# Automation of Contouring and Planning in Radiotherapy

Peter W.J. Voet

Propositions accompanying the thesis

# Automation of Contouring and Planning in Radiotherapy

1. Although manual editing of autocontours is often required, autocontouring software can significantly reduce the overall contouring workload. (this thesis)

2. In contrast to many other innovations in medicine, automated treatment planning can improve treatment quality, while at the same time reducing costs. (this thesis)

3. The value of beam angle optimization and addition of noncoplanar beam directions is on average small for head-and-neck cancer patients, yet significant for individual cases. (this thesis)

4. In prostate cancer patients with metal hip prostheses, blocking only beam paths (beamlets) that transverse metal before reaching the target, instead of the common full exclusion of beam directions that include such paths, can improve organ at risk sparing to a larger extent than increasing the number of treatment beams. (this thesis)

5. Fully automated, high quality VMAT planning for prostate cancer patients is feasible and has been adopted as a standard clinical routine. (this thesis)

6. Reporting effect sizes may be more informative than reporting *p*-values.

7. The quality of an espresso is determined by the weakest link in the production chain. This is similar to the impact of radiotherapy dose delivery on a serial organ at risk.

8. Although smoking is bad for human health, taxes on selling tobacco also finances the treatment for lung cancer.

9. Mining of the platinum needed for a particulate filter for a truck diesel engine can result in the same amount of pollution as if that same truck would drive around for 30 years without the particulate filter.

10. The job contents of radiation dosimetrists differs substantially among countries, mainly originating from differences in responsibilities in the treatment planning process. The Dutch system makes practicing the profession attractive.

11. There is a clear relationship between the experienced travel distance and the available time.

Peter W.J. Voet, 12-6-2014

AUTOMATION OF CONTOURING AND PLANNING IN RADIOTHERAPY

Automation of contouring and planning in radiotherapy

Layout: Zink Typografie (www.zinktypografie.nl)
Printing: Smart Printing Solutions, Gouda
Cover design by: Henrie van der Est and Peter Voet; photo was taken at Fraser Island, Australia.


ISBN: 978-94-6228460-9

# Automation of Contouring and Planning in Radiotherapy

## Automatisering van het intekenen en van 'planning' in de radiotherapie

PROEFSCHRIFT

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

Prof.dr. H.A.P. Pols

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

donderdag 12 juni 2014 om 13.30 uur

door

Peter Willy Josephus Voet

geboren te Terneuzen

**Promotiecommissie:**

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

**Overige leden:** Prof.dr. A. van der Lugt
Prof.dr. M. Alber
Prof.dr. R.J. Baatenburg-de Jong

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

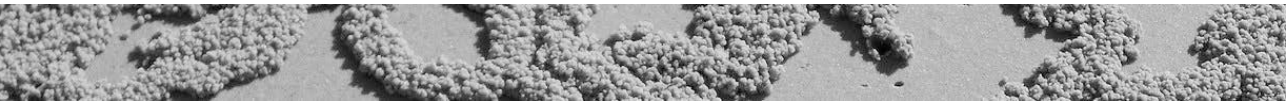
The work described in this thesis was performed at the Department of Radiation Oncology at the Erasmus MC, Cancer Institute, Rotterdam, The Netherlands.

Email for correspondence: peterwjvoet@gmail.com

# contents

# chapter 1

## Introduction



# 1.1 Cancer and radiotherapy

The World Health Organization (WHO) has identified cancer as one of the four leading threats to human health. Cancer is responsible for about 20% of deaths in the European region, being the most important cause after cardiovascular diseases. In Europe, there are more than 3 million new cases every year. For The Netherlands, the Dutch Cancer Society has predicted an increase in the total number of new cancer patients from 100.000 in 2011 to 125.000 in 2030. Radiotherapy is one of the main treatment modalities for cancer, along with surgery and chemotherapy. According to the European Organization for Radiation Oncology (ESTRO), roughly 50% of all new cancer patients are treated with radiotherapy. Generally, radiotherapy is used as a local treatment, irradiating body areas (targets) with proven or suspected disease with ionizing radiation. Most patients are treated with external beam radiotherapy, using a linear accelerator for generation of the applied X-rays.

## 1.2 Current contouring and external beam radiotherapy planning in the clinical workflow

The goal of radiotherapy is to deliver a high dose of radiation to tumor cells while minimizing the risk of treatment-induced complications by keeping doses to healthy radiosensitive tissues surrounding the tumor (i.e., organs-at-risk, OARs) as low as possible. To this purpose, an individualized treatment plan is generated for each patient, aiming at defining the optimal settings of the treatment unit for the intended treatment. Starting point for plan generation is an accurate definition of the tumor and OARs in 3 dimensions (contours) in a computer tomography scan (CT-scan) of the patient, acquired prior to treatment. Plan generation is performed by a dosimetrist in collaboration with the treating physician, using the treatment prescription as a basis. In this prescription, the physician defines treatment objectives in terms of dose to be delivered to the tumor and limitations in doses that may be received by the various OARs. For high dose, curative treatments, intensity modulated radiotherapy (IMRT) or volumetric modulated arc therapy (VMAT) is often used to obtain the best possible dose distribution for the patient.

In order to optimally shape the dose distribution using IMRT (or VMAT), *accurate* delineations (contours) of *all* relevant organs in the planning CT scan are mandatory. Compared to 3D conformal radiotherapy (3DCRT), IMRT does generally require more contouring time. For head-and-neck cancer patients, manual contouring of the individual neck levels I-V [1] (both necks) and 20 OARs (salivary glands, chewing and swallowing muscles, and spinal cord/brainstem) may take 3 hours [see Chapter2]). Despite well-described guidelines for manual contouring of neck levels and OARs [1,2,3,4], contouring still results in substantial intra- and interobserver variations [5,6,7].

Currently, plan generation is a manual, interactive trial-and-error procedure in which the dosimetrist tries to steer the treatment planning system (TPS) toward a plan that satisfies the treatment prescription, with minimal dose delivery to the most important OARs. With this procedure there is no guarantee that, even after many hours of planning workload, an alternative approach in steering the TPS software would not have resulted in a significantly improved treatment plan. In other words, it always remains unclear whether more time and effort should be invested to further search for a better therapy. On top of this, well-known (substantial) variations in the skills, experience and drive of planners for generation of high quality treatment plans (even within a single department), and the quality of interaction of the treating clinician with the planners may seriously impact plan quality for individual patients. Another frequent threat for generation of high quality plans in current clinical practice is lack of planning time, sometimes related to lack of planning resources (personnel or workstations and software licenses). In case of time limitations, even a planner full of ideas and motivated for further

plan improvement may not be able to achieve the best dose distribution for an individual patient. Apart from the existing unfavourable variability in plan quality, the current human factor in plan generation does generally also result in a high cost because of the involved workload, especially for difficult planning cases with many hours of planning time.

## 1.3 Outline of the thesis

The work presented in this thesis focuses on *automation* of 1) contouring in planning CT-scans and, 2) treatment plan generation. The goal is to resolve problems with the current manual procedures, as described above.

For autocontouring we investigated the use of the commercially available atlas-based autosegmentation software, ABAS (Elekta). In atlas-based autosegmentation, the CT-scan of a new patient is segmented using segmented scans of previously treated patients, called atlases.

The basis of the investigations on automated plan generation was Erasmus-iCycle [8], an in-house developed algorithm for integrated beam angle and beam profile optimization. Erasmus-iCycle automatically generates for each individual patient one or more Pareto-optimal IMRT plans, based on constraints and prioritized objectives specified for the corresponding disease/tumor site.

In chapter 2, we performed a clinical validation study on atlas-based autosegmentation for target volumes and normal tissues for head-and-neck cancer patients. The accuracy of ABAS, based on a single and multiple atlases was assessed. Manual contouring times were compared with the time needed to edit the autocontours.

In chapter 3 we investigated the necessity of editing automatically generated contours for head-and-neck cancer patients by evaluating dosimetric parameters for the neck levels and OARs in generated treatment plans.

In chapter 4, Erasmus-iCycle was used for automated plan generation in head-and-neck cancer to compare different beam angle selection strategies. 7- and 9-beam equiangular beam configurations were compared with 7- and 9-beam plans with optimized beam angles. In addition to coplanar beam arrangements, noncoplanar beam set-ups were studied as well. The optimal couch angle in coplanar plans was derived to be used as basic input for volumetric modulated arc therapy (VMAT) planning.

In chapter 5, a prospective clinical study is presented to demonstrate the clinical

value of automated treatment planning. For a group of 20 patients, treating physicians had to choose between a manually optimized IMRT plan, made by dosimetrists, and a plan based on automatic plan generation with Erasmus-iCycle. For the latter plan, the Erasmus-iCycle dose distribution was manually reconstructed in the clinical treatment planning system to obtain a clinically deliverable plan. The patient was treated with the preferred plan.

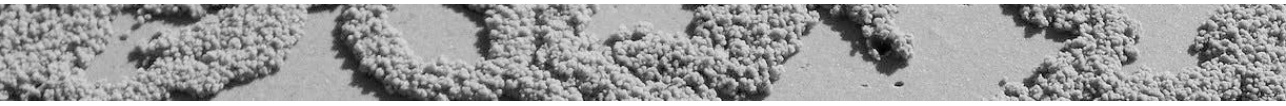
In chapter 6, Erasmus-iCycle was used to study different planning strategies for prostate cancer patients with metal hip prostheses. Plans with beam directions that fully avoided the metal prostheses were compared with plans in which all beam directions were allowed, but parts of the beam irradiating through metal in front of the tumor were blocked. The investigations did include an assessment of the added value of noncoplanar beam arrangements.

In chapter 7 we present a method to fully automatically generate clinical treatment plans for prostate cancer patients, i.e. the manual reconstruction of Erasmus-iCycle plans by dosimetrists was now also automated (compare with chapter 5). Fully automated plan generation was investigated for prostate cancer patients treated with VMAT.

In chapter 8 the results of this thesis are discussed followed by a discussion on future perspectives.

chapter 2

# Clinical validation of atlas-based autosegmentation of multiple target volumes and normal tissue (swallowing/mastication) structures in the head & neck



David N. Teguh, MD, PhD, Peter C. Levendag, MD, PhD, Peter W.J. Voet, RTT, Abrahim Al-Mamgani, MD, PhD, Xiao Han, PhD, Theresa K. Wolf, BSc, Lyndon S. Hibbard, PhD, Peter Nowak, MD, PhD, Hafid Akhiat, BSc, Maarten L.P. Dirkx, PhD, Ben J.M. Heijmen, PhD, Mischa S. Hoogeman, PhD

*International Journal of Radiation Oncology Biology Physics; 15:81(4): 950–7 (2011)*

# Abstract

**Purpose:** To validate and clinically evaluate autocontouring by use of atlas-based autosegmentation (ABAS) of computed tomography images.
**Methods and Materials:** The data from 10 head-and-neck patients were selected as input for ABAS and the neck levels I-V, and 20 organs-at-risk were manually contoured according to published guidelines. The total contouring times were recorded. Two different ABAS strategies, multiple- and single subject, were evaluated and the similarity of the autocontours with the atlas contours was assessed using Dice coefficients and mean distances, using the leave-one-out method. For 12 clinically treated patients, 5 experienced observers edited the autosegmented contours. The editing times were recorded. The Dice coefficients and mean distances were calculated among clinically used contours, autocontours and edited autocontours. Finally, an expert panel scored all autocontours as well as the edited autocontours regarding their adequacy relative to the published atlas.
**Results:** The time to autosegment all the structures by ABAS was 7 minutes/patient. No significant differences were observed in the autosegmentation accuracy for N0 and N+ patients. The multi-subject atlas performed best, with a Dice coefficient and mean distance of 0.74 and 2 mm, 0.67 and 3 mm, 0.71 and 2 mm, 0.50 and 2 mm, and 0.78 and 2 mm, for salivary glands, neck levels, chewing muscles, swallowing muscles, and cord-brainstem, respectively. The mean Dice coefficient and mean distance of the autocontours vs. clinical contours was 0.8 and 2.4 mm for the neck levels and salivary glands, respectively. For the autocontours vs. edited autocontours, the mean Dice coefficient and mean distance was 0.9 and 1.6 mm, respectively. The expert panel scored 100% of the autocontours as "minor deviation, editable" or better. The expert panel scored 88% of the edited contours as good opposed to 83% for the clinical contours. The total editing time was 66 min.
**Conclusions:** Multiple-subject ABAS of CT images proved to be a useful novel tool in rapid delineation of target and normal tissues. Although editing of the autocontours is inevitable, substantial time reduction was achieved using editing, instead of manual contouring (180 vs. 66 min.).

# 2.1 Introduction

The large numbers of target and normal tissue structures that require manual delineation in head-and-neck cancer patients make contouring tedious and time consuming. In addition, optimal sparing conditions in the head-and-neck (H&N) IMRT require accurate delineation of those structures. To ensure consistent delineation of the target volume, a CT-based atlas of the neck levels I-V and guidelines for critical organs-at-risk (OARs) were developed and now used [1,2.3,4] However, contouring still suffers from intra- and inter-observer variations [5,6,7].

A promising new tool is autocontouring by atlas-based autosegmentation (ABAS) of CT-images [9,10,11]. This tool automatically creates the contours for the neck levels and organs-at-risk in the CT-images of a new patient. Atlas-based autosegmentation has the potential to lower the contouring burden and thus allow more normal tissues to be included in inverse treatment planning for high-dose IMRT to fully exploit knowledge on dose-volume effects. Atlas-based autocontouring also potentially reduces the intra-observer and inter-observer variability in contouring.

In this study, we have quantified the accuracy of autocontouring using ABAS and assess the clinical applicability of this tool. This is to our knowledge the first paper describing the validation of ABAS (Elekta) of target tissues (including the neck levels I-V) and all possible normal tissue structures (including the mastication and swallowing muscles) in the head-and-neck. We determined the accuracy and time reduction of contouring. The first part of the paper assesses the geometrical accuracy of ABAS. Two ABAS approaches were evaluated (1) selection of the atlas patient with the greatest similarity metric and (2) combining multiple segmentations of all atlas patients into one segmentation. The comparison of a multiple-subject atlas against the use of a single-subject atlas was quantified. The second part of the study addressed the clinical implementation. The differences among the clinically used contours, autocontours, and edited autocontours were quantified, and the quality of all contours was scored by an expert panel.

# 2.2 Material and Methods

## Description of ABAS

Atlas-based autosegmentation is the process of performing segmentation on novel data using the knowledge of a prior segmentation, a dataset that had the structures of interest already labelled. The registration strategy incorporates objects' shape information in the atlas to help improve the registration efficiency and robustness, while still being able to account for large intersubject shape differences. The key component of atlas-based autosegmentation is a database, (i.e., the so-called atlas), containing image data (e.g. CT scan data) with delineated contours of the structures (organs) of interest. These atlas contours must be transferred to the image data of a new patient who undergoing radiotherapy. The transfer is accomplished by non-rigidly registering the image data of the atlas to the image data of the new subject. Having obtained the transformation vectors from the atlas image data to the new subject image data, it is possible to transform the atlas-contours to the new patient. The implementation that was here tested uses for the non-rigid registration a hierarchical approach (ABAS version 1.1). This approach makes the non-rigid registration procedure efficient and robust, while still being able to register large shape differences that are present between different patients. Hierarchical atlas registration consists of three major steps: linear registration; object-driven 'poly-smooth' non-linear registration; and shape-constrained dense deformable registration [10,11].

## Atlas patients

CT data of 10 H&N cancer patients (4 No and 6 N+) was used. An experienced staff member manually contoured individual levels I-V (both necks) and twenty OARs (salivary glands, chewing and swallowing muscles, and cord/brainstem) (Table 2.1).

First, the contouring was done by a research fellow and subsequently corrected by the supervisor. Contouring of the neck levels are according to the international consensus guidelines of Gregoire and Levendag and for the normal tissues according to our guidelines for delineating OARs [2,3,4,12,13]. Those contours were used to construct the atlas and were regarded as a gold standard. The total contouring time that was needed to create the atlas was recorded. The contours of levels with invaded muscle in N+ patients

**Table 2.1:** Tissues delineated for the atlas.

| Tissue | n |
|---|---|
| Submandibular glands | 2 |
| Parotid glands | 2 |
| Mastication muscles | 5×2 |
| Swallowing muscles | 5 |
| Spinal cord | 1 |
| Neck levels | 5×2 |

were not used for autosegmentation, because in our guideline the anatomical border of those levels is differently defined as for levels with a non-invaded muscle. Apart from the levels with invaded muscle, all other contours of N+ patients were used by the ABAS tool.

## Atlas Selection

In theory, atlas-based autosegmentation requires just one set of images from a patient as the atlas. In practice, however, the difference in the anatomy of patients merits the use of various atlas patients. In this paper, we evaluated two different atlas selection strategies to determine the best approach (figure 2.1).

The first approach is a selection of the atlas patient based on the highest similarity metric between all the atlas patients and the new patient. The global mutual information similarity after the global linear registration was used to choose the best atlas. The global Correlation Coefficient is only good for linear registration. The local correlation coefficient was also evaluated. The Local Correlation Coefficient (LCC) is the usual correlation coefficient computed within a small neighbourhood of each image voxel. The second approach for an atlas selection strategy is to apply multiple atlas data sets to the CT data of a new patient, thereby generating multiple sets of autosegmentation sets. For the fusion of multiple single-atlas autosegmentations, to one multiple-subject autosegmentation, the STAPLE algorithm was used [14]. The STAPLE (Simultaneous Truth And Performance Level Evaluation) algorithm was introduced by Warfield et al. which offers a more sophisticated strategy for multiple segmentation fusion in that it automatically estimates the segmentation quality of each classifier and at the same time derives a weighted combination of the multiple classifiers based on their estimated segmentation quality.

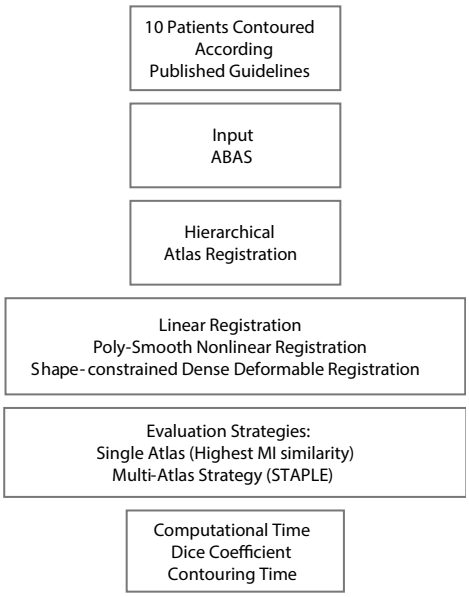



**Figure 2.1:** Flow chart input ABAS and strategy evaluation

STAPLE method was applied for each structure separately. For each structure, it takes as input a collection of segmentation results, one for each atlas. It then simultaneously computes: (1) a probabilistic estimate of the 'true' segmentation and (2) a measure of the performance level for each individual atlas result.

## Atlas evaluation

For evaluation the mean distance and standard deviation around the mean distance between the autosegmented structure and the reference structure and the Dice coefficient were calculated. Mean distance measure is the mean distance between corresponding points in the surfaces of $A$, $B$. The surfaces are represented by triangular meshes consisting of sets of node points (triangle vertices), and edges (triangle sides) between the node points. Since the surfaces have generally different numbers of nodes, the distances $d(A, B)$ are taken from the nodes on one surface $A$ to the nearest node on the other

structure $B$. The mean distance is defined as:

$$M(A,B) = \frac{1}{N(A)} \sum_{i=0}^{N(A)} d(A,B)$$

where $d(A,B) = \sqrt{(a_i - b_i)^2}$ is the Euclidean distance from the $i$-th point on $A$, $a_i$, to $b_i$, the point on $B$ closest to $a_i$, and $N(A)$ is the number of surface mesh nodes on structure $A$. The standard deviation of the distances is defined as

$$SD(d(A,B)) = \sqrt{\frac{1}{N(A)-1} \sum_{i=0}^{N(A)} (d(A,B) - M(d(A,B)))^2}.$$

This is a measure of the degree to which the distances are spread out over their range. If the structures are in good agreement, the mean and standard deviation are both small. The Dice coefficient is defined as:

$$\text{Dice} = 2\frac{|A \cap B|}{|A| + |B|}$$

where $A$ and $B$ are the two structures evaluated. This formula represents the size of the union of 2 structures divided by the average size of the two sets. A value of 0 indicates no overlap; a value of 1 indicates perfect agreement.

The leave-one-out cross validation method was used to remove bias, i.e. the patient for whom autocontours were generated was temporarily removed from the atlas. Levels with an invaded muscle were excluded from the geometrical validation, because the border of these levels differ by definition from the levels without invaded muscle. Levels with invaded muscle were evaluated in the clinical validation of ABAS (see next section).

## Clinical Validation of ABAS

First, for 12 No and N+ patients who had undergone intensity modulated radiotherapy by 10 experienced clinicians and residents, the clinically applied contours (neck levels and salivary glands) were evaluated by an expert panel and compared with the published atlases [2,3,4,12,13] using the following scores: 0=poor, 1=moderate, 2=good (figure 2.2).

Then, autocontours were generated for those patients by using the multiple-subject ABAS tool. The accuracy of autocontours relative to the published atlas was evaluated by

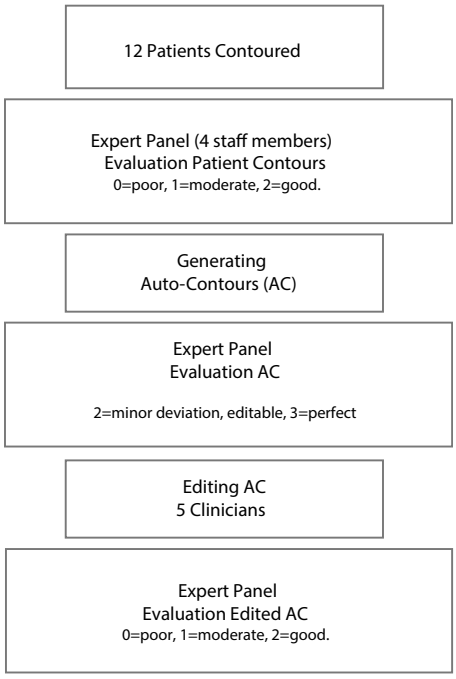



**Figure 2.2:** Flow chart clinical validation

the expert panel. For the autocontours, the following scoring system was used: 0=poor, 1=major deviation, editable, 2=minor deviation, editable, 3=perfect. Generally , the expert panel scored it as "minor deviation, editable" when the structures need to be edited on maximum 3 CT-slices, otherwise it was scored as major deviation, editable. Finally, for each patient those autocontours were offered to 2 of a group of 5 experienced observers to edit, if needed, and the editing times were recorded. Editing the contours were done in Focal version 4.3.3 (Elekta). Those edited autocontours were also scored by the expert panel. The Dice coefficients and mean distances were calculated to quantify the differences among the clinical contours, autocontours, and edited autocontours of the observers (figure 2.2). In addition, other contours, i.e. mastication muscles, swallowing muscles, and cord-brainstem were autocontoured and edited and editing time was also recorded.

# 2.3 Results

## Geometrical validation of ABAS

Time to autocontour all the structures by ABAS was approximately 7 minutes/patient. The initial contouring time for the 30 structures delineated (Table 2.1) was on average 180 minutes per patient. Figure 2.3 shows an example of an autosegmentation of a neck level, mastication and swallowing muscles using a single-subject and a multiple-subject atlas.

In this example, the multiple-subject atlas contours were in better agreement with the reference contours than the selected single-subject atlas contours. The comparison for all patients and structures was summarized in figure 2.4.

This figure shows the Dice coefficients and mean distances for the single-subject and multiple-subject approaches. The multiple-subject atlas method consistently performed superior to the selected single-subject atlas. Those results were in agreement with the fact that the tested similarity metrics did not or moderately correlate with the accuracy of the autosegmentation (median $R^2$ of 0.2, range 0.1–0.6). Excluding the levels with invaded muscle, no significant differences were observed in the autosegmentation accuracy for No and N+ patients.

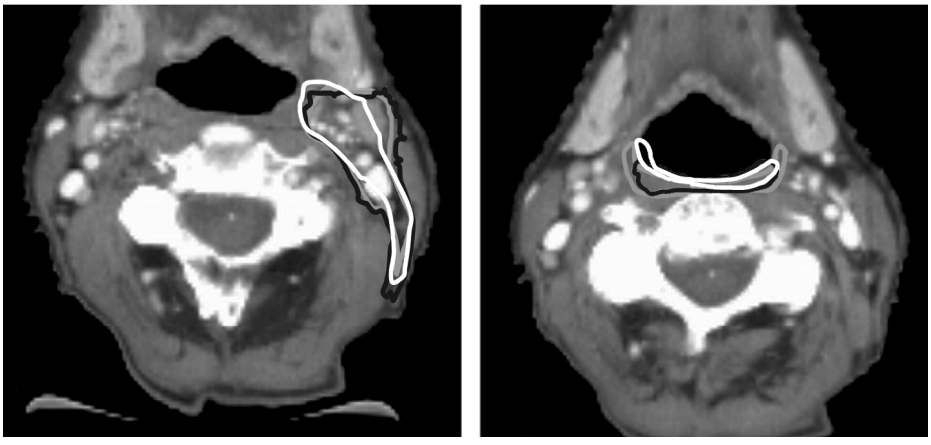


**Figure 2.3:** Examples of single-subject autosegmentations (white), multiple-subject autosegmentations (black) together with the reference contours (golden standard) (grey). Left: level II, Right: Swallowing muscle.

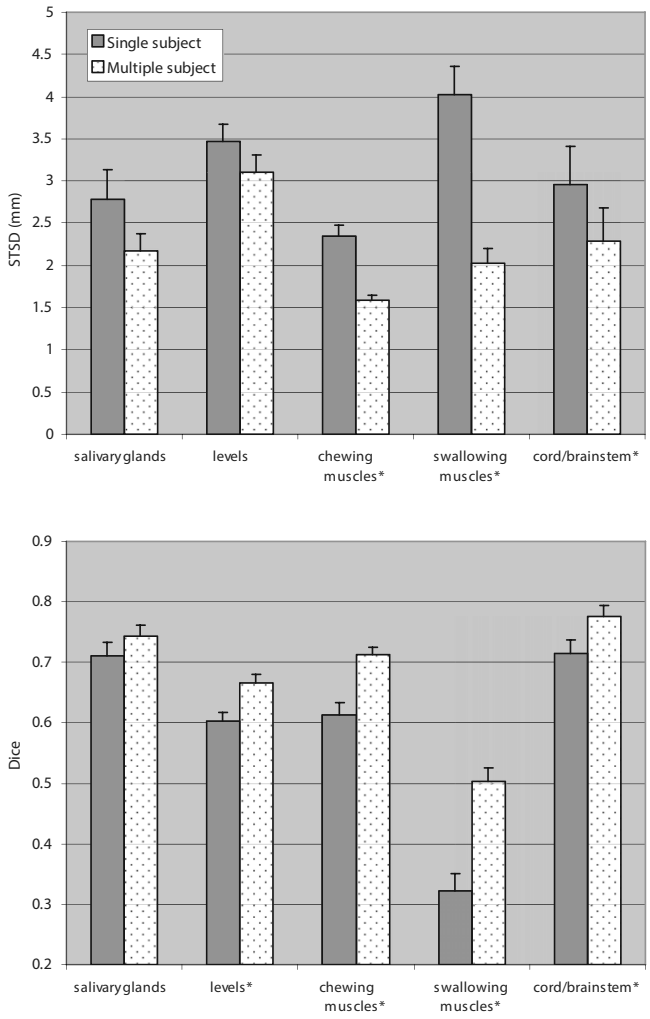



**Figure 2.4:** Comparing mean distance and Dice coefficient for the single-subject and multiple-subject approach. Both metrics quantify the agreement between autocontours and the gold standard. The error bar equals the standard error and the asterisk denotes whether the difference for a subgroup of structures is significant at the 0.01 level (paired t-test). Overall, the multiple-subject method performed better ($p <$ 0.001; paired t-test).

**Table 2.2:** Dice coefficient and mean distance for multiple-subject autocontours vs. reference (golden standard) contours.

| Variable | Neck levels | Parotid glands | Submandibular glands | Chewing muscles | Swallowing muscles | Spinal cord - brainstem |
|---|---|---|---|---|---|---|
| Dice coefficient | 0.67 | 0.79 | 0.70 | 0.71 | 0.50 | 0.78 |
| Mean distance * (mm) | 3.4 ± 3.1 | 2.5 ± 2.8 | 1.9 ± 1.4 | 1.6 ± 1.4 | 2.0 ± 1.9 | 2.3 ± 1.4 |

* Mean ± standard deviation.

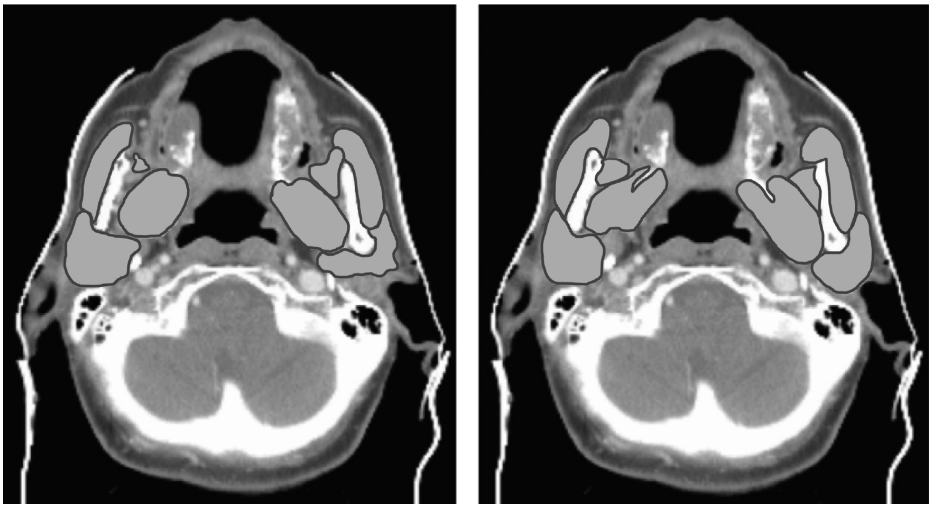


**Figure 2.5:** Autocontouring of organs-at-risk (mastication muscles, parotid gland) (left) vs. reference contours (golden standard)(right).

For all patients, the mean Dice coefficient and mean distances (mm) of the multiple-subject method were listed in Table 2.2.

Other results of the multiple-subject autosegmentations are shown in figures 2.5, 2.6 and 2.7.

In figure 2.5, an example of the application of auto-ABAS for the contouring of OARs (masseter muscles, pterygoid muscles, temporalis muscles and parotid glands) is shown. Figure 2.6 shows a coronal view of the autocontours and reference contours of levels II to IV. Figure 2.7 shows the autosegmentation of levels vs. observer contouring. Figure 2.8 shows an autosegmentation example for a neck level with invaded muscle (reference contour left, and multiple-subject autocontour right).

The muscle was not included in the autocontour, because the algorithm used the level

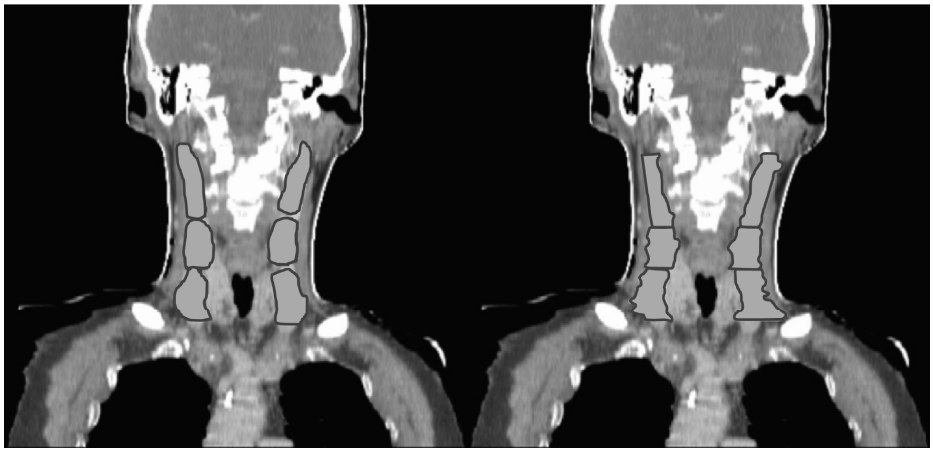


**Figure 2.6:** Coronal view of autocontours of neck levels (left) vs. reference contours (golden standard) (right).

of No atlas patients to autosegment the involved level of the N+ patient. At the medial site, some vessels were not encompassed. Both issues can be quickly fixed by editing the autocontour.

## Clinical validation

The mean editing time for neck levels, parotid glands, submandibular glands, mastication muscles, swallowing muscles, and cord-brainstem was 31, 7, 6, 14, 7, and 1 min respectively. The Dice coefficients and the mean distance of the clinical contours vs. the autocontours, autocontours vs. edited autocontours, and the observer 1 vs. observer 2, were listed in Table 2.3.

Most variation was found for the levels and in particular for the clinical contours of the levels and the autocontours of the levels. All autocontours (100%) were scored as "minor deviations/editable" or better by the expert panel. The expert panel scored 88% of the edited contours as good, where 83% of the clinically used contours were scored as good. The automatically generated contours still need to be edited, but the editing time is much less than the time needed for manual delineation from scratch.

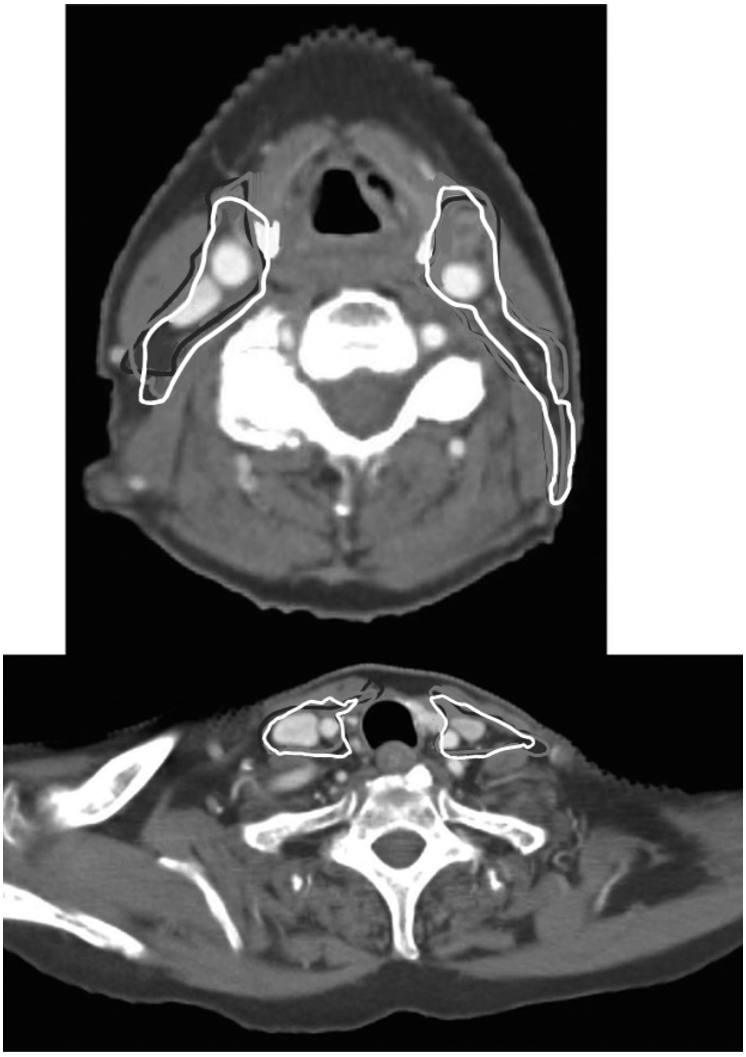


**Figure 2.7:** Autocontouring of levels (white contour) vs. observer contouring (black and grey contours).

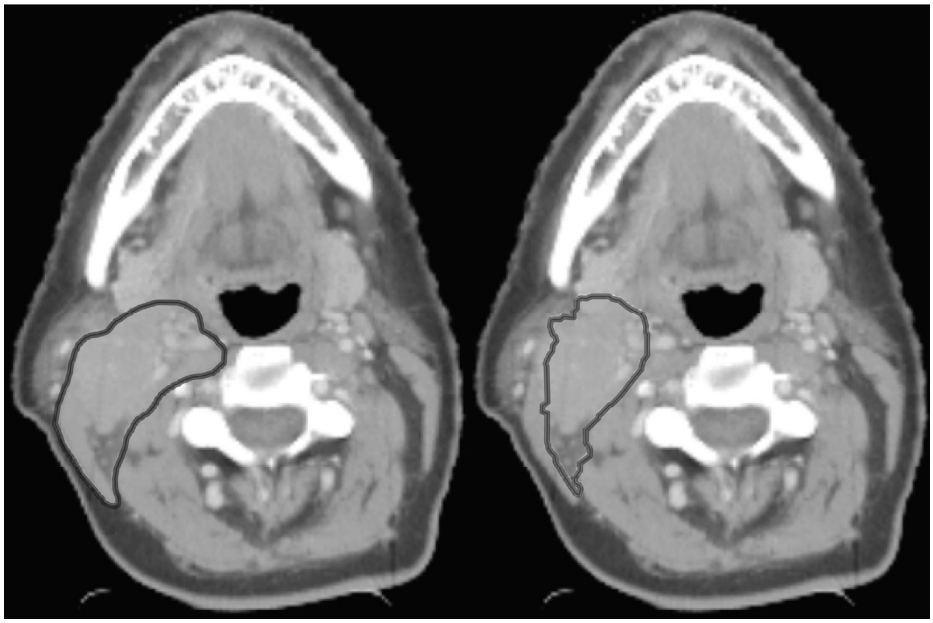


**Figure 2.8:** Autosegmentation example for a neck level with invaded muscle (reference contour left and multiple-subject autocontour in right). The muscle was not included in the autocontour, because the algorithm used the level of N0 atlas patients to autosegment the involved level of the N+ patient. At the medial site, some vessels were not encompassed. Both issues can be quickly fixed by editing the autocontour.

**Table 2.3:** Dice coefficient and mean distance and SD around the mean distance of different structures comparing clinical contours, autocontours (AC) and edited autocontours (e-AC). e-$AC_1$ denotes the edited autocontour of observer 1 and e-$AC_2$ denotes the edited autocontour of observer 2.

| Variable | Clinical vs. AC | Clinical vs. e-AC | AC vs. e-AC | e-AC1 vs. e-AC2 |
|---|---|---|---|---|
| **Dice coefficient** | | | | |
| Neck levels | 0.73 | 0.79 | 0.83 | 0.81 |
| Parotid glands | 0.80 | 0.81 | 0.91 | 0.89 |
| Submandibular glands | 0.72 | 0.77 | 0.83 | 0.86 |
| **Mean distance ± SD (mm)** | | | | |
| Neck levels | 3.2 ± 3.6 | 2.2 ± 2.8 | 1.5 ± 2.8 | 1.8 ± 2.5 |
| Parotid glands | 2.3 ± 3.8 | 2.1 ± 3.0 | 0.8 ± 1.2 | 1.1 ± 1.3 |
| Submandibular glands | 1.6 ± 1.4 | 1.2 ± 1.2 | 1.4 ± 1.8 | 1.0 ± 1.1 |

Abbreviations: AC = autocontours; e-AC = edited autocontours; e-AC1 = e-AC of observer 1; e-AC2 = e-AC of observer 2; SD = standard deviation.

## 2.4 Discussion

This study described the development, validation and use of the ABAS tool. As seen in figure 2.4 comparison with the reference delineation clearly shows the advantage of using a multiple-subject atlas for segmentation. The multiple-subject atlas Dice coefficients and mean distances are more satisfactory than those with the single-subject atlas. As demonstrated by figures 2.3–2.6, the organs-at-risk, parotid glands and mastication muscles are accurately segmented. From figure 2.6, resemblance is apparent, but the autocontours still required editing to be used in the treatment planning process. The Dice coefficients between the different structures are listed in Table 2.3. From these results around 80–90% agreement between the autocontours and the edited autocontours was found, similar or slightly better than other published data [15,16,17]. Although our data showed a considerable inter-patient variability in intravenous contrast uptake, head pose, dental artifacts, and use of a tongue depressor, all the autocontours were scored as "minor deviations, editable" or better by the expert panel, a promising result. The expert panel scored 88% of the edited contours as good, where 83% of the clinically used contours were scored as good. From these data we can state that the edited contours (88% good) were closer to the published atlas than the ones used in clinic (83% good); the former might make the delineations for radiotherapy plans for cancer patients more accurate. Chao et al. used a computer-assisted target volume delineation system and reference templates. They found that the variation was significantly reduced but not necessarily the accuracy, which was beyond the scope of their study [16].It should be remembered that the border in N+ patients with invaded muscle in the present group was defined differently than those without invaded muscle. Therefore, the invaded muscle has to be manually added to the autocontour of the level [13].

Castadot et al. [15] compared 12 voxel-based deformable registration strategies in adaptive radiation therapy for the treatment of head-and-neck tumors. His dataset contained 5 patients. They concluded that level-set Demon's algorithm (voxel-intensity-based registration) implemented in multi-resolution is a good strategy for head-and-neck adaptive radiation therapy as it is the best compromise in terms of median and inter-quartile range for both the Dice similarity index and correlation coefficient. Sims et al. tried to establish the accuracy of ABAS, such that a priori information was used to delineate a limited set of organs of interest (brainstem, parotid glands and mandible). The Dice coefficients for all OARs were 0.68 ± 0.25 for a first center and 0.82 ± 0.13 for a second center. Systematic oversegmentation of the parotids and undersegmentation of the brainstem occurred that required careful review and editing in most cases [18]. They concluded that the autocontours and the substantial time reduction in contouring time proved that atlas-based autosegmentation would be a useful novel tool in the rapid

delineation of neck levels and the limited number of organs-at-risk evaluated. Wang et al. mapped contours from the planning CT onto daily CT or four-dimensional CT images using an image intensity-based deformable registration algorithm. Here also only a limited set of contours was evaluated (CTV, parotid glands and brain stem). The volume overlap index ($A \cap B/(A+B)/2$)) and mean absolute surface-to-surface distance was 83% and 1.3 mm, respectively. They concluded that a final review by physicians is highly recommended [19]. Commowick et al. presented a method for creating an anatomical atlas of the head-and-neck region from a database of 45 manually delineated CT images [9]. They constructed a mean CT image set with atlas contours from a database of patients beforehand. The constructed mean atlas is then applied to a new patient. The evaluation of the built atlas has shown good results both qualitatively and quantitatively, although some important structures were not included in their database. Our study described an autosegmentation system with all possibly important (normal tissue) structures needed for treating head-an-neck cancer included. Zhang et al. showed for 7 patients that atlas-based image segmentation can automatically delineate the organs-at-risk (mandible, brainstem and parotids only) on daily CT images [17]. Quantitative validations demonstrated that the method was robust. Their Dice coefficients were slightly lower than ours.

Other studies [20,21] have shown anatomical changes in parotid and submandibular glands during radiotherapy, as assessed with non-rigid registration. The non-rigid registration framework of Vásquez Osorio [21] is an effective method to simultaneously register anatomical changes of multiple organs with very different magnitudes and complexity. General shrinkage and deformation of irradiated glands were observed. The spared glands showed few changes. Anatomical changes can be accommodated by repeat planning (adaptive radiotherapy), however repeat planning requires re-contouring of the new CT image or cone beam CT data set, which is tedious and time consuming. In the future, the ABAS tool might be used to ease adaptive radiotherapy.

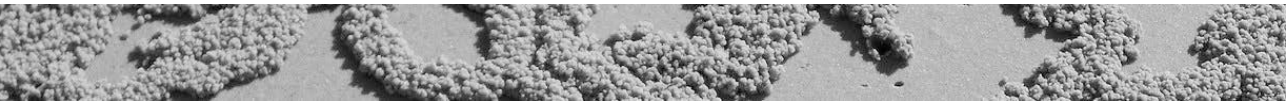
## 2.5 Conclusions

Despite the large inter-patient variability in the study population, neck levels and OARs could be accurately contoured, using the ABAS tool. The multi-subject atlas performed better than the best single-subject atlas. Although editing of the autocontours is inevitable (66 min for 30 structures), substantial time reduction was achieved by editing instead of manually contouring. Therefore, the new algorithm for autosegmentation could

substantially reduce the clinical workload spent on organ delineation. This is even more relevant, because the edited contours were scored as having similar or better quality than the clinically used contours.

chapter 3

# Does atlas-based autosegmentation of neck levels require subsequent manual contour editing to avoid risk of severe target underdosage? A dosimetric analysis



Peter W.J. Voet, RTT, Maarten L.P. Dirkx, PhD, David N. Teguh, MD, PhD,
Mischa S. Hoogeman, PhD, Peter C. Levendag, MD, PhD, Ben J.M. Heijmen, PhD

*Radiotherapy Oncology; 98(3):373–7 (2011)*

# Abstract

**Background and Purpose:** To investigate the *dosimetric* impact of not editing autocontours of the elective neck and organs-at-risk (OAR), generated with atlas-based autosegmentation (ABAS) (Elekta software) for head-and-neck cancer patients.
**Materials and Methods:** For nine patients ABAS autocontours and autocontours edited by two observers were available. Based on the non-edited autocontours clinically acceptable IMRT plans were constructed (designated 'ABAS plans'). These plans were then evaluated for the two edited structure sets, by quantifying the percentage of the neck-PTV receiving more than 95% of the prescribed dose ($V_{95\%}$) and the near-minimum dose ($D_{99\%}$) in the neck PTV. Dice coefficients and mean contour distances were calculated to quantify the similarity of ABAS autocontours with the structure sets edited by observer 1 and observer 2. To study the dosimetric importance of editing OAR autocontours a new IMRT plan was generated for each patient-observer combination, based on the observer's edited CTV and the non-edited salivary gland autocontours. For each plan mean doses for the non-edited glands were compared with doses for the same glands edited by the observer.
**Results:** For both observers, edited neck CTVs were larger than ABAS autocontours ($p \leq 0.04$), by a mean of 8.7%. When evaluating ABAS plans on the PTVs of the edited structure sets, $V_{95\%}$ reduced by 7.2% ± 5.4% (1 SD) ($p < 0.03$). The mean reduction in $D_{99\%}$ was 14.2 Gy (range 1–54 Gy). Even for Dice coefficients >0.8 and mean contour distances <1 mm, reductions in $D_{99\%}$ up to 11 Gy were observed. For treatment plans based on observer PTVs and non-edited autocontoured salivary glands, the mean doses in the edited glands differed by only −0.6 Gy ± 1.0 Gy ($p = 0.06$).
**Conclusions:** Editing of autocontoured neck CTVs generated by ABAS is required to avoid large underdosages in target volumes. Often used similarity measures for evaluation of autocontouring algorithms, such as Dice coefficients, do not predict well for expected PTV underdose. Editing of salivary glands is less important as mean doses achieved for non-edited glands predict well for edited structures.

# 3.1 Introduction

Modern intensity modulated radiotherapy (IMRT) treatment planning systems, especially those with biologically based IMRT optimisation, allow users to generate treatment plans with tight constraints on many organs-at-risk (OARs), while maintaining good target coverage [22]. A basic prerequisite for IMRT planning is that all OARs for which constraints are applied needs to be contoured. In addition, due to steep dose gradients outside the planning target volume (PTV), the accuracy of delineation is becoming of greater importance. For head-and-neck cancer patients manual contouring of the clinical target volume (CTV), consisting of neck levels to be treated electively, and relevant OARs typically takes 3 h in our clinic [5]. Often, the delineations are sensitive to inter-observer variations. To reduce these inter-observer variations, guidelines for delineation of all neck levels and OARs have been developed and are currently used in clinical routine [1,2,3,12,13]. To reduce the time needed for delineation, there is a growing interest in the use of autocontouring software. Huykens et al. [23] validated the use of an autosegmentation tool for prostate cancer patients based on CT datasets. The accuracy of the automatically contoured prostate was scored as 'good', indicating that only minor manual adjustments were needed. For the autocontoured bladder they reported a score between 'excellent' and 'good'. Isambert et al. [24] concluded that an atlas-based autosegmentation tool was robust and reliable for automatic delineation of large structures in the brain; however, manual contouring was still needed for smaller structures. A clinical validation study on atlas-based autosegmentation for head-and-neck cancer patients was recently published by Teguh et al. [25]. They showed that *geometric* differences between autocontours generated with ABAS (Elekta software) and corresponding edited contours were relatively small. Dice coefficients of 0.8 for the CTV neck levels and 0.9 for the salivary glands, and mean contour distances of 1.5 mm and 1.1 mm, respectively, were reported. This study also showed that by editing autocontours, the time needed for delineation compared to manual contouring was significantly reduced from 180 to 66 min on average.

When geometric differences between autocontours and manually edited autocontours are limited, like in the study of Teguh et al. [25], it is questionable whether the impact of adjusting the autocontours on the final dose distribution is clinically relevant. In the PTV some of the dose differences might be diminished by the application of CTV-to-PTV margins to account for set-up accuracy and by the unavoidable non-conformity of the dose distribution. The primary aim of this study is to investigate *dosimetric* implications of not editing autocontoured neck levels and OARs derived with ABAS for head-and-neck cancer patients. The quality of autocontouring algorithms is often evaluated

with Dice coefficients and mean contour distances between edited and non-edited structure sets. The secondary goal of this paper is to investigate whether observed favourable Dice coefficients or mean contour distances do indeed predict for low PTV underdosage.



## 3.2 Materials and Methods

### ABAS and required input

ABAS is a program that segments organs in a CT dataset of a new patient by using one atlas patient or multiple atlas patients. An atlas patient consists of a CT scan with predefined organs of interest, both target volumes and OARs. A detailed description of the method has been published by Han et al. [26]. First, non-rigid registration is used to transform the CT scan of an atlas patient to the CT scan of the new patient. Specific models for e.g., head-and-neck, prostate and brain are available in the software, taking structure-specific information, like elasticity, into account. Then, using the obtained transformation, autocontours are generated by mapping the atlas contours to the CT scan of the new patient. In theory, ABAS just requires one atlas patient to contour a new subject. However, due to differences in patients' anatomies, the accuracy of ABAS improves when a multiple-subject segmentation approach is used [26]. With this method, a set of autocontours is generated for a new patient, based on multiple atlas patients. To combine the different structure sets to one, the STAPLE algorithm is used [14].

For this study, atlases of 10 randomly chosen head-and-neck cancer patients, with tumors in the nasopharynx, oropharynx or base of the tongue were used. Four of these patients were neck-node negative, the others had one or more positive neck nodes. For each patient neck levels I - V were manually delineated on both sides of the neck according to international consensus guidelines [1,12,13]. In addition, 20 OARs were contoured: both parotid and submandibular glands, five swallowing muscles, six chewing muscles, four adjacent structures (coronoid process, condylar process) and the spinal cord. The delineations of the neck levels and OARs were checked and approved by an expert panel of physicians before being used as atlas.

**Table 3.1:** Characteristics of patients used in our study

| Patient | Tumor type | TNM | Treated neck levels | Dose interval | Followed by |
|---|---|---|---|---|---|
| 1 | Squamous cell Ca. right cheek | pT2N0M0 | II-IV ri | 60 Gy PORT* | - |
| 2 | Hypopharynx Ca. | T2N0M0 | II-IV ri+le | 46 Gy | 24 Gy on primary tumor |
| 3 | Nasopharynx/cavum nasi Ca. | T2N2M0 | RPS**, Ib-V ri+le | 70 Gy | - |
| 4 | Sinus piriformis Ca. | T2N0M0 | II-IV ri+le | 46 Gy | 24 Gy on primary tumor |
| 5 | Tongue Ca. | T2N0M0 | Ib-IV ri | 46 Gy | Brachy therapy boost |
| 6 | Supra Glotic Larynx Ca. | T3N0M0 | II-IV ri+le | 46 Gy | 24 Gy on primary tumor |
| 7 | Glotic Larynx Ca. | T2N0M0 | II-IV ri+le | SIB *** | - |
| 8 | Oropharynx Ca. | T1N2bM0 | RPS, Ib-IV ri | 46 Gy | BT + ND **** |
| 9 | Base of tongue Ca. | T3N1M0 | Ib-IV ri, Ib-V le | 46 Gy | BT + ND **** |

* PORT: post-operative RT

** RPS: retropharyngeal space

*** SIB: Simultaneous integrated boost, 30×2.3 Gy primary tumor, 30×1.7 Gy neck levels

**** BT + ND: Brachytherapy boost + neck dissection

## Dosimetric impact of editing autocontours on PTV coverage

In Table 3.1 characteristics of the nine patients used in this study are given. For each of the patients a physician manually contoured the primary tumor.

In addition, based on the 10 atlas patients, showing no overlap with our current study patients, autocontours for the elective neck levels and the OARs were generated using the head-and-neck algorithm in ABAS. Subsequently, two well-trained physicians (OBS1, OBS2) edited the ABAS contours for the elective neck CTVs and salivary glands to a clinically acceptable structure set. For each patient, the primary tumor and the elective neck levels were expanded by 5 mm to produce the PTV in each of the three structure sets ($PTV_{ABAS}$, $PTV_{OBS1}$, and $PTV_{OBS2}$).

For each of the nine patients a clinically acceptable IMRT plan, designated ABAS plan, was constructed based on the ABAS autocontours using Monaco (Elekta software, version 1.0). For these ABAS plans, the dose coverage of the $PTV_{OBS1}$ and $PTV_{OBS2}$ was evaluated by quantifying the percentage of the volume receiving more than 95% of the prescribed dose ($V_{95\%}$), and the minimum dose delivered to 99% of the $PTV_{OBS1,2}$, i.e., the near- minimum dose ($D_{99\%}$). To analyse effects of small geometrical differences, this analysis was repeated for PTVs that were contracted by 1, 2 and 3 mm, respectively.

To quantify the similarity between CTVs derived with ABAS and corresponding edited CTVs by either OBS1 or OBS2, Dice coefficients and mean contour distances were calculated. The Dice coefficient is a measure of the overlap of two structures $A$, $B$ [27]:

$$\text{Dice} = 2\frac{|A \cap B|}{|A| + |B|}$$

where $|A|$, $|B|$ represent the volumes of the structures and $|A \cap B|$ is the volume of intersection. The Dice coefficient can range from 0 (no overlap) to 1 (perfect overlap). The mean contour distance is defined as:

$$M(A,B) = \frac{1}{N(A)} \sum_{i=0}^{N(A)} d(A_i, B_i)$$

where $d(A_i, B_i)$ is the shortest distance from point $i$ on surface $A$ to the closest point on surface $B$, and $N(A)$ is the total number of surface points on contour $A$. In case of perfect overlap the mean contour distance = 0. A non-parametric Spearman's correlation test was used to assess the correlation between differences in $V_{95\%}$, $D_{99\%}$ and either the Dice coefficients or the mean contour distances.



### Dosimetric impact of editing OAR contours on dose in salivary glands

To study the dosimetric impact of editing autocontoured salivary glands, 18 other IMRT plans were made based on either $PTV_{OBS1}$ or $PTV_{OBS2}$ and the non-edited autocontours for the salivary glands. For each plan the mean doses in the (non-edited) autocontoured glands were compared with the mean doses in the parotid and submandibular glands edited by either OBS1 or OBS2. A non-parametric Wilcoxon signed rank test was used to determine whether the observed differences in mean doses were statistically significant.

## 3.3 Results

### Dosimetric impact of editing autocontours on PTV coverage

In Table 3.2 the *geometric* overlap between the autocontoured CTVs and the edited CTVs, as quantified by the Dice coefficient and the mean contour distance are summarized.

Between both observers no statistically significant differences were observed ($p > 0.57$). However, the edited neck CTVs were larger than the corresponding autocontoured neck CTVs, with mean differences of 9.9% and 7.5% for OBS1 and OBS2, respectively. In a non-parametric Wilcoxon signed rank test, these volume increases were statistically signifi-

**Table 3.2:** Summary of the *geometric* overlap between the autocontoured CTVs and the edited CTVs for the 9 study patients. ΔVolume quantifies the volume increase of the edited CTVs compared to the corresponding ABAS contours. For each of the parameters no statistically significant differences are observed between observer 1 and 2.

| | Observer 1 mean (range) | Observer 2 mean (range) | Wilcoxon signed rank test Observer 1 vs Observer 2 |
|---|---|---|---|
| Dice coefficient | 0.82 (0.69–0.95) | 0.81 (0.64–0.89) | $p = 0.57$ |
| Mean distance (mm) | 1.7 (0.6–4.9) | 1.7 (0.8–3.9) | $p = 0.77$ |
| ΔVolume CTV(%) | 9.9 (0.7–32.8) | 7.5 (−1.9–17.5) | $p = 0.77$ |

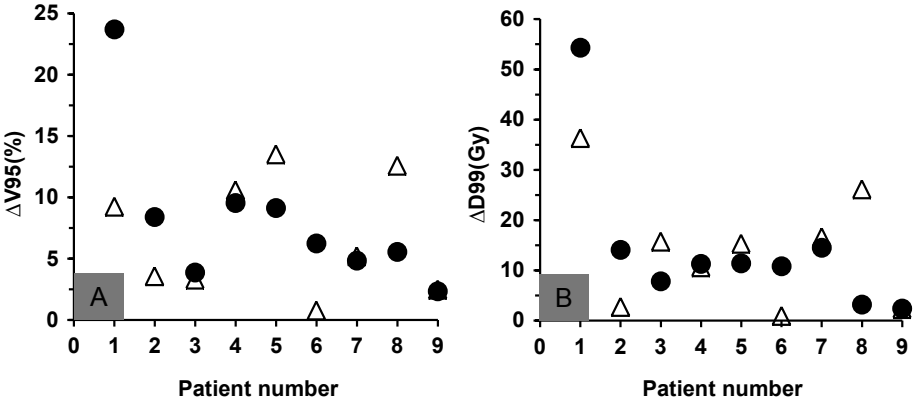



**Figure 3.1:** Reduction in PTV volume receiving at least 95% of the prescribed dose (A) and in the near-minimum dose in this volume (B) when evaluating the ABAS plan on the PTV of OBS 1 (△) and OBS2 (•) respectively.

cant ($p = 0.01$ and $p = 0.02$, respectively). During editing, the observers mainly enlarged the CTVs at the lateral borders. For the ABAS plans the mean $V_{95\%}$ of $PTV_{ABAS}$ was 97.1% ± 1.5% (1 SD). Deviations from the prescription that at least 99% of the PTV should receive 95% of the prescribed dose were mainly observed in air cavities inside the PTV and/or in superficial parts of the PTV close to the body surface (i.e., in the build-up region). When evaluating the ABAS plans for $PTV_{OBS1}$ and $PTV_{OBS2}$, $V_{95\%}$ reduced by 6.5% ± 4.4% and by 7.8% ± 6.5% respectively (see Figure 3.1A). Again, for both observers, this reduction was statistically significant ($p = 0.01$ and $p = 0.02$, respectively). The mean reduction in $D_{99\%}$ was 14.1 Gy (range 1–36 Gy) for $PTV_{OBS1}$ and 14.4 Gy (range 2–54 Gy) for $PTV_{OBS2}$ (see Figure 3.1B).

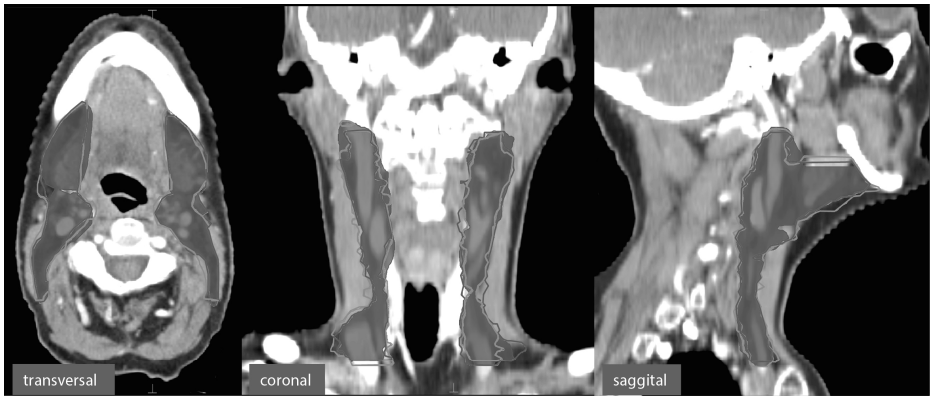



**Figure 3.2:** Example of a patient (patient 9 in Table 3.1) showing a good geometric similarity between the autocontoured neck CTV generated with ABAS (filled grey area) and the edited CTV contours of OBS1 and OBS2 (grey lines). While the two PTVs based on the edited autocontoured CTVs were not fully covered by the ABAS plan, no underdosages were observed after contracting these volumes by 2 mm.

The large deviations for patient 1 are caused by the fact that observer 2 considered partial inclusion of neck level IV in the CTV essential, thereby enlarging the volume at the caudal side by 3 cm compared to the autocontoured neck CTV.

Only for patient 6 the dose coverage for $PTV_{OBS1}$, as quantified by $V_{95\%}$ and $D_{99\%}$, was similar to the one for $PTV_{ABAS}$ without contracting the PTV contour (as described in materials and methods). In six of the eighteen cases, contracting $PTV_{OBS1,2}$ by 2 mm resulted in a similar or better dose coverage compared to $PTV_{ABAS}$ (see Figure 3.2 for an example).

But in eight cases contracting of $PTV_{OBS1,2}$ by 3 mm was still not sufficient to get a similar dose coverage. For this subgroup, the mean differences in $V_{95\%}$ in the contracted PTV were still 3.3% (range: 1–19%).

In Figure 3.3, the observed reductions in $V_{95\%}$ and $D_{99\%}$ are plotted for all study patients versus the Dice coefficients between the ABAS autocontours and the edited CTV contours by OBS1 and OBS2.

A statistically significant correlation was found in both cases ($p < 0.001$ and $p = 0.003$ respectively). The correlation between the reductions in both $V_{95\%}$ and $D_{99\%}$ and the mean contour distances (not shown) was not statistically significant ($p = 0.06$ and 0.57, respectively). A large spread around the linear correlations is observed, implying that relatively high Dice coefficients (>0.8) or small mean contour distances (<1 mm) can still be accompanied by large PTV underdosages.

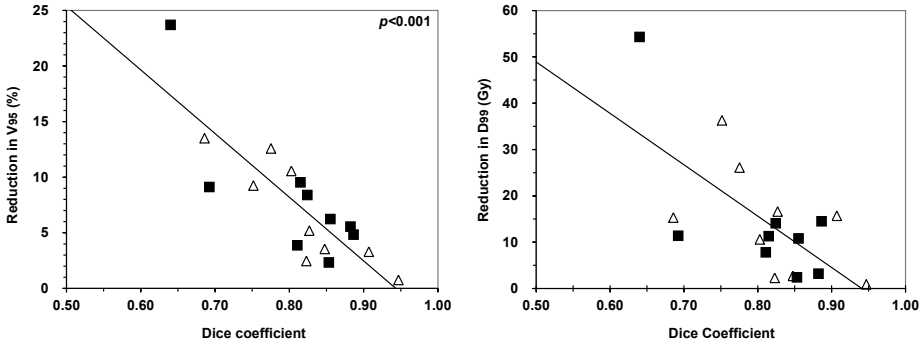



**Figure 3.3:** Correlation between the Dice coefficient and the reduction in PTV receiving at least 95% of the prescribed dose (A) and the reduction in near-minimum dose in this volume (B) when evaluating the ABAS plan on the PTV of OBS 1 (△) and OBS2 (■) respectively. The plotted trend lines are tested statistically significant in a Spearman's correlation test.

## Dosimetric impact of editing OAR contours on dose in salivary glands

When evaluating IMRT plans based on either $PTV_{OBS1}$ or $PTV_{OBS2}$ and non-edited ABAS salivary glands, the mean doses in the edited parotid glands differed by −0.8 Gy ± 1.1 Gy (1SD) and in the edited submandibular glands by −0.2 Gy ± 0.7 Gy (1SD). These relatively small differences were not statistically significant ($p = 0.12$ for OBS1 and $p = 0.29$ for OBS2).

# 3.4 Discussion

In this study we have demonstrated that the use of ABAS autocontours generally results in large underdosages in the target volumes. The fact that in eight of the eighteen cases contracting the PTVs by 3 mm did not yet result in adequate dose coverage indicates that local differences are often quite large. Editing of the neck CTVs resulted in seventeen of the eighteen cases in a volume increase of the CTV, indicating either that observers tend to enlarge a neck CTV rather than diminish it, or that the output of ABAS is too small

for the majority of cases. The main differences are observed at the lateral borders of level II, III and IV. According to the atlas guidelines [1,12,13] this lateral border should be at the exact transition of the fat space and the sterno-cleido mastoid muscle, but in this study we observed that physicians tend to include up to 3 mm of this muscle in the CTV as well. In addition, some adjustments in autocontours, like for patient 1, were driven by the physician's judgement of patient specific information. To improve the accuracy of the output of ABAS, a future revision of the applied atlases might be considered. However, before changing any atlas we plan to investigate a larger patient cohort first because intersubject variations are large in head-and-neck cancer patients and wrongly adapted atlases could cause systematic errors. Moreover, adaptation of the atlases should preferentially be based on adaptation of the consensus guidelines.

In a recent study, Tsuji et al. [28] evaluated the dosimetric impact of autosegmentation for adaptive IMRT for head-and-neck cancer patients. In this study, structures in the mid-treatment CT were derived by transforming manually delineated contours in the initial planning CT, based on deformable registration between both CT sets. When using the autocontoured structures to create an adaptive treatment plan, a significantly lower mean dose coverage of the manually delineated GTV and the CTV in the mid-treatment CT was observed. A similar result was found in our study, when autocontouring a new patient based on ten atlas patients. In contrast to the findings of Tsuji et al., our data showed a statistically significant correlation between Dice coefficients and target coverage. However, in clinical practice, this correlation has a very limited value, since Dice coefficients and mean contour distances cannot be determined before editing the autocontours. So, the amount of underdosages of the target cannot be assessed beforehand. In addition, the scatter in the results shown in Figure 3.3 indicates that both Dice coefficients and mean contour distances have little predictive value regarding the impact on dose coverage. In some publications the similarity of two volumes is considered good if a Dice coefficient is 0.7 or higher (see e.g., [29]). But for a tumor volume, such a geometric overlap does not necessarily result in acceptable dose coverage. Even for Dice coefficients of 0.8, underdosages in the PTV of up to 11 Gy were observed in this study.

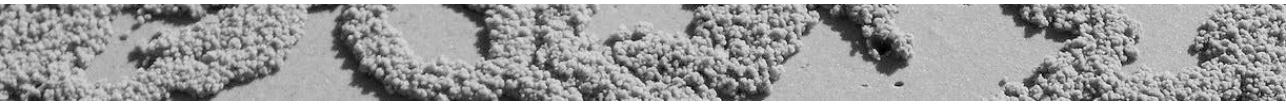
Editing of salivary glands, taking 13 min on average [25], generally leads to small differences in the mean dose in these organs, indicating that this might be omitted, especially when the mean dose is expected to be much lower or higher than the objective. Also re-planning, e.g., after anatomical changes during treatment, can generally be performed using non-edited salivary gland contours because it is already known from the initial plan whether the mean doses are critical. Editing is only essential when mean doses in OARs get close to their objectives.

## 3.5 Conclusion

Editing of CTV neck contours generated by ABAS is essential to avoid large (local) underdosages in the target volume. When applying highly conformal IMRT plans with steep gradients toward OARs and the surrounding tissues, small geometrical differences in the CTV volume will often have a large dosimetric impact on the dose coverage of the target. However, editing salivary glands has only a small impact on the mean dose in these OARs and might often be avoided. Dice coefficients and mean contour distances are useful to quantify the geometric overlap of different structure sets, but are of limited value to predict dosimetric target coverage and plan quality.

chapter 4

# Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head-and-neck IMRT and implications for VMAT



Peter W.J. Voet, RTT, Sebastiaan Breedveld, PhD, Maarten L.P. Dirkx, PhD,
Peter C. Levendag, MD, PhD, Ben J.M. Heijmen, PhD

*Medical Physics; 39(8):4858–65 (2012)*

# Abstract

**Purpose:** To quantify improved salivary gland sparing for head-and-neck cancer patients using intensity-modulated radiotherapy (IMRT) plans based on integrated computerized optimization of beam orientations and intensity profiles. To assess if optimized nonzero couch angles also improve VMAT plans.

**Methods:** Our in-house developed algorithm iCycle was used for automated generation of multicriterial optimized plans with optimized beam orientations and intensity profiles, and plans with optimized profiles for preselected beam arrangements. For 20 patients, five IMRT plans, based on one 'wishlist', were compared: i), ii) seven and nine beam equiangular coplanar plans ($iCycle_{7equi}$, $iCycle_{9equi}$), iii), iv) nine-beam plans with optimized coplanar and noncoplanar beam orientations ($iCycle_{copl}$, $iCycle_{noncopl}$) and v) a nine-beam coplanar plan with optimized gantry angles and one optimized couch rotation ($iCycle_{couch}$). VMAT plans without and with this optimized couch rotation were evaluated.

**Results:** $iCycle_{noncopl}$ resulted in the best salivary gland sparing, while $iCycle_{couch}$ yielded similar results for 18 patients. For $iCycle_{7equi}$, submandibular gland NTCP values were on average 5% higher. $iCycle_{9equi}$ performed better than $iCycle_{7equi}$. $iCycle_{copl}$ showed further improvement. Application of the optimized couch angle from $iCycle_{couch}$ also improved NTCP values in VMAT plans.

**Conclusions:** iCycle allows objective comparison of competing planning strategies. Integrated optimization of beam profiles and angles can significantly improve normal tissue sparing, yielding optimal results for $iCycle_{noncopl}$.

# 4.1 Introduction

Intensity-modulated radiotherapy (IMRT) with computer-optimized beam profiles has become standard care for curative treatments of head-and-neck cancer patients. Until now, it has been common practice that a dosimetrist selects the applied number of beams and their orientations in a trial-and-error process. This is not straightforward, especially for complex treatment sites involving many organs-at-risk (OARs). The quality of the final plan may thus depend on the skills and experience of the planner. Alternatively, a template solution is often used per target site to derive clinical IMRT plans.

For head-and-neck cancer patients, published data on the impact of (computer) optimized and patient-individualized beam arrangements on the quality of treatment plans are limited. In most comparative studies for head-and-neck cancer IMRT very few patients were included and equispaced coplanar beam arrangements were used as reference. Meedt et al. [30] reported on coplanar beam angle-optimization (BAO) for one laryngeal-cancer patient. BAO plans with six or seven beams were compared to a manually defined nine-field plan and a plan with 15 equiangular beams. Target coverage was improved for both BAO plans, and mean doses in OARs were reduced. Djajaputra et al. [31] used an extensive search from different sets of beams to derive an optimal beam angle configuration. For one head-and-neck cancer patient they showed that the use of seven or nine noncoplanar beams with optimized angles improved the plan quality over that of nine equiangular coplanar beams. Wang et al. [32] investigated the effectiveness of noncoplanar BAO in ten paranasal sinus carcinoma patients. Five beams with optimized beam angles generally performed better than plans with nine equiangular coplanar beams. Nutting et al. [33] determined the added value of noncoplanar BAO plans over coplanar BAO plans for 6 parotid gland carcinoma patients. Each plan consisted of 3 or 4 fields. For the noncoplanar plans, the inhomogeneity in the PTV was greater, and a higher normal brain volume received a dose of at least 54 Gy. They reported no significant advantage for noncoplanar BAO. A limitation of their study was the small number of beams used; most studies in the literature suggest that at least 7 beams are required to achieve an optimal IMRT plan.

In previous publications on BAO our group mainly investigated in-house developed algorithms for beam angle, weight and shape optimization in 3DCRT, with a special focus on liver SBRT. For IMRT, only a simple segmented approach was investigated [34], or beam angle and beam profile optimization were performed consecutively [35]. The latter approach started with selection of optimal beam angles for 3DCRT, using an in-house BAO algorithm. Intensity profiles for these angles were then optimized with a commercial treatment planning system (TPS). It was demonstrated that with a noncoplanar beam arrangement OARs could be better spared than with coplanar beams only. BAO is now

clinically applied for 3DCRT plan generation in liver SBRT [36]. Recently, we developed a multicriterial plan optimization algorithm for *integrated* beam angle and beam intensity optimization, called iCycle [8,37,38] Outputs of iCycle are Pareto-optimal IMRT plans with optimized beam setups. With Pareto-optimal we mean that none of the objectives in the applied 'wishlist' (see materials and methods section) can be improved any further without deteriorating one or more constraints or higher prioritized objectives. For preselected fixed beam arrangements, iCycle can also be used to optimize beam profiles only. In this study, iCycle was used to systematically compare IMRT plans for various beam angle selection strategies in 20 randomly selected head-and-neck cancer patients. Each plan aimed at maximum salivary gland sparing while obtaining the prescribed high tumor coverage. Due to the proximity of several OARs to the PTV, we expected that the use of noncoplanar beam arrangements could be beneficial, like for liver SBRT patients [35]. Variations in the procedure for beam angle selection resulted in seven- and nine-beam equiangular coplanar plans, coplanar plans with optimized gantry angles, fully noncoplanar plans with optimized couch and gantry angles, and coplanar plans with optimized gantry angles for various non-zero couch angles. To assess the effect of these nonzero couch angles in VMAT plans, we used a commercial TPS to generate two different VMAT plans for 2 patients. For the first VMAT plan we used the regular 0° couch angle; for the second plan the couch angle corresponding to the nine-beam coplanar plan generated with iCycle, showing the lowest mean NTCP in the salivary glands, was used.



## 4.2 Materials and Methods

### Study patients and clinical treatment plans

Twenty recently treated head-and-neck cancer patients with various tumor sites were randomly selected from our clinical database (Table 4.1).

For treatment, IMRT plans were made using Monaco (version 2.04, Elekta). The intention was to draw up the best clinically acceptable treatment plan for each patient on the basis of target coverage, sparing of organs-at-risk, and number of beams. The main focus was put on minimization of the mean dose in the salivary glands (both parotids and submandibular glands) with the intend to maintain salivary flow as much as possible [39,40], while assuring that at least 98.5% of the PTV was treated with at least 95% of the

**Table 4.1:** Characteristics of patients used in this study.

| Patient | Tumor type | TNM | Dose interval | Followed by |
|---|---|---|---|---|
| 1 | Palatum Molle Ca. | T2N2cM0 | 46 Gy | Cyberknife boost |
| 2 | Tonsillar fossa Ca. | T2N0M0 | 46 Gy | Cyberknife boost |
| 3 | Tonsillar fossa Ca. | T3N0M0 | 46 Gy | Cyberknife boost |
| 4 | Tonsillar fossa Ca. | T2N0M0 | 46 Gy | Cyberknife boost |
| 5 | Tonsillar fossa Ca. | T2N0M0 | 46 Gy | Cyberknife boost |
| 6 | Tonsillar fossa Ca. | T2N0M0 | 46 Gy | Brachytherapy boost |
| 7 | Tonsillar fossa Ca. | T1N2aM0 | 46 Gy | Brachytherapy boost |
| 8 | Tonsillar fossa Ca. | TxN0M0 | 46 Gy | Brachytherapy boost |
| 9 | Tonsillar fossa Ca. | T2N0M0 | 46 Gy | Brachytherapy boost |
| 10 | Oropharynx Ca. | T3N0M0 | 46 Gy | Cyberknife boost |
| 11 | Oropharynx Ca. | T2N0M0 | 46 Gy | Cyberknife boost |
| 12 | Oropharynx Ca. | T1N0M0 | 46 Gy | Brachytherapy boost |
| 13 | Oropharynx Ca. | T2N0M0 | 46 Gy | Brachytherapy boost |
| 14 | Base of tongue Ca. | T4aN1M0 | 46 Gy | Cyberknife boost |
| 15 | Base of tongue Ca. | T1N1M0 | 46 Gy | Cyberknife boost |
| 16 | Base of tongue Ca. | cT1N2bM0 | 46 Gy | Brachytherapy boost |
| 17 | Cheek Ca. | T4aN0M0 | 46 Gy | 24 Gy primary tumor |
| 18 | Floor of mouth Ca. | pT4aN0M0 | 66 Gy PORT | |
| 19 | Parotid Ca. | pT2N0M0 | 66 Gy PORT | |
| 20 | Nasopharynx Ca. | T1N2M0 | 70 Gy | Brachytherapy boost |

PORT: post-operative RT

prescribed dose and keeping the maximum dose in critical serial organs below a fixed threshold value. (e.g., 50 Gy for the spinal cord). In most cases this resulted in a seven-beam IMRT plan.

## iCycle: Integrated beam angle and profile optimization

iCycle is a novel in-house developed algorithm for integrated beam angle and beam profile optimization [8,37,38]. The algorithm is described in detail in [8]. Here we will briefly summarize the main characteristics. Core of iCycle is the 2-phase ε-constraint (2pεc) algorithm for generating optimal IMRT plans for preselected (fixed) beam arrangements. This algorithm has a multicriterial optimization approach that generates a single Pareto-

**Table 4.2:** Applied wishlist for all study patients. SMG = submandibular gland

| **Constraints** | | | |
|---|---|---|---|
| | Volume | Type | Limit |
| | PTV | maximum | 107% of prescribed dose |
| | Spinal cord | maximum | 38 Gy (*48 Gy) |
| | Unspecified tissue | maximum | 107% of prescribed dose |

*48 Gy for patients 18–20.

| **Objectives** | | | | | |
|---|---|---|---|---|---|
| Priority | Volume | Type | Goal | Sufficient | Parameters |
| 1 | PTV | LTCP | 1 | 0.5 | $\alpha = 0.8$ |
| 2 | Parotid/SMG | mean | 39 Gy | | |
| 3 | Parotid/SMG | mean | 20 Gy | | |
| 4 | Parotid/SMG | mean | 10 Gy | | |
| 5 | Parotid/SMG | mean | 2 Gy | | |
| 6 | PTV shell 1 cm | maximum | 35 Gy | | |
| 7 | PTV shell 2 cm | maximum | 30 Gy | | |
| 8 | PTV shell 3 cm | maximum | 25 Gy | | |
| 9 | PTV shell 4 cm | maximum | 20 Gy | | |
| 10 | PTV shell 5 cm | maximum | 15 Gy | | |

optimal plan for the preselected beam angles. Basis is a user-defined 'wishlist' containing the clinical plan objectives that all have an ascribed priority and hard constraints to be strictly obeyed during plan generation (see Table 4.2 for an example).

In iCycle, beam profile optimization is integrated in an iterative procedure for selection of optimal beam directions. Starting with zero beams, new beams with optimal directions are consecutively selected from a list of candidate beams and added to the plan. For selection of a next beam orientation, first, all candidate directions not yet selected are temporarily added to the configuration established in the previous iteration, and the IMRT optimization problem is solved. In the end, the orientation with the best score is added to the beam configuration. For each orientation added this results in one new Pareto-optimal IMRT plan. Identical wishlists can be used for groups of patients, e.g., for all head-and-neck cancer patients with the main focus on sparing of salivary glands, making plan generation fully automatic and user independent.

The 2pec algorithm consists of two phases. First, all objectives are consecutively minimized within hard constraints, trying to reach their predefined goal values (but not

below), starting with the highest prioritized objective (see Table 4.2). After minimization, the objective is turned into a constraint, applying the attained value. This method ensures that in the following minimizations of lower prioritized objectives, the attained value of the just optimized objective will at least be maintained. Another advantage is that for each patient one can start with the same wishlist with demanding objectives. In the second phase, objectives that could have been reduced below their defined goal as used in the first phase are sequentially further minimized, again starting with the highest prioritized objective. Except for objectives for the target dose, in the second phase all objectives are minimized to the full extent. Minimization of the target objective is stopped at a user-defined 'sufficient' value, accepting small deviations from 100% coverage of the PTV with 95% of the prescribed dose to leave room for minimization of lower prioritized objectives.

Apart from the wishlist, the only input that iCycle requires to generate plans is a prescribed number of beams, and a list with candidate beam directions that may be added to the plan. This list may be restricted for generation of coplanar plans with couch angle 0°, but it may also contain a large set of orientations with various couch and gantry angles that can be delivered without crashes between the gantry and the couch or patient.

As mentioned above, in the process of generating a plan with the prescribed number of beams, Pareto-optimal plans will also be generated for smaller beam numbers. The plan with the highest number of beams has the best plan quality regarding the highest prioritized objective that is minimized to its full extent (see above). However, differences with plans with fewer beams may be small, making a plan with fewer beams attractive because of reduced QA workload or treatment time. On the other hand, if the evaluation of plan quality as a function of beam number shows that more beams than initially prescribed would result in a better plan quality, iCycle can easily be instructed to generate more plans with higher beam numbers.

## iCycle plan generation for the study patients

The applied wishlist is depicted in Table 4.2. It contains hard constraints on the maximum allowed dose in the PTV, spinal cord and unspecified tissue. For patients 1–17, the maximum allowed dose in the spinal cord was set to 38 Gy, i.e., lower than our clinically used constraint, to leave sufficient room for optimizing the subsequent boost plan. The objective with priority 1 was to achieve adequate target coverage. The target dose was optimized by minimizing the Logarithmic Tumor Control Probability (LTCP) as described by Alber et al. [41]. A prescription dose and an alpha parameter, affecting the penalty given to underdosed voxels in the tumor, are inputs to derive the LTCP. On

the basis of the first five study patients, alpha was set at 0.8, which ensured that at least 98.5% of the PTV volume received at least 95% of the prescribed dose for all treatment plans. Objectives 2–5 aimed at a balanced reduction of the mean dose in each of the individual salivary glands. Since the PTV objective was higher prioritized, target dose-coverage never reduced. Finally, to achieve a steep dose fall-off outside the target volume, maximum dose objectives were defined for a set of shells with 0.3 cm thickness at 1, 2, 3, 4 and 5 cm of the PTV.

For coplanar BAO, beams were selected from 72 equiangular spaced candidate beams. For noncoplanar BAO, we added noncoplanar candidate beams with a 10° separation, resulting in approximately 320 beams in total. The coarser separation of the noncoplanar candidate beams reduced the time required for BAO without compromising plan quality. In this study, we used a maximum noncoplanar beam angle in cranial direction of 45°. Beams outside this range would never be selected due to the applied maximum dose objectives outside the PTV (objectives 6–10 in Table 4.2). We also omitted candidate beams at both sides of the sphere that could result in collisions between the gantry and the patient or the treatment couch. Consequently, the list of feasible candidate beams used by iCycle for noncoplanar BAO depends on target site and treatment machine.

Using iCycle, the following plans were derived for the patients in this study:

1. a seven-beam coplanar equiangular plan at couch angle 0° ($iCycle_{7equi}$).
2. a nine-beam coplanar equiangular plan at couch angle 0° ($iCycle_{9equi}$).
3. a nine-beam coplanar plan with optimized gantry angles at couch angle 0° ($iCycle_{copl}$).
4. a nine-beam noncoplanar plan with optimized gantry and couch angles ($iCycle_{noncopl}$).
5. a nine-beam coplanar plan with optimized gantry angles and an optimal couch angle ($iCycle_{couch}$). This plan was selected from 9 sequentially optimized iCycle plans, generated for fixed couch angles at 340°, 345°, 350°, 355°, 5°, 10°, 15°, 20°, or 0° (i.e., $iCycle_{copl}$), based on the lowest mean NTCP for the salivary glands.

$iCycle_{couch}$ was investigated because this plan requires one single couch rotation at the start of treatment only; next, all beams can be delivered without the technicians having to enter the treatment room in between. In contrast, treatment of fully noncoplanar plans is more labor-intensive and time-consuming, as technicians have to enter the treatment room to manually execute prescribed couch rotations.

### Impact of couch angle on VMAT plans

For the two study patients who benefited most from using an optimized couch angle for the coplanar plan (i.e., showing the largest difference in mean salivary gland NTCP when comparing $iCycle_{copl}$ and $iCycle_{couch}$), we investigated whether VMAT plans could be improved by using this optimized couch angle, instead of the commonly used zero couch angle. Single, full-arc VMAT-plans were generated with Monaco for both situations. Differences in salivary gland sparing were quantified.

### Comparison of treatment plans

Like in our clinically used plans, plan evaluation was mainly focused on the achieved salivary gland sparing. We compared mean doses in each of the glands. Moreover, these mean doses were converted into NTCP values, using published dose-response models for parotid [39] and submandibular glands [40]. Both NTCP models are based on the probability of Grade 4 toxicity (i.e., salivary flow rate <25% of baseline pre-radiotherapy). A two-sided Wilcoxon matched-pair signed-rank test [42] was used to compare the different iCycle plans for the patient group.

## 4.3 Results

### Comparison of iCycle plans

With the applied wishlist, at least 98.5% of the PTV received 95% of the prescribed dose for all plans. For 18 of the 20 patients this applied for even more than 99% of the PTV. Due to applied hard constraint in the wishlist, observed maximum doses in the spinal cord remained below 38 Gy for patients 1–16, and below 48 Gy for patients 17–20. For patient 20, the observed maximum dose for brainstem was 48 Gy, while the maximum doses in the optical chiasm, optical nerves and eyes were less than 40 Gy. In all other patients the maximum doses in the brainstem were below 38 Gy and the maximum doses in the optical chiasm, optical nerves and eyes below 25 Gy.

Figure 4.1 compares mean salivary gland doses for $iCycle_{noncopl}$ with the corresponding doses in the other iCycle plans.

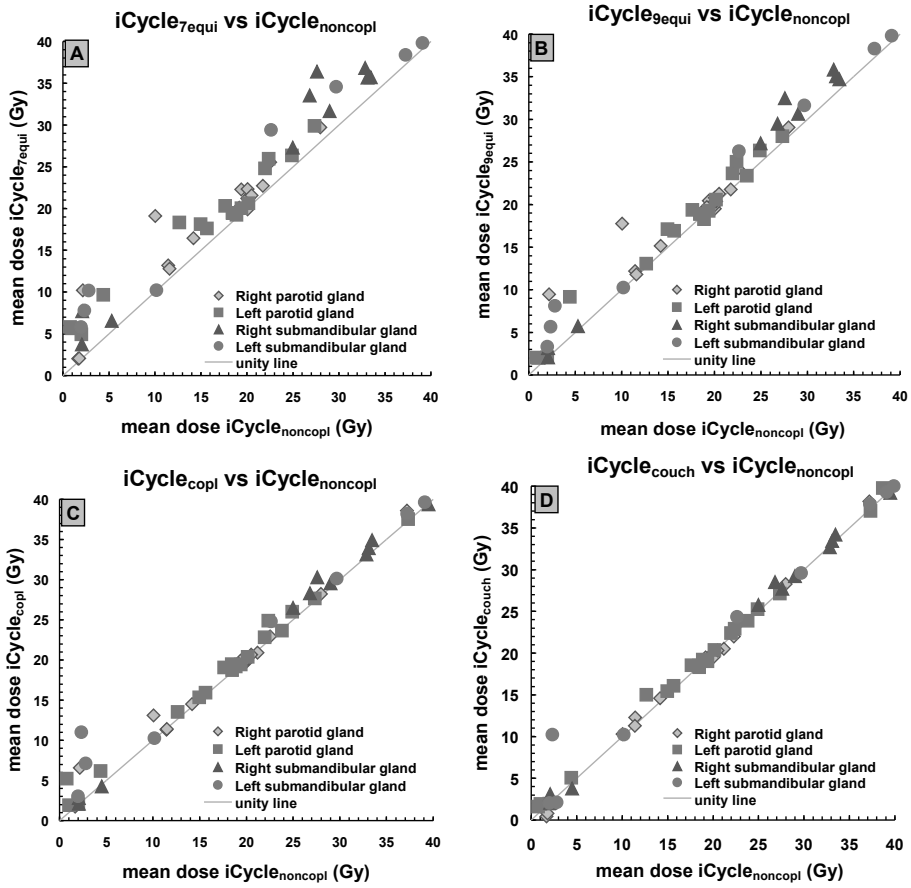



**Figure 4.1:** Mean salivary gland doses for all coplanar techniques compared to $iCycle_{noncopl}$.

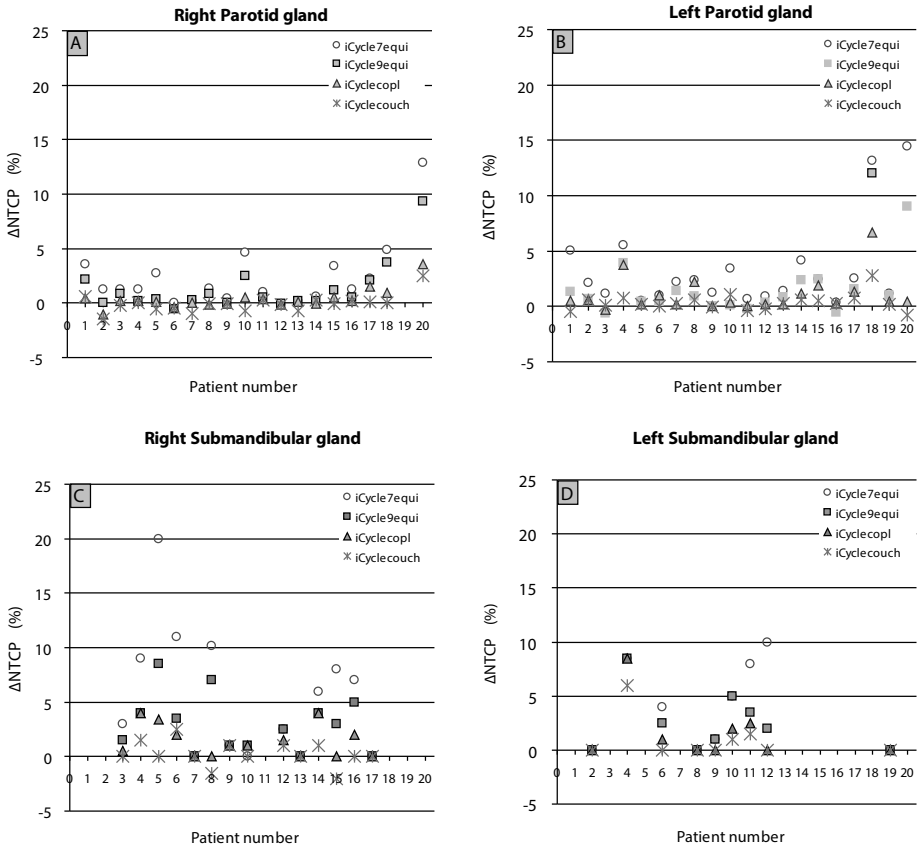



**Figure 4.2:** Differences in NTCP values for all plans relatively to $iCycle_{noncoplan}$, visualized per salivary gland. NTCP values of glands that could not be spared, because they were entirely enclosed in the PTV, were omitted.

Figure 4.2 shows the NTCP differences between $iCycle_{noncopl}$ and the other plans.

In Table 4.3, these differences are summarized. The lowest mean salivary gland doses were observed for $iCycle_{noncopl}$. Differences with $iCycle_{7equi}$, $iCycle_{9equi}$, and $iCycle_{copl}$ were statistically significant ($p < 0.001$).

Compared to $iCycle_{couch}$, only the differences in mean submandibular gland dose were statistically significant ($p = 0.05$). For $iCycle_{7equi}$, $iCycle_{9equi}$ and $iCycle_{copl}$, the observed increases in mean salivary gland doses compared to $iCycle_{noncopl}$ also trans-

**Table 4.3:** Differences between $iCycle_{noncopl}$ and all other plans in mean salivary gland dose and NTCP, averaged over the 20 study patients, for both parotid and submandibular glands. For $iCycle_{noncopl}$, the mean parotid and submandibular gland doses were 17.5 Gy (range: 0.8–38.7 Gy) and 24.5 Gy (range: 2.0–41.0 Gy), respectively. Mean NTCPs were 7% (range: 1–47%) and 30% (range: 10–65%).

| | | $iCycle_{7equi}$ | $iCycle_{9equi}$ | $iCycle_{copl}$ | $iCycle_{couch}$ |
|---|---|---|---|---|---|
| Parotid glands | Δ mean dose*(Gy) | 2.5 | 1.4 | 0.7 | 0.1 |
| | SD (Gy) | 2.2 | 1.9 | 1.2 | 0.7 |
| | Range (Gy) | −0.2–9.1 | −0.6–7.7 | −0.8–4.4 | −1.2–2.3 |
| | *p*-value* | <0.001 | <0.001 | <0.001 | 0.203 |
| | Δ NTCP*(%) | 2.8 | 1.6 | 0.7 | 0.1 |
| | SD (%) | 3.5 | 2.7 | 1.4 | 0.8 |
| | Range (%) | −0.2–14.4 | −0.7–12.0 | −1.1–6.7 | −1.5–2.8 |
| | *p*-value* | <0.001 | <0.001 | 0.005 | 0.629 |
| Submandibular glands | Δ mean dose*(Gy) | 3.1 | 1.7 | 1.2 | 0.6 |
| | SD (Gy) | 2.6 | 1.5 | 2.0 | 1.7 |
| | Range (Gy) | −0.5–8.8 | 0.0–5.3 | −1.1–8.6 | −1.5–7.9 |
| | *p*-value* | <0.001 | <0.001 | <0.001 | 0.05 |
| | Δ NTCP*(%) | 5.0 | 2.8 | 1.5 | 0.5 |
| | SD (%) | 5.1 | 2.7 | 2.0 | 1.5 |
| | Range (%) | 0–20.0 | 0–8.5 | 0–8.5 | −2.0–6.0 |
| | *p*-value* | <0.001 | <0.001 | 0.001 | 0.085 |

lated in statistically significant increased NTCPs (Table 4.3). $iCycle_{7equi}$ yielded higher mean salivary gland doses and NTCPs than each of the nine-beam techniques. BAO for $iCycle_{copl}$ resulted in better salivary gland sparing than achieved with $iCycle_{9equi}$.

The observed mean NTCP differences between different planning techniques are small. However, one should keep in mind that the mean NTCP values for $iCycle_{noncopl}$ were 7% for the parotid glands and 30% for the submandibular glands. Therefore, in a relative sense, $iCycle_{noncopl}$ had a far lower risk for damaging parotid glands than e.g., $iCycle_{7equi}$. Moreover, the gain of applying a more advanced planning technique was patient dependent (see Table 4.3). For example, the mean difference in NTCP for the submandibular glands between $iCycle_{7equi}$ and $iCycle_{9equi}$ was 2.2%; whereas, in patient 5 the use of nine-beams reduced the NTCP for the right submandibular gland from 40% to 29%. For the parotids, the mean NTCP difference between $iCycle_{9equi}$ and $iCycle_{copl}$ was 0.9%, but in patient 20 the NTCPs reduced from 53% to 47% and 44%, respectively.

Compared to $iCycle_{copl}$, use of the optimal couch angle ($iCycle_{couch}$) improved the NTCP for parotid glands by up to 4% and for submandibular glands by up to 3%. For 18 of the 20 patients, NTCPs for $iCycle_{couch}$ were similar to the values for $iCycle_{noncopl}$.

**Table 4.4:** Mean doses in salivary glands (Gy) for VMAT plans with a couch angle of $0^0$ and with the optimized couch angle from $iCycle_{couch.}$ Submandibular glands for patient 18 were omitted in the table since they were fully included in the PTV.

| | Patient 4 | | Patient 18 | |
|---|---|---|---|---|
| | VMAT 0° | VMAT 350° | VMAT 0° | VMAT 340° |
| Right parotid | 22.7 | 20.6 | 13.9 | 9.2 |
| Left parotid | 26.6 | 24.8 | 42.5 | 39.1 |
| Right submandibular gland | 44.3 | 42.7 | – | – |
| Left submandibular gland | 35.6 | 35.2 | – | – |

## Calculation times for generating iCycle plans

The calculation time for the beam angle optimization phase in iCycle scales linearly with the number of candidate beams. For the multicriterial optimization, the required time grows to the power of three with the number of beams. Moreover, it depends on the number of constraints and objectives in the wishlist. For this part, calculation times are similar for coplanar and noncoplanar iCycle plans.

Using a modern eight-core server, calculation times were typically 3 hours for the nine-beam coplanar BAO plans and 12 hours for the noncoplanar plans using a voxel size of 0.98 x 0.98 x 2.5$mm^3$, identical to the CT resolution. For small structures like optical nerves, all voxels were included in the optimization. For larger structures, approximately 5000 voxels were selected based on a Hammersley sequence sampling. For all individual plans, a calculation gridsize of 5 x 10 mm was used (based on an MLC with 1 cm leaf width).

## Impact of couch optimization on VMAT plans

Patients 4 and 18 had the largest reduction in mean salivary gland NTCP when changing from $iCycle_{copl}$ to $iCycle_{couch}$. For these patients, the optimal couch angle was 340° and 350° respectively. Table 4.4 shows salivary gland doses for VMAT plans with couch angles 0° and 340°/350°. Like observed for $iCycle_{couch}$, use of the optimal couch angle improved the salivary gland sparing for VMAT plans as well.

## 4.4 Discussion

To our knowledge, this is the first systematic study on the impact of beam angles on salivary gland sparing in head-and-neck cancer patients, based on integrated computerized optimization of beam angles and intensity profiles. By using a single wishlist with plan objectives and constraints for all patients and techniques, plan generation became fully user independent, ensuring objective plan comparisons. Nine beam plans performed better than seven beam plans, and coplanar plans with optimized gantry angles had better gland sparing than equiangular plans. Noncoplanar plans with optimized gantry and couch angles had the lowest gland NTCPs, although differences with coplanar plans with an optimized couch angle were negligible for most patients.

iCycle generates multicriterial optimized treatment plans. Other algorithms for multicriterial plan optimization have been evaluated clinically, showing a benefit for plan quality and plan efficiency [43,44]. But in these studies the number of treatment beams and their orientations were still manually selected. The regular labor-intensive trial-and-error process for selecting beam orientations and number of beams in IMRT treatment planning may result in variations in plan quality, depending on the skills and experience of the dosimetrist. Class solutions for beam arrangements may be suboptimal for individual patients. iCycle is an algorithm for fully automated generation of plans with optimized beam angles and intensity profiles. Therefore, the workload for generating plans is almost negligible. Prior to the start of the investigations described in this paper, for 10 of the 20 selected patients we compared the clinical plan as generated with the Monaco TPS with the coplanar iCycle plan with optimized gantry angles ($iCycle_{copl}$). For each of these patients, salivary gland sparing in the iCycle plan was superior. On average, the mean gland NTCP reduced from 29.2% to 19.4%.

iCycle has been developed in a research setting and not for clinical use; a segmentation algorithm is currently lacking. However, it turned out that using the optimal beam angles and mean gland doses as established with iCycle as input, Monaco could be steered to generate a clinically acceptable plan that was almost identical to the iCycle plan with respect to the obtained gland NTCPs. The procedure for converting an iCycle plan into a Monaco plan has recently been automated. As a result, high quality IMRT plans for head-and-neck cancer patients can now be generated in two automated steps. The first step is generation of a plan in iCycle; the second is the conversion into a clinically deliverable Monaco plan. We recently performed a prospective study for head-and-neck patients, showing a clear benefit for this automated plan generation approach [45]. In that study, we took into account more relevant critical structures, like oral cavity, larynx and swallowing muscles, during plan optimization.

We demonstrated that iCycle$_{couch}$ plans were sometimes superior to the corresponding plans for couch angle 0°. This was especially true in unilateral neck treatments, in which rotation of the couch often resulted in a free projection of the contralateral salivary glands, yielding reductions in delivered gland doses. As demonstrated for two patients, also with VMAT the optimal non-zero couch angles resulted in lower salivary gland doses than plans for couch angle 0°.

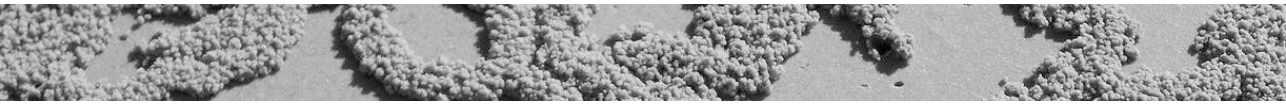
All iCycle plans in this study were automatically generated based on one identical wishlist. This allows for an objective comparison of competing treatment strategies for individual patients. For none of them, clinically relevant plan improvements were observed when using more than nine beams. Patients with unilateral targets showed very little improvement in plan quality after adding beams eight or nine. However, adding extra beams never deteriorated plan results. The investigations revealed that the gain of more complex or time-consuming treatment techniques was patient dependent. As planning with iCycle is fully automated, a possible future clinical application is to generate for each patient nine-beam coplanar plans for couch angle 0° and several non-zero angles, and a nine-beam noncoplanar plan. Based on the differences in NTCP and involved treatment time, one could then choose for each individual patient the most appropriate treatment technique.

## 4.5 Conclusions

iCycle is a novel plan optimization algorithm for user-independent generation of treatment plans with optimized beam angles and intensity profiles, allowing objective comparison of competing planning strategies. For head-and-neck cancer patients, plan generation with iCycle can result in improved salivary gland sparing with a largely reduced planning workload. Observed sparing was best for fully noncoplanar plans. VMAT with an optimized couch angle can improve salivary gland sparing compared to the commonly applied zero couch angle.

chapter 5

# Toward fully automated multicriterial plan generation: a prospective clinical study



Peter W.J. Voet, RTT, Maarten L.P. Dirkx, PhD, Sebastiaan Breedveld, PhD,
Dennie Fransen, RTT, Peter C. Levendag, MD, PhD, Ben J.M. Heijmen, PhD

*International Journal of Radiation Oncology Biology Physics; 85(3):866–72 (2013).*

## Summary

A randomized, prospective study was performed comparing fully automatically and manually generated IMRT plans for head-and-neck cancer patients. Of the two plans, the treating physician selected the plan to be used for treatment. In 97% of cases, the automatically generated plan was chosen because of the superior quality. In addition to better sparing of the large majority of critical structures, improved target coverage was achieved for most patients.



## Abstract

**Purpose:** To prospectively compare plans generated with iCycle, an in-house developed algorithm for fully automated multicriterial intensity modulated radiotherapy (IMRT) beam profile and beam orientation optimization, with plans manually generated by dosimetrists with the clinical treatment planning system.
**Methods and Materials:** For 20 randomly selected head-and-neck cancer patients with various tumor locations (of whom 13 received sequential boost treatments) we offered the treating physician the choice between an automatically generated iCycle plan and a manually optimized plan using standard clinical procedures. While iCycle used a fixed 'wish-list' with hard constraints and prioritized objectives, the dosimetrists manually selected the beam configuration and fine-tuned the constraints and objectives for each IMRT

plan. Dosimetrists were not informed in advance whether a competing iCycle plan was made. The two plans were simultaneously presented to the physician, who then selected the plan to be used for treatment. For the patient group, we also quantified differences in PTV coverage and sparing of critical tissues.
**Results:** In 32/33 plan comparisons the physician selected the iCycle plan for treatment. This highly consistent preference for automatically generated plans was mainly caused by improved sparing for the large majority of critical structures. With iCycle, the normal tissue complication probabilities (NTCPs) for parotid and submandibular glands were reduced by 2.4% ± 4.9% (maximum: 18.5%, $p = 0.001$) and 6.5% ± 8.3% (maximum: 27%, $p = 0.005$), respectively. The reduction in mean oral cavity dose was 2.8 Gy ± 2.8 Gy (maximum: 8.1 Gy, $p = 0.005$). For swallowing muscles, esophagus and larynx, the mean dose reduction was 3.3 Gy ± 1.1 Gy (maximum: 9.2 Gy, $p < 0.001$). In addition, for 15 of the 20 patients, the target coverage was improved as well.
**Conclusions:** In 97% of cases, the automatically generated plan was selected for treatment because of superior quality. Apart from improved plan quality, automatic plan generation is economically attractive because of reduced workload.

## 5.1 Introduction

Intensity modulated radiotherapy (IMRT) is commonly applied as curative radiotherapy treatment for head-and-neck cancer patients. As the input for computerized optimization of the beam profiles, the number of beams, their configuration, the constraints, and objectives need to be defined by the dosimetrist. It is often difficult for dosimetrists to select these input parameters and to assess whether and how a plan can be further optimized by modifying them for a second attempt. This can result in a time consuming trial-and-error process. Moreover, the realized plan quality can depend highly on the complexity of the case, the time available for plan generation, and the skills and ambition of the dosimetrist.

In the present study we prospectively compared 2 plans for a group of randomly selected head-and-neck patients: 1 generated by dosimetrists using the clinical treatment planning system (Monaco, Elekta AB, Sweden), and 1 using automatic plan generation by iCycle. iCycle [8,37,38] is an in-house developed algorithm for multicriterial optimization of beam profiles and gantry angles. The patients were treated according to the best plan, as selected by the treating physician. Plans were compared based on physicians' preferences and quantitative assessments.

**Table 5.1:** Patient characteristics.

| Patient | Tumor type | TNM | Dose intervals (Gy) | Followed by |
|---|---|---|---|---|
| 1 | Hypopharynx Ca. | T4aN2bM0 | 0–66 | |
| 2 | Oral Cavity Ca. | T4aN0M0 | 0–46 + 46–66 | |
| 3 | Tonsil Ca. | T2N0M0 | 0–46 | brachytherapy |
| 4 | Larynx Ca. | T4N2bMx | 0–46 + 46–70 | |
| 5 | Larynx Ca. | T4aN1M0 | 0–46 + 46–70 | |
| 6 | Nasopharynx Ca. | T1N1Mx | 0–46 + 46–70 | |
| 7 | Oral Cavity Ca. | T2N0M0 | 0–46 | Cyberknife boost |
| 8 | Base of tongue Ca. | T2N2bM0 | 0–46 | Cyberknife boost |
| 9 | Hypopharynx Ca. | T2N2bM0 | 0–46 + 46–70 | |
| 10 | Sinus piriformis Ca. | TxN2bMx | 0–46 + 46–66 | |
| 11 | Tonsil Ca. | T3N2bMx | 0–46 | Cyberknife boost |
| 12 | Hypopharynx Ca. | T4aN0M0 | 0–46 + 46–70 | |
| 13 | Hypopharynx Ca. | T2N0M0 | 0–46 + 46–70 | |
| 14 | Tongue Ca. | T4N2cM0 | 0–46 + 46–66 | |
| 15 | Larynx Ca. | T3N0M0 | 0–46 + 46–70 | |
| 16 | Base of tongue Ca. | T2N0M0 | 0–46 + 46–70 | |
| 17 | Lip Ca. | TxN0M0 | 0–45 + 45–62.5 | |
| 18 | Left neck(Mesothelioma) | | 0–44 | |
| 19 | Oral Cavity Ca. | T2N0M0 | 0–66 | |
| 20 | Larynx Ca. | T3N0M0 | 0–46 + 46–70 | |



## 5.2 Methods and Materials

### Patients

An overview of the 20 patients included in our study is provided in Table 5.1.

The treatments for the indicated dose intervals were delivered using an Elekta Synergy accelerator and 6 MV photon beams. For 13 of the 20 patients, a sequential boost followed the initial treatment ≤45 or 46 Gy. Because boosts delivered with brachytherapy or Cyberknife were excluded in our analyses, the treating physicians made in total 33 plan comparisons and selections.

Before plan generation, the treating physician contoured the CT-scan of the patient, largely assisted by the autocontouring program ABAS [24]. Apart from the clinical target

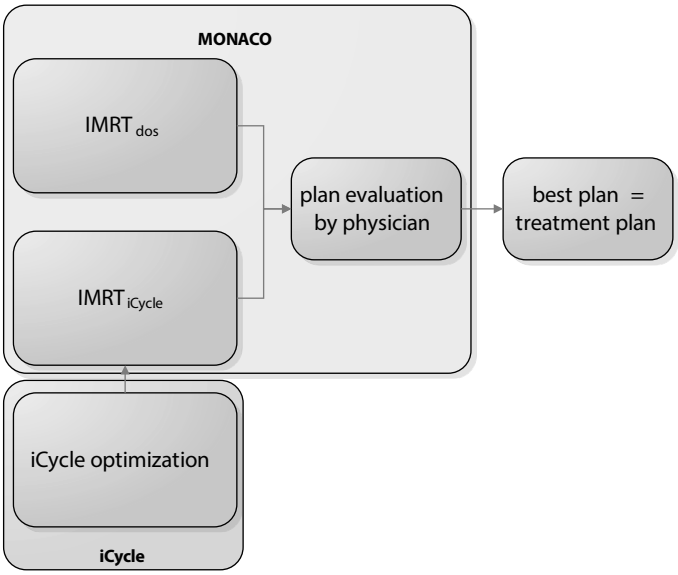



**Figure 5.1:** Workflow of this study. Note that both $IMRT_{dos}$ and $IMRT_{iCycle}$ were presented to the treating physician in the same program (Monaco).

volumes (CTV), the parotid and submandibular glands, oral cavity, swallowing muscles, first centimeter of the esophagus [4], larynx, spinal cord, and brain stem were delineated. For planning, CTVs were extended with a margin of 5 mm to generate planning target volumes (PTV).

## Study design

The design of our study is schematically presented in figure 5.1.

During the study period, the standard protocol for IMRT of head-and-neck patients included "manual" generation of a treatment plan ($IMRT_{dos}$) by a dosimetrist using our clinical treatment planning system Monaco, version 2.04. For each IMRT plan the beam configuration was manually selected and the constraints and objectives were fine tuned. For the study patients, a second plan ($IMRT_{iCycle}$) was generated, using an automated plan generation with iCycle [8] (see also next section). Neither the involved dosimetrist,

nor the treating physician, knew in advance whether a competing iCycle plan would be developed.

iCycle is not commissioned for clinical use. Therefore, 1 of the investigators (PV) used Monaco for conversion of each iCycle plan into a highly similar, clinically deliverable plans ($IMRT_{iCycle}$). First, a Monaco plan template was automatically generated, including the optimized beam angles and achieved plan parameters for organs-at-risk (OARs) in iCycle. With minimal user interference, these values could then be reproduced in Monaco within a 1-Gy difference for all OARs. To prevent any bias the investigator (PV) did not have access to the $IMRT_{dos}$ plan while generating the $IMRT_{iCycle}$ plan. $IMRT_{dos}$ and $IMRT_{iCycle}$ were together presented to the treating physician for plan selection, using identical lay-outs of plotted dose distributions, and dose-volume histograms. Seven dosimetrists, experienced in IMRT planning, were involved in generating the standard plans using Monaco. Seven head-and-neck radiation oncologists entered the patient data and evaluated competing $IMRT_{iCycle}$ and $IMRT_{dos}$ plans to select the best plan for treatment. In addition to physician preference, we compared plans using quantitative analyses of the achieved PTV coverage and OAR sparing. For patients with a sequential boost, not delivered with brachytherapy or the Cyberknife, summed dose distributions were evaluated. The target dose was quantified by the dose delivered to 99% of the target volume ($D_{99\%}$). For the spinal cord and brainstem, maximum doses were compared. The mean doses were evaluated for the other OARs. For salivary glands, the mean doses were converted into normal tissue complication probabilities (NTCP) values, using published dose-response models for parotid [39] and submandibular glands [40]. Both models are based on the probability of Grade 4 toxicity (i.e., salivary flow rate <25% of baseline pre-radiotherapy). For the group of patients, 2-sided Wilcoxon matched-pair signed-rank tests were used to derive the statistical significance of observed differences in plan parameters.

For both $IMRT_{dos}$ and $IMRT_{iCycle}$, the required hands-on planning time was logged. Time for optimization, segmentation, and dose calculation, which did not require any manual interference, was excluded.

## iCycle plan generation

iCycle is an algorithm for automated multicriterial optimization of both beam orientations and IMRT fluence profiles. Its features have been previously described in detail [8]. In brief, the basis for plan optimization is an a-priori defined 'wishlist' containing hard constraints to be strictly obeyed and prioritized objectives. Per patient category, a fixed wishlist is used. Treatment plan generation is fully automated (i.e., without any

user interaction such as tweaking of the objectives). The core of iCycle is the 2-phase ε-constraint algorithm for generating Pareto-optimal plans for preselected (fixed) beam arrangements [37,38]. In generated Pareto-optimal plans, none of the objectives can be improved any further without deteriorating ≥1 constraints or higher prioritized objectives. In iCycle, beam profile optimization is integrated in an iterative procedure for the selection of optimal beam directions. Starting with 0 beams, new beams with optimal directions are consecutively selected from a list of candidate beams and added to the plan. For the selection of the next beam orientations, all candidate beams not yet selected are temporarily added to the configuration established in the previous iteration, and the IMRT optimization problem is solved. In the end, the orientation with the best score is added to the beam configuration. For each orientation added, this results in 1 new Pareto-optimal IMRT plan. The addition of a new beam improves plan quality regarding the highest prioritized objective that can still be improved on. In the present study, a total of 9 directions were used for plan generation for each patient, resulting in Pareto-optimal plans with 9, 8, …, 1 beams. With more than 9 beams, clinically relevant improvements were not observed. Using a minimum gain of 0.5 Gy per added beam in at least one of the objectives, we determined the number of beams for generating the $IMRT_{iCycle}$ plans. In the present study, iCycle was used to generate Pareto-optimal coplanar plans at couch angle 0°. The beam directions were automatically selected from 72 equi-spaced candidate directions (5° separation), including 0°.

The applied wishlist for the first dose interval of each patient is listed in Table 5.1. It was established using the clinical protocol trial runs of iCycle for a small group of patients, and discussions with 2 of the 7 involved clinicians. iCycle has a mechanism to reduce, if possible, the objective functions in the wishlist to values lower than the indicated goal values, with an accent on the objectives with the highest priorities. PTV coverage had the greatest priority. To enforce that the PTV dose delivered to 99% of the PTV was ≥95% of the prescribed dose, the logarithmic tumor control probability (LTCP) was used [37,41]. For a homogeneous target dose equal to the prescribed dose, LTCP equaled 1. A lower target dose would result in a strongly enhanced LTCP, and greater target doses would decrease the LTCP to a minimum value of 0 for a clinically infeasible, infinite dose. To guarantee good target coverage, an alpha value of 0.8 was used. A higher alpha parameter puts a higher penalty on cold spots in the PTV. A reduction of the mean salivary gland doses was performed using a multi-level approach [24], i.e. by repeated use of the objective function with decreasing priorities and goal values (Table 5.2).

iCycle first tried to minimize the mean dose in each of the salivary glands to 39 Gy (corresponding to an NTCP of about 50%, objective 2), and then to 20 Gy (NTCP of about 10%, objective 3). Before minimizing the dose in the salivary glands even further (priorities 9 and 11), the mean oral cavity dose (priority 4), the maximum dose in spinal

**Table 5.2:** Applied wishlist for first dose interval of each patient, containing hard constraints, and prioritized objectives.

| **Constraints** | | | |
|---|---|---|---|
| | Volume | Type | Limit |
| | PTV | max | 107% of prescribed dose |
| | Spinal Cord | max | 48 Gy* |
| | Unspecified tissue | max | 107% of prescribed dose |

| **Objectives** | | | |
|---|---|---|---|
| Priority | Volume | Type | Goal |
| 1 | PTV | ↓LTCP | 1 |
| 2 | Parotid/SMG | ↓mean | 39 Gy |
| 3 | Parotid/SMG | ↓mean | 20 Gy |
| 4 | Oral Cavity | ↓mean | 39 Gy |
| 5 | Spinal Cord/Brain Stem | ↓max | 30 Gy |
| 6 | External ring** | ↓max | 90% of prescribed dose |
| 7 | Larynx + swallowing muscles | ↓mean | 75% of prescribed dose |
| 8 | PTV shell 1 cm*** | ↓max | 75% of prescribed dose |
| 9 | Parotid/SMG | ↓mean | 10 Gy |
| 10 | PTV shell 4 cm*** | ↓max | 40% of prescribed dose |
| 11 | Parotid/SMG | ↓mean | 2 Gy |

Abbreviations: LTCP = logarithmic tumor control probability; PTV = planning target volume; SMG = submandibular gland.

* For patients treated with a sequential boost technique, a spinal cord constraint of 38 Gy was applied to leave room for optimizing the boost plan.

** Structure 2 cm distance interior from patient surface, preventing high superficial doses in incident beams.

*** PTV shells 1 and 4 cm from PTV to control dose gradient outside PTV.

cord and brainstem (priority 5) and the mean dose in the larynx and swallowing muscles (priority 7) were optimized. When for instance the dose in the oral cavity would not be considered first, the salivary glands might be spared at the cost of an unacceptable high dose in the oral cavity. The volumes 'External ring, PTV shell 1 cm and PTV shell 4 cm' aimed at reduction of the entrance dose, and steering the dose gradient outside the PTV, respectively.

For patients with a sequential boost (Table 5.1) we used the same wishlist for the boost. Only the constraints for the spinal cord and brainstem were adjusted according to the dose interval.

# 5.3 Results

$IMRT_{iCycle}$ and $IMRT_{dos}$ plans consisted of 6–9 beams (average 8.6), and 5–9 beams (average 7.3), respectively. In 32 of 33 plan comparisons, the physician selected $IMRT_{iCycle}$ for treatment, in the vast majority of cases because of reduced dose delivery to OARs, with negligible loss in PTV coverage, equal coverage, or even improved coverage. For patient 11, $IMRT_{iCycle}$ was selected notwithstanding the higher OAR doses. In this case, $IMRT_{dos}$ had a too low PTV coverage, which was avoided in $IMRT_{iCycle}$.

The obtained OAR doses are compared in figure 5.2.

In the case of a patient having plans for two dose intervals (see Table 5.1), data for the summed plan were provided. For the iCycle plans the mean NTCP was reduced by 2.4% (maximum: 18.5%, $p = 0.001$) for parotid glands and by 6.5% (maximum: 27%, $p = 0.005$) for submandibular glands. The mean dose in the oral cavity reduced by on average 2.8 Gy (maximum: 8.1 Gy, $p = 0.005$). A mean dose reduction in swallowing muscles, esophagus

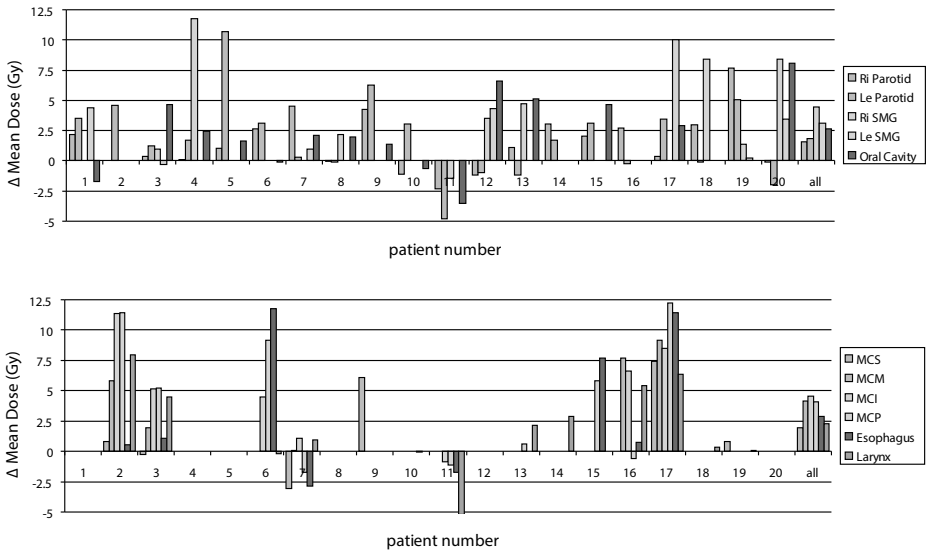



**Figure 5.2:** Differences in OAR mean doses between $IMRT_{dos}$ and $IMRT_{iCycle}$ for the 20 study patients with at the right end ('all') the patient group averages. Positive values indicate lower OAR doses for $IMRT_{iCycle}$. Results for OARs that could not be spared, because they are completely embedded in the PTV, were omitted. MCS = musculus constrictor superior; MCM = musculus constrictor medius; MCI = musculus constrictor inferior; MCP = musculus constrictor cricopharyngeus. SMG = submandibular gland.

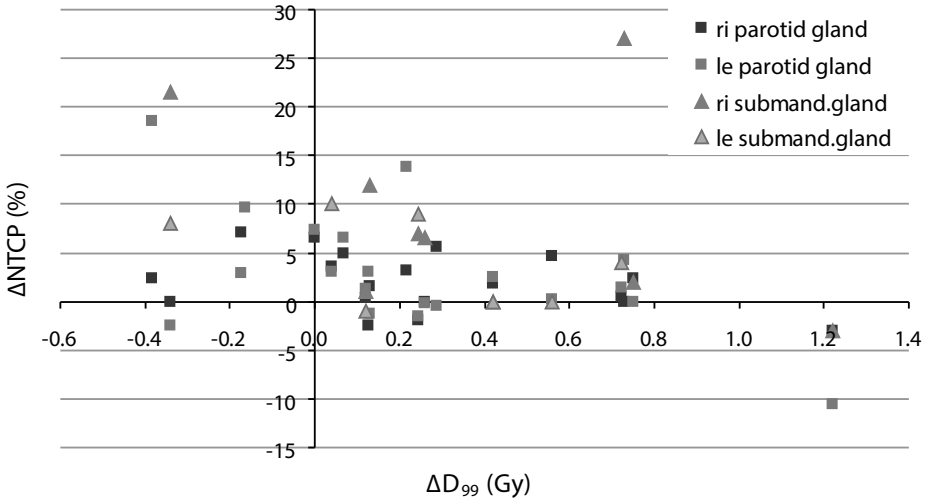



**Figure 5.3:** Differences between $IMRT_{dos}$ and $IMRT_{iCycle}$ in $D_{99\%}$ and NTCP values for the individual salivary glands, for the 20 study patients. Results for salivary glands that could not be spared, because they were completely embedded in the PTV, were omitted. For points in the upper right quadrant, $IMRT_{iCycle}$ showed both a better target coverage and improved OAR sparing. SMG = submandibular gland.



and larynx was observed of 3.3 Gy ± 1.1 Gy (maximum 9.2 Gy, $p < 0.001$). For $IMRT_{iCycle}$ the maximum doses in the spinal cord and in the brain stem were well below tolerance with average values of 34.8 Gy and 21.8 Gy, respectively. Relative to $IMRT_{dos}$ the maximum doses in the spinal cord and the brain stem were reduced by 3.3 Gy ± 4.0 Gy ($p = 0.04$) and 1.1 Gy ± 6.4 Gy ($p = 0.531$), respectively.

Figure 5.3 illustrates the balance between target coverage and NTCP for the salivary glands.

Overall, the target coverage was slightly better for the $IMRT_{iCycle}$ plans. The mean improvement in $D_{99\%}$ was 0.24 Gy ± 0.4 Gy ($p = 0.07$). Of salivary glands, 75% had lower NTCPs with the $IMRT_{iCycle}$ plan (upper left and right quadrants in figure 5.3); 59% of the glands are in the upper right quadrant, indicating that $IMRT_{iCycle}$ had the lowest mean gland dose, and improved target coverage.

For each patient, we estimated both for $IMRT_{dos}$ and $IMRT_{iCycle}$ the probability that the function of at least one parotid gland could be preserved, and the probability that at least one submandibular gland could be spared. These probabilities, P, were calculated according to $P = (1 - NTCP_{left} \times NTCP_{right})$. Figure 5.4 shows clearly that this probability was greater for the $IMRT_{iCycle}$ plans, except for patient 11 (see figure 5.2). In cases in which $IMRT_{dos}$ already has a high sparing probability (i.e., close to 100%), there is

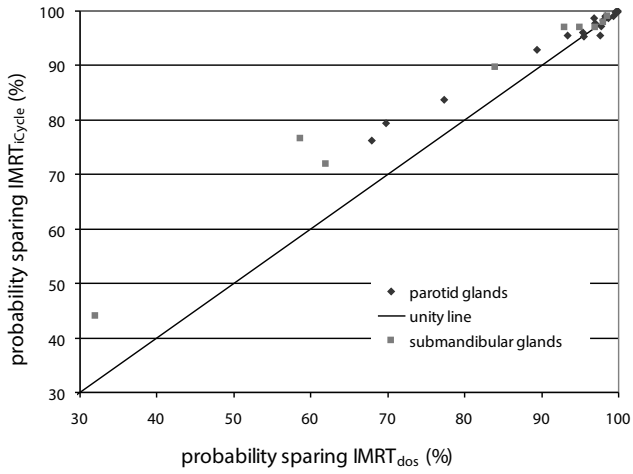



**Figure 5.4:** Probability of sparing at least one of the salivary glands when using either IMRT$_{iCycle}$ or IMRT$_{dos}$.

naturally limited room for improvement with IMRT$_{iCycle}$. However, for patients with a lower probability of sparing at least one of the glands, sparing could be substantially enhanced using IMRT$_{iCycle}$.

The average hands-on time spent by dosimetrists on generating an IMRT$_{dos}$ plan was 3.3 hours (range 1–6.5 hours). As described in the Methods and Materials section, generating an iCycle plan is fully automated. Conversion into a Monaco plan to arrive at the clinically applicable IMRT$_{iCycle}$ plan, as performed in this study, took on average 1.5 hours (range 1–2.5 hours).

## 5.4 Discussion

To our knowledge, this is the first prospective clinical study evaluating the possibilities for automated treatment plan generation. We included patients with a broad range of head-and-neck tumor sites in the present study. For each study patient and dose interval, a plan generated with iCycle, our in-house algorithm for *automated*, multicriterial optimization of beam angles and profiles, was compared with a plan manually generated by the dosimetrists, following the routine clinical protocol. In 32 out of 33 cases, physicians selected the automatically generated plan for treatment. Generally,

$IMRT_{iCycle}$ plans showed higher quality for a very broad range of plan parameters (see figure 5.2). This could explain the high consistence in the physician's preferences for the automatically generated plans. Only once did a physician select the $IMRT_{dos}$ plan for treatment (patient 9; 0–46 Gy). In this case, $IMRT_{iCycle}$ showed larger hotspots (107% of the prescribed dose) in parts of the PTV adjacent to the parotid glands. That these hotspots resulted in reduced parotid gland mean doses was considered less important. With some minor adjustments in the standard wishlist used by iCycle, these hotspots could have been avoided [37]. For the sequential boost of this patient, the physician preferred the $IMRT_{iCycle}$ plan for treatment. As demonstrated in previous studies, iCycle can also be used to fully optimize noncoplanar IMRT plans [8]. Compared with coplanar beam arrangements, minor improvements in OAR sparing were observed for most head-and-neck cancer patients. Because delivery of noncoplanar plans would be more labor intensive and time consuming, as technicians must enter the room to manually execute the prescribed couch rotations, coplanar beam arrangements were used in the present study.

Using the applied wishlist, with multi-level objectives for the salivary glands, the sparing of different OARs was better balanced with $IMRT_{iCycle}$. For some patients, this resulted in less sparing for a lower-prioritized OAR to the benefit of higher prioritized ones. As an example, in patient 1, the mean dose in the oral cavity was 1.7 Gy greater with the $IMRT_{iCycle}$ plan; however, the sparing of the parotids and left submandibular gland was improved (mean dose reductions of 2.2 Gy, 3.5 Gy and 4.4 Gy, respectively).

For the patient group, the differences between $IMRT_{dos}$ and $IMRT_{iCycle}$, appeared to be small. Nevertheless, individual patients could benefit highly from automated plan generation with iCycle, as shown by differences in NTCP values up to 18.5% for parotid glands and up to 27% for submandibular glands (see figure 5.3).

The basis of automated plan generation with iCycle is the automated steering of plan optimizations, using a wishlist with plan criteria that is identical for all patients. The observed high clinician preference and the favorable plan parameters for the $IMRT_{iCycle}$ plans have demonstrated that automated plan generation can be successfully performed. As mentioned also in the Methods and Materials section, the wishlist used in this study was compiled using discussions with 2 of the 7 clinicians participating in the study. However, with the input from only 2 of the clinicians, it was possible to generate a wishlist that served the needs of the others.

Other recent reports also addressed the possibility of (semi-)automatic treatment planning. Thieke et al. [43] described a multicriterial optimization technique with an interactive plan navigation tool. They automatically generated a database with Pareto-optimal IMRT plans. By interactively exploring this database the optimal treatment plan was identified. Teichert et al. [46] enhanced the algorithm such that the user could also navigate between plans with different beam configurations. In contrast to iCycle, the number of treatment beams and the beam configurations were manually predefined.

Craft et al. [44] reported on multicriterial plan generation for glioblastoma and pancreatic cancer patients, showing clear benefits in treatment planning efficiency and plan quality compared with manually generated treatment plans. They also used a predefined beam arrangement.

Another limitation of both multicriterial optimization approaches is that manually selecting an optimal plan from the Pareto front might be very difficult and subjective, especially when many OARs are involved.

Zhang et al. [47] described an algorithm for automatic intensity-modulated radiation treatment planning for lung cancer. In their work the beam angle configuration was selected from an expert database, depending on tumor size and position. A drawback of this approach is that it remains unclear whether this configuration is optimal for the next patient.

As mentioned in the Methods and Materials section, iCycle currently cannot be used as a direct method for generating clinical plans. Therefore, in the present study, fully automatically generated iCycle plans were 'manually' converted into corresponding Monaco plans ($IMRT_{iCycle}$) of highly similar quality. From the superiority of the $IMRT_{iCycle}$ plans compared with the manually generated plans ($IMRT_{dos}$), we have concluded that fully automated planning with iCycle will be feasible, once the system has been prepared for direct clinical use. For the present study, however, there was also an advantage with the conversion of the iCycle plans. Because both plans were available in Monaco, the layout of the presented dose distributions and dose-volume histograms was equal for $IMRT_{iCycle}$ and $IMRT_{dos}$, avoiding plan selection bias.

In an ongoing project, we aim at the direct use of iCycle plans in the clinic, avoiding conversion to Monaco plans. We also have started to routinely apply iCycle for treatment of head-and-neck cancer patients as described in the present report (i.e., the use of iCycle for generation of an IMRT plan, followed by conversion to a deliverable plan with highly similar quality in Monaco). More comparative studies are needed to investigate the importance of automated treatment planning for other tumor sites.

## 5.5 Conclusions

Compared to plans generated by dosimetrists with Monaco, the plans automatically generated with iCycle were highly preferred by treating physicians. The results from quantitative plan analyses were in line with this preference. By routinely applying iCycle,

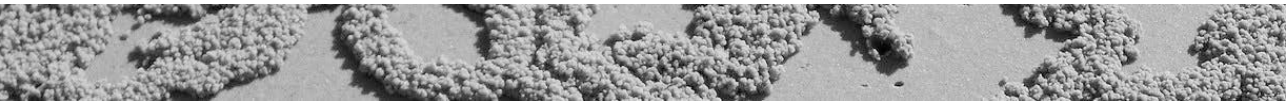
optimized, patient specific treatment plans can be generated for large groups of patients with minimal user interaction.

## 5.6 Acknowledgements

The authors would like to thank the dosimetrists who generated the $IMRT_{dos}$ plans and the physicians for entering patients in the study and evaluating the treatment plans.

chapter 6

# Automated generation of IMRT treatment plans for prostate cancer patients with metal hip prostheses: comparison of different planning strategies



Peter W.J. Voet RTT, Maarten L.P. Dirkx, PhD,
Sebastiaan Breedveld PhD, Ben J.M. Heijmen PhD

*Medical Physics; 40(7):071704 (2013)*

# Abstract

**Purpose**: To compare IMRT planning strategies for prostate cancer patients with metal hip prostheses.
**Methods and Materials:** All plans were generated fully automatically (i.e., no human trial-and-error interactions) using iCycle, our in-house developed algorithm for multi-criterial selection of beam angles and fluence profiles, allowing objective comparison of planning strategies. For 18 prostate cancer patients (eight with bilateral hip prostheses, ten with a right-sided unilateral prosthesis), two planning strategies were evaluated: i) full exclusion of beams containing beamlets that would deliver dose to the target after passing a prosthesis ($IMRT_{remove}$), and ii) exclusion of those beamlets only ($IMRT_{cut}$). Plans with optimized coplanar and noncoplanar beam arrangements were generated. Differences in PTV coverage and sparing of organs-at-risk (OARs) were quantified. The impact of beam number on plan quality was evaluated.
**Results:** Especially for patients with bilateral hip prostheses, $IMRT_{cut}$ significantly improved rectum and bladder sparing compared to $IMRT_{remove}$. For 9-beam coplanar plans, rectum $V_{60\ Gy}$ reduced by 17.5% ± 15.0% (maximum 37.4%, $p = 0.036$) and rectum $D_{mean}$ by 9.4% ± 7.8% (maximum 19.8%, $p = 0.036$). Further improvements in OAR sparing were achievable by using noncoplanar beam set-ups, reducing rectum $V_{60\ Gy}$ by another 4.6% ± 4.9% ($p = 0.012$) for noncoplanar 9-beam $IMRT_{cut}$ plans. Large reductions in rectum dose delivery were also observed when increasing the number of optimized beam directions in the plans. For bilateral implants, the rectum $V_{60\ Gy}$ was 37.3% ± 12.1% for coplanar 7-beam plans and reduced on average by 13.5% (maximum 30.1%, $p = 0.012$) for 15 directions.
**Conclusions:** iCycle was able to automatically generate high quality plans for prostate cancer patients with prostheses. Excluding only beamlets that passed through the prostheses ($IMRT_{cut}$ strategy) significantly improved OAR sparing. Noncoplanar beam arrangements and, to a larger extent, increasing the number of treatment beams, further improved plan quality.

# 6.1 Introduction

With the aging of the population, an increasing number of patients with osteoarthritis will get a hip replacement. In The Netherlands, the incidence of hip implants increased by 50% between 1995 and 2005, and is expected to increase by another 149% till 2030 [48]. Consequently, the number of prostate cancer patients with a metal hip implant referred for radiotherapy is expected to grow extensively. For these patients, the search space for beam angle selection is severely limited. The American Association of Physicists in Medicine (AAPM) Task Group 63 recommends using beam arrangements that completely avoid dose delivery through metal hip prostheses [49]. Especially for patients with bilateral prostheses, this generally results in less favorable dose distributions since primarily only anterior and posterior beams can be selected. Recently, our group described a new planning technique that allowed the use of default beam set-ups for patients with prostheses [50]. Tumor dose delivery through prostheses was avoided by manually delineating prosthesis–avoiding-volumes (PAVs) outside the patient's body contour, and putting strict dose constraints on them. Compared to the prostheses-avoiding technique as suggested by AAPM, better sparing of the organs-at-risk (OARs) and/or an improved target coverage was obtained. A limitation of this work was that only one patient with a unilateral and one with a bilateral prosthesis were evaluated. In addition, it is not clear to what extent the plan quality was affected by the used default set of coplanar beam orientations.

In this study, different treatment planning strategies for patients with metal implants were systematically evaluated for 18 prostate cancer patients. For plan generation we used iCycle, an in-house developed multicriterial plan generation algorithm with integrated beam angle and beam intensity optimization [8]. Due to the deterministic approach, iCycle can be used for objective comparison of different planning strategies. Plan generation with iCycle is fully automated, i.e. there is no manual trial-and-error tweaking by a dosimetrist to generate an acceptable plan. In a recent prospective clinical study we reported for a group of head-and-neck cancers patients superiority of plans automatically generated with iCycle, compared to plans made by dosimetrists [45]. In this study, the automated plan generation allowed objective comparison of different strategies for coping with metal hip implants in prostate cancer patients. For each patient, we investigated how close plan quality could get to the plan for the patient without prostheses. Plans with both coplanar and noncoplanar beam arrangements were evaluated. In addition, the impact of the number of treatment beams on plan quality was quantified.

# 6.2 Materials and Methods

## Patients

From our clinical database, 18 prostate cancer patients treated in our department with radiotherapy between January 2009 and January 2012 were selected. Eight of them had bilateral hip prostheses; the other ten patients had a right-sided unilateral hip prosthesis. Prior to treatment, a computed tomography (CT) scan was acquired using a Somatom Sensation Open multi-slice CT scanner (Siemens, Erlangen, Germany). For some patients, also an MR scan was made and registered to the planning CT. The prostate, seminal vesicles, bladder and rectum were delineated according to the Radiation Therapy Oncology Group (RTOG) guidelines [51]. In addition, the metal hip prosthesis was contoured. All streak artifacts in the CT images, caused by the metal implants, were given an electron density override of 1.0 to achieve a more realistic and accurate dose calculation [52]. The planning target volume (PTV) was defined by expanding the prostate with a margin of 5 mm (7 mm in caudal direction) and the seminal vesicles with a uniform margin of 8 mm. A total dose of 78 Gy was prescribed to the PTV. For treatment planning, 10 MV photon beams were used, as we clinically use to do for all our pelvic patients treated with IMRT.

## iCycle plan generation

iCycle is an in-house developed algorithm for multicriterial optimization of both IMRT fluence profiles and coplanar or noncoplanar beam arrangements, described in detail in [8,37,38]. Starting point for plan optimization is an a-priori defined "wishlist", containing hard constraints to be strictly obeyed and prioritized objectives, and a list with candidate beam orientations. Per patient category, one fixed wishlist is used. For coplanar plans, 72 equi-angular spaced beams at couch angle zero were used as candidate beams. For noncoplanar plan generation, feasible noncoplanar candidate beams were added to this list. Simulation at the treatment unit showed that the maximum couch rotation that avoided any collision was ±20°. In practice, this resulted in a total of 166 noncoplanar candidate beams.

Treatment plan generation is fully automated, i.e., without any user interaction such as tweaking of the objective weights and mutual priorities. Core of iCycle is the 2-phase-

**Table 6.1:** Applied wishlist for all study patients. LTCP = Logarithmic Tumor Control Probability.

| **Constraints** | | | |
|---|---|---|---|
| | Volume | Type | Limit |
| | PTV | max | 104% of prescribed dose |
| | PTV shell 50 mm | max | 60% of prescribed dose |
| | Unspecified tissue | max | 104% of prescribed dose |

| **Objectives** | | | |
|---|---|---|---|
| Priority | Volume | Type | Goal |
| 1 | PTV | ↓LTCP | 0.5 |
| 2 | PTV shell 5 mm | ↓max | 95% of prescribed dose |
| 3 | Rectum | ↓mean | 33% of prescribed dose |
| 4 | PTV shell 15 mm | ↓max | 70% of prescribed dose |
| 5 | PTV shell 25 mm | ↓max | 50% of prescribed dose |
| 6 | External ring | ↓max | 50% of prescribed dose |
| 7 | Bladder | ↓mean | 60% of prescribed dose |
| 8 | Unspecified tissue | ↓mean | 10 Gy |

ε-constraint (2pεc) algorithm for generating Pareto-optimal IMRT plans for preselected (fixed) beam arrangements. In these Pareto-optimal plans, none of the objectives can be improved any further without deteriorating higher prioritized objectives or violating constraints. In iCycle, beam profile optimization is integrated in an iterative procedure for selection of beam orientations. Starting with zero beams, new beams are consecutively selected from a list of candidate beams. First, all candidate directions not yet selected are temporarily added to the configuration established in the previous iteration, and the IMRT optimization problem is solved. In the end, the best scoring orientation is added to the beam configuration, resulting in one new Pareto-optimal plan [8].

Table 6.1 shows the wishlist that was used for all plan generations in this study. PTV coverage had the highest priority. To enforce that the dose delivered to 99% of the PTV was at least 95% of the prescribed dose, we used the Logarithmic Tumor Control Probability (LTCP) [37,41]. For a homogeneous target dose, equal to the prescribed dose, LTCP equals 1. A lower target dose results in a strongly increased LTCP, while higher target doses decrease the LTCP to a minimum value of 0 for a clinically infeasible, infinite dose. To guarantee good target coverage, an alpha value of 0.8 was used. A higher alpha puts a higher penalty on cold spots in the PTV. Objectives with decreasing priorities were set to

the mean rectum and bladder doses. A constraint was used to avoid any doses higher than 60% of the prescribed dose at 50 mm from the PTV. In addition, to achieve steep dose fall-off outside the PTV and suppress high doses in the rectum, objectives for three shells around the PTV (with priorities 2, 4 and 5) were defined. To minimize entrance dose we applied an objective for an external ring, 2 cm inside the patient body contour. In the first phase of the 2pec optimization, objectives are optimized to their goal, if achievable, and then constrained. In the second phase, all objectives are processed again in priority, but now to their fullest extent [8].

## Planning strategies

For patients with bilateral hip prostheses, iCycle generated coplanar and noncoplanar IMRT plans for two planning strategies. In accordance with AAPM task group report 63 [49], coplanar and noncoplanar $IMRT_{remove}$ plans were generated in which beams that (partly) entered through a metal hip prosthesis before PTV dose delivery were excluded from the list of feasible candidate beams used in the plan generation process (see previous section). For the second strategy, $IMRT_{cut}$ plans were generated in which all feasible candidate beam directions were allowed, but any fluence delivery through prosthetic devices at the entrance side of the PTV was prevented in order to limit uncertainties in dose delivery to the target and organs-at-risk due to inaccuracies in the dose calculation algorithm caused by high-density material. To achieve this, iCycle automatically set the fluence of beamlets, corresponding to the projection of the prostheses in front of the PTV, including a margin of 5 mm to account for set-up uncertainties and rounding errors as a result of the finite fluence grid size, to zero. Depending on the size of the prosthesis and the PTV, and the mutual distance, only 25–40 coplanar candidate beams (from the original 72 beams) remained available for coplanar $IMRT_{remove}$ and 55–82 noncoplanar beams (from the original 166) for noncoplanar $IMRT_{remove}$. In the $IMRT_{cut}$ strategy, also beamlets in a margin of 5 mm around the PTV projection were excluded. By using this additional margin any deviations in dose delivery resulting from irradiation through a prosthesis is avoided in case of errors in patient positioning. A 5 mm margin was considered acceptable due to the daily performed inter- and intra-fraction set-up corrections for prostate cancer patients, based on implanted gold markers and the intra-fraction StereoGraphic Targeting (iSGT) approach [53,54]. For less strict procedures for prostate set-up verification and correction, the margin should be enlarged.

For comparison, iCycle was also used to generate coplanar and noncoplanar plans in which a bone override (with a uniform relative electron density of 1.7) was assigned to

the metal implants ($IMRT_{bone}$). In this way we were able to benchmark the $IMRT_{remove}$ and $IMRT_{cut}$ plans against plans that could be achievable if the patient would not have a metal hip implant.

The above described coplanar $IMRT_{remove}$, $IMRT_{cut}$, and $IMRT_{bone}$ plans were also generated for patients with a unilateral hip prosthesis. In addition, for 3 patients non-coplanar $IMRT_{cut}$ plans were generated. All iCycle plans were generated with a maximum of 9 beam orientations, except for the coplanar $IMRT_{cut}$ strategy for which up to 15 beams were used.

### Plan evaluation

For all plans, the volume of the PTV receiving at least 95% of the prescribed dose ($V_{95\%}$) was assessed. For the rectum, we evaluated the mean dose ($D_{mean}$) and the percentage rectum volume receiving more than 60 Gy ($V_{60\ Gy}$). The latter is associated with the risk of Grade ≥2 rectum toxicity, and rectal bleeding [55]. For the bladder, the mean dose was scored and the bladder volume receiving more than 65 Gy ($V_{65\ Gy}$) and 70 Gy ($V_{70\ Gy}$) respectively [56]. A Wilcoxon signed rank test was used to determine the statistical significance of the differences between different planning approaches.

## 6.3 Results

### Patients with bilateral hip prostheses; 9-beam plans.

In figure 6.1 achieved dose distributions are presented for one of the patients with bilateral metal hip prostheses. Clearly, for the coplanar $IMRT_{remove}$ technique, a wide range of lateral treatment beams could not be selected in the plan, resulting in a high dose region that stretched into antero-posterior direction. This also applied for the noncoplanar $IMRT_{remove}$ plan (not shown). For coplanar $IMRT_{cut}$, more dose was delivered from lateral directions (although complete lateral beams were not selected either), yielding steeper dose fall-off toward the rectum and the surrounding areas. With noncoplanar $IMRT_{cut}$ the dose to the rectum could be further reduced. When comparing this plan

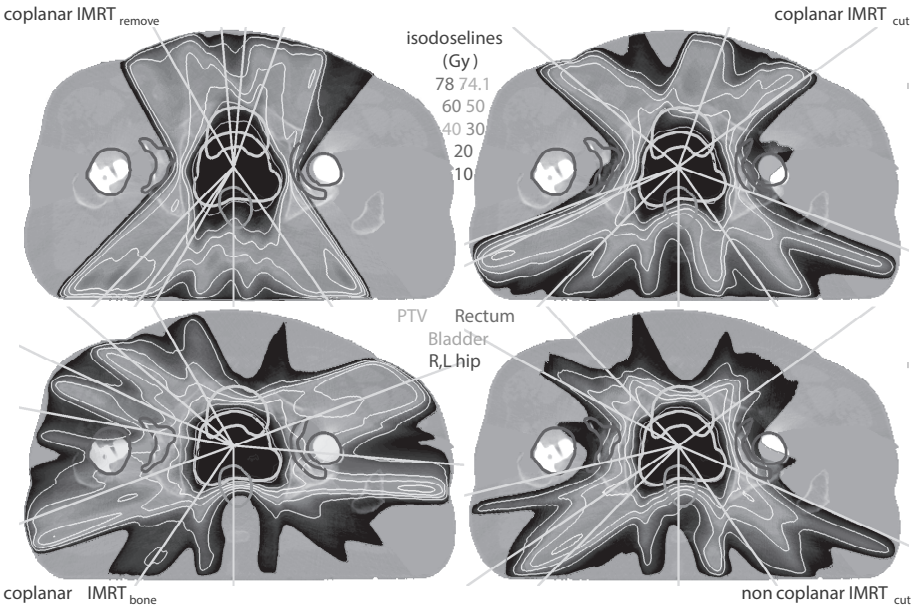



**Figure 6.1:** Dose distributions of 9-beam plans for a patient with bilateral metal hip prostheses. Each line represents a selected beam orientation.

with the noncoplanar $IMRT_{bone}$ plan, differences in PTV coverage and OAR sparing were fairly small.

The obtained $V_{95\%}$ for the PTV was at least 99%, except for one case. For this patient, due to the restrictions in selectable beam orientations, $V_{95\%}$ was only 22.3% for coplanar $IMRT_{remove}$, compared to 99.1% for coplanar $IMRT_{cut}$. Improvement in target coverage was only possible with a higher dose in all OARs. For the other patients, rectum and bladder sparing were significantly improved when using coplanar $IMRT_{cut}$ instead of coplanar $IMRT_{remove}$ (see figure 6.2a). On average, rectum $V_{60\ Gy}$ was reduced by 17.5% ± 15.0% (maximum reduction 37.4%, $p = 0.036$) and rectum $D_{mean}$ by 9.4 Gy ± 7.8 Gy (maximum reduction 19.8 Gy, $p = 0.036$) (see figure 6.3). Due to the lower bladder priority in the wishlist (see Table 6.1), improvements in bladder sparing were smaller than for rectum. Nevertheless, differences in bladder $D_{mean}$ up to 10.8 Gy and in $V_{65\ Gy}$ and $V_{70\ Gy}$ up to 23.5% were observed. OAR sparing could be further improved, although by a smaller amount, by using optimized noncoplanar instead of coplanar beam arrangements. For rectum $V_{60\ Gy}$ and rectum $D_{mean}$, and bladder $V_{65\ Gy}$ the differences with

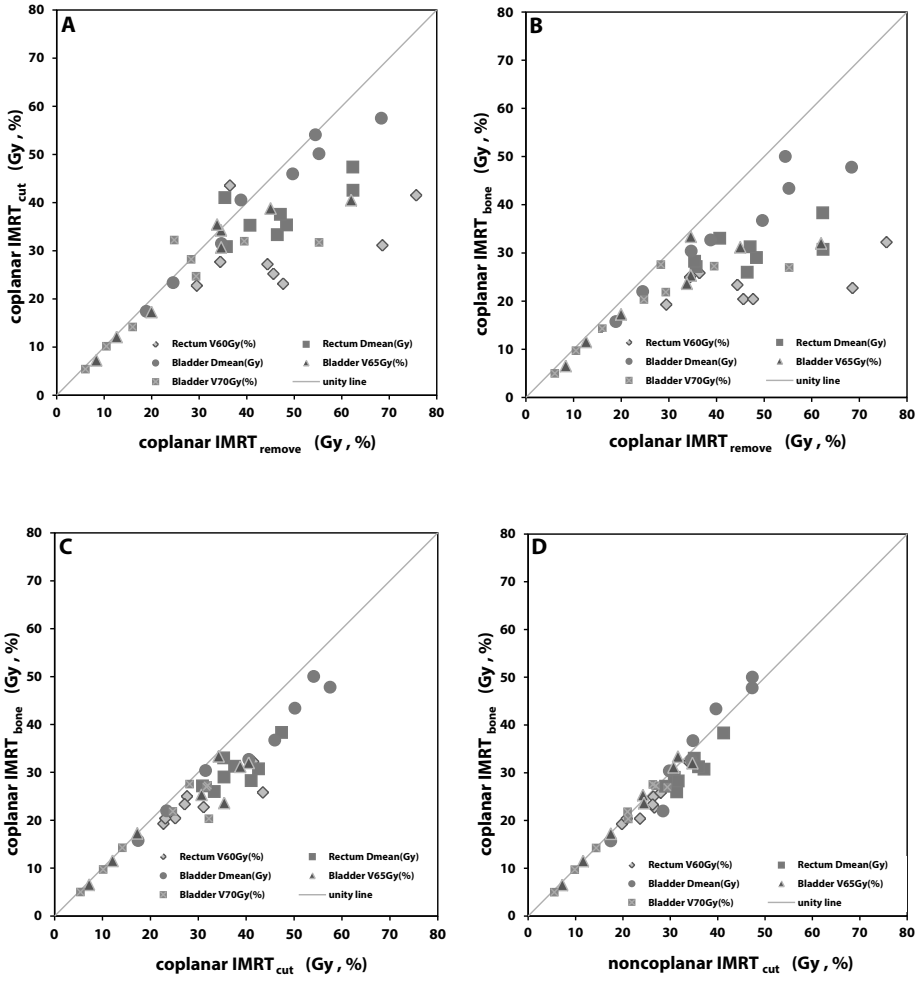



**Figure 6.2:** Comparison of rectum and bladder doses for coplanar $IMRT_{cut}$, $IMRT_{remove}$, and $IMRT_{bone}$, and noncoplanar $IMRT_{cut}$ plans. All markers represent an OAR value for one of the eight patients with bilateral hip prostheses.

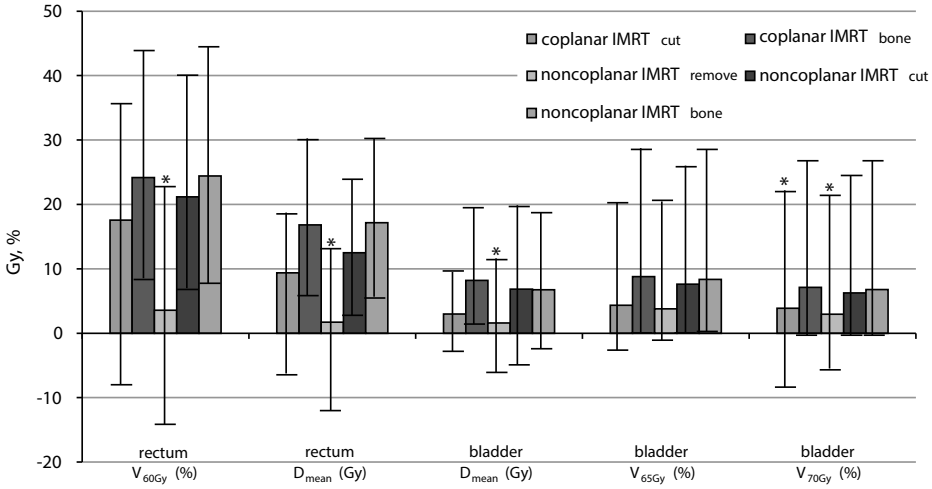



**Figure 6.3:** Improvements in rectum and bladder sparing for the different approaches with respect to coplanar $IMRT_{remove}$ for patients with bilateral hip prostheses. The error bars indicate the observed maximum and minimum values. All differences marked by * were *not* statistically significant ($p > 0.05$).

coplanar $IMRT_{cut}$ arrangements were statistically significant. On average, rectum $V_{60\,Gy}$ was reduced by 4.6% ± 4.9% (maximum reduction 15.5%, $p = 0.018$), rectum $D_{mean}$ by 3.9 Gy ± 3.0 Gy (maximum reduction 9.4 Gy, $p = 0.017$), and bladder $V_{65\,Gy}$ by 4.3% ± 4.2% (maximum reduction 11%, $p = 0.043$).

Figure 6.2 demonstrates that while moving from coplanar $IMRT_{remove}$ (b) to coplanar $IMRT_{cut}$ (c) to noncoplanar $IMRT_{cut}$ (d), rectum and bladder dose-volume parameters gradually decrease toward coplanar $IMRT_{bone}$ values, i.e. the values that could be obtained for a patient without hip implant.

## Patients with a unilateral hip prosthesis; 9-beam plans.

When comparing coplanar $IMRT_{cut}$ with coplanar $IMRT_{remove}$ plans minor differences were observed. Only for one patient bladder $D_{mean}$ was reduced by 1.9 Gy; for the other patients the differences for the evaluated parameters for rectum and bladder sparing remained within 0.8 Gy or 1%. The results were also very close to the values for coplanar $IMRT_{bone}$, showing group average differences in rectum and bladder $D_{mean}$ within 1.4 Gy

and rectum $V_{60\,Gy}$ within 1%. Only for two patients the differences in rectum $D_{mean}$ exceeded 0.8 Gy and in rectum $V_{60\,Gy}$ 0.7%. For the patient showing the largest differences between coplanar $IMRT_{cut}$ and coplanar $IMRT_{bone}$, and two other, randomly selected patients, the impact of using an optimized noncoplanar beam setup was investigated. Only for the first patient, a clear improvement in plan quality was observed, with differences in rectum $V_{60\,Gy}$ of 11% and in rectum $D_{mean}$ of 13.3 Gy. Compared to the other patients, this patient had a relatively large PTV, combined with a substantial amount of metal around the acetabulum.

### Impact of the number of beams on plan quality

Figure 6.4 shows the dependence of dose-volume parameters for PTV, rectum and bladder on the number of selected optimized beam directions in coplanar $IMRT_{cut}$ plans. For patients with bilateral hip prostheses the objective for target coverage could not be achieved in 2 of the 8 patients when using 5 beams only, yielding a $V_{95\%}$ for the PTV of 89.3% and 86.7%, respectively. Using 9 beams instead of 7 beams resulted in a reduction in rectum $V_{60\,Gy}$ of 7.0 ± 6.6% ($p = 0.012$) and in rectum $D_{mean}$ of 5.3 ± 2.3 Gy ($p = 0.017$). Adding another 6 beams resulted in a further improvement in rectum sparing, reducing $V_{60\,Gy}$ by 6.4 ± 4.7% ($p = 0.012$) and $D_{mean}$ by 6.9 ± 1.4 Gy ($p = 0.012$). Due to the lower bladder priority in the wishlist (see Table 6.1), the number of beams had a smaller impact on bladder sparing.

For patients with a unilateral hip prosthesis adequate target coverage was already achieved for the 5-beam $IMRT_{cut}$ plans. Nevertheless, also for this group, adding more beams resulted in a better rectum sparing. When using 9 beams instead of 7, rectum $V_{60\,Gy}$ was reduced by 3.1 ± 0.5% ($p = 0.005$), and rectum $D_{mean}$ by 3.0 ± 0.1 Gy ($p = 0.007$). When using 15 instead of 9 beams these values got lower by 4.3 ± 0.7% ($p = 0.005$) and 4.9 ± 0.3 Gy ($p = 0.005$).

## 6.4 Discussion

Radiotherapy treatment of prostate cancer patients with unilateral or bilateral metal hip prosthesis is challenging. According to recommendations of the AAPM taskgroup 63 [49], the use of beams that irradiate the PTV through a prosthesis should preferably be

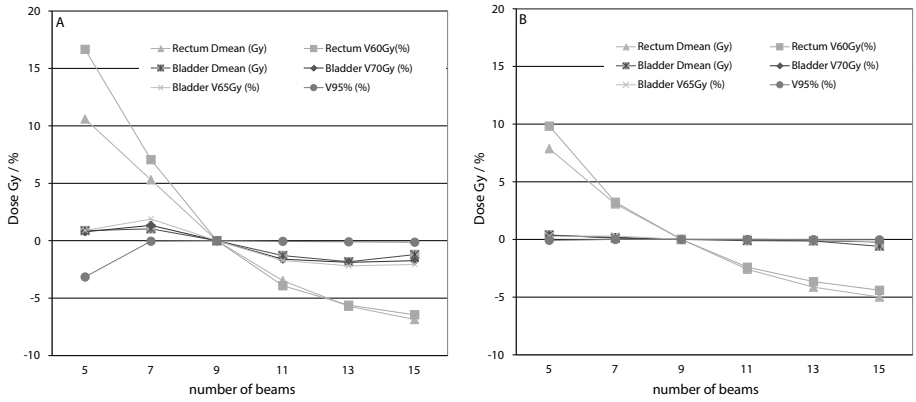



**Figure 6.4:** Dependence of dose-volume parameters for PTV, rectum and bladder on the number of optimized beam directions in coplanar $IMRT_{cut}$ plans. Mean differences with respect to 9-beam plans are presented for patients with bilateral hip prostheses (A) and a unilateral hip prosthesis (B).

avoided, due to uncertainties in the dose calculation in the vicinity of prostheses for most commercially available treatment planning systems (TPS). In this study we systematically compared this planning technique ($IMRT_{remove}$) with a strategy that only excluded those beamlets from the optimization that would deliver dose to the PTV while first passing through an implant ($IMRT_{cut}$). With this latter strategy significant reductions in dose delivery to the rectum and bladder could be obtained (see figures 6.2 and 6.3), being attributed to the enhanced freedom in selecting beamlets with largely lateral PTV irradiation.

As shown in figure 6.4, increasing the number of treatment beams will improve rectum sparing. Compared to 9-beam noncoplanar plans, a similar rectum sparing could be achieved by using 11–13 coplanar beams for all patients. With a larger number of beams, the patient volume treated to a low dose will increase, but we consider this a minor limitation. The number of monitor units does not necessarily increase for larger numbers of treatment beams, but treatment delivery times will generally be longer due to the time needed for setting up and verification of each beam. To even a larger extent this also applies when using noncoplanar beam set-ups, since technicians have to enter the room for couch adjustments. For prostate cancer patients a longer treatment delivery time may have a large impact on intra-fraction motion [54]. Therefore, in our clinic, a maximum of 9 coplanar beams is used for treatment of prostate cancer patients, unless we expect a clear, clinically relevant benefit with respect to rectum sparing of the use of larger numbers of coplanar beams. In our study we did not evaluate VMAT, allowing

much shorter treatment delivery times, because iCycle does not yet support this. We are currently working on this.

As explained in the materials and method section, during plan generation iCycle consecutively adds new beam orientation to the plan, based on the already available configuration. In other papers, we have demonstrated that very good plans may be obtained in this way. However, there is no proof that this method will result in the overall best plan. Certainly, from a theoretical point of view there are reasons to believe that there will be plans that are suboptimal. To what extent this is problematic and how often it occurs is part of further study. Anyway, because of the high quality of the iCycle plans compared to the manually generated plans [8], we believe that this issue is of minor importance for clinical practice. Because in this comparative study all iCycle plans were generated in the same way, it does not affect the conclusion of our work either.

Recently, Brooks et al. [57] compared coplanar and noncoplanar IMRT plans for one prostate cancer patient, using beams that fully avoided the prosthesis but not always covered the entire PTV. For the noncoplanar plan, a better rectum sparing was reported, but a larger bladder volume was exposed to doses up to 50 Gy. With our $IMRT_{cut}$ strategy the dose delivered to both rectum and bladder could be reduced for most patients.

Kling et al. reported the comparison of 7 static fields IMRT with 2–4 partial arc TomoTherapy plans for patients with bilateral metal hip prostheses [52]. For all plans, a completely metal-avoiding strategy was used, like our $IMRT_{remove}$ approach. For both techniques they observed a vast antero-posterior dose delivery, resulting in unacceptably large amount of high dose outside the PTV. Previously, for two patients we simulated irradiation with the $IMRT_{cut}$ strategy in a commercial TPS (Monaco 3.1, Elekta, AB), using PAVs to effectively block beamlets that would first pass prostheses and then deliver dose to the PTV [50]. For both patients, the simulated $IMRT_{cut}$ strategy, using the default beam setup for prostate cancer patients without implants, resulted in improved OAR sparing compared to $IMRT_{remove}$ with manually selected beam orientations. In this paper, iCycle was used for integrated beam angle selection and IMRT optimization, giving opportunities to further enhance OAR sparing with $IMRT_{cut}$ compared to the default beam set-up (not investigated here). For each patient iCycle automatically provided a patient-specific set of beam orientations, depending of the patient anatomy and the size of the prostheses (see figure 6.1 for an example). Even for patients with a unilateral hip prosthesis, we were not able to extract a class solution for the beam configuration from these iCycle results.

Although in the $IMRT_{cut}$ strategy fluence delivery is avoided for beamlets that would deliver dose to the PTV after passing a prosthesis, the dose delivered to the prostheses might be underestimated in case the MLC dose leakage is not properly taking into account in the TPS. Since in all our plans the maximum dose in the prostheses remained

below 44 Gy and the dose contribution due to leakage was always less than 1 Gy, we do not consider this a risk for the patient.

Especially for patients with bilateral hip prostheses, delineation of the prostate is often hampered due to severe streaking artifacts in the planning CT scans. To improve delineation accuracy, fusion with an MR scan is highly recommended. Not for all patients in this study this was performed. But we think this does not impact our main conclusions regarding the comparison of the planning strategies.

In our study we used a 10 MV photon beam for all plans. Previously, we have seen that when using more than five treatment beams for prostate cancer, the selected beam energy is of minor importance. This is also supported by literature (eg. [58]). Therefore, we expect similar improvements in plan quality with $IMRT_{cut}$ when using either lower or higher photon energies. Attenuation of the treatment couch was not taken into account during plan generation. By taking into account the optimized fluences for some beam directions would somewhat change, but the final dose distribution would be similar.

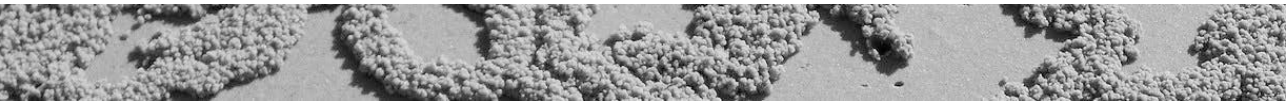
We demonstrated for each of the eight patients with bilateral implants clear superiority of the $IMRT_{cut}$ planning strategy compared to the $IMRT_{remove}$ approach. In principle, $IMRT_{cut}$ can be simulated in a commercial TPS with the previously described PAVs. However, this is time-consuming, so we encourage TPS vendors to implement an $IMRT_{cut}$ solution as described in this paper.

## 6.5 Conclusions

For prostate cancer patients with metal hip implants, iCycle was able to automatically generate high quality, patient-specific IMRT plans with optimized coplanar or noncoplanar beam arrangements. Excluding only beamlets that delivered dose to the PTV after passing through prostheses significantly improved OAR sparing for patients with bilateral implants as opposed to the AAPM taskgroup 63 recommendation to fully exclude beams containing such beamlets. Noncoplanar beam arrangements and, to a larger extent, increasing the number of optimized treatment beams further improved plan quality.

chapter 7

# Fully automated volumetric modulated arc therapy plan generation for prostate cancer patients



Peter W.J. Voet RTT, Maarten L.P. Dirkx PhD, Sebastiaan Breedveld PhD,
Abrahim Al-Mamgani MD, PhD, Luca Incrocci MD, PhD, Ben J.M. Heijmen PhD

*International Journal of Radiation Oncology Biology Physics; 88(5): 1175–79 (2014).*

## Summary

A system for fully automated volumetric modulated arc therapy (VMAT) treatment planning for prostate cancer patients was developed and evaluated. Plan quality in automatically generated VMAT plans was similar to manually generated VMAT plans, while saving at least one hour hands-on time. Based on these results, clinical implementation has been started.

## Abstract

**Purpose**: To develop and evaluate fully automated VMAT treatment planning for prostate cancer patients, avoiding manual trial-and-error tweaking of plan parameters by dosimetrists.
**Methods and Materials:** A system was developed for fully automated generation of VMAT plans with our commercial clinical TPS, linked to the in-house developed Erasmus-iCycle multicriterial optimizer for pre-optimization. For 30 randomly selected patients, automatically generated VMAT plans ($VMAT_{auto}$) were compared with VMAT plans generated manually by one expert dosimetrist in the absence of time pressure ($VMAT_{man}$). For all treatment plans, planning target volume (PTV) coverage and sparing of organs-at-risk (OARs) were quantified.

**Results:** All generated plans were clinically acceptable and had similar PTV coverage ($V_{95\%}$ > 99%). For $VMAT_{auto}$ and $VMAT_{man}$ plans the OAR sparing was similar as well, while only the former plans were generated without any planning workload.
**Conclusions:** Fully automated generation of high quality VMAT plans for prostate cancer patients is feasible and has recently been implemented in our clinic.

## 7.1 Introduction

For several years, intensity modulated radiotherapy (IMRT) has been used at a large scale as curative treatment for prostate cancer patients. More recently, volumetric modulated arc therapy (VMAT) is increasingly being used as an effective and time efficient alternative [59]. Due to shorter treatment delivery times, application of VMAT might also reduce the potential impact of intra-fraction prostate motion [54]. Like for static gantry IMRT planning, generation of an optimal VMAT plan is not straightforward and often time consuming. Dosimetrists have to select the optimization parameters manually and modify them in a trial-and-error process. The quality of the final treatment plan depends on the skills and experience of the dosimetrist, the complexity of the case, and the available time for plan generation.

Recently, we developed Erasmus-iCycle, an algorithm for fully automated, multicriterial beam profile optimization and beam angle selection for coplanar and noncoplanar IMRT [8,45,60,61]. Mainly because of (small) inaccuracies in the final patient dose distribution, Erasmus-iCycle can currently not be used for automated generation of clinically deliverable plans. In a prospective clinical study, we compared an Erasmus-iCycle based semi-automated planning procedure for treatment of head-and-neck cancer patients with the regular manual plan generation by dosimetrists [45]. For semi-automated planning, Erasmus-iCycle was used for pre-optimization by fully automatically generating a plan with optimized fluence profiles and beam angles. One of the investigators (P.W.J. Voet) then reconstructed the plan in the clinical treatment planning system (TPS). In parallel, for each patient, a manual IMRT plan was created by a dosimetrist without input of Erasmus-iCycle, using the clinical TPS and following the clinical procedures. In 97% of cases, physicians selected the semi-automatically generated plan for treatment. The procedure both resulted in a higher plan quality and reduced the planning hands-on time by 50% compared to manual planning. Improved plan quality with Erasmus-iCycle was also observed for prostate SBRT (stereotactic body radiation therapy) [61]. In

that study we also demonstrated higher plan quality for optimized, noncoplanar beam arrangements compared to coplanar.

In this study we develop and evaluate a fully automated VMAT planning strategy for prostate cancer patients. Erasmus-iCycle is still used as pre-optimization for generation of VMAT plans in the clinical TPS, but we realized a fully automated reconstruction of Erasmus-iCycle plans in Monaco, avoiding manual plan reconstruction by dosimetrists as used in [45]. For 30 prostate cancer patients we compared automatically generated VMAT plans ($VMAT_{auto}$) with VMAT plans generated by one expert dosimetrist (P.W.J. Voet) ($VMAT_{man}$).

# 7.2 Materials and Methods

## Study patients

From our clinical database, we randomly selected 30 prostate cancer patients, 10 from each risk groups (see Table 7.1), who were irradiated in our department with curative intent between January 2012 and January 2013. All patients had 4 implanted gold fiducials, used for both inter- and intra-fraction set-up corrections [54].

For all patients, a computed tomography (CT) scan was acquired using a Somatom Sensation Open multi-slice CT scanner (Siemens, Erlangen, Germany). Prostate, seminal vesicles (for groups 2 and 3 patients), bladder, rectum, anus, and hips were delineated according to the Radiation Therapy Oncology Group (RTOG) guidelines [51]. The PTV was defined by expanding the prostate with a margin of 5 mm (7 mm in caudal direction)

**Table 7.1:** Grouping criteria for prostate cancer patients, based on the risk of tumor involvement in the seminal vesicles.

| | | T1-T3a | | T3b, T4 |
|---|---|---|---|---|
| Gleason score | PSA ≤ 4 | 4 ≤ PSA ≤ 10 | PSA ≥ 20 | any PSA |
| 2–6 | group 1 | group 1 | group 2 | group 3 |
| 7 | group 1 | group 2 | group 2 | group 3 |
| 8–10 | group 2 | group 2 | group 3 | group 3 |

PSA = prostate-specific antigen

and, for groups 2 and 3 patients, the seminal vesicles with a uniform margin of 8 mm, as derived by Mutanga et al. [62]. All patients were irradiated with a step-and-shoot IMRT technique. Patients in groups 1 and 3 were treated with 78 Gy to the PTV, delivered in 39 fractions. For group 2 patients, a simultaneous integrated boost technique was used, delivering in 39 fractions 78 Gy to the prostate PTV and 72.2 Gy to the seminal vesicles PTV.

## VMAT plan generation

*Automated VMAT plan generation using Erasmus-iCycle*

After delineation of the organs-at-risk, prostate, and, if applicable, seminal vesicles, and applying CTV-PTV margins, $VMAT_{auto}$ plans were generated fully automatically in a two-step approach. First, Erasmus-iCycle was used to automatically generate a 23-beam, equiangular IMRT plan with 10 MV photon beams. Based on achieved constraints and objective values in this plan, a VMAT plan template for our clinical TPS (Monaco version 3.3, Elekta AB, Stockholm, Sweden) was generated. Next, the patient was automatically selected in Monaco. Based on the patient-specific template, a clinically deliverable single arc VMAT plan was generated without any user interaction. For the final dose calculation in Monaco, the implemented Monte Carlo dose calculation engine was applied. After finishing, $VMAT_{auto}$ plans were automatically saved in Monaco for review.

In [8] Erasmus-iCycle is described in detail. Fully automated plan generation with Erasmus-iCycle is based on a planning "wishlist", containing hard constraints to be strictly obeyed and prioritized objectives, and a list with candidate beam orientations. Core of Erasmus-iCycle is the 2-phase ε-constraint (2pεc) algorithm for generating Pareto-optimal IMRT plans for preselected (fixed) beam arrangements. In this study, this configuration consisted of 23 equi-angular beams. For a subset of prostate cancer patients we found that by using more than 23 beams plan quality did not improve, while calculation times increased.

All Erasmus-iCycle plans in this study were optimized using one fixed wishlist (see Table 7.2). This wishlist was a priori established/tuned in a collaboration between treating physicians, dosimetrists and physicists by iterative planning, plan evaluation, and wishlist adjustments for a small group of test patients. First priority was PTV coverage. To ensure that the dose delivered to 99% of the PTV was at least 95% of the prescribed dose, we used the Logarithmic Tumor Control Probability (LTCP, [41]), with a fixed sensitivity value (alpha) of 0.8. For a homogeneous target dose, equal to the prescribed dose, LTCP equals 1. A lower target dose results in a strongly increased LTCP, while higher target doses decrease the LTCP to a minimum value of 0 for a clinically infeasible, infinite

**Table 7.2:** Applied wishlist for all study patients. LTCP = logarithmic tumor control probability. gEUD = generalized equivalent uniform dose.

| **Constraints** | | | |
|---|---|---|---|
| | Volume | Type | Limit |
| | PTV | max | 104% of prescribed dose |
| | PTV shell 50 mm | max | 60% of prescribed dose |
| | Unspecified tissue | max | 104% of prescribed dose |
| | Right + left hip | max | 40 Gy |

| **Objectives** | | | |
|---|---|---|---|
| Priority | Volume | Type | Goal |
| 1 | PTV | ↓LTCP | 0.5 |
| 2 | Rectum | ↓gEUD (parameter 12) | 40% of prescribed dose |
| 3 | Rectum | ↓gEUD (parameter 8) | 25% of prescribed dose |
| 4 | Rectum | ↓mean | 33% of prescribed dose |
| 5 | External ring | ↓max | 40% of prescribed dose |
| 6 | PTV shell 5 mm | ↓max | 93% of prescribed dose |
| 7 | Anus $D_{mean}$ | ↓mean | 10% of prescribed dose |
| 8 | PTV shell 15 mm | ↓max | 70% of prescribed dose |
| 9 | PTV shell 25 mm | ↓max | 50% of prescribed dose |
| 10 | Bladder | ↓mean | 60% of prescribed dose |
| 11 | Right + left hip | ↓mean | 25% of prescribed dose |
| 12 | Unspecified tissue | ↓mean | 10 Gy |

dose. Objectives with decreasing priorities were set to the rectum, anus, bladder and hips. To minimize the rectum volume treated to a high dose, a generalized equivalent uniform dose objective (gEUD, [63]) with volume parameter $a$=12 was used (priority 2). The volume treated to intermediate to high doses was further suppressed by minimizing a gEUD with $a$=8 (priority 3). Using a mean dose objective, the rectum volume treated to lower doses was minimized as well (priority 4). To realize a steep dose fall-off outside the PTV, one constraint and three objectives (with priorities 6, 8 and 9) for shells around the PTV were defined. In addition, an objective for an external ring, 2 cm inside the patient body contour, was used to minimize the entrance dose (priority 5). In the first phase of the 2pec optimization, starting with the priority 1 objective, objective functions were consecutively minimized, but not below the prescribed goal values as mentioned in Table 7.2, and then fixed as a constraint for minimization of lower prioritized objectives. In the second phase, all objectives were again consecutively minimized according priority, but now to the fullest extent.

*Manual VMAT plan generation*
As a benchmark for $VMAT_{auto}$ plans, one expert dosimetrist (PV) used the clinical TPS to manually generate $VMAT_{man}$ plans, without any knowledge of the $VMAT_{auto}$ results. Like for $VMAT_{auto}$, a single arc rotation with a 10 MV photon beam was used. Group based templates with dose constraints used for generation of clinical IMRT plans were selected as starting point for $VMAT_{man}$ plan generation. During optimization, the template parameters were manually tweaked to improve plan quality. In contrast to the general clinical practice, the expert dosimetrist could generate the $VMAT_{man}$ plans without a strong clinical time pressure. A maximum of 4 hours of hands-on planning time was used.

## Plan evaluation

For all plans, the volume of the PTV receiving at least 95% of the prescribed dose ($V_{95\%}$) and the volume receiving more than 107% ($V_{107\%}$) were determined. For the rectum, the mean dose ($D_{mean}$) and the volume percentages receiving more than 60 Gy ($V_{60\ Gy}$) and more than 75 Gy ($V_{75\ Gy}$) were evaluated. Both $V_{60\ Gy}$ and $V_{75\ Gy}$ are associated with the risk of Grade ≥2 rectum toxicity, and rectal bleeding [55]. For the bladder, $D_{mean}$ and the volume receiving more than 65 Gy ($V_{65\ Gy}$) were quantified [56]. The mean dose in the anus and the maximum dose in the hips were also derived. Two-sided Wilcoxon matched-paired signed rank tests were used to assess statistical significance of observed plan parameters differences between $VMAT_{auto}$ and $VMAT_{man}$.

# 7.3 Results

For all plans, $V_{95\%}$ for the PTV was at least 99%, and PTV doses higher than 107% did not occur. Statistically significant differences in target coverage were not observed between $VMAT_{man}$, and $VMAT_{auto}$ plans ($p \geq 0.44$).

Between $VMAT_{auto}$ and $VMAT_{man}$ only small differences in rectum, anus and bladder plan parameters were observed (Figure 7.1 and Table 7.3). While for patients in groups 1 and 2, rectum $V_{60\ Gy}$ and $D_{mean}$ were slightly lower for $VMAT_{auto}$, differences were statistically insignificant for all other parameters.

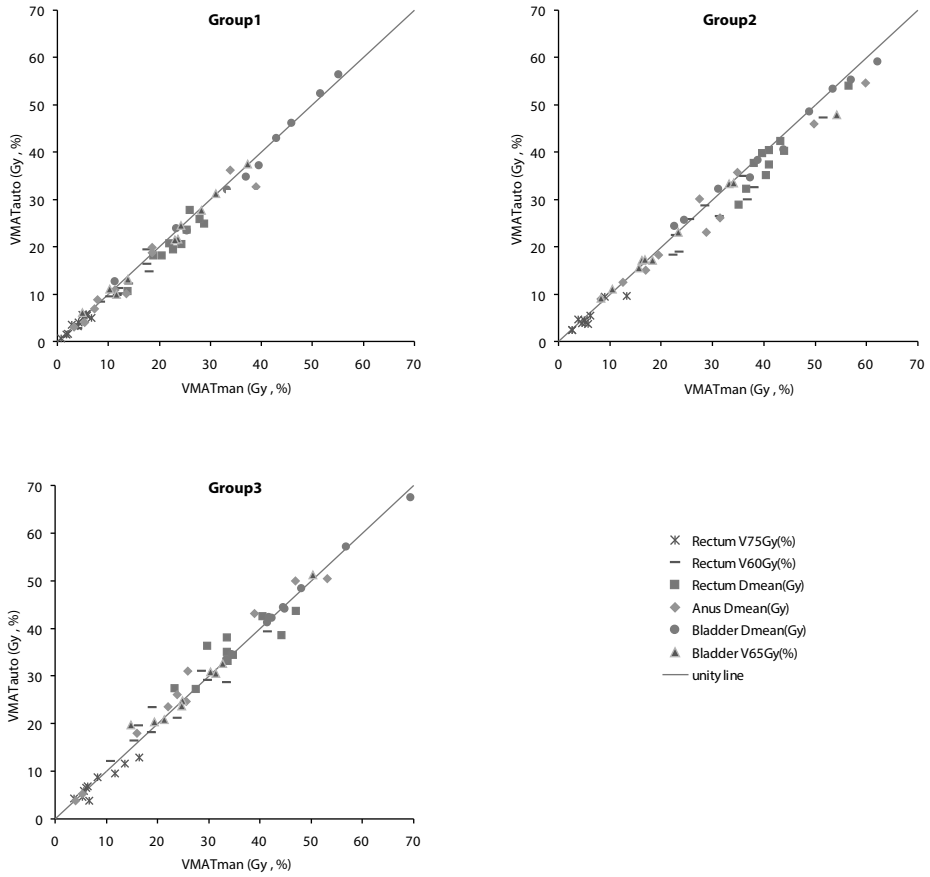



**Figure 7.1:** Comparison of rectum, bladder and anus plan parameters between $VMAT_{auto}$ and $VMAT_{man}$ for the three study groups. Each marker represents a parameter comparison for one of the study patients. For data points to the right of the unity line $VMAT_{auto}$ yielded better sparing.

**Table 7.3:** Comparison of plan parameters for the organs-at-risk for each of the study groups, showing mean differences, standard deviations, difference ranges and *p*-values. Positive values point at lower parameter values for $VMAT_{auto}$ when compared to $VMAT_{man}$. NS = statistically not significant, i.e., $p>0.05$.

| $VMAT_{man}$ − $VMAT_{auto}$ | GROUP 1 | | | GROUP 2 | | | GROUP 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean ± 1SD | [Range] | *p*-value | Mean ± 1SD | [Range] | *p*-value | Mean ± 1SD | [Range] | *p*-value |
| Rectum $V_{75\ Gy}$ (%) | 0.3 ± 0.7 | [-0.8 , 0.7] | NS | 0.8 ± 1.3 | [-0.7 , 3.7] | NS | 0.9 ± 1.6 | [-0.6 , 3.6] | NS |
| Rectum $V_{60\ Gy}$ (%) | 1.2 ± 1.4 | [-1.8 , 3.2] | 0.028 | 3.1 ± 2.6 | [-0.2 , 6.9] | 0.012 | -0.2 ± 2.9 | [-4.4 , 4.8] | NS |
| Rectum $D_{mean}$ (Gy) | 2.1 ± 1.7 | [-1.6 , 4.0] | 0.013 | 2.8 ± 2.3 | [0.0 , 6.4] | 0.008 | -0.8 ± 3.7 | [-6.5 , 5.7] | NS |
| Anus $D_{mean}$ (Gy) | 0.8 ± 2.5 | [-2.4 , 6.3] | NS | 1.9 ± 3.0 | [-2.7 , 5.8] | NS | -1.4 ± 2.4 | [-5.2 , 2.7] | NS |
| Bladder $V_{65\ Gy}$ (%) | 0.4 ± 1.1 | [-1.2 , 2.0] | NS | 0.6 ± 2.1 | [-1.0 , 6.3] | NS | -0.1 ± 2.2 | [-4.8 , 4.3] | NS |
| Bladder $D_{mean}$ (Gy) | 0.4 ± 1.5 | [-1.3 , 2.4] | NS | 0.8 ± 1.8 | [-1.8 , 3.2] | NS | 0.2 ± 0.7 | [-0.5 , 2.0] | NS |
| Right Hip $D_{max}$ (Gy) | -0.3 ± 1.8 | [-4.8 , 2.2] | NS | 1.2 ± 2.7 | [-4.5 , 5.0] | NS | 1.7 ± 3.0 | [-3.3 , 7.5] | NS |
| Left Hip $D_{max}$ (Gy) | 0.2 ± 1.7 | [-1.9 , 3.0] | NS | 1.7 ± 2.4 | [-0.7 , 6.6] | NS | 0.8 ± 1.2 | [-1.0 , 2.7] | NS |

Compared to $VMAT_{man}$ plans, $VMAT_{auto}$ plans were generated with a vast reduction in workload. Excluding the time for contouring, generation of the $VMAT_{man}$ plans took on average 5 hours, including 3 hours hands-on time. In contrast, with our current hardware and software, generation of the $VMAT_{auto}$ plans took about two hours and did not require any user interaction. Half of the time was spent on the pre-optimization in Erasmus-iCycle and half of the time on reconstruction of a clinically deliverable plan in Monaco. As explained before, both steps were sequentially performed in an automated way.

## 7.4 Discussion

The purpose of this study was development and evaluation of a procedure for fully automated VMAT plan generation for prostate cancer patients to replace manual IMRT planning as previously used in our clinic. While in a previous study [45] dosimetrists still had to manually reconstruct the automatically generated Erasmus-iCycle plan in the clinical TPS to obtain a clinically deliverable treatment plan, this manual plan reconstruction was now also replaced by automation. As demonstrated in Figure 7.1 and Table 7.3, differences between $VMAT_{auto}$ and $VMAT_{man}$ plans were small and not consistent. The conclusion from this observation is that even an expert dosimetrist, in the absence of time pressure, could not beat the procedure for fully automated plan generation.

We also compared the $VMAT_{auto}$ plans to the clinically applied step-and-shoot IMRT plans. In all patients a large and consistent improvement in plan quality was observed for the VMAT plans. Rectum $V_{75\,Gy}$, $V_{60\,Gy}$, and $D_{mean}$, and anus $D_{mean}$ were reduced by up to 6.1%, 16.7%, 15.4 Gy, and 18.4 Gy, respectively. For most patients, doses in the bladder and the hips were largely reduced as well, showing differences up to 12 Gy in bladder mean dose and maximum dose in the hips. We believe that the observed large differences in plan quality are mainly caused by variability in planning skills and experience among our dosimetrists and by limitations in available planning time in the clinical setting. Previously, we also observed this in our prospective study with Erasmus-iCycle for head-and-neck cancer patients [45].

Based on the results in this paper, we have recently started clinical implementation of the fully automated VMAT plan generation. For the first ten patients, we asked dosimetrists to attempt improving the $VMAT_{auto}$ plans by manually tweaking the plan objectives in Monaco in a trial-and-error procedure, as normally done in clinical prac-

tice. A maximum hands-on time of 1 hour was allowed, corresponding to total planning times of up to 3 hours. For all patients, the $VMAT_{auto}$ and manually tweaked VMAT plans had similar PTV coverage; the mean $V_{95\%}$ was 99.5% ± 0.2% (1SD), $p = 1.0$. Also differences in rectum $V_{75\,Gy}$ and $V_{60\,Gy}$ (0.1% ± 0.5% (1SD), maximum 1%) were statistically insignificant ($p \geq 0.54$). In case of a slight reduction in $V_{75\,Gy}$ or $V_{60\,Gy}$ by manual tweaking for an individual patient, this minor improvement generally resulted in an increased dose delivery to other, lower prioritized OARs and/or in a deteriorated dose fall-off in unspecified tissues. For all patients, the treating physician selected the $VMAT_{auto}$ plan for treatment. Because of the success of the $VMAT_{auto}$ procedure, we have decided that manual VMAT planning will not become part of our clinical routine for prostate cancer patients.

As explained in detail in the materials and methods section, the $VMAT_{auto}$ procedure developed in this study is a two-step process, using sequentially two independent optimizers. In on-going research projects, automated plan generation with a single optimizer is pursued. In this study, 23-beam equi-angular IMRT plans, automatically generated with Erasmus-iCycle, were used as input for $VMAT_{auto}$ plan generation in the clinical TPS. Fixed beam plans were used because VMAT plan generation is currently not yet supported by Erasmus-iCycle. In previous studies, significant improvements in plan quality were observed when increasing the number of beam directions [60,61]. In this study the optimal number of beams was 23; with less beams plan quality was deteriorated for some patients, and the use of more beams only increased calculation times in Erasmus-iCycle. For prostate cancer patients treated with static field IMRT, Quan et al. [64] also reported an improvement in plan quality when increasing the number of beams up to 24. In their study, static gantry IMRT plans with 8 beams showed a similar plan quality as VMAT plans. To guarantee a smooth transition from Erasmus-iCycle to the clinical TPS, in this study, the Erasmus-iCycle wishlist was constructed with types of constraints and objectives that were also available in the clinical TPS. In principle this could result in a sub-optimal wishlist, but we did not find any evidence for this. When comparing the output of Erasmus-iCycle with the final $VMAT_{auto}$ plans, only minor differences in achieved values for the two highest prioritized objectives were observed (see Table 7.2). For the other objectives the achieved values often showed a small deterioration. Some changes could be expected due to differences in the dose calculation algorithm, and because a 23 static beam IMRT plan, used for pre-optimization, was converted in a single arc VMAT plan in Monaco. No improvements in plan quality were observed when allowing a dual arc VMAT delivery.

To evaluate whether the $VMAT_{auto}$ plans could be accurately delivered, we performed dose measurements at an Elekta Synergy linac (Elekta AB, Sweden) using a 2D array in an Octavius phantom (PTW, Germany) for 6 study patients, 2 from each group. Using reference criteria for gamma evaluation of 3% global and 3 mm distance to agreement, a

pass-rate of at least 95.7% was observed. We are convinced that Erasmus-iCycle can be used as a pre-optimizer for other commercial TPS as well. In addition to a list of obtained values for constraints and objectives, also a Dicom RT-Dose file can be exported from Erasmus-iCycle to be fed into a TPS to derive clinically acceptable and deliverable VMAT plans. Further research to assess the feasibility and accuracy of this approach will be performed.

Recently, Wu et al. [65] published an automated VMAT planning procedure for head-and-neck cancer patients based on a database with IMRT plans and overlap volume histograms. In terms of target coverage and sparing of organs-at-risk the generated VMAT plans were comparable to the IMRT plans. A rather large database with previously treated IMRT patients was required as input to assure proper matching for a new patient. Zhang et al. [47] described an algorithm for fully automated IMRT planning for lung cancer. The planning objectives and constraints were selected from an expert database according to tumor size and location. In terms of tumor coverage and normal tissue sparing the automatically generated plans were intended to be better than, or at least equivalent to, those manually designed by dosimetrists.

This study focused on development and evaluation of a method for automated VMAT plan generation for prostate cancer patients. Extension of this technique to other patient categories is under investigation.

## 7.5 Conclusions

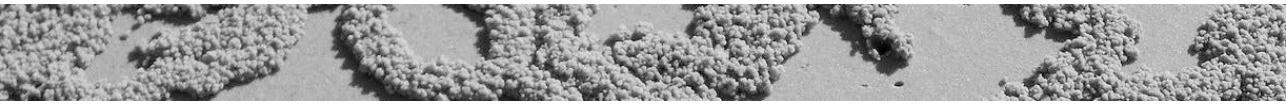
A system for fully automated VMAT plan generation for prostate cancer patients was developed and evaluated, avoiding manual trial-and-error tweaking of plan parameters by dosimetrists. The system has recently been implemented in our clinic to improve plan quality and reduce workload.

## 7.6 Acknowledgement

The authors would like to thank Dennis Grofsmid for performing the dose QA measurements for the $VMAT_{auto}$ plans.

# chapter 8

## Discussion



# 8.1 Introduction

This thesis focuses on automation of tasks in the preparation phase of a radiotherapy treatment, resulting in equal, or preferentially higher, treatment quality with less human interference. Due to technical limitations, tumor doses needed for high local control rates can currently not always be delivered, and successful tumor eradication may be accompanied by serious side effects. On the other hand, increased efficiency and cost reduction are getting of great importance to keep high quality healthcare available for all, and even allow for improvements, e.g. by adaptive treatment approaches. In this thesis we have demonstrated that automation of tasks may result in higher quality treatments (including those based on adaptive re-planning), and lower costs.

# 8.2 Current application of autocontouring in radiotherapy

After the clinical validation study for head-and-neck cancer patients in 2010 (chapter 2), atlas-based autosegmentation (ABAS) has been used in our institution for all head-and-neck cancer patients. Since then, ABAS has also been evaluated and implemented in daily routine for other treatment sites. Currently, ABAS is an essential tool in our daily clinical workflow. In 2013, about 2700 ABAS runs were performed (on a total of ~4500

newly treated patients with external beam radiotherapy) for new breast, head-and-neck, thoracic, brain, cervical, and prostate cancer. For head-and-neck cancer, ABAS was used for autocontouring of organs-at-risk (OARs) and elective neck targets. For the other sites, autocontouring was limited to OARs. For another 800 patients, ABAS was used for autosegmentation of a new scan based on a previously contoured CT-scan of the same patient, which was then used as single atlas ('intra-patient autosegmentation' or 're-contouring'). Re-scanning and re-contouring was frequently performed for head-and-neck cancer patients that lost weight during treatment, requiring a new treatment plan. In our clinical experience, for these patients, intra-patient autosegmentation resulted in higher quality autocontours than the use of contoured scans of other, previously treated patients as atlases. Only minor editing of autocontours was generally needed (~15 minutes editing time). Intra-patient autosegmentation was also routinely applied for contouring of 4DCT scans of patients with tumors in the thoracic region. In these cases, the treating physician first manually contoured a 3D scan belonging to a single respiratory phase. Next, ABAS was used for propagating these manually delineated contours to the other phases.

In literature, the use of autocontouring software for delineation of OARs and target volumes has been described for various treatment sites [e.g., 66–69]. In line with our findings (chapters 2, 3, and clinical experience), careful evaluation of generated autocontours and manual editing were recommended. Simmad et al. [69] tested two commercially available atlas-based autosegmentation software tools for prostate cancer patients (ABAS, Elekta, Sweden (also used in our studies), and iPlan RT Image, Brainlab, Germany). For both systems, they concluded that editing of autocontours was mandatory for clinical use. They also investigated the use of the contoured planning CT scan of a patient for autocontouring of CBCTs acquired during treatment. They concluded that, due to the poor CBCT image quality, the required accuracy for organ delineation and time savings, needed for (online) adaptive radiotherapy, were not realized.

Despite the required inspection and editing of autocontours, substantial time savings have been reported compared to manual contouring, ranging from 26% [68]), to 50% [69], or even 60% [67 and chapter 2 of this thesis].

The required accuracy of autocontours generally depends on the involved structures. In chapter 3 we showed that editing of parallel OARs, like parotid glands, did not largely affect the planned mean dose, often being used for plan evaluation. Nevertheless, it is to be expected that exact definition of the boundaries of very small parallel OARs, like submandibular glands, and swallowing muscles, is more critical. Also, accurate editing of elective neck level CTVs turned out to be important in order to prevent large underdosages (chapter 3). Careful inspection, and possibly manual editing, of automatically generated contours of serial organs, like the spinal cord, is also needed. Small inaccuracies in the definition of the boundaries of these organs may translate in large deviations in

the percentage of volume treated to a high dose, and/or in the maximum dose delivered to the organ. This effect is similar to the dosimetric consequences we observed for elective neck level CTV contours in chapter 3.

In our studies for head-and-neck cancer patients described in chapters 2 and 3, and in current clinical practice, we did not use the autocontouring software for delineation of the primary tumor. First, the locations and shapes of the primary tumors do largely differ between patients. Consequently, it is hardly possible to define a set of atlases to be used for accurate autosegmentation of the primary tumor. Besides this, the CTV of the primary tumor is generally based on more information than available in the planning CT only. Physicians use patient specific information obtained from multiple imaging modalities, which may be supplemented by surgical reports, or visual examinations.

In the evolving field of online-adaptive planning, accurate and fast autocontouring is a prerequisite. An online-adaptive procedure should take a couple of minutes at maximum. In this limited timeframe, the entire workflow should be finalized: acquisition of a new CT, (semi-) automated contouring of target volumes and OARs, and generation and QA of the new treatment plan. Li et al. [70] set up a clinical adaptive workflow for prostate cancer patients treated with IMRT, using an in-room diagnostic CT. The structures delineated on the initial planning CT scan were mapped to the second CT scan using deformable image registration. For this purpose, a segmentation and registration toolkit was incorporated in the treatment planning system (TPS). Compared to manual prostate contouring in daily scans, Dice coefficients of 0.84 were obtained for autosegmentation. To avoid the need for online manual reviewing and editing, while maintaining adequate target coverage, they automatically expanded the prostate autocontours by 3 mm, prior to automated generation of the plan-of-the-day.

## 8.3 Current clinical use and impact of Erasmus-iCycle

The first clinical application of Erasmus-iCycle was during our prospective head-and-neck cancer study (Chapter 5). For the participating patients, the research optimization algorithm, Erasmus-iCycle, was used to fully automatically generate a treatment plan, which was then accurately reconstructed in the clinical TPS by one of the investigators, resulting in a clinically deliverable plan, highly similar to the original Erasmus-iCycle plan. Based on the observed strong preference of treating physicians for the reconstructed Erasmus-iCycle treatment plans compared to fully manually generated plans,

we implemented Erasmus-iCycle in our standard routine for plan generation for head-and-neck cancer patients (in 2013: 450 plans for 300 patients). In the same period, some initial tests were performed using Erasmus-iCycle for cervical cancer patients. Also for this treatment site, Erasmus-iCycle proved to be useful as pre-optimization step, resulting in higher plan quality compared to manual planning from scratch. As from June 2011, Erasmus-iCycle is used for clinical plan generation for all cervical cancer patients (31 patients treated in 2013). In November 2013, we implemented fully automated VMAT planning for prostate cancer patients, as described in chapter 7. For this application, Erasmus-iCycle is still used as pre-optimization for generation of VMAT plans in the clinical TPS, but we now realized a fully automated reconstruction of Erasmus-iCycle plans in Monaco, avoiding manual plan reconstruction by dosimetrists as described in chapter 5 for head-and-neck cancer patients. Till February 1st, 2014, 25 patients were treated with this procedure. Erasmus-iCycle is currently also being used for linac-based SBRT for liver.

From the studies described in this thesis and our clinical experiences so far, we have concluded that semi-automated or fully automated plan generation, based on Erasmus-iCycle pre-optimization, results in equal or better plan quality than obtained with the common manual trial-and-error planning procedure. In all cases the planning workload is substantially reduced. In on-going projects, expansion of Erasmus-iCycle based planning toward new treatment sites is being prepared, and further use of fully automated instead of semi-automated plan generation is being explored. In the meantime, a user-friendly user interface has been developed to support broad clinical application.

## 8.4 Development of wishlists for Erasmus-iCycle plan generation

Basic input for each Erasmus-iCycle plan generation is the wishlist in which clinical objectives with assigned priorities and constraints are defined. Wishlists are generated for groups of patients with a similar treatment site and dose prescription. Each wishlist fully defines the automated plan generation for all patients in the group; no tweaking for individual patients is performed.

The workflow for development of a wishlist for a new treatment site is briefly summarized in Figure 8.1. The involved team consists of an expert dosimetrist or physicist and a group of radiation oncologists, treating these specific patients. In the first step, the

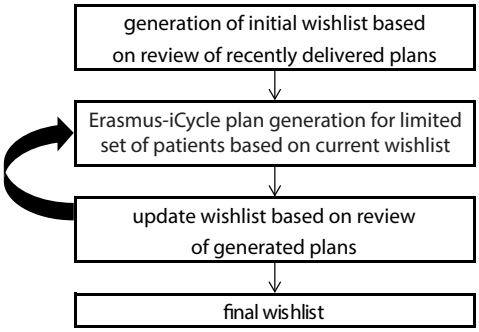



**Figure 8.1:** Workflow for generating a wishlist.

planning expert and the physicians review plans of recently treated patients. The goal is that the expert gets a clear view on applied hard constraints and treatment objectives, including their priorities. In addition, the expert tries to find out what physicians like and do not like in the current plans, what they would like to improve if they had the possibility, what their ideas are on the balance between OAR sparing and dose fall-off in unspecified tissues, etc. Based on this input, the initial wishlist is generated by the expert, and then further optimized in an iterative procedure (Figure 8.1). In this procedure, the wishlist is repeatedly updated based on Erasmus-iCycle plan generation with the current wishlist for a group of ~5 patients, followed by plan review (depending on the iteration with or without physicians). Initially, the focus is on obtaining clinically acceptable plans, while later iterations are focused on improvements compared to current clinical plan generation. At some point in the iterative procedure, the initially involved ~5 patients used for development of the wishlist, are substituted by ~5 new patients for an independent evaluation of the wishlist.

Clinical implementation of fully automated VMAT planning as described in chapter 7 included a final test of the quality of the wishlist, generated with the procedure described above. For the first 10 clinical patients, dosimetrists were asked to manually tweak the prescription parameters in the Monaco plan to investigate whether they could improve on the quality of the automatically generated plan. Only small improvements in rectum $V_{75\ Gy}$ (up to 0.8%) could be achieved at the cost of unfavorable dose increases in lower prioritized OARs and/or in a deteriorated dose fall-off in the unspecified tissues. When starting automated planning for a new treatment site, such a QA procedure is highly recommended for the first patients. Involvement of clinicians, dosimetrists and physicists that were not involved in establishment of the wishlist allows for an independent, complimentary check on plan quality.

## 8.5 (Possible) limitations of Erasmus-iCycle and future work

### Pareto navigation and clinical optimality

The 2pεc multicriterial optimizer in Erasmus-iCycle generates for each patient a single Pareto-optimal plan for the final selected beam configuration, the so-called wishplan. For every patient, the wishlist, as established for the patient group, is the driving force for generation of this particular wishplan, instead of another Pareto-optimal solution. In our experience, the wishplans have very high quality compared to manual planning, and there are no indications that plans with (substantially) improved quality do exist. On the other hand, we did not systematically investigate this for large groups of patients: neither do we currently have tools available to carefully perform these investigations. To allow this type of analyses, we are currently extending Erasmus-iCycle with a sensitivity analysis tool, i.e. a Pareto-navigation tool, like described by Monz et al. [71] and Thieke et al. [43]. For future research, we intend to use the wishplan generated by Erasmus-iCycle as starting point for navigation, since we believe that this plan is already of good quality. One of the research objectives is to find out whether Pareto navigation can be limited to an area close to the wishplan.

### Faster alternative for the current 2pεc optimizer in Erasmus-iCycle

Currently, the 2pεc algorithm is used for stepwise optimization of individual objectives in the wishlist according to their priorities. For more complex cases like in head-and-neck cancer, a wishlist may consist of more than 20 objectives, requiring a large number (>20) of consecutive optimizations in Erasmus-iCycle. In Erasmus-iCycle, the typical clinical optimization time in Erasmus-iCycle for a plan with 23 fixed equiangular beams, as used in the pre-optimization phase of automated prostate VMAT planning (chapter 7), is 35 minutes. Optimization of treatment plans for head-and-neck cancer patients, including beam angle optimization, takes about 1 hour for 9 coplanar beams, and up to 12 hours for 9 noncoplanar beams (both techniques using BAO). Recently, we have evaluated the lexicographic reference point method (LRPM) [72] as alternative for the 2pεc algorithm. The main advantage of LRPM is that it requires only one single optimization, even for a large wishlist. By adding lexicographic ordering of the objectives

based on the assigned priorities in the wishlist, prioritized optimization was retained with LRPM, like in the 2pεc method. Application of LRPM was tested retrospectively for the 30 prostate cancer patients used in chapter 7. Differences in obtained plan parameters compared to 2pεc optimization were neither clinically, nor statistically, significant. With LRPM, optimization times reduced by a factor of nearly 12 (from 34.9 to 3.0 minutes on average). Currently, we are further evaluating the application of LRPM.

## 8.6 Changing role of dosimetrists

For several treatment sites we have demonstrated that automated plan generation results in a consistent, high plan quality. It is expected that this will hold for most treatment sites. Large scale clinical introduction will result in important shifts in the activities of dosimetrists. They will get responsibilities for generating and tuning wishlists for difficult, non-standard patients, and for adapting existing wishlists when clinical treatment protocols are modified. Overall, for a fixed patient load, the required planning manpower will reduce. This does not mean that the need for highly qualified technicians (often (partly) working as dosimetrists in The Netherlands) will reduce. Instead of treatment planning, they will have to play an important role in the development and execution of complex novel treatment techniques such as adaptive radiotherapy. Recently, Bondar et al. [73] presented an online-adaptive IMRT technique for cervical cancer patients, based on a pretreatment established individualized plan library, generated with variable bladder filling CT scans. Since the beginning of the clinical implementation, in June 2011, 85 patients have been treated according to this protocol. The technicians at the linac now have to select in each fraction the most suitable plan, based on in-room imaging. Ultimately, for real-time online-adaptive treatment, a plan-of-the-day treatment will consist of daily acquisition of an in room (cone beam) CT or MR, real-time online re-contouring, re-planning, and plan evaluation and QA, which should all take place in a time-frame of several minutes. Highly trained technicians are indispensable for safe and effective application of such protocols.

## 8.7 Conclusion

In this thesis we have described some major advantages of automatic contouring and treatment planning. In the coming years, the interest and clinical application of these techniques are expected to grow extensively. Important new applications are expected in the upcoming field of dose-adaptive radiotherapy, which might be applied on a daily base.

# references


1. Levendag P, Braaksma M, Coche E, van der Est H, Hamoir M, Muller K, Noever I, Nowak P, van Sörnsen de Koste J, Gregoire V:" Rotterdam and Brussels CT-based neck nodal delineation compared with the surgical levels as defined by the American Academy of Otolaryngology-Head and Neck Surgery" in *Int. J. Radiat. Oncol. Biol. Phys.* 58(1): 113–123 (2004).
2. Teguh D, Levendag P, Voet PWJ, van der Est H, Noever I, de Kruijf W, van Rooij P, Schmitz P, Heijmen BJM: "Trismus in patients with oropharyngeal cancer: "relationship with dose in structures of mastication apparatus" in *Head Neck.* 30(5): 622–630 (2008).
3. Teguh D, Levendag P, Noever I, van Rooij P, Voet PWJ, van der Est H, Sipkema D, Sewnaik A, de Jong RJ, de la Bije D, Schmitz PI: "Treatment techniques and site considerations regarding dysphagia-related quality of life in cancer of the oropharynx and nasopharynx" in *Int. J. Radiat. Oncol. Biol. Phys.* 72(4): 1119–1127 (2008).
4. Levendag P, Teguh D, Voet PWJ, van der Est H, Noever I, de Kruijf W, Kolkman-Deurloo IK, Prevost JB, Poll J, Schmitz P, Heijmen BJM: "Dysphagia disorders in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscle: a dose-effect relationship" in *Radiother. Oncol.* 85(1): 64–73 (2007).
5. Hong T, Tome W, Chappell R, Harari P: "Variations in target delineation for head and neck IMRT: An international multi-institutional study" in *Int. J. Radiat. Oncol. Biol. Phys.* 60: S157-S158 (2004).
6. Jeanneret-Sozzi W, Moeckli R, Valley J, Zouhair A, Ozsahin E, Mirimanoff R: "The reasons for discrepancies in target volume delineation: a SASRO study on head-and-neck and prostate cancers" in *Strahlenther. Onkol.* 182(8): 450–457 (2006).
7. Steenbakkers R, Duppen J, Fitton I. Deurloo K, Zijp L, Eisbruch A, Nowak P, van Herk M, Rasch C: "Observer variation in delineation of nasopharyngeal carcinoma for radiotherapy, A 3-D analysis" in *Int. J. Radiat. Oncol. Biol. Phys.* 60: S160-S161 (2004).
8. Breedveld S, Storchi P, Voet PWJ, Heijmen BJM: "iCycle: integrated beam angle and profile optimization for generation of Pareto-optimal coplanar and noncoplanar IMRT plans" in *Med. Phys.* 2012; 39(2):951
9. Commowick O, Gregoire V, Malandain G: "Atlas-based delineation of lymph node levels in head and neck computed tomography images" in *Radiother. Oncol.* 87(2): 281–289 (2008).

10. Han, X Hoogeman M, Levendag P, Hibbard L, Teguh D, Voet PWJ, Cowen A, Wolf T: "Atlas-based autosegmentation of head and neck CT images" in *Med. Image Comput. Comput. Assist. Interv.* 11: 434–441 (2008).
11. Levendag P, Hoogeman M, Teguh D, Wolf T, Hibbard L, Wijers O, Heijmen BJM, Nowak P, Vásquez Osorio E, Han X: "Atlas Based Autosegmentation of CT Images: Clinical Evaluation of using Autocontouring in High-dose, High-precision Radiotherapy of Cancer in the Head and Neck" in *Int. J. Radiat. Oncol. Biol. Phys.* 72: S401 (2008).
12. Gregoire V, Levendag P, Ang K, Bernier J, Braaksma M, Budach V, Chao, C, Coche E, Cooper J, Cosnard G, Eisbruch A, El Sayed S, Emami B, Grau C, Hamoir M, Lee N, Maingon P, Muller K, Reychler H: "CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC,RTOG consensus guidelines" in *Radiother. Oncol.* 69(3): 227–236 (2003).
13. Levendag P, Al-Mamgani A, Teguh D: "Contouring in Head & Neck Cancer". *Elsevier Professional Education.* ISBN: 978-3-437-59904-0 (2009).
14. Warfield S, Zou K, H., Wells W: "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation" in *IEEE Trans. Med. Imaging* 23(7): 903–921 (2004).
15. Castadot P, Lee J, Parraga A, Geets X, Macq B, Gregoire V: "Comparison of 12 deformable registration strategies in adaptive radiation therapy for the treatment of head and neck tumors" in *Radiother. Oncol.* 89(1): 1–12 (2008).
16. Chao K, Bhide S, Chen H, Asper J, Bush S, Franklin G, Kavadi V, Liengswangwong V, Gordon W, Raben A, Strasser J, Koprowski C, Frank S, Chronowski G, Ahamad A, Malyapa R, Zhang L, Dong L: "Reduce in variation and improve efficiency of target volume delineation by a computer-assisted system using a deformable image registration approach" in *Int. J. Radiat. Oncol. Biol. Phys.* 68(5): 1512–1521 (2007).
17. Zhang T, Chi Y, Meldolesi E, Yan D: "Automatic delineation of on-line head-and-neck computed tomography images: toward on-line adaptive radiotherapy" in *Int. J. Radiat. Oncol. Biol. Phys.* 68(2): 522–530 (2007).
18. Sims R, Isambert A, Gregoire V, Bidault F, Fresco L, Sage J, Mills J, Bourhis J, Lefkopoulos D, Commowick O, Benkebil M, Malandain G: "A pre-clinical assessment of an atlas-based automatic segmentation tool for the head and neck" in *Radiother. Oncol.* 93(3): 474–8 (2009).
19. Wang H, Garden A, Zhang L, Wei X, Ahamad A, Kuban D. A, Komaki R, O'Daniel J, Zhang Y, Mohan R, Dong L: "Performance evaluation of automatic anatomy segmentation algorithm on repeat or four-dimensional computed tomography images using deformable image registration method" in *Int. J. Radiat. Oncol. Biol. Phys.* 72(1): 210–219 (2008).
20. Vásquez Osorio E, Hoogeman M, Al Mamgani A, Teguh D, Levendag P, Heijmen BJM: "Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration" in *Int. J. Radiat. Oncol. Biol. Phys.* 70(3): 875–882 (2008).
21. Vásquez Osorio E, Hoogeman M, Bondar L, Levendag P, Heijmen BJM: "A novel flexible framework with automatic feature correspondence optimization for nonrigid registration in radiotherapy" in *Med. Phys.* 36(7): 2848–2859 (2009).
22. Qi X, Semenenko V, Li X: "Improved critical structure sparing with biologically based IMRT optimization" in *Med. Phys.* 36(5): 1790–9 (2009).
23. Huyskens D, Maingon P, Vanuytsel L, Remouchamps V, Roques T, Dubray B, Haas B, Kunz P, Coradi T, Bühlman R, Reddick R, Esch AV, Salamon E: "A qualitative and a quantitative analysis of an autosegmentation module for prostate cancer" in *Radiother. Oncol.* 90(3): 337–345 (2009).

24. Isambert A, Dhermain F, Bidault F, Commowick O, Bondiau P, Malandain G, Lefkopoulos D: "Evaluation of an atlas-based automatic segmentation software for the delineation of brain organs at risk in a radiation therapy clinical context" in *Radiother. Oncol.* 87(1): 93–99 (2008).
25. Teguh D, Levendag P, Voet PWJ, Al-Mamgani A, Han X, Wolf T, Hibbard L, Nowak P, Akhiat A, Dirkx MLP, Heijmen BJM, Hoogeman M: "Clinical Validation of Atlas-based Autosegmentation of Multiple Target Volumes and Normal Tissue Structures (swallowing and mastication) in the Head & Neck" in *Int. J. Radiat. Oncol. Biol. Phys.* 15:81(4): 950–7 (2011).
26. Han X, Hoogeman M, Levendag P, Hibbard L, Teguh D, Voet PWJ, Cowen A, Wolf T: "Atlas-based of head and neck CT images" in *Med Image Comput. Comput. Assist. Interv*. 11: 434–41 (2008).
27. Dice L: "Measures of the Amount of Ecologic Association Between Species" in *Ecology* 26(3): 297–302 (1945).
28. Tsuji S, Hwang A, Weinberg V, Yom S, Quivey J, Xia P: "Dosimetric Evaluation of Automatic Segmentation for Adaptive IMRT for Head-and-Neck Cancer" in *Int. J. Radiat. Oncol. Biol. Phys.* 77(3):707–14 (2010).
29. Zijdenbos A, Dawant B, Margolin M, Palmer A: "Morphometric analysis of white matter lesions in MR images: method and validation" in *IEEE Trans. Med. Imaging* 13: 716–24 (1994).
30. Meedt G, Alber A, Nüsslin F: "Noncoplanar beam direction optimization for intensity modulated radiotherapy" in *Phys. Med. Biol.*48(18):2999–3019 (2003).
31. Djajaputra D, Wu Q, Wu Y, Mohan R: "Algorithm and performance of a clinical IMRT beam-angle optimization system" in *Phys. Med. Biol.* 48(19):3191–3212 (2003).
32. Wang X, Zhang X, Dong L, Liu H, Gillin M, Ahamad A, Ang K, Mohan R: "Effectiveness of noncoplanar IMRT planning using a parallelized multiresolution beam angle optimization method for paranasal sinus carcinoma" in *Int. J. Radiat. Oncol. Biol. Phys.* 63(2):594–601 (2005).
33. Nutting C, Rowbottom C, Cosgrove V, Henk J, Dearnaley D, Robinson M, Conway J, Webb S: "Optimization of radiotherapy for carcinoma of the parotid gland: a comparison of conventional, three dimensional conformal and intensity modulated techniques" in *Radiother. Oncol.*;60(2):163–172 (2001).
34. Woudstra E, Storchi P: "Constrained treatment planning using sequential beam selection" in *Phys. Med. Biol.*45(8):2133–49 (2000).
35. De Pooter J, Méndez Romero A, Wunderink W, Storchi P, Heijmen BJM: "Automated noncoplanar beam direction optimization improves IMRT in SBRT of liver metastasis" in *Radiother. Oncol.*88(3):376–381 (2008).
36. De Pooter J, Méndez Romero A, Jansen W, Storchi P, Woudstra E, Levendag P, Heijmen BJM: "Computer optimization of noncoplanar beam setups improves stereotactic treatment of liver tumors" in *Int. J. Radiat. Oncol. Biol. Phys.*66(3):913–922 (2006).
37. Breedveld S, Storchi P, Keijzer M, Heemink A, Heijmen BJM: "A novel approach to multicriteria inverse planning for IMRT" in *Phys. Med. Biol.*52(20):6339–6353 (2007).
38. Breedveld S, Storchi P, Heijmen BJM: "The equivalence of multicriteria methods for radiotherapy plan optimization" in *Phys. Med. Biol.*54(23)**:**7199–7209 (2009).
39. Dijkema T, Raaijmakers C, Ten Haken R, Roesink J, Braam P, Houweling A, Moerland M, Eisbruch A, Terhaard C: "Parotid gland function after radiotherapy: the combined Michigan and Utrecht experience" in *Int. J. Radiat. Oncol. Biol. Phys.*78(2):449–453 (2010).
40. Murdoch-Kinch C, Kim H, Vineberg K, Ship J, Eisbruch A: "Dose-effect realtionships for the submandibular salivary glands and implications for their sparing by intensity modulated radiotherapy" in *Int. J. Radiat. Oncol. Biol. Phys.* 72(2): 373–382 (2008).
41. Alber M, Reemtsen R: "Intensity modulated radiotherapy treatment planning by use of a barrier-penalty multiplier method" in *Optim Methods Softw* 22:391–411 (2007).

42. McDonald J: "Handbook of Biological Statistics (2nd ed.)". *Sparky House Publishing, Baltimore, Maryland* pages 198–201 (2009).
43. Thieke C, Küfer K, Monz M, Scherrer A, Alonso F, Oelfke U, Huber P, Debus J, Bortfeld T: "A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation" in *Radiother. Oncol.* 85(2):292–298 (2007).
44. Craft D, Hong S, Shih H, Bortfeld T: "Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy" in *Int. J. Radiat. Oncol. Biol. Phys.* 82(1):e83–90 (2012).
45. Voet PWJ, Dirkx MLP, Breedveld S, Fransen D, Levendag P, Heijmen BJM: "Toward fully automated multicriterial plan generation: a prospective clinical study" in *Int J Radiat Oncol Biol Phys*; 85(3): 866–872 (2013).
46. Teichert K, Süss P, Serna J, Monz M, Küfer K, Thieke C: "Comparative analysis of Pareto surfaces in multicriteria IMRT planning" in *Phys. Med. Biol.*56(12): 3669–3684 (2011).
47. Zhang X, Li X, Quan E, Pan X, Li Y: "A methodology for automatic intensity-modulated radiation treatment planning for lung cancer" in *Phys. Med. Biol.*56(13): 3873–3893 (2011).
48. Otten R , van Roermund P, Picavet H: "Trend in the number of knee and hip arthroplasties: considerably more knee and hip prostheses due to ostheoarthritis in 2030" in *Ned. Tijdschr. Geneeskd*. 154:A1534 (2010).
49. Reft C, Alecu R, Das I, Gerbi B, Keall P, Lief E, Mijnheer B, Papanikolaou N, Sibata C, Dyk J: "Dosimetric considerations for patients with HIP prostheses undergoing pelvic irradiation. Report of the AAPM Radiation Therapy Committee Task Group 63" in *Med. Phys.*; 30(6):1162–1182 (2003).
50. Van der Est H, Prins P, Heijmen BJM, Dirkx MLP: "Intensity modulated radiatio therapy for patients with a metal hip prosthesis based on class solutions" in *Practical Radiation Oncology* 2:35–40 (2012).
51. Gay H, Barthold H, O'Meara E, Bosch W, el Naqa I, Al-Lozi R, Rosenthal S, Lawton C, Lee W, Sandler H, Zietman A, Meyerson R, Dawson L, Willett C, Kachnic L, Jhingran A, Portelance L, Ryu J, Small jr W, Gaffney D, Viswanathan A, Michalski J: "Pelvic normal tissue contouring guidelines for radiation therapy: a radiation therapy oncology group consensus panel atlas" in *Int. J. Radiat. Oncol. Biol. Phys.*; 83(3):353–362 (2012).
52. Kling J, Patel K: "Prostate treatment with helical TomoTherapy in patients with bilateral hip prostheses - two case studies" in *Med. Dos.* 38: 30–34 (2013).
53. Mutanga T, de Boer H, van der Wielen G, Wentzler D, Barnhorn J, Incrocci L, Heijmen BJM: "Stereographic targeting in prostate radiotherapy: speed and precision by daily automatic positioning corrections using kilovoltage/megavoltage image pairs" in *Int. J. Radiat. Oncol. Biol. Phys.*; 71(4):1074–1083 (2008).
54. Mutanga T, de Boer H, Rajan V, Dirkx MLP, Incrocci L, Heijmen BJM: "Day-to-day reproducilbility of prostate intrafraction motion assessed by multiple kV and MV imaging of implanted markers during treatment" in *Int. J. Radiat. Oncol. Biol. Phys.*; 83(1):400–407 (2012).
55. Michalski J, Gay H, Jackson A, Tucker S, Deasy J: "Radiation dose-volume effects in radiation-induced rectal injury" in *Int. J. Radiat. Oncol. Biol. Phys.*; 76(3):123–129 (2010).
56. Viswanathan A, Yorke E, Marks L, Eifel P, Shipley W: "Radiation dose-volume effects of the unrinary bladder" in *Int. J. Radiat. Oncol. Biol. Phys.*;76(3):116–122 (2010).
57. Brooks C, Cheung R, Kudchadker R: "Intensity-modulated radiation therapy with noncoplanar beams for treatment of prostate cancer in patients with bilateral hip prosthesis" in *Med. Dos.* 35(2): 87–91 (2009).
58. Sun M, Ma I: "Treatment of exceptionally large prostate cancer patients with low-energy intensity modulated photons" in *J. Appl. Clin. Med. Phys.* 7(4), 43–49 (2006).
59. Hall W, Fox TH, Jiang X, Prabhu R, Rossi P, Godette K, Jani A: "Treatment efficiency of volumetric modulated arc therapy in comparison with intensity-modulated radiotherapy in the treatment of prostate cancer" in *J. Am. Coll. Radiol.* 10(2): 128–134 (2013).

60. Voet PWJ, Dirkx MLP, Breedveld S, Heijmen BJM: "Automated generation of IMRT treatment plans for prostate cancer patients with metal hip prostheses: comparison of different planning strategies." in *Med. Phys.*2013; 40(7):071704.
61. Rossi L, Breedveld S, Heijmen BJM, Voet PWJ, Lanconelli N, Aluwini S: "On the beam direction search space in computerized non-coplanar beam angle optimization for IMRT-prostate SBRT" in *Phys. Med. Biol.* 57(17):5441–5458 (2012).
62. Mutanga TF, de Boer HC, van der Wielen GJ, Hoogeman M, Incrocci L, Heijmen BJM: "Margin evaluation in the presence of deformation, rotation, and translation in prostate and entire seminal vesicle irradiation with daily marker-based setup corrections." in *Int. J. Radiot. Oncol. Biol. Phys.*2011; 81(4): 1160–1167.
63. Niemierko A, Goitein M: "Modeling of normal tissue response to radiation: The critical volume model." in *Int. J. Radiot. Oncol. Biol. Phys.*1993; 25(1): 135–145.
64. Quan EM, Li X, Li Y, Wang X, Kudchadker R, Johnson J, Kuban D, Lee A, Zhang X: "A comprehensive comparison of IMRT and VMAT plan quality for prostate cancer treatment." in *Int. J. Radiat. Oncol. Biol. Phys.*2012; 83(4): 1169–1178.
65. Wu B, Pang D, Simari P, Taylor R, Sanquineti G, McNutt T: "Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study" in *Med. Phys.* 40(2): 021714 (2013).
66. Anders L, Stieler F, Siebenlist K, Schäfer J, Lohr F, Wenz F: "Performance of an atlas based autosegmentation software for delineation of target volumes for radiotherapy for breast and anorectal cancer" in *Radiother. Oncol.* 102(1), 68–73 (2011).
67. Gambacorta M, Valentini C, Dinapoli N, Boldrini L, Caria N, Barba MC, Mattiucci GC, Pasini D, Minsky B, Valentini V: "Clinical validation of atlas-based auto-segmentation of pelvic volumes and normal tissue in rectal tumors using auto-segmentation computed system" in *Acta Oncol* 52(8): 1676–81 (2013).
68. Young A, Wortham A, Wernick I, Evans A, Ennis R: "Atlas-based segmentation improves consistency and decreases time required for contouring postoperative endometrial cancer nodal volumes" in *Int. J. Radiat. Oncol. Biol. Phys.* 79(3):943–7 (2011).
69. Simmat I, Georg P, Georg D, Birkfellner W, Goldner G, Stock M: "Assessment of accuracy and efficiency of atlas-based autosegmentation for prostate radiotherapy in a variety of clinical conditions" in *Strahlenther. Onkol.*188(9):807–815 (2012).
70. Li X, Quan E, Li Y, Pan X, Zhou Y, Wang X, Du W, Kudchadker R, Johnson J, Kuban D, Lee A, Zhang X: "A fully automated method for CT-on-rails-guided online adaptive planning for prostate cancer intensity modulated radiation therapy" in *Int. J. Radiat. Oncol. Biol. Phys.* 86(5): 835–841 (2013).
71. Monz M, Küfer KH, Bortfeld T, Thieke C: "Pareto navigation: algorithmic foundation of interactive multi-criteria IMRT planning" in *Phys. Med. Biol.*53(4):985–98 (2008).
72. Ogryczak W, Kozlowski B: "Reference point method with importance weighted ordered partial achievements" in *TOP: Aan Official Journal of the Spanish Society of Statistics and Operations Research* 19(2):380–401 (2011).
73. Bondar M, Hoogeman M, Mens JW, Quint S, Ahmad R, Dhawtal G, Heijmen BJM: "Individualized non-adaptive and online-adaptive intensity-modulated radiotherapy treatment strategies for cervical cancer patients based on pretreatment acquired variable bladder filling computed tomography scans" in *Int. J. Radiat. Oncol. Biol. Phys.*83(5):1617–1623 (2012).

# list of peer-reviewed publications

## First author publications

1. Does atlas-based autosegmentation of neck levels require subsequent manual contour editing to avoid risk of severe target underdosage? A dosimetric analysis.
   **Voet PWJ**, Dirkx MLP, Teguh DN, Hoogeman MS, Levendag PC, Heijmen BJM.
   Radiother Oncol. 98(3):373–7 (2011).
2. Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head-and-neck IMRT and implications for VMAT.
   **Voet PWJ**, Breedveld S, Dirkx MLP, Levendag PC, Heijmen BJM.
   Med Phys.39(8):4858–65 (2012).
3. Toward fully automated multicriterial plan generation: a prospective clinical study.
   **Voet PWJ**, Dirkx MLP, Breedveld S, Fransen D, Levendag PC, Heijmen BJM.
   Int J Radiat Oncol Biol Phys. 85(3):866–72 (2013).
4. Automated generation of IMRT treatment plans for prostate cancer patients with metal hip prostheses: comparison of different planning strategies.
   **Voet PWJ**, Dirkx MLP, Breedveld S, Heijmen BJM.
   Med Phys. 40(7):071704 (2013).
5. Fully automated VMAT plan generation for prostate cancer patients.
   **Voet PWJ**, Dirkx MLP, Breedveld S, Al-Mamgani A, Incrocci L, Heijmen BJM.
   Int J Radiat Oncol Biol Phys. 88(5):1175–79 (2014).

## Second author publications

1. An evaluation of two techniques for beam intensity modulation in patients irradiated for stage III non-small cell lung cancer.
   van Sörnsen de Koste J, **Voet PWJ**, Dirkx MLP, van Meerbeeck J, Senan S; Rotterdam Oncological Thoracic Study Group.
   Lung Cancer. 32(2):145–53 (2001).
2. Curative radiotherapy for a second primary lung cancer arising after pneumonectomy – techniques and results.
   Lagerwaard FJ, **Voet PWJ**, van Meerbeeck JP, Burgers SA, Senan S; Rotterdam Oncological Thoracic Study Group.
   Radiother Oncol. 62(1):21–5 (2002).
3. Four-dimensional stereotactic radiotherapy for early stage non-small cell lung cancer: a comparative planning study.
   Prevost JB, **Voet PWJ**, Hoogeman M, Praag J, Levendag P, Nuyttens JJ.
   Technol Cancer Res Treat. 7(1):27–33 (2008).

## Other publications

1. Dosimetric consequences of tumor mobility in radiotherapy of stage I non-small cell lung cancer -an analysis of data generated using 'slow' CT scans.
   van Sörnsen de Koste JR, Lagerwaard FJ, Schuchhard-Schipper RH, Nijssen-Visser MR, **Voet PWJ**, Oei SS, Senan S.
   Radiother Oncol. 61(1):93–9 (2001).
2. Can elective nodal irradiation be omitted in stage III non-small-cell lung cancer? Analysis of recurrences in a phase II study of induction chemotherapy and involved-field radiotherapy.
   Senan S, Burgers S, Samson MJ, van Klaveren RJ, Oei SS, van Sörnsen de Koste J, **Voet PWJ**, Lagerwaard FJ, Maarten van Haarst J, Aerts JG, van Meerbeeck JP.
   Int J Radiat Oncol Biol Phys. 54(4):999–1006 (2002).
3. Can errors in reconstructing pre-chemotherapy target volumes contribute to the inferiority of sequential chemoradiation in stage III non-small cell lung cancer (NSCLC)?
   Lagerwaard FJ, van de Vaart PJ, **Voet PWJ**, Nijssen-Visser MR, Schuchhard-Schipper RH, Joosten HP, Oei SS, Senan S.
   Lung Cancer. 38(3):297–301 (2002).

4. Intraoperative validation of CT-based lymph nodal levels, sublevels IIa and IIb: is it of clinical relevance in selective radiation therapy?
   Levendag P, Gregoire V, Hamoir M, **Voet PWJ**, van der Est H, Heijmen BJM, Kerrebijn J.
   Int J Radiat Oncol Biol Phys. 62(3):690–9 (2005).
5. Dysphagia disorders in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscle: a dose-effect relationship.
   Levendag PC, Teguh DN, **Voet PWJ**, van der Est H, Noever I, de Kruijf WJ, Kolkman-Deurloo IK, Prevost JB, Poll J, Schmitz PI, Heijmen BJM.
   Radiother Oncol. 85(1):64–73 (2007).
6. Atlas-based auto-segmentation of head and neck CT images.
   Han X, Hoogeman MS, Levendag PC, Hibbard LS, Teguh DN, **Voet PWJ**, Cowen AC, Wolf TK.
   Med Image Comput Comput Assist Interv.11(Pt 2):434–41 (2008).
7. Trismus in patients with oropharyngeal cancer: relationship with dose in structures of mastication apparatus.
   Teguh DN, Levendag PC, **Voet PWJ**, van der Est H, Noever I, de Kruijf W, van Rooij P, Schmitz PI, Heijmen BJM.
   Head Neck. 30(5):622–30 (2008).
8. Treatment techniques and site considerations regarding dysphagia-related quality of life in cancer of the oropharynx and nasopharynx.
   Teguh DN, Levendag PC, Noever I, van Rooij P, **Voet PWJ**, van der Est H, Sipkema D, Sewnaik A, Baatenburg de Jong RJ, de la Bije D, Schmitz PI.
   Int J Radiat Oncol Biol Phys. 72(4):1119–27 (2008).
9. Analysis of the motion of oropharyngeal tumors and consequences in planning target volume determination.
   Prévost JB, de Boer H, Pöll J, **Voet PWJ**, Levendag P.
   Radiother Oncol. 87(2):268–73 (2008).
10. Results of fiberoptic endoscopic evaluation of swallowing vs. radiation dose in the swallowing muscles after radiotherapy of cancer in the oropharynx.
    Teguh DN, Levendag PC, Sewnaik A, Hakkesteegt MM, Noever I, **Voet PWJ**, van der Est H, Sipkema D, van Rooij P, Baatenburg de Jong RJ, Schmitz PI.
    Radiother Oncol. 89(1):57–63 (2008).
11. Early hyperbaric oxygen therapy for reducing radiotherapy side effects: early results of a randomized trial in oropharyngeal and nasopharyngeal cancer.
    Teguh DN, Levendag PC, Noever I, **Voet PWJ**, van der Est H, van Rooij P, Dumans AG, de Boer MF, van der Huls MP, Sterk W, Schmitz PI.
    Int J Radiat Oncol Biol Phys. 75(3):711–6 (2009).

12. Clinical validation of atlas-based auto-segmentation of multiple target volumes and normal tissue (swallowing/mastication) structures in the head and neck.
Teguh DN, Levendag PC, **Voet PWJ**, Al-Mamgani A, Han X, Wolf TK, Hibbard LS, Nowak P, Akhiat H, Dirkx MLP, Heijmen BJM, Hoogeman MS.
Int J Radiat Oncol Biol Phys. 15;81(4):950–7 (2011).
13. IMRT for image-guided single vocal cord irradiation.
Osman SO, Astreinidou E, de Boer HC, Keskin-Cambay F, Breedveld S, **Voet PWJ**, Al-Mamgani A, Heijmen BJM, Levendag PC.
Int J Radiat Oncol Biol Phys. 82(2):989–97 (2012).
14. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.
Breedveld S, Storchi PR, **Voet PWJ**, Heijmen BJM.
Med Phys. 39(2):951–63 (2012).
15. On the beam direction search space in computerized non-coplanar beam angle optimization for IMRT-prostate SBRT.
Rossi L, Breedveld S, Heijmen BJM, **Voet PWJ**, Lanconelli N, Aluwini S.
Phys Med Biol. 57(17):5441–58 (2012).
16. A margin-of-the-day online adaptive intensity-modulated radiotherapy strategy for cervical cancer provides superior treatment accuracy compared to clinically recommended margins: A dosimetric evaluation.
Ahmad R, Bondar L, **Voet PWJ**, Mens JW, Quint S, Dhawtal G, Heijmen BJM, Hoogeman M.
Acta Oncol. 52(7):1430–6 (2013).

# summary

This thesis focuses on the automation of tasks in the preparation phase of a radiotherapy treatment, resulting in equal, or preferentially higher, treatment quality with less human interference. Chapter 1 is a general introduction into radiotherapy treatment, and gives a short description of the performed research projects. Chapter 2 describes a clinical validation study of atlas-based autosegmentation (ABAS) for head- and-neck cancer patients. In chapter 3, we address the question if it is necessary to review and edit structure sets generated by autocontouring software before using them for treatment planning. Chapter 4 describes the use of Erasmus-iCycle for an objective investigation on the impact of integrated, computerized optimization of beam orientations and IMRT fluence profiles on salivary gland sparing in head-and-neck cancer patients. Based on the results of this work, we performed a randomized, prospective clinical study comparing automatically and manually generated intensity modulated radiotherapy treatment (IMRT) plans for head-and-neck cancer patients. Chapter 5 summarizes the results of this study. In chapter 6 we compare different IMRT planning strategies for prostate cancer patients with metal hip prostheses using Erasmus-iCycle. Chapter 7 describes fully automated volumetric modulated arc therapy (VMAT) plan generation for prostate cancer patients. Chapter 8 is a general discussion related to the main topics described in this thesis and describes some directions for further research.

## Chapter 2

We performed a clinical validation study on ABAS for delineation of multiple clinical target volumes and normal tissue structures for head-and-neck cancer patients. Two dif-

ferent ABAS strategies, using either a single atlas or atlases from multiple subjects, were evaluated. Using the leave-one-out method, the similarity of the generated autocontours and the (input) atlas contours was assessed using Dice coefficients and mean distances to agreement. For 12 clinically treated patients, 5 experienced observers edited the autosegmented contours. The editing times were recorded. Dice coefficients and mean distances were calculated among clinically used contours, autocontours and edited autocontours. Finally, an expert panel scored all autocontours and the edited autocontours regarding their adequacy relative to the published atlas.

ABAS needed 7 minutes per patient to autosegment all structures. Differences in the autosegmentation accuracy were not significantly different between No and N+ patients. The multi-subject atlas performed best, yielding a Dice coefficient and mean distance of 0.74 and 2 mm for salivary glands, 0.67 and 3 mm for neck levels, 0.71 and 2 mm for chewing muscles, 0.50 and 2 mm for swallowing muscles, and 0.78 and 2 mm for the cord and brainstem. For the neck levels and salivary glands the mean Dice coefficient and the mean distance between the autocontours and clinically applied contours was 0.8 and 2.4 mm, respectively. Between the autocontours and the edited autocontours, those values were 0.9 and 1.6 mm, respectively. The expert panel scored all autocontours as "minor deviation, editable" or better. 88% of the edited contours were scored as "good" opposed to 83% for the clinically used contours.

Multiple-subject ABAS of CT images proved to be a useful novel tool for fast delineation of targets and normal tissues. Although editing of the autocontours turned out to be inevitable, a substantial time reduction was achieved when compared to manual contouring from scratch (66 versus 180 min).

## Chapter 3

In this study, the dosimetric impact of editing autocontours of elective neck CTVs and organs-at-risk (OARs), generated by ABAS, was investigated for head-and-neck cancer patients. For nine patients, structure sets generated by ABAS and edited autocontours by two observers were available. Based on the non-edited autocontours clinically acceptable IMRT plans (designated 'ABAS plans') were generated. These plans were then evaluated for the two edited structure sets. Dice coefficients and mean contour distances were calculated to quantify the similarity of ABAS autocontours and the structure sets edited by observer 1 and observer 2. To study the dosimetric importance of editing OAR autocontours a new IMRT plan was generated for each patient-observer combination, based on the CTV edited by the observer and the non-edited salivary gland autocontours. For each plan, mean doses for the non-edited and edited glands were compared.

For both observers, edited neck CTVs were larger than ABAS autocontours, by 8.7% on average ($p \leq 0.04$). When evaluating ABAS plans on the PTVs of the edited structure sets, $V_{95\%}$ reduced by 7.2% ± 5.4% (1 SD) ($p < 0.03$). The mean reduction in $D_{99\%}$ was 14.2 Gy (range 1–54 Gy). Even for Dice coefficients >0.8 and mean contour distances <1 mm, reductions in $D_{99\%}$ up to 11 Gy were observed. For treatment plans based on observer PTVs and non-edited autocontoured salivary glands, the mean doses in the edited glands differed by only -0.6 Gy ± 1.0 Gy ($p = 0.06$).

This shows that editing of autocontoured neck CTVs generated by ABAS is required to avoid large underdosage in target volumes. Often used similarity measures for evaluation of autocontouring algorithms, such as Dice coefficients, do not predict well for expected PTV underdose. Editing of salivary glands is less important, as mean doses achieved for non-edited glands predict well for edited structures.

## Chapter 4

In this study, we quantified improved salivary gland sparing for head-and-neck cancer patients when using IMRT plans based on integrated computerized optimization of beam orientations and intensity profiles. The effect on plan quality using optimized non-zero couch angles in VMAT plans was assessed as well. Erasmus-iCycle was used for automated generation of multicriterial optimized plans with optimized beam orientations and intensity profiles, and plans with optimized profiles for preselected beam arrangements. For 20 patients, 5 IMRT plans, based on one 'wishlist', were compared: i), ii) seven- and nine-beam equiangular coplanar plans ($iCycle_{7equi}$, $iCycle_{9equi}$), iii), iv) nine-beam plans with optimized coplanar and noncoplanar beam orientations ($iCycle_{copl}$, $iCycle_{noncopl}$) and v) a nine-beam coplanar plan with optimized gantry angles and one optimized couch rotation ($iCycle_{couch}$). VMAT plans without and with this optimized couch rotation were evaluated.

Erasmus-iCycle allowed objective comparison of the different planning strategies. $iCycle_{noncopl}$ resulted in the best salivary gland sparing, but $iCycle_{couch}$ yielded similar results for 18 patients. For $iCycle_{7equi}$, submandibular gland NTCP values were on average 5% higher. $iCycle_{9equi}$ performed better than $iCycle_{7equi}$. $iCycle_{copl}$ showed further improvement, indicating the benefit of integrated optimization of beam angles and beam profiles. Application of the optimized couch angle from $iCycle_{couch}$ also improved NTCP values in VMAT plans.

## Chapter 5

In this prospective clinical study, we made a comparison of plans generated with Erasmus-iCycle and plans manually generated by dosimetrists with the clinical treatment planning system. For 20 randomly selected head-and-neck cancer patients with various tumor locations (of whom 13 received sequential boost treatments) we offered the treating physician the choice between an automatically generated Erasmus-iCycle plan and a manually optimized plan using standard clinical procedures. While Erasmus-iCycle used a fixed 'wishlist' with hard constraints and prioritized objectives, the dosimetrists manually selected the beam configuration and fine-tuned the constraints and objectives for each IMRT plan. Dosimetrists and treating physicians were not informed in advance whether a competing Erasmus-iCycle plan was made. The two plans were simultaneously presented to the physician, who then selected the preferred plan to be used for patient treatment. For the patient group, we also quantified differences in PTV coverage and sparing of critical tissues.

In 32/33 plan comparisons the physician selected the Erasmus-iCycle plan for treatment. This highly consistent preference for automatically generated plans was mainly caused by improved sparing for the large majority of critical structures. With Erasmus-iCycle, the normal tissue complication probabilities (NTCPs) for parotid and submandibular glands were reduced by 2.4% ± 4.9% (maximum: 18.5%, $p = 0.001$) and 6.5% ± 8.3% (maximum: 27%, $p = 0.005$), respectively. The reduction in mean oral cavity dose was 2.8 Gy ± 2.8 Gy (maximum: 8.1 Gy, $p = 0.005$). For swallowing muscles, esophagus and larynx, the mean dose reduction was 3.3 Gy ± 1.1 Gy (maximum: 9.2 Gy, $p < 0.001$). In addition, for 15 of the 20 patients, the target coverage was improved as well. Apart from improved plan quality, automatic plan generation is economically attractive because of reduced workload.

## Chapter 6

In this study, different IMRT planning strategies for prostate cancer patients with metal hip prostheses were compared. All plans were generated fully automatically (i.e., without human trial-and-error interactions) using Erasmus-iCycle, allowing objective comparison of planning strategies. For 18 prostate cancer patients (8 with bilateral hip prostheses, 10 with a right-sided unilateral prosthesis), two planning strategies were evaluated: i) full exclusion of beams containing beamlets that would deliver dose to the target after passing a prosthesis ($IMRT_{remove}$), and ii) exclusion of those beamlets only ($IMRT_{cut}$). Plans with optimized coplanar and noncoplanar beam arrangements were generated. Differences in

PTV coverage and sparing of OARs were quantified. The impact of beam number on plan quality was evaluated.

Especially for patients with bilateral hip prostheses, $IMRT_{cut}$ significantly improved rectum and bladder sparing compared to $IMRT_{remove}$. For 9-beam coplanar plans, rectum $V_{60\ Gy}$ reduced by 17.5% ± 15.0% (maximum 37.4%, $p = 0.036$) and rectum $D_{mean}$ by 9.4% ± 7.8% (maximum 19.8%, $p = 0.036$). Further improvements in OAR sparing were achievable by using noncoplanar beam set-ups, reducing rectum $V_{60\ Gy}$ by another 4.6% ± 4.9% ($p = 0.012$) for noncoplanar 9-beam $IMRT_{cut}$ plans. Large reductions in rectum dose delivery were also observed when increasing the number of optimized beam directions in the plans. For bilateral implants, the rectum $V_{60\ Gy}$ was 37.3% ± 12.1% for coplanar 7-beam plans and reduced on average by 13.5% (maximum 30.1%, $p = 0.012$) for 15-beam plans.

Erasmus-iCycle was able to automatically generate high quality plans for prostate cancer patients with metal hip prostheses. Excluding only beamlets that passed through the prostheses ($IMRT_{cut}$ strategy) significantly improved OAR sparing. Noncoplanar beam arrangements and, to a larger extent, increasing the number of treatment beams, further improved plan quality.

## Chapter 7

A system for fully automated VMAT treatment planning for prostate cancer patients was developed, evaluated, and clinically implemented. Erasmus-iCycle was directly linked with our clinical, commercially available TPS to fully automatically generate VMAT plans ($VMAT_{auto}$). For 30 randomly selected patients, $VMAT_{auto}$ plans were compared with VMAT plans generated manually by one expert dosimetrist in the absence of time pressure ($VMAT_{man}$). For all treatment plans, PTV coverage and sparing of OARs were quantified.

All generated plans were clinically acceptable and had similar PTV coverage ($V_{95\%}$ > 99%). For $VMAT_{auto}$ and $VMAT_{man}$ plans the OAR sparing was similar as well, while only the former plans were generated without any manual planning workload. For the first ten clinically treated patients, dosimetrists were asked to attempt improving the $VMAT_{auto}$ plans by manually tweaking the plan objectives in Monaco in a trial-and-error procedure, as normally done in clinical practice. Rectum $V_{75\ Gy}$ an rectum $V_{60\ Gy}$ were on average slightly lower in the tweaked plans (maximum difference 1%), but this minor improvement in rectum parameters generally resulted in an increased dose delivery to the other, lower prioritized OARs and/or in a deteriorated dose fall-off in unspecified tissues.

The results of this study clearly show that fully automated generation of high quality VMAT plans for prostate cancer patients is feasible. The procedure is currently routinely applied in our clinic.

## Chapter 8

This chapter presents a general discussion on the main research topics in this thesis. Current clinical applications of automated contouring and treatment planning are summarized. Amongst others, topics for further research are related to speeding up both procedures, evaluating them for new treatment sites and making them more widely applicable.

# samenvatting

Het onderzoek dat in dit proefschrift wordt beschreven heeft betrekking op de automatisering van verschillende taken in het voorbereidingstraject van een radiotherapie behandeling, met als doel om een gelijke, of bij voorkeur betere, kwaliteit van behandeling te verkrijgen met minder handmatige handelingen. Hoofdstuk 1 is een algemene introductie in de radiotherapie, en geeft een korte beschrijving van de uitgevoerde onderzoeksprojecten. Hoofdstuk 2 beschrijft de resultaten van een klinische validatiestudie van het gebruik van op atlassen gebaseerde (*atlas-based*) autosegmentatie software (ABAS) voor hoofd-hals kankerpatiënten. In hoofdstuk 3 beantwoorden we de vraag in hoeverre het noodzakelijk is om de met deze software automatisch ingetekende structuren te controleren en aan te passen voordat deze worden gebruikt voor het maken van een behandelplan. Hoofdstuk 4 onderzoekt, gebruik makend van Erasmus-iCycle, in hoeverre afname in de speekselklierfunctie bij hoofd-hals kankerpatiënten kan worden beperkt door behalve de intensiteitgemoduleerde fluentieprofielen van de behandelbundels ook de bundeloriëntaties automatisch te optimaliseren. Gebaseerd op de resultaten van deze studie hebben we een gerandomiseerde prospectieve klinische studie uitgevoerd voor hoofd-hals patiënten waarin we automatisch gegenereerde intensiteitgemoduleerde radiotherapie (IMRT) behandelplannen hebben vergeleken met handmatig gemaakte plannen. De resultaten van deze studie zijn samengevat in hoofdstuk 5. In hoofdstuk 6 vergelijken we verschillende IMRT planningstrategieën voor prostaatkankerpatiënten met metalen heupprothesen, gebruikmakend van Erasmus-iCycle. Hoofdstuk 7 beschrijft een volledig automatische procedure om volumetrisch gemoduleerde rotatietherapie (VMAT) plannen te maken voor prostaatkankerpatiënten. Hoofdstuk 8 is een algemene discussie over de onderwerpen in dit proefschrift, en beschrijft richtingen voor toekomstig onderzoek.

## Hoofdstuk 2

We hebben een klinische validatiestudie uitgevoerd om te bepalen in hoeverre met ABAS klinische doelvolumina en gezonde weefselstructuren voor hoofd-hals patiënten nauwkeurig en snel kunnen worden ingetekend. We hebben daarbij twee verschillende ABAS strategieën geëvalueerd, gebruikmakend van een enkele atlas en gebruikmakend van atlassen van meerdere patiënten. Gebruik makend van de *leave-one-out* methode werd de overeenkomst tussen een gegenereerde autocontour en een (input) atlascontour gekwantificeerd door de Dice coëfficiënt en de gemiddelde afstand tussen beide contouren te bepalen. Voor 12 klinisch behandelde patiënten hebben 5 ervaren artsen de automatisch gegenereerde contouren gecontroleerd en desgewenst handmatig aangepast. De tijd die ze daarvoor per patiënt nodig hadden is bijgehouden. Dice coëfficiënten en gemiddelde afstanden zijn berekend tussen de klinisch gebruikte contouren, de automatisch gegenereerde contouren (autocontouren), en de aangepaste autocontouren. Tenslotte is door een expert panel de nauwkeurigheid van enerzijds de autocontouren en anderzijds de handmatig aangepaste autocontouren beoordeeld in vergelijking met een gepubliceerde atlas.

Per patiënt had ABAS 7 minuten nodig om alle structuren te segmenteren. Er was geen significant verschil in de nauwkeurigheid van de segmentatie voor patiënten met en zonder positieve halsklieren. De nauwkeurigheid van autosegmentatie was het beste wanneer gebruik gemaakt werd van atlassen van meerdere patiënten. Hierbij was de gemiddelde Dice coëfficiënt en de gemiddelde afstand tussen autocontouren en de input data 0,74 en 2 mm voor de speekselklieren, 0,67 en 3 mm voor de verschillende halsklier niveaus, 0,71 en 2 mm voor de kauwspieren, 0,50 en 2 mm voor de slikspieren, en 0,78 en 2 mm voor het ruggenmerg en de hersenstam. De gemiddelde Dice coëfficiënt en de gemiddelde afstand van de autocontouren ten opzichte van de klinisch gebruikte contouren was 0,8 en 2,4 mm voor de halsklier niveaus en de speekselklieren. Wanneer de autocontouren werden vergeleken met de handmatig aangepaste autocontouren was de gemiddelde Dice coëfficiënt 0,9 en de gemiddelde afstand 1,6 mm. Het expert panel scoorde alle autocontouren als "weinig afwijkend, goed aan te passen" of beter. 88% van de aangepaste autocontouren werd gescoord als "goed", ten opzichte van 83% voor de klinisch gebruikte contouren.

Uit deze studie blijkt dat ABAS zeer goed bruikbaar is om doelvolumina en gezonde weefsels snel en automatisch in te laten tekenen wanneer gebruik gemaakt wordt van atlassen van meerdere patiënten. Hoewel de autocontouren meestal nog handmatig moesten worden aangepast, was hiervoor veel minder tijd nodig dan wanneer de contouren volledig handmatig werden ingetekend (66 in plaats van 180 minuten).

## Hoofdstuk 3

In deze studie hebben we het dosimetrische effect onderzocht van het aanpassen van autocontouren voor electief te bestralen halsklieren (CTVs) en risico-organen die met ABAS werden ingetekend bij hoofd-hals patiënten. Voor 9 patiënten waren datasets met ABAS autocontouren en door twee artsen aangepaste autocontouren beschikbaar. Gebaseerd op de niet-aangepaste autocontouren zijn klinisch acceptabele IMRT plannen gemaakt (ABAS-plannen genaamd). Deze plannen zijn vervolgens geëvalueerd voor de twee aangepaste contour sets. Dice coëfficiënten en gemiddelde afstanden zijn berekend om de overeenkomst tussen ABAS autocontouren en de door de artsen aangepaste contoursets te bepalen. Om het dosimetrische effect van het aanpassen van autocontouren voor risico-organen te bepalen werd een nieuw IMRT plan gegenereerd, gebaseerd op de door de artsen aangepaste CTVs en de niet aangepaste autocontouren voor de speekselklieren. Voor ieder plan is de gemiddelde dosis van de automatisch ingetekende speekselklieren vergeleken met de gemiddelde dosis van de door een arts aangepaste intekening van speekselklieren.

Voor beide artsen waren de aangepaste CTVs groter dan de ABAS autocontour, met een gemiddelde van 8,7% ($p \leq 0{,}04$) Wanneer we de ABAS plannen evalueerden op het te bestralen doelvolume (PTV) voor de aangepaste contour sets, nam het volume dat bestraald werd tot 95% van de voorgeschreven dosis ($V_{95\%}$) af met 7,2% ± 5,4% (1 SD) ($p < 0{,}03$). De gemiddelde afname in de minimale dosis in 99% van het volume ($D_{99\%}$) was 14,2 Gy (bereik 1–54 Gy). Zelfs voor Dice coëfficiënten >0,8 en gemiddelde afstanden tussen contouren <1 mm, kwamen reducties in $D_{99\%}$ tot 11 Gy voor. Voor de behandelplannen gebaseerd op de PTVs die door de artsen waren gedefinieerd en de niet-aangepaste, automatisch ingetekende speekselklieren was de gemiddelde dosis in de aangepaste speekselklieren slechts 0,6 Gy ± 1,0 Gy ($p = 0{,}06$) lager.

Dit laat zien dat het aanpassen van autocontouren van hals CTVs die door ABAS zijn gegenereerd noodzakelijk is om grote onderdosering in doelvolumina voor de bestraling te voorkomen. Veelgebruikte maten voor de overeenkomst tussen volumina, zoals de Dice coëfficiënt, hebben geen voorspellende waarde voor de te verwachten onderdoseringen in het PTV. Het aanpassen van de intekening van de speekselklieren bleek minder belangrijk omdat het verschil in gemiddelde dosis voor de automatisch ingetekende en de aangepaste klieren gering is.

## Hoofdstuk 4

In deze studie hebben we onderzocht in hoeverre de afname van de speekselklierfunctie tijdens bestraling van hoofd-hals patiënten kan worden beperkt, door gebruik te maken van IMRT plannen waarbij, behalve de intensiteitsprofielen, ook de bundelhoeken van de bestralingsvelden automatisch worden geoptimaliseerd. Daarnaast is onderzocht of VMAT plannen kunnen worden verbeterd door gebruik te maken van een geoptimaliseerde tafeldraaiing. We hebben Erasmus-iCycle gebruikt voor het automatisch genereren van multicriteria IMRT plannen met geoptimaliseerde bundelhoeken en intensiteitsprofielen, en van plannen met geoptimaliseerde intensiteitsprofielen maar vooraf vastgelegde bundelhoeken. Voor 20 patiënten werden 5 IMRT plannen vergeleken gebaseerd op één identieke wishlist: i), ii) plannen met zeven en negen equidistante coplanaire bundels ($iCycle_{7equi}$, $iCycle_{9equi}$), iii), iv) negen bundel plannen met geoptimaliseerde coplanaire en niet-coplanaire bundelhoeken ($iCycle_{copl}$, $iCycle_{noncopl}$) en v) een coplanair negen bundelplan met geoptimaliseerde bundelhoeken en één geoptimaliseerde tafelhoek ($iCycle_{couch}$). VMAT plannen met en zonder deze geoptimaliseerde tafelhoek zijn geëvalueerd.

Door gebruik te maken van Erasmus-iCycle zijn we in staat om de verschillende planningstrategieën op een objectieve manier te vergelijken. $iCycle_{noncopl}$ resulteerde in de laagste dosis in de speekselklieren, maar $iCycle_{couch}$ liet vergelijkbare resultaten zien voor 18 patiënten. Het voorspelde risico op bijwerkingen (NTCP) voor de submandibulaire speekselklieren was gemiddeld 5% hoger voor $iCycle_{7equi}$. $iCycle_{9equi}$ deed het beter dan $iCycle_{7equi}$ en $iCycle_{copl}$ liet een verdere verbetering zien. Het toepassen van de geoptimaliseerde tafeldraaiing verbeterde ook de NTCP waardes in VMAT plannen.

## Hoofdstuk 5

In deze prospectieve klinische studie hebben we plannen die automatisch met Erasmus-iCycle werden gegeneerd en handmatig gegeneerde plannen met het klinische planningsysteem vergeleken. Voor 20 willekeurig geselecteerde hoofd-hals patiënten met verschillende tumorlocaties (waarvan 13 een sequentiële boostbehandeling kregen) hebben we de behandelend arts laten kiezen tussen een automatisch gegenereerd Erasmus-iCycle plan en een handmatig geoptimaliseerd plan volgens de standaard klinische procedures. Erasmus-iCycle maakte gebruik van een vaste 'wishlist' bestaande uit harde voorwaarden en geprioriteerde doelen voor het uiteindelijke behandelplan. De laboranten daarentegen kozen de bundelhoeken, en optimaliseerden de voorwaarden en doelen handmatig voor iedere IMRT planning. Noch de laborant, noch de behandelend arts werden vooraf

geïnformeerd of er een concurrerend Erasmus-iCycle plan gemaakt zou worden. De twee plannen werden gelijktijdig aan de arts getoond, waarna deze één van beide plannen selecteerde voor patiëntbehandeling. Voor de patiëntgroep hebben we het verschil in de minimale dosisdekking van het PTV en in de afgegeven dosis in risico-organen bepaald.

In 32/33 planningen selecteerde de arts het Erasmus-iCycle plan voor behandeling. Deze zeer consistente voorkeur voor de automatisch gegenereerde plannen werd in hoge mate ingegeven door een lagere dosisafgifte in een groot aantal risico-organen. Zo werd met Erasmus-iCycle het voorspelde risico op bijwerkingen voor de grote en kleine speekselklieren verminderd met respectievelijk 2,4% ± 4,9% (maximum: 18,5%, $p$ = 0,001) en 6,5% ± 8,3% (maximum: 27%, $p$ = 0,005). De dosisreductie in de mondholte was 2,8 Gy ± 2,8 Gy (maximum: 8,1 Gy, $p$ = 0,005). Voor slikspieren, de slokdarm en het strottenhoofd was de afname in gemiddelde dosis 3,3 Gy ± 1,1 Gy (maximum: 9,2 Gy, $p$ < 0,001). Daarnaast was voor 15 van de 20 patiënten de minimale dosisdekking van het PTV verbeterd. Behalve dat door het automatisch genereren van plannen de plankwaliteit verbetert, is dit ook economisch aantrekkelijk doordat de werklast voor planningslaboranten afneemt.

## Hoofdstuk 6

In deze studie werden verschillende IMRT strategieën vergeleken voor prostaatkankerpatiënten met metalen heupprothesen. Alle plannen werden volledig automatisch gegenereerd met Erasmus-iCycle, dus zonder handmatige interacties, zodat een objectieve planvergelijking mogelijk was. Er werden plannen met geoptimaliseerde coplanaire en niet-coplanaire bundelhoeken gemaakt. Voor 18 prostaatkankerpatiënten (8 met tweezijdige heupprothesen, 10 met een enkelzijdige, rechterprothese) zijn er twee planningstrategieën geëvalueerd. In de eerste strategie werden bundelrichtingen van waaruit dosis in een prothese werd afgegeven voordat het doelvolume werd bereikt niet geselecteerd ($IMRT_{remove}$). In de tweede strategie mochten deze bundelrichtingen wel worden geselecteerd, en werd alleen het deel van de bundel afgedekt dat directe dosis aan het doelvolume zou afgeven nadat het door de prothese is gegaan ($IMRT_{cut}$). Verschillen in de minimale dosisdekking van het PTV en de dosisafgifte in risico-organen werden gekwantificeerd.

Voor patiënten met tweezijdige heupprothesen was er een significante verlaging te zien in de afgegeven dosis in het rectum en de blaas wanneer $IMRT_{cut}$ werd gebruikt in plaats van $IMRT_{remove}$. Voor 9-bundel coplanaire plannen reduceerde het volume van het rectum dat tot 60 Gy werd bestraald ($V_{60\ Gy}$) met 17,5% ± 15,0% (maximum 37,4%, $p$ = 0,036) en de gemiddelde rectumdosis met 9,4% ± 7,8% (maximum 19,8%, $p$ = 0,036).

Een verdere verlaging van dosis in de risico-organen was haalbaar door gebruik te maken van niet-coplanaire bundelconfiguraties. Dit resulteerde voor niet-coplanaire 9-bundel $IMRT_{cut}$ plannen tot een extra afname in $V_{60\ Gy}$ met 4,6% ± 4,9% ($p$ = 0,012). Een grote dosisafname in het rectum werd ook gezien wanneer het aantal geoptimaliseerde bundelrichtingen in de planningen werd vergroot. Voor patiënten met dubbelzijdige heupprothesen was de rectum $V_{60\ Gy}$ 37,3% ± 12,1% voor coplanaire 7-bundel plannen; voor plannen met 15 bundels reduceerde dit gemiddeld met 13,5% (maximum 30,1%, $p$ = 0,012).

Deze studie heeft aangetoond dat we met Erasmus-iCycle automatisch plannen van hoge kwaliteit kunnen maken voor prostaatkankerpatiënten met heupprothesen. Door alleen het deel van de bundel af te dekken dat eerst directe dosis afgeeft in de heupprothese ($IMRT_{cut}$ strategie) wordt een significante afname in de dosisafgifte in risico-organen bereikt. Het gebruik van niet-coplanaire bundelrichtingen en in hogere mate het vergroten van het aantal bundelrichtingen gaf een verdere verbetering van de plankwaliteit.

## Hoofdstuk 7

In dit onderzoek is een systeem voor het automatisch genereren van VMAT behandelplannen voor prostaatkankerpatiënten ontwikkeld, geëvalueerd, en klinisch geïmplementeerd. We hebben Erasmus-iCycle direct gekoppeld aan het in onze kliniek gebruikte commerciële dosisplanningsysteem om volledig automatisch VMAT plannen te kunnen maken ($VMAT_{auto}$). Voor 30 willekeurig geselecteerde patiënten werden $VMAT_{auto}$ plannen vergeleken met plannen die handmatig, zonder tijdsdruk, door één expertlaborant gemaakt zijn ($VMAT_{man}$). Voor alle plannen werd de minimale dosisdekking van het PTV en de dosisafgifte in risico-organen gekwantificeerd.

Alle gegenereerde plannen waren klinisch acceptabel en hadden een vergelijkbare minimale dosisdekking van PTV ($V_{95\%}$ > 99%). De dosis in de risico-organen was vergelijkbaar voor $VMAT_{auto}$ en $VMAT_{man}$ plannen, maar de eerstgenoemden werden vervaardigd zonder enige werklast voor de laboranten. Voor de eerste 10 klinisch behandelde patiënten is aan laboranten gevraagd om te pogen om de $VMAT_{auto}$ plannen te verbeteren door handmatig, via een trial-and-error methode, te sleutelen aan de planningsdoelen in Monaco, zoals nu nog gebruikelijk is in de klinische praktijk. In deze handmatig aangepaste plannen waren de rectum $V_{75\ Gy}$ en de rectum $V_{60\ Gy}$ gemiddeld iets lager (maximaal verschil 1%), maar deze beperkte verbetering in de dosisafgifte in het rectum resulteerde meestal in een toename van de dosis in andere, lager geprioriteerde risico-organen en/of een verslechterde dosisafval in de omliggende weefsels.

De resultaten van deze studie laten zien dat het volledig automatisch genereren van VMAT plannen van hoge kwaliteit haalbaar is voor prostaatkankerpatiënten. Deze werkwijze wordt nu standaard toegepast in onze kliniek.

## Hoofdstuk 8

Dit hoofdstuk geeft een algemene discussie over de onderzoeksonderwerpen die in dit proefschrift gepresenteerd zijn. De huidige toepassingen van het automatisch genereren van contouren en plannen in onze kliniek worden beschreven. Toekomstig onderzoek zal zich onder meer richten op het sneller en breder toepasbaar maken van beide procedures en op de evaluatie van deze methodes voor nieuwe behandelingen.

# acknowledgements

Scientific research is a team sport! Nothing can be achieved without the direct support of colleagues around you. I am grateful that I was part of a strong and inspiring team. Now I can look back on a very exciting period, and a great experience. I would like to thank all of my colleagues of the department, family, and friends for supporting me in this period. Clearly, the main focus was on my research work, which I luckily greatly enjoyed.

Professor dr. Heijmen, dear Ben, a huge thank you for giving me the opportunity of starting this PhD research in the first place, and believing in me. Your endless support, enthusiastic open attitude when discussing my research projects always energized me. Your ability of reading draft versions of manuscripts and being able to find the scientific difficulties in a split second kept on surprising me. Despite your busy agenda, you always made time for me.

Dr.ir. Dirkx, dear Maarten, thanks for your patient guidance on this research journey. Thank you for all the detailed discussions we had regarding ideas, analyses of data, or changes in the focus of the research. Thank you for helping me to transfer exciting scientific results into readable publications. I really appreciated having you as my direct supervisor.

Sebastiaan Breedveld, our collaboration was very special. From the moment I started using Erasmus-iCycle for my research projects we had a "direct 24-7 hotline". You were always willing to listen to my wild ideas about improvements, or requests for additional functionality. We officially abolished the nine-to-five mentality, which even resulted into a nice Christmas-colored-CT. Thanks to you, I considered my research a great hobby, rather than work. I also enjoyed the discussions we had about coffee and tea. Although coffee is still by far number 1, you showed me the world of tea. It was a great pleasure working together.

I would also like to thank Steven Petit for reviewing the initial discussion of this thesis and for giving some helpful directions. Wouter Wunderink, thank you for transferring a nice postcard into the cover of this book.

A special thanks to some of my RTT colleagues: Gert Korevaar and Sandra Quint. Despite my renegade to G2, we always stayed in contact; your support is very much appreciated. Dennie Fransen, thank you for your open attitude toward changes in the field of treatment planning and thank you for your support in helping performing a clinical prospective study. Also your support in finalizing some figures of the thesis is highly appreciated. I would like to thank John van Sörnsen de Koste for injecting me with the research virus. You taught me to be critical, and not to take anything for granted in research. Henrie van der Est, we were dosimetrist colleagues for many years, and we participated in the head-and-neck research group together; your critical opinion, but above all your creative way of working and thinking always motivated me a lot! Also thanks for helping me designing the cover of the thesis.

The endless support of the 'Matterhorn-guys', Wilco Schillemans, Davy Wentzler, and Andras Zolnay is highly appreciated; your solutions made life a lot easier. Hafid Akhiat, Bart Kanis, and Bert van der Leije, a big thank you for all the help in solving my annoying IT-related problems.

Iskandar bin Kassim† and Abdul Wahab Sharfo (Buddi). I found out how important roommates are in dealing with every day hurdles. Thanks for being around.

Jolanda Jacobs and Jacqueline van der Valk, thanks for helping me with all kind of things, and thanks for being there whenever I needed it. Being C.E.B. is sometimes tough, but your support made life much easier ☺.

I want to thank my parents for supporting me, believing in me, and above all encouraging me to get the best out of everything.
I tried my very best to finish in time: '*sorry dad, we both did what we could*'.

A very special thank you to Daniëlla. You faced all the rough weather of my PhD journey. I am blessed having you by my side. Voor Chantal: '*Papa heeft éindelijk zijn boekje klaar hoor, dankjewel voor je steun. Je bent een kanjer!!*'

# PhD portfolio

Name PhD student: Peter W.J. Voet
Erasmus MC Department: Radiotherapy
Research School: Molecular Medicine
PhD period: 2010–2014
Promotor: Prof. dr. Ben J.M. Heijmen
Copromotor: Dr. ir. Maarten L.P. Dirkx

| | |
|---|---|
| **1. PhD training** | |
| **General courses** | |
| • Statistics (MolMed) | 2010 |
| • Biomedical English Writing and Communication (Erasmus MC) | 2011 |
| **Specific courses** | |
| • ESTRO teaching course: Advanced Imaging for physicists | 2011 |
| • ESTRO teaching course: Basic clinical radiobiology | 2013 |
| • ESTRO teaching course: Clinical particle therapy | 2014 |
| **International conferences** | |
| • ESTRO 29, Spain (oral presentation) | 2010 |
| • ESTRO 30, UK (oral presentation) | 2011 |
| • AAPM 2012, USA (poster presentation) | 2012 |
| • ESTRO 32, Switzerland (poster discussion presentation) | 2013 |
| • ESTRO 33, Austria (oral presentation) | 2014 |
| **National conferences** | |
| • Wetenschappelijke vergadering NVRO, Utrecht (oral presentation) | 2011 |
| • 27e NVKF conferentie, Zeist (oral presentation) | 2012 |
| • Wetenschappelijke Kringdag klinische fysica radiotherapie, Utrecht (oral presentation) | 2013 |

Continued on the next page

| | |
|---|---|
| **Teaching** | |
| Faculty member/Teacher at ESTRO course "Advanced Treatment Planning" in: | |
| • Italy | 2011 |
| • Czech Republic | 2012 |
| • The Netherlands | 2013 |
| **Presentations as invited speaker:** | |
| 23 times on "Monaco treatment planning software and ABAS autosegmentation software" in Australia, China, Croatia, Germany, Italy, Malaysia, Portugal, Singapore, Spain, Taiwan, Thailand, USA, UK | 2010–2013 |
| **Lecturing** | |
| Workshops on the Monaco treatment planning system, 10 times in China, Germany, Malaysia, The Netherlands, Portugal, Singapore, Taiwan, Thailand, UK, USA | 2010–2013 |
| **In-house Presentations:** | |
| Physics Journal Club | |
| • "Latest on VMAT, IMRT, 3D-CRT" | 2010 |
| • "Pareto front analysis in clinical practice: what is it, and what does it bring?" | 2011 |
| Refereeravonden | 2010, 2011 (2x), 2012, 2013 |
| **Other** | |
| Guest teacher in-service opleiding radiotherapeutisch laborant | |
| • "Treatment planning: from simple to complex" | 2012, 2013 |
| Guest teacher Hogeschool InHolland, Haarlem, The Netherlands | |
| • "IMRT: future challenges …" | 2012, 2013 |
| Guest teacher NVRO module treatment planning and delivery | |
| • "IMRT conceptueel" | 2013 |

# curriculum vitae

Peter Voet was born on July 11, 1975 in Terneuzen, The Netherlands. After finishing the HAVO (school of higher general secondary education), at Zeldenrust College, Terneuzen in 1992, he started in 1993 the in-service training *Radiotherapeutisch Laborant* (RTT/Dosimetrist) at the Daniel den Hoed kliniek/Opleidings instituut Erasmus Medical Center in Rotterdam. He graduated in 1996.

In 1998 he started working as *Research Laborant* at the Scanditronix racetrack MM-50 research unit. Peter joined the Rotterdam Oncological Thoracic Study group (ROTS) from 1998–2002, and he was part of the head-and-neck research group from 2002–2010. In the period 2004–2009, he worked as coordinator ICT, research and development at the radiotherapy department. In this period he worked mainly on the practical implementation of IMRT, and the clinical implementation, including staff training, of Elekta software products (ABAS, Focal, XiO, Monaco).

In November 2009, he started his PhD research at the Erasmus Medical Center under the supervision of prof. dr. Ben Heijmen and dr. ir. Maarten Dirkx, which eventually resulted in this thesis.

Peter is married to Daniëlla; they are the proud parents of Chantal.

This thesis focuses on automated contouring of planning CT-scans and automated treatment plan generation in radiotherapy to improve treatment quality and / or efficiency compared to current manual procedures.

For automated contouring we have evaluated a commercially available atlas-based autosegmentation tool. With this tool, the CT-scan of a new patient can be segmented based on segmented scans of previously treated patients, called atlases, thereby significantly reducing the manual contouring time.

For automated plan generation we have investigated Erasmus-iCycle, an in-house developed algorithm for multicriterial beam angle and profile optimization. For individual patients, Erasmus-iCycle can fully automatically generate Pareto-optimal plans, based on a wish-list with constraints and prioritized objectives specified for the corresponding disease / tumor site. Compared to manually generated plans, Erasmus-iCycle plans are superior or equal in quality. Moreover, cost savings are possible, as time consuming manual planning by dosimetrists can be avoided.

Feasibility of high quality automated contouring and planning has been demonstrated, which is also highly relevant for the rapidly evolving field of adaptive daily re-planning.