beam selection search-space consisted of 300 candidate beams, separated by about 10 degrees, and homogeneously distributed across the part of the sphere where collisions between the patient/couch and the gantry were avoided, as verified at the treatment unit.

All plans in this study were generated for an Elekta Synergy® treatment machine equipped with a Versa HD™ collimator with a leaf width of 5 mm. 10 MV photon beams were used. Dose calculations in Monaco® (version 5.10) were performed with a dose grid resolution of 3 mm. The total number of control points in all plans in this study was kept fixed at 300 for all investigated techniques (i.e. VMAT+CS, VMAT, VMAT+, and 30-NCP) to eliminate potential bias. IMRT was delivered with dynamic multi-leaf collimation with a maximum of 10 control points per involved beam, in line with our clinical practice. A 5 degrees collimator angle was used for all arcs/beams, and the maximum dose rate was 600 MU/min. For generation of the VMAT+5 and VMAT+CS plans, VMAT and the non-coplanar IMRT beams were optimized simultaneously, both in Erasmus-iCycle and in Monaco®.

8.2.3 Workflow for generation of the non-coplanar beam angle class solution (CS)

The final CS to supplement dual-arc VMAT in VMAT+CS treatments was developed in a stepwise approach, based on automatically generated plans:

(i) For each of the 20 study patients, the optimal VMAT+5 plan with individualized beam angles was generated.

(ii) Based on an analysis of the angular distribution of the selected 20x5 non-coplanar beams, candidate class solutions, CS_i, were defined as described in the Results section, all including a small number of frequently selected beam directions, and accounting for the left-right symmetry in the patients' anatomies.

(iii) For a sub-group of 6 randomly selected patients, VMAT+CS_i plans were generated for all CS_i. These plans were then compared with corresponding VMAT plans, and the CS_i resulting in the most favorable plan quality increases relative to VMAT (focusing on rectum dose parameters) was selected as final CS.

8.2.4 Dosimetric comparisons of VMAT+CS plan parameters with VMAT, VMAT+5 and 30-NCP

Automatically generated VMAT+CS, VMAT, VMAT+5 and 30-NCP plans were compared for the 20 study patients. Prior to the analyses, all generated 80 plans were normalized to have identical PTV dose coverage (V_{38Gy} = 95%, as requested in clinical practice). Next, compliance with the clinically applied dose constraints (Table 8.1) was verified for all plans. Finally, plan parameter differences were analyzed. Paired two-sided Wilcoxon signed-rank tests were performed to assess clinical significance of observed differences ($p < .05$).

8.2.5 Plan deliverability, treatment time and MU for VMAT and VMAT+CS

For a sub-group of 5 patients with the largest plan quality gains achieved with VMAT+CS compared to VMAT, both VMAT and VMAT+CS plans were delivered at an Elekta Synergy linac (Elekta AB, Sweden) while irradiating a PTW 2D-Array seven29™ and Octavius™ phantom (PTW, Freiburg, Germany). The measurements were compared to Monaco® TPS predictions using a commercial QA software package (PTW VeriSoft® version 6.2) with 5% cutoff, 3% global maximum dose and 1 mm distance to agreement (3%/1 mm) criteria, and 95% Gamma passing rate threshold. For delivery time comparisons, we separately measured *i)* beam-on-times, *ii)* gantry-travel-times (times to rotate the gantry from one fixed angle to another while the beam is off), and *iii)* couch-travel-times (times required to rotate the treatment couch in between beams, including time needed for entering the room). Additionally, the VMAT and VMAT+CS plans were compared regarding the total number of monitor units (MU).

8.3 Results

8.3.1 Establishment of the final CS to define VMAT+CS

Based on the four clusters of frequently selected angles in the generated VMAT+5 plans (Figure 8.1), four principal directions (gantry angle, couch angle) for class solution definition were derived: A = (65°, -30°), B = (295°, 30°), C = (65°, 30°) and D = (295°, -30°). These are visually displayed in Figure 8.2. Based on these directions, three candidate CS_i were defined: CS_1, containing all four directions, CS_2 consisting of directions A and B, and CS_3 with directions C and D. Clearly, VMAT+CS_3 resulted in the most favorable plan quality improvements for the rectum dose parameters compared to VMAT (Figure 8.3).
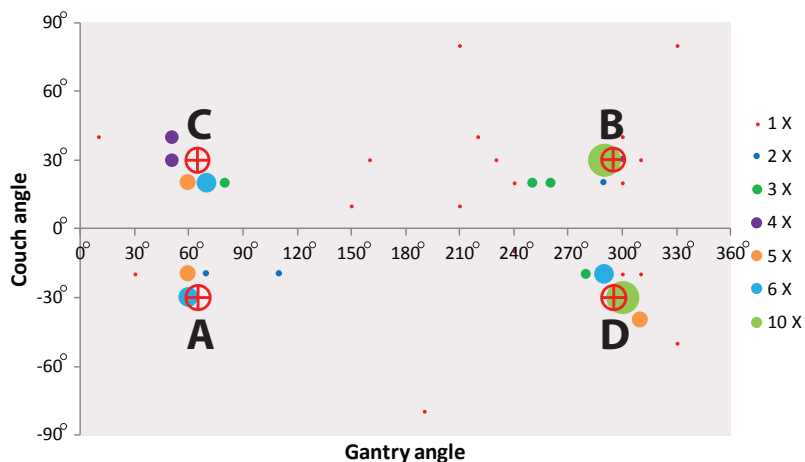
Figure 8.1: Angular distribution of the 100 non-coplanar IMRT beams in the VMAT+5 plans of the 20 study patients. Based on these results four principle beam directions A = (65°, -30°), B = (295°, 30°), C = (65°, 30°) and D = (295°, -30°) were derived to develop the final class solution for VMAT+CS.

Therefore, CS_3 was selected as the final CS for further evaluation of VMAT+CS with respect to VMAT, VMAT+5 and 30-NCP for all 20 study patients.

8.3.2 Comparisons of VMAT+CS plan parameters with VMAT, VMAT+5 and 30-NCP

All 20x4 plans included in the analyses fulfilled the clinically applied dose constraints. In general, all plans with non-coplanar beam arrangements (VMAT+CS, VMAT+5, and 30-NCP) resulted in substantial reductions in doses in healthy tissues and dose bath compared to VMAT (Figure 8.4 and Table 8.2). Differences between VMAT+CS and VMAT+5 were generally small and clinically negligible, while the former had less non-coplanar beams (yielding smaller overall treatment delivery time), and did not require individualized beam-angle-optimization (yielding reduced optimization time). Remarkably, also the differences between VMAT+CS and 30-NCP, the latter with much enhanced degrees of freedom in plan optimization, were relatively small. Actually, in rectum and bladder Dmean there was even a small advantage for VMAT+CS. For rectum D_{1cc}, V_{40GyEq}, V_{60GyEq}, bladder D_{1cc} and all urethra parameters, differences between VMAT+CS and 30-NCP were considered clinically irrelevant. There was a significant improvement in dose bath with 30-NCP, but clear disadvantages of this technique are the long optimization times (~100 hours per patient) and the very long delivery times (see Section 8.3.3).
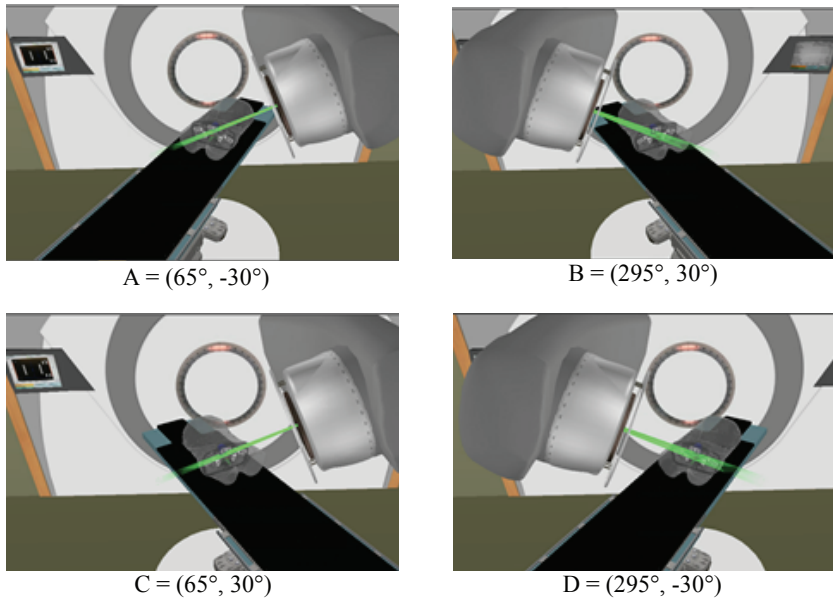
Figure 8.2: Four principle beam directions, characterized by (gantry angle, couch angle), from Figure 8.1, used to develop the final class solution for VMAT+CS. This final CS consisted of directions C and D.

8.3.3   Plan deliverability, treatment time and MU for VMAT and VMAT+CS

All delivered VMAT and VMAT+CS plans passed the QA tests (Gamma passing rate > 95%) with an average Gamma passing rate of 98.3% ± 1.0% [97.5%, 100%] for VMAT plans and 98.3% ± 0.7% [97.5%, 99.4%] for VMAT+CS plans. Compared to VMAT, the average total delivery time of VMAT+CS plans increased from 9.1 ± 0.7 min to 11.0 ± 0.3 min (see Figure 8.5 for details). VMAT+CS plans required 3% less MU (4055 ± 191 compared to 4186 ± 398), which was not significant (p = .375).

8.3.4   Optimization time reduction

Using a class solution instead of individualized beam angle selection in VMAT+5 resulted in a substantial reduction in optimization time by a factor of 25. Optimization for VMAT+CS plans in Erasmus-iCycle took on average 1 hour.

8
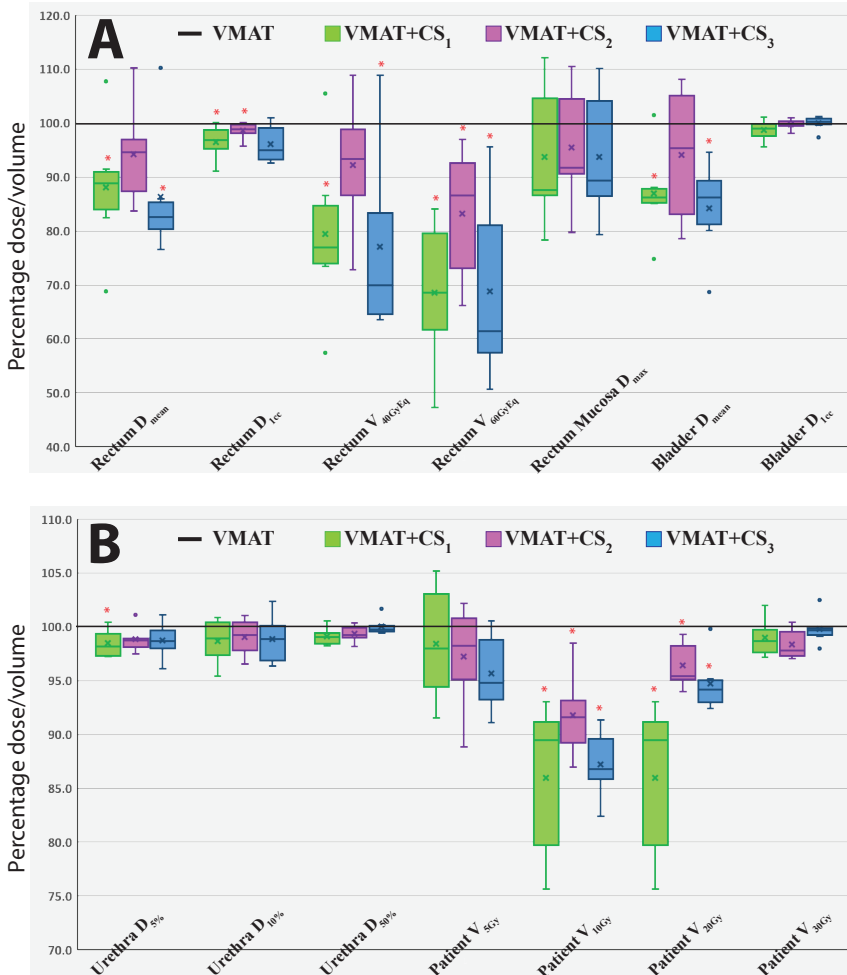
Figure 8.3: Healthy tissue dose parameters for VMAT+CS_1, VMAT+CS_2, and VMAT+CS_3 plans relative to VMAT (VMAT always 100%) for the 6 randomly selected patients used to compare the class solutions. All plans were normalized to deliver the clinically required PTV coverage (V_{38Gy} = 95%). The central line of each box represents the median value, and its upper and lower edge the 25th and 75th percentiles, respectively. The x mark inside the box represents the mean value. The whiskers extend to the minimum and maximum values, or to 1.5 times the inter-quartile range from the top/bottom of the box. Values outside this range are plotted individually as outliers ('°'). ∗ : Difference with VMAT is statistically significant.
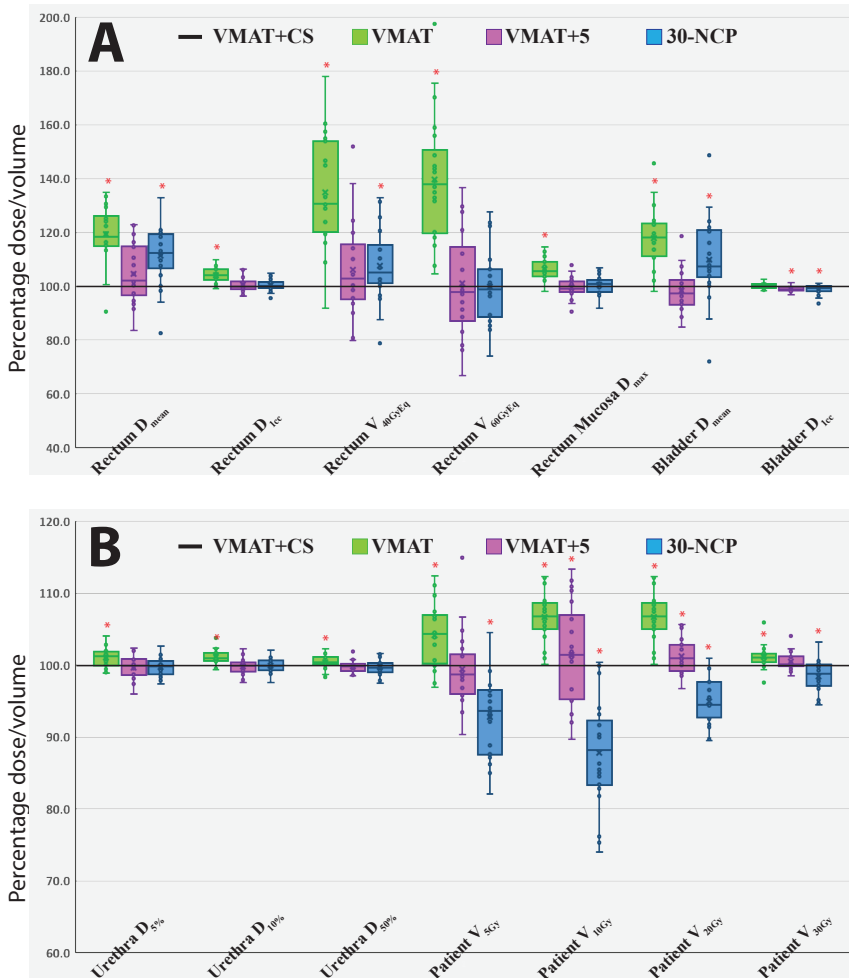
Figure 8.4: Healthy tissue dose parameters for VMAT, VMAT+5, and 30-NCP plans relative to VMAT+CS (VMAT+CS always 100%) for all study patients. All plans were normalized to deliver the clinically required PTV coverage (V_{38Gy} = 95%). The central line of each box represents the median value, and its upper and lower edge the 25th and 75th percentiles, respectively. The whiskers extend to the minimum and maximum values, or to 1.5 times the inter-quartile range from the top/bottom of the box. Hollow points are the discrete data points of the 20 study patients ('$\circ$'). Positive values indicate lower doses for VMAT+CS.

∗ : Difference with VMAT+CS is statistically significant.

Table 8.2: Comparison of dosimetric plan parameters of VMAT+CS with VMAT, VMAT+5, and 30-NCP plans for all study patients. While the VMAT+CS column contains absolute values of the plan parameters, the VMAT, VMAT+5 and 30-NCP columns show percentage differences with respect to VMAT+CS. Positive differences hint at an advantage of VMAT+CS.

	VMAT+CS	VMAT – (VMAT+CS)		(VMAT+5) – (VMAT+CS)		30NCP – (VMAT+CS)	
	Mean ± SD [range]	Mean ± SD [range]	p-value	Mean ± SD [range]	p-value	Mean ± SD [range]	p-value
PTV							
$V_{100\%}$ (%)	95.0 ± 0.0 [94.9, 95.1]	0.0 ± 0.1 [-0.1, 0.1]	1	0.0 ± 0.0 [-0.1, 0.0]	1	0.0 ± 0.0 [-0.1, 0.1]	1
$D_{98\%}$ (Gy)	35.1 ± 0.7 [33.8, 36.3]	1.1 ± 1.5 [-1.0, 1.5]	**0.004**	-0.1 ± 1.6 [-2.8, 3.8]	0.245	0.6 ± 1.7 [-1.5, 4.3]	0.368
$CI^{‡}$	1.11 ± 0.04 [1.04, 1.20]	-1.6 ± 1.9 [-6.8, 2.5]	**0.001**	-0.5 ± 2.3 [-5.1, 3.4]	0.522	-1.5 ± 1.9 [-4.7, 2.0]	**0.006**
Rectum							
D_{mean} (Gy)	5.4 ± 1.0 [3.7, 7.5]	19.4 ± 10.6 [-9.3, 35.0]	**<0.001**	4.6 ± 10.9 [-16.5, 23.0]	0.097	11.5 ± 11.2 [-17.4, 32.9]	**0.001**
D_{1cc} (Gy)	27.6 ± 2.6 [23.7, 32.8]	4.2 ± 2.7 [-1.0, 9.8]	**<0.001**	0.5 ± 2.8 [-3.5, 6.5]	0.312	0.4 ± 2.1 [-4.3, 4.9]	0.596
V_{40GyEq} (%)	3.5 ± 1.3 [1.8, 6.1]	34.9 ± 20.3 [-8.1, 78.0]	**<0.001**	6.1 ± 18.5 [-20.1, 52.0]	0.231	7.6 ± 13.5 [-21.2, 32.8]	**0.03**
V_{60GyEq} (%)	1.1 ± 0.7 [0.2, 2.8]	39.7 ± 23.2 [4.6, 97.5]	**<0.001**	1.1 ± 19.9 [-33.2, 36.6]	0.784	-0.2 ± 13.6 [-25.8, 27.8]	0.49
Rectum Mucosa							
D_{max} (Gy)	26.2 ± 2.6 [20.9, 31.6]	6.3 ± 3.9 [-1.9, 14.7]	**<0.001**	-0.3 ± 4.0 [-9.3, 8.0]	0.763	0.6 ± 3.6 [-8.2, 7.0]	0.577
Bladder							
D_{mean} (Gy)	6.6 ± 1.3 [4.5, 8.8]	17.9 ± 11.0 [-1.8, 45.6]	**<0.001**	-1.5 ± 7.7 [-15.2, 18.6]	0.841	9.8 ± 15.7 [-28.0, 48.6]	**0.006**
D_{1cc} (Gy)	36.7 ± 1.2 [34.1, 38.5]	0.3 ± 1.1 [-1.5, 2.7]	0.409	-1.0 ± 1.0 [-3.0, 1.3]	**0.036**	-1.1 ± 1.8 [-6.4, 1.1]	**0.017**
Urethra							
$D_{5\%}$ (Gy)	40.1 ± 0.8 [38.6, 41.5]	1.1 ± 1.4 [-1.1, 4.0]	**0.003**	-0.4 ± 1.6 [-4.0, 2.4]	0.261	-0.2 ± 1.3 [-2.7, 2.7]	0.409
$D_{10\%}$ (Gy)	39.6 ± 0.7 [38.4, 41.1]	1.1 ± 1.0 [-0.6, 3.8]	**<0.001**	-0.2 ± 1.1 [-2.4, 2.3]	0.571	-0.1 ± 1.0 [-2.4, 2.1]	0.596
$D_{50\%}$ (Gy)	38.0 ± 0.5 [37, 38.9]	0.4 ± 0.9 [-1.7, 2.3]	**0.036**	-0.2 ± 0.8 [-1.5, 1.9]	0.927	-0.3 ± 1.1 [-2.5, 1.6]	0.312
left femur head							
D_{max} (Gy)	13.8 ± 1.6 [10.7, 17.9]	10.9 ± 8.4 [-8.8, 23.3]	**<0.001**	1.6 ± 8.3 [-15.5, 19.3]	0.133	-34.2 ± 15.6 [-65.4, -6.8]	**<0.001**
Right femur head							
D_{max} (Gy)	14.0 ± 1.7 [10.7, 17.5]	7.4 ± 5.3 [-3.4, 16.9]	**<0.001**	-1.3 ± 4.7 [-9.2, 7.3]	0.701	-32.3 ± 13.9 [-50.0, -7.0]	**<0.001**
Patient							
V_{2Gy} (cc)	5325 ± 1007 [4151, 7790]	9.3 ± 5.7 [-1.6, 20.5]	**<0.001**	18.6 ± 6.8 [6.3, 30.3]	**0.008**	32.9 ± 6.9 [14.7, 42.2]	**<0.001**
V_{5Gy} (cc)	3444 ± 687 [2692, 5134]	4.1 ± 4.4 [-3.1, 12.4]	**0.001**	-0.6 ± 5.2 [-9.6, 14.9]	0.546	-7.2 ± 5.5 [-17.9, 4.5]	**<0.001**
V_{10Gy} (cc)	1332 ± 339 [924, 2022]	19.0 ± 7.3 [4.2, 30.4]	**<0.001**	1.4 ± 7.0 [-10.3, 13.3]	**0.008**	-12.2 ± 7.6 [-26.0, 0.3]	**<0.001**
V_{20Gy} (cc)	317 ± 83 [213, 481]	6.7 ± 3.3 [0.1, 12.3]	**<0.001**	1.2 ± 2.4 [-3.2, 5.6]	**0.003**	-5.1 ± 3.2 [-10.5, 1.0]	**<0.001**
V_{30Gy} (cc)	156 ± 44 [102, 246]	1.1 ± 1.6 [-2.5, 5.9]	**0.001**	0.5 ± 1.3 [-1.5, 4.1]	0.09	-1.6 ± 2.2 [-5.5, 3.2]	**0.008**

‡CI = *Conformity Index* (= patient volume receiving the prescribed dose / PTV volume receiving prescribed dose)

8
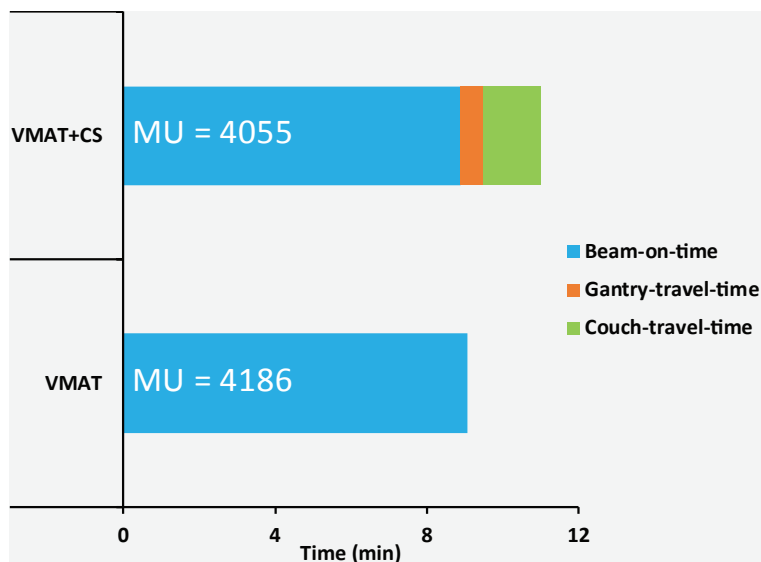
Figure 8.5: Average measured beam-on-times, gantry-travel-times and couch-travel-times for five VMAT and VMAT+CS plans. The numbers in the bars represent mean delivered MUs.

8.4 Discussion

In this study, we have developed and evaluated a novel treatment approach for prostate SBRT at a C-arm linac, consisting of dual-arc coplanar VMAT supplemented with a non-coplanar beam-angle class solution (CS) consisting of two IMRT beams (VMAT+CS). Initial aim of the study was to explore opportunities for enhancement of the plan quality for prostate SBRT, as obtained with VMAT, by adding non-coplanar beams. To keep the treatment delivery time short, no more than 5 non-coplanar beams were added. The study was inspired by the success for liver SBRT, where substantial plan quality enhancement compared to VMAT could be obtained by adding 1-5 non-coplanar beams with patient-specific, computer-optimized orientations (VMAT+) (Sharfo *et al.* 2017). For the prostate case studied in this paper, the distribution of selected non-coplanar orientations for VMAT+5 in the twenty study patients pointed at a possibility for the use of a CS, which was then successfully further explored. A fixed set of two non-coplanar orientations (CS) resulted for all patients in substantial plan quality enhancements compared to VMAT, while the increase in treatment time was very moderate (from 9.1 min to 11.0 min).

Remarkably, the quality of VMAT+CS plans was highly similar to that of VMAT+5 plans, the latter with more, and also patient-specific beam orientations. Most likely, because the pelvic anatomies of prostate cancer patients are highly similar, high quality

plans could be generated for all patients with a fixed set of two non-coplanar beams supplementing VMAT. Interestingly, also the quality of plans with 30 non-coplanar beams with individually optimized orientations (30-NCP) was highly similar to that of VMAT+CS, while optimization and treatment times were largely enhanced. Apparently, adding only two well-selected orientations to the patient plans was enough for a substantial gain in plan quality. Adding more non-coplanar beams and patient-specific optimization of the orientations of the non-coplanar beams did not result in significantly better plans, especially when also considering the involved increases in optimization and delivery times.

Recently, Rossi *et al.* (2018) showed a clear advantage for non-coplanar CyberKnife® planning compared to coplanar VMAT for prostate SBRT. Also in that study, all plans were fully automatically generated, using the same autoplanning system and configuration as applied in this study. Also the patient group was the same. Interestingly, the quality of the VMAT+CS plans generated here is very similar to that of the previously generated CyberKnife® plans. On the other hand, delivery times with the CyberKnife® were much longer (45 min vs. 11 min in this study). There is a high similarity here with the above comparison between VMAT+CS and 30-NCP; apparently in prostate SBRT there is no need for using a large amount of non-coplanar beams to get a very high plan quality.

Currently, we have not yet an idea whether the CS developed for the planning protocol used in our center would also work for SBRT planning approaches in other centers. This is a topic of further research. However, as a small side study, we tested the developed CS for 15 prostate cancer patients treated with our regular 20 x 3 Gy scheme for homogeneous target dose delivery (non SBRT). Also for this conventional delivery scheme, VMAT+CS showed an improvement in plan quality compared to VMAT (< 10% in rectum, anus and bladder mean doses, data not presented), but no improvements in the high rectum and bladder doses. Possibly, the applied larger PTV margins and/or higher PTV dose homogeneity (with shallower dose fall-off towards the healthy tissues) reduced the impact of the added CS beams.

Addition of a few, well-selected, non-coplanar beams to VMAT (the VMAT+ approach) has now turned out to be successful for both liver SBRT and prostate SBRT. New studies on other tumor sites are part of a future project. We are able to do this type of work due to availability of our in-house developed optimizer for fully automated and integrated multi-criterial beam angle selection and IMRT beam profile optimization. The success of our work on VMAT+ is an indication for the need of advanced algorithms for integrated beam angle and beam profile optimization in commercial treatment planning systems. In this context, one should keep in mind that with the VMAT+CS approach, BAO is not needed for new patients. However, development of the CS was only possible with the use of integrated optimization of beam angles and profiles.

To the best of our knowledge, this paper and our paper on liver SBRT (Sharfo *et al.* 2017) are the only papers that use autoplanning to systematically investigate the addition of non-coplanar IMRT beams to fast coplanar VMAT for enhancement of the plan quality achieved with VMAT, while keeping treatments fast. Several recent publications (Audet *et al.* 2011; Clark *et al.* 2012; Thomas *et al.* 2014; Woods *et al.* 2016; Slosarek *et al.* 2018) investigated the use of non-coplanar arcs to enhance plan quality without prohibitively prolonged treatment times. Selection of the non-coplanar arcs was performed manually. Clark *et al.* (2012) and Thomas *et al.* (2014) showed that three non-coplanar arcs combined with one coplanar arc produced clinically equivalent conformity and dose spillage compared with GammaKnife for multiple cranial brain metastases while increasing the delivery efficiency due to its reduced treatment time. Since, this class solution has been incorporated in the Eclipse treatment planning system as HyperArc and has further proven to improve delivery efficiency and reduce dose to normal brain tissue when compared to coplanar VMAT (Ohira *et al.* 2018). More specific studies are needed to compare such approaches with the proposed VMAT+.

In conclusion, using an algorithm for fully automated, integrated multi-criterial beam profile and beam angle optimization, we have derived a two-beam non-coplanar class solution to supplement VMAT for prostate SBRT. Adding the CS beams to conventional coplanar VMAT resulted in substantial improvement in treatment plan quality with a minimal increase in treatment delivery time. Due to the use of a non-coplanar beam-angle class solution, time-consuming individualized beam angle optimization can be avoided.

8

First Fully Automated Planning Solution for Robotic Radiosurgery - Comparison with Automatically Planned Volumetric Arc Therapy for Prostate Cancer

Linda Rossi, Abdul Wahab M. Sharfo, Shafak Aluwini, Maarten L.P. Dirkx, Sebastiaan Breedveld and Ben J.M. Heijmen

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

Abstract

Background: For conventional radiotherapy treatment units, automated planning can significantly improve plan quality. For robotic radiosurgery, systems for automatic generation of clinically deliverable plans do not yet exist. For prostate stereotactic body radiation therapy (SBRT), few studies have systematically compared VMAT with robotic treatment.

Material and Methods: The multi-criteria autoplanning optimizer, developed at our institute, was coupled to the commercial treatment planning system of our robotic treatment unit, for fully automated generation of clinically deliverable plans (autoROBOT). The system was then validated by comparing autoROBOT plans with manually generated plans. Next, the autoROBOT system was used for systematic comparisons between autoROBOT plans and VMAT plans, that were also automatically generated (autoVMAT). CTV-PTV margins of 3 mm were used for autoROBOT (clinical routine) and autoVMAT plan generation. For autoVMAT, an extra plan was generated with 5 mm margin (often applied for VMAT). Plans were generated for a 4×9.5 Gy fractionation scheme.

Results: Compared to manual planning, autoROBOT improved rectum D_{1cm^3} (16%), V_{60GyEq} (75%) and D_{mean} (41%), and bladder D_{mean} (37%) (all $p \leq .002$), with equal PTV coverage. In the autoROBOT and autoVMAT comparison, both with 3 mm margin, rectum doses were lower for autoROBOT by 5% for rectum D_{1cm^3} ($p = .002$), 33% for V_{60GyEq} ($p = .001$) and 4% for D_{mean} ($p = .05$), with comparable PTV coverage and other OAR sparing. With 5 mm margin for VMAT, 18/20 plans had a PTV coverage lower than requested ($< 95\%$) and all plans had higher rectum doses than autoROBOT (mean percentage differences of 13% for D_{1cm^3}, 69% for V_{60GyEq} and 32% for D_{mean} (all $p < .001$)).

Conclusions: The first system for fully automated generation of clinically deliverable robotic plans was built. Autoplanning did largely enhance robotic plan quality, compared to manual planning. Using autoplanning for both the robotic system and VMAT,

superiority of non-coplanar robotic treatment compared to coplanar VMAT for prostate SBRT was demonstrated.

9

9.1   Introduction

In prostate stereotactic body radiation therapy (SBRT), patients are treated with large fraction doses, requiring high accuracy delivery and with image-guided dose delivery (Brenner & Hall 1999; Madsen *et al.* 2007; Fuller *et al.* 2010; Boike *et al.* 2011; Freeman & King 2011; Jabbari *et al.* 2012; Aluwini *et al.* 2013; Katz *et al.* 2013; King *et al.* 2013; van de Water *et al.* 2014). Both C-arm linacs (Willoughby *et al.* 2006; Kupelian *et al.* 2007; Dong *et al.* 2014; Keall *et al.* 2016) and robotic units (Fuller *et al.* 2010; Freeman & King 2011; Aluwini *et al.* 2013; Katz *et al.* 2013; King *et al.* 2013; van de Water *et al.* 2014) have been used for prostate SBRT.

Recent findings on the potential added value of non-coplanar setups for prostate SBRT instead of coplanar treatment are contradictory (Rossi *et al.* 2012; Lin *et al.* 2014b; MacDougall *et al.* 2014; van de Water *et al.* 2014). Two recent studies have compared robotic treatment and VMAT for prostate SBRT (Lin *et al.* 2014b; MacDougall *et al.* 2014). MacDougall *et al.* (2014) found no discernible dosimetric differences, based on only six patients. Lin *et al.* (2014b) concluded that VMAT was preferable because of reduced treatment time and superior dose distribution conformality. In both studies, all plans were generated manually, and clinically delivered plans were retrospectively compared with an alternative plan. Both the manual planning and retrospective comparisons may have introduced bias and noise in the technique comparisons.

Recently, several systems have been proposed for planning automation (Zhang *et al.* 2011; Breedveld *et al.* 2012; Quan *et al.* 2012b; Purdie *et al.* 2014; Voet *et al.* 2014; Wu *et al.* 2014; Zarepisheh *et al.* 2014; Krayenbuehl *et al.* 2015; Tol *et al.* 2016), all for treatment with C-arm linacs. In this work, we have developed the first system for automatic generation of deliverable plans for non-coplanar robotic treatment (autoROBOT). Basis of the autoROBOT planning system is a multi-criterial optimizer that was also the core of a recently developed system for automatic VMAT plan generation for C-arm linacs (Voet *et al.* 2014; Sharfo *et al.* 2016). The developed autoROBOT planning system was first evaluated by comparing manually generated prostate SBRT plans with autoROBOT plans. We then used the autoROBOT and autoVMAT planning systems to systematically compare robotic and VMAT treatment for prostate SBRT. The use of exactly the same plan optimization scheme for autoROBOT and autoVMAT (described in Section 9.2.2) allowed bias-free technique comparisons and allowed generation of new input for the on-going debate (Rossi *et al.* 2012; Lin *et al.* 2014b; MacDougall *et al.* 2014; van de Water *et al.* 2014) on potential added value of non-coplanar prostate SBRT, compared to coplanar treatment.

9

9.2   Material and methods

9.2.1   Patients

In this study, contoured CT-scans of 20 prostate SBRT patients, previously treated with the robotic M6 CyberKnife® (Accuray Inc., Sunnyvale, CA, USA), were used. A planning target volume (PTV) with a 3 mm isotropic margin around the prostate (PTV_{3mm}) was used for clinical planning (van de Water *et al.* 2014). In the investigations, both autoROBOT and autoVMAT plans were generated for PTV_{3mm}. AutoVMAT plans were also generated for PTV_{5mm}, as often applied for C-arm linac prostate SBRT. Average PTV_{3mm} and PTV_{5mm} sizes were 91.2 cm^3 [57.8–142.3 cm^3] and 109.5 cm^3 [71.1–165.7 cm^3], respectively.

Contoured OARs were rectum (outer contour), rectal mucosa (3 mm wall), bladder, urethra, femoral heads, scrotum and penis. All plans simulated delivery of 38 Gy in four fractions, with highly heterogeneous dose distributions to emulate high dose-rate brachytherapy dosimetry (Rossi *et al.* 2012).

Five patients were used for configuration of the autoROBOT and autoVMAT planning systems (Section 9.2.2.2). The automated work-flows were then applied to all 20 patients. For validation of the autoROBOT planning system, autoROBOT plans for the first 10 study patients were compared to the manually generated and clinically delivered plans. For all 20 patients, autoROBOT plans were compared with autoVMAT$_{3mm}$ plans and autoVMAT$_{5mm}$ plans.

9.2.2   Automated plan generation

9.2.2.1   The autoVMAT and autoROBOT planning systems

Basis of autoROBOT and autoVMAT plan generation was the Erasmus-iCycle multi-criterial optimizer for generating *Pareto-optimal* and clinically favorable plans (Breedveld *et al.* 2012). For practical and legal reasons, Erasmus-iCycle plans cannot be directly used clinically. However, we have recently coupled Erasmus-iCycle to the Monaco® treatment planning system (TPS) (Elekta AB, Stockholm, Sweden) for fully automated, multi-criterial generation of IMRT and VMAT plans for clinical delivery at a linac; based on the Erasmus-iCycle dose distribution, a patient-specific Monaco® template is automatically produced, to be used for automated final plan generation. Effectively, Erasmus-iCycle first optimizes the plan, while Monaco® converts it into a clinically deliverable plan, see (Voet *et al.* 2014) for details. The resulting plan quality is equal, or superior to the quality of manually generated plans, and the system is now in routine clinical use (Voet *et al.* 2013b, 2014; Heijkoop *et al.* 2014; Sharfo *et al.* 2015; Della Gala *et al.* 2017).

For this study, we have configured the system for generating dual, full-arc autoVMAT plans for prostate SBRT according to our clinical protocol, deliverable at an Elekta linac equipped with an Agility™ MLC. Final plans were generated with Monaco® version 5.10 (Elekta AB, Stockholm, Sweden).

For automated multi-criterial generation of autoROBOT plans, a special version of Erasmus-iCycle was prepared for plan optimization for the IRIS™ variable aperture collimator (Accuray AB, Sunnyvale, CA, USA), mounted on the CyberKnife®. Basis was a previously developed version for optimization with fixed cone diameters and non-coplanar beam set-ups (van de Water *et al.* 2011a). This system was modified to handle the available non-coplanar beam directions (nodes) of our novel M6 CyberKnife® systems and the IRIS™ collimator, i.e., 117 node positions from the full body path. For fully automated generation of final, deliverable plans, this Erasmus-iCycle version was coupled to the MultiPlan® TPS (version 5.1.3) that comes with the CyberKnife®, similar to the system built for linacs (above). Similar to the linac solution, automatically produced individualized planning templates were used as intermediate between Erasmus-iCycle and MultiPlan®, aiming at generating clinically deliverable plans that dosimetrically mimicked the initial Erasmus-iCycle plans. As in clinical practice, the goal was to keep the delivery time below 45 min. Apertures from 10 to 40 mm diameter could be selected, as used clinically for manually generated plans.

9.2.2.2   Configuration of autoVMAT and autoROBOT planning

As described in Section 9.2.2.1, both for autoVMAT and autoROBOT planning, Erasmus-iCycle is used for plan optimization, while the respective clinical planning systems are used for mimicking the Erasmus-iCycle plan. Plan generation with Erasmus-iCycle is based on a '*wish-list*', containing the hard planning constraints and planning objectives with their goal values and assigned priorities (Breedveld *et al.* 2012). For each treatment site/ treatment technique, a dedicated wish-list is configured, which is then used for automated plan generation for all involved patients, without further change.

In this study, a single wish-list was generated and applied for both autoVMAT and autoROBOT planning.

Using the same wish-list for both techniques is a key aspect of this study, since it allowed to perform a fair like for like comparison of the two delivery techniques. Technical details on the developed wish-list for prostate SBRT are presented in Section 9.6.

9

9.2.3 Plan evaluation and comparison

In this study, plan comparisons were mainly focused on our clinical aims. For the PTV, the near-minimum dose ($D_{98\%}$) and the coverage ($V_{100\%}$) were evaluated. A coverage of 95% is requested for clinical plans ($V_{100\%} = 95\%$), while a coverage between 93% and 95% is still acceptable if necessary to fulfill OAR constraints. Rectum is considered the most important OAR, focusing at the high doses with $D_{1cm^3} < 32.3$ Gy. For bladder, the D_{1cm^3} requirement is < 38 Gy. Urethra $D_{50\%}$ and $D_{5\%}$ constraint values are 40 and 45.5 Gy, respectively.

Apart from these clinically used plan parameters, we also evaluated and compared D_{mean} for both rectum and bladder, V_{40Gy} and V_{60Gy} (2 Gy/fx equivalent dose) for rectum, as suggested by QUANTEC (Michalski *et al.* 2010), as well as the dose bath, looking at patient volumes receiving > 30, > 20, > 10, > 5 and > 3 Gy, as 5% of maximum dose.

When PTV coverage was achieved (> 95%) for both plans, the plan with the slightly higher coverage was re-normalized to the value of the other plan. This approach minimized bias in comparison of OAR doses, related to different PTV coverages. Two-sided Wilcoxon signed-rank tests were performed to compare plan parameters, using $p < .05$ as cut-off for statistical significance.

Apart from plan quality comparisons based on DVH metrics, for each patient, autoROBOT and autoVMAT plans were also compared by the participating clinician (S.A.), who scored quality differences using visual analogue scales (VASs) as presented in Section 9.3. PTV, rectum, bladder, urethra and overall quality were scored separately. In total, 40 of these plan comparisons (20 patients; autoROBOT vs. autoVMAT$_{3mm}$ and autoROBOT vs. autoVMAT$_{5mm}$) were performed in a random order. In each comparison, the two plans were presented side-by-side to the clinician, who did not know which plan was presented on the left and which on the right of the screen (also here random ordering).

To investigate clinical deliverability of automatically generated plans, dosimetric quality assurance (QA) was performed, as done in our clinical routine. To this purpose, for five arbitrarily selected patients, independent dose calculations were performed for the autoROBOT plans, and measurements for autoVMAT plans with 3 and 5 mm margin. For the autoROBOT plans, beam directions and weights were used to re-calculate the entire 3D dose distribution with the Monte Carlo dose computation software SciMoCa™ (Scientific RT, Munich, Germany). For autoVMAT, plans were delivered while irradiating a 2D-array in an Octavius™ phantom (PTW, Freiburg, Germany). 3D (autoROBOT) and 2D (autoVMAT) γ-analyses were performed with 5% cutoff, 3% global maximum dose and 1 mm distance to agreement (3%/1 mm) criteria, and 95% passing rate threshold.

9.3 Results

9.3.1 autoROBOT vs. manual robotic planning

All manually and automatically generated plans for robotic treatment fulfilled clinical requirements.

Automated planning improved plan quality compared to the manually generated plans used for patient treatments, as visible in population average DVHs in Figure 9.1. Differences in PTV coverage were negligible (95.0% and 95.2% for manual and autoROBOT plans (p = .9)), but large differences in OAR doses were observed; each patient plan improved with automated planning compared to manual planning. On average, rectum D_{1cm^3} was reduced from 31.2 to 26.3 Gy (16% reduction, p = .002), V_{60Gy} from 2.4 to 0.6% (75% reduction, p = .002) and rectum D_{mean} from 10.4 to 6.1 Gy (41% reduction, p = .002). Bladder D_{mean} was improved from 14.0 (manual planning) to 6.1 Gy with automated planning (56% reduction, p = .002).
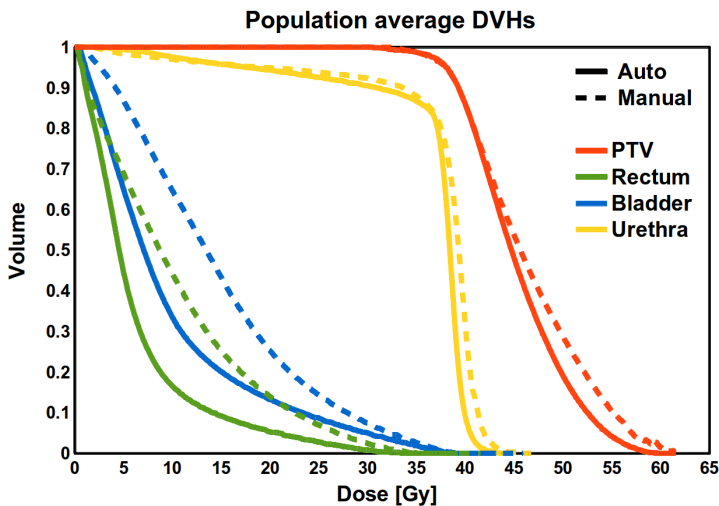

Figure 9.1: Population average DVHs for automatically generated robotic plans (autoROBOT, solid lines) and manually generated robotic plans (manual, dashed lines), the latter used for patient treatment.

Table 9.1: For all 20 patients, comparisons of autoROBOT with autoVMAT$_{3mm}$ and autoVMAT$_{5mm}$ plans.

	autoROBOT	autoVMAT$_{3mm}$				autoVMAT$_{5mm}$			
			autoVMAT$_{3mm}$ - autoROBOT (%)*				autoVMAT$_{5mm}$ - autoROBOT (%)*		
	Mean [Range]	Mean	Mean [Range]	p-value	#Pts†	Mean	Mean [Range]	p-value	#Pts†
PTV									
$V_{100\%}$ (%)	95.2 [95.0,95.5]	95.3	0 [-1,0]	.3	8	92.7	3 [0,7]	< .001	20
$D_{98\%}$ (Gy)	36.1 [35.2,36.9]	35.8	1 [-2,3]	.01	13	33.7	7 [3,13]	< .001	20
$CI^{\ddagger}$	1.1 [1.1,1.1]	1.2	6 [3,10]	< .001	20	1.2	6 [2,10]	< .001	20
Rectum									
D_{1cm^3} (Gy)	28.0 [23.0,33.5]	29.4	5 [-3,18]	.002	16	32.2	13 [1,23]	< .001	20
V_{60GyEq} (%)	1.1 [0.3, 2.6]	1.5	32 [-26,77]	.001	17	3.3	69 [15,89]	< .001	20
V_{40GyEq} (%)	3.8 [1.9, 6.0]	4.9	22 [-13,56]	< .001	17	9.2	58 [24,76]	< .001	20
D_{mean} (Gy)	6.3 [4.2, 7.7]	6.6	4 [-18,25]	.05	15	9.3	32 [14,45]	< .001	20
Bladder									
D_{1cm^3} (Gy)	37.4 [36.4,39.1]	37.2	-1 [-6,2]	.3	9	37.6	0 [-3,3]	.4	12
D_{mean} (Gy)	9.7 [6.5,13.0]	8.4	-18 [-45,7]	< .001	2	9.3	-6 [-32,15]	.1	7
Urethra									
$D_{5\%}$ (Gy)	40.4 [39.4,42.3]	40.9	1 [-4,6]	.06	13	41.6	3 [-3,8]	.001	17
$D_{50\%}$ (Gy)	38.3 [37.5,39.2]	38.5	1 [-3,3]	.2	14	39.1	2 [-1,6]	< .001	17
Patient									
V_{3Gy} (cm^3)	4910 [3428,7064]	4378	-12 [-30,11]	.001	4	4669	-5 [-22,16]	.05	6
V_{5Gy} (cm^3)	3143 [2147,4779]	3538	12 [-5,28]	< .001	18	3864	19 [3,35]	< .001	20
V_{10Gy} (cm^3)	1137 [737,1872]	1583	29 [19,37]	< .001	20	1789	37 [24,47]	< .001	20
V_{20Gy} (cm^3)	293 [203,442]	342	14 [5,20]	< .001	20	392	25 [17,34]	< .001	20
V_{30Gy} (cm^3)	150 [103,229]	159	5 [0,9]	< .001	20	182	18 [12,24]	< .001	20

* Percentage differences are expressed as ±|100 ∗ (autoVMAT − autoROBOT)/autoVMAT| with positive differences representing better performance for robotic.
† Number of patients with superior plan parameter quality for robotic treatment.
‡ CI = *Conformity Index* (= patient volume receiving the prescribed dose / PTV volume receiving prescribed dose).

9

9.3.2 autoROBOT vs. autoVMAT plan quality

9.3.2.1 autoROBOT vs. autoVMAT$_{3mm}$

Both the autoROBOT and autoVMAT$_{3mm}$ plans with $V_{100\%} > 95\%$ could be generated for all patients, as visible in Table 9.1. The near-minimum PTV dose was on average slightly higher for autoROBOT plans and the *CI* was lower (Table 9.1, Figure 9.2).

For the rectum (highest priority OAR), all parameters were on average lower for the autoROBOT with reduction of 5% for D_{1cm^3}, 32% for V_{60GyEq}, 22% for V_{40GyEq} and 4% for D_{mean} (Table 9.1, Figure 9.2). Superiority in rectum dose parameters was observed in 15–17 of the 20 study patients (Table 9.1), where differences were considered to have real clinical impact for eight patients (see clinical scoring below). For the 3–5 patients with a rectum dose advantage for autoVMAT$_{3mm}$, the differences with the robotic system were always small and only for one patient considered clinically significant (see clinical scoring below).

AutoVMAT$_{3mm}$ performed significantly better for bladder D_{mean}, but the difference in the most important parameter, D_{1cm^3} , was small (1%) and statistically insignificant (Table 9.1). Differences in urethra dose parameters were statistically insignificant.

For all patients, the autoROBOT was superior regarding patient volumes receiving > 5, > 10, > 20 and > 30 Gy (Table 9.1 and Figure 9.2), with percentage mean differences of 12% for V_{5Gy}, 29% for V_{10Gy}, 14% for V_{20Gy} and 5% for V_{30Gy}. AutoVMAT$_{3mm}$ performed better for patient volumes receiving > 3 Gy, with mean percentage improvement of 12%.

Figure 9.3 shows axial dose distributions for patient 17, who demonstrated the largest advantages for autoROBOT in rectum plan parameters compared to autoVMAT (see also Figure 9.2), and for patient 13, with the largest rectum advantages for VMAT. Apart from a better rectum sparing, in patient 17, autoROBOT plan also showed better dose conformality, in agreement with Figure 9.2.

All autoROBOT and autoVMAT$_{3mm}$ plans were clinically acceptable for the participating clinician. The comparisons as presented in the upper panel of Figure 9.4, are in line with the plan parameter evaluations above. PTV doses were found of equal quality for all patients. Apart from one patient with a small advantage for autoVMAT$_{3mm}$, rectum dose was considered equal or superior for autoROBOT. For bladder there was a balance, with only equal plan quality or small differences scored. For the urethra, the clinician had a slight preference for the autoROBOT. Overall, for 13 patients the clinician preferred autoROBOT, for two patients he preferred the autoVMAT$_{3mm}$ plan, and for five patients he scored equal quality.

9
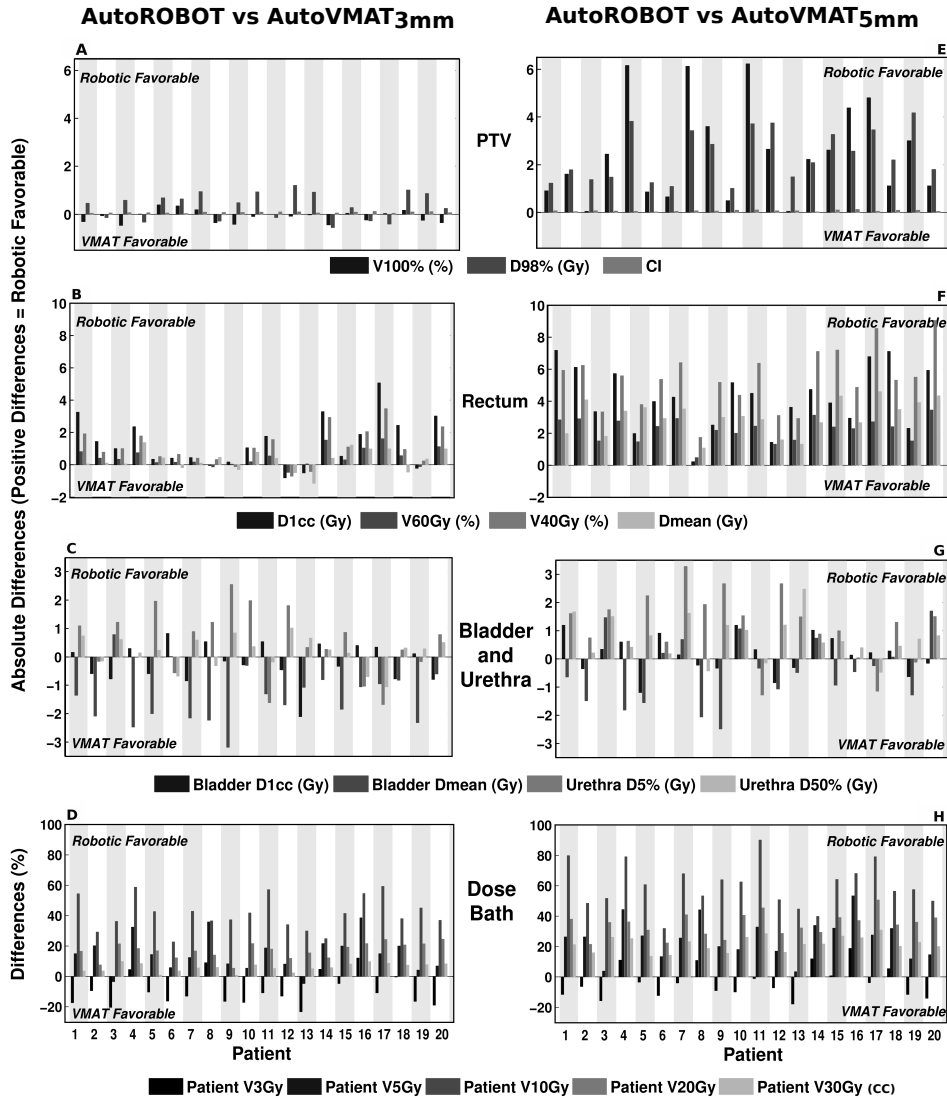
Figure 9.2: For all 20 patients, differences between autoROBOT and autoVMAT$_{3mm}$ (left panels), or autoVMAT$_{5mm}$ (right panels), expressed as $\pm|$autoVMAT $-$ autoROBOT$|$ with positive values representing better quality for autoROBOT. For dose bath percentage differences as $\pm|100 *$ (autoVMAT $-$ autoROBOT)/autoVMAT$|$ are expressed to compensate for differences in volumes (cm^3) range between parameters. *CI = Conformity Index* (= patient volume receiving 38 Gy/PTV receiving 38 Gy)

Patient 17: Biggest rectum advantage with autoROBOT

			38 Gy
			30 Gy
			25 Gy
			20 Gy
			15 Gy
			10 Gy
			5 Gy

autoROBOT   autoVMAT₃ₘₘ   autoVMAT₅ₘₘ
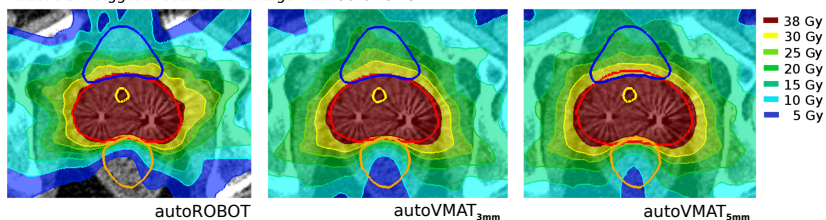
Patient 13: Biggest rectum advantage with autoVMAT

autoROBOT   autoVMAT₃ₘₘ   autoVMAT₅ₘₘ
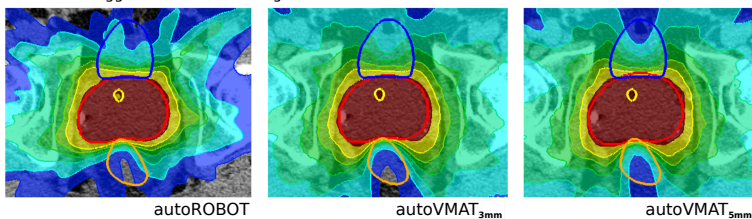
Figure 9.3: Axial dose distributions for autoROBOT, autoVMAT$_{3mm}$ and autoVMAT$_{5mm}$, for patients 17 (upper panels) and 13 (lower panels). These patients demonstrated the most pronounced advantage in rectum dose for autoROBOT instead of autoVMAT (patient 17), and the most pronounced advantage using autoVMAT compared to autoROBOT (patient 13) (see also Figure 9.2). Red contour: PTV (3 mm or 5 mm), orange contour: rectum, blue contour: bladder and yellow contour: urethra.

AutoROBOT   vs.   AutoVMAT₃ₘₘ

AutoROBOT better   Equal   AutoVMAT better

	3	2	1	0	1	2	3
PTV				20			
Rectum	2	6	7	4	1		
Bladder			6	8	6		
Urethra		2	8	4	6		
Overall		4	9	5	2		

AutoROBOT   vs.   AutoVMAT₅ₘₘ

AutoROBOT better   Equal   AutoVMAT better

	3	2	1	0	1	2	3
PTV	4	6	8	2			
Rectum	16	2	2				
Bladder		3	6	11			
Urethra	3	7	5	3	2		
Overall	11	8	1				

0 = no preference, similar plan quality
1 = preferred plan is better, probably low impact for patient
2 = preferred plan is better, possibly important impact for patient
3 = preferred plan is better, real impact for patient

Figure 9.4: Visual analogue scale (VAS) used for blind side-by-side plan comparisons by the treating clinician, and clinician scoring with the values representing numbers of plans (for each line, the sum values equal to 20).

9.3.2.2 autoROBOT vs. autoVMAT$_{5mm}$

While for autoROBOT and autoVMAT$_{3mm}$, a PTV coverage ≥ 95% was obtained for all 20 patients, with autoVMAT$_{5mm}$ this was only achieved for two patients, due to OAR constraints. Seven other patients obtained a clinically still acceptable coverage between 95% and 93%, while for the remaining 11 patients coverage was clinically unacceptable (< 93%), with a minimum of 88.8%. Also the near-minimum PTV doses were lower in the autoVMAT$_{5mm}$ plans, while the *CI* was higher.

Notwithstanding the lower PTV coverage for autoVMAT$_{5mm}$, rectum sparing was also unfavorable compared to autoROBOT, with mean percentage differences of 13% for rectum D_{1cm^3}, 69% for V_{60GyEq}, 58% for V_{40GyEq} and 32% for D_{mean}. Differences in bladder and urethra plan parameters were statistically insignificant. Dose bath was also favorable for autoROBOT plans, with reductions of patient irradiated volumes of 19% for V_{5Gy}, 37% for V_{10Gy}, 25% for V_{20Gy} and 18% for V_{30Gy}. V_{3Gy} was 5% lower for autoVMAT$_{5mm}$, but this was not statistically significant. Details on the comparisons between autoROBOT and autoVMAT$_{5mm}$ are presented in the right part of Table 9.1 and Figure 9.2. Favorable plan quality for autoROBOT compared to autoVMAT$_{5mm}$ is also observed in Figure 9.3 (right panels) and superiority of autoROBOT was also confirmed by the clinician scoring (Figure 9.4, lower panel).

For all 20 patients, the overall quality of the autoROBOT plans was considered better than for autoVMAT$_{5mm}$. For 11 patients, the clinician expected a real clinical impact of choosing the autoROBOT plan instead of the autoVMAT$_{5mm}$ plan, for other eight patients a possibly important impact was expected, and for one patient a quality gain with probably low impact was scored.

9.3.3 Dosimetric QA

All plans passed the QA tests, with average Gamma-passing rates of 98.7 ± 0.6% for autoROBOT, 99.8 ± 0.2% for autoVMAT$_{3mm}$ and 99.6 ± 0.8% for autoVMAT$_{5mm}$.

9.4 Discussion

In this study, we have presented the first system for fully automated generation of clinically deliverable treatment plans for a commercial robotic treatment unit. Automated planning, including non-coplanar beam angle selection, showed to improve plan quality, compared to manual planning. With equal PTV coverage, autoROBOT plans were superior to manual plans for all patients in sparing of the rectum and bladder, with negligible (but still superior) differences for all other clinical requirements. These findings are in line with results that we obtained on automated planning for regular

linacs, using a similar approach for automatic plan generation, see Section 9.2 and (Voet *et al.* 2013b, 2014; Sharfo *et al.* 2015; Della Gala *et al.* 2017). Apparently, interactive, manual planning is so complex and dependent on the planners' skills and allotted planning time, that an optimal planning solution can often not be guaranteed. The applied wish-list approach for automated planning features for each individual patient a systematic search for finding the dosimetric parameters of a *Pareto-optimal* plan with clinically desirable trade-offs between all objectives. A commercial planning system is then used to realize a clinically deliverable plan, using the attained plan parameters as constraints, without any further trial-and-error planning. As described in (Breedveld *et al.* 2012) and Section 9.6, a wish-list for a treatment site is developed based on the clinical treatment protocol and a few (typically 5) plans of recently treated patients. A wish-list configuration entails repeated automatic plan generations for the five patients, each time followed by an update of the wish-list that aims at a still higher plan quality in the next iteration. This iterative process is stopped if further improvements are considered not feasible. Specifically advantageous for autoROBOT planning, the upfront knowledge of feasible constraints allows the use of high resolution optimization grid in the commercial TPS for generating the deliverable plan.

Also for other systems, improvements in VMAT/IMRT plan quality by using automated planning have been reported (Fogliata *et al.* 2014; Hansen *et al.* 2016). Nelms *et al.* (2012) observed large plan quality variations between 125 manual planners from various institutes, even with a very detailed and quantitative description of planning goals. Berry *et al.* (2016a,b) showed large inter-planner variations in quality of plans that were manually generated within a single institution. Automated planning assisted in reducing the variations (Berry *et al.* 2016a,b). Clearly, further investigations on inconsistencies in manual planning and the potential role for automated planning are warranted.

Strong points of our comparison of robotic surgery with VMAT for prostate SBRT are *i)* the use of validated automated multi-criterial planning for both techniques (validation by systematic comparison with manual planning, see (Voet *et al.* 2014) for autoVMAT and Section 9.3 for autoROBOT) and *ii)* the use of the same TPS and exactly the same optimization scheme for initial plan optimization for the two techniques (wish-list, see Section 9.2). Due to these features, a bias-free comparison between robotic treatment and VMAT could be made, based on consistent, high quality plans.

Technique comparisons were performed using dosimetric (DVH) evaluations and by blind side-by-side plan scoring by the clinician responsible for prostate SBRT in our center. The clinician scoring has important added value compared to dosimetric analyses only, as it gives integrated views, considering the full dose distribution to OARs and PTV and the global clinical quality of the plan for each individual patient. In a

9

clinical setting, a clinician would never accept a plan comparison that is only based on DVH parameters.

AutoROBOT plans showed significant advantages compared to autoVMAT, both in the DVH analyses and the clinician's scoring. This was found for equal, 3 mm, CTV-PTV margins, and even stronger when comparing autoROBOT plans with 3 mm margin with autoVMAT plans with 5 mm margin. For 11 of the 20 patients, the autoVMAT plan with 5 mm margin was clinically unacceptable because of low PTV coverage (< 93%). On top of that, rectum, bladder and urethra doses were significantly higher compared to autoROBOT. For all patients, the autoROBOT plan had a largely reduced dose bath compared to both autoVMAT$_{3mm}$ and autoVMAT$_{5mm}$. The latter may especially be important for avoidance of secondary tumors in the increasing fraction of younger prostate cancer patients, related to Prostate-Specific Antigen screening.

A limitation of the study is that we did not clinically compare robotic treatment with VMAT, as needed for final conclusions, which we considered out of the scope of this paper. Another (practically unavoidable) limitation is that the autoVMAT and autoROBOT plans were calculated with different dose calculation engines as implemented in the corresponding TPS. Although both systems were thoroughly tested prior to clinical introduction, this might cause some bias in the comparisons.

Neither of the two recent studies that compared robotic treatment and VMAT for prostate SBRT (Lin *et al.* 2014b; MacDougall *et al.* 2014) observed the potential of a large plan quality improvement for robotic treatment, as observed in our study. Both the manual planning and retrospective plan comparisons, as used in these studies, may have introduced bias in the technique comparisons. MacDougall *et al.* (2014) used a 3 mm CTV-PTV margin for robotic treatment and a 5 mm margin for VMAT, and found no discernible dosimetric differences based on only six patients. Lin *et al.* (2014b) used a 3 mm margin for robotic treatment and for VMAT 5 mm in all directions, except 3 mm in posterior direction. They concluded that VMAT was preferable because of reduced treatment time and superior dose distribution conformality. The study showed however large and systematic differences between robotic treatment and VMAT in PTV dose inhomogeneity and PTV coverage, which could have influenced the conclusions.

Dong *et al.* (2014) compared VMAT with non-coplanar treatment at a C-arm linac, using with the 4π non-coplanar delivery approach involving both gantry rotations and couch displacements. For both techniques, the CTV-PTV margin was 5 mm with a reduction to 3 mm toward the rectum. As in our study, they observed clear plan quality advantages for non-coplanar treatment compared to coplanar VMAT. Automated plan generation was however only used for the non-coplanar planning, which could possibly have introduced some bias in the comparisons, favoring non-coplanar treatment. For robotic couch translations and rotations, Linthout *et al.* (2007) observed patient motion of up to 3 mm and 2°. Nonetheless, Dong *et al.* (2014) used the same CTV-PTV planning

margin for VMAT and non-coplanar treatment, possibly resulting in some study bias in favor of non-coplanar linac treatment. In our study, we investigated isotropic 3 mm and 5 mm CTV-PTV margins for autoVMAT. As our autoROBOT plans were already superior to VMAT with isotropic 3 mm margins, the same (and probably to a larger extent) is expected to hold for 5 mm margins with a reduction to 3 mm toward the rectum.

Delivery times of the autoROBOT plans generated in this study were around 45 min (Section 9.2), as used in our clinical practice for treatment with an IRIS™ variable aperture collimator, while VMAT treatments times were much shorter (~8–10 min). Most of the VMAT$_{5mm}$ plans were clinically unacceptable, and robotic treatment would anyway be preferable, also with the prolonged treatment time. For the other VMAT$_{5mm}$ plans, quality gains with robotic have to be weighed against the prolonged treatment duration. The same holds for VMAT$_{3mm}$ plans, that might be applicable at linacs with novel systems for intra-fraction motion correction (Willoughby *et al.* 2006; Keall *et al.* 2016). In this study, we have generated robotic plans for the variable aperture IRIS™ collimator. Currently, an MLC is available for the investigated robotic treatment unit (Asmerom *et al.* 2016; Fürweger *et al.* 2016), probably resulting in reduced delivery times (van de Water *et al.* 2011b; McGuinness *et al.* 2015; Jang *et al.* 2016; Kathriarachchi *et al.* 2016).

As described in Section 9.2, for robotic prostate SBRT plans, we try to mimic HDR brachytherapy dose distribution with intentionally inhomogeneous PTV dose delivery, with high peak doses inside the PTV. The urethra dose is minimized by dose–volume constraints. As the robot corrects for rotational tumor displacements, no PRV planning margin around the urethra is clinically used. C-arm linacs are not equipped with a system for rotation correction, implying that a PRV margin around the urethra may be needed for the inhomogeneous dose distributions studied in this paper. This could then possibly result in an enhanced percentage of patients with an unacceptably low PTV coverage. The need and implications of the use of a urethra PRV margin at a C-arm linac have not been investigated in this study.

9.5   Conclusions

The first system for fully automated generation of clinically deliverable plans for non-coplanar robotic treatment has been presented. The system features multi-criterial beam profile and beam angle optimization, resulting in plans with clinically favorable trade-offs between all treatment aims. For prostate SBRT, clinically acceptable, high quality plans could be generated that highly outperformed manually generated plans. Automatically generated robotic plans had consistently higher quality than automatically generated plans for VMAT at a linac. Further research on improvement of plan quality and plan consistency, including the role of automated planning, is warranted.

9.6 Wish-list generation for prostate SBRT

For each patient, Erasmus-iCycle automatically generates a *Pareto-optimal* plan with clinically favourable trade-offs between treatment objectives. Input for Erasmus-iCycle plan generation is a contoured CT-scan and a wish-list.

A wish-list contains the hard constraints, which always need to be fulfilled in plan generation, and treatment objectives with assigned priorities. Objectives are planning aims that need to be met as closely as possible (or superseded, if possible). Starting with the highest priority, the objectives in the wish-list are sequentially minimized, each time followed by adding the attained objective value as a novel constraint for the next optimization problem to ensure that high priority goal values will not be deteriorated in the minimization of lower priority objective functions (see (Breedveld *et al.* 2009, 2012) for more details). The wish-list is generated in an iterative procedure, starting with a first 'guess' of the wish-list by an experienced planner. This wish-list is then used for automated plan generation for a small group of (5-10) patients, followed by plan evaluation to update the wish-list for plan generation in a next iteration. This process with repeated wish-list updates stops if no further enhancement of plan quality is feasible. For groups of patients (e.g. all prostate patients treated with SBRT) this list is fixed, i.e. for all patients in the group the plan is fully automatically generated with the same wish-list.

As described in more details in Section 9.2, in this study, two versions of Erasmus-iCycle were used, one for non-coplanar plan generation for a robotic system equipped with a variable aperture collimator, the other for VMAT pre-optimization for an Elekta linac with MLC. Using the above described iterative procedure in parallel for the two Erasmus-iCycle versions, a single wish-list was generated for both robotic and VMAT pre-optimization.

Table 9.2 shows the wish-list used for all robotic and VMAT automated plan generations. All applied constraint and objective convex functions were used for the automated multi-criterial plan generation. These functions were selected to generate plans in line with the (not always convex) clinical planning aims (see Section 9.2.3). The SE function (Sum of Exponentials) defined by the equation below is basically a sum of exponentials of differences between attained voxel doses d_j and D_c, a user-defined critical dose.

$$\text{SE} = \frac{1}{m} \sum_{j=1}^{m} \exp^{\alpha(d_j - D_c)} \tag{9.1}$$

where m is the number of voxels in the structure and α is the sensitivity parameter.

For tumors, the parameter α is positive, and SE is equal to the Logarithmic Tumor Control (LTCP), as introduced by Alber & Reemtsen (2007), with D_c equal to the

Table 9.2: Applied wish-list for all study patients.

Constraints				
Structure	**Type**	**Limit**	**Parameters**	
PTV_{opt} [a]	Max	61.5 Gy		
Rectum	Max	36.5[b] Gy		
Rectum	gEUD[c]	28 Gy	$\alpha = 20$	
Rectum Mucosa	Max	27 Gy		
Overlap (Rectum, PTV+3mm)	Max	38 Gy		
Bladder	Max	39.5 Gy		
Bladder	gEUD	30.7 Gy	$\alpha = 20$	
Overlap (Bladder, PTV+3mm)	Max	41.8 Gy		
Urethra	Max	50 Gy		
Urethra	gEUD	39 Gy	$\alpha = 3$	
Penis Scrotum	Max	1.5 Gy		
PTV Shell 3 mm	Max	38 Gy		
PTV Shell 30 mm	Max	20 Gy		
Entrance Dose[d]	Max	20 Gy		

Objectives				
Priority	**Structure**	**Type**	**Goal**	**Parameters**
1	PTV_{opt}	↓ SE[e]	optimized[f]	$D_c = 37$ Gy, $\alpha = 0.9$, sufficient = as goal
2	PTV_{opt}	↓ SE	2.2	$D_c = 57$ Gy, $\alpha = 0.07$, sufficient = 2.2
3	CTV	↑ Min	34 Gy	sufficient = 34 Gy
4	Rectum	↓ SE	0 Gy	$D_c = 28$ Gy, $\alpha = -0.3$
5	Bladder	↓ SE	0 Gy	$D_c = 34$ Gy, $\alpha = -0.1$
6	Rectum	↓ Mean	0 Gy	
7	Bladder	↓ Mean	0 Gy	
8	Urethra	↓ Mean	0 Gy	
9	Dose Bath[g]	↓ Max	15 Gy	
10	Left Femoral Head	↓ Max	24 Gy	
	Right Femoral Head	↓ Max	24 Gy	

[a] PTV_{opt} is the PTV excluding overlaps with rectum, bladder and urethra.
[b] Maximum dose constraints were set lower than clinical requirements to account for voxel sampling for the optimizations.
[c] *Generalized Equivalent Uniform Dose* (Niemierko 1997).
[d] Dose in 2 cm thick layer inside the body contour.
[e] SE = *Sum of Exponentials*.
[f] Values are automatically set to ensure a PTV coverage of 95%, if feasible within the constraints, see text.
[g] Dose in patient volume in between shells at 3 cm from the PTV and 2 cm from the body contour.

prescribed tumor dose. The attractive characteristic of SE is that tumor underdosage is heavily penalized, while overdose has a relatively low impact on the function value. In Table 9.2, SE is used in priorities 1 and 2 to obtain clinically favorable PTV dose distributions. To limit for each patient both positive and negative deviations from the clinically requested 95% PTV coverage ($V_{100\%} = 95\%$), the goal value for SE in priority 1 is set automatically (Table 9.2). To this purpose, for priority 1, two plans are first generated with relatively small and large goal values, respectively. For both plans, the PTV coverage is then calculated and the final goal value for SE is determined by exponential interpolation. The aim of SE in priority 2 is creating a large dose inhomogeneity in the PTV like in HDR brachytherapy.

For prostate SBRT, it is extremely important that especially the high doses in rectum and bladder are avoided as much as possible. To this purpose, a SE function with negative α is used in priorities 4 and 5, highly favoring avoidance of doses higher than the defined critical values D_c (Table 9.2).

10

Late Toxicity in the Randomized Multi-Center HYPRO Trial for Prostate Cancer Analyzed with Automated Treatment Planning

Abdul Wahab M. Sharfo[1]**, Maarten L.P. Dirkx**[1]**, Rik G. Bijman**[1]**, Wilco Schillemans**[1]**, Sebastiaan Breedveld**[1]**, Shafak Aluwini**[1]**, Floris Pos**[2]**, Luca Incrocci**[1]** and Ben J.M. Heijmen**[1]

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

[2] Department of Radiation Oncology, Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands

Abstract

Purpose/Objective: Assess to what extent the use of automated treatment planning would have reduced organ-at-risk dose delivery observed in the randomized HYPRO trial for prostate cancer, and estimate related toxicity reductions. Investigate to what extent improved plan quality for hypo-fractionation scheme as achieved with automated planning can potentially reduce observed enhanced toxicity for the investigated hypo-fractionation scheme to levels observed for conventional-fractionation scheme.

Material/Methods: For 725 trial patients, VMAT plans were generated with an algorithm for automated multi-criterial plan generation (autoVMAT). All clinically delivered plans (CLINICAL), generated with commonly applied interactive trial-and-error planning were also available for the investigations. Analyses were based on dose-volume histograms (DVH) and predicted normal tissue complication probabilities (NTCP) for late gastrointestinal (GI) toxicity.

Results: Compared to CLINICAL, autoVMAT plans had similar or higher PTV coverage, while large and statistically significant OAR sparing was achieved. Mean doses in the rectum, anus and bladder were reduced by 7.8 ± 4.7 Gy, 7.9 ± 6.0 Gy and 4.2 ± 2.9 Gy, respectively ($p < .001$). NTCPs for late grade ≥ 2 GI toxicity, rectal bleeding and stool incontinence were reduced from 23.3 ± 9.1% to 19.7 ± 8.9%, from 9.7 ± 2.8% to 8.2 ± 2.8%, and from 16.8 ± 8.5% to 13.1 ± 7.2%, respectively ($p < .001$). Reductions in rectal bleeding NTCP were observed for all published Equivalent Uniform Dose volume parameters, n. AutoVMAT allowed hypo-fractionation with predicted toxicity similar to conventional-fractionation with CLINICAL plans.

Conclusions: Compared to CLINICAL, autoVMAT had superior plan quality, with meaningful NTCP reductions for both conventional-fractionation and hypo-fractionation schemes. AutoVMAT plans might reduce toxicity for hypo-fractionation to levels that were clinically observed (and accepted) for conventional-fractionation. This may be relevant when considering clinical use of the investigated hypo-fractionation schedule with relatively high fraction dose (3.4 Gy).

10.1   Introduction

Prostate cancer is the second most common cancer in men worldwide (Torre *et al.* 2015). External beam radiation therapy (EBRT) is one of the primary treatment modalities for patients with localized or locally advanced prostate cancer (Resnick *et al.* 2013; Hamdy *et al.* 2016). In the last decade, substantial improvements in EBRT techniques have been made, resulting in significant improvements in the treatment of prostate cancer patients (Peeters *et al.* 2006; Dearnaley *et al.* 2007; Bekelman *et al.* 2011; Zelefsky *et al.* 2012). Dose escalation has significantly improved treatment outcome in patients with localized prostate cancer (Peeters *et al.* 2006; Dearnaley *et al.* 2007). However, this was often associated with a significant increase in toxicity (Peeters *et al.* 2005; Dearnaley *et al.* 2007), including rectal bleeding, fecal incontinence and changes in bowel habits (Michalski *et al.* 2010). Between 2007 and 2010, the randomized phase 3 multi-center HYPRO trial for intermediate- or high-risk localized prostate cancer investigated whether hypo-fractionated EBRT (19 fractions of 3.4 Gy) could improve relapse-free survival without increasing toxicity, compared to conventionally-fractionated radiotherapy (39 fractions of 2.0 Gy) (Aluwini *et al.* 2015, 2016; Incrocci *et al.* 2016).

Since the end of patient accrual for the HYPRO trial, important improvements have been reported in treatment planning, in particular related to automated treatment plan generation. In Rotterdam, Erasmus-iCycle has been developed for fully automated, multi-criterial generation of *Pareto-optimal* plans with clinically desired balances between all treatment objectives (Breedveld *et al.* 2007, 2012). The optimized Erasmus-iCycle plans are automatically re-constructed with the Monaco® treatment planning system (TPS) (Elekta AB, Stockholm, Sweden) to generate clinically deliverable plans. Compared to manual trial-and-error planning, automated planning has improved plan quality and consistency and drastically reduced treatment planning workload (Voet *et al.* 2013b; Sharfo *et al.* 2016; Della Gala *et al.* 2017; Heijmen *et al.* 2018).

Aluwini *et al.* (2016) observed enhanced late toxicity in the hypo-fractionation arm of the HYPRO trial compared to conventional-fractionation. In this study, Erasmus-iCycle/Monaco was used to automatically generate deliverable treatment plans for HYPRO patients and compare them with the actually delivered, manually generated plans. As hypo-fractionation has logistic advantages for both the patient and the treatment center, we also investigated whether replacement of CLINICAL planning for hypo-fractionation by automated planning could avoid toxicity increases relative to accepted toxicity for conventional-fractionation with CLINICAL planning. Apart from using dose-volume histogram (DVH) parameters, analyses were also based on calculated normal tissue complication probabilities (NTCP) for late gastrointestinal (GI) toxicity.

10.2 Materials and methods

10.2.1 Patients and clinical treatment plans

Between 2007 and 2010, in total 820 patients with intermediate- or high-risk prostate cancer were included in the HYPRO trial by seven Dutch radiotherapy centers. Patients were randomized to receive 78 Gy in 39 fractions (5fr/wk) (conventional-fractionation), or 64.6 Gy in 19 fractions (3fr/wk) (hypo-fractionation) (Aluwini *et al.* 2015, 2016; Incrocci *et al.* 2016). Three treatment groups were defined based on the risk of seminal vesicles' involvement. For group 1, the clinical target volume (CTV) consisted of the prostate only to be treated to the prescribed dose. For group 2, the prostate was treated to the prescribed dose, while the seminal vesicles were treated to a lower dose. For group 3, both the prostate and the seminal vesicles were treated up to the prescribed dose (Aluwini *et al.* 2015). Patients were treated with a simultaneous-integrated-boost technique (SIB), either for dose reduction in the seminal vesicles (group 2) and/or for delivery of part of the dose with a reduced planning target volume (PTV) margin (all groups). For the large target (PTV_2), the prescribed dose was then reduced to 72.15 Gy in the conventional-fractionation arm or 57.76 Gy in case of hypo-fractionation, instead of 78 Gy or 64.6 Gy as used for the boost (PTV_1). Depending on the set-up verification and correction strategy used in each participating institute, margins of 3–10 mm were added to the clinical target volume (CTV) (equal in both groups), yielding the PTVs. For the boost, it was allowed to reduce the margin toward the rectum to 0 mm. All centers applied the same dose constraints. The rectal volume receiving ≥ 65 Gy (EQD_{2Gy} for α/β = 3 Gy) had to be < 50%, and the anal mean dose < 60 Gy (EQD_{2Gy}). For bladder and femoral heads no constraints were specified (Aluwini *et al.* 2015).

A total of 725 HYPRO patients had evaluable clinical treatment plans, including scans, doses, and contoured structures, and could be included in this study. 361 patients were treated with conventional-fractionation (75 in group 1, 219 in group 2, and 67 in group 3), and 364 patients were treated with hypo-fractionation (73 in group 1, 219 in group 2, and 72 in group 3). Of the included patients, 95.6% was treated using intensity-modulated radiotherapy (IMRT) with 5–15 beams (median: 7 beams); the remainder with volumetric-modulated arc therapy (VMAT). In all cases, a multi-leaf collimator (MLC) with a leaf width of 1 cm was used. In 96.4% of the plans, 10 MV photon beams were applied, in 3.2% 18 MV photon beams, and for the remainder, 6 MV photon beams. The treatment plans used in this trial ('CLINICAL' plans) were generated with conventional, interactive trial-and-error planning (designated in this paper by 'manual planning'), using the Monaco® or Pinnacle TPS.

10

10.2.2 autoVMAT vs. CLINICAL

For each of the included patients, Erasmus-iCycle/Monaco was used to fully automatically generate a treatment plan for a modern treatment technique and modern equipment, i.e. VMAT at an Elekta Synergy® treatment machine (Elekta AB, Sweden), equipped with an Agility™ MLC with 160 leaves of 0.5 cm in width ('autoVMAT'). Like for the vast majority of CLINICAL plans (above), a 10 MV photon beam was used. The autoVMAT plans were compared with the CLINICAL plans.

Although most of the included HYPRO patients were treated with an IMRT plan (above), for automated planning we chose for VMAT, as this is currently the common treatment approach for prostate cancer in the Netherlands. For a sub-group of 60 patients, we investigated to what extent our conclusions regarding plan quality differences between CLINICAL and automated planning depended on the difference in treatment technique. For this analysis, we used a sub-group of 60 patients treated with IMRT, consisting of 10 randomly selected patients for the 6 combinations of treatment group/fractionation arm. Erasmus-iCycle/Monaco was also used to automatically generate an IMRT plan for the clinical beam configuration ('autoIMRT'). Because of clinically applied 1 cm leaves (above), for the same sub-group of 60 patients, VMAT plans for the MLCi2 collimator with a leaf width of 1 cm (Elekta AB, Sweden) were automatically generated as well, and compared with the autoVMAT plans for the Agility™ MLC with 0.5 cm leaves. For all included patients, the automatically generated plan(s) had the same clinical intent as the clinical plan, i.e., same fractionation, risk-group, etc.

Generally applied dosimetric parameters such as PTV coverage, near-minimum and near-maximum PTV doses, and mean and maximum doses in organs-at-risk (OARs) were used for plan evaluation and comparison. In addition, assuming an α/β-ratio of 3 Gy, we evaluated the percentage of rectum volume receiving equivalent doses in 2 Gy fractions (EQD_{2Gy}) of more than 65 Gy and 75 Gy (V_{65GyEq} and V_{75GyEq}, respectively). The latter parameters are associated with grade ≥ 2 GI toxicity and rectal bleeding (Michalski *et al.* 2010). Differences in predicted NTCPs for late GI toxicity were also quantified (see below). To apply the models, for both fractionation schemes, rectal doses were first converted into equivalent doses for 2 Gy fractions (EQD_{2Gy}), using α/β = 3 Gy.

10.2.3 Hypo-fractionation with autoVMAT vs. conventional-fractionation with CLINICAL Plans

For comparison of plans for the two fractionation schemes, all rectal and bladder doses were first expressed in terms of EQD_{2Gy} using an α/β-ratio of 3 Gy. For the rectum, these doses were then used to compare predicted NTCPs, V_{65GyEq}, V_{75GyEq}, and D_{mean}. For the bladder, D_{mean} was calculated and compared.

10.2.4 NTCP modeling for late GI toxicity

Aluwini *et al.* (2016) have described the late GI toxicity (follow-up > 3 months) as observed in the HYPRO trial. Late toxicity was scored at 6 months and at 1, 2, 3, 4 and 5 years after treatment by case report form according to criteria from the Radiation Therapy Oncology Group and the European Organization for Research and Treatment of Cancer (RTOG/EORTC) (Cox *et al.* 1995) and by patients' self-assessment questionnaires. Details are provided in a previous publication (Aluwini *et al.* 2016).

NTCP models for toxicity observed in the HYPRO trial were generated for grade ≥ 2 late GI toxicity, stool incontinence, stool frequency, rectal bleeding and GI proctitis, using delivered rectal dose distributions and reported toxicities of the included 725 HYPRO patients. For NTCP modeling, all rectal DVHs were first expressed in terms of EQD_{2Gy} using an α/β-ratio of 3 Gy, in order to combine patients from the two treatment arms in the NTCP modeling. Rectum dosimetric parameters gEUD(n) with the volume-effect parameter, n, ranging from 0 to 1 in steps of 0.01 were extracted from the converted rectal DVHs. These dosimetric parameters, along with the clinical covariates (baseline GI toxicity, diabetes, treatment group, Gleason, age, prostate volume, abdominal surgery, etc.) were considered in the establishment of NTCP models with multi-variate logistic regression, i.e.

$$NTCP = \frac{1}{1 + e^{-s}} \text{ with } s = \beta_0 \sum_{i=1}^{n} \beta_i . x_i$$

with x_i representing the different input variables and b_i the corresponding regression coefficients. For generation of all NTCP models we strictly followed the approach by Cicchetti *et al.* (2016). Model parameters are presented in Table 10.1.

The derived model for rectal bleeding had EUD ($n = 0.13$) as the dosimetric input variable (Table 10.1). In the literature, reported n values for rectal bleeding vary between 0.03 and 0.18 (Rancati *et al.* 2011; Gulliford *et al.* 2012). To investigate the impact of n on differences between automated and manual planning in calculated NTCP-values, we also derived two alternative NTCP-models from the clinical data, one fixing n at 0.03 and the other for $n = 0.18$. For stool incontinence, rectal EUD ($n = 0.20$) fitted best the clinical observations. In the literature $n = 1$ (mean dose) has been reported (Rancati *et al.* 2011). Again, to explore the sensitivity on n, a separate NTCP-model for $n = 1$ was generated.

10.2.5 Automated treatment planning using Erasmus-iCycle/ Monaco

Automated plan generation with Erasmus-iCycle is based on a tumor site-specific '*wish-list*' with hard constraints to be strictly obeyed, and planning objectives with ascribed priorities, used to ensure clinically favorable trade-offs between treatment

10

Table 10.1: NTCP model information for late GI symptoms.

Variable	Coefficient	p-value
Late GI Grade ≥ 2 Toxicity	*GOF (p = .465)*	
Baseline	2.083	**.015**
Diabetes	0.608	**.012**
T-Group 3vs12	0.406	**.032**
EUD(0.13)	0.084	**< .001**
Constant	-6.362	**< .001**
Stool Incontinence	*GOF (p = .192)*	
Baseline	1.453	.082
Age	0.059	**.001**
Diabetes	0.572	**.034**
EUD(0.2)	0.089	**< .001**
Constant	-10.579	**<0.001**
Rectal Bleeding	*GOF (p = .364)*	
EUD(0.13)	0.072	**.033**
Gleason	-0.201	.147
Constant	-5.025	**.017**
Stool Frequency	*GOF (p = .029)*	
Gleason	0.281	.073
Abdominal Surgery	-0.936	.054
EUD(0.11)	0.069	.114
Constant	-8.815	**.002**
Proctitis	*GOF (p = .332)*	
Prostate Volume	0.011	.075
Diabetes	1.203	.056
T-Group 2vs13	-0.917	**.012**
EUD(0.14)	0.096	**.043**
Constant	-12.198	**< .001**

Abbreviations: NTCP = *Normal Tissue Complication Probability*, GI = *Gastro-Intestinal*, T-Group 3vs12 = treatment group 3 vs. 1&2, T-Group 2vs13 = treatment group 2 vs. 1&3. GOF = *Goodness of Fit*, EUD = *Equivalent Uniform Dose*.

goals in a multi-criterial optimization (Voet *et al.* 2013b; Sharfo *et al.* 2016; Della Gala *et al.* 2017; Heijmen *et al.* 2018). In this study, two comparable wish-lists were developed for automated plan generation, one for conventional-fractionation, the other for hypo-fractionation. Conversion into deliverable clinical plans was performed with Monaco®, version 5.11 (Elekta AB, Stockholm, Sweden). Achieving adequate PTV coverage (i.e., $V_{95\%} \geq 99\%$) had highest priority in plan generation, followed by minimization of the doses to the rectum, anus, bladder and femoral heads. The wish-list used in this study is presented in Table 10.2. Dose calculations in Monaco® were performed with a Monte Carlo dose engine using a 1% dose variance, and a dose grid resolution of 3 mm. For VMAT plans, a maximum of 150 control points per arc was allowed. Multi-criterial plan optimization with Erasmus-iCycle took between 15 and 60 min, depending on the patient geometry. Subsequent plan reconstruction with the Monaco® TPS took between 10 and 30 min.

10.2.6 Statistical analysis

To assess statistical significance of observed differences in treatment plans, the paired samples *t-test* was applied in the comparisons of autoVMAT vs. CLINICAL, and the

Table 10.2: Applied wish-list for automatic VMAT plan generation for conventional- and hypo-fractionated EBRT.

Constraints			
Structure	**Type**	**Limit**	
PTV_P	Max	105% of D_{high}	
PTV_P	Mean	101% of D_{high}	
$PTV_{SV} - PTV_{P+3mm}$	Max	95% of D_{high}	
Rectum & Anus	Max	102% of D_{high}	
PTV_{SV} Shell 50 mm	Max	50% of D_{high}	
Unspecified Tissues	Max	105% of D_{high}	

Objectives				
Priority	**Structure**	**Type**	**Goal**	**Parameters**
1	PTV_P	↓ LTCP	0.8	$D_{high} = 78^*, 64.6^\dagger$ Gy, $\alpha = 0.8$
2	PTV_{SV}	↓ LTCP	0.4	$D_{low} = 72.2^*, 57.8^\dagger$ Gy, $\alpha = 1$
3	Rectum	↓ EUD	20 Gy	$k = 12$
4	Rectum	↓ EUD	10 Gy	$k = 8$
5	PTV_{SV} Shell 5 mm	↓ Max	80% of D_{low}	
	Skin Ring 20 mm	↓ Max	20% of D_{low}	
6	Rectum	↓ Mean	5 Gy	
7	Anus	↓ Mean	5 Gy	
8	Bladder	↓ Mean	5 Gy	
9	PTV_{SV} Shell 15 mm	↓ Max	50% of D_{low}	
	PTV_{SV} Shell 25 mm	↓ Max	30% of D_{low}	
10	Left & Right Femoral Head	↓ Max	50% of D_{high}	

Abbreviations: PTV_P = planning target volume of the prostate, PTV_{SV} = planning target volume of the seminal vesicles. D_{high}, D_{low} = high and low prescribed doses, PTV_{P+3mm} = a 3mm transition region within PTV_{SV}, LTCP = *Logarithmic Tumor Control Probability*, EUD = Equivalent Uniform Dose, α = cell sensitivity, k = volume parameter.
*, †: Applied values for the conventional- and hypo- fractionation regimen, respectively.

independent samples *t-test* was used to compare hypo-fractionation with autoVMAT with conventional-fractionation based on CLINICAL plans (SPSS® v.21; SPSS Inc., Chicago, USA). Differences were considered statistically significant for $p < 0.05$. Goodness of fit analyses were performed to assess the quality of established NTCP-models, with $p > .05$ pointing at adequate prediction performance.

10.3 Results

10.3.1 autoVMAT vs. CLINICAL

Differences between autoVMAT and CLINICAL in coverage of PTV_1 and PTV_2 were generally small (Figure 10.1, Figure 10.2, and Table 10.3). Coverage of PTV_1 by at least 95% of the prescribed dose was enhanced for autoVMAT, which was more pronounced for patients treated in the hypo-fractionation arm (0.4 ± 1.6%, $p < .001$), compared to the conventional-fractionation arm (0.1 ± 0.9% $p = .005$). For PTV_2, dose coverage by at least 95% of the prescribed dose was on average equivalent for the autoVMAT and the CLINICAL plans in the hypo-fractionation arm (0.0 ± 0.7%, $p = .213$), but statistically significantly increased with autoVMAT in the conventional arm (0.1 ± 1.1%, $p = .04$). The near-minimum dose ($D_{98\%}$) in PTV_1 was slightly higher with autoVMAT
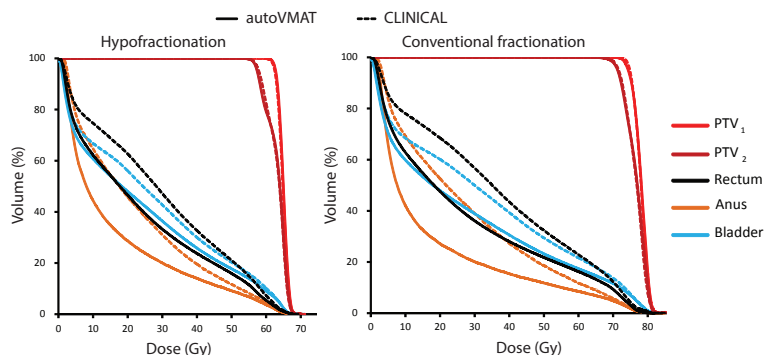
Figure 10.1: Comparisons of population-averaged dose-volume histograms for the manually generated (CLINICAL) and automatically generated (autoVMAT) plans, for both hypo-fractionation (left, 364 patients) and conventional-fractionation (right, 361 patients).

for both fractionation arms (0.1 ± 0.9 Gy, p = .022). In contrast, for PTV_2, $D_{98\%}$ was statistically significantly reduced (-0.4 ± 1.9 Gy, $p < .001$). For the autoVMAT plans, the near-maximum dose ($D_{2\%}$) in PTV_1 was enhanced by 0.5 ± 1.1 Gy ($p < .001$). Conformality was significantly better for autoVMAT, while the CLINICAL plans had better PTV dose homogeneity (Table 10.3). Overall, observed homogeneity differences between autoVMAT and CLINICAL are considered of minor clinical importance, certainly in the context of the large differences observed for OARs, as described in the next paragraph. AutoVMAT showed largely reduced OAR doses (Figure 10.1, Figure 10.2, and Table 10.3). Figure 10.1 shows superiority of autoVMAT in population-averaged DVHs for rectum anus and bladder. Mean doses in the rectum, anus and bladder were on average reduced by 7.8 ± 4.7 Gy, 7.9 ± 6.0 Gy and 4.2 ± 2.9 Gy ($p < .001$), respectively. Rectum V_{65GyEq} and V_{75GyEq} were reduced by 4.2 ± 3.9 and 1.5 ± 2.4% ($p < .001$), and maximum doses in the left and right femoral heads by 8.3 ± 6.1 Gy and 9.1 ± 5.8 Gy ($p < .001$). Due to the reduced dose delivery to the rectum in the autoVMAT plans, the rectal NTCPs were significantly lower for all treatment groups in both fractionation arms (Figure 10.3 and Table 10.4). Averaged over both fractionation arms and all treatment groups, NTCPs were reduced from 23.3 ± 9.1% to 19.7 ± 8.9%, from 16.8 ± 8.5% to 13.1 ± 7.2%, from 9.7 ± 2.8% to 8.2 ± 2.8%, from 5.6 ± 2.9% to 4.9 ± 2.5%, and from 3.4 ± 3.3% to 2.6 ± 2.5% for late grade ≥ 2 GI toxicity, stool incontinence, rectal bleeding, stool frequency, and proctitis, respectively. The reported reduction of 1.5% (above) in rectal bleeding NTCP for n = 0.13 was lower for n = 0.03 (0.8%) and higher for n = 0.18 (1.7%). For stool incontinence, n = 0.2 was found for the HYPRO data, with the overall NTCP gain for autoVMAT of 3.7% as reported above. When assuming n = 1 for stool incontinence as reported in the literature, the NTCP gain was slightly lower, i.e. 3.2%. All reported differences were statistically significant. For more details on NTCP comparisons, see Figure 10.3 and Table 10.4.

10

Table 10.3: Comparison of dosimetric parameters for automatically generated VMAT plans (AUTO) and clinically delivered manually generated (CLINICAL) plans.

| | | Group 1 | | Group 2 | | Group 3 | | All Groups | |
| | Fractionation | AUTO | CLINICAL - AUTO | AUTO | CLINICAL - AUTO | CLINICAL - AUTO | AUTO | AUTO | CLINICAL - AUTO |
		Avg. ± SD	Δmean ± SD	Avg. ± SD	Δmean ± SD	Δmean ± SD	Avg. ± SD	Avg. ± SD	Δmean ± SD
PTV_1									
$V_{95\%}$ (%)	Conventional	99.5 ± 0.4	-0.1 ± 0.6‡	99.6 ± 0.6	-0.2 ± 0.9*	-0.1 ± 1.1‡	99.1 ± 0.7	99.5 ± 0.6	-0.1 ± 0.9*
	Hypo	99.6 ± 0.5	-0.5 ± 1.2*	99.4 ± 0.6	-0.4 ± 1.7	-0.5 ± 1.8*	99.3 ± 0.6	99.4 ± 0.6	-0.4 ± 1.6
$D_{98\%}$ (Gy)	Conventional	75.1 ± 0.4	0.2 ± 1.0‡	75.3 ± 0.5	-0.1 ± 1.0‡	0.0 ± 0.8‡	74.9 ± 0.4	75.2 ± 0.5	0.0 ± 1.0‡
	Hypo	62.3 ± 0.4	-0.2 ± 0.8‡	62.2 ± 0.4	-0.1 ± 0.8*	-0.2 ± 0.9‡	62.2 ± 0.4	62.2 ± 0.4	-0.1 ± 0.8*
$D_{2\%}$ (Gy)	Conventional	81.3 ± 0.2	-0.8 ± 1.0	81.2 ± 0.3	-0.6 ± 1.0	-0.8 ± 1.0	81.3 ± 0.2	81.2 ± 0.3	-0.7 ± 1.0
	Hypo	67.5 ± 0.1	-1.2 ± 0.7	66.6 ± 0.6	0.0 ± 1.0‡	-0.9 ± 0.8	67.6 ± 0.1	67.0 ± 0.6	-0.4 ± 1.0
$CI_{95\%}$	Conventional	1.17 ± 0.26	0.13 ± 0.22	1.27 ± 0.30	0.12 ± 0.27	0.00 ± 0.17‡	1.29 ± 0.25	1.25 ± 0.28	0.10 ± 0.25
	Hypo	1.17 ± 0.09	0.11 ± 0.17	1.25 ± 0.16	0.08 ± 0.18	0.10 ± 0.17	1.22 ± 0.16	1.23 ± 0.15	0.09 ± 0.18
$CI_{50\%}$	Conventional	4.46 ± 1.05	1.95 ± 1.33	5.96 ± 1.54	2.22 ± 1.77	2.25 ± 1.42	4.56 ± 0.95	5.39 ± 1.53	2.17 ± 1.62
	Hypo	4.20 ± 0.78	2.13 ± 1.49	6.13 ± 1.50	1.89 ± 1.94	2.06 ± 1.36	4.51 ± 1.03	5.42 ± 1.56	1.97 ± 1.75
HI	Conventional	0.08 ± 0.01	-0.01 ± 0.02*	0.07 ± 0.01	-0.01 ± 0.01	-0.01 ± 0.01	0.08 ± 0.01	0.08 ± 0.01	-0.01 ± 0.01
	Hypo	0.08 ± 0.01	-0.01 ± 0.01	0.07 ± 0.01	0.00 ± 0.01‡	-0.01 ± 0.02	0.08 ± 0.01	0.07 ± 0.01	0.00 ± 0.02
PTV_2									
$V_{95\%}$ (%)	Conventional	99.8 ± 0.3	0.0 ± 0.6‡	99.7 ± 0.6	-0.2 ± 1.1	0.1 ± 1.3‡	99.3 ± 0.5	99.7 ± 0.6	-0.1 ± 1.1*
	Hypo	99.8 ± 0.2	0.0 ± 0.3‡	99.7 ± 0.3	-0.1 ± 0.8*	0.2 ± 0.5*	99.6 ± 0.3	99.7 ± 0.3	0.0 ± 0.7‡
$D_{98\%}$ (Gy)	Conventional	71.3 ± 0.8	0.3 ± 1.7‡	70.5 ± 2.2	0.3 ± 2.7‡	0.2 ± 1.9‡	70.6 ± 1.0	70.6 ± 1.9	0.3 ± 2.4‡
	Hypo	57.8 ± 1.0	0.5 ± 1.3*	56.6 ± 0.7	0.4 ± 1.1	0.9 ± 1.9*	57.3 ± 0.9	56.9 ± 0.9	0.5 ± 1.3
Rectum									
D_{mean} (Gy)	Conventional	21.4 ± 5.7	7.7 ± 3.9	27.9 ± 6.5	10.3 ± 4.8	8.7 ± 4.6	28.4 ± 7.4	26.7 ± 7.1	9.4 ± 4.7
	Hypo	19.6 ± 5.0	3.6 ± 3.0	23.4 ± 6.0	7.3 ± 4.2	4.9 ± 3.1	25.6 ± 5.3	23.1 ± 6.0	6.1 ± 4.1
V_{75GyEq} (%)	Conventional	2.9 ± 2.7	0.8 ± 1.8	2.8 ± 2.2	1.3 ± 2.6	1.4 ± 2.8	3.9 ± 5.3	3.0 ± 3.1	1.2 ± 2.5
	Hypo	4.3 ± 3.6	1.3 ± 1.7	4.2 ± 3.7	1.9 ± 2.4	1.6 ± 1.9	7.3 ± 6.0	4.8 ± 4.4	1.7 ± 2.2
V_{65GyEq} (%)	Conventional	9.8 ± 3.6	2.8 ± 2.9	14.1 ± 5.6	5.7 ± 4.6	4.5 ± 3.5	14.0 ± 7.9	13.2 ± 6.0	4.8 ± 4.2
	Hypo	8.7 ± 3.6	2.2 ± 2.5	11.7 ± 5.6	4.1 ± 3.8	3.2 ± 2.9	13.1 ± 6.7	11.4 ± 5.7	3.5 ± 3.5
Anus									
D_{mean} (Gy)	Conventional	15.7 ± 7.4	9.5 ± 5.2	16.8 ± 9.3	10.6 ± 6.8	8.0 ± 5.9	17.7 ± 10.2	16.8 ± 9.1	9.9 ± 6.4
	Hypo	17.1 ± 7.9	5.5 ± 4.9	15.3 ± 7.7	6.5 ± 5.0	4.9 ± 4.3	15.7 ± 8.8	15.7 ± 8.0	6.0 ± 4.9
Bladder									
D_{mean} (Gy)	Conventional	23.6 ± 12.0	4.7 ± 2.5	28.2 ± 12.7	5.9 ± 3.2	4.7 ± 3.0	28.6 ± 10.3	27.3 ± 12.3	5.4 ± 3.1
	Hypo	19.5 ± 9.4	2.5 ± 1.7	25.0 ± 10.7	3.1 ± 2.2	2.7 ± 1.9	25.1 ± 9.0	23.9 ± 10.4	2.9 ± 2.1
Right Femoral Head									
D_{max} (Gy)	Conventional	30.3 ± 3.4	10.5 ± 4.5	31.2 ± 4.6	8.8 ± 6.8	9.9 ± 6.4	30.8 ± 3.2	30.9 ± 4.1	9.4 ± 6.3
	Hypo	24.2 ± 3.1	10.0 ± 4.9	24.8 ± 3.3	8.5 ± 5.5	8.1 ± 4.9	24.7 ± 3.4	24.7 ± 3.3	8.8 ± 5.3
Left Femoral Head									
D_{max} (Gy)	Conventional	30.9 ± 4.0	10.3 ± 5.4	32.5 ± 4.9	8.3 ± 7.0	9.8 ± 5.9	32.1 ± 3.3	32.1 ± 4.5	9.0 ± 6.6
	Hypo	24.9 ± 3.7	9.2 ± 4.9	26.6 ± 3.5	6.9 ± 6.0	7.5 ± 4.7	26.6 ± 4.0	26.2 ± 3.7	7.6 ± 5.6

Abbreviations: SD = *Standard Deviation*, CI = *Conformity Index*, HI = *Homogeneity Index*.
All differences without mark are statistically significantly with $p < .001$. Differences marked with * are statistically significant with $.001 < p < .05$. Differences marked with † are statistically significant with $.001 < p < .05$. Differences marked with ‡ are not statistically significant $p > .05$.
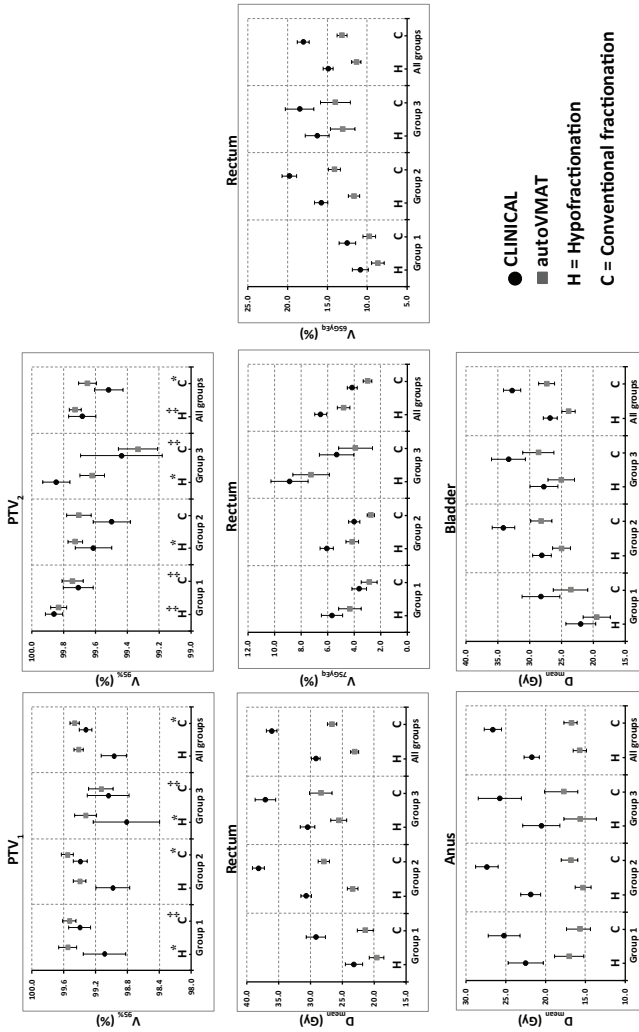
10

Figure 10.2: Comparison of dosimetric parameters for autoVMAT and the manually generated (CLINICAL) plans in the hypo-fractionation (H) and the conventional-fractionation (C) treatment arms. Error bars indicate the 95% confidence interval of the average values observed in the involved plans. Averages for hypo-fractionation and conventional arms were calculated for 364 patients (73 in group 1, 219 in group 2, and 72 in group 3) and 361 patients (75 in group 1, 219 in group 2, and 67 in group 3), respectively. For group 1, the clinical target volume (CTV) consisted of the prostate only to be treated to the prescribed dose. For group 2, the prostate was treated to the prescribed dose, while the seminal vesicles were treated to a lower dose. For group 3, both the prostate and the seminal vesicles were treated up to the prescribed dose. All differences without mark are statistically significant with p < .001. Differences marked with * are statistically significant with .001 < p < .05. Differences marked with ‡ are not statistically significant (p > .05).
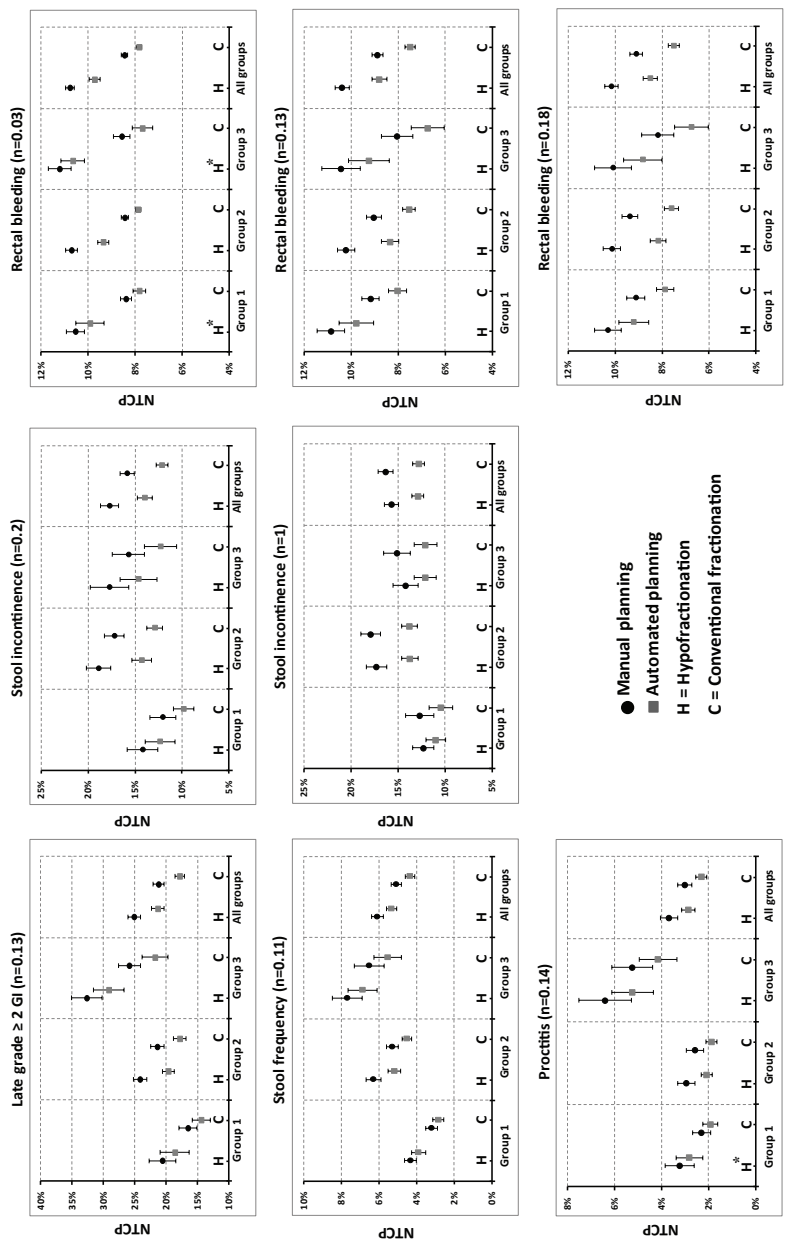
Figure 10.3: Comparison of calculated normal tissue complication probabilities (NTCP) for GI toxicity for automatically (autoVMAT) and manually (CLINICAL) generated plans for patients in the hypo-fractionation (H) and conventional-fractionation (C) treatment arms. Error bars indicate the 95% confidence interval of the average values observed in the involved plans. Averages for hypo-fractionation and conventional arms were calculated for 364 patients (73 in group 1, 219 in group 2, and 72 in group 3) and 361 patients (75 in group 1, 219 in group 2, and 67 in group 3), respectively. All differences were statistically significant in favor of automated planning with *p* < .05.

10

Table 10.4: Comparison of calculated normal tissue complication probabilities (NTCP) for gastrointestinal (GI) toxicity between automatically generated VMAT plans (AUTO) and the clinically delivered manually generated plans (CLINICAL).

	Fractionation	Group 1		Group 2		Group 3		All Groups	
		AUTO	CLINICAL - AUTO	AUTO	CLINICAL - AUTO	AUTO	CLINICAL - AUTO	AUTO	CLINICAL - AUTO
		Avg. ± SD	Δmean ± SD	Avg. ± SD	Δmean ± SD	Avg. ± SD	Δmean ± SD	Avg. ± SD	Δmean ± SD
Late GI NTCP (%)	Conventional	14.5 ± 6.3	2.1 ± 2.2	17.9 ± 7.4	3.6 ± 3.1	21.9 ± 8.7	4.1 ± 3.6	17.9 ± 7.8	3.4 ± 3.1
(*n* = 0.13)	Hypo	18.7 ± 10.1	2.0 ± 3.8	19.7 ± 7.5	4.5 ± 4.1	29.2 ± 10.7	3.5 ± 3.5	21.4 ± 9.6	3.8 ± 4.1
Bleeding NTCP (%)	Conventional	7.8 ± 1.2	0.6 ± 0.9	7.9 ± 0.8	0.6 ± 1.3	7.7 ± 1.8	0.9 ± 1.2	7.9 ± 1.1	0.6 ± 1.3
(*n* = 0.03)	Hypo	9.9 ± 2.6	0.6 ± 2.1*	9.4 ± 1.8	1.4 ± 1.7	10.7 ± 2.1	0.6 ± 1.3*	9.7 ± 2.1	1.0 ± 1.7
Bleeding NTCP (%)	Conventional	8.0 ± 1.7	1.1 ± 1.2	7.5 ± 2.1	1.5 ± 1.3	6.8 ± 2.9	1.3 ± 1.2	7.5 ± 2.2	1.4 ± 1.3
(*n* = 0.13)	Hypo	9.8 ± 3.3	1.1 ± 2.0	8.4 ± 2.7	1.9 ± 1.8	9.3 ± 3.8	1.2 ± 1.3	8.8 ± 3.1	1.6 ± 1.8
Bleeding NTCP (%)	Conventional	7.9 ± 1.6	1.2 ± 1.2	7.6 ± 2.2	1.8 ± 1.3	6.8 ± 2.9	1.4 ± 1.2	7.5 ± 2.3	1.6 ± 1.3
(*n* = 0.18)	Hypo	9.2 ± 2.7	1.1 ± 1.7	8.2 ± 2.6	2.0 ± 1.7	8.8 ± 3.5	1.3 ± 1.3	8.5 ± 2.9	1.7 ± 1.7
Stool incontinence NTCP (%)	Conventional	9.9 ± 4.9	2.2 ± 2.4	12.9 ± 6.2	4.3 ± 3.5	12.3 ± 7.0	3.4 ± 2.6	12.2 ± 6.2	3.7 ± 3.3
(*n* = 0.2)	Hypo	12.4 ± 7.0	1.8 ± 3.2	14.3 ± 8.1	4.6 ± 4.3	14.7 ± 8.4	3.1 ± 3.0	14.0 ± 8.0	3.7 ± 4.0
Stool incontinence NTCP (%)	Conventional	10.5 ± 5.5	2.2 ± 1.7	13.8 ± 6.3	4.1 ± 3.5	12.1 ± 5.0	3.0 ± 2.6	12.8 ± 6.1	3.5 ± 3.2
(*n* = 1)	Hypo	11.0 ± 4.5	1.3 ± 1.3	13.8 ± 6.8	3.6 ± 2.7	12.1 ± 5.2	2.1 ± 3.0	12.9 ± 6.2	2.8 ± 2.7
Stool frequency NTCP (%)	Conventional	2.9 ± 1.3	0.4 ± 0.5	4.5 ± 2.0	0.8 ± 0.8	5.6 ± 3.0	1.0 ± 1.0	4.4 ± 2.3	0.7 ± 0.8
(*n* = 0.11)	Hypo	3.9 ± 1.6	0.4 ± 0.8	5.2 ± 2.5	1.1 ± 1.3	6.9 ± 3.3	0.8 ± 1.1	5.4 ± 2.7	0.8 ± 1.3
Proctitis NTCP (%)	Conventional	1.9 ± 1.4	0.4 ± 0.6	1.9 ± 1.8	0.7 ± 1.3	4.2 ± 3.3	1.1 ± 1.1	2.3 ± 2.3	0.7 ± 1.1
(*n* = 0.14)	Hypo	2.8 ± 2.5	0.4 ± 1.1*	2.1 ± 1.8	0.9 ± 1.2	5.3 ± 3.9	1.2 ± 1.5	2.9 ± 2.7	0.8 ± 1.3

Abbreviations: SD = *Standard Deviation.*
All differences without mark are statistically significantly with $p < .001$. Differences marked with * are statistically significant with $.001 < p < .05$.

10

10.3.2 autoVMAT vs. autoIMRT

PTV coverage for autoVMAT and autoIMRT were similar (0.0 ± 0.5%, p = .743). Figure 10.4a shows for autoVMAT and autoIMRT, population-mean DVHs for PTV_1, PTV_2, rectum, anus and bladder. Clearly the differences between autoVMAT and autoIMRT in OAR doses are much smaller than differences between autoVMAT and CLINICAL (compare Figure 10.1 and Figure 10.4a). Rectum and anus mean doses and rectum V_{75GyEq} were slightly (but not statistically significant) higher for autoIMRT; 0.3 ± 0.9 Gy (p = .073), 0.1 ± 1.0 Gy (p = .712), 0.2 ± 0.8% (p = .282), respectively. Bladder mean dose and rectum V_{65GyEq} were also slightly enhanced for autoIMRT and this was statistically significant; 0.9 ± 2.0 Gy (p = .014), and 0.4 ± 0.7% (p = .007), respectively.

10.3.3 0.5 cm leaves vs. 1 cm leaves

Differences in PTV coverage for autoVMAT with Agility™ MLC (i.e., 0.5 cm leaves) and autoVMAT with MLCi2 (i.e., 1 cm leaves) were not significant (0.1 ± 0.4%, p = .19). Figure 10.4b shows for the Agility™ and MLCi2 MLC, population-mean DVHs for PTV_1, PTV_2, rectum, anus and bladder. The differences between Agility™ and MLCi2 in OAR doses were minor. Rectum and anus mean doses and rectum V_{65GyEq}, V_{75GyEq} were not statistically significant; 0.2 ± 0.7 Gy (p = .125), 0.1 ± 1.4 Gy (p = .828), 0.1 ± 0.8% (p = .696), and 0.0 ± 0.6% (p = .771), respectively. While the difference in mean bladder mean dose was small, it was statistically significant in favor of the Agility™ MLC; 0.4 ± 0.8 Gy (p = .008).

10.3.4 Hypo-fractionation with autoVMAT vs. conventional-fractionation with CLINICAL plans

As shown in Figure 10.3 and Table 10.4, for the CLINICAL plans, calculated NTCPs for hypo-fractionation were higher than for conventional-fractionation. This is in line with previous late toxicity analyses of the HYPRO trial (Aluwini *et al.* 2016). This difference is also observed for the autoVMAT plans. Interestingly, delivered rectum EQD_{2Gy} doses for hypo-fractionated treatment with autoVMAT plans are lower than for conventional-fractionation with the CLINICAL plans: After converting delivered rectum doses to EQD_{2Gy}, the population-average mean rectum dose for hypo-fractionated autoVMAT plans was 28% lower than for conventional-fractionation with CLINICAL plans (23.7 ± 6.9 Gy vs. 31.4 ± 7.6 Gy, p < .001). For rectum V_{65GyEq}, the average absolute reduction was 3% (11.3 ± 5.7% vs. 14.3 ± 5.8%, p < .001), whereas V_{75GyEq} slightly increased by 1.8% (4.8 ± 4.4% vs. 3.0 ± 3.2%, p < .001). For stool incontinence this translated into a statistically significant lower NTCP for hypo-fractionation with autoVMAT compared to conventional-fractionation with CLINICAL plans (relative

10
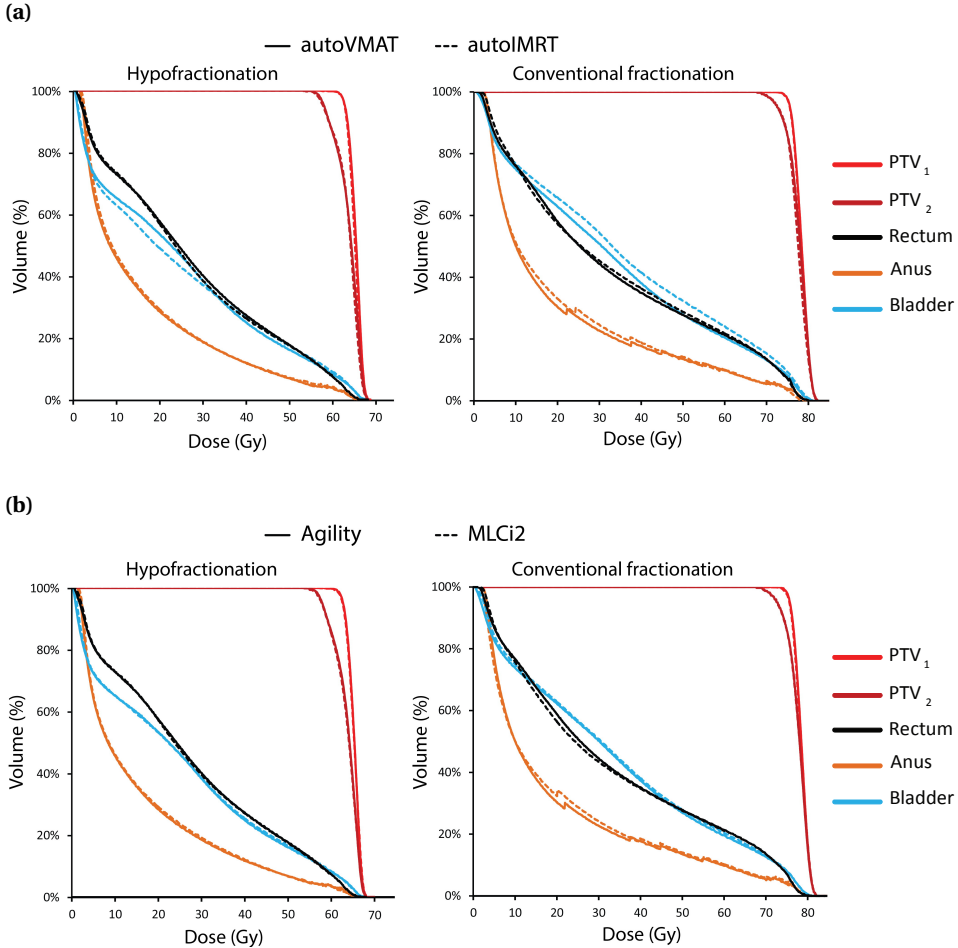
Figure 10.4: (a) Comparisons of population-averaged dose-volume histograms (DVHs) for the automatically generated VMAT and IMRT plans for 30 patients treated in the hypo-fractionation arm (left) and 30 patients treated in the conventional-fractionation arm (right). **(b)** Comparisons of population-averaged DVHs for automatically generated VMAT plans with the Agility™ MLC and with the MLCi2 for 30 patients treated in the hypo-fraction arm (left) and 30 patients treated in the conventional-fractionation arm (right).

10

reduction 12.6% for $n = 0.2$; $p = .001$). For the other symptoms there was no difference: late grade ≥ 2 GI 0.4% ($p = .901$), rectal bleeding -0.9% ($n = 0.13$, $p = .702$), stool frequency 5.1% ($p = .157$), and for proctitis -4.1% ($p = .580$).

After converting delivered bladder doses to EQD_{2Gy}, population-average mean bladder doses for hypo-fractionated autoVMAT plans were 13% lower than for conventional-fractionation with CLINICAL plans (25.2 ± 12.0 Gy vs. 28.7 ± 12.4 Gy, $p < .001$). Therefore, also for GU toxicity there is an indication that the combination of hypo-fractionation plus autoVMAT could be favorable instead of conventional-fractionation with the CLINICAL plans.

10.4   Discussion

For each of the 725 HYPRO patients included in this study, Erasmus-iCycle/Monaco was used to automatically generate a treatment plan for a modern treatment technique and modern equipment, i.e. VMAT at an Elekta Synergy® treatment machine (Elekta AB, Sweden), equipped with an Agility™ MLC with 160 leaves of 0.5 cm in width. With similar PTV dose delivery, these autoVMAT plans had clearly superior OAR dose delivery compared to the CLINICAL plans, resulting in clinically meaningful reductions in calculated rectum NTCPs.

Most of the CLINICAL treatments were performed with IMRT and with MLCs with a leaf width of 1 cm (Section 10.2). To exclude that the enhanced plan quality of autoVMAT compared to CLINICAL was related to the choice of VMAT instead of IMRT, or 0.5 cm vs. 1 cm, we performed two sub-group analyses. It was clearly demonstrated that these technique differences could not explain the large improvements in plan quality with autoVMAT (compare Figure 10.4a and Figure 10.4b with Figure 10.1).

Reported GI toxicity in the HYPRO trial was used to derive rectum NTCP models applied in this study, including clinical covariates and an EUD volume parameter, n, depending on the specific endpoint (Table 10.1). For rectal bleeding, $n = 0.13$ was found. However, for this endpoint two alternative models were also generated with imposed n-values of 0.03 and 0.18, as found in the literature (Rancati *et al.* 2011; Gulliford *et al.* 2012). For the whole range of investigated n-values, reduced NTCPs were found for the autoVMAT plans. The advantage for autoVMAT increased with increasing n. In a very recent inter-institutional analysis, Thor *et al.* (2018) emphasized the importance of sparing the low- to intermediate- rectal dose regions to prevent late rectal bleeding following prostate radiotherapy, resulting in an n-value of 0.3. With $n = 0.3$, the predicted NTCP reductions with automated planning compared to manual planning would be even bigger and more clinically meaningful.

10

Care should be taken in the interpretation of the observed superiority of the automatically generated plans. Configuration of the system for automated planning was based on the planning constraints provided in the trial protocol and on application of the ALARA principle, i.e. try to reduce doses in healthy tissues, and particularly OARs, as much as possible. Also the clinical planning was based on the trial protocol and Dutch planners are trained in working according to the ALARA principle. Apparently, the common basis for automated and manual plan generation did not prevent large quality differences. Superiority of the automatically generated plans may have been caused by inherent advantages of automated planning for consistently finding optimal solutions. The on average higher quality of autoVMAT plans could also have been caused by variations in the skills of the planners who generated the CLINICAL plans, or variations in the quality of CLINICAL plans due to variations in complexity of the cases. Insufficient time in the clinical setting to obtain the best plans, or limited perseverance of planning teams in always finding the best solution for each individual patient could also have contributed. Patient inclusion for the HYPRO trial started in 2007. Therefore, other potential explanations for the superiority of the recently generated autoVMAT plans are the use of TPSs with less advanced options for generating the CLINICAL plans, or limited experience in IMRT planning at the time of patient inclusion. However, to the best of our knowledge, there are no publications demonstrating improved plan quality for prostate IMRT in recent years due to more sophisticated planning tools or enhanced planning experience. There is also no evidence that this occurred in the centers participating in the HYPRO trial. Recent publications on cervical cancer and lung cancer have reported superiority of automated planning with Erasmus-iCycle/Monaco compared to manual planning, using the same, recent Monaco® version for both (Sharfo *et al.* 2016; Della Gala *et al.* 2017). Heijmen *et al.* (2018) observed in an international validation study with four European centers, superiority of automated planning for prostate cancer compared to manual planning, both performed with the same Monaco® version. In another recent study (Voet *et al.* 2014), we have compared for 30 prostate cancer patients autoVMAT with manual planning by the most competent planner in our institute without any time constraint. His task was to just do everything to generate the best possible plan in line with our clinical practice. Interestingly, in that study, the manual planner did on average obtain the same quality as autoVMAT. This is attributed to the non-clinical setting for this study; only the most competent planner did the manual planning and he could take all the time he needed to make the best possible plan. In the current study, comparisons are made with realistic manual planning in a clinical setting, 7 instead of 1 center were involved, and conclusions were based on 725 patients (by far the largest study on automated planning so far).

The HYPRO trial showed enhanced late toxicity for the investigated hypo-fractionation schedule with relatively high fractions doses (3.4 Gy), compared

10

to conventional-fractionation (Aluwini *et al.* 2016). Our simulations show that automatically generated plans for the hypo-fractionation schedule could result in toxicity levels, below or similar to those observed (and accepted) for conventional-fractionation. Given the logistic and economic advantages of the more than 50% reduction in number of treatment fractions with hypo-fractionation, for both the patient and the treating institution, this observation on toxicity may be considered in discussions on clinical usefulness of the investigated hypo-fractionation scheme. In these discussions, the modest predictive accuracy of the applied NTCP models (see Table 10.1), as generally observed in the literature on prostate cancer, has to be taken into account.

In this study, NTCP-models were used for plan evaluation. Possibly, direct use of these models in plan optimization could result in further plan quality increases. Further study on this is warranted.

In conclusion, in this study we have used automated planning to analyze late toxicity in the randomized HYPRO trial, that compared hypo-fractionation with conventional-fractionation for prostate cancer patients. Plans of 725 patients were analyzed. For both fractionation schedules, autoplanning resulted in reduced OAR doses and toxicity, compared to the trial results obtained with manual planning. Evidence was obtained that autoplanning could reduce toxicity resulting from hypo-fractionation to a level equivalent to that observed in the HYPRO trial for conventional-fractionation, possibly allowing a reduction of 50% in treatment fractions while keeping toxicity below the currently accepted level.

10.5   Acknowledgments

10

The TRENDY Multi-Center Randomized Trial on Hepatocellular Carcinoma - Trial QA Including Automated Treatment Planning and Benchmark-Case Results

Steven J.M. Habraken[1], Abdul Wahab M. Sharfo[1], Jeroen Buijsen[2], Wilko F.A.R. Verbakel[3], Cornelis J.A. Haasbeek[3], Michel C. Öllers[2], Henrike Westerveld[4], Niek van Wieringen[4], Onne Reerink[5], Enrica Seravalli[5], Pètra M. Braam[6], Markus Wendling[6], Thomas Lacornerie[7], Xavier Mirabel[7], Reinhilde Weytjens[8], Lieselotte Depuydt[9], Stephanie Tanadini-Lang[10], Oliver Riesterer[10], Karin Haustermans[11], Tom Depuydt[11], Roy S. Dwarkasing[12], François E.J.A. Willemssen[12], Ben J.M. Heijmen[1] and Alejandra Méndez Romero[1]

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

[2] MAASTRO Clinic, Department of Radiation Oncology, Maastricht, The Netherlands

[3] VU University Medical Center, Radiation Oncology, Amsterdam, The Netherlands

[4] Academic Medical Center, Radiotherapy, Amsterdam, The Netherlands

[5] University Medical Center Utrecht, Department of Radiotherapy, Utrecht, The Netherlands

[6] Radboud University Medical Center, Radiation Oncology, Nijmegen, The Netherlands

[7] Oscar Lambret Comprehensive Cancer Center, Academic Radiation Therapy Department, Lille, France

[8] University Hospital Antwerp, Iridium Kankernetwerk, Radiotherapy, Antwerp, Belgium

[9] GZA Ziekenhuizen/Iridium Kankernetwerk, Radiotherapy, Wilrijk, Belgium

[10] University Hospital Zurich, Department of Radiation Oncology, Zurich, Switzerland

[11] University Hospitals Leuven, Department of Radiation Oncology, Belgium

[12]Erasmus MC, Department of Radiology and Nuclear Medicine, Rotterdam, The Netherlands

11

Abstract

Background and Purpose: The TRENDY trial is an international multi-center phase-II study, randomizing hepatocellular carcinoma (HCC) patients between transarterial chemoembolization (TACE) and stereotactic body radiation therapy (SBRT) with a target dose of 48-54 Gy in six fractions. The radiotherapy quality assurance (QA) program, including prospective plan feedback based on automated treatment planning, is described and results are reported.

Materials and Methods: Scans of a single patient were used as a benchmark case. Contours submitted by nine participating centers were compared with reference contours. The subsequent planning round was based on a single set of contours. A total of 20 plans from participating centers, including 12 from the benchmark case, 5 from a clinical pilot and 3 from the first study patients, were compared to automatically generated VMAT plans.

Results: For the submitted liver contours, Dice Similarity Coefficients (DSC) with the reference delineation ranged from 0.925 to 0.954. For the GTV, the DSC varied between 0.721 and 0.876. For the 12 plans on the benchmark case, healthy liver normal-tissue complication probabilities (NTCPs) ranged from 0.2% to 22.2% with little correlation between NCTP and PTV-$D_{95\%}$ ($R^2 < 0.3$). Four protocol deviations were detected in the set of 20 treatment plans. Comparison with coplanar autoVMAT QA plans revealed these were due to too high target dose and sub-optimal planning. Overall, autoVMAT resulted in an average liver NTCP reduction of 2.2 percent point (range: 16.2 percent point to -1.8 percent point, $p = .03$), and lower doses to the healthy liver ($p < .01$) and gastrointestinal organs-at-risk ($p < .001$).

Conclusions: Delineation variation resulted in feedback to participating centers. Automated treatment planning can play an important role in clinical trials for prospective plan QA as sub-optimal plans were detected.

11.1 Background

Quality assurance (QA) is essential to clinical trials involving radiation therapy, as protocol violations may seriously impact trial outcome (Williams *et al.* 2007; Fairchild *et al.* 2012; Weber *et al.* 2012; Fairchild *et al.* 2013; Ibbott *et al.* 2013; Ohri *et al.* 2013; Moore *et al.* 2015). This holds for all trial aspects, including delineation (Steenbakkers *et al.* 2006; Brouwer *et al.* 2012; Joye *et al.* 2014; Berry *et al.* 2016b) and planning (Nelms *et al.* 2012; Berry *et al.* 2016a; Li *et al.* 2017).

The TRENDY trial (registered as NCT02470533 on https://clinicaltrials.gov/) is an international multi-center clinical trial in which patients with hepatocellular carcinoma (HCC) are randomized between transarterial chemoembolization (TACE) with drug eluting beads in the standard arm and stereotactic body radiation therapy (SBRT) in the experimental arm. Radiotherapy is delivered in six fractions with a total target dose of 48–54 Gy. The primary endpoint is time to progression. Secondary endpoints include time to local recurrence, response rate (partial or complete), overall survival, toxicity and quality of life. Patients are accrued from eleven centers in four different countries. Radiotherapy is delivered in ten of these centers, all experienced with liver SBRT; patients from one center are referred to Erasmus MC for radiotherapy.

Due to underlying liver disease (cirrhosis), the patient population in the TRENDY trial is prone to developing hepatic toxicity after radiotherapy, and sparing of the healthy liver is crucial in these patients (Dawson *et al.* 2002). As radiotherapy of the liver is technologically challenging, an extensive QA program, comparable to that of other recent radiotherapy trials on primary liver cancer (Dawson *et al.* 2012; Hong & Dawson 2015), has been developed. QA guidelines and recommendations are outlined in a QA protocol, included in (Habraken *et al.* 2017) and summarized in Table 11.1. The trial QA is coordinated by a QA team, consisting of two radiation oncologists and two medical physicists from different participating centers. The QA team is supported by two radiologists for advice on target definition. Prior to inclusion of patients, centers are supposed to recently have had an external beam output audit. Also, they need to fill out a trial-specific facility questionnaire explaining in detail technical aspects of the applied procedures for treating HCC patients. In addition, centers have to prepare delineation and treatment planning on a benchmark case (Melidis *et al.* 2014). During patient accrual, the QA protocol accommodates patient-specific prospective feedback (prior to treatment) on target definition, organ-at-risk (OAR) delineation and treatment planning. It defines minor and major protocol deviations. When a major deviation is detected during prospective monitoring by the QA team, the participating center needs to first improve the contouring and/or planning before start of treatment. Retrospectively detected major deviations have to be avoided for future patients from the same center.

11

Table 11.1: Minor and major protocol violations.

	PROTOCOL	Minor deviation	Major deviation
GTV definition	GTV is defined on a planning breathhold CT (expiration phase) or on a single breathing phase of a 4D-CT, both acquired in a hepatic arterial dominant phase		Non-compliance
Tumor delineation	Generous		
Liver, kidneys, stomach, heart and Gallbladder	Full delineation	No doubt on minor or major deviation in planning criteria (below)	Doubt on minor or major deviation in planning criteria (below)
Spinal cord, esophagus, duodenum, and bowel	Partial delineation including tissue in high dose region		Non-compliance
Left adrenal gland	Only when the right adrenal gland lies in the high-dose region		Non-compliance
CTV-to-GTV margin	0 mm		Non-compliance
Reduction of PTV margin to spare OAR	Not allowed		Non-compliance
Fractionation	6 fractions		Non-compliance
PTV*	95% ≥ 48 Gy	93-95% ≥ 48 Gy	< 93% ≥ 48 Gy
PTV	$D_{max} \leq 1.33 \times D_{95\%,PTV}$	$1.33 \times D_{95\%,PTV} < D_{max} \leq 1.4 \times D_{95\%,PTV}$	$D_{max} > 1.4 \times D_{95\%,PTV}$
Liver-GTV	$D_{mean} \leq 22$ Gy		$D_{mean} > 22$ Gy
Liver-GTV	> 800 ml less than 23.4 Gy		< 800 ml
Liver-GTV	NTCP ≤ 5%	5% < NTCP < 10%	NTCP ≥ 10%
Stomach/ Small and large bowel	$D_{max} < 39$ Gy, $V_{30Gy} \leq 5$cc	$V_{>39Gy} < 1$cc	$V_{>39Gy} \geq 1$cc
Esophagus	$D_{max} \leq 36$ Gy	$V_{>36Gy} < 1$cc	$V_{>36Gy} \geq 1$cc
Spinal cord	$D_{max} < 24$ Gy	$D_{max} < 25.5$ Gy	$D_{max} \geq 25.5$ Gy
Kidney	$D_{2/3} < 19.2$ Gy	19.2 Gy < $D_{2/3}$ < 21 Gy	$D_{2/3} \geq 21$ Gy
Avoidable hot spots outside PTV	Not allowed		Non-compliance

* Whenever possible, it is highly recommended to aim at 95% of PTV ≥ 54Gy.

Minor deviations should be avoided when possible. For each submitted patient, a volumetric-modulated arc therapy (VMAT) plan is generated at Erasmus MC on behalf of the QA team, using fully automated multi-criterial plan generation (autoVMAT) (Voet *et al.* 2014; Sharfo *et al.* 2016). AutoVMAT planning results are provided to the treating center. The aim is to provide prospective feedback for as many patients as possible, but at the very least for the first three patients from each center. Depending on feedback on the first patients, subsequent patients can be reviewed either prospectively or retrospectively (after start of treatment). To make prospective feedback practically feasible, participating centers are invited to submit at least three days prior to the planned start of treatment all images, contours and the treatment plan.

This paper contains quantitative analyses of the contouring and treatment planning benchmark-case results and reports on the first experiences with treatment plan feedback based on autoVMAT.

11.2   Materials and Methods

11.2.1   Target definition and contouring guidelines

The target lesion(s) should be visible on contrast-enhanced CT imaging and are primarily delineated on the arterial contrast phase of the CT-scan. Additional diagnostic imaging (MR/PET) is recommended when available. In regions with poor visibility of tumor edges, generous delineation is required to avoid tumor miss. General recommendations for OAR delineation are outlined in the study protocol. Additional published guidelines were supplied to support delineation (Jabbour *et al.* 2014). OARs that have to be fully delineated in 3D are: liver, kidneys, stomach, heart and gallbladder. For the required calculation of the liver normal-tissue complication probability (NTCP, see below), a healthy liver structure has to be created, defined as the full liver minus the gross tumor volume (GTV), and minus parts of the liver that are not functional such as regions previously treated with radio-frequency ablation (RFA) (if present). Partial delineations are allowed for spinal cord, duodenum, esophagus, and bowel (in case areas of the small and/or large bowel are located close to the tumor or could be located in high dose gradients).

11.2.2   Planning constraints and objectives

In treatment planning, the goal is to irradiate the full PTV with at least 48 Gy, or, if possible within planning constraints, up to 54 Gy, in six fractions. General recommendations for treatment planning and planning constraints are outlined in the trial protocol and summarized in the QA protocol, where also minor and major deviations are defined. The healthy liver is an important OAR. Several planning constraints are used to guarantee safe liver dose delivery, in particular, NTCP ≤ 5%. An NTCP between 5% and 10% qualifies as a minor deviation. For NTCP calculation, the healthy liver dose-volume histogram (DVH) is first converted into a DVH for 1.5 Gy/fraction dose delivery in every voxel using $\alpha/\beta =$ 2.5 Gy. Then the NTCP is calculated using $TD_{50/5} = 39.8$ Gy, $m = 0.12$, $n = 0.97$ (Dawson *et al.* 2002, 2006). A software implementation for NTCP calculation has been developed and distributed among the participating radiotherapy centers. In addition, the average dose to the healthy parts of the liver should be less than 22 Gy (Dawson *et al.* 2006) and, based on (Son *et al.* 2010) and $\alpha/\beta = 3$ Gy, at least 800 cc should receive less than 23.4 Gy.

11.2.3   Benchmark case

For the benchmark case, data from a 64-year-old male patient, treated at Erasmus MC, was used. Initially, he had an HCC in segment V-VI, for which he underwent surgery. Three years later, there was a 4 cm lesion visible on CT, compatible with HCC, close

to the right branch of the portal vein. There were no enlarged lymph nodes, nor lung metastases. The AFP blood level was 3. MRI showed a 4 cm lesion with wash-out surrounded by a capsule, compatible with HCC. The tumor was located in segment V, transition to VIII and toward segment IV.

In the delineation round, participating centers were asked to define the target (GTV), and delineate the liver for this patient. Arterial and venous phase contrast CT-scans (with 2.5 mm slice separation), diagnostic MR images (with 2 mm and 7.7 mm slice separation), and a description of the relevant medical history of the patient were provided. The centers were asked to delineate the liver and GTV according to the trial protocol. The contours were evaluated based on consensus among the medical doctors in the QA team, supported by a radiologist for the GTV. Quantitative comparisons of 3D volumes generated from submitted and reference contours, as defined by the QA team, were based on; the volume difference, the average distance in 3D, the center-of-mass (COM) shift, and the Dice Similarity Coefficient (DSC). The DSC of two volumes is defined as the overlap, divided by the average volume.

In the subsequent planning round, the centers generated for the same patient a treatment plan according to the study protocol, based on contours used for the initial treatment of the patient at Erasmus MC. The results of the treatment planning for the benchmark case were primarily evaluated based on the treatment planning protocol, see the dosimetric parameters listed in Table 11.2. Also, correlations between the achieved target dose and the NTCP and OAR doses were investigated.

11.2.4   Automated treatment planning

Based on the trial constraints and objectives, fully automated plan generation for HCC was implemented in the in-house developed Erasmus-iCycle software for fully automatic, prioritized multi-criteria optimization. Erasmus-iCycle was used to generate an individualized planning template for fully automatic generation of a clinically deliverable coplanar VMAT plan with the Monaco® TPS (Elekta AB, Stockholm, Sweden) (Voet *et al.* 2014; Sharfo *et al.* 2016). AutoVMAT was primarily used to assess the quality of treatment planning, both for the benchmark case and for clinical trial patients. To this end, fully automatically generated coplanar VMAT plans were compared to 20 treatment plans for 9 patients; 12 for the benchmark case (1 patient, see below), 5 from a clinical pilot study (5 patients), and 3 from the first 3 patients included in the trial so far. AutoVMAT plans were scaled to the clinical target dose (PTV-$D_{95\%}$, dose to 95% of the PTV). The target dose for the benchmark case and one of the trial patients was 54 Gy in six fractions. For the other 7 patients, it was around 48 Gy, also in six fractions.

11

11.2.5 Dose escalation with autoVMAT

Since the trial protocol recommends to cover as much of the PTV as possible with 54 Gy in six fractions, in addition, autoVMAT was also used to assess for how many patients dose escalation to 54 Gy was possible without violating any planning constraint.

11.3 Results

11.3.1 Benchmark case – GTV and liver delineation

All ten participating radiotherapy centers completed the contouring. Two institutes primarily contoured on the venous instead of the arterial contrast phase of the CT. All contours were considered acceptable, although variation in both GTV and liver contours was statistically significant (see below). Two axial CT slices with submitted contours are shown in Figure 11.1, where also the reference contours are displayed. Several centers delineated parts of the gallbladder, vena cava and diaphragm as liver.
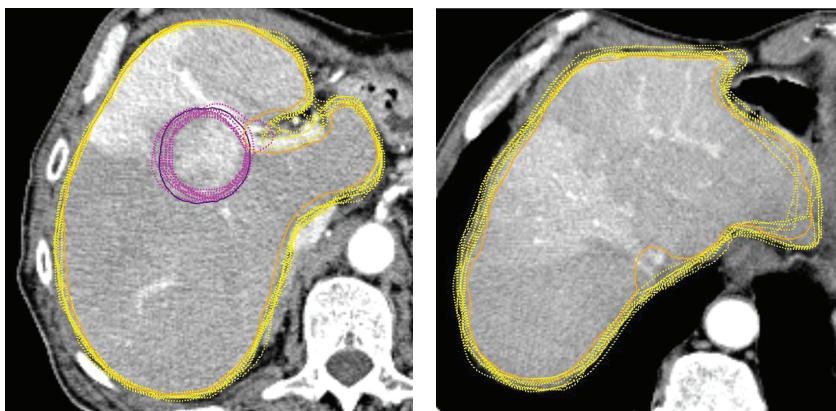

Figure 11.1: Axial slices of the benchmark-case CT scan with liver contours (submitted: yellow, dashed; reference: orange, solid) and GTV contours (submitted: pink, dashed; reference: purple, solid).

Table 11.2: Benchmark-case results and protocol requirements on treatment planning. Roman numbers (I, II, . . .) refer to the institutes and revised plans from the same institutes are indicated with (*).

Item	Protocol	I	II	III	III*	IV	V	VI	VII	VII*	VIII	IX	X	Auto
Modality		VMAT	VMAT	IMRT	VMAT	CK	VMAT	CK	VMAT	VMAT	VMAT	VMAT	VMAT	VMAT
Energy (MV)		10	6	6	6	6	10	6	10	10	10	10	6	10
Beams/Arcs		2	2	9	4		2		1	1	2	1	2	2
Gantry Angles		180 - 40	180 - 0	20, 55, 145, 170, 195, 220, 265, 318, 330	180 - 180		180 - 0		180 - 30	280 - 40	180 - 90	180 - 50	180 - 180	190 - 175
MUs		1775	1640	1552	1889		1413		1459	1320	1778	2489	3236	1938
D_{max} PTV (Gy)	≤ 72	70.5	65.6	67.5	69.5	70.1	65.0	68.5	71.6	68.5	69.6	70.3	71.4	70.2
$V_{PTV} \geq 48$ Gy (%)	> 95	100	100	100	100	100	100	100	100	99.3	100	99.6	100	100
$V_{PTV} \geq 54$ Gy (%)		99.6	99.6	99.6	98.7	99.4	98.4	93.1	100	80.5	90.1	93.5	98.1	94.9
D_{max} Liver–GTV (Gy)	≤ 22	15.8	15.7	17.0	14.9	15.9	15.4	17.1	18.4	14.2	17.2	14.7	14.6	15.0
$V_{Liver-GTV} < 23.4$ Gy (cc)	> 800	925	924	897	965	962	934	926	856	948	904	955	962	924
$NTCP_{Liver-GTV}$ (%)	≤ 5	2.4	3.2	7.0	0.8	2.1	1.2	3.8	22.2	0.2	4.4	0.5	0.4	0.9
D_{max} Stomach/Bowel (Gy)	< 39	16.2	16.3	21.4	13.9	18.6	16.4	17.4	18.0	14.3	18.3	15.2	18.6	11.9
V_{30Gy} Stomach/Bowel (cc)	< 5	0	0	0	0	0	0	0	0	0	0	0	0	0
D_{max} Esophagus (Gy)	≤ 36	1.5	1.9	1.3	1.5	1.5	1.7	7.3	4.2	3.3	3.3	1.4	1.8	1.6
D_{max} Spinal Cord (Gy)	< 24	6.7	8.7	14.7	11.8	5.1	7.0	6.2	6.0	5.5	8.6	7.4	17.8	8.0
$D_{2/3}$ Right Kidney (Gy)	< 19.2	0.9	0.9	3.8	1	5.4	0.7	3.6	1.4	1.1	2.0	0.7	1.1	1.0
$D_{2/3}$ Left Kidney (Gy)	< 19.2	2.9	3.7	0.5	2.7	2.9	3.0	1.6	3.2	3.8	3.8	2.8	5.7	3.6
Avoidable hotspots	None	None	None	None	None	None	None	None	None	None	None	None	None	None

11

The results of the quantitative comparisons with the reference contours, are shown in Figure 11.2. Submitted liver volumes are on average 101.4 cc = 8.0% larger than the reference (range 2.2% to 12.0%, $p < .001$), while submitted GTVs are on average 12.6 cc = 24% smaller than the reference (range -2.8% to 43.2%, $p < .001$). The average distance in 3D between submitted and reference volumes ranges from 1.7 to 3.5 mm for the liver and from 1.2 to 2.4 mm for the GTV. Center-of-mass shifts range from 0.8 to 2.4 mm and from 1.7 to 4.6 mm, respectively. DSCs of submitted and reference volumes range from 0.925 to 0.954 for the liver and from 0.721 to 0.876 for the GTV.

11.3.2 Benchmark case – Treatment planning

Treatment planning results are summarized in Table 11.2. Two centers (III and VII) did not meet the NTCP constraint initially and re-planned the benchmark case after feedback had been provided. The revised plans are labeled III* and VII* respectively. For reference, also an autoVMAT plan with a target dose of 54 Gy to 95% of the PTV was generated, see Table 11.2.

PTV dose homogeneity and conformity vary, with some institutes aiming at a high target dose allowing for large dose gradients in the GTV-PTV margin, and others optimizing for a smoother, more homogeneous, dose distribution. Within the set of 12 plans for the benchmark case, all based on the same CT and structure set, there is little correlation between PTV-$D_{95\%}$ and the achieved NTCP value ($R^2 = 0.3$, $R^2 < 0.1$ for the 10 protocol compliant plans), and hardly any between PTV-$D_{95\%}$ and OAR doses ($R^2 < 0.2$ for esophagus, stomach, duodenum and bowel), see Figure 11.3. This suggests that the observed differences in dose distribution, especially those within the sub-set of protocol compliant plans, are largely due to variations in treatment planning and treatment technique.

11.3.3 Plan QA with autoVMAT

Within the set of 20 plans, two major and two minor deviations were detected, all with respect to the NTCP. With autoVMAT, one major NTCP deviation of 22.2% was reduced to a minor deviation of 7.5%, while the other was reduced from 28.7% to 12.5%, which is still not protocol compliant. For both these plans, the deviation was caused in part by the dose to the PTV being higher than feasible within the NTCP constraint. The autoVMAT plans could be scaled to fully compliant plans with NTCP values of 0.9% and 3.8%, respectively, by reducing the PTV-$D_{95\%}$ from 58.8 Gy to 54.0 Gy and from 50.6 Gy to 48.0 Gy. The major NTCP deviations were detected in one plan from the benchmark case and one from the clinical pilot. No patients were treated with these plans. In addition, with autoVMAT, two minor NTCP deviations of respectively 8.3% and 7.0%, one in another
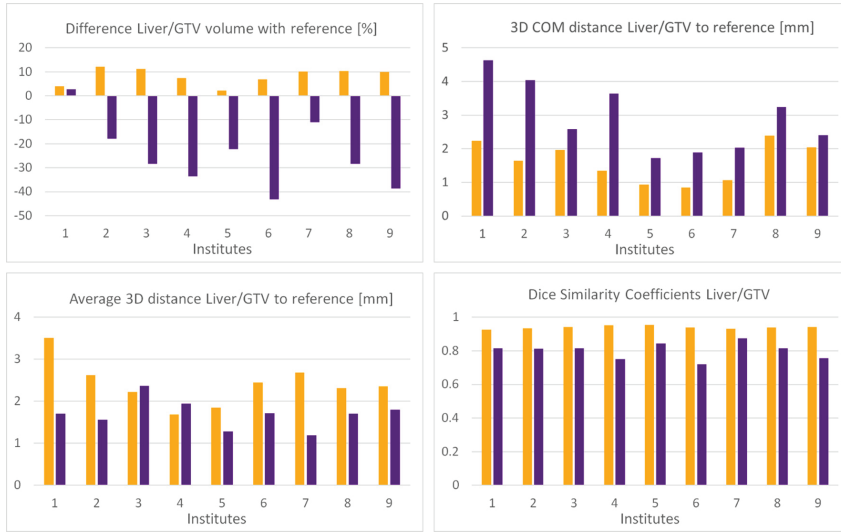
11

Figure 11.2: Differences between submitted contours and reference contours for all centers, excluding Erasmus MC. Liver: orange, GTV: purple (COM = center of mass). Distances are based on 3D volumes and the DSC of two volumes is defined as the overlap, divided by the average volume.
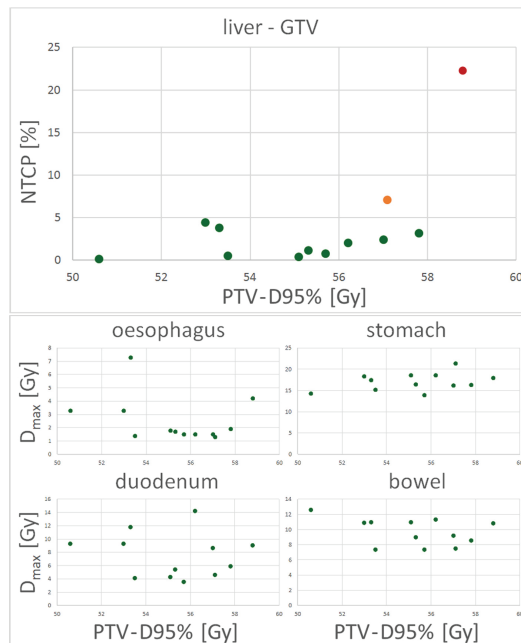

Figure 11.3: (a) Liver-GTV NTCP (%) versus $D_{95\%}$ for the PTV for 12 plans of the benchmark case. Minor protocol deviations are represented in orange, major deviations in red, **(b)** Gastrointestinal OAR doses versus $D_{95\%}$ for the PTV for 12 plans of the benchmark case.
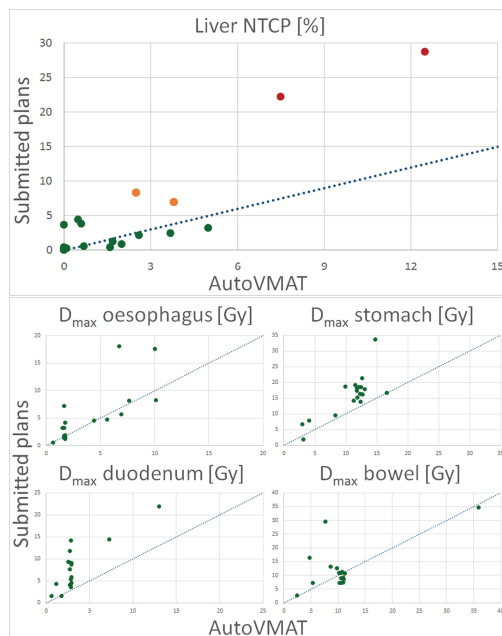
Figure 11.4: (a) Liver NTCP for submitted treatment plans versus autoVMAT. Minor protocol deviations are represented in orange, major deviations in red., **(b)** Gastrointestinal OAR doses for submitted treatment plans versus autoVMAT.

plan from the benchmark case and one in another plan from the clinical pilot, could be reduced to protocol compliant values of 2.5% and 3.8%, respectively, with the same PTV-$D_{95\%}$.

Although some plans using substantially more modulation/monitor units (MUs) and some non-coplanar (CyberKnife®) plans outperform the corresponding coplanar autoVMAT plan, in most cases, autoVMAT resulted in a superior plan, with an average NTCP reduction of 2.2 percent point (range: 16.2 percent point to -1.8 percent point, $p =$ .03), a lower mean dose to the healthy liver ($p < .01$) and lower doses to gastrointestinal OARs ($p < .001$), see Figure 11.4. For 7 of the 20 plans, including 4 with a protocol deviation, autoVMAT achieved a reduction of the liver NTCP by more than 2 percent point without compromising any other OAR. A reduction of the gastrointestinal (GI) maximum dose(s) of more than 5.0 Gy was achieved for 4 additional plans, without sacrificing the NTCP or increasing the dose delivered to other OARs.

11.3.4   Potential for dose escalation analyzed with autoVMAT

For 7 out of the 9 patients considered here, the clinical target dose was around 48 Gy in six fractions. Using autoVMAT, we investigated if dose escalation to 54 Gy in six fractions

was possible within planning constraints for the OARs. For one patient, a target dose larger than 48 Gy was not possible due to an overlap between the PTV and part of the bowel. For another patient, it was not possible to achieve more than 48 Gy without violating the NTCP constraint. For the other 5 patients that were clinically treated to 48 Gy, dose escalation to 54 Gy could be achieved with autoVMAT. In practice, however, clinical considerations other than what can be achieved in treatment planning play a role in deciding on target dose.

11.4   Discussion

As part of the TRENDY multi-center randomized trial, an extensive QA program has been implemented, including a benchmark case and prospective (prior to treatment delivery) and retrospective feedback on delineation and treatment planning. Prospective and retrospective feedback during patient accrual promotes the quality and uniformity of SBRT treatment in the trial, but has the possible disadvantage of the trial being less representative for normal clinical practice, especially during the start-up phase when revision of delineation and treatment planning is more likely to be required. Since both the trial protocols, and all QA results are available, other centers should, however, be able to reproduce all (future) trial results.

Submitted benchmark-case delineations were evaluated by comparison with a reference, based on consensus within an expert panel. The submitted delineations of liver and GTV are small in volume compared to the total volume represented by the CT-scan, in which case the DSC becomes comparable to kappa values for inter-observer variation in delineation (see Section 11.6), as reported for HCC in (Hong *et al.* 2014; Jabbour *et al.* 2014; Gkika *et al.* 2017). Based on the widely accepted and used interpretation of kappa values in inter-observer studies (Landis & Koch 1977; Allozi *et al.* 2010), we conclude that there is almost perfect agreement (DSC between 0.8 and 1.0) between all submitted liver contour sets and six of nine GTV contour sets, and that there is substantial agreement (DSC between 0.6 and 0.8) between the other three GTV contour sets and the reference. The DSCs for the liver contours are comparable in magnitude to the kappa values reported in inter-observer studies (Hong *et al.* 2014; Jabbour *et al.* 2014; Gkika *et al.* 2017). Both individual feedback and general recommendations, based on the observed delineation variations, have been provided to participating centers.

Feedback on treatment planning was based both on the trial protocol and results obtained with fully automated treatment planning. The added value of autoVMAT for plan QA is twofold. It contributes to understanding the nature of protocol deviations in treatment planning, which are often due to sub-optimal planning combined with (too) much priority being given to target dose, or to some OAR(s) at the cost of others. In such

11

cases, dosimetric autoVMAT results provide a solution approach, thereby reducing both the number of deviations and patient loss in the trial. In addition, autoVMAT results can be used to identify, quantify and monitor variations in protocol compliant plans, even for an individual patient. Although re-planning is not mandatory in such cases, this contributes to the objective and systematic evaluation of submitted treatment plans. AutoVMAT results show what can be achieved in treatment planning, and individual feedback based (in part) on these results, may promote the quality and uniformity of treatment planning for future patients from the same center.

When more patients from multiple centers are included in the trial, automated treatment planning will be applied to identify systematic variations in treatment planning between participating centers and to monitor improvements in treatment planning. Although it is expected that variation in submitted dose distributions is mostly due to treatment planning, at this point it is, strictly speaking, not possible to separately study the impact of treatment planning and treatment technique. In the longer run, dedicated fully automated treatment planning adjusted to treatment delivery techniques in participating centers could be implemented as a part of clinical trials.

As has been demonstrated previously, knowledge-based (Moore *et al.* 2015) and automated treatment planning (Sharfo *et al.* 2018) may, in retrospect, reveal sub-optimal treatment planning in clinical trials. To the best of our knowledge, the present manuscript constitutes the first report on the use of automated treatment planning for prospective and retrospective plan QA within a clinical trial (while patients are accrued). We have demonstrated that, because of the quality and consistency of automatically generated plans, they can assist in evaluating constraint violations in submitted plans (maybe these can be avoided by changing the plan). Automated planning can also identify sub-optimal treatment plans that do fulfill all constraints, but with PTV or OAR dose delivery that could be made better with plan adjustments. Due to the automation, fast feedback on plans submitted by participating centers is feasible.

11.5   Acknowledgements

11

11.6 Derivation regarding Cohen's kappa parameter and the Dice Similarity Coefficient

11.6.1 Introduction and objective

Cohen's kappa parameter (kappa) is the appropriate statistical measure to quantify the variation in inter-observer delineation studies. Our study is based on a comparison with reference contours, based on consensus within an expert panel, in which case the Dice similarity coefficient (DSC) is more appropriate. The DSC of two volumes A and B is defined as the overlap divided by the average volume, i.e., $2(A \cap B)/(A + B)$.

In cases in which delineated volumes are small compared to the total volume of the scan, kappa becomes similar to the DSC. For kappa, a widely used and accepted interpretation exists (Landis & Koch 1977; Allozi *et al.* 2010), which, in this case, can also be applied to the DSC. Moreover, DSCs can be compared to values of kappa reported in literature (Jabbour *et al.* 2014; Hong *et al.* 2014; Gkika *et al.* 2017).

11.6.2 Derivation

Consider two observers (I and II) delineating the same 3-dimensional structure as A and B on the same image set of the same patient. Let V be the total scanned volume. The relevant overlap volumes can be calculated as listed in the table below.

	Delineated by I	Not Delineated by I	**Total**
Delineated by II	$A \cap B$	$B - A \cap B$	**B**
Not Delineated by II	$A - A \cap B$	$V - (A + B) + A \cap B$	**V − B**
Total	**A**	**V - A**	

Here $\cap$ denotes the overlap between A and B. Cohen's kappa is defined as

$$\kappa = \frac{\rho_o - \rho_e}{1 - \rho_e} \tag{11.1}$$

where ρ_o is the observed probability of agreement and ρ_e is the (hypothetical) probability of chance agreement, as estimated based on the available data.

The observed probability of agreement is defined as the total volume delineated by both observers $A \cap B$ and the total volume not delineated by any of the observers $V - (A + B) + A \cap B$, divided by the total volume V

$$\rho_o = \frac{A \cap B + V - (A + B) + A \cap B}{V} = \frac{2(A \cap B) + V - (A + B)}{V} \tag{11.2}$$

The probabilities of observers I and II delineating a volume element are given by A/V and B/V respectively, whereas the probabilities of the observers not delineating a

11

volume element are given by given by (V-A)/V and (V-B)/V respectively. The probability of chance agreement is equal to the probability of both observers delineating or not delineating a volume, hence

$$\rho_e = \frac{A}{V} * \frac{B}{V} + \frac{V-A}{V} * \frac{V-B}{V} = \frac{2(A*B) - V(A+B) + V^2}{V^2} \qquad (11.3)$$

Combining all this, we find

$$\kappa = \frac{\rho_o - \rho_e}{1 - \rho_e} = \frac{2V(A \cap B) - 2(A*B)}{V(A+B) - 2(A*B)} \qquad (11.4)$$

which, in the limit of V much larger than A and B, reduces to the Dice similarity coefficient $2(A \cap B)/(A+B)$, as quoted above.

Discussion

12.1   Introduction

This thesis focuses on evaluation and application of automated treatment planning in radiotherapy. Throughout this work, we demonstrated that automated treatment planning based on Erasmus-iCycle has been able to produce a consistently higher plan quality than commonly used manual trial-and-error planning. Furthermore, automated planning was useful in decision-making regarding preference of one treatment approach/technique over another, while largely avoiding bias as observed in published studies based on manual planning. We also showed that Erasmus-iCycle may be used as a tool to assure high plan quality in multi-institutional trials, even if a large variation in treatment approaches, applied treatment planning systems (TPSs), etc. is allowed. In this chapter, the performed studies are discussed in a broader context.

12.2   Automated treatment plan generation with Erasmus-iCycle – an overview

Erasmus-iCycle has no CE-marking and is not applied in a direct way to generate clinically deliverable plans. Instead, fully automated multi-criterial plan generation with Erasmus-iCycle is always followed by re-construction of the Erasmus-iCycle plan in a commercial TPS (generally Monaco®). Initially, such plan re-construction was performed manually by generating a new plan in Monaco®, guided by the dose distribution

obtained with Erasmus-iCycle. In 2012, Voet *et al.* (2013b) performed a prospective study for head-and-neck cancer patients to evaluate this semi-automated treatment planning approach. In their study, there was an overwhelming preference for the semi-automatically generated plans compared to manually generated plans by the dosimetrists. As described in chapter 6 of this thesis, this semi-automated planning was also used for comparison of various IMRT and VMAT approaches for cervical cancer. In 2014, Voet *et al.* (2014) presented a method for automated re-construction of Erasmus-iCycle treatment plans in Monaco®, based on patient-specific Monaco® planning templates that were automatically generated from Erasmus-iCycle dose distributions. Because of the enhanced plan quality and the significant reduction in workload compared to manual planning, Erasmus-iCylce/Monaco was then introduced for fully automated clinical planning for head-and-neck cancer (also for re-planning after anatomical changes during treatment) and prostate cancer, followed by locally-advanced cervical cancer in October 2014 (Chapter 2), and advanced non-small cell lung cancer in March 2016 (Della Gala *et al.* 2017). Due to the huge reduction in manual planning workload, automated treatment planning opened new doors for the extension of plan-libraries for plan-of-the-day adaptive radiotherapy in locally-advanced cervical cancer (Nováková *et al.* 2017). In 2017, automatic treatment plan generation with Erasmus-iCycle/Monaco was used for linac-based treatment of 291 head-and-neck cancer patients, 360 prostate cancer patients, 54 cervical cancer patients (plan-of-the-day adaptive therapy with 2 or 3 plans/patient), and 219 NSCLC patients. Unrelenting work is on-going to clinically introduce automated treatment planning for other treatment sites such as Hodgkin's lymphoma, breast cancer and rectal cancer.

In collaboration with the University Hospitals of Mannheim and Vienna, the performance of Erasmus-iCycle/Monaco for fully automated multi-criteria VMAT planning was validated for spinal metastases (Chapter 3), advanced gastric cancer (Chapter 4), and sequential-boost whole pelvic prostate cancer (Chapter 5). In general, the plan quality of the automatically generated plans was superior to the plans generated by local expert dosimetrists in absence of time pressure. In a study by Heijmen *et al.* (2018), the Erasmus-iCycle/Monaco planning approach was overall superior to routine manual planning for prostate cancer patients treated in four leading European university hospitals.

Recently, Erasmus-iCycle was configured and coupled to the clinical MultiPlan® TPS for fully automated generation of clinically deliverable non-coplanar CyberKnife® (CK) plans. This system was successfully tested for prostate SBRT (Chapter 9), and vestibular schwannoma (Rossi *et al.* 2017). A dedicated (non-clinical) research version of Erasmus-iCycle was developed for intensity-modulated proton therapy (IMPT), mainly used to investigate mitigation of the impact of treatment uncertainties (Arts *et al.* 2017; Bijman *et al.* 2017). Further development on this front is on-going. Automated planning

12

for brachytherapy and for very high electron energy are also active fields of research at Erasmus MC (Breedveld *et al.* 2019; Oud *et al.* 2020).

12.3 Automated planning for treatment technique comparison

12.3.1 Bias reduction and enhancement of patient numbers

In clinical practice, appropriate choices between available delivery techniques, such as VMAT or IMRT (Chapter 6), coplanar or non-coplanar (Chapters 7-9) (Voet *et al.* 2012, 2013a; Rossi *et al.* 2012), prone or supine (Heijkoop *et al.* 2016), or photons versus protons (Arts *et al.* 2017; Bijman *et al.* 2017; Florijn *et al.* 2020) are essential to maximize the treatment quality for individual patients and to make optimal use of available resources. Generally, the performance of clinical studies to answer such questions for patient populations is considered unfeasible. Alternatively, techniques can be compared in treatment planning studies using planning CT data of a group of patients. However, treatment planning studies based on manual trial-and-error planning may suffer from considerable and often unknown bias. Sources of bias can be *i)* differences in available planning skills for the investigated treatment techniques, *ii)* inconsistencies in planning among patients, *iii)* use of different TPSs with more or less advanced planning options, *iv)* upfront preference for one of the techniques. Moreover, because of the workload in manual planning, plan numbers in these studies are generally low. When using Erasmus-iCycle for automated multi-criterial planning in treatment technique comparisons, bias can largely be reduced because *i)* the plan quality is not affected by the skills of the dosimetrists, *ii)* Erasmus-iCycle can be used for planning for all investigated treatment techniques using the same decision-making strategy ('*wish-list*'), *iii)* automated planning excludes upfront preferences, *iv)* due to the very limited manual workload, large numbers of patients may be included.

12.3.2 Treatment technique decision-making for individual patients

Treatment technique/modality comparisons based on Erasmus-iCycle automated planning can be incorporated into the decision-making process, allowing that the most optimal, personalized treatment will be delivered to a specific patient. In this context, we have recently performed a computer simulation study using automated treatment planning to establish percentages of prostate cancer patients that might be treated with a hypo-fractionation regimen, as a function of the accepted increase in normal tissue complication probabilities (NTCPs) relative to conventional-fractionation (Bijman *et al.* 2018). 50% of the studied patient cohort could be treated with hypo-fractionated EBRT

12

for an accepted increase in the gastrointestinal complications of less than 2.5% compared to conventional-fractionation. Similarly, in view of the relatively high treatment costs and limited capacity of proton therapy, automated treatment planning may be used to assist in making choices, so that the treatment modality with the highest cost-effectiveness will be applied. In the Netherlands, a patient-specific comparison between photon and proton treatment is mandatory for specific tumor groups. Automated treatment planning will enable objective and efficient treatment modality selection. Erasmus-iCycle was used to automatically generate IMRT and IMPT plans for oropharyngeal, cervical and endometrial cancer patients to determine which patients would benefit most from IMPT compared to IMRT (van de Sande *et al.* 2016; Arts *et al.* 2017). In a recent study (Florijn *et al.* 2020), Erasmus-iCycle was used to fully automatically generated IMPT, VMAT and IMRT plan for skull-base meningioma patients with target diameter > 3 cm in order to investigate the potential advantages for proton therapy. The results supported serious consideration of IMPT for treatment, because of significantly reduced dose delivery to the hippocampi and other critical organs. Ultimately, improved technique decisions will improve patient care and allow effective and efficient use of available clinical resources.

12.3.3 Non-coplanar beam set-ups in SBRT

Coplanar VMAT with a single couch position is often considered an interesting treatment option because of the involved short delivery times. On the other hand, because of limited availability of beam angle optimization options in commercial treatment planning systems, high quality comparisons between coplanar VMAT with non-coplanar treatment are scarce. Erasmus-iCycle features integrated multi-criterial beam profile and beam angle optimization. Based on obtained optimal beam angles and dose distributions, high quality, clinically deliverable non-coplanar plans can then be automatically re-constructed in Monaco® (see above). Using this planning approach, we have demonstrated that non-coplanar beam arrangements can significantly improve dose distributions when the tumor is wrapped around multiple OARs (Chapters 7-9). The clinical value of non-coplanar treatment became especially clear for liver metastases patients, allowing for dose escalation compared to coplanar VMAT plans (Chapter 7), and for prostate SBRT by enhancing OAR sparing with the robotic CyberKnife® compared to VMAT treatment (Chapter 9). Notwithstanding the substantial gains in average plan quality for non-coplanar treatment, large differences between patients were observed, with some patients hardly gaining anything. Non-coplanar treatments result in prolonged delivery times with an associated increased cost, an enhanced risk of intra-fraction tumor motion and a potential loss in patient comfort. Indeed, in a study on lung cancer SBRT, Rossi *et al.* (2016) observed an increased intra-fraction tumor motion for non-coplanar IMRT with treatment times of on average 32 min

12

compared to 12 min for VMAT. However, this increased intra-fraction motion could be compensated by an increase in CTV-PTV margin of only 0.3 mm. A possible clinical scenario could be to automatically generate for each individual patient a coplanar VMAT plan and a non-coplanar plan. Based on differences in plan quality and treatment time, individualized decisions could then be taken.

Another complimentary approach could be to reduce the number of couch positions in non-coplanar treatment. In Chapter 7, we have proposed and tested a novel treatment modality for liver metastases, designated VMAT+, that substantially reduced treatment delivery times (~65%) compared to multiple non-coplanar IMRT beams treatment, while marginally sacrificing dose quality and maintaining possible dose escalation to the target. For prostate SBRT, we developed a non-coplanar two-beam class solution to complement the coplanar VMAT treatment, resulting in substantial plan quality gain with a low increase in treatment delivery time (Chapter 8).

12.4   Mechanisms for obtaining high plan quality with Erasmus-iCycle

Plan quality gains observed with Erasmus-iCycle automated planning are often larger than reported for other systems such as the auto-planning module in Pinnacle (Philips) and the knowledge-based automated planning module, as implemented in RapidPlan™ (Varian). In part, this can be explained from differences in approaching the multi-criterial aspect of plan generation. Erasmus-iCycle is the only system that generates plans being guaranteed *Pareto-optimal* for individual patients, resulting from the use of the $2p\epsilon c$ optimizer (Breedveld *et al.* 2009, 2012). In addition, differences also arise from different approaches used for configuration of the systems. E.g., in RapidPlan™, achievable DVHs are predicted using models describing relationships between achieved dose and anatomy as observed in previous clinical plans, normally generated with manual planning. Therefore, for new patients this system basically reproduces the quality of the plans of previously treated patients. There is no inherent drive to exceed the quality of the best manual plans used for training. In contrast to this so-called knowledge-based approach, configuration of Erasmus-iCycle for a treatment site, i.e., generation of the related wish-list, has an explicit drive to improve on the training plans. As visualized in Figure 12.1, generation of the wish-list is an iterative process using typically 5 training patients. In each iteration, plans are automatically generated based on a wish-list that is updated based on the plan results obtained in the previous iteration. Initially, the aim is always to re-produce the quality of the training plans, i.e., the clinical plan quality. When that is achieved, the planning team defines specific targets for the next iterations
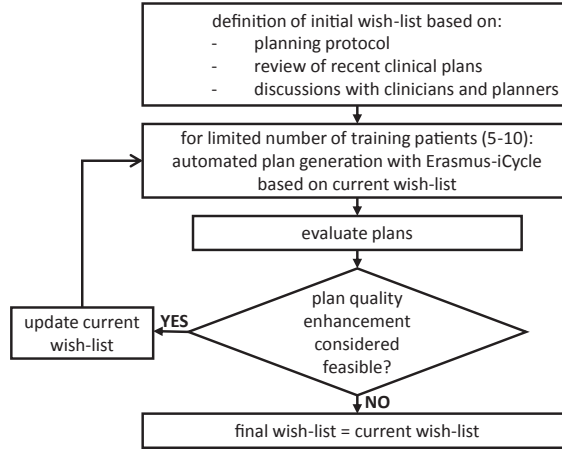
12

Figure 12.1: Scheme summarizing the tuning of a wish-list to be used by Erasmus-iCycle/Monaco. In the tuning process the wish-list is iteratively improved, aiming at plan quality improvements for the training patients. The wish-list updates aim at a final quality of automatically generated plans that is superior to the quality of the 5-10 training plans. The involved wish-list updates can consist of addition or removal of hard constraints or objectives, changes in constraint values or objective goal values, and changes in objective priorities. Automated generation of 5-10 plans is the basis for each wish-list update, allowing multiple updates without excessive workload.

to further improve. Due to the use of automated planning in the wish-list generation, many iterations can be performed without excessive planning workload.

12.5 Automated treatment planning in clinical trials

In (randomized) clinical trials, the quality of treatment plans of enrolled patients may be suboptimal (Abrams *et al.* 2012; Weber *et al.* 2012; Ohri *et al.* 2013; Moore *et al.* 2015; Sharfo *et al.* 2018). On one hand, this is not good for patients participating in the trials since they may be vulnerable to excess risk of radiation-induced toxicity or a reduced probability for curing the disease. On the other hand, this may also result in debatable answers in studies comparing radiotherapy with other modalities, or studies investigating enhanced toxicity when adding radiotherapy to the treatment. Eventually, this may seriously affect the trial outcome. To avoid this, there is a strong need for using automated treatment planning in (randomized) clinical trials when radiotherapy is involved. However, full and fast introduction of high-quality automated planning may be hampered by practical problems. In the meantime, automated treatment planning can be used to verify individualized plan quality in clinical trials. In Chapter 11, automated treatment planning with Erasmus-iCycle/Monaco was proposed for the 'TRENDY'

international multi-center randomized trial for hepatocellular carcinoma as part of the quality assurance (QA) program, featuring prospective plan QA. Prior to the start of treatment of a new trial patient in a participating center, a VMAT plan is automatically generated at Erasmus MC. This VMAT plan is then sent to the treatment center. Based on comparison with their own plan, the treating team can decide to adapt their treatment plan if improvements seem possible and relevant. In the TRENDY trial, patients are treated with a broad range of treatment techniques and equipment (depending on the center), while the automatically generated plan is always for VMAT on an Elekta unit. Therefore, direct use of the latter plan is not possible. Nevertheless, it allows comparison of the local plan for the local treatment approach with a high quality VMAT plan. In Chapter 10, we provided evidence that automated planning could substantially have reduced the risk of late rectal toxicities in the randomized multi-center HYPRO trial. In a follow-up analysis of the data presented in Chapter 10, large variations in clinical plan quality amongst participating institutions were observed (not published), while the automated plans benchmark showed a clear consistency in plan quality amongst patients in the two arms of the trial. This shows that automated treatment planning and/or automated plan QA might be required to better guarantee high quality treatment, make the treatment amongst participating centers in multi-center trials more uniform, and identify and correct sub-optimal treatment plans or protocol violations prospectively.

12.6 Current and future developments of automated treatment planning

For all patients of a treatment group, automated treatment plan generation with Erasmus-iCycle is based on an *a priori* defined generic 'wish-list' that represents a fixed decision-making strategy. The challenge of wish-list creation is to find a general configuration that results in high-quality, clinically acceptable plans for all patients in the treatment cohort. Generation of a widely applicable wish-list is often not straightforward, requiring substantial tuning (Figure 12.1) to obtain appropriate parameters, including choice of hard constraints and objectives with their priorities, goal values for the objectives, etc.. The effort required for creating a wish-list also depends on the experience and skills of the involved team. Currently, simplification and automation of the wish-list generation process is an active research area in our department. Moreover, automated personalization of initial wish-list parameters during automated planning is being explored as well (Rossi *et al.* 2018). With this approach, the actual parameters defined in the wish-list become less critical, so tuning of the wish-list to make it applicable for all patients will become easier and less time-intensive. Incorporation of clinically validated NTCP models in plan optimization to directly steer the optimization of dose distributions

12

has proven to be feasible in the TRENDY trial (Chapter 11). Further investigation on the real clinical value of NTCP-based treatment planning optimization is a research topic at our department.

As explained above, in our current work-flow automated generation of a clinically deliverable VMAT/IMRT or CK treatment plan is performed in two steps: after generation of the Erasmus-iCycle plan based on a tumor site-specific wish-list, this plan is automatically re-constructed in our clinical TPS (Monaco® or MultiPlan®) using *patient-specific* planning templates that are automatically created from the Erasmus-iCycle plan. Generation of these templates is non-trivial because of *i)* dissimilarities of the available cost-functions and their functionality between Erasmus-iCycle and the clinical TPS, *ii)* differences in handling hard constraints and objective priorities, *iii)* differences in the dose calculation engine (pencil beam vs. Monte Carlo), and *iv)* plan quality loss between the fluence-optimized Erasmus-iCycle plan and the deliverable, segmented plan in the clinical TPS. To minimize differences in the dose calculation between Erasmus-iCycle and Monaco® (point *iii* above), especially in low-density tissues like lung, we have recently coupled the fast Monte Carlo dose calculation package (GPUMCD), as available in Monaco®, to Erasmus-iCycle. In a separate project, Erasmus MC is collaborating with Elekta AB on the implementation of Erasmus-iCycle functionality (i.e., automated multi-criteria optimization based on wish-lists) in their TPS. An aim of this project is to have a smooth transition from the multi-criterial fluence map optimization to the segmentation phase, reducing the re-construction issues mentioned above. This would allow worldwide use of high-quality automated multi-criterial planning with possibilities to exchange wish-lists between centers for gradual improvement of treatment planning. In the end, this may also contribute to more standardization in radiotherapy dose delivery.

Due to limitations of the two-step approach in our current automated planning work-flow (above), it may not always be possible to fully re-produce Erasmus-iCycle plans in the commercial TPS, possibly yielding a sub-optimal plan quality. In our department, we have developed and implemented a QA tool to enable detection of those outliers (Wang *et al.* 2017). Based on overlap volume histograms between OARs and the PTV, deviations between realized and expected DVH parameters for OARs may be detected and a warning for the need of manual fine-tuning of the plan may be given in case of big variations.

Based on the applied wish-list and the selected beam geometry, Erasmus-iCycle generates for each patient a single *Pareto-optimal* plan with trade-offs among treatment objectives that are considered clinically favorable. For a specific patient, a physician might be interested in evaluating different clinical trade-offs, like the balance between target coverage and organ at risk sparing. Instead of generating another *Pareto-optimal* plan (i.e., using another wish-list with different trade-offs), we have started a project

12

that investigates Pareto-navigation in a small region in the solution space around the Erasmus-iCycle plan, the so-called 'wish-point'. The aim is to evaluate the navigation tool in clinical practice. Observed deviations between final plans and wish-point plans could result in changes in wish-lists.

Erasmus-iCycle uses the $2p\epsilon$-constraint lexicographic optimization algorithm for multi-criterial generation of a *Pareto-optimal* plan (Breedveld *et al.* 2009, 2012). Although this algorithm ensures high quality plans, it requires many consecutive minimizations of objective functions and therefore the calculation times may be in the order of 2-4 hours for plan generations based on an extensive wish-list with many OARs involved (e.g., for head-and-neck cancer patients). In recent years, our group has developed and implemented a lexicographic reference point method (LRPM) which can replace the $2p\epsilon c$ optimization by a single optimization problem, thereby significantly reducing the calculation times while maintaining the current high plan quality (van Haveren *et al.* 2017). For prostate cancer patients, calculation times of ~1 minute were achieved and for head-and-neck cancer patients of ~5 minutes (compared to 12.4 and 210 min with our current method). Due to these huge reductions in calculation times, the LRPM approach allows for a more effective and time-efficient clinical work-flow, and possibly in the near future fast and daily on-line treatment planning based on the patient geometry of that day. The LRPM approach has also been investigated for automatic re-optimization of IMPT treatment plans, taking 3 minutes only, thereby making a big leap towards real-time adaptive proton therapy (Jagt *et al.* 2018, 2019). Similar to the current wish-list approach in Erasmus-iCycle, the LRPM requires a set of trade-off configuration parameters which should be determined per treatment site in an iterative trial-and-error procedure. Unrelenting on-going research has recently shown the feasibility to automate the optimization/tuning of these parameters and thus drastically reducing the time and workload for configuration (van Haveren *et al.* 2019, 2020).

Recently, we have successfully expanded our automated multi-criterial planning approach to dwell time optimization for high dose rate (HDR) brachytherapy plans (Breedveld *et al.* 2018, 2019; Oud *et al.* 2020). Erasmus-iCycle was coupled to the clinical Oncentra® TPS for generation of clinically deliverable plans. With this system we achieved improved plan quality and shorter optimization times (~1 sec) compared to clinically delivered plans that were manually optimized. Recent investigations focused on automated planning for locally-advanced cervical cancer patients (Oud *et al.* 2020).

12

12.7 Role of automation of treatment planning in high-, low- and middle-income countries

Cancer has been identified as the second leading cause of death among non-communicable diseases (Hunter & Reddy 2013). According to the World Health Organization (WHO), the global cancer incidence between 2015 and 2035 is predicted to rise by ~65% from 15 million to more than 24 million (Bray *et al.* 2018). Two-thirds of the cases are expected in low- and middle-income countries (LMICs) (Bray *et al.* 2018). Because of its effectiveness and relatively low cost, radiotherapy is currently one of the most widely used treatment options for cancer around the world. However, as the incidence of cancer increases, the costs of diagnosis and treatment may become a heavy burden, especially in LMICs where the required need for more radiation therapy infrastructure and human resources (i.e., radiation oncologists, medical physicists, and radiation therapy technologists) could become financially unbearable. Moreover, required high-qualified medical, scientific and technical expertise is scarce in many countries, which could restrict the number of patients who can benefit from high-quality care. Automated treatment plan generation could possibly play an important role in improving this situation. It could drastically reduce the number of well-trained dosimetrists needed for large-scale generation of high-quality plans. Use of wish-lists of more experienced centers could to a large extent circumvent the need for specific expertise to generate them. Nevertheless, there are various challenges with this concept. Just considering the planning as a black box may result in dangerous situations. For safe usage, the local team has to be able to interpret generated plans and recognize problematic cases. It should also be made sure that wish-lists generated elsewhere are indeed applicable to the local situation, e.g., the effectiveness of a wish-list may depend on the applied CTV-PTV margin.

12

References

ABRAMS, R.A., WINTER, K.A., REGINE, W.F., SAFRAN, H., HOFFMAN, J.P., LUSTIG, R. *et al.* (2012). Failure to adhere to protocol specified radiation therapy guidelines was associated with decreased survival in RTOG 9704 — A phase III trial of adjuvant chemotherapy and chemoradiotherapy for patients with resected adenocarcinoma of the pancreas. *Int. J. Radiat. Oncol. Biol. Phys.*, **82**(2), 809–816. 190

AHAMAD, A., D'SOUZA, W., SALEHPOUR, M., IYER, R., TUCKER, S.L., JHINGRAN, A. *et al.* (2005). Intensity-modulated radiation therapy after hysterectomy: Comparison with conventional treatment and sensitivity of the normal-tissue–sparing effect to margin size. *Int. J. Radiat. Oncol. Biol. Phys.*, **62**(4), 1117–1124. 78

AHMAD, R., HOOGEMAN, M.S., BONDAR, M., DHAWTAL, V., QUINT, S., DE PREE, I. *et al.* (2011). Increasing treatment accuracy for cervical cancer patients using correlations between bladder-filling change and cervix–uterus displacements: Proof of principle. *Radiother. Oncol.*, **98**(3), 340–346. 10

ALBER, M., & REEMTSEN, R. (2007). Intensity modulated radiotherapy treatment planning by use of a barrier-penalty multiplier method. *Optim. Methods Softw.*, **22**(3), 391–411. 48, 71, 80, 144

ALDUHAIBY, E.Z., BREEN, S., BISSONNETTE, J.P., SHARPE, M., MAYHEW, L., TYLDESLEY, S. *et al.* (2012). A national survey of the availability of intensity-modulated radiation therapy and stereotactic radiosurgery in Canada. *Radiat. Oncol.*, **7**(1), 18. 37

ALLOZI, R., LI, X.A., WHITE, J., APTE, A., TAI, A., MICHALSKI, J.M. *et al.* (2010). Tools for consensus analysis of experts' contours for radiotherapy structure definitions. *Radiother. Oncol.*, **97**(3), 572–578. 179, 181

ALONGI, F., COZZI, L., ARCANGELI, S., IFTODE, C., COMITO, T., VILLA, E., LOBEFALO, F., NAVARRIA, P., REGGIORI, G., MANCOSU, P., CLERICI, E., FOGLIATA, A., TOMATIS, S., TAVERNA, G., GRAZIOTTI, P., & SCORSETTI, M. (2013). Linac based SBRT for prostate cancer in 5 fractions with VMAT and flattening filter free beams: preliminary report of a phase II study. *Radiat. Oncol.*, **8**(1), 171. 114

ALUWINI, S., VAN ROOIJ, P., HOOGEMAN, M., BANGMA, C., KIRKELS, W.J., INCROCCI, L., & KOLKMAN-DEURLOO, I.K. (2010). CyberKnife Stereotactic Radiotherapy as Monotherapy for Low- to Intermediate-Stage Prostate Cancer: Early Experience, Feasibility, and Tolerance. *J. Endourol.*, **24**(5), 865–869. 115

ALUWINI, S., VAN ROOIJ, P., HOOGEMAN, M.S., KIRKELS, W., KOLKMAN-DEURLOO, I.K., & BANGMA, C. (2013). 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.*, **8**(1), 84. 131

ALUWINI, S., POS, F., SCHIMMEL, E., VAN LIN, E., KROL, S., VAN DER TOORN, P.P. *et al.* (2015). Hypofractionated versus conventionally fractionated radiotherapy for patients with prostate cancer (HYPRO): Acute toxicity results from a randomised non-inferiority phase 3 trial. *Lancet Oncol.*, **16**(3), 274–283. 4, 150, 151

ALUWINI, S., POS, F., SCHIMMEL, E., KROL, S., VAN DER TOORN, P.P., DE JAGER, H. *et al.* (2016). Hypofractionated versus conventionally fractionated radiotherapy for patients with prostate cancer (HYPRO): Late toxicity results from a randomised, non-inferiority, phase 3 trial. *Lancet Oncol.*, **17**(4), 464–474. 4, 150, 151, 153, 161, 165

AMINI, A., JONES, B.L., YEH, N., RUSTHOVEN, C.G., ARMSTRONG, H., & KAVANAGH, B.D. (2015). Survival outcomes of whole-pelvic versus prostate-only radiation therapy for high-risk prostate cancer patients with use of the national cancer data base. *Int. J. Radiat. Oncol. Biol. Phys.*, **93**(5), 1052–1063. 60

APPENZOLLER, L.M., MICHALSKI, J.M., THORSTAD, W.L., MUTIC, S., & MOORE, K.L. (2012). Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Med. Phys.*, **39**(12), 7446–7461. 45

ARTS, T., BREEDVELD, S., DE JONG, M.A., ASTREINIDOU, E., TANS, L., KESKIN-CAMBAY, F. *et al.* (2017). The impact of treatment accuracy on proton therapy patient selection for oropharyngeal cancer patients. *Radiother. Oncol.*, **125**(3), 520–525. 186, 187, 188

ASMEROM, G., BOURNE, D., CHAPPELOW, J., GOGGIN, L. M., HEITZ, R., JORDAN, P. *et al.* (2016). The design and physical characterization of a multileaf collimator for robotic radiosurgery. *Biomed. Phys. Eng. Exp.*, **2**(1), 017003. 143

AUDET, C., POFFENBARGER, B.A., CHANG, P., JACKSON, P.S., LUNDAHL, R.E., RYU, S.I., & RAY, G.R. (2011). Evaluation of volumetric modulated arc therapy for cranial radiosurgery using multiple noncoplanar arcs. *Med. Phys.*, **38**(11), 5863–5872. 125

BEDFORD, J.L., THOMAS, M.D.R., & SMYTH, G. (2013). Beam modeling and VMAT performance with the Agility 160-leaf multileaf collimator. *J. Appl. Clin. Med. Phys.*, **14**(2), 172–185. 12, 88

BEKELMAN, J.E., MITRA, N., EFSTATHIOU, J., LIAO, K., SUNDERLAND, R., YEBOA, D.N. *et al.* (2011). Outcomes after intensity-modulated versus conformal radiotherapy in older men with nonmetastatic prostate cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, **81**(4), e325–e334. 150

BERRY, S.L., MA, R., BOCZKOWSKI, A., JACKSON, A., ZHANG, P., & HUNT, M. (2016a). Evaluating inter-campus plan consistency using a knowledge based planning model. *Radiother. Oncol.*, **120**(2), 349–355. 141, 170

BERRY, S.L., BOCZKOWSKI, A., MA, R., MECHALAKOS, J., & HUNT, M. (2016b). Interobserver variability in radiation therapy plan output: Results of a single-institution study. *Pract. Radiat. Oncol.*, **6**(6), 442–449. 45, 141, 170

BIJMAN, R.G., BREEDVELD, S., ARTS, T., ASTREINIDOU, E., DE JONG, M.A., GRANTON, P.V. *et al.* (2017). Impact of model and dose uncertainty on model-based selection of oropharyngeal cancer patients for proton therapy. *Acta Oncol.*, **56**(11), 1444–1450. 186, 187

BIJMAN, R.G., SHARFO, A.W.M., SCHILLEMANS, W., HEEMSBERGEN, W., WITTE, M., POS, F. *et al.* (2018). NTCP-model based patient selection for hypofractionated prostate treatment - A computer simulation. *Radiother. Oncol.*, **127**(1_suppl), S157–S158. 187

BODA-HEGGEMANN, J., WEISS, C., SCHNEIDER, V., HOFHEINZ, R.D., HANEDER, S., MICHAELY, H. *et al.* (2013). Adjuvant IMRT/XELOX radiochemotherapy improves long-term overall-and disease-free survival in advanced gastric cancer. *Strahlenther. Onkol.*, **189**(5), 417–423. 45

BOIKE, T.P., LOTAN, Y., CHO, L.C., BRINDLE, J., DEROSE, P., XIE, X.J. *et al.* (2011). Phase I dose-escalation study of stereotactic body radiation therapy for low- and intermediate-risk prostate cancer. *J. Clin. Oncol.*, **29**(15), 2020–2026. 131

BOKRANTZ, R. (2012). Multicriteria optimization for volumetric-modulated arc therapy by decomposition into a fluence-based relaxation and a segment weight-based restriction. *Med. Phys.*, **39**(11), 6712–6725. 70

BONDAR, M.L., HOOGEMAN, M.S., MENS, J.W., QUINT, S., AHMAD, R., DHAWTAL, G. *et al.* (2012). Individualized nonadaptive and online-adaptive intensity-modulated radiotherapy treatment strategies for cervical cancer patients based on pretreatment acquired variable bladder filling computed tomography scans. *Int. J. Radiat. Oncol. Biol. Phys.*, **83**(5), 1617–1623. 10, 11, 79

BORTFELD, T., & WEBB, S. (2008). Single-arc IMRT? *Phys. Med. Biol.*, **54**(1), N9. 88

BORTFELD, T. (2010). The number of beams in IMRT—theoretical investigations and implications for single-arc IMRT. *Phys. Med. Biol.*, **55**(1), 83–97. 78, 85

BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R.L., TORRE, L.A., & JEMAL, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.*, **68**(6), 394–424. 194

BREEDVELD, S., STORCHI, P.R.M., KEIJZER, M., HEEMINK, A.W., & HEIJMEN, B.J.M. (2007). A novel approach to multi-criteria inverse planning for IMRT. *Phys. Med. Biol.*, **52**(20), 6339–53. 13, 21, 28, 45, 78, 80, 150

BREEDVELD, S., STORCHI, P.R.M., & HEIJMEN, B.J.M. (2009). The equivalence of multi-criteria methods for radiotherapy plan optimization. *Phys. Med. Biol.*, **54**(23), 7199–7209. 13, 21, 28, 78, 80, 144, 189, 193

BREEDVELD, S., STORCHI, P.R.M., VOET, P.W.J., & HEIJMEN, B.J.M. (2012). iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Med. Phys.*, **39**(2), 951–963. 12, 13, 14, 21, 28, 37, 45, 48, 54, 60, 63, 64, 78, 79, 80, 94, 95, 96, 115, 116, 131, 132, 133, 141, 144, 150, 189, 193

BREEDVELD, S., BENNAN, A., ALUWINI, S., SCHAART, D., KOLKMAN-DEURLOO, I.K., & HEIJMEN, B.J.M. (2018). Fast automated multi-criteria planning for HDR brachytherapy explored for prostate cancer. *Radiother. Oncol.*, **127**(1_suppl), S1240. 193

BREEDVELD, S., BENNAN, A.B.A., ALUWINI, S., SCHAART, D.R., KOLKMAN-DEURLOO, I.K.K., & HEIJMEN, B.J.M. (2019). Fast automated multi-criteria planning for HDR brachytherapy explored for prostate cancer. *Physics in Medicine & Biology*, **64**(20), 205002. 187, 193

BRENNER, D.J., & HALL, E.J. (1999). Fractionation and protraction for radiotherapy of prostate carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 131

BROUWER, C.L., STEENBAKKERS, R.J.H.M., VAN DEN HEUVEL, E., DUPPEN, J.C., NAVRAN, A., BIJL, H.P. *et al.* (2012). 3D Variation in delineation of head and neck organs at risk. *Radiat. Oncol.*, **7**(1), 32. 170

BUERGY, D., LOHR, F., BAACK, T., SIEBENLIST, K., HANEDER, S., MICHAELY, H. *et al.* (2012). Radiotherapy for tumors of the stomach and gastroesophageal junction – A review of its role in multimodal therapy. *Radiat. Oncol.*, **7**(1), 192. 45, 54

BUERGY, D., SHARFO, A.W.M., HEIJMEN, B.J.M., VOET, P.W.J., BREEDVELD, S., WENZ, F. *et al.* (2017). Fully automated treatment planning of spinal metastases – A comparison to manual planning of volumetric modulated arc therapy for conventionally fractionated irradiation. *Radiat. Oncol.*, **12**(1), 33. 60

BURGHELEA, M., VERELLEN, D., POELS, K., HUNG, C., NAKAMURA, M., DHONT, J., GEVAERT, T., VAN DEN BEGIN, R., COLLEN, C., MATSUO, Y., KISHI, T., SIMON, V., HIRAOKA, M., & DE RIDDER, M. (2016). Initial characterization, dosimetric benchmark and performance validation of Dynamic Wave Arc. *Radiat. Oncol.*, **11**(1), 63. 114

BUSCHMANN, M., SEPPENWOOLDE, Y., WIEZOREK, T., WEIBERT, K., & GEORG, D. (2016). Advanced optimization methods for whole pelvic and local prostate external beam therapy. *Phys. Med.*, **32**(3), 465–473. 70, 116

BYLUND, K.C., BAYOUTH, J.E., SMITH, M.C., HASS, A.C., BHATIA, S.K., & BUATTI, J.M. (2008). Analysis of interfraction prostate motion using megavoltage cone beam computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.*, **72**(3), 949–956. 70

CANTIN, A., GINGRAS, L., LACHANCE, B., FOSTER, W., GOUDREAULT, J., & ARCHAMBAULT, L. (2015). Dosimetric evaluation of three adaptive strategies for prostate cancer treatment including pelvic lymph nodes irradiation. *Med. Phys.*, **42**(12), 7011–7021. 70

CHANG, D.T., SWAMINATH, A., KOZAK, M., WEINTRAUB, J., KOONG, A.C., KIM, J. *et al.* (2011). Stereotactic body radiotherapy for colorectal liver metastases. *Cancer*, **117**(17), 4060–4069. 94

CHEN, M.F., TSENG, C.J., TSENG, C.C., KUO, Y.C., YU, C.Y., & CHEN, W.C. (2007). Clinical outcome in posthysterectomy cervical cancer patients treated with concurrent Cisplatin and intensity-modulated pelvic radiotherapy: Comparison with conventional radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, **67**(5), 1438–1444. 10

CICCHETTI, A., RANCATI, T., EBERT, M., FIORINO, C., PALORINI, F., KENNEDY, A. *et al.* (2016). Modelling late stool frequency and rectal pain after radical radiotherapy in prostate cancer patients: Results from a large pooled population. *Phys. Med.*, **32**(12), 1690–1697. 153

CLARK, G.M., POPPLE, R.A., PRENDERGAST, B.M., SPENCER, S.A., THOMAS, E.M., STEWART, J.G., GUTHRIE, B.L., MARKERT, J.M., & FIVEASH, J.B. (2012). Plan quality and treatment planning technique for single isocenter cranial radiosurgery with volumetric modulated arc therapy. *Pract. Radiat. Oncol.*, **2**(4), 306–313. 125

CORBIN, K.S., HELLMAN, S., & WEICHSELBAUM, R.R. (2013). Extracranial oligometastases: A subset of metastases curable with stereotactic radiotherapy. *J. Clin. Oncol.*, **31**(11), 1384–1390. 36

COSGROVE, V.P., THOMAS, M.D.R., WESTON, S.J., THOMPSON, M.G., REYNAERT, N., EVANS, C.J. *et al.* (2009). Physical characterization of a new concept design of an Elekta radiation head with integrated 160-leaf multi-leaf collimator. *Int. J. Radiat. Oncol. Biol. Phys.*, **75**(3), S722–S723. 12

COX, J.D., STETZ, J.A., & PAJAK, T.F. (1995). Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European organization for research and treatment of cancer (EORTC). *Int. J. Radiat. Oncol. Biol. Phys.*, **31**(5), 1341–1346. 153

COZZI, L., DINSHAW, K.A., SHRIVASTAVA, S.K., MAHANTSHETTY, U., ENGINEER, R., DESHPANDE, D.D. *et al.* (2008). A treatment planning study comparing volumetric arc modulation with RapidArc and fixed field IMRT for cervix uteri radiotherapy. *Radiother. Oncol.*, **89**(2), 180–191. 78, 88

CRAFT, D., HALABI, T., SHIH, H.A., & BORTFELD, T. (2007). An approach for practical multiobjective IMRT treatment planning. *Int. J. Radiat. Oncol. Biol. Phys.*, **69**(5), 1600–1607. 60

CRAFT, D., & MONZ, M. (2010). Simultaneous navigation of multiple Pareto surfaces, with an application to multicriteria IMRT planning with multiple beam angle configurations. *Med. Phys.*, **37**(2), 736–741. 28, 37, 60, 70

CUNNINGHAM, D., ALLUM, W.H., STENNING, S.P., THOMPSON, J.N., VAN DE VELDE, C.J.H., NICOLSON, M. *et al.* (2006). Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N. Engl. J. Med.*, **355**(1), 11–20. 45

DAMAST, S., WRIGHT, J., BILSKY, M., HSU, M., ZHANG, Z., LOVELOCK, M. *et al.* (2011). Impact of dose on local failure rates after image-guided reirradiation of recurrent paraspinal metastases. *Int. J. Radiat. Oncol. Biol. Phys.*, **81**(3), 819–826. 37

DAWSON, L.A., NORMOLLE, D., BALTER, J.M., MCGINN, C.J., LAWRENCE, T.S., & TEN HAKEN, R.K. (2002). Analysis of radiation-induced liver disease using the Lyman NTCP model. *Int. J. Radiat. Oncol. Biol. Phys.*, **53**(4), 810–821. 49, 170, 172

DAWSON, L.A., ECCLES, C., & CRAIG, T. (2006). Individualized image guided iso-NTCP based liver cancer SBRT. *Acta Oncol.*, **45**(7), 856–864. 172

DAWSON, L.A., KAVANAGH, B.D., PAULINO, A.C., DAS, S.K., MIFTEN, M., LI, X.A. *et al.* (2010). Radiation-associated kidney injury. *Int. J. Radiat. Oncol. Biol. Phys.*, **76**(3), S108–S115. 49

DAWSON, L.A., ZHU, A., KNOX, J. *et al.* (2012). Randomized phase III study of sorafenib versus stereotactic body radiation therapy followed by sorafenib in hepatocellular carcinoma. *Radiation Therapy Oncology Group*, **1112**. 170

DE POOTER, J.A., MÉNDEZ ROMERO, A., JANSEN, W.P.A., STORCHI, P.R.M., WOUDSTRA, E., LEVENDAG, P.C. *et al.* (2006). Computer optimization of noncoplanar beam setups improves stereotactic treatment of liver tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, **66**(3), 913–922. 94

DE POOTER, J.A., MÉNDEZ ROMERO, A., WUNDERINK, W., STORCHI, P.R.M., & HEIJMEN, B.J.M. (2008). Automated non-coplanar beam direction optimization improves IMRT in SBRT of liver metastasis. *Radiother. Oncol.*, **88**(3), 376–381. 37, 94, 104

DEARNALEY, D.P., SYDES, M.R., GRAHAM, J.D., AIRD, E.G., BOTTOMLEY, D., COWAN, R.A. *et al.* (2007). Escalated-dose versus standard-dose conformal radiotherapy in prostate cancer: first results from the MRC RT01 randomised controlled trial. *Lancet Oncol.*, **8**(6), 475–487. 150

DELLA GALA, G., DIRKX, M.L.P., HOEKSTRA, N., FRANSEN, D., LANCONELLI, N., VAN DE POL, M. *et al.* (2017). Fully automated VMAT treatment planning for advanced-stage NSCLC patients. *Strahlenther. Onkol.*, **193**(5), 402–409. 45, 46, 48, 54, 60, 132, 141, 150, 154, 164, 186

DIKKEN, J.L., VAN SANDICK, J.W., MAURITS SWELLENGREBEL, H.A., LIND, P.A., PUTTER, H., JANSEN, E.P.M. *et al.* (2011). Neo-adjuvant chemotherapy followed by surgery and chemotherapy or by surgery and chemoradiotherapy for patients with resectable gastric cancer (CRITICS). *BMC Cancer*, **11**(1), 329. 45

DONG, P., LEE, P., RUAN, D., LONG, T., ROMEIJN, E., YANG, Y. *et al.* (2013). 4π Non-coplanar liver SBRT: A novel delivery technique. *Int. J. Radiat. Oncol. Biol. Phys.*, **85**(5), 1360–1366. 94, 104

DONG, P., NGUYEN, D., RUAN, D., KING, C., LONG, T., ROMEIJN, E. *et al.* (2014). Feasibility of prostate robotic radiation therapy on conventional C-arm linacs. *Pract. Radiat. Oncol.*, **4**(4), 254–260. 114, 131, 142

EISBRUCH, A., HARRIS, J., GARDEN, A.S., CHAO, C.K.S., STRAUBE, W., HARARI, P.M. *et al.* (2010). Multi-institutional trial of accelerated hypofractionated intensity-modulated radiation therapy for early-stage oropharyngeal cancer (RTOG 00-22). *Int. J. Radiat. Oncol. Biol. Phys.*, **76**(5), 1333–1338. 54

EZZELL, G.A., BURMEISTER, J.W., DOGAN, N., LOSASSO, T.J., MECHALAKOS, J.G., MIHAILIDIS, D. *et al.* (2009). IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. *Med. Phys.*, **36**(11), 5359–5373. 98

FAIRCHILD, A., AIRD, E., FENTON, P.A., GREGOIRE, V., GULYBAN, A., LACOMBE, D. *et al.* (2012). EORTC radiation oncology group quality assurance platform: Establishment of a digital central review facility. *Radiother. Oncol.*, **103**(3), 279–286. 170

FAIRCHILD, A., STRAUBE, W., LAURIE, F., & FOLLOWILL, D. (2013). Does quality of radiation therapy predict outcomes of multicenter cooperative group trials? A literature review. *Int. J. Radiat. Oncol. Biol. Phys.*, **87**(2), 246–260. 170

FERLAY, J., SOERJOMATARAM, I., DIKSHIT, R., ESER, S., MATHERS, C., REBELO, M. *et al.* (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer*, **136**(5), E359–E386. 54

FERNANDEZ-OTS, A., & CROOK, J. (2013). The role of intensity modulated radiotherapy in gynecological radiotherapy: Present and future. *Rep. Pract. Oncol. Radiother.*, **18**(6), 363–370. 10

FEUVRET, L., NOËL, G., MAZERON, J.J., & BEY, P. (2006). Conformity index: A review. *Int. J. Radiat. Oncol. Biol. Phys.*, **64**(2), 333–342. 15, 98

FLORIJN, M.A., SHARFO, A.W.M., WIGGENRAAD, R.G.J., VAN SANTVOORT, J.P.C., PETOUKHOVA, A.L., HOOGEMAN, M.S. *et al.* (2020). Lower doses to hippocampi and other brain structures for skull-base meningiomas with intensity modulated proton therapy compared to photon therapy. *Radiother. Oncol.*, **142**(2), 147–153. 187, 188

FOGLIATA, A., BELOSI, F., CLIVIO, A., NAVARRIA, P., NICOLINI, G., SCORSETTI, M. *et al.* (2014). On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer. *Radiother. Oncol.*, **113**(3), 385–391. 54, 55, 141

FOGLIATA, A., NICOLINI, G., CLIVIO, A., VANETTI, E., LAKSAR, S., TOZZI, A. *et al.* (2015a). A broad scope knowledge based model for optimization of VMAT in esophageal cancer: Validation and assessment of plan quality among different treatment centers. *Radiat. Oncol.*, **10**(1), 220. 37, 38, 54, 55

FOGLIATA, A., NICOLINI, G., BOURGIER, C., CLIVIO, A., DE ROSE, F., FENOGLIETTO, P. *et al.* (2015b). Performance of a knowledge-based model for optimization of volumetric modulated arc therapy plans for single and bilateral breast irradiation. *PLoS ONE*, **10**(12), e0145137. 37

FREEMAN, D.E., & KING, C.R. (2011). Stereotactic body radiotherapy for low-risk prostate cancer: Five-year outcomes. *Radiat. Oncol.*, **6**(1), 3. 131

FULLER, D.B., MARDIROSSIAN, G., WONG, D., & MCKELLAR, H. (2010). Prospective evaluation of stereotactic radiotherapy for low and intermediate risk prostate cancer: Emulating HDR brachytherapy dose distribution. *Int. J. Radiat. Oncol. Biol. Phys.*, **78**(3), S358–S359. 131

FÜRWEGER, C., PRINS, P., COSKAN, H., & HEIJMEN, B.J.M. (2016). Characteristics and performance of the first commercial multileaf collimator for a robotic radiosurgery system. *Med. Phys.*, **43**(5), 2063–2071. 143

GKIKA, E., TANADINI-LANG, S., KIRSTE, S., HOLZNER, P.A., NEEFF, H.P., RISCHKE, H.C. *et al.* (2017). Interobserver variability in target volume delineation of hepatocellular carcinoma. *Strahlenther. Onkol.*, **193**(10), 823–830. 179, 181

GLICKSMAN, R.M., TJONG, M.C., NEVES-JUNIOR, W.F.P., SPRATT, D.E., CHUA, K.L.M., MANSOURI, A., CHUA, M.L.K., BERLIN, A., WINTER, J.D., DAHELE, M., SLOTMAN, B.J., BILSKY, M., SHULTZ, D.B., MALDAUN, M., SZERLIP, N., LO, S.S., YAMADA, Y., VERA-BADILLO, F.E., MARTA, G.N., & MORAES, F.Y. (2020). Stereotactic Ablative Radiotherapy for the Management of Spinal Metastases. *JAMA Oncol.*, **6**(4), 567. 114

GONDI, V., CUI, Y., MEHTA, M.P., MANFREDI, D., XIAO, Y., GALVIN, J.M. *et al.* (2015). Real-time pretreatment review limits unacceptable deviations on a cooperative group radiation therapy technique trial: Quality assurance results of RTOG 0933. *Int. J. Radiat. Oncol. Biol. Phys.*, **91**(3), 564–570. 54

GOODMAN, K.A., WIEGNER, E.A., MATUREN, K.E., ZHANG, Z., MO, Q., YANG, G. *et al.* (2010). Dose-escalation study of single-fraction stereotactic body radiotherapy for liver malignancies. *Int. J. Radiat. Oncol. Biol. Phys.*, **78**(2), 486–493. 94

GUCKENBERGER, M., RICHTER, A., KRIEGER, T., WILBERT, J., BAIER, K., & FLENTJE, M. (2009). Is a single arc sufficient in volumetric-modulated arc therapy (VMAT) for complex-shaped target volumes? *Radiother. Oncol.*, **93**(2), 259–265. 78, 85

GUCKENBERGER, M., ANDRATSCHKE, N., DIECKMANN, K., HOOGEMAN, M.S., HOYER, M., HURKMANS, C., TANADINI-LANG, S., LARTIGAU, E., MÉNDEZ ROMERO, A., SENAN, S., & VERELLEN, D. (2017). ESTRO ACROP consensus guideline on implementation and practice of stereotactic body radiotherapy for peripherally located early stage non-small cell lung cancer. *Radiother. Oncol.*, **124**(1), 11–17. 114

GULLIFORD, S.L., PARTRIDGE, M., SYDES, M.R., WEBB, S., EVANS, P.M., & DEARNALEY, D.P. (2012). Parameters for the Lyman Kutcher Burman (LKB) model of normal tissue complication probability (NTCP) for specific rectal complications observed in clinical practise. *Radiother. Oncol.*, **102**(3), 347–351. 153, 163

HABRAKEN, S.J.M., SHARFO, A.W.M., BUIJSEN, J., VERBAKEL, W.F.A.R., HAASBEEK, C.J.A., ÖLLERS, M.C. *et al.* (2017). The TRENDY multi-center randomized trial on hepatocellular carcinoma – Trial QA including automated treatment planning and benchmark-case results. *Radiother. Oncol.*, **125**(3), 507–513. 54, 170

HAMDY, F.C., DONOVAN, J.L., LANE, J.A., MASON, M., METCALFE, C., HOLDING, P. *et al.* (2016). 10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. *N. Engl. J. Med.*, **375**(15), 1415–1424. 150

HANEDER, S., MICHAELY, H.J., SCHOENBERG, S.O., KONSTANDIN, S., SCHAD, L.R., SIEBENLIST, K. *et al.* (2012). Assessment of renal function after conformal radiotherapy and intensity-modulated radiotherapy by functional 1 H-MRI and 23 Na-MRI. *Strahlenther. Onkol.*, **188**(12), 1146–1154. 54

HANSEN, C.R., BERTELSEN, A., HAZELL, I., ZUKAUSKAITE, R., GYLDENKERNE, N., JOHANSEN, J. *et al.* (2016). Automatic treatment planning improves the clinical quality of head and neck cancer treatment plans. *Clin. Transl. Radiat. Oncol.*, **1**, 2–8. 141

HAZELL, I., BZDUSEK, K., KUMAR, P., HANSEN, C.R., BERTELSEN, A., ERIKSEN, J.G. *et al.* (2016). Automatic planning of head and neck treatment plans. *J. Appl. Clin. Med. Phys.*, **17**(1), 272–282. 37

HEIJKOOP, S.T., LANGERAK, T.R., QUINT, S., BONDAR, L., MENS, J.W.M., HEIJMEN, B.J.M. *et al.* (2014). Clinical implementation of an online adaptive plan-of-the-day protocol for nonrigid motion management in locally advanced cervical cancer IMRT. *Int. J. Radiat. Oncol. Biol. Phys.*, **90**(3), 673–679. 10, 11, 132

HEIJKOOP, S.T., WESTERVELD, H., BIJKER, N., FEIJE, R., SHARFO, A.W.M., VAN WIERINGEN, N. *et al.* (2016). Optimal patient positioning (prone versus supine) for VMAT in gynecologic cancer: A dosimetric study on the effect of different margins. *Int. J. Radiat. Oncol. Biol. Phys.*, **96**(2), 432–439. 96, 187

HEIJMEN, B.J.M., VOET, P.W.J., FRANSEN, D., PENNINKHOF, J., MILDER, M., AKHIAT, H. *et al.* (2018). Fully automated, multi-criterial planning for volumetric modulated arc therapy – An international multi-center validation for prostate cancer. *Radiother. Oncol.*, **128**(2), 343–348. 116, 150, 154, 164, 186

HERFARTH, K.K., DEBUS, J., LOHR, F., BAHNER, M.L., RHEIN, B., FRITZ, P. *et al.* (2001). Stereotactic single-dose radiation therapy of liver tumors: Results of a phase I/II trial. *J. Clin. Oncol.*, **19**(1), 164–170. 94

HERON, D.E., GERSZTEN, K., SELVARAJ, R.N., KING, G.C., SONNIK, D., GALLION, H. *et al.* (2003). Conventional 3D conformal versus intensity-modulated radiotherapy for the adjuvant treatment of gynecologic malignancies: A comparative dosimetric study of dose–volume histograms. *Gynecol. Oncol.*, **91**(1), 39–45. 78

HONG, T.S., BOSCH, W.R., KRISHNAN, S., KIM, T.K., MAMON, H.J., SHYN, P. *et al.* (2014). Interobserver variability in target definition for hepatocellular carcinoma with and without portal vein thrombus: Radiation therapy oncology group consensus guidelines. *Int. J. Radiat. Oncol. Biol. Phys.*, **89**(4), 804–813. 179, 181

HONG, T.S., & DAWSON, L.A. (2015). Randomized phase III study of focal radiation therapy for unresectable, localized intrahepatic cholangiocarcinoma. *Oncology NRG-GI001 version date 6/30/2015*, **6**, 30. 170

HOOGEMAN, M.S., NUYTTENS, J.J., LEVENDAG, P.C., & HEIJMEN, B.J.M. (2008). Time dependence of intrafraction patient motion assessed by repeat stereoscopic imaging. *Int. J. Radiat. Oncol. Biol. Phys.*, **70**(2), 609–618. 20

HOOVER, D.A., MACFARLANE, M., WONG, E., BATTISTA, J.J., & CHEN, J.Z. (2015). Feasibility of a unified approach to intensity-modulated radiation therapy and volume-modulated arc therapy optimization and delivery. *Med. Phys.*, **42**(2), 726–734. 107

HUNTER, D.J., & REDDY, K.S. (2013). Noncommunicable diseases. *N. Engl. J. Med.*, **369**(14), 1336–1343. 194

HUSSEIN, M., SOUTH, C.P., BARRY, M.A., ADAMS, E.J., JORDAN, T.J., STEWART, A.J. *et al.* (2016). Clinical validation and benchmarking of knowledge-based IMRT and VMAT treatment planning in pelvic anatomy. *Radiother. Oncol.*, **120**(3), 473–479. 54, 55, 70

IBBOTT, G.S., HAWORTH, A., & FOLLOWILL, D.S. (2013). Quality assurance for clinical trials. *Front. Oncol.*, **3**. 170

INCROCCI, L., WORTEL, R.C., ALEMAYEHU, W.G., ALUWINI, S., SCHIMMEL, E., KROL, S. *et al.* (2016). Hypofractionated versus conventionally fractionated radiotherapy for patients with localised prostate cancer (HYPRO): Final efficacy results from a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol.*, **17**(8), 1061–1069. 4, 150, 151

JABBARI, S., WEINBERG, V.K., KAPREALIAN, T., HSU, I.C., MA, L., CHUANG, C. *et al.* (2012). Stereotactic body radiotherapy as monotherapy or post–external beam radiotherapy boost for prostate cancer: Technique, early toxicity, and PSA response. *Int. J. Radiat. Oncol. Biol. Phys.*, **82**(1), 228–234. 131

JABBOUR, S.K., HASHEM, S.A., BOSCH, W., KIM, T.K., FINKELSTEIN, S.E., ANDERSON, B.M. *et al.* (2014). Upper abdominal normal organ contouring guidelines and atlas: A radiation therapy oncology group consensus. *Pract. Radiat. Oncol.*, **4**(2), 82–89. 172, 179, 181

JAGT, T., BREEDVELD, S., VAN HAVEREN, R., HEIJMEN, B.J.M., & HOOGEMAN, M.S. (2018). An automated planning strategy for near real-time adaptive proton therapy in prostate cancer. *Phys. Med. Biol.*, **63**(13), 135017. 193

JAGT, T.Z., BREEDVELD, S., VAN HAVEREN, R., NOUT, R.A., ASTREINIDOU, E., HEIJMEN, B.J.M., & HOOGEMAN, M.S. (2019). Plan-library supported automated replanning for online-adaptive intensity-modulated proton therapy of cervical cancer. *Acta Oncologica*, **58**(10), 1440–1445. 193

JANG, S.Y., LALONDE, R., OZHASOGLU, C., BURTON, S., HERON, D., & HUQ, M.S. (2016). Dosimetric comparison between cone/Iris-based and InCise MLC-based CyberKnife plans for single and multiple brain metastases. *J. Appl. Clin. Med. Phys.*, **17**(5), 184–199. 143

JANG, W.I., BAE, S.H., KIM, M.S., HAN, C.J., PARK, S.C., KIM, S.B., CHO, E.H., CHOI, C.W., KIM, K.S., HWANG, S., KIM, J.H., CHANG, A.R., PARK, Y., KIM, E.S., KIM, W.C., JO, S., & PARK, H.J. (2020). A phase 2 multicenter study of stereotactic body radiotherapy for hepatocellular carcinoma: Safety and efficacy. *Cancer*, **126**(2), 363–372. 114

Joye, I., Lambrecht, M., Jegou, D., Hortobágyi, E., Scalliet, P., & Haustermans, K. (2014). Does a central review platform improve the quality of radiotherapy for rectal cancer? Results of a national quality assurance project. *Radiother. Oncol.*, **111**(3), 400–405. 170

Kachnic, L.A., Winter, K., Myerson, R.J., Goodyear, M.D., Willins, J., Esthappan, J. *et al.* (2013). RTOG 0529: A phase 2 evaluation of dose-painted intensity modulated radiation therapy in combination with 5-Fluorouracil and Mitomycin-C for the reduction of acute morbidity in carcinoma of the anal canal. *Int. J. Radiat. Oncol. Biol. Phys.*, **86**(1), 27–33. 54

Kang, S.W., Chung, J.B., Kim, J.S., Kim, I.A., Eom, K.Y., Song, C., Lee, J.W., Kim, J.Y., & Suh, T.S. (2017). Optimal planning strategy among various arc arrangements for prostate stereotactic body radiotherapy with volumetric modulated arc therapy technique. *Radiol. Oncol.*, **51**(1), 112–120. 114

Kathriarachchi, V., Shang, C., Evans, G., Leventouri, T., & Kalantzis, G. (2016). Dosimetric and radiobiological comparison of CyberKnife M6 InCise multileaf collimator over IRIS variable collimator in prostate stereotactic body radiation therapy. *J. Med. Phys.*, **41**(2), 135. 143

Katz, A.J., Santoro, M., Diblasio, F., & Ashley, R. (2013). Stereotactic body radiotherapy for localized prostate cancer: Disease control and quality of life at 6 years. *Radiat. Oncol.*, **8**(1), 118. 131

Kavanagh, B.D., Pan, C.C., Dawson, L.A., Das, S.K., Li, X.A., Ten Haken, R.K. *et al.* (2010). Radiation dose–volume effects in the stomach and small bowel. *Int. J. Radiat. Oncol. Biol. Phys.*, **76**(3), S101–S107. 15, 20, 80

Keall, P.J., Ng, J.A., Juneja, P., O'Brien, R.T., Huang, C.Y., Colvill, E. *et al.* (2016). Real-time 3D image guidance using a standard LINAC: Measured motion, accuracy, and precision of the first prospective clinical trial of kilovoltage intrafraction monitoring–guided gating for prostate cancer radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, **94**(5), 1015–1021. 131, 143

Kearney, V., Cheung, J.P., McGuinness, C., & Solberg, T.D. (2017). CyberArc: a non-coplanar-arc optimization algorithm for CyberKnife. *Phys. Med. Biol.*, **62**(14), 5777–5789. 114

Kearney, V., Descovich, M., Sudhyadhom, A., Cheung, J.P., McGuinness, C., & Solberg, T.D. (2018). A continuous arc delivery optimization algorithm for CyberKnife m6. *Med. Phys.*, **45**(8), 3861–3870. 114

King, C.R., Freeman, D., Kaplan, I., Fuller, D., Bolzicco, G., Collins, S. *et al.* (2013). Stereotactic body radiotherapy for localized prostate cancer: Pooled analysis from a multi-institutional consortium of prospective phase II trials. *Radiother. Oncol.*, **109**(2), 217–221. 131

Kirkpatrick, J.P., van der Kogel, A.J., & Schultheiss, T.E. (2010). Radiation dose–volume effects in the spinal cord. *Int. J. Radiat. Oncol. Biol. Phys.*, **76**(3), S42–S49. 37

KRAYENBUEHL, J., NORTON, I., STUDER, G., & GUCKENBERGER, M. (2015). Evaluation of an automated knowledge based treatment planning system for head and neck. *Radiat. Oncol.*, **10**(1), 226. 131

KUPELIAN, P., WILLOUGHBY, T., MAHADEVAN, A., DJEMIL, T., WEINSTEIN, G., JANI, S. *et al.* (2007). 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.*, **67**(4), 1088–1098. 131

LANDIS, J.R., & KOCH, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, **33**(1), 159. 179, 181

LEE, E.K., FOX, T., & CROCKER, I. (2006). Simultaneous beam geometry and intensity map optimization in intensity-modulated radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, **64**(1), 301–320. 37

LEE, J., LIM, D.H., KIM, S., PARK, S.H., PARK, J.O., PARK, Y.S. *et al.* (2012). Phase III trial comparing Capecitabine plus Cisplatin versus Capecitabine plus Cisplatin with concurrent Capecitabine radiotherapy in completely resected gastric cancer with D2 lymph node dissection: The ARTIST trial. *J. Clin. Oncol.*, **30**(3), 268–273. 45

LEE, J., SHIN, I.S., YOON, W.S., KOOM, W.S., & RIM, C.H. (2020). Comparisons between radiofrequency ablation and stereotactic body radiotherapy for liver malignancies: Meta-analyses and a systematic review. *Radiother. Oncol.*, **145**(apr), 63–70. 114

LI, N., CARMONA, R., SIRAK, I., KASAOVA, L., FOLLOWILL, D., MICHALSKI, J. *et al.* (2017). Highly efficient training, refinement, and validation of a knowledge-based planning quality-control system for radiation therapy clinical trials. *Int. J. Radiat. Oncol. Biol. Phys.*, **97**(1), 164–172. 170

LI, Z., ZENG, J., WANG, Z., ZHU, H., & WEI, Y. (2014). Dosimetric comparison of intensity modulated and volumetric arc radiation therapy for gastric cancer. *Oncol. lett.*, **8**(4), 1427–1434. 46, 55

LIN, Y.W., LIN, K.H., HO, H.W., LIN, H.M., LIN, L.C., LEE, S.P., & CHUI, C.S. (2014a). Treatment plan comparison between stereotactic body radiation therapy techniques for prostate cancer: Non-isocentric CyberKnife versus isocentric RapidArc. *Phys. Med.*, **30**(6), 654–661. 114

LIN, Y.W., LIN, K.H., HO, H.W., LIN, H.M., LIN, L.C., LEE, S.P. *et al.* (2014b). Treatment plan comparison between stereotactic body radiation therapy techniques for prostate cancer: Non-isocentric CyberKnife versus isocentric RapidArc. *Phys. Med.*, **30**(6), 654–661. 131, 142

LINTHOUT, N., VERELLEN, D., TOURNEL, K., REYNDERS, T., DUCHATEAU, M., & STORME, G. (2007). Assessment of secondary patient motion induced by automated couch movement during on-line 6 dimensional repositioning in prostate cancer treatment. *Radiother. Oncol.*, **83**(2), 168–174. 142

LIU, E., STENMARK, M.H., SCHIPPER, M.J., BALTER, J.M., KESSLER, M.L., CAOILI, E.M. *et al.* (2013). Stereotactic body radiation therapy for primary and metastatic liver tumors. *Trans. Oncol.*, **6**(4), 442–446. 94

LOHR, F., DOBLER, B., MAI, S., HERMANN, B., TIEFENBACHER, U., WIELAND, P. *et al.* (2003). Optimization of dose distributions for adjuvant locoregional radiotherapy of gastric cancer by IMRT. *Strahlenther. Onkol.*, **179**(8), 557–563. 54

LOW, D.A., HARMS, W.B., MUTIC, S., & PURDY, J.A. (1998). A technique for the quantitative evaluation of dose distributions. *Med. Phys.*, **25**(5), 656–661. 98

LUTZ, S., BERK, L., CHANG, E., CHOW, E., HAHN, C., HOSKIN, P. *et al.* (2011). Palliative radiotherapy for bone metastases: An ASTRO evidence-based guideline. *Int. J. Radiat. Oncol. Biol. Phys.*, **79**(4), 965–976. 36

LYU, Q., YU, V.Y., RUAN, D., NEPH, R., O'CONNOR, D., & SHENG, K. (2018). A novel optimization framework for VMAT with dynamic gantry couch rotation. *Phys. Med. Biol.*, **63**(12), 125013. 114

MACDONALD, J.S., SMALLEY, S.R., BENEDETTI, J., HUNDAHL, S.A., ESTES, N.C., STEMMERMANN, G.N. *et al.* (2001). Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. *N. Engl. J. Med.*, **345**(10), 725–730. 45, 54

MACDOUGALL, N.D., DEAN, C., & MUIRHEAD, R. (2014). Stereotactic body radiotherapy in prostate cancer: Is Rapidarc a better solution than CyberKnife? *Clin. Oncol.*, **26**(1), 4–9. 131, 142

MACFARLANE, M., HOOVER, D.A., WONG, E., READ, N., PALMA, D., VENKATESAN, V. *et al.* (2016). Evaluation of unified intensity-modulated arc therapy for the radiotherapy of head-and-neck cancer. *Radiother. Oncol.*, **119**(2), 331–336. 107

MADSEN, B.L., HSI, R.A., PHAM, H.T., FOWLER, J.F., ESAGUI, L., & CORMAN, J. (2007). Stereotactic hypofractionated accurate radiotherapy of the prostate (SHARP), 33.5 Gy in five fractions for localized disease: First clinical trial results. *Int. J. Radiat. Oncol. Biol. Phys.*, **67**(4), 1099–1105. 131

MANTINI, G., TAGLIAFERRI, L., MATTIUCCI, G.C., BALDUCCI, M., FRASCINO, V., DINAPOLI, N. *et al.* (2011). Effect of whole pelvic radiotherapy for patients with locally advanced prostate cancer treated with radiotherapy and long-term androgen deprivation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, **81**(5), e721–e726. 60

MARKS, L.B., BENTZEN, S.M., DEASY, J.O., KONG, F.M.S., BRADLEY, J.D., VOGELIUS, I.S. *et al.* (2010a). Radiation dose–volume effects in the lung. *Int. J. Radiat. Oncol. Biol. Phys.*, **76**(3), S70–S76. 49

MARKS, L.B., YORKE, E.D., JACKSON, A., TEN HAKEN, R.K., CONSTINE, L.S., EISBRUCH, A. *et al.* (2010b). Use of normal tissue complication probability models in the clinic. *Int. J. Radiat. Oncol. Biol. Phys.*, **76**(3), S10–S19. 29

MATUSZAK, M.M., STEERS, J.M., LONG, T., MCSHAN, D.L., FRAASS, B.A., ROMEIJN, E. *et al.* (2013). FusionArc optimization: A hybrid volumetric modulated arc therapy (VMAT) and intensity modulated radiation therapy (IMRT) planning strategy. *Med. Phys.*, **40**(7), 071713. 107

MAYLES, W.P.M. (2010). Survey of the availability and use of advanced radiotherapy technology in the UK. *Clin. Oncol.*, **22**(8), 636–642. 28, 37

MCGUINNESS, C.M., GOTTSCHALK, A.R., LESSARD, E., NAKAMURA, J.L., PINNADUWAGE, D., POULIOT, J. *et al.* (2015). Investigating the clinical advantages of a robotic linac equipped with a multileaf collimator in the treatment of brain and prostate cancer patients. *J. Appl. Clin. Med. Phys.*, **16**(5), 284–295. 143

MCVICAR, N., POPESCU, I.A., & HEATH, E. (2016). Techniques for adaptive prostate radiotherapy. *Phys. Med.*, **32**(3), 492–498. 70

MELIDIS, C., BOSCH, W.R., IZEWSKA, J., FIDAROVA, E., ZUBIZARRETA, E., ULIN, K. *et al.* (2014). Global harmonization of quality assurance naming conventions in radiation therapy clinical trials. *Int. J. Radiat. Oncol. Biol. Phys.*, **90**(5), 1242–1249. 170

MÉNDEZ ROMERO, A., WUNDERINK, W., HUSSAIN, S.M., DE POOTER, J.A., HEIJMEN, B.J.M., NOWAK, P.C.J.M. *et al.* (2006). Stereotactic body radiation therapy for primary and metastatic liver tumors: A single institution phase I-II study. *Acta Oncol.*, **45**(7), 831–837. 94

MICHALSKI, J.M., GAY, H., JACKSON, A., TUCKER, S.L., & DEASY, J.O. (2010). Radiation dose–volume effects in radiation-induced rectal injury. *Int. J. Radiat. Oncol. Biol. Phys.*, **76**(3), S123–S129. 64, 134, 150, 152

MONDLANE, G., GUBANSKI, M., LIND, P.A., UREBA, A., & SIEGBAHN, A. (2017). Comparison of gastric-cancer radiotherapy performed with volumetric modulated arc therapy or single-field uniform-dose proton therapy. *Acta Oncol.*, **56**(6), 832–838. 46, 55

MONZ, M., KÜFER, K.H., BORTFELD, T.R., & THIEKE, C. (2008). Pareto navigation—algorithmic foundation of interactive multi-criteria IMRT planning. *Phys. Med. Biol.*, **53**(4), 985–998. 45

MOORE, K.L., SCHMIDT, R., MOISEENKO, V., OLSEN, L.A., TAN, J., XIAO, Y. *et al.* (2015). Quantifying unnecessary normal tissue complication risks due to suboptimal planning: A secondary study of RTOG 0126. *Int. J. Radiat. Oncol. Biol. Phys.*, **92**(2), 228–235. 170, 180, 190

MORGAN, S.C., HOFFMAN, K., LOBLAW, D.A., BUYYOUNOUSKI, M.K., PATTON, C., BAROCAS, D., BENTZEN, S., CHANG, M., EFSTATHIOU, J., GREANY, P., HALVORSEN, P., KOONTZ, B.F., LAWTON, C., LEYRER, C.M., LIN, D., RAY, M., & SANDLER, H. (2018). Hypofractionated Radiation Therapy for Localized Prostate Cancer: Executive Summary of an ASTRO, ASCO, and AUA Evidence-Based Guideline. *Pract. Radiat. Oncol.*, **8**(6). 114

MORIKAWA, L.K., & ROACH, M. (2011). Pelvic nodal radiotherapy in patients with unfavorable intermediate and high-risk prostate cancer: Evidence, rationale, and future directions. *Int. J. Radiat. Oncol. Biol. Phys.*, **80**(1), 6–16. 60

MUNDT, A.J., LUJAN, A.E., ROTMENSCH, J., WAGGONER, S.E., YAMADA, S.D., FLEMING, G. *et al.* (2002). Intensity-modulated whole pelvic radiotherapy in women with gynecologic malignancies. *Int. J. Radiat. Oncol. Biol. Phys.* 10, 78

NELMS, B.E., ROBINSON, G., MARKHAM, J., VELASCO, K., BOYD, S., NARAYAN, S. *et al.* (2012). Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Pract. Radiat. Oncol.*, **2**(4), 296–305. 38, 141, 170

NIEMIERKO, A. (1997). Reporting and analyzing dose distributions: A concept of equivalent uniform dose. *Med. Phys.*, **24**(1), 103–110. 49, 71, 145

NOVÁKOVÁ, E., HEIJKOOP, S.T., QUINT, S., ZOLNAY, A.G., MENS, J.W.M., GODART, J. *et al.* (2017). What is the optimal number of library plans in ART for locally advanced cervical cancer? *Radiother. Oncol.*, **125**(3), 470–477. 186

OHIRA, S., UEDA, Y., AKINO, Y., HASHIMOTO, M., MASAOKA, A., HIRATA, T., MIYAZAKI, M., KOIZUMI, M., & TESHIMA, T. (2018). HyperArc VMAT planning for single and multiple brain metastases stereotactic radiosurgery: a new treatment planning approach. *Radiat. Oncol.*, **13**(1), 13. 125

OHRI, N., SHEN, X., DICKER, A.P., DOYLE, L.A., HARRISON, A.S., & SHOWALTER, T.N. (2013). Radiotherapy protocol deviations and clinical outcomes: A meta-analysis of cooperative group clinical trials. *J. Natl. Cancer Inst.*, **105**(6), 387–393. 170, 190

OHTAKARA, K., HAYASHI, S., & HOSHI, H. (2012). The relation between various conformity indices and the influence of the target coverage difference in prescription isodose surface on these values in intracranial stereotactic radiosurgery. *Br. J. Radiol.*, **85**(1014), e223–e228. 32

OLIVER, M., ANSBACHER, W., & BECKHAM, W.A. (2009). Comparing planning time, delivery time and plan quality for IMRT, RapidArc and tomotherapy. *J. Appl. Clin. Med. Phys.*, **10**(4), 117–131. 37

ONAL, C., DÖLEK, Y., & AKKUŞ YILDIRIM, B. (2018). Dosimetric comparison of 3-dimensional conformal radiotherapy, volumetric modulated arc therapy, and helical tomotherapy for postoperative gastric cancer patients. *Jpn. J. Radiol.*, **36**(1), 30–39. 46, 55

OUD, M., KOLKMAN-DEURLOO, I.K., MENS, J.W., LATHOUWERS, D., PERKÓ, Z., HEIJMEN, B.J.M., & BREEDVELD, S. (2020). Fast and fully-automated multi-criterial treatment planning for adaptive HDR brachytherapy for locally advanced cervical cancer. *Radiotherapy and Oncology*, **148**(jul), 143–150. 187, 193

PAIK, E.K., KIM, M.S., CHOI, C.W., JANG, W.I., LEE, S.H., CHOI, S.H. *et al.* (2015). Dosimetric comparison of volumetric modulated arc therapy with robotic stereotactic radiation therapy in hepatocellular carcinoma. *Radiat. Oncol. J.*, **33**(3), 233. 104

PARK, S.H., SOHN, T.S., LEE, J., LIM, D.H., HONG, M.E., KIM, K.M. *et al.* (2015). Phase III trial to compare adjuvant chemotherapy with Capecitabine and Cisplatin versus concurrent chemoradiotherapy in gastric cancer: Final report of the adjuvant chemoradiotherapy in stomach tumors trial, including survival and subset analyses. *J. Clin. Oncol.*, **33**(28), 3130–3136. 45

PEETERS, S.T.H., HEEMSBERGEN, W.D., VAN PUTTEN, W.L.J., SLOT, A., TABAK, H., MENS, J.W.M. *et al.* (2005). Acute and late complications after radiotherapy for prostate cancer: Results of a multicenter randomized trial comparing 68 Gy to 78 Gy. *Int. J. Radiat. Oncol. Biol. Phys.*, **61**(4), 1019–1034. 150

PEETERS, S.T.H., HEEMSBERGEN, W.D., KOPER, P.C.M., VAN PUTTEN, W.L.J., SLOT, A., DIELWART, M.F.H. *et al.* (2006). Dose-response in radiotherapy for localized prostate cancer: Results of the Dutch multicenter randomized phase III trial comparing 68 Gy of radiotherapy with 78 Gy. *J. Clin. Oncol.*, **24**(13), 1990–1996. 150

PENG, C., AHUNBAY, E., CHEN, G., ANDERSON, S., LAWTON, C., & LI, X.A. (2011). Characterizing interfraction variations and their dosimetric effects in prostate cancer radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, **79**(3), 909–914. 70

PORTELANCE, L., CHAO, K.S.C., GRIGSBY, P.W., BENNET, H., & LOW, D. (2001). Intensity-modulated radiation therapy (IMRT) reduces small bowel, rectum, and bladder doses in patients with cervical cancer receiving pelvic and para-aortic irradiation. *Int. J. Radiat. Oncol. Biol. Phys.*, **51**(1), 261–266. 78

POTREBKO, P.S., MCCURDY, B.M.C., BUTLER, J.B., & EL-GUBTAN, A.S. (2008). Improving intensity-modulated radiation therapy using the anatomic beam orientation optimization algorithm. *Med. Phys.*, **35**(5), 2170–2179. 37

PURDIE, T.G., DINNIWELL, R.E., FYLES, A., & SHARPE, M.B. (2014). Automation and intensity modulated radiation therapy for individualized high-quality tangent breast treatment plans. *Int. J. Radiat. Oncol. Biol. Phys.*, **90**(3), 688–695. 131

QI, P., POULIOT, J., ROACH, M., & XIA, P. (2014). Offline multiple adaptive planning strategy for concurrent irradiation of the prostate and pelvic lymph nodes. *Med. Phys.*, **41**(2), 021704. 70

QUAN, E.M., LI, X., LI, Y., WANG, X., KUDCHADKER, R.J., JOHNSON, J.L. *et al.* (2012a). A comprehensive comparison of IMRT and VMAT plan quality for prostate cancer treatment. *Int. J. Radiat. Oncol. Biol. Phys.*, **83**(4), 1169–1178. 21, 78, 85

QUAN, E.M., CHANG, J.Y., LIAO, Z., XIA, T., YUAN, Z., LIU, H. *et al.* (2012b). Automated volumetric modulated arc therapy treatment planning for stage III lung cancer: How does it compare with intensity-modulated radiotherapy? *Int. J. Radiat. Oncol. Biol. Phys.*, **84**(1), e69–e76. 21, 131

RANCATI, T., FIORINO, C., FELLIN, G., VAVASSORI, V., CAGNA, E., CASANOVA BORCA, V. *et al.* (2011). Inclusion of clinical risk factors into NTCP modelling of late rectal toxicity after high dose radiotherapy for prostate cancer. *Radiother. Oncol.*, **100**(1), 124–130. 153, 163

RAO, M., YANG, W., CHEN, F., SHENG, K., YE, J., MEHTA, V. *et al.* (2010). Comparison of Elekta VMAT with helical tomotherapy and fixed field IMRT: Plan quality, delivery efficiency and accuracy. *Med. Phys.*, **37**(3), 1350–1359. 37

RESNICK, M.J., KOYAMA, T., FAN, K.H., ALBERTSEN, P.C., GOODMAN, M., HAMILTON, A.S. *et al.* (2013). Long-term functional outcomes after treatment for localized prostate cancer. *N. Engl. J. Med.*, **368**(5), 436–445. 150

ROACH, M., DESILVIO, M., LAWTON, C., UHL, V., MACHTAY, M., SEIDER, M.J. *et al.* (2003). Phase III trial comparing whole-pelvic versus prostate-only radiotherapy and neoadjuvant versus adjuvant combined androgen suppression: Radiation therapy oncology group 9413. *J. Clin. Oncol.*, **21**(10), 1904–1911. 60

ROBAR, J.L., & THOMAS, C. (2012). HybridArc: A novel radiation therapy technique combining optimized dynamic arcs and intensity modulation. *Med. Dosim.*, **37**(4), 358–368. 107

ROESKE, J.C., BONTA, D., MELL, L.K., LUJAN, A.E., & MUNDT, A.J. (2003). A dosimetric analysis of acute gastrointestinal toxicity in women receiving intensity-modulated whole-pelvic radiation therapy. *Radiother. Oncol.*, **69**(2), 201–207. 12, 79

ROSSI, L., BREEDVELD, S., HEIJMEN, B.J.M., VOET, P.W.J., LANCONELLI, N., & ALUWINI, S. (2012). On the beam direction search space in computerized non-coplanar beam angle optimization for IMRT - Prostate SBRT. *Phys. Med. Biol.*, **57**(17), 5441–5458. 28, 78, 80, 96, 114, 131, 132, 187

ROSSI, L., BREEDVELD, S., ALUWINI, S., & HEIJMEN, B.J.M. (2015). Noncoplanar beam angle class solutions to replace time-consuming patient-specific beam angle optimization in robotic prostate stereotactic body radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, **92**(4), 762–770. 96

ROSSI, L., MÉNDEZ ROMERO, A., MILDER, M., DE KLERCK, E., BREEDVELD, S., & HEIJMEN, B.J.M. (2017). Automated planning to reduce integral dose in robotic radiosurgery for benign tumors. *Radiother. Oncol.*, **123**(1_suppl), S449–S450. 186

ROSSI, L., SHARFO, A.W.M., ALUWINI, S., DIRKX, M.L.P., BREEDVELD, S., & HEIJMEN, B.J.M. (2018). First fully automated planning solution for robotic radiosurgery – Comparison with automatically planned volumetric arc therapy for prostate cancer. *Acta Oncol.*, **57**(11), 1490–1498. 114, 116, 124, 191

ROSSI, M.M.G., PEULEN, H.M.U., BELDERBOS, J.S.A., & SONKE, J.J. (2016). Intrafraction motion in stereotactic body radiation therapy for non-small cell lung cancer: Intensity modulated radiation therapy versus volumetric modulated arc therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, **95**(2), 835–843. 188

RUSTHOVEN, K.E., KAVANAGH, B.D., CARDENES, H., STIEBER, V.W., BURRI, S.H., FEIGENBERG, S.J. *et al.* (2009). Multi-institutional phase I/II trial of stereotactic body radiation therapy for liver metastases. *J. Clin. Oncol.*, **27**(10), 1572–1578. 94

SCHEFTER, T.E., KAVANAGH, B.D., TIMMERMAN, R.D., CARDENES, H.R., BARON, A., & GASPAR, L.E. (2005). A phase I trial of stereotactic body radiation therapy (SBRT) for liver metastases. *Int. J. Radiat. Oncol. Biol. Phys.*, **62**(5), 1371–1378. 94

SCHNEIDER, B.J., DALY, M.E., KENNEDY, E.B., ANTONOFF, M.B., BRODERICK, S., FELDMAN, J., JOLLY, S., MEYERS, B., ROCCO, G., RUSTHOVEN, C., SLOTMAN, B.J., STERMAN, D.H., & STILES, B.M. (2018). Stereotactic Body Radiotherapy for Early-Stage Non–Small-Cell Lung Cancer: American Society of Clinical Oncology Endorsement of the American Society for Radiation Oncology Evidence-Based Guideline. *J. Clin. Oncol.*, **36**(7), 710–719. 114

SCHREIBMANN, E., LAHANAS, M., XING, L., & BALTAS, D. (2004). Multiobjective evolutionary optimization of the number of beams, their orientations and weights for intensity-modulated radiation therapy. *Phys. Med. Biol.*, **49**(5), 747–770. 28, 37

SHARFO, A.W.M., VOET, P.W.J., BREEDVELD, S., MENS, J.W.M., HOOGEMAN, M.S., & HEIJMEN, B.J.M. (2015). Comparison of VMAT and IMRT strategies for cervical cancer patients using automated planning. *Radiother. Oncol.*, **114**(3), 395–401. 13, 28, 37, 38, 60, 96, 132, 141

SHARFO, A.W.M., BREEDVELD, S., VOET, P.W.J., HEIJKOOP, S.T., MENS, J.W.M., HOOGEMAN, M.S. *et al.* (2016). Validation of fully automated VMAT plan generation for library-based plan-of-the-day cervical cancer radiotherapy. *PLoS ONE*, **11**(12), e0169202. 45, 46, 48, 54, 60, 96, 116, 131, 150, 154, 164, 171, 173

SHARFO, A.W.M., DIRKX, M.L.P., BREEDVELD, S., ROMERO, A.M., & HEIJMEN, B.J.M. (2017). VMAT plus a few computer-optimized non-coplanar IMRT beams (VMAT+) tested for liver SBRT. *Radiother. Oncol.*, **123**(1), 49–56. 114, 123, 125

SHARFO, A.W.M., DIRKX, M.L.P., BIJMAN, R.G., SCHILLEMANS, W., BREEDVELD, S., ALUWINI, S. *et al.* (2018). Late toxicity in the randomized multicenter HYPRO trial for prostate cancer analyzed with automated treatment planning. *Radiother. Oncol.*, **128**(2), 349–356. 116, 180, 190

SHIKAMA, N., TSUJINO, K., NAKAMURA, K., & ISHIKURA, S. (2014). Survey of advanced radiation technologies Used at designated cancer care hospitals in Japan. *Jpn. J. Clin. Oncol.*, **44**(1), 72–77. 28, 37

SHIRAISHI, S., TAN, J., OLSEN, L. a, & MOORE, K.L. (2015). Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery. *Med. Phys.*, **42**(2), 908–917. 21

SLOSAREK, K., BEKMAN, B., WENDYKIER, J., GRZĄDZIEL, A., FOGLIATA, A., & COZZI, L. (2018). In silico assessment of the dosimetric quality of a novel, automated radiation treatment planning strategy for linac-based radiosurgery of multiple brain metastases and a comparison with robotic methods. *Radiat. Oncol.*, **13**(1), 41. 125

SMALLEY, S.R., BENEDETTI, J.K., HALLER, D.G., HUNDAHL, S.A., ESTES, N.C., AJANI, J.A. *et al.* (2012). Updated analysis of SWOG-directed intergroup study 0116: A phase III trial of adjuvant radiochemotherapy versus observation after curative gastric cancer resection. *J. Clin. Oncol.*, **30**(19), 2327–2333. 45, 54

SON, S.H., CHOI, B.O., RYU, M.R., KANG, Y.N., JANG, J.S., BAE, S.H. *et al.* (2010). Stereotactic body radiotherapy for patients with unresectable primary hepatocellular carcinoma: Dose-volumetric parameters predicting the hepatic complication. *Int. J. Radiat. Oncol. Biol. Phys.*, **78**(4), 1073–1080. 172

STEENBAKKERS, R.J.H.M., DUPPEN, J.C., FITTON, I., DEURLOO, K.E.I., ZIJP, L.J., COMANS, E.F.I. *et al.* (2006). Reduction of observer variation using matched CT-PET for lung cancer delineation: A three-dimensional analysis. *Int. J. Radiat. Oncol. Biol. Phys.*, **64**(2), 435–448. 170

STIELER, F., WOLFF, D., SCHMID, H., WELZEL, G., WENZ, F., & LOHR, F. (2011). A comparison of several modulated radiotherapy techniques for head and neck cancer and dosimetric validation of VMAT. *Radiother. Oncol.*, **101**(3), 388–393. 28

TEOH, M., CLARK, C.H., WOOD, K., WHITAKER, S., & NISBET, A. (2011). Volumetric modulated arc therapy: A review of current literature and clinical use in practice. *Br. J. Radiol.*, **84**(1007), 967–996. 78, 85, 88

THOMAS, E.M., POPPLE, R.A., WU, X., CLARK, G.M., MARKERT, J.M., GUTHRIE, B.L., YUAN, Y., DOBELBOWER, M.C., SPENCER, S.A., & FIVEASH, J.B. (2014). Comparison of Plan Quality and Delivery Time Between Volumetric Arc Therapy (RapidArc) and Gamma Knife Radiosurgery for Multiple Cranial Metastases. *Neurosurgery*, **75**(4), 409–418. 125

THOR, M., JACKSON, A., ZELEFSKY, M.J., STEINECK, G., KARLSDÒTTIR, A., HØYER, M. *et al.* (2018). Inter-institutional analysis demonstrates the importance of lower than previously anticipated dose regions to prevent late rectal bleeding following prostate radiotherapy. *Radiother. Oncol.*, **127**(1), 88–95. 163

TIWANA, M.S., BARNES, M., KIRALY, A., & OLSON, R.A. (2016). Utilization of palliative radiotherapy for bone metastases near end of life in a population-based cohort. *BMC Palliat. Care*, **15**(1), 2. 28

TOL, J.P., DELANEY, A.R., DAHELE, M., SLOTMAN, B.J., & VERBAKEL, W.F.A.R. (2015). Evaluation of a knowledge-based planning solution for head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, **91**(3), 612–620. 21

TOL, J.P., DAHELE, M., DELANEY, A.R., DOORNAERT, P., SLOTMAN, B.J., & VERBAKEL, W.F.A.R. (2016). Detailed evaluation of an automated approach to interactive optimization for volumetric modulated arc therapy plans. *Med. Phys.*, **43**(4), 1818–1828. 131

TORRE, L.A., BRAY, F., SIEGEL, R.L., FERLAY, J., LORTET-TIEULENT, J., & JEMAL, A. (2015). Global cancer statistics, 2012. *CA Cancer J. Clin.* 150

VAN DE BUNT, L., VAN DER HEIDE, U.A., KETELAARS, M., DE KORT, G.A.P., & JÜRGENLIEMK-SCHULZ, I.M. (2006). Conventional, conformal, and intensity-modulated radiation therapy treatment planning of external beam radiotherapy for cervical cancer: The impact of tumor regression. *Int. J. Radiat. Oncol. Biol. Phys.*, **64**(1), 189–196. 10, 78

VAN DE SANDE, M.A.E., CREUTZBERG, C.L., VAN DE WATER, S., SHARFO, A.W.M., & HOOGEMAN, M.S. (2016). Which cervical and endometrial cancer patients will benefit most from intensity-modulated proton therapy? *Radiother. Oncol.*, **120**(3), 397–403. 188

VAN DE WATER, S., HOOGEMAN, M.S., BREEDVELD, S., & HEIJMEN, B.J.M. (2011a). Shortening treatment time in robotic radiosurgery using a novel node reduction technique. *Med. phys.*, **38**(3), 1397–1405. 133

VAN DE WATER, S., HOOGEMAN, M.S., BREEDVELD, S., NUYTTENS, J.J.M.E., SCHAART, D.R., & HEIJMEN, B.J.M. (2011b). Variable circular collimator in robotic radiosurgery: A time-efficient alternative to a mini-multileaf collimator? *Int. J. Radiat. Oncol. Biol. Phys.*, **81**(3), 863–870. 143

VAN DE WATER, S., VALLI, L., ALUWINI, S., LANCONELLI, N., HEIJMEN, B.J.M., & HOOGEMAN, M.S. (2014). Intrafraction prostate translations and rotations during hypofractionated robotic radiation surgery: Dosimetric impact of correction strategies and margins. *Int. J. Radiat. Oncol. Biol. Phys.*, **88**(5), 1154–1160. 131, 132

VAN DER POOL, A.E.M., MÉNDEZ ROMERO, A., WUNDERINK, W., HEIJMEN, B.J.M., LEVENDAG, P.C., VERHOEF, C. *et al.* (2010). Stereotactic body radiation therapy for colorectal liver metastases. *Br. J. Surg.*, **97**(3), 377–382. 94

VAN HAGEN, P., HULSHOF, M.C.C.M., VAN LANSCHOT, J.J.B., STEYERBERG, E.W., HENEGOUWEN, M.I. VAN BERGE, WIJNHOVEN, B.P.L. *et al.* (2012). Preoperative chemoradiotherapy for esophageal or junctional cancer. *N. Engl. J. Med.*, **366**(22), 2074–2084. 45

VAN HAVEREN, R., OGRYCZAK, W., VERDUIJN, G.M., KEIJZER, M., HEIJMEN, B.J.M., & BREEDVELD, S. (2017). Fast and fuzzy multi-objective radiotherapy treatment plan generation for head and neck cancer patients with the lexicographic reference point method (LRPM). *Phys. Med. Biol.*, **62**(11), 4318–4332. 193

VAN HAVEREN, R., HEIJMEN, B.J.M., & BREEDVELD, S. (2019). Automatically configuring the reference point method for automated multi-objective treatment planning. *Physics in Medicine & Biology*, **64**(3), 035002. 193

VAN HAVEREN, R., HEIJMEN, B.J.M., & BREEDVELD, S. (2020). Automatic configuration of the reference point method for fully automated multi-objective treatment planning applied to oropharyngeal cancer. *Medical Physics*, **47**(4), 1499–1508. 193

VERHEIJ, M., JANSEN, E.P.M., CATS, A., VAN GRIEKEN, N.C.T., AARONSON, N.K., BOOT, H. *et al.* (2016). A multicenter randomized phase III trial of neo-adjuvant chemotherapy followed by surgery and chemotherapy or by surgery and chemoradiotherapy in resectable gastric cancer: First results from the CRITICS study. *J. Clin. Oncol.*, **34**(15_suppl), 4000–4000. 45

VIDETIC, G.M.M., DONINGTON, J., GIULIANI, M., HEINZERLING, J., KARAS, T.Z., KELSEY, C.R., LALLY, B.E., LATZKA, K., LO, S.S., MOGHANAKI, D., MOVSAS, B., RIMNER, A., ROACH, M., RODRIGUES, G., SHIRVANI, S.M., SIMONE, C.B., TIMMERMAN, R., & DALY,

M.E. (2017). Stereotactic body radiation therapy for early-stage non-small cell lung cancer: Executive Summary of an ASTRO Evidence-Based Guideline. *Pract. Radiat. Oncol.*, **7**(5), 295–301. 114

VOET, P.W.J., BREEDVELD, S., DIRKX, M.L.P., LEVENDAG, P.C., & HEIJMEN, B.J.M. (2012). Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head and neck IMRT and implications for VMAT. *Med. Phys.*, **39**(8), 4858–4865. 28, 78, 80, 94, 96, 187

VOET, P.W.J., DIRKX, M.L.P., BREEDVELD, S., & HEIJMEN, B.J.M. (2013a). Automated generation of IMRT treatment plans for prostate cancer patients with metal hip prostheses: Comparison of different planning strategies. *Med. Phys.*, **40**(7), 071704. 70, 78, 96, 187

VOET, P.W.J., DIRKX, M.L.P., BREEDVELD, S., FRANSEN, D., LEVENDAG, P.C., & HEIJMEN, B.J.M. (2013b). Toward fully automated multicriterial plan generation: A prospective clinical study. *Int. J. Radiat. Oncol. Biol. Phys.*, **85**(3), 866–872. 10, 37, 38, 45, 46, 48, 54, 60, 78, 80, 87, 132, 141, 150, 154, 186

VOET, P.W.J., DIRKX, M.L.P., BREEDVELD, S., AL-MAMGANI, A., INCROCCI, L., & HEIJMEN, B.J.M. (2014). Fully automated volumetric modulated arc therapy plan generation for prostate cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.*, **88**(5), 1175–1179. 10, 13, 14, 21, 28, 29, 37, 38, 45, 46, 48, 54, 60, 79, 89, 94, 96, 115, 131, 132, 141, 164, 171, 173, 186

WANG, P.M., HSU, W.C., CHUNG, N.N., CHANG, F.L., JANG, C.J., FOGLIATA, A. *et al.* (2014). Feasibility of stereotactic body radiation therapy with volumetric modulated arc therapy and high intensity photon beams for hepatocellular carcinoma patients. *Radiat. Oncol.*, **9**(1), 18. 94, 104

WANG, Y., HEIJMEN, B.J.M., & PETIT, S.F. (2017). Prospective clinical validation of independent DVH prediction for plan QA in automatic treatment planning for prostate cancer patients. *Radiother. Oncol.*, **125**(3), 500–506. 192

WEBER, D.C., TOMSEJ, M., MELIDIS, C., & HURKMANS, C.W. (2012). QA makes a clinical trial stronger: Evidence-based medicine in radiation therapy. *Radiother. Oncol.*, **105**(1), 4–8. 170, 190

WIELAND, P., DOBLER, B., MAI, S., HERMANN, B., TIEFENBACHER, U., STEIL, V. *et al.* (2004). IMRT for postoperative treatment of gastric cancer: Covering large target volumes in the upper abdomen: A comparison of a step-and-shoot and an arc therapy approach. *Int. J. Radiat. Oncol. Biol. Phys.*, **59**(4), 1236–1244. 54

WILKE, L., ANDRATSCHKE, N., BLANCK, O., BRUNNER, T.B., COMBS, S.E., GROSU, A.L., MOUSTAKIS, C., SCHMITT, D., BAUS, W.W., & GUCKENBERGER, M. (2019). ICRU report 91 on prescribing, recording, and reporting of stereotactic treatments with small photon beams: Statement from the DEGRO/DGMP working group stereotactic radiotherapy and radiosurgery. *Strahlenther. Onkol.*, **195**(3). 114

WILLIAMS, M.J., BAILEY, M.J., FORSTNER, D., & METCALFE, P.E. (2007). Multicentre quality assurance of intensity-modulated radiation therapy plans: A precursor to clinical trials. *Australas. Radiol.*, **51**(5), 472–479. 170

WILLOUGHBY, T.R., KUPELIAN, P.A., POULIOT, J., SHINOHARA, K., AUBIN, M., ROACH, M. *et al.* (2006). Target localization and real-time tracking using the Calypso 4D localization system in patients with localized prostate cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, **65**(2), 528–534. 131, 143

WONG, K.K., ALL, S., WAXER, J., OLCH, A.J., VENKATRAMANI, R., DHALL, G. *et al.* (2017). Radiotherapy after high-dose chemotherapy with autologous hematopoietic cell rescue: Quality assessment of head start III. *Pediatr. Blood Cancer*, **64**(10), e26529. 54

WOODS, K., NGUYEN, D., TRAN, A., YU, V.Y., CAO, M., NIU, T. *et al.* (2016). Viability of noncoplanar VMAT for liver SBRT compared with coplanar VMAT and beam orientation optimized 4π IMRT. *Adv. Radiat. Oncol.*, **1**(1), 67–75. 94, 104, 125

WU, B., PANG, D., SIMARI, P., TAYLOR, R., SANGUINETI, G., & MCNUTT, T. (2013). Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: A head-and-neck case study. *Med. Phys.*, **40**(2), 021714. 21

WU, B., PANG, D., LEI, S., GATTI, J., TONG, M., MCNUTT, T. *et al.* (2014). Improved robotic stereotactic body radiation therapy plan quality and planning efficacy for organ-confined prostate cancer utilizing overlap-volume histogram-driven planning methodology. *Radiother. Oncol.*, **112**(2), 221–226. 131

YCHOU, M., BOIGE, V., PIGNON, J.P., CONROY, T., BOUCHÉ, O., LEBRETON, G. *et al.* (2011). Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: An FNCLCC and FFCD multicenter phase III trial. *J. Clin. Oncol.*, **29**(13), 1715–1721. 45

YUAN, L., GE, Y., LEE, W.R., YIN, F.F., KIRKPATRICK, J.P., & WU, Q.J. (2012). Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. *Med. Phys.*, **39**(11), 6868–6878. 45

ZAREPISHEH, M., LONG, T., LI, N., TIAN, Z., ROMEIJN, E., JIA, X. *et al.* (2014). A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning. *Med. Phys.*, **41**(6), 061711. 131

ZELEFSKY, M.J., KOLLMEIER, M., COX, B., FIDALEO, A., SPERLING, D., PEI, X. *et al.* (2012). Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, **84**(1), 125–129. 150

ZHANG, T., LIANG, Z.W., HAN, J., BI, J.P., YANG, Z.Y., & MA, H. (2015). Double-arc volumetric modulated therapy improves dose distribution compared to static gantry IMRT and 3D conformal radiotherapy for adjuvant therapy of gastric cancer. *Radiat. Oncol.*, **10**(1), 114. 46, 55

ZHANG, X., LI, X., QUAN, E.M., PAN, X., & LI, Y. (2011). A methodology for automatic intensity-modulated radiation treatment planning for lung cancer. *Phys. Med. Biol.*, **56**(13), 3873–3893. 21, 131

ZHAO, N., YANG, R., JIANG, Y., TIAN, S., GUO, F., & WANG, J. (2015a). A hybrid IMRT/VMAT technique for the treatment of nasopharyngeal cancer. *BioMed. Res. Int.*, **2015**. 107

ZHAO, N., YANG, R., WANG, J., ZHANG, X., & LI, J. (2015b). An IMRT/VMAT technique for nonsmall cell lung cancer. *BioMed. Res. Int.*, **2015**. 107

ZHU, W.G., XUA, D.F., PU, J., ZONG, C.D., LI, T., TAO, G.Z. *et al.* (2012). A randomized, controlled, multicenter study comparing intensity-modulated radiotherapy plus concurrent chemotherapy with chemotherapy alone in gastric cancer patients with D2 resection. *Radiother. Oncol.*, **104**(3), 361–366. 45

List of Publications

Journal Articles

1. Bijman, R., Rossi, L., **Sharfo, A.W.M.**, Heemsbergen, W., Incrocci, L., Breedveld, S., & Heijmen, B.J.M. (**2020**). Automated Radiotherapy Planning for Patient-Specific Exploration of the Trade-Off Between Tumor Dose Coverage and Predicted Radiation-Induced Toxicity—A Proof of Principle Study for Prostate Cancer. *Frontiers in Oncology*, **10**:943.

2. Florijn, M.A., **Sharfo, A.W.M.**, Wiggenraad, R.G.J., van Santvoort, J.P.C., Petoukhova, A.L., Hoogeman, M.S., Mast, M.E., & Dirkx, M.L.P. (**2020**). Lower Doses to Hippocampi and Other Brain Structures for Skull-Base Meningiomas with Intensity Modulated Proton Therapy Compared to Photon Therapy. *Radiotherapy and Oncology*, **142**:147-153.

3. **Sharfo, A.W.M.**, Stieler, F., Kupfer, O., Heijmen, B.J.M., Dirkx, M.L.P., Breedveld, S., Wenz, F., Lohr, F., Boda-Heggemann, J., & Buergy, D. (**2018**). Automated VMAT Planning for Postoperative Adjuvant Treatment of Advanced Gastric Cancer. *Radiation Oncology*, **13**(1):74.

4. Rossi, L., **Sharfo, A.W.M.**, Aluwini, S., Dirkx, M.L.P., Breedveld, S., & Heijmen, B.J.M. (**2018**). First Fully Automated Planning Solution for Robotic Radiosurgery – Comparison with Automatically Planned Volumetric arc Therapy for Prostate Cancer. *Acta Oncologica*, **57**(11):1490-1498.

5. **Sharfo, A.W.M.**, Dirkx, M.L.P., Bijman, R.G., Schillemans, W., Breedveld, S., Aluwini, S., Pos, F., Incrocci, L., & Heijmen, B.J.M. (**2018**). Late Toxicity in the Randomized Multicenter HYPRO Trial for Prostate Cancer Analyzed with Automated Treatment Planning. *Radiotherapy and Oncology*, **128**(2):349-356.

6. Buschmann, M., **Sharfo, A.W.M.**, Penninkhof, J., Seppenwoolde, Y., Goldner, G., Georg, D., Breedveld, S., & Heijmen, B.J.M. (**2018**). Automated Volumetric

Modulated Arc Therapy Planning for Whole Pelvic Prostate Radiotherapy. *Strahlentherapie und Onkologie*, **194**(4):333-342.

7. Habraken, S.J., **Sharfo, A.W.M.**, Buijsen, J., Verbakel, W.F., Haasbeek, C.J., Öllers, M.C., Westerveld, H., van Wieringen, N., Reerink, O., Seravalli, E., Braam, P.M., Wendling, M., Lacornerie, T., Mirabel, X., Weytjens, R., Depuydt, L., Tanadini-Lang, S., Riesterer, O., Haustermans, K., Depuydt, T., Dwarkasing, R.S., Willemssen, F.E., Heijmen, B.J.M., & Méndez Romero, A. (**2017**). The TRENDY Multi-Center Randomized Trial on Hepatocellular Carcinoma – Trial QA Including Automated Treatment Planning and Benchmark-Case Results. *Radiotherapy & Oncology*, **125**(3):507-513.

8. Buergy, D., **Sharfo, A.W.M.**, Heijmen, B.J.M., Voet, P.W.J., Breedveld, S., Wenz, F., Lohr, F., & Stieler, F. (**2017**). Fully Automated Treatment Planning of Spinal Metastases – A Comparison to Manual Planning of Volumetric Modulated Arc Therapy for Conventionally Fractionated Irradiation. *Radiation Oncology*, **12**(1):33.

9. **Sharfo, A.W.M.**, Dirkx, M.L.P., Breedveld, S., Méndez Romero, A., & Heijmen, B.J.M. (**2017**). VMAT Plus a Few Computer-Optimized Non-Coplanar IMRT Beams (VMAT+) Tested for Liver SBRT. *Radiotherapy & Oncology*, **123**(1):49-56.

10. **Sharfo, A.W.M.**, Breedveld, S., Voet, P.W.J., Heijkoop, S.T., Mens, J.W.M., Hoogeman, M.S., & Heijmen, B.J.M. (**2016**). Validation of Fully Automated VMAT Plan Generation for Library-Based Plan-of-the-Day Cervical Cancer Radiotherapy. *PLoS One*, **11**(12):e0169202.

11. Heijkoop, S.T., Westerveld, H., Bijker, N., Feije, R., **Sharfo, A.W.M.**, van Wieringen, N., Mens, J.W.M., Stalpers, L.J., & Hoogeman, M.S. (**2016**). Optimal Patient Positioning (Prone Versus Supine) for VMAT in Gynecologic Cancer: A Dosimetric Study on the Effect of Different Margins. *International Journal of Radiation Oncology*Biology*Physics*, **96**(2):432-439.

12. van de Sande, M.A., Creutzberg, C.L., van de Water, S., **Sharfo, A.W.M.**, & Hoogeman, M.S. (**2016**). Which Cervical and Endometrial Cancer Patients Will Benefit Most from Intensity-Modulated Proton Therapy? *Radiotherapy and Oncology*, **120**(3):397-403.

13. **Sharfo, A.W.M.**, Voet, P.W.J., Breedveld, S., Mens, J.W.M., Hoogeman, M.S., & Heijmen, B.J.M. (**2015**). Comparison of VMAT and IMRT Strategies for Cervical Cancer Patients Using Automated Planning. *Radiotherapy & Oncology*, **114**(3):395-401.

Summary

Erasmus-iCycle/Monaco has been developed and introduced at our institute for fully automated multi-criterial treatment planning. This thesis focuses on *1)* assessment of the clinical advantages of automated treatment planning with Erasmus-iCycle/Monaco compared to the conventional manual treatment planning (Chapters 2-5), *2)* use of automated treatment planning in decision-making on superiority of one treatment approach/technique over another (Chapters 6-9), and *3)* evaluation of the use of automated treatment planning as a quality assurance tool in multi-institutional clinical trials (Chapters 10 and 11).

In Chapter 2, we describe a validation study on the use of Erasmus-iCycle/Monaco for fully automated generation of VMAT plan-libraries for plan-of-the-day (PotD) adaptive radiotherapy in locally-advanced cervical cancer. For 34 cervical cancer patients (24 '*non-movers*' with, apart from the back-up plan, one IMRT plan, and 10 '*movers*' with a back-up plan and two IMRT plans) automatically generated dual-arc VMAT plans (autoVMAT) were benchmarked against: *1)* dual-arc VMAT plans generated manually by an expert dosimetrist in absence of time pressure (manVMAT), *2)* manually generated, clinically delivered 9-beam static IMRT plans (CLINICAL), and *3)* automatically generated, fixed-beam IMRT-DMLC plans with 20 equi-angular beams (autoIMRT). All treatment plans were visually inspected for clinical acceptability and compared regarding differences in PTV coverage and doses in OARs. Furthermore, treatment delivery times of the automatically generated VMAT and IMRT plans were compared. All generated plans had a highly similar PTV coverage ($V_{95\%} > 99.5\%$) and obeyed the clinical constraints. After plan normalization for exactly equal median PTV doses in corresponding plans, all evaluated OAR parameters in the autoVMAT plans were on average lower than in the CLINICAL and manVMAT plans, showing average reductions in small bowel V_{45Gy} (the highest prioritized OAR planning objective) of 34.6% ($p < .001$) and 30.3% ($p < .001$), respectively. AutoIMRT reduced the small bowel V_{45Gy} by another 2.7% compared to autoVMAT, at the cost of a prolonged delivery time. Compared to manual planning, autoVMAT plans showed an enhanced positive impact for small bowel sparing for empty bladder PTVs compared to full bladder PTVs, due to differences in concavity of the PTVs.

Based on the results of this study, automated VMAT plan generation completely replaced the previously applied 3D-CRT and IMRT techniques for treatment of cervical cancer patients. This did not only result in a significant improvement in plan quality but also in a major reduction in treatment delivery time from 12 minutes for IMRT to less than 5 minutes for VMAT, which limits the impact of intra-fraction bladder filling and patient motion. Due to the achieved workload reduction with automated planning, extension of plan-libraries for PotD adaptive radiotherapy has become feasible.

In Chapter 3, automated VMAT planning for spinal metastases was compared to manual planning by expert dosimetrists in collaboration with the University Hospital of Mannheim. Plans for 42 targets (22 at the level of the kidneys, 17 at the level of the lungs, 3 on both levels) were analyzed. Plans were compared with respect to target coverage, OARs sparing and treatment time. All Erasmus-iCycle/Monaco plans were considered clinically acceptable. Compared to manually generated plans, the PTV coverage and OAR sparing were significantly improved at the expense of minimally increased treatment delivery times (7% on average).

For postoperative adjuvant treatment of advanced gastric cancer, automatically generated plans were compared to plans generated by expert dosimetrists in the University Hospital of Mannheim (Chapter 4). For 20 patients, autoVMAT and manVMAT plans were compared based on dose-volume histograms and predicted normal tissue complication probability (NTCP) analyses. All plans were considered clinically acceptable and had similar target coverage. With autoVMAT, substantial improvements in OAR sparing, dose conformity and integral-dose were achieved. This resulted in statistically significant reductions in predicted NTCPs for the left and right kidney and the liver by on average 11.3%, 12.8% and 7%, respectively ($p \leq$ .001). AutoVMAT plans had a higher modulation degree which resulted in a prolonged treatment delivery time compared to manVMAT of 19 seconds ($p = .006$).

In Chapter 5, automatically generated VMAT plans for whole-pelvic prostate radiotherapy (WPRT) were compared to plans that were manually generated in the University Hospital of Vienna by expert dosimetrists in the absence of time constraints. Using a cohort of 5 training patients, the Erasmus-iCycle/Monaco system was configured for the generation of autoVMAT plans for treatment with a sequential-boost technique, and for a simultaneous-integrated-boost technique. The automated work-flow was then evaluated for another cohort of 30 randomly selected patients by comparison with the manually generated plans. For 27 of the 30 patients, the physician preferred the autoVMAT plan over the manVMAT plan due to substantial improvements in OAR sparing and dose conformity. On average, mean doses in rectum, bladder and small bowel were reduced by 4.5 ± 2.2 Gy, 10.7 ± 3.2 Gy and 2.1 ± 1.2 Gy, respectively (all $p <$ .001). In addition, the planning workload per patient was reduced by > 70 minutes.

Different treatment delivery strategies for locally-advanced cervical cancer patients were compared in Chapter 6 using automated treatment planning. For 10 patients, fully automated plan generation with Erasmus-iCycle was used to assist an expert planner in generating five clinically deliverable plans with the Monaco® TPS: single- and dual-arc VMAT plans, and IMRT plans with 9, 12, and 20 equi-angular beams. For each patient, all plans were clinically acceptable with similar PTV coverage. OAR sparing increased when going from 9 to 12 to 20 IMRT beams, and from single- to dual-arc VMAT. 12 and 20 beam IMRT plans were superior to both single- and dual-arc VMAT, with substantial variations in gain in OAR sparing among the study patients. On the other hand, delivery of VMAT plans was significantly faster than delivery of IMRT plans. This complex trade-off between plan quality and treatment delivery is especially interesting when fast multi-leaf collimators are available.

In Chapter 7, a novel fast treatment delivery approach, VMAT+, combining VMAT with a few computer-optimized non-coplanar IMRT beams, was proposed and evaluated for liver SBRT. For 15 patients with liver metastases, VMAT+ was investigated as an alternative for *i)* dual-arc coplanar VMAT and *ii)* 25-beam non-coplanar treatment with computer-optimized beam orientations (25-NCP). All plans were generated fully automatically using Erasmus-iCycle/Monaco for delivery of the highest feasible Biologically Effective Dose (BED). OAR doses, dose spillage, dose compactness, and measured delivery times were evaluated. For 25-NCP and VMAT+, the maximum achievable tumor BED was equal. Improvements in OAR sparing with 25-NCP were generally modest and/or statistically insignificant, while treatment delivery times were on average 20.5 minutes longer. VMAT resulted in a lower maximum achievable tumor BED in 5 patients. Compared to VMAT, VMAT+ significantly improved OAR sparing, dose compactness and intermediate-dose spillage, at the cost of a modest increase in treatment delivery time (12.5 ± 3.8 min instead of 8.4 ± 2.3 min).

In Chapter 8, we used automated treatment planning with integrated multi-criterial beam profile and beam angle optimization (BAO) to derive a two-beam non-coplanar *class solution* (CS) to supplement coplanar VMAT for prostate SBRT. For 20 patients, Erasmus-iCycle was first used to fully automatically generate VMAT plans supplemented with five patient-specific computer-optimized non-coplanar IMRT beams (VMAT+5). Based on an analysis of the most preferred non-coplanar beam directions in VMAT+5 of the 20 patients, a class solution was derived to be applied for each patient. The VMAT+CS plans were then compared to automatically generated *i)* dual-arc coplanar VMAT plans, *ii)* VMAT+5 plans, and *iii)* 30-beam non-coplanar IMRT plans with computer-optimized beam orientations (30-NCP). Differences in PTV coverage, OAR sparing, computation time and treatment delivery time were quantified. Compared to VMAT, VMAT+CS significantly reduced doses to rectum, bladder and the dose-bath. Mean relative differences in rectum D_{mean}, D_{1cc}, V_{40GyEq} and V_{60GyEq} were 19.4 ± 10.6%, 4.2 ± 2.7%,

34.9 ± 20.3%, and 39.7 ± 23.2%, respectively (all $p < .001$). Moreover, VMAT+CS required 3% less MU on average, with a minor increase in total delivery time (11.0 ± 0.3 min instead of 9.1 ± 0.7 min). The quality of VMAT+CS plans was similar to VMAT+5 plans, while optimization times were reduced by a factor of 25 due to the avoidance of BAO. Compared to 30-NCP, VMAT+CS was considered clinically equivalent with respect to high doses in the rectum and bladder and much faster to deliver, while 30-NCP had a more favorable dose bath but with much enhanced optimization and delivery times.

In Chapter 9, a fully automated treatment planning solution for robotic radiosurgery was developed and evaluated for prostate SBRT. For this purpose, Erasmus-iCycle was coupled to MultiPlan®, the TPS of the CyberKnife® system, to fully automatically generate clinically deliverable non-coplanar plans (autoCK). The system was first validated for 10 prostate SBRT patients by comparing autoCK plans to manually generated, clinically delivered plans. Additionally, for 20 patients autoCK plans were compared to automatically generated, clinically deliverable autoVMAT plans with CTV-PTV margins of 3 mm (as used for autoCK) and 5 mm (as clinically feasible on a standard linac). The same wish-list was used for optimization of CyberKnife® and VMAT plans to maximally avoid bias in plan generation. Manual and autoCK plans had equal PTV coverage and fulfilled clinical requirements. With autoCK, sparing of the OARs was significantly improved compared to manual plans with reductions in rectum V_{60GyEq} and D_{mean} of 75% and 41%, ($p \leq .002$), respectively. AutoCK and autoVMAT plans with 3 mm CTV-PTV margin had similar PTV coverage, but the rectum dose was significantly lower in the autoCK plans and the low dose bath was clearly reduced. For 18 of the 20 autoVMAT plans with 5 mm CTV-PTV margin,the PTV coverage was lower than requested (< 95%), while the rectum doses were significantly higher. Clinician scoring underlined overall plan quality superiority of autoCK plans. In case of a clinically acceptable VMAT plan, the reduced treatment time could outweigh plan quality advantages of CyberKnife® treatment.

The extent to which automated treatment planning could have reduced OAR dose delivery as observed in the randomized multi-institutional HYPRO trial, comparing conventional and hypo-fractionated external beam radiotherapy for prostate cancer, was investigated in Chapter 10. For 725 patients, differences in plan quality between automatically generated VMAT plans (autoVMAT) and the manually generated, clinically delivered plans (CLINICAL) were quantified using DVH parameters and NTCP models for late gastrointestinal toxicity, derived from reported toxicities in the HYPRO trial. Compared to CLINICAL, autoVMAT plans had similar or higher PTV coverage, while large and statistically significant OAR sparing was achieved. Mean doses in the rectum, anus and bladder were reduced by 7.8 ± 4.7 Gy, 7.9 ± 6.0 Gy and 4.2 ± 2.9 Gy, respectively ($p < .001$). The reduced rectum dose delivery resulted in significant and meaningful reductions in NTCPs for late gastrointestinal toxicity grade ≥ 2, rectal bleeding and stool incontinence. With autoVMAT, the predicted toxicity for

hypo-fractionation could be reduced to levels that were clinically observed (and accepted) for conventional-fractionation.

In Chapter 11, we investigated the use of automated treatment planning for prospective feedback on plan quality in a multi-institutional randomized trial on hepatocellular carcinoma. A total of 20 plans from participating centers, including 12 from a benchmark case, 5 from a clinical pilot, and 3 from the first patients enrolled in the trial, were compared to automatically generated VMAT plans (autoVMAT). Four protocol deviations were detected in the set of 20 treatment plans, related to violation of the NTCP constraint imposed on the healthy liver. Overall, compared to the submitted plans, autoVMAT resulted in an average liver NTCP reduction of 2.2% (range: 16.2%; -1.8%, $p = .03$), and lower-doses to the healthy liver ($p < .01$) and gastrointestinal organs ($p < .001$). Additionally, the potential for dose escalation with autoVMAT was investigated. For 5 of the 7 patients who were clinically treated to 48 Gy, dose escalation to 54 Gy was achieved with autoVMAT within planning constraints for the OARs. Currently, an autoVMAT plan is generated prior to the start of treatment of every new trial patient as benchmark for the treatment plan generated in the participating institute (prospective plan QA). Based on a comparison of dose distributions, the participating center can continue with the initially generated plan, or use the autoVMAT plan to further improve their own plan.

In Chapter 12, automated treatment planning is discussed in a broader context.

Samenvatting

Erasmus-iCycle / Monaco is ontwikkeld en geïntroduceerd in ons instituut voor het volledig geautomatiseerd genereren van behandelplannen op basis van multi-criteria optimalisatie. Dit proefschrift richt zich op *1)* de evaluatie van klinische voordelen van het geautomatiseerd berekenen van behandelplannen met Erasmus-iCycle / Monaco in vergelijking met de conventionele aanpak van handmatig plannen (hoofdstuk 2-5), *2)* het gebruik van geautomatiseerd plannen in de besluitvorming over superioriteit van een behandelingsaanpak / -techniek ten opzichte van een andere (hoofdstuk 6-9), en *3)* evaluatie van het gebruik van geautomatiseerd plannen als hulpmiddel voor kwaliteitsborging in multi-institutioneel klinisch onderzoek (hoofdstuk 10 en 11).

In hoofdstuk 2 beschreven we een validatieonderzoek naar het gebruik van Erasmus-iCycle / Monaco voor het volledig geautomatiseerd maken van volumetrisch gemoduleerde rotatietherapie (VMAT) plan-bibliotheken voor plan-van-de-dag (PotD) -adaptieve radiotherapie bij lokaal uitgebreide baarmoederhalskanker. Voor 34 patiënten met baarmoederhalskanker (24 '*niet-bewegers*' met, afgezien van het back-up plan, een intensiteitsgemoduleerd radiotherapie (IMRT) plan en 10 '*bewegers*' met een back-up plan en twee IMRT plannen) werden automatisch gegenereerde VMAT plannen met een dubbele rotatie (autoVMAT) vergeleken met: *1)* VMAT plannen met een dubbele rotatie die handmatig werden gegenereerd door een ervaren planner zonder tijdsdruk (manVMAT), *2)* handmatig gegenereerde, klinisch toegepaste statische IMRT plannen met 9-bundelrichtingen (KLINISCH) en *3)* automatisch gegenereerde IMRT-plannen gebruikmakend van dynamische multileaf segmentatie (DMLC) met een vaste bundelopzet van 20 equidistante bundelhoeken (autoIMRT). Alle behandelplannen werden visueel beoordeeld op klinische aanvaardbaarheid en vergeleken op basis van verschillen in dekking van het 'planning-target volume' (PTV) en doses in de risico-organen ('organs-at-risk', OARs). Bovendien werden de behandeltijden van de automatisch gegenereerde VMAT- en IMRT plannen vergeleken. Alle gegenereerde plannen hadden een zeer vergelijkbare PTV-dekking ($V_{95\%} > 99,5\%$) en voldeden aan de klinische eisen. Nadat de plannen waren genormeerd op exact gelijke mediane PTV-doses bleken alle geëvalueerde OAR-parameters in de autoVMAT plannen

gemiddeld lager te zijn dan in de klinische en manVMAT plannen, met een gemiddelde afname in de dunne darm V_{45Gy} (OAR planningsdoel met de hoogste prioriteit) van 34,6% ($p < 0,001$) en 30,3% ($p < 0,001$), respectievelijk. Wanneer autoIMRT werd vergeleken met autoVMAT was de dunne darm V_{45Gy} nog eens 2,7% lager maar afgifte van deze plannen vergde een veel langere behandeltijd. Vergeleken met handmatige dosisplanning werd er voor autoVMAT plannen meer dunne darm sparing gezien voor lege blaas PTVs dan voor volle blaas PTVs, door verschillen in de concaviteit van de PTVs. Op basis van de resultaten van deze studie heeft geautomatiseerde VMAT planning de eerder toegepaste 3D-CRT (3D conformatie radiotherapie) en IMRT-techniek voor de behandeling van patiënten met baarmoederhalskanker volledig vervangen. Dit resulteerde niet alleen in een significante verbetering van de kwaliteit van de behandelplannen, maar ook in een aanzienlijke afname van de behandeltijd van 12 minuten voor IMRT tot minder dan 5 minuten voor VMAT, waardoor de invloed van intrafractionele blaasvulling en beweging van de patiënt wordt beperkt. Vanwege de gerealiseerde werklastvermindering met geautomatiseerde dosisplanning is uitbreiding van plan-bibliotheken voor PotD adaptieve radiotherapie nu haalbaar geworden.

In hoofdstuk 3 werd geautomatiseerde VMAT-planning voor spinale metastasen vergeleken met handmatige dosisplanning door ervaren planners in samenwerking met het Academisch Ziekenhuis van Mannheim. Plannen voor 42 doelgebieden (22 op het niveau van de nieren, 17 op het niveau van de longen, 3 op beide niveaus) werden geanalyseerd en vergeleken op basis van PTV dekking, afname van OAR doses en behandeltijd. Alle Erasmus-iCycle / Monaco plannen werden als klinisch aanvaardbaar beoordeeld. In vergelijking met de handmatige plannen waren de PTV dekking en OAR-sparing aanzienlijk verbeterd, ten koste van een minimale toename in behandeltijd (gemiddeld met 7%).

Voor postoperatieve adjuvante behandeling van vergevorderd maagkanker werden automatisch gegenereerde plannen vergeleken met plannen gegenereerd door ervaren planners in het Academisch Ziekenhuis van Mannheim (hoofdstuk 4). Voor 20 patiënten werden de opties autoVMAT en manVMAT vergeleken op basis van dosisvolume histogrammen (DVH) en de voorspelde kans op normale weefselcomplicatie (NTCP). Alle plannen werden als klinisch aanvaardbaar beschouwd en hadden een vergelijkbare dekking van het doelgebied. Met autoVMAT werden substantiële verbeteringen in OAR-sparing, dosisconformiteit en integrale dosis bereikt. Dit resulteerde in een statistisch significante afname van de voorspelde NTCPs voor de linker- en rechternier en de lever met respectievelijk gemiddeld 11,3%, 12,8% en 7% ($p \le 0,001$). AutoVMAT plannen hadden een hogere modulatiegraad, wat resulteerde in een 19 seconde ($p = 0,006$) langere behandeltijd dan voor manVMAT.

In hoofdstuk 5 werden automatisch gegenereerde VMAT plannen voor prostaatpatiënten die ook op lymfklieren in het bekkengebied bestraald worden

vergeleken met plannen die handmatig werden gemaakt door ervaren planners in het Academisch Ziekenhuis van Wenen zonder tijdsbeperkingen. Aan de hand van een groep van 5 trainingspatiënten werd het Erasmus-iCycle / Monaco-systeem geconfigureerd voor het genereren van autoVMAT plannen voor behandeling met een sequentiële boost techniek, en voor een simultaan geïntegreerde boost techniek. De geautomatiseerde workflow werd vervolgens geëvalueerd voor een andere groep van 30 willekeurig geselecteerde patiënten en vergeleken met de handmatig gegenereerde plannen. Voor 27 van de 30 patiënten gaf de arts de voorkeur aan het autoVMAT plan boven het manVMAT plan vanwege aanzienlijke verbeteringen in de OAR-sparing en dosisconformiteit. Gemiddeld over alle patiënten waren de gemiddelde doses in de endeldarm, de blaas en de dunne darm verminderd met respectievelijk 4,5 ± 2,2 Gy, 10,7 ± 3,2 Gy en 2,1 ± 1,2 Gy (alle $p < 0,001$). Bovendien werd de werklast per patiënt voor dosisplanning met meer dan 70 minuten verminderd.

Verschillende behandelstrategieën voor patiënten met lokaal uitgebreide baarmoederhalskanker werden op basis van geautomatiseerde dosisplanning vergeleken in hoofdstuk 6. Voor 10 patiënten werden op basis van volledig geautomatiseerde dosisplanning met Erasmus-iCycle gebruikt door een ervaren planner vijf klinisch uitvoerbare behandelplannen gemaakt met het planningsysteem Monaco®: VMAT plannen met een enkele en dubbele rotatie en IMRT plannen met 9, 12 en 20 equidistante bundels. Voor elke patiënt bleken alle plannen klinisch acceptabel te zijn en was de PTV-dekking vergelijkbaar. De OAR sparing verbeterde wanneer het aantal IMRT bundels toenam van 9 naar 12 naar 20, en voor een VMAT plannen met een dubbele in plaats van een enkele rotatie. IMRT plannen met 12 en 20 bundels waren superieur ten opzichte van VMAT plannen met een enkele en met een dubbele rotatie, waarbij de mate van verbetering in OAR-sparing sterk varieerde tussen de studiepatiënten. Aan de andere kant was de afgifte van VMAT plannen aanzienlijk sneller dan die van IMRT plannen. Deze complexe afweging tussen plankwaliteit en behandeltijd is vooral interessant als er snelle multileaf collimatoren beschikbaar zijn.

In hoofdstuk 7 werd een nieuwe, snelle behandeltechniek, VMAT+, waarbij VMAT wordt gecombineerd met een aantal computer-geoptimaliseerde niet-coplanaire IMRT bundels, voorgesteld en geëvalueerd voor stereotactische bestraling (SBRT) van gemetastaseerde levertumoren. Voor 15 patiënten werd VMAT+ onderzocht als alternatief voor *i)* coplanaire VMAT met een dubbele rotatie en *ii)* niet-coplanaire bestraling met 25 bundels met computer-geoptimaliseerde richtingen (25-NCP). Alle plannen werden volledig automatisch gegenereerd met behulp van Erasmus-iCycle / Monaco met als doel om een zo'n hoog mogelijke biologisch effectieve dosis (BED) te bereiken. OAR-doses, dosis buiten het doelgebied, dosis compactheid en gemeten behandeltijden werden vergeleken. Voor 25-NCP en VMAT+ was de maximaal haalbare tumor-BED identiek. Verbeteringen in OAR-sparing met 25-NCP waren over het

algemeen beperkt en/of statistisch niet significant, terwijl de behandeltijden gemiddeld 20,5 minuten langer waren. VMAT resulteerde bij 5 patiënten in een lagere maximaal haalbare tumor-BED. In vergelijking met VMAT gaf VMAT+ een aanzienlijk verbeterde OAR-sparing, dosis compactheid en minder dosis buiten het doelgebied, ten koste van een beperkte toename in behandeltijd (12,5 ± 3,8 min in plaats van 8,4 ± 2,3 min).

In hoofdstuk 8 werd geautomatiseerde dosisplanning met geïntegreerde bundelhoekoptimalisatie gebruikt om een standaardoplossing ('*class-solution*': CS), bestaande uit twee niet-coplanaire IMRT bundels als aanvulling van coplanaire VMAT, te ontwikkelen voor prostaat SBRT. Voor 20 patiënten werd Erasmus-iCycle eerst gebruikt om volledig automatisch VMAT plannen te genereren die waren aangevuld met vijf patiënt-specifieke, computer-geoptimaliseerde, niet-coplanaire IMRT bundels (VMAT+5). Op basis van de meest geselecteerde niet-coplanaire bundelrichtingen in deze 20 patiëntplannen werd een standaardoplossing met twee vaste niet-coplanaire bundels afgeleid die vervolgens voor elke patiënt werd toegepast. De VMAT+CS plannen werden vergeleken met *i)* coplanaire VMAT plannen met twee rotaties, *ii)* VMAT+5 plannen en *iii)* niet-coplanaire IMRT plannen met 30 bundels met computer-geoptimaliseerde bundelhoeken (30-NCP). Verschillen in PTV-dekking, OAR-sparing, rekentijd en behandeltijd werden gekwantificeerd. Wanneer VMAT+CS werd vergeleken met VMAT waren de doses in de endeldarm en de blaas, en het dosisbad significant lager. Voor de endeldarm waren de verschillen in gemiddelde dosis, D_{1cc}, V_{40GyEq} en V_{60GyEq} gemiddeld 19,4 ± 10,6%, 4,2 ± 2,7%, 34,9 ± 20,3%, en 39,7 ± 23,2% lager (alle $p < 0,001$). Bovendien vereiste VMAT+CS gemiddeld 3% minder MU, met een kleine toename in de totale behandeltijd (11,0 ± 0,3 min in plaats van 9,1 ± 0,7 min). Over het algemeen was de kwaliteit van de VMAT+CS plannen vergelijkbaar met de VMAT+5 plannen, terwijl de rekentijden een factor 25 korter waren omdat hoekoptimalisatie hierbij overbodig was. Vergeleken met 30-NCP werd VMAT+CS als klinisch equivalent beoordeeld met betrekking tot hoge doses in de endeldarm en de blaas. Weliswaar was het dosisbad gunstiger voor 30-NCP, maar de optimalisatie en afgifte van deze plannen was veel tijdsintensiever.

In hoofdstuk 9 werd een volledig geautomatiseerde dosisplanningsmethode voor robot-radiochirurgie beschreven en geëvalueerd voor prostaat SBRT. Voor dit doel werd Erasmus-iCycle gekoppeld aan MultiPlan®, het planningsysteem van het CyberKnife®, om volledig automatisch klinisch uitvoerbare niet-coplanaire plannen (autoCK) te genereren. Het systeem werd eerst gevalideerd voor 10 prostaat SBRT patiënten, door autoCK plannen te vergelijken met handmatig gegenereerde, klinisch gebruikte plannen. Vervolgens werden autoCK plannen voor 20 patiënten vergeleken met automatisch gegenereerde, klinisch uitvoerbare autoVMAT plannen met CTV-PTV marges van 3 mm (zoals gebruikt voor autoCK) en 5 mm (zoals klinisch uit te voeren op een standaard bestralingstoestel). Voor de optimalisatie van het CyberKnife® plan en het VMAT plan

werd dezelfde iCycle 'wishlist' gebruikt, om bias bij het genereren van plannen zo goed mogelijk te vermijden. Handmatige en automatische gegenereerde CK plannen hadden een gelijke PTV-dekking en voldeen beide aan de klinische eisen. Met autoCK werden de OARs aanzienlijk beter gespaard in vergelijking met de handmatige gegenereerde plannen, met een afname in de V_{60GyEq} en gemiddelde dosis in de endeldarm van respectievelijk 75% en 41%, ($p \leq 0{,}002$). AutoCK en autoVMAT plannen met een 3 mm CTV-PTV-marge hadden een vergelijkbare PTV-dekking, maar de dosis in de endeldarm was aanzienlijk lager in de autoCK plannen en het bad met lage dosis was duidelijk kleiner. Voor 18 van de 20 autoVMAT plannen met 5 mm CTV-PTV marge was de PTV dekking lager dan gewenst (< 95%), terwijl de dosis in de endeldarm aanzienlijk hoger was. Artsen beoordeelden de plankwaliteit van de autoCK plannen over het algemeen als superieur. In geval van een klinisch acceptabel VMAT plan kan de verkorte behandeltijd opwegen tegen de voordelen van de CyberKnife-behandeling.

In hoofdstuk 10 werd onderzocht in welke mate de OAR-dosisafgifte die werd waargenomen in de gerandomiseerde, multi-institutionele HYPRO-studie, waarin conventionele en hypo -gefractioneerde externe radiotherapie voor prostaatkanker werden vergeleken, kon worden verlaagd met geautomatiseerde dosisplanning. Voor 725 patiënten werden de verschillen in plankwaliteit tussen automatisch gegenereerde VMAT plannen (autoVMAT) en de handmatig gegenereerde, klinisch toegepaste plannen (CLINICAL) gekwantificeerd aan de hand van DVH-parameters en NTCP-modellen voor late gastro-intestinale toxiciteit die waren gebaseerd op de gerapporteerde toxiciteit in de HYPRO-studie. In vergelijking met CLINICAL hadden de autoVMAT plannen een vergelijkbare of hogere PTV-dekking, terwijl grote en statistisch significante verbetering in OAR-sparing werd bereikt. De gemiddelde doses in de endeldarm, de anus en de blaas werden verlaagd met respectievelijk 7,8 ± 4,7 Gy, 7,9 ± 6,0 Gy en 4,2 ± 2,9 Gy ($p < 0{,}001$). De verminderde dosis in de endeldarm resulteerde in een significante en klinisch relevante afname in de NTCPs voor late gastro-intestinale toxiciteit graad ≥ 2, rectale bloeding en incontinentie. Met autoVMAT kon de voorspelde toxiciteit voor hypofractionering worden verlaagd tot niveaus die klinisch werden waargenomen (en geaccepteerd) voor de conventionele fractionering.

In hoofdstuk 11 hebben we het gebruik van geautomatiseerde dosisplanning onderzocht om prospectief terugkoppeling te geven over plankwaliteit in een multi-institutioneel gerandomiseerd onderzoek naar hepatocellulair carcinoom. 20 dosisplannen van deelnemende centra (12 plannen van een benchmark casus, 5 klinische testplannen en 3 plannen voor de eerste patiënten die deelnamen aan de studie) werden vergeleken met automatisch gegenereerde VMAT plannen (autoVMAT). Voor deze 20 plannen werden vier afwijkingen van het studieprotocol gevonden die betrekking hadden op een overschrijding van de maximale NTCP-waarde die werd toegestaan voor de gezonde lever. Ten opzichte van de ingediende plannen was de gemiddelde NTCP voor

de lever in de autoVMAT plannen 2,2% lager (bereik: 16,2%, -1,8%, $p = 0,03$) en werd er ook minder dosis afgegeven in de gezonde lever ($p < 0,01$) en in de gastro-intestinale organen ($p < 0,001$). Tevens werd de mogelijkheid van dosisescalatie met autoVMAT onderzocht. Voor 5 van de 7 patiënten die klinisch werden behandeld met 48 Gy kon met autoVMAT een dosisverhoging naar 54 Gy worden bereikt, binnen de planningseisen voor de OARs. Momenteel wordt voorafgaand aan de start van de behandeling voor elke nieuwe studiepatiënt een autoVMAT plan gegenereerd als benchmark voor het behandelplan dat in het deelnemende instituut wordt gemaakt (prospectieve plan QA). Op basis van een vergelijking van de dosisverdelingen kan het deelnemende centrum besluiten om door te gaan met het initieel gegenereerde plan, of kan het autoVMAT plan worden gebruikt om het eigen plan verder te verbeteren.

In hoofdstuk 12 wordt geautomatiseerde dosisplanning besproken in een bredere context.

Acknowledgements

The genesis of the work described in this thesis was accompanied by the support of incredibly helpful colleagues and friends, whom I would like to acknowledge here.

First and foremost, I would like to give my sincerest and everlasting thanks to my promoter Prof. dr. Ben Heijmen for giving me the opportunity to conduct this research at the Erasmus MC. Dear Ben, I cannot thank you enough for lending me your constant support, encouragement, motivation and invaluable guidance during this journey. You have always been very helpful by pushing my projects in the right direction to meet high scientific standards, while providing the trust and freedom necessary to accomplish real progress. You and your wife Alessandra have spared no effort to make me and my wife feel home in the Netherlands. I personally enjoyed exchanging views on every aspect during the mutual visits.

I offer my utmost and heart-felt thanks and respect to my co-promoter dr.ir. Maarten Dirkx. Dear Maarten, your wisdom, guidance, and constructive notes on my projects were essential to the success of my work and were very much appreciated. You have helped me grow professionally and personally.

I express my sincere gratitude to Prof. dr. R.A. Nout, Prof. dr. J.J. Sonke, and Prof. dr. D. Verellen for serving on my PhD defense committee.

I am indebted to Peter Voet for providing me with the ABC of treatment planning. Dear Peter, all this fun has started with you! You paved the way to me and made treatment planning my hobby. You have played a significant part in the completion of my thesis by translating its summary into Dutch. Your friendship, compassion and support have helped me through many challenges during my life in the Netherlands…

I am extremely obliged to Sebastiaan Breedveld, who as the founder of Erasmus-iCycle has significantly contributed not only to this work but also to the treatment process at our institute. I also enjoyed sharing views and experiences on life with you.

Peter and Sebastiaan, I am really glad that both of you agreed to become my paranymphs. With you the journey began, and with you it ended...

I am thankful to Linda Rossi for her contribution which resulted two chapters in this thesis.

I am very grateful to Rik Bijman for his immense help in statistical analysis by creating the NTCP models, thus giving clinical meaning to mere numbers and credibility to the HYPRO article.

I am immensely grateful to Joan Penninkhof who guided and helped me throughout the duration of the clinical project and the collaboration with Vienna.

Special thanks appertain to Prof. Frank Lohr, Daniel Buergy, Florian Stieler and their colleagues from the Medical Faculty in Mannheim, Germany. As well as Martin Buschmann, Yvette Seppenwoolde and their colleagues from the Medical University of Vienna, Austria for the fruitful collaboration which yielded three chapters in this thesis.

I am also grateful to my former and current colleagues in the Medical Physics group at Erasmus MC for a rewarding research environment. I particularly enjoyed working with Sabrina Heijkoop, Steven Habraken, Mireille Florijn, and Mariska van de Sande. I also owe a gratitude to Vincent Riemersma, and Hafid Akhiat with whom I had the pleasure to interact over work and personal matters.

I profusely thank my colleagues, Andras Zolnay, Xander van Doorn, Wilco Schillemans, and Erik-Jan Tromp for providing help and solutions related to Matterhorn. I am also very grateful to Hafid Akhiat, Ruud Cools, Martijn Hol, Bart Kanis and Erik Loeff for unflinchingly helping me on countless occasions performing my measurements on the treatment units.

I would like to single out a special thank to Steven Petit and Azad Shikh Murad for showing support and help when needed.

I would like to show my gratitude to Mischa Hoogeman, Sandra de Wringer and my colleagues in the TT- and KF-Group for giving me the opportunity to expand my clinical experience and grow in my work.

I thank my 'former' roommates Peter Voet, Mariska van de Sande, Tine Arts, Rik Bijman and Mireille Florijn for contributing to positive work atmosphere in our office.

I owe a big shout-out to my ESTRO party buddies, especially, Yibing Wang and Thyrza Jagt. With you I had amazing memories!

I would also like to thank Jacqueline van der Valk and Jolanda Jacobs for helping me and showing support when needed.

Finally, though they should come first in the list! With warmhearted gratefulness I would like to thank my parents, my big family and friends in Syria for their unconditional love and encouragement. Sorry dad, I couldn't make it on time, you left us early, but I know you are there with me.

And to my wife Ranim for all her love, dedication and support through many challenges during this work. I am truly blessed to have you and our son Jude in my life. Jude, I wanted to make your birthday-anniversary extra special!

أسـرتي، أصدقائي، أحبّتي ... شـكراً من القلب على كلّ الـدّعم و التشـجيع الذي منحتموني إيّاه، أتمنّى أن أكون قد مثّلتكم ومثّلت بلدي الأول سـوريا خير تمثيل...

PhD Portfolio

Name PhD Student: Abdul Wahab M. Sharfo   PhD Period: 2013 – 2018
Erasmus MC Department: Radiation Oncology   Promotor: Prof. Dr. Ben J.M. Heijmen
Research School: Molecular Medicine   Co-Promotor: Dr. Ir. Maarten L.P. Dirkx

1. PhD Training

General Courses

Workshop on Photoshop & Illustrator CS6 for PhD-students & other researchers	2013
Workshop on InDesign CS6 for PhD-students & other researchers	2013
Scientific Integrity	2014
Workshop on Presenting Skills for junior researchers	2016
Basic Introduction Course on SPSS (Statistics)	2016
Biomedical English Writing Course for MSc & PhD-students	2016
CPO course Patient Oriented Research: Design, conduct and analysis	2017
Research management for PhD-students	2018

Specific Courses

ESTRO Course: Image Guided Radiotherapy in Clinical Practice	2017
ESTRO Course: Radiobiological Basis of Clinical Particle Therapy	2018

Seminars and Workshops

2nd Daniel den Hoed Day, Erasmus MC	2013
PhD-Day, Erasmus MC	2014, 2016
Symposium the Science in Transition Movement, Erasmus MC	2014
R&D Physics meeting, department of Radiation Oncology	2013 - 2019
Refereeravond, department of Radiation Oncology	2013 - 2019
Research Round, department of Radiation Oncology	2015 - 2018
Research Day, department of Radiation Oncology	2015 - 2019
Journal Club, department of Radiation Oncology	2013 - 2019

International Conferences

ESTRO 33, Vienna, Austria (Poster Presentation)	2014
3rd ESTRO Forum, Barcelona, Spain (Poster Presentation)	2015
ESTRO 35, Turin, Italy (Oral Presentation)	2016
ESTRO 36, Vienna, Austria (Oral Presentation)	2017
ESTRO 37, Barcelona, Spain (Oral Presentation, 2 Poster Presentations)	2018
ESTRO 38, Milan, Italy (Oral Presentation)	2019

National Conferences

NVKF conferentie, Zeist (Oral Presentation)	2014, 2017, 2019
Wetenschappelijke Kringdag Klinische Fysica Radiotherapie, Utrecht (Oral Presentation)	2015, 2017

In-house Oral Presentations

Journal Club	2014, 2015, 2016 (2x)
Werkbespreking fysica	2014, 2015, 2016, 2017
Refereeravond	2014, 2015, 2016, 2017
Research Day	2015
Research Round: Make Way for Youth	2016
Bijscholing laboranten: Automatisch treatment planning	2015, 2017, 2019 (x5)
Bijscholing laboranten: Advanced Monaco® treatment planning	2019 (4 dagen)

2. Miscellaneous

Peer-reviewed 9 manuscripts for 5 journals	2015 - 2019
Abstract selected as one of the best five ESTRO 36 abstracts (physics track)	2017
Invited Speaker ESTRO 37 Symposium	2018
Clinical project: Update automatisch plannen klinisch	2017
Part in organizing committee of Erasmus MC Research Day II and III	2016, 2017
Planning competition: AAMD/RSS SBRT Prostate Plan Study (highest excellence level III, "Best in VMAT")	2016
Planning competition: QAD TG244 H&N Plan Study (highest excellence level II, Written Interview "Best in IMRT")	2017
Collaboration on automated planning with manufacturer of treatment planning systems	2013 - 2019

Curriculum Vitae

Abdul Wahab Sharfo was born on October 22, 1984 in Aleppo, Syria, where he finished high school in 2002. In the same year he started the Bachelor of Science at the University of Aleppo studying Physics. He obtained his Bachelor's degree in 2006 with 'Summa Cum Laude' distinction and continued with the Master of Solid-State Physics. After which he worked as a teacher assistant at the faculty of Medicine at the University of Aleppo and the University of Damascus, where he obtained a scholarship to pursue his postgraduate studies abroad.

In 2010, he moved to the Netherlands and bridged his knowledge in Medical Physics at the Delft University of Technology. In November 2013, he started his Ph.D. research at the department of Radiation Oncology at the Erasmus MC Cancer Institute under the supervision of Prof. dr. Ben J.M. Heijmen and dr. ir. Maarten L.P. Dirkx, which eventually resulted in this thesis.

Currently, Abdul Wahab works as a physicist at the Radiotherapy department in Erasmus MC Cancer Institute, where he next to his clinical duties provides assistance to other researchers.

Abdul Wahab is married to Ranim; and with her, the proud parents of Jude.

Automated treatment planning has been proposed as a tool to substantially increase radiotherapy plan quality and reduce planning workload.
In Rotterdam, Erasmus-iCycle has been developed for fully-automated multi-criterial optimization (MCO) of intensity profiles and beam angles.
Erasmus-iCycle generates for each patient a single Pareto-optimal plan.
For each tumor site, a configuration of the system is needed to ensure that these plans are also clinically relevant.

In this thesis, automated planning with Erasmus-iCycle was validated for several tumor sites by comparisons with clinical manual planning, and manual planning by an expert planner in the absence of time limitations.
In all cases, automatically generated plans had large dosimetric advantages compared to manual planning.
Automated planning was also successfully applied for development of novel treatment techniques and for 'bias-free' comparisons of techniques.