

# The impact of plan complexity on the accuracy of VMAT for the treatment of head and neck cancer

A thesis submitted in partial fulfilment of the requirements for the

Degree of Master of Science in Medical Physics

by

Thomas William Scott Satherley

Department of Physics and Astronomy

University of Canterbury

2015

# Abstract

**Purpose:** At the Wellington Blood and Cancer Centre (WBCC), Volumetric Modulated Arc Therapy (VMAT) is used to treat a variety of head and neck (H&N) cancers. Presently, the complexity of plans is limited to ensure the accuracy of patient treatment within the range of the departmental experience. The complexity limitation is applied through use of a monitor unit (MU) constraint during plan optimisation. Plans of higher complexity can be obtained by loosening the MU constraint, and setting more stringent optimisation objectives on organs at risk (OAR) and target volumes (PTV). This could potentially yield higher quality treatment plans but may also degrade the accuracy of the TPS calculation or the plan delivery at the treatment machine. The aim of this study is to investigate the level of plan complexity that results in accurate treatment plan calculation and delivery, and quantify the corresponding gain in plan quality.

**Methods:** Five previously treated H&N patients were selected for the study. Each patient's clinical plan was used as the lowest complexity level and labelled C1. Subsequently, an approximate pareto-optimal plan (C3) was created that focused equally on sparing spinal cord, brain stem and parotid gland while maintaining, or improving on, the previously obtained target coverage. Next, a C2 plan was created such that the plan quality was in between C1 and C3. Plan quality of each complexity level was assessed in terms of OAR sparing and PTV coverage. The average leaf pair opening (LPO), critical leaf pair opening (%LPO<1cm) and mean leaf travel were used as plan complexity metrics.

The calculation and delivery accuracy of each complexity level using Varian TrueBeam LINAC/Eclipse TPS was verified using time resolved point dose measurements (TRPD), EBT film measurements (Ashland Inc.) and ArcCheck measurements (Sun Nuclear Corp.). A comprehensive uncertainty analysis was carried out including a quantification of the measurement and delivery reproducibility.

**Results:** Increasing plan complexity from C1 to C3 reduced the Spinal Cord  $D_{1cc}$ , Brain Stem  $D_1$  and Parotid Gland  $D_{mean}$  up to 14.7 Gy, 7.1 Gy and 7.8 Gy, respectively. In addition, C3 plans improved the target coverage compared to C1 plans, with the PTV66 and PTV54  $D_{98}$  increasing up to 1.0 Gy and 0.6 Gy, respectively. The verification measurements showed that the plan calculation and delivery for all complexity levels was well within clinical acceptance levels (Table 1). TRPD showed that VMAT dose delivery itself was repeatable within 0.1% (1 S.D.) over 10 consecutive deliveries for both C1 and C3 complexity levels.

Table 1: TRPD, EBT film and ArcCheck results;  $\gamma$ -analysis was performed using a {2%; 2 mm} criterion

Complexity level	Mean point dose deviation $\pm 1$ S.D.	Mean EBT film $\gamma$ pass-rate $\pm 1$ S.D.	Mean ArcCheck $\gamma$ pass-rate $\pm 1$ S.D.
C1	-0.1% $\pm$ 1.0%	95% $\pm$ 3%	96% $\pm$ 4%
C2	-0.1% $\pm$ 1.2%	95% $\pm$ 2%	95% $\pm$ 3%
C3	0.0% $\pm$ 1.2%	96% $\pm$ 2%	94% $\pm$ 5%

**Discussion & Conclusions:** This study has shown that increasing the plan complexity can provide significant dosimetric advantages for the treatment of H&N cancer. Verification measurement results indicated that this did not noticeably degrade the calculation and delivery accuracy of VMAT using a Varian TrueBeam LINAC and our Eclipse TPS beam model. H&N VMAT at the WBCC can now be developed further with greater confidence in the dosimetric accuracy of higher complexity plans.

# Acknowledgements

I would like to thank my clinical supervisors Dr. Rob Louwe and Ms. Rebecca Day for their extraordinary support during this project. Their considerable experience and expertise in this field has helped me greatly. Thank you both.

I would also like to thank the various WBCC staff who were accessible and helped to answer any questions I had during this learning experience. Finally, thank you to Ms. Lynne Greig and the Wellington Blood and Cancer Centre for providing the time and equipment required to conduct this project.

# Table of Contents

Glossary .....	vii
1. Introduction.....	1
1.1. Head and neck cancer.....	1
1.2. 3D conformal radiotherapy .....	3
1.3. Intensity modulated radiotherapy.....	3
1.3.1. Inverse planning.....	5
1.3.2. IMRT Optimisation methods .....	6
1.3.3. IMRT for H&N treatments .....	6
1.4. Volumetric modulated arc therapy.....	7
1.4.1. VMAT Optimisation.....	8
1.4.2. VMAT vs. IMRT .....	9
1.4.3. VMAT uncertainties and quality control .....	9
1.3.4. Plan complexity .....	10
1.5. Purpose of this study .....	11
2. Materials & Methods .....	12
2.1. Treatment Planning .....	12
2.1.1. Selected cases.....	12
2.1.2. Definition of dose-volume histogram metrics .....	12
2.1.3. Definition of planning volumes .....	13
2.1.4. Plan Quality .....	15
2.1.5. Complexity levels .....	16
2.1.6. Progressive Resolution Optimiser.....	17
2.1.7. Generating C3 plans.....	20
2.1.8. Generating C2 plans.....	23
2.1.9. Final dose calculation .....	23
2.1.10. Plan delivery .....	23
2.1. Quantification of plan complexity .....	24
2.2.1. Average Leaf Pair Opening (ALPO) .....	24
2.2.2. Critical Leaf Pair Opening (CLPO).....	25
2.2.3. Mean Leaf Travel (LT) .....	26
2.2.4. Monitor Units (MU).....	26
2.3. Verification measurement methods.....	27
2.3.1. Time resolved point dose analysis .....	27

2.3.2.	EBT3 Gafchromic Film .....	33
2.3.3.	ArcCheck .....	38
3.	Results.....	41
3.1.	Plan quality.....	41
3.1.1.	Target coverage.....	41
3.1.2.	OAR sparing .....	42
3.1.3.	Case Study .....	43
3.2.	Plan complexity.....	47
3.2.1.	Complexity metrics .....	47
3.2.2.	ALPO, CLPO, LT relationships with MU .....	50
3.2.3.	Case study .....	52
3.2.4.	TPS version 11 upgrade .....	52
3.3.	Plan dosimetry verification .....	54
3.3.1.	VMAT delivery repeatability.....	54
3.3.2.	Impact of dose grid resolution .....	56
3.3.3.	Point dose verification .....	56
3.3.4.	Film dose verification .....	59
3.3.5.	ArcCheck dose verification.....	59
3.3.6.	Case study .....	60
4.	Discussion .....	63
4.1.	Plan Quality.....	64
4.1.1.	Spinal cord dose.....	65
4.1.2.	Brain stem dose.....	66
4.1.3.	Parotid gland dose.....	66
4.1.4.	Advantages and disadvantages of the complexity level system .....	67
4.2.	Plan complexity.....	67
4.2.1.	MU & Gantry speed.....	68
4.2.2.	Field size .....	69
4.2.3.	Leaf travel .....	70
4.2.4.	Complexity variation with case .....	71
4.2.5.	Differences in complexity between CCW & CW beams.....	71
4.2.6.	Relationships between ALPO/CLPO/LT and MU .....	72
4.2.7.	TPS version 11 upgrade .....	72
4.2.8.	Limitations of the complexity level planning system .....	73
4.2.9.	Other considerations .....	74
4.3.	Plan dosimetry verification .....	75
4.3.1.	LINAC dose delivery repeatability .....	76

4.3.2.	Impact of dose grid resolution .....	77
4.3.3.	Point dose verification .....	78
4.3.4.	Film dose verification .....	80
4.3.5.	ArcCheck dose verification.....	81
4.3.6.	VMAT insensitivity to errors.....	82
5.	Conclusion .....	84
	Bibliography .....	85

# Glossary

3DCRT	<i>3-Dimensional conformal radiotherapy</i>
AAA	<i>Analytical anisotropic algorithm</i>
ALPO	<i>Average leaf pair opening</i>
BS	<i>Brain stem</i>
C1	<i>Complexity level 1</i>
C2	<i>Complexity level 2</i>
C3	<i>Complexity level 3</i>
CLPO	<i>Critical leaf pair opening</i>
CT	<i>Computed tomography</i>
CTV	<i>Clinical target volume</i>
DAO	<i>Direct aperture optimisation</i>
DICOM RT	<i>Digital Imaging and Communications in Medicine file format for radiation therapy</i>
DTA	<i>Distance to agreement</i>
DVH	<i>Dose-volume histogram</i>
DVO	<i>Dose-volume objective</i>
GTV	<i>Gross tumour volume</i>
H&N	<i>Head and neck</i>
ICRU	<i>International commission on radiation units and measurements</i>
IMRT	<i>Intensity modulated radiotherapy</i>
LINAC	<i>Linear accelerator</i>
LPO	<i>Leaf pair opening</i>
LT	<i>Mean leaf travel</i>
MLC	<i>Multi-leaf collimator</i>
MR	<i>Magnetic resonance (imaging)</i>
MU	<i>Monitor Units</i>
NTCP	<i>Normal tissue complication probability</i>
NTO	<i>Normal tissue objective</i>
OAR	<i>Organ at risk</i>
OD	<i>Optical density</i>
PET	<i>Positron emission tomography</i>
PPTV	<i>Planning target volume (for VMAT optimisation)</i>
PRO	<i>Progressive resolution optimiser</i>
PRV	<i>Planning risk volume</i>
PTV	<i>Planning target volume</i>
SC	<i>Spinal cord</i>
TCP	<i>Tumour control probability</i>
TPS	<i>Treatment planning system</i>
VMAT	<i>Volumetric modulated arc therapy</i>
WBCC	<i>Wellington Blood and Cancer Centre</i>

# 1. Introduction

## 1.1. Head and neck cancer

Head and neck (H&N) cancer is carcinoma (commonly epithelial squamous-cell carcinoma) that affects the oral cavity, nasal cavity, pharynx, larynx, salivary glands and paranasal sinuses [1]. Around two thirds of patients present with later stage disease that has commonly spread to the regional lymph nodes [2]. The important risk factors for H&N cancer have been attributed to cigarette smoking, alcohol consumption [3, 4, 5] and exposure to human papilloma virus [6]. The incidence of H&N cancer is greatest for those aged over 50 but the incidence for younger adults has recently increased [7, 8]. As of 2013, approximately 520 new cases of H&N cancer are registered with the ministry of health in New Zealand each year [9]. Standard treatment of H&N cancer includes surgery, chemotherapy and radiotherapy. Either surgery or radiotherapy is used for treatment of stage I or II disease with the intent of curing the disease. Radiotherapy can also be used as a postoperative treatment [10]. For locally advanced disease, chemotherapy is often administered concurrently with radiotherapy [11].

Radiotherapy as an effective treatment of H&N cancer requires that the targeted volumes are correctly irradiated to a prescribed dose level. The effectiveness of the treatment depends on its ability to eliminate the tumour cells where the probability of doing this with a certain dose is dictated by the tumour control probability (TCP) [12]. However, the safe application of radiotherapy also requires that the organs at risk (OAR) are spared from radiation dose levels that may cause damage (i.e. radiation

induced toxicities). The probability that a certain dose will damage an OAR to some end point is defined by the normal tissue complication probability (NTCP) [13]. For example, delivering a mean dose above ~26 Gy to the parotid glands greatly increases the NTCP for xerostomia where salivary production by the parotid glands is severely reduced [14]. Additionally, delivering a 50 Gy or higher maximum dose to the spinal cord can result in myelopathy where at its worst grade results in paralysis, complete loss of sensory function and complete incontinence [15]. Radiotherapy is a trade-off between increasing TCP by irradiating the targeted volumes with a high dose while ensuring low NTCP to avoid radiation toxicities that impact on the patient's quality of life. To do this, radiotherapy treatments are fractionated where the total dose is delivered in multiple sessions to spare slow responding tissues (like many OARs) more than fast responding tissues (like many tumours) [16]. Additionally, conformal treatment delivery techniques that selectively irradiate the target volumes while sparing OARs have been pursued.

Radiotherapy of the H&N is difficult because the targeted volumes are often in close proximity to several critical OARs. Specifically, H&N tumours are often concave in shape and can wrap around in close proximity to the spinal cord and/or brain stem. Additionally, for treatments where the regional lymph nodes are involved, the parotid glands can be in close proximity to the targeted volumes. Furthermore, the sheer number of additional OARs located in the H&N such as the oral cavity and mandible, block many of the potential paths that radiation beams could use to irradiate target volumes.

## **1.2. 3D conformal radiotherapy**

Over the past few decades, multiple treatment techniques have been developed in attempt improve the quality of radiotherapy in general. 3-dimensional Conformal Radiotherapy (3DCRT) uses a linear accelerator (LINAC) to irradiate targeted volumes with multiple photon beams directed from various gantry angles around the patient. Planning of the beams is performed using 3D patient anatomy data derived from a multiple slice computed tomography (CT) scan. The LINAC is fitted with a computer controlled multi-leaf collimator (MLC) consisting of two banks of opposing tungsten leaves where each leaf is independently positioned to block certain areas of the photon beam and create a customised aperture. The planner, with the guidance of 3D anatomical data, uses the MLC to shape the photon beams to conform to the contour of the target volumes and avoid OARs. Multiple beams directed from different gantry angles are usually required to produce an acceptable dose distribution. 3DCRT H&N techniques improve target volume coverage and OAR sparing compared to older 2D planning methods [17]. However, H&N 3DCRT techniques are not often able to adequately spare the parotid glands, resulting in severe xerostomia. Furthermore, the dose distributions generated by 3DCRT can only be convex in shape, often not suiting the geometry of H&N planning volumes. Practically, the process of generating high quality 3DCRT H&N plans is a time-consuming because the process requires multiple manually created plan iterations to find the best plan.

## **1.3. Intensity modulated radiotherapy**

Intensity modulated radiotherapy (IMRT) represents a large technological improvement for LINAC based radiotherapy treatment. Like 3DCRT, IMRT also uses multiple photon beams directed from various gantry angles to produce a conformal

dose distribution around the target volumes. However, the key difference between 3DCRT and IMRT is the method of dose distribution generation. 3DCRT uses beams with a uniform or wedged fluence distribution within the MLC defined aperture, while IMRT uses beams where the fluence distribution is heterogeneous within the aperture [18]. The level of conformality that can be achieved with IMRT is much higher than can be achieved with 3DCRT and this allows greater sparing of OARs without compromising the dose delivery to the target volumes. Intensity modulation of beams can be achieved in two ways:

- Each beam is comprised of multiple subfields called segments (step and shoot approach),
- Dynamic MLC motion is used during the beam delivery (dynamic MLC or ‘sliding window’ approach).

In the step and shoot approach, each segment has a specified MLC defined aperture and an aperture weighting. The MLC defined aperture determines which regions are irradiated, while the aperture weight defines the quantity of radiation to be delivered through the aperture. The integral dose delivered by a beam comprised of multiple segments is calculated as the sum of the dose contributions from each segment. The desired intensity modulation for a beam is therefore created by choosing the appropriate combination of segments. In the ‘sliding window’ approach, the two banks of MLC leaves travel across the field defining a narrow window during continuous irradiation. The exact shape and width of the window varies as it travels across the field and the LINAC dose rate also varies. In this fashion, the intensity modulated distribution is built up. Regardless of the delivery approach used, the

overall IMRT plan is comprised of multiple intensity modulated beams directed from various gantry angles. Typically there are five or seven beams.

### **1.3.1. Inverse planning**

In theory, a human planner could select the appropriate combination of segments for multiple beams to produce the desired dose distribution. This planning approach is referred to as forward planning. However in practice, the selection of the appropriate segments may not be obvious and it is likely difficult for a planner to generate an IMRT plan within an acceptable time frame using a forward planning approach. Instead, IMRT plans are generated with an inverse planning approach where the segments are determined by computed iterative optimisation [18, 19]. In inverse planning, the planner defines dose-volume objectives (DVO) for target and OAR volumes representing the dosimetric goals to be achieved by the plan. An example of a DVO could be that 98% of the target volume should receive at least a minimum of 62.7 Gy (i.e.  $V_{62.7\text{Gy}} \geq 98\%$ ). The degree in which a plan accomplishes all of the defined DVOs is quantified by an objective function. For a single DVO, the objective function is commonly formulated as the square of the difference between the defined DVO and the dose/volume achieved by the plan. The total objective function is a weighted sum of each of the individual DVO contributions where the weighting factors are used to set the relative priority or importance of each DVO. The objective function is at a minimum when there is maximum agreement between the defined DVOs and the dose distribution that the plan actually produces. An optimisation algorithm generates an IMRT plan by iteratively finding the combination of segments that minimise the value of the objective function.

### **1.3.2. IMRT Optimisation methods**

There are two primary IMRT optimisation methods: beamlet based optimisation [20] and direct aperture optimisation (DAO) [21]. Beamlet based optimisation divides each beam into a matrix of smaller beamlets and the beamlet weights are optimised to minimise the objective function and produce the specified dose distribution. A constraint is imposed during beamlet optimisation to ensure that the optimiser does not use negative beamlet weights which are meaningless in reality. The final deliverable plan is generated after the optimised beams are translated from the theoretical combination of beamlet weights to deliverable aperture shapes using a leaf sequencing algorithm. MLC constraints such as leaf positioning restrictions are taken into account during the leaf sequencing phase.

Direct aperture optimisation algorithms optimise aperture shapes and weights directly, removing the need for a leaf sequencing step. DAO iteratively creates IMRT beams by randomly selecting a parameter (MLC leaf position or an aperture weight) and changing that parameter by a random amount. A change is accepted where the change reduces the objective function relative to the previous iteration. To avoid becoming trapped in the objective function's local minima, changes that increase the objective function are also accepted with a probability that decreases with the number of successful iterations. Changes that violate MLC constraints are not accepted and another parameter is selected at random.

The process is repeated through many plan iterations until a plan of acceptable quality is generated and the algorithm is halted.

### **1.3.3. IMRT for H&N treatments**

Multiple studies have shown that IMRT produces superior dose distributions for H&N treatments compared to 3DCRT [22, 23]. These studies showed that IMRT improved

target coverage characteristics compared to 3DCRT by significantly increasing the target volume  $V_{95\%}$  and  $V_{90\%}$ . The same studies also showed that IMRT can reduce the maximum dose delivered to the spinal cord. Additionally, IMRT has been reported to reduce the mean dose delivered parotid glands and significantly reduce xerostomia related complications [24, 25]. However, H&N IMRT techniques can have long treatment delivery times on the LINAC depending on how many beams and segments are used by the plan [26]. Additionally, it has been reported that H&N IMRT plans can require up to 2.6 times more monitor units (MU) than 3DCRT for dynamic IMRT delivery [25]. A potential downside to increased monitor unit utilisation is that the patient can be exposed to higher out of field dose. Hall [27] has reported that increased to out of field patient exposure due to leakage and transmission radiation is a factor that can increase the probability of inducing secondary cancers after radiotherapy.

## **1.4. Volumetric modulated arc therapy**

Volumetric modulated arc therapy (VMAT) [28] is an evolution of static gantry IMRT. Multiple radiotherapy vendors have an implementation of VMAT. For the purposes of this study, future use of the acronym VMAT refers to the Varian Medical Systems (Palo Alto, USA) implementation of VMAT marketed as RapidArc®. Instead of delivering multiple intensity modulated beams from static gantry angles, VMAT delivers multiple segments while the gantry is rotated around the patient in a continuous arc. As such, in the concept of a static gantry IMRT beam comprised of segments is replaced in VMAT by an arc comprised of segments defined at various angles around the arc called control points. Like IMRT, each VMAT segment has a defined MLC aperture and weight. VMAT is delivered dynamically where the MLC

leaves move in synchronisation with the gantry rotation to deliver the planned apertures at each control point. The correct aperture weight is delivered by modulating the LINAC's dose rate between 0 MU/min and 600 MU/min. Additionally, high aperture weights may require a LINAC dose rate that exceeds 600 MU/min and the gantry rotation speed is slowed to compensate, ensuring that the correct aperture weight is delivered.

### **1.4.1. VMAT Optimisation**

VMAT plans are generated using an inverse planning optimisation algorithm first developed by Otto [28]. The algorithm uses DAO for optimisation of the segments and a progressive sampling method to control the addition of control points to the optimisation. Because a VMAT beam is delivered in a continuous arc, efficiency constraints for maximum MLC motion and maximum beam weight per degree of gantry rotation must be taken into account during optimisation to ensure smooth delivery. The MLC motion constraint, determined by the maximum MLC leaf speed and the speed of gantry rotation ensures that MLC leaf position changes between control points are possible without slowing the gantry rotation speed. The beam weight constraint, determined by the maximum LINAC dose rate and the gantry rotation speed ensures that beam weights can be delivered without often needing to slow the gantry. Otto realised that optimisation of a particular segment in an arc is limited by the two segments on control points either side of the optimised segment. This is because the random MLC leaf position and beam weight changes made by DAO often exceed the efficiency constraints imposed by the adjacent segments. The progressive sampling method used in the VMAT optimisation algorithm overcomes this problem by initially optimising only a few segments coarsely spaced around the beam. By doing this, the initial optimisation of segments is not restricted by the

efficiency constraints and there is a greater chance that the optimiser can find an optimal arrangement of segments. However, the integral dose delivered by the beam can only be accurately calculated when there are sufficient segments positioned around the arc. Therefore, the progressive sampling method begins by optimising 10 segments and iterates through four further resolution levels where segments are added for optimisation. Each successive resolution level doubles the number of segments plus one. In the final resolution level, 177 segments are used to approximate a 360 degree arc.

### **1.4.2. VMAT vs. IMRT**

Multiple studies have assessed the differences between VMAT and IMRT for the treatment of H&N cancer [29, 30, 31]. It has been reported that VMAT is at least equivalent to IMRT in terms of target coverage and OAR sparing. On average VMAT treatment delivery times have been reported to be around half that of IMRT.

Additionally, VMAT H&N plans require around 50% less MU than equivalent IMRT plans.

### **1.4.3. VMAT uncertainties and quality control**

VMAT is much more complex than 3DCRT in its delivery on the LINAC and planning on the treatment planning system (TPS). VMAT delivery involves a complex interplay between the gantry rotation, MLC configuration and dose rate. The TPS dose calculation for VMAT plans is complex because multiple aspects must be correctly modelled such as the MLC characteristics and the dose calculation model for small aperture sizes (and more). Additionally VMAT adds many 'black boxes' to the planning process such as the VMAT optimiser. Therefore there is increased uncertainty relating to the agreement between the TPS calculated dose and dose

delivered by the LINAC (dosimetric accuracy). As a result, plan specific quality control (QC) tests are usually performed to assess whether the predicted dose calculated by the TPS is accurately reproduced by the LINAC delivery. These dose verification tests of individual plans are usually end-to-end style tests where the calculated dose is compared to measurement of the delivered dose using ionisation chambers, gafchromic film and diode arrays. The end-to-end tests are global tests of the VMAT process from plan optimisation and dose calculation, transfer of the plan to the LINAC and delivery of the plan on the LINAC. Additional mechanical QC tests are also performed separately to monitor the performance of the LINAC's VMAT delivery such as dynamic MLC and beam output. The agreement between the TPS calculated dose and the LINAC delivered dose is one aspect of the total dosimetric uncertainty of a VMAT treatment. Other dosimetric uncertainties result from patient positioning, inter-fraction patient setup variation, intra-fraction motion and tumour growth and shrinkage [32]. Further uncertainty can result from the variability and inaccuracies present in contouring of the anatomy and planning volumes by the radiation oncologist [33].

#### **1.3.4. Plan complexity**

Complex IMRT beams have been associated with high MU utilisation, large number of segments, small aperture sizes, complicated aperture shapes, high aperture area variability and high dose modulation [34, 35, 36, 37]. The negative aspects of high complexity IMRT beams are that they take longer to deliver, and the patient is exposed to higher amounts of radiation leaking through and scattering off the MLC leaves. Additionally, higher complexity beams can degrade the agreement between the calculated and measured dose. However, a certain amount of plan complexity is required to create a dose distribution of acceptable quality [38]. How much

complexity is needed often depends on the difficulty of the planning problem which is determined by the anatomy and planning structures. For example IMRT plans for H&N cases are normally more complex than IMRT plans for prostate cases. Complexity of VMAT plans have also been associated to the beam characteristics previously mentioned for IMRT. Additionally, complex VMAT beams have been associated with increased mean leaf travel [39]. Unlike IMRT, the negative aspects of high complexity VMAT plans are not so related to delivery inefficiencies but rather the degradation of VMAT dosimetric accuracy [39, 40].

## **1.5. Purpose of this study**

VMAT is currently used to treat the majority of H&N patients at the Wellington Blood and Cancer Centre (WBCC). A total of 113 H&N patients have been treated with VMAT since 2012. The initial implementation of WBCC's H&N VMAT planning protocol has been conservatively designed and purposefully limits plan complexity to reduce the likelihood that VMAT dosimetric accuracy is degraded. The level of complexity that degrades VMAT dosimetric accuracy is currently not known for the WBCC VMAT system. Therefore, one of the objectives of this study is to determine whether there is a level of beam complexity that degrades VMAT dosimetric accuracy below clinically acceptable tolerances. To do this, beam complexity needs to be quantified in some way. As such, this study will also investigate the use of metrics derived from beam parameters to quantify complexity. Finally, this study will determine whether higher complexity VMAT beams can improve the quality of dose distributions used to treat H&N cancer.

# 2. Materials & Methods

## 2.1. Treatment Planning

### 2.1.1. Selected cases

Five patients previously treated with H&N VMAT at the WBCC were selected at random for this study. Disease site, TNM disease staging [41] and planning target volume (PTV) volumes are listed for each case (table 2.1). The selected patients were a representative sample of the H&N cases commonly treated at the WBCC.

Table 2.1: Summary of the patients included in this study. Disease site, TNM disease staging and PTV volumes are listed for each case.

Patient	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5
Disease Site	Oropharynx	Oropharynx + Tongue	Hypopharynx	Larynx	Nasopharynx
Stage (TNM)	T4a,N1,M0	T2,N0,M0	T2,N0,M0	T2,N0,M0	T1,N0,M0
PTV66 volume (cm <sup>3</sup> )	190.3	237.5	84.9	158.9	287.6
PTV54 volume (cm <sup>3</sup> )	724.6	907.9	421	637.4	750.5

### 2.1.2. Definition of dose-volume histogram metrics

Dose-volume histograms (DVH) were used to quantify the dose delivered to each planning volume. The following metrics were extracted from the DVH to quantify certain characteristics of the dose distribution such as target coverage and OAR dose.  $D_x$  type metrics indicate the minimum dose delivered to a percentage of a volume (x). Additionally,  $D_{x\text{cc}}$  type metrics indicate the minimum dose delivered to a volume (x) measured in cubic centimetres (cc). Type  $V_y$  type metrics indicate the % of a volume that receives a minimum dose of at least (y).  $D_{\text{mean}}$  type metrics represent the mean dose delivered to the entire volume of interest.

### **2.1.3. Definition of planning volumes**

#### **Target volumes**

Delineation of target volumes and OARs was performed by a radiation oncologist with CT imaging and additional guidance from positron emission tomography (PET) and magnetic resonance (MR) imaging. Delineation of target volumes was in accordance with consensus guidelines [42]. Definition of volumes was in accordance with International commission on radiation units and measurements (ICRU) guidelines [43].

The demonstrable extent of the tumour and involved lymph nodes were defined as a gross tumour volume (GTV). Potential areas of local extension were defined as a clinical target volume (CTV66). Both GTV and CTV66 were prescribed 66 Gy in 30 fractions. A secondary CTV54 target volume, extending inferiorly on either one or both sides of the neck was prescribed 54 Gy in 30 fractions to electively treat potential microscopic disease present in the lymphatic pathways. Planning target volumes (PTV66 and PTV54) were defined as 5 mm expansions around the corresponding CTVs to account for the uncertainty in patient positioning. For the purpose of optimisation, a PPTV66 volume was defined as a subtraction of the GTV from the PTV66 and cropped 3mm inside the skin. Similarly, a PPTV54 volume was defined as a subtraction of the PTV66 from the PTV54 and cropped 3mm within the skin. The PPTV54 volume was also cropped from the PTV66 volume to create a 5mm margin between the two volumes. As the PTV54 overlaps the PTV66, the PPTV volumes were required to ensure contradicting DVOs were not set that would hinder the optimisation. Additionally, the PTVs were cropped 3 mm inside the skin to ensure that the optimiser would not attempt to deliver dose to the dose build up regions where it is difficult to deposit dose with entrance fluence. VMAT optimisation was

performed on the PPTV66, PPTV54 and GTV volumes. WBCC clinical tolerances for the PTV volumes are summarised in table 2.2.

Table 2.2: Clinical tolerances for target volumes. Tolerances are expressed in terms of percentages of the prescribed dose for the respective PTV66 and PTV54 volumes. Starting optimisation priority values listed are for both target volumes. For the definitions of DVH metrics see section 2.1.2.

<b>PPTV DVH metric</b>	<b>WBCC Tolerance</b>	<b>starting optimisation priority</b>
D <sub>98</sub>	≥ 95 %	120
D <sub>50</sub>	100% ± 1.5 %	DVO not used
D <sub>2</sub>	≤ 107 %	120

### **Organs at Risk**

The OARs prioritised to be spared were the spinal cord (SC), brain stem (BS) and parotid glands. For oropharyngeal tumours, the SC was the most prominent OAR because its 50 Gy tolerance dose level was often in close proximity. A spinal cord planning risk volume (SC<sub>PRV</sub>) was defined as a 5mm expansion around the SC, except for the superior edge of the cord where no expansion was made as it directly connects to the BS at this position. For nasopharyngeal treatment sites, the BS was often a prominent OAR where the target volumes were in close proximity. A brain stem planning risk volume (BS<sub>PRV</sub>) was defined as a 5 mm extension around the BS, except for the inferior edge where no extension was made as it directly connects to the SC at this position. PRV volumes were defined for the SC and BS to take into account uncertainty and variation in their position because overdosing either could result in serious complications. The parotid glands were defined without an extension. The parotid glands often partially overlapped with the PTV54 and PTV66 volumes. In this situation, the portion of the gland overlapping into the PTV was not used during plan optimisation to avoid contradicting DVOs. The portion of the gland not overlapping with the PTV was contoured separately and used for optimisation to minimise the mean dose to the parotid glands. Additional to sparing the SC, BS and parotid glands,

the dose to the oral cavity, mandible, larynx and pharyngeal constrictors was also minimised during optimisation but with lower priority.

Table 2.3: DVH optimisation objectives for the prioritised and non-prioritised OAR volumes. Starting optimisation priority values are listed. For definition of DVH metrics see section 2.1.2.

OAR	WBCC Tolerance	Starting optimisation priority
Spinal Cord PRV	$SC_{PRV} D_{1cc} \leq 50 \text{ Gy}$	70
Brain Stem PRV	$BS_{PRV} D_1 \leq 54 \text{ Gy}$	70
Parotid Glands	$D_{mean} \leq 26 \text{ Gy}$ (Bilateral) $D_{mean} \leq 39 \text{ Gy}$ (Ipsilateral) $D_{mean} \leq 25 \text{ Gy}$ (Contralateral)	70
Oral Cavity	$D_{mean} \leq 45 \text{ Gy}$ , $D_1 < 65 \text{ Gy}$	50
Mandible	$D_{mean} \leq 50 \text{ Gy}$ , $D_1 < 70 \text{ Gy}$	50
Larynx	$D_{mean} \leq 45 \text{ Gy}$ , $V_{50Gy} < 66 \%$	50
Inferior Pharyngeal Constrictor	$V_{60Gy} < 12\%$	50
Middle Pharyngeal Constrictor	$V_{60Gy} < 75\%$	50
Superior Pharyngeal Constrictor	$V_{65Gy} < 33\%$	50
Post Mean	Volume should exclude 35 Gy isodose	50

These additional OARs were optimised only when they were originally contoured in the clinical plan. A post mean volume, contoured posterior to the SC was used to reduce the dose delivered to that region to maintain drainage of the lymph from the lymph nodes. The WBCC clinical tolerance doses for the prioritised and non-prioritised OARs are summarised in table 2.3.

#### 2.1.4. Plan Quality

Plan quality was defined in terms of the ability of a plan to deliver the prescribed dose to each of the target volumes while simultaneously minimising the dose delivered to the OARs. Plan quality was primarily assessed from the dose delivered to both PTV volumes as well as the dose delivered to the SC, BS and parotid glands. The dose delivered to target and OAR volumes were assessed with the DVH metrics outlined in tables 2.2 and table 2.3 respectively. The 3D-dose distribution was also verified using the dose colour wash function in eclipse to identify undesirable dose distribution features not identified by the DVH.

### 2.1.5. Complexity levels

A three tier system was developed where plans of increasing complexity were created for each case. Only three plans were able to be created for each case to keep the project feasible within the allotted time frame. The primary objective of this system was to create plans with a wide range of complexity in a well-defined way to investigate the relationship between the dosimetric accuracy of plans and plan complexity. Additionally, with this approach it was possible to relate the plan complexity to the quality of the plan. The clinical plan that was used to treat the patient was defined as complexity level one (C1). It was assumed that the clinical plans were of relatively low complexity as they were created following a conservative planning protocol that used a restrictive MU constraint during optimisation. MUs were restricted in the clinical plans as a workaround to minimise the number of segments with a MLC leaf opening below the minimum allowed field size of  $4 \text{ cm}^2$ . For each patient, two more plans of increasing complexity were created, labelled complexity levels two (C2) and complexity level three (C3). These complexity levels were defined in terms of relative plan quality. C3 plans approximated pareto-optimal plans that prioritised coverage of both PTV54 and PTV66 target volumes as well as sparing of the SC, BS and parotid glands. Pareto-optimal plans are those where further improving one aspect of the plan is impossible without deteriorating another aspect of the plan as a consequence [44]. Multiple plan iterations would have to be created to find the set of plans that make up the pareto-frontier to completely prove pareto-optimality of the plan. However, the C3 plans could not be proven to be truly pareto-optimal as creating multiple iterations of the plan was not feasible within the allocated time. Instead of selecting a plan from the pareto-frontier, C3 plans were selected when it appeared that further plan quality improvements were no longer

possible without beginning to compromise other objectives. The C2 plans were created to provide a middle ground in plan quality and plan complexity between the C1 and C3 plans. The goal for C2 plans was to achieve OAR sparing that was mid-way between C1 and C3 for the prioritised OARs. This approach assumed that the resulting plan complexity would also be half way between C1 and C3 plans. The target coverage planning goals for both C2 and C3 plans was to at a minimum match the coverage achieved by the C1 plan. Small improvements in target coverage in C2 and C3 plans were also accepted. As the clinical C1 plans all met the clinical target coverage objectives, larger improvements in target coverage were generally not possible.

#### **2.1.6. Progressive Resolution Optimiser**

VMAT plans were generated using the Progressive Resolution Optimiser (PRO) [28] module within the Eclipse treatment planning system (figure 2.1).

The clinical C1 plans were optimised using version 10.0.28 while the majority of C2 and C3 plans were optimised using version 11.0.31 after a departmental TPS version update during this study. For case #5, all C1, C2 and C3 plans were optimised using version 10.0.28. Differences between optimiser versions were investigated and the results are reported in section 3.2.4. A real-time plot of the DVHs (figure 2.1. : 1) was used to assess the dose delivered to each volume involved in the optimisation. During optimisation, plan improvements were achieved by manipulating the shape of the DVHs with the application of DVOs for each volume involved in the optimisation (figure 2.1. : 2). Upper DVOs were used to limit the maximum dose delivered to a specified percentage of a volume while lower DVOs were used to ensure a minimum dose was delivered to a specified percentage of a volume. Mean DVOs were used to

manipulate the mean dose given to the entire volume. Priority numbers (figure 2.1 : 2) were used to weight the relative importance of the competing DVOs.

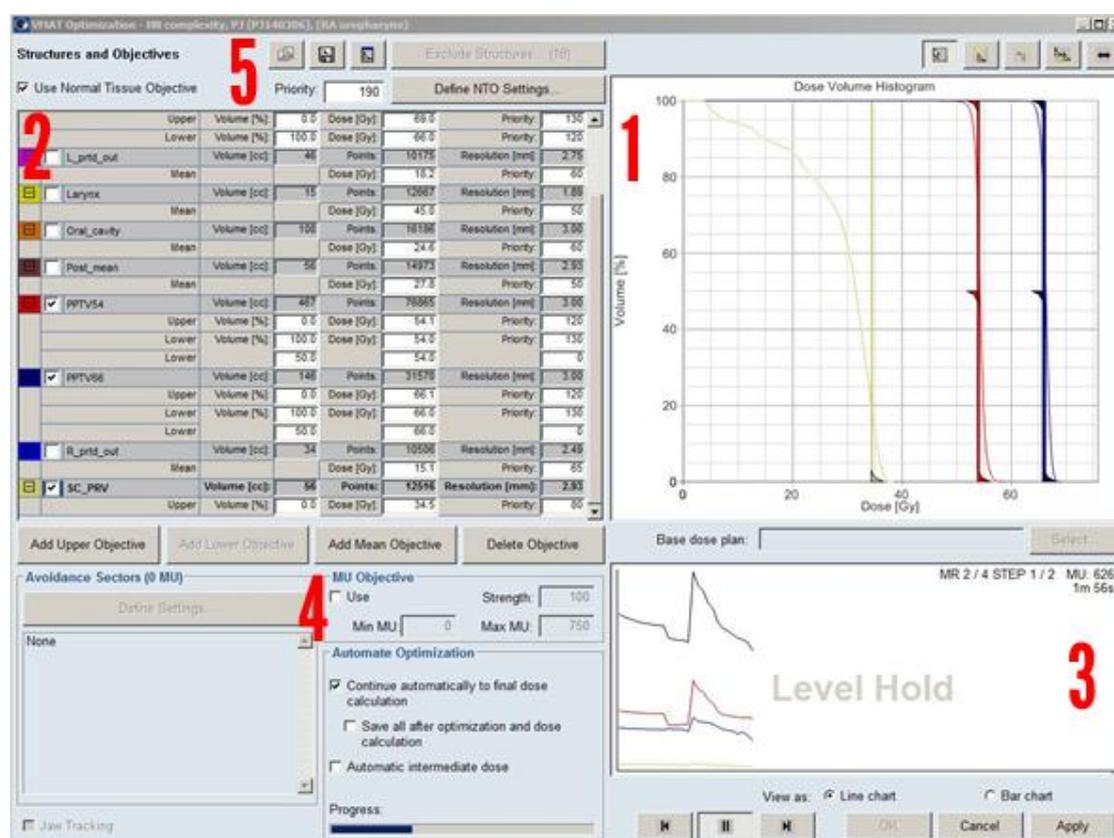


Figure 2.1: The Progressive Resolution Optimiser. Important functions used for plan optimisation are numbered 1 to 5. 1: The real time DVH plot. 2: The list of DVOs for the optimisation volumes. 3: the real time objective function line plot. 4: the MU objective. 5: the normal tissue objective function.

An objective function line plot available in Eclipse was used to visualise the contribution of each volume to the objective function (figure 2.1: 3). Although the quantity on the ordinate of this plot in eclipse is not clearly defined or quantified in contrast to some other TPS, it represents the value of the optimisation objective function. This quantity was used to monitor the progress of the optimisation and to assess whether a change to a DVO was achievable by the optimiser indicated by a downwards trend of objective function. The controls below the objective function line plot were used to select the resolution level to optimise on. Optimisation could be performed on a single resolution level using the pause function that stopped the optimiser automatically moving to the next resolution level. Resolution levels could

also be navigated by using the forward and backwards buttons. The MU used by the plan, displayed in the top right corner of the objective function line plot, was used as a rough measure of plan complexity during optimisation. The MU objective (figure 2.1 : 4) was used during the optimisation of the C1 plans, essentially limiting the MU allowed to be used by those plans. By limiting the MU used by the plan, it was possible to coarsely restrict the complexity of the plan. For C1 plans, the MU objective was set to 500 MU for cases 1, 2, 3, 5 and 400 MU for case 4 as defined by the WBCC clinical protocol for each treatment site. The priority number of the MU objective was set to 100 for all cases. The MU objective was not used for the optimisation of the C2 and C3 plans. The Normal Tissue Objective (NTO) (fig 2.1 : 5) was used to ensure that there was adequate dose fall off into normal tissue that was not explicitly contoured but located within the body contour. The normal tissue objective (NTO) was set to 'automatic sparing of normal tissue' and had a priority of 190 which was unchanged for all plans. The field parameters (table 2.4) used by the C1 plans remained the same for the C2 and C3 plans.

Final dose calculation was performed using the Analytical Anisotropic Algorithm (AAA) [45, 46] version 11.0.31 for all cases. The Truebeam LINAC beam model was calculated from in house WBCC commissioning data. The focal spot size, MLC leaf transmission and dosimetric leaf gap beam model have been reported to be the important adjustable parameters that affect the AAA dose calculation accuracy of small fields [67], like those expected to be found in the generated VMAT plans. The focal spot size parameter was set to 0.0 mm, MLC model leaf transmission parameter was set to 1.35% and the dosimetric leaf gap parameter was set to 2.0 mm.

Table 2.4: The field parameters used by all plans.

<b>Beam</b>	<b>Energy</b>	<b>Gantry rotation</b>	<b>Collimator angle</b>
CCW	6 MV Photons	179° to 181° CCW	20°
CW	6 MV Photons	181° to 179° CCW	340°

### 2.1.7. Generating C3 plans

C3 plans were generated by re-starting the VMAT optimisation of a copied C1 plan.

The plan was optimised in a number of iterations whereby the optimisation parameters were adapted using the following approach:

- Prior to starting the optimisation, the monitor unit constraint was turned off as to not restrict the complexity of the plan.
- In the initial optimisation stage, upper and lower objectives were set for the PTV66, PTV54 and GTV volumes. DVOs were not defined for OAR volumes to first achieve maximum target coverage. For each target volume, a lower objective was added at the prescribed dose for 100% of the volume and an upper objective at 0.1 Gy above the prescribed dose for 0% of the volume after which the optimisation was started.
- After the initial optimisation stage, target coverage was usually satisfactory with the  $D_{98}$ ,  $D_{50}$  and  $D_2$  metrics all well within clinical tolerance. Major changes to the DVOs were made with the optimiser paused at resolution level two, where the optimiser was still responsive to user changes of DVOs.
- DVOs were added to the non-prioritised OAR volumes so that they were spared similarly to the C1 plan. Subsequently, DVOs were added to the prioritised OAR volumes one at a time. Upper objectives that limit the  $D_{max}$  were used for  $SC_{PRV}$  and  $BS_{PRV}$  volumes while mean objectives were used for parotid glands. Setting more stringent DVOs was preferred over changing the priority of the objectives.

However, the DVO priority was increased if the optimiser became unresponsive to DVO setting changes. The priority values for OAR DVOs generally remained at 80 or below. The priority values for the target volumes were allowed to increase up to 170 where required. Further optimisation of the non-prioritised OARs was attempted where applicable, provided that sparing of the prioritised OAR was not compromised as a result.

- As more stringent DVOs were iteratively applied to the OAR volumes, OAR doses were decreased and the plan MU was increased. Often at a certain point, OAR doses were not able to be further reduced, even though there did not appear to be any trade-offs with target coverage. In this situation, the optimiser was allowed to progress through to completion and the dose distribution was calculated using the AAA algorithm using a 2.5x3.0 mm dose grid resolution. Subsequent optimisations utilising the calculated previously calculated AAA dose often allowed further OAR dose reductions to be achieved, possibly by moving the current plan away from a local minimum. After one or two re-optimisations, the plan usually approached an approximate pareto-optimal solution. This was indicated by the observation that further reductions to OAR doses could be made, but not without a corresponding degradation of target coverage or increased dose to other OARs.

Depending on the case, one of two following methods was used to select the final C3 plan. Both methods tried to achieve an equal balance between OAR sparing and target coverage.

1. For cases where minimum clinically acceptable  $D_{98}$  was easily achieved, the C3 plan was selected when the PTV66 and PTV54  $D_{98}$  were made to exceed the minimum tolerance with some added margin. For example, a PTV66  $D_{98}$  of more than 63.5 Gy was preferred over the minimum 62.7 Gy.
2. For cases where the minimum PTV66 and PTV54  $D_{98}$  requirements were fundamentally difficult to achieve, the C3 plan was selected when the  $D_{98}$  met the minimum clinical tolerance (without added margin).

Concurrently, the PTV  $D_{50}$  and  $D_2$  metrics were kept within clinical tolerance. Plans violating the clinical tolerances were re-optimised to correct the problem(s) rather than applying a global dose renormalisation to the plan.

### **Generating high quality plans**

To achieve the largest possible OAR dose reductions and best possible target volume coverage, it was necessary to use unrealistic DVOs, often resulting in a counter intuitive increase in the value of the objective function. For example while optimising the C3 plan for case #3, setting the  $SC_{PRV} D_{max}$  upper DVO to an easily achievable 45 Gy resulted in a plan that produced an actual  $D_{max}$  of around 46 Gy. There was close agreement between the prescribed objective and the achieved dose, resulting in a low contribution to the objective function. However, to dramatically decrease the  $SC_{PRV}$  dose, a more difficult  $D_{max}$  objective of ~25 Gy was used, resulting in an actual achieved  $D_{max}$  of around 35 Gy. As the agreement between the prescribed objective and achieved dose was poor, the contribution to the objective function was large. Therefore, lower OAR doses were often achieved despite increasing the value of the objective function. Generally, the rule was to set DVOs well beyond the desired goals because the optimiser would never fully achieve the prescribed DVO. As such, focus

was placed on the absolute quality of the plan rather the value of the objective function itself.

### **2.1.8. Generating C2 plans**

Optimisation of the C2 plans were performed similarly to the procedure described above. However, instead of aiming to find a pareto-optimal plan, the DVOs were adjusted such that the SC  $D_{1cc}$ , BS  $D_1$ , and Parotid Gland  $D_{mean}$  were half way between that achieved by the C1 and C3 plans. Target coverage characteristics were kept similar to the C3 plan. When a plan had achieved these goals, optimisation was completed and the C2 plan was selected.

### **2.1.9. Final dose calculation**

After finalising the C2 and C3 plans, the final dose distributions were calculated with the AAA algorithm using a 1.0 mm dose grid resolution. For all cases, the assessment of plan quality was made with the final dose calculation using AAA version 11.0.31.

### **2.1.10. Plan delivery**

VMAT plans were delivered using a Varian Truebeam LINAC equipped with a 120 leaf Millennium MLC. Leaf width was 5 mm for the central 20 cm of the field and 10 mm for the outer 20 cm of the field. Maximum MLC leaf speed was 2.76 cm/s [68], maximum machine dose rate was 600 MU/min and maximum gantry rotation speed was 6 deg/s.

## 2.1. Quantification of plan complexity

### 2.2.1. Average Leaf Pair Opening (ALPO)

A modified version of the average leaf pair opening (ALPO) as defined by Zygmanski *et al.* [47] was used as a measure of the average field size as defined by the MLC leaves. First, a leaf pair opening (LPO) was defined as the distance between pairs of opposing MLC leaves. Fully closed MLC leaf pairs (LPO=0) were not included in the calculation of the ALPO. The ALPO for an individual segment (ALPO<sub>s</sub>) was then calculated using the following equation 2.1:

$$ALPO_s = \frac{\sum_j |X_R - X_L|_j}{N_s^{open\ pairs}} \quad (\text{Equation 2.1})$$

where  $X_R$  and  $X_L$  are the MLC positions of the two opposing MLC leaves (right bank/left bank),  $N_s^{open\ pairs}$  are the number of open pairs in the segment,  $j$  is the index for leaf pairs with a non-zero LPO and  $s$  is the segment index. Up to 60 leaf pairs could be used to collimate the beam with the Truebeam's Millennium MLC.

Subsequently, the ALPO for a beam (ALPO<sub>b</sub>) consisting of  $s$  segments was calculated with equation 2.2:

$$ALPO_b = \frac{\sum_s ALPO_s}{N_b^{segments}} \quad (\text{Equation 2.2})$$

where  $N_b^{segments}$  is the number of segments in the beam,  $s$  is the segment index and  $b$  is the beam index. All beams in this study consisted of 177 segments.

Finally, the ALPO for a plan consisting of  $b$  beams was calculated with equation 2.3:

$$ALPO = \frac{\sum_b ALPO_b}{N_p^{beams}} \quad (\text{Equation 2.3})$$

where  $N_p^{beams}$  is the number of beams in the plan. All plans in this study consisted of two beams.

### 2.2.2. Critical Leaf Pair Opening (CLPO)

The Critical leaf pair opening (CLPO) metric quantified the proportion of LPOs that were smaller than a threshold of 10 mm present in a VMAT plan. Fully closed leaf pairs were not included in the calculation. The CLPO for a segment ( $CLPO_s$ ) was calculated using equation 2.4:

$$CLPO_s = \frac{N_{crit, s}^{LPO < 10}}{N_s^{open\ pairs}} \quad (\text{Equation 2.4})$$

where  $N_{crit, s}^{LPO < 10}$  are the number of LPOs smaller than 10mm in the segment,

$N_s^{open\ pairs}$  are the number of open pairs in the segment and s is the segment index.

Subsequently, the CLPO for a beam ( $CLPO_b$ ) consisting of s segments was calculated with equation 2.5:

$$CLPO_b = \frac{\sum_s CLPO_s}{N_b^{segments}} \quad (\text{Equation 2.5})$$

where  $N_b^{segments}$  is the number of segments in the beam and b is the beam index.

Finally, the CLPO for a plan was calculated with equation 2.6:

$$CLPO = \frac{\sum_b CLPO_b}{N_p^{beams}} \quad (\text{Equation 2.6})$$

where  $N_p^{beams}$  is the number of beams in the plan and b is the beam index.

### 2.2.3. Mean Leaf Travel (LT)

The mean leaf travel (LT) metric quantified the mean distance travelled by MLC leaves during delivery of a beam. Mean distance travelled by all MLC leaves between two segments, s-1 and s ( $LT_s$ ) was calculated with equation 2.7:

$$LT_s = \frac{\sum_l |x_{l,s} - x_{l,s-1}|}{N_s^{MLC\ leaves}} \quad (\text{Equation 2.7})$$

where  $x_s$  and  $x_{s-1}$  are the MLC positions for segment s and segment s-1 respectively,  $N_s^{MLC\ leaves}$  is the number of MLC leaves in the segment and l is the leaf index.

Subsequently, LT for a beam was calculated with equation 2.8:

$$LT_b = \frac{\sum_s LT_s}{N_b^{segments}} \quad (\text{Equation 2.8})$$

where  $N_p^{segments}$  is the number of segments in the beams and s is the segment index.

Finally, LT for a plan was calculated with equation 2.9:

$$LT = \frac{\sum_b LT_b}{N_p^{beams}} \quad (\text{Equation 2.9})$$

where  $N_p^{beams}$  is the number of beams in the plan and b is the beam index.

### 2.2.4. Monitor Units (MU)

The MU used by each plan was recorded from eclipse after dose calculation was performed for the finalised plans. Beam MU was correlated against ALPO, CLPO and LT.

## 2.3. Verification measurement methods

### 2.3.1. Time resolved point dose analysis

#### Ionisation chambers and measurement phantom

Time resolved point dose measurements [48] were made using 0.015 cc and 0.030 cc PTW pinpoint ionisation chambers (figure 2.2) [49]. The ionisation chambers were cross-calibrated with the WBCC local reference in a 6 MV photon beam in accordance with the TRS-398 code of practice [50]. The physical dimensions of the chambers and the correction factors used for the calculation of absolute dose are displayed in table 2.5:

Table 2.5: Summary of the physical dimensions and absolute dose calculation / correction factors of the 0.015cc and 0.030cc PTW Pinpoint ionisation chambers.

Ionisation chamber	Active volume Dimensions length/diameter (mm)	Dose to water calibration factor Gy/nC ( $N_{D,w,Q0}$ )	Ion recombination correction factor ( $K_s$ )	Polarity correction factor ( $K_{pol}$ )	Beam quality correction factor ( $K_{Q,Q_0}$ )	Water to plastic water correction factor ( $C_w$ )
PTW 0.015cc Pinpoint	5.00/1.00	1.204	1.000	0.996	0.993	0.998
PTW 0.030cc Pinpoint	5.00/1.45	2.414	1.000	0.999	0.993	0.998

Measurements were made using a 30x30x20 cm plastic water stack (“The original”, Computerized Imaging Reference Systems) made up of multiple slabs varying from 5.0 cm to 0.1 cm in thickness. A cavity drilled into the side of a 2.0 cm thick slab was used to position the ionisation chamber at the centre of this slab. Points posterior or anterior to the LINAC isocentre were measured by placing the cavity slab either higher or lower in the stack. Points superior or inferior to the LINAC isocentre were positioned by applying sagittal couch shifts using the couch position read out as a

reference. Only one cavity position was available in the cavity slab and the phantom was always positioned isocentrically. As a result only points in the sagittal plane through the isocentre were measured.



Figure 2.2: 0.015 cc (top) and 0.030 cc (bottom) PTW pinpoint ionisation chambers

### **Time resolved measurement acquisition**

A PTW T10016 Tandem electrometer provided a polarisation voltage of -300V to the ionisation chamber and collected the ionisation chamber signal. Time resolved measurement data was obtained by operating the electrometer in continuous streaming mode with a time period of 0.1 s using in-house control software [51]. In this mode, the electrometer reported the mean measured current ( $I_{t,av}$ ) in amperes over each 0.1 s time period ( $t$ ), allowing for the calculation of time resolved charge in coulombs ( $Q_t$ ) with equation 2.10.

$$Q_t = I_{t,av} \times 0.1 \text{ s} \quad (\text{Equation 2.10})$$

The control software saved the time resolved charge data to a comma separated values file that would be imported into the analysis software. Prior to, and immediately after delivery of the VMAT beam, data was acquired for 60s to later be used for leakage

correction of the measurement. Temperature and pressure were monitored for  $K_{tp}$  correction [50] of the measurement using the following equation 2.11:

$$K_{tp} = \frac{(273.15 + T_{amb}) \times 101.325}{293.15 \times P_{amb}} \quad (\text{Equation 2.11})$$

Where  $T_{amb}$  is the measured temperature ( $^{\circ}\text{C}$ ) and  $P_{amb}$  is the measured pressure (kPa).

### TPS dose calculation

A verification plan was created by copying the patient VMAT plan onto a CT data set of the plastic water phantom (figure 2.3).

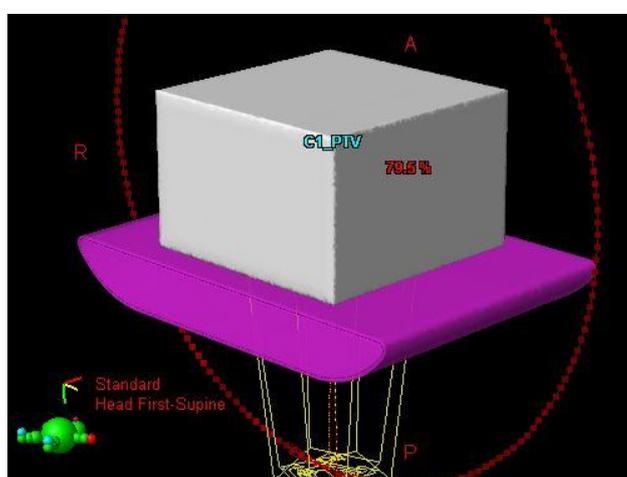


Figure 2.3: A screenshot of the verification plan used for TPS dose calculation of point dose and film measurements. The HU values of the plastic water structure were overridden to 0 HU. Values of -970 HU and -450 HU were assigned to the body and surface of the virtual couch structure respectively

The HU values of the plastic water structure were overridden to 0 HU. The centre of the plastic water phantom was determined by finding the intersection of three radiopaque markers placed on the phantom during the CT scan and was aligned to the treatment isocentre. A virtual couch structure was used to model the attenuation of the beam through the treatment couch. Based on WBCC experimental data, values of -970 HU and -450 HU were assigned to the body and surface of the couch respectively. Measurement points were selected by creating reference points in the verification plan. The AAA TPS dose was calculated with a 1.5 mm dose grid

resolution. A work around was used to extract the TPS calculated dose per segment as eclipse does not provide this data. The verification plan was exported to in-house software that converted the 177 segments of each VMAT beam into two new plans containing 177 separate static beams each (one for each VMAT beam). The MU weight of each static beam was defined as the MU delivered between adjacent control points. The MLC configuration of each static beam was defined as the average leaf position between adjacent control points. The converted plans were imported into the TPS and dose for both of the reconstructed plans was calculated with a 1.5 mm dose grid resolution. After dose calculation, DICOM RT plans now containing segment resolved TPS calculated dose data ( $D_{s,tps}$ ) for the defined reference points were exported to the in-house point dose analysis software.

### **Segment dose analysis**

Pre-irradiation charge leakage was approximated with linear model fitted to the 60 s of data acquired before the VMAT beam delivery and was subtracted from the time resolved charge data. Post-irradiation charge leakage was approximated with an exponential decay model fitted to the 60 s of data acquired after the VMAT beam was delivered and was subtracted from the time resolved charge data [48]. The measured time resolved dose ( $D_{t,exp}$ ) with equation 2.12:

$$D_{t,exp} = Q_{t,corr} \times N_{d,w,q0} \times K_{q,q0} \times K_s \times K_{pol} \times C_w \times K_{tp} \times OC \quad (\text{Equation 2.12})$$

where  $Q_{t,corr}$  is the leakage corrected time resolved charge,  $N_{d,w,q0}$  is the dose to water calibration factor for pinpoint ionisation chamber determined in a C-60 beam,  $K_{q,q0}$  is the beam quality correction factor converting the calibration factor from C-60 beam quality to the LINAC 6 MV beam quality,  $K_s$  is the ion recombination correction factor to correct for incomplete charge collection due to ion recombination,  $K_{pol}$  is the

polarity correction factor to correct the ionisation chamber response due to a change in polarity of the polarisation voltage,  $C_w$  is the water-plastic water correction factor to account for the measurement being performed in plastic water,  $K_{tp}$  is the temperature/pressure correction factor and OC is the correction for LINAC output, measured for each session.

Data contained in the LINAC log files was used to determine the elapsed time where the delivery of each segment was completed ( $T_s$ ). The measured segment resolved dose ( $D_{s,exp}$ ) was then determined by selecting  $D_{t,exp}$  for times where  $t=T_s$ . An additional time shift was required to compensate for the difference between the time clocks of the various computer systems. As a final step,  $D_{s,exp}$  and  $D_{s,tps}$  were plotted as a function of control point to verify the synchronisation and compare the measured and calculated dose for all segments (figure 2.4). This method enabled measurement of the delivered dose at each control point of the treatment plan. In this way, observed deviations can be correlated to specific machine parameters such as field size, gantry angle, and position of the detector relative to the field edge.

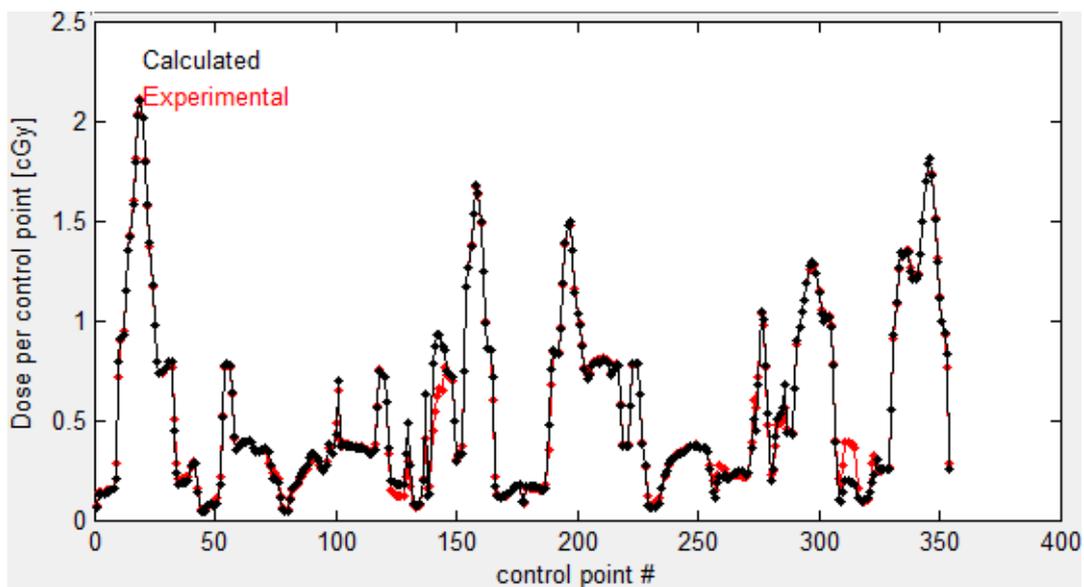


Figure 2.4: Segment resolved measured (red) and TPS calculated dose (black) plotted together to assess the agreement between measured and calculated segment doses.

### **VMAT dose delivery repeatability**

Segment resolved point dose measurements with a PTW 0.015 cc Pinpoint chamber were used to quantify the repeatability [52] of VMAT dose delivery for a C1 plan and a C3 plan. The chamber was placed in the plastic water stack phantom and positioned at the LINAC isocentre using the patient setup lasers and LINAC cross-wires.

Repeatability was determined over ten consecutive measurements without altering the setup of the phantom and chamber. Phantom temperature and atmospheric pressure were monitored throughout the session for  $K_{tp}$  correction of each measurement.

Segment dose delivery repeatability for each plan was determined by calculating the standard deviation for each of the measured segment doses over ten measurements.

Integral dose delivery repeatability for each plan was determined by calculating the standard deviation of ten integral dose measurements.

### **End-to-end integral dose verification**

The measured integral dose ( $D_{exp}$ ) was calculated by summing  $D_{s,exp}$  for all 354 segments. Using the patient structure set, two measurement points located in PTV and an OAR were selected for each plan. The selected points were transferred to the plastic water phantom verification plan. Ideally, measurement points were positioned in areas where the 3D dose gradients were small to minimise the impact of detector positioning errors. This was not always possible for points positioned near OAR volumes. A cylindrical contour modelling the size of the pinpoint ionisation chambers' active volume was drawn around each measurement point. The TPS calculated integral dose ( $D_{tps}$ ) was defined as the mean dose delivered to the chamber contour to simulate ionisation chamber volume averaging [53, 54]. AAA dose calculation was performed with a 1.0 mm dose grid resolution. The integral dose deviation ( $\Delta D_{int}$ ) was calculated with equation 2.13 for all plans.

$$\Delta D_{int} = \frac{(D_{exp} - D_{tps})}{D_{tps}} \quad (\text{Equation 2.13})$$

For each complexity level, the integral dose deviations of OAR and PTV points of interest were averaged separately to assess the VMAT dosimetric accuracy for each region of interest. Additionally, for each complexity level, integral dose deviations for OAR and PTV points were combined into a single average to assess overall VMAT dosimetric accuracy.

### **Sensitivity to phantom positioning**

To assess the sensitivity of point dose measurements to small phantom positioning errors, multiple measurements were made with artificially induced setup errors in three directions.

## **2.3.2. EBT3 Gafchromic Film**

### **Equipment**

EBT3 Gafchromic film dosimetry [55] was used to measure 2D dose distributions at coronal dose planes in a 30x30x20cm plastic water stack made up of multiple slabs varying from 5.0 cm to 0.1 cm in thickness. In-house developed Matlab software was used for analysis of the film dosimetry results. Films were scanned at 72 DPI in transmission mode with an Epson 10000XL scanner. A typical scan of an irradiated film is displayed in figure 2.5.

### **Film scanning protocol**

Films were scanned at least 12 hours after the film was irradiated. The irradiated films were handled with gloves to avoid contamination and kept in low light conditions. A frame was used to ensure films were reproducibly positioned in the scanner.

The effect of scanner output variability was reduced with the use of a scanning protocol that exploited the warm up characteristics of the scanner as follows:

- The scanner was warmed up with 30 consecutive scans performed in quick succession.
- Each film was scanned three times in quick succession to compensate for the cooling down of the scanner after opening the lid to replace films.

The third film scan was selected for analysis because the scanner output typically stabilised to a reproducible level for this scan.

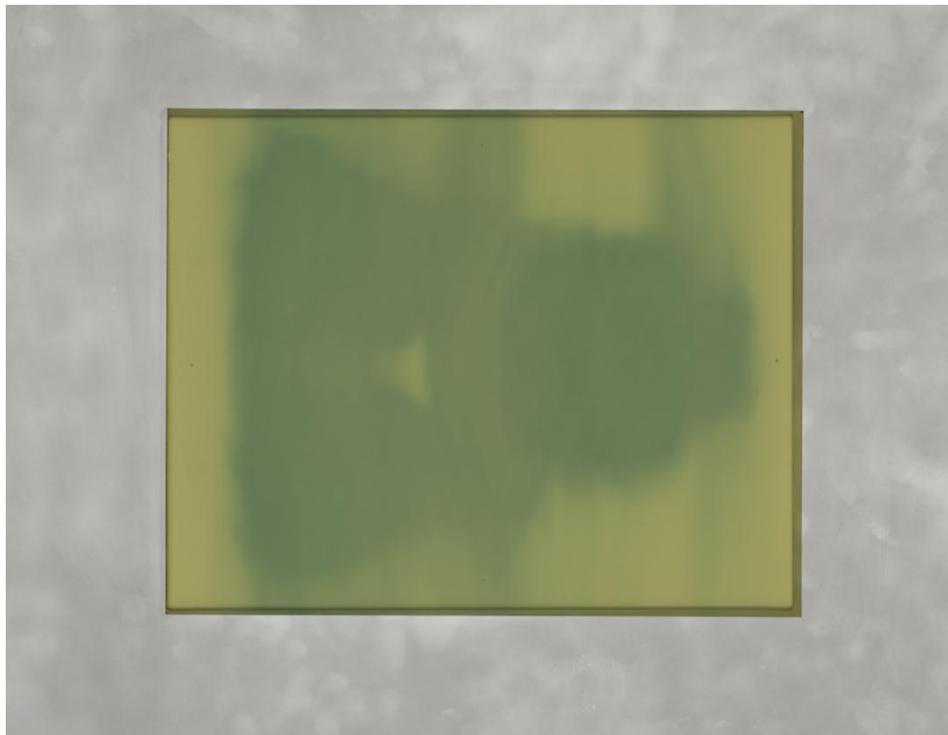


Figure 2.5: A typical scan of an irradiated EBT3 film. The delivered dose distribution is revealed by the darkened areas of the film. A frame (grey) was used to ensure films were reproducibly positioned in the scanner.

### Optical density calculation

Film optical density ( $OD_{film}$ ) was calculated with equation 2.14:

$$OD_{film}^c = -\log_{10} \left( \frac{PV_{film}^c}{PV_{empty}^c} \right) \quad (\text{Equation 2.14})$$

where  $PV_{film}$  are the pixel values of the scanned film,  $PV_{empty}$  are the pixel values of a scan without a film present (empty scan) and  $c$  is the colour channel (red, green, blue).

The net OD using the red channel ( $OD_{net}$ ) was calculated with equation 2.15:

$$OD_{net} = \frac{OD_{film}^r}{OD_{film}^b} - \frac{OD_{blank}^r}{OD_{blank}^b} \quad (\text{Equation 2.15})$$

where  $OD_{blank}$ , representing the OD of a non-irradiated film was subtracted from the experimental film OD ( $OD_{film}$ ). Additionally, film specific thickness variation was corrected by dividing red channel OD by the blue channel OD. The net OD was subsequently used for dose calculation.

### Sensitometric curve

A calibration film was created by irradiating a film at 5cm depth in plastic water (95cm SSD) with a jaw defined step-wedge plan producing ten increasing dose levels ranging from ~0.0 Gy to ~3.9 Gy along the superior/inferior axis (figure 2.6). The reference calibration dose was generated from the step-wedge TPS AAA dose calculation after the plan was modified to match central axis ionisation chamber measurements of the dose levels.



Figure 2.6: An example of a step-wedge calibration film. Ten dose levels from ~0.0 to ~3.9 Gy were created with a jaw defined step-wedge plan.

The sensitometric curve relating dose to net OD was modelled with a gamma distributed single hit model [56] (equation 2.16.):

$$D(OD)_i = a_i e^{\frac{-\log(1-OD_i/c_i)}{b_i}} - 1 \quad (\text{Equation 2.16})$$

where  $D(OD)_i$  is the sensitometric curve,  $a_i$ ,  $b_i$ , and  $c_i$  are the curve fitting parameters and  $OD_i$  is the net OD. The model was extended such that each lateral position of the film ( $i$ ) had an individually calculated sensitometric curve allowing for the effective removal of lateral position scanning artefacts. The fitting parameters ( $a_i, b_i, c_i$ ), described by second degree polynomial functions with dependence on lateral position were found by optimising the fit between the measured and reference calibration dose using the Levenberg-Marquardt least-squares curve fitting algorithm [57, 58]. The generated fitting parameters were saved as a calibration set to be referenced for the dose calculation of measurement films.

### **End-to-end integral dose verification**

Film analysis used the same verification plan that was used for point dose analysis, described in section 2.3.1. Two film measurement planes were defined by creating reference points at the depth of the plane of interest in the verification plan. Film planes were selected at the same depths where point doses were measured so dose measured by the two methods could be compared. The AAA TPS dose was calculated with a 1.0 mm dose grid resolution.

The measurement film, blank film, empty scan, calibration set and TPS plan/dose matrix were imported into the film analysis software. TPS calculated dose was corrected to account for the LINAC output on the day of the film measurement.

Measured film dose planes were compared to TPS calculated dose planes with gamma analysis [59] using dose deviation of 2% ( $\Delta D = 2\%$ ) and distance to agreement of 2 mm ( $DTA = 2$  mm) criteria.  $\Delta D$  values were calculated relative to the average dose delivered to the high dose area of the film (80% of the 99th percentile film dose) that approximated the PTV in the dose plane. An isodose threshold of 50% of the maximum plane dose was used for gamma analysis of both PTV and OAR planes.

Film results of different plans were compared with gamma pass rates and mean gamma values. For each complexity level, gamma pass rates and mean gamma values for OAR and PTV planes were averaged separately to assess VMAT dosimetric accuracy for each region of interest. Additionally, for each complexity level, gamma pass rates and mean gamma values for OAR and PTV planes were combined into averages to assess overall VMAT dosimetric accuracy.

### 2.3.3. ArcCheck

#### Equipment

End-to-end integral dose verification measurements were performed using the ArcCheck cylindrical diode array phantom (figure 2.7) [60, 61]. The ArcCheck had 1386 diode detectors arranged helically at a depth of 2.9 cm in a cylindrical acrylic (PMMA) phantom. Adjacent diodes were spaced 1.0 cm apart. An acrylic insert was used to fill the air gap located in the centre of the cylinder. The SNC Patient software [61] was used to control the device during data acquisition and was also used to perform gamma analysis between the measured and calculated diode doses.

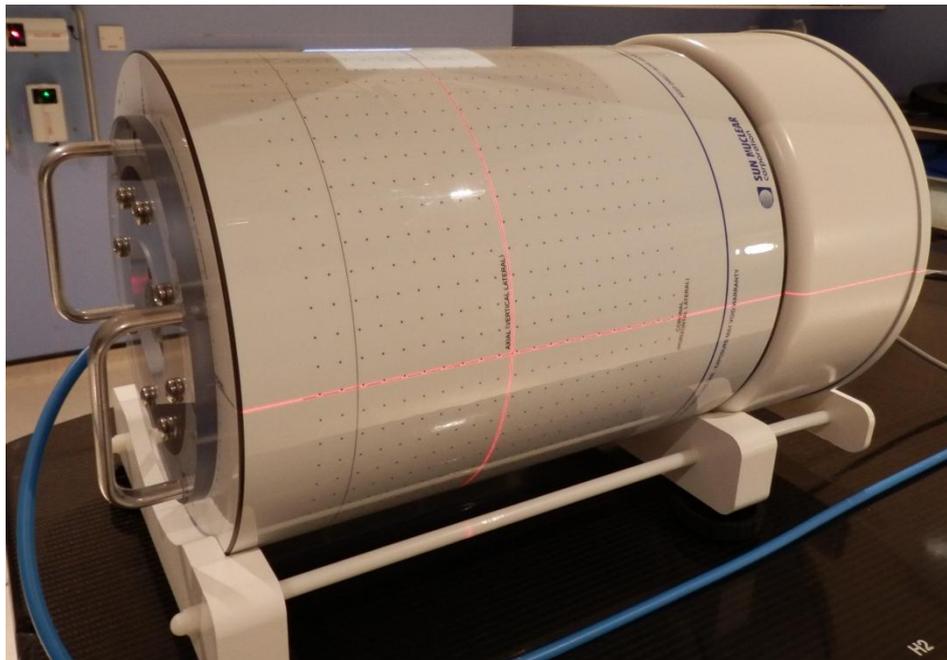


Figure 2.7: The ArcCheck diode array. Visible is the central acrylic insert and the external alignment markers used to position the phantom at isocentre.

#### ArcCheck Calibration

The ArcCheck dose calibration was performed by irradiating two calibration diodes with a 200MU, 10x10cm 6MV photon field. AAA TPS dose calculation of the calibration setup was used to determine that 267cGy/200MU was delivered to the calibration diodes. Immediately after irradiating the ArcCheck with the calibration

field, the TPS calculated calibration dose was entered into the SNC patient software to convert the calibration diode meter readings into dose, creating an absolute dose calibration file. The remaining diodes were calibrated relative to the calibration diodes using an array calibration file. The stability of the array calibration was monitored with by irradiating the ArcCheck with a conformal arc each month.

### **TPS dose calculation**

A verification plan was created by copying a VMAT plan onto a CT scan of the ArcCheck phantom. A TPS HU override was not applied to the phantom structure and the ArcCheck heterogeneity correction was turned off in SNC patient. The central reference point of the ArcCheck was determined by finding the intersection of three radiopaque markers placed on the ArcCheck's external positioning lines during the CT scan and was aligned to the treatment isocentre. A virtual couch structure was used to model the attenuation of the beam through the treatment couch. Based on WBCC experimental data, values of -970 HU and -450 HU were assigned to the body and surface of the couch respectively. AAA dose calculation was performed using a 1.0 mm dose grid resolution. The calculated dose matrix and RT plan were exported in DICOMRT format to be imported into the SNC patient software.

### **Sensitivity to phantom positioning**

Systematic phantom shifts were applied in three directions to assess the impact of positioning errors on ArcCheck results.

### **Plan measurement**

The ArcCheck was positioned at the LINAC isocentre and calibrated. All 30 plans were measured in a single three hour session. To minimise variation of the results due

to phantom positioning, measurements were performed on all 30 plans without altering the setup of the phantom between measurements.

### **Analysis**

TPS calculated dose was imported into the SNC patient software. Agreement between measured and calculated dose was quantified with gamma analysis using gamma criterion  $\Delta D=2\%$  and  $DTA = 2\text{mm}$  and a dose threshold of 10%. Absolute dose mode was used with the Van Dyk  $\Delta D$  calculation turned on. Software calculated alignment shifts were not used. Gamma pass rates were averaged for C1, C2 and C3 plans to assess VMAT accuracy for the three complexity levels.

## 3. Results

For the following results, differences between complexity levels were analysed with a series of statistical significance tests. With the small sample size of five cases, differences with statistical significance ( $p < 0.05$ ) could not be proven. However, clear trends in the data were observed and these are reported where the result was consistently observed for all cases. It is estimated that three more cases displaying similar results would allow for significance to be proven.

### 3.1. Plan quality

In the following section, plan quality indicated by DVH metrics (described in section 2.1.2.) for target coverage and OAR sparing are reported for all complexity levels. DVH metrics are an efficient method to compare plans, but important spatial details of the dose distribution may be missed. Therefore, in addition, typical features of the dose distributions created by C1, C2 and C3 plans are illustrated.

#### 3.1.1. Target coverage

PTV66 and PTV54  $D_{98}$ ,  $D_{50}$ , and  $D_2$  DVH metrics are displayed for all complexity levels in table 3.1. Relative to the C1 plans, the C2 and C3 plans consistently increased the PTV66  $D_{98}$  for all cases, but the magnitude of the increase varied from case to case. On average, C2 plans increased the PTV66  $D_{98}$  by  $0.7 \pm 0.4$  Gy (1 S.D.), while C3 plans increased the PTV66  $D_{98}$  by  $0.6 \pm 0.3$  Gy. Despite all three complexity levels having similar PTV54  $D_{98}$  and  $D_2$  metrics, C2 and C3 plans consistently produced PTV54  $D_{50}$  that was closer to 54 Gy than the C1 plans indicating increased

dose uniformity. No other consistent trends with complexity level were observed for the other PTV54 and PTV66 DVH metrics.

Table 3.1: Target coverage DVH metrics comparison between C1, C2 and C3 plans for PTV66 and PTV54 volumes. Metrics are expressed as mean (range). Wilcoxon matched-pair signed rank test (two tailed) p-values are listed for C1 vs. C2 and C1 vs. C3. For definition of DVH metrics see section 2.1.2.

Complexity level	C1	C2	C3	C1/C2 p	C1/C3 p
<b>PTV66</b>					
D <sub>98</sub> (Gy)	62.9 (62.4-63.3)	63.6 (62.5-64.1)	63.5 (62.7-64.0)	0.063	0.058
D <sub>50</sub> (Gy)	66.2 (65.8-67.0)	66.2 (65.9-66.7)	66.1 (65.9-66.9)	0.892	0.462
D <sub>2</sub> (Gy)	68.1 (67.5-69.6)	68.1 (67.6-69.0)	68.1 (67.7-69.5)	0.685	0.855
<b>PTV54</b>					
D <sub>98</sub> (Gy)	51.6 (51.3-52.0)	51.9 (51.2-52.4)	51.8 (51.0-52.3)	0.201	0.269
D <sub>50</sub> (Gy)	54.5 (54.2-55.0)	54.3 (53.9-55.1)	54.2 (53.8-54.8)	0.136	0.058
D <sub>2</sub> (Gy)	56.7 (56.0-57.6)	56.7 (56.0-58.0)	56.7 (56.2-57.5)	0.855	1.000

### 3.1.2. OAR sparing

Spinal cord, brain stem and parotid gland DVH metrics for all complexity levels are displayed in table 3.2. Relative to C1 plans, C2 and C3 plans reduced the SC D<sub>1cc</sub> by on average  $5 \pm 2$  Gy and  $10 \pm 3$  Gy respectively. Additionally, on average C2 and C3 plans increased the minimum distance from the SC<sub>PRV</sub> to the critical 50 Gy isodose by 1.1 mm (0.5 mm to 2.0 mm) and 1.8 mm (1.5 mm to 6.0 mm) respectively. For cases where the target volumes were in close proximity to the brain stem, C2 and C3 plans decreased the BS D<sub>1</sub> on average by  $6 \pm 1$  Gy and  $10 \pm 3$  Gy. For cases where the brain stem was not in close proximity to the target volumes, the brain stem dose was not explicitly optimised and no further dose reduction was observed. For cases where the C1 treatment plans produced mean parotid doses greater than 20 Gy, the corresponding C3 plans reduced parotid gland mean dose by 6 to 8 Gy. C3 plans made smaller parotid dose reductions (2 to 4 Gy) where the corresponding C1 plans produced doses less than 15 Gy. Improved sparing of the prioritised OARs made by C2 and C3 plans was consistently observed for all cases.

Table 3.2: OAR DVH metrics comparison between C1, C2 and C3 plans for Spinal Cord, Brainstem, left Parotid Gland and right Parotid Gland. Metrics are expressed as mean (range). Wilcoxon matched-pair signed rank test (two tailed) p-values are listed for C1 vs. C2 and C1 vs. C3. For definition of DVH metrics see section 2.1.2.

Complexity level	C1	C2	C3	C1/C2 p	C1/C3 p
<b>Spinal Cord PRV</b>					
D <sub>1cc</sub> (Gy)	42.8 (41.5-43.5)	38.6 (37.0-42.0)	31.6 (28.8-35.4)	0.063	0.063
Min. Distance to 50 Gy isodose (mm)	1.4 (1.0-3.0)	2.5 (1.5-4.0)	3.2 (1.5-6.0)	0.054	0.058
<b>Brainstem PRV</b>					
D <sub>1</sub> (Gy)	28.8 (5.0-49.7)	26.3 (5.7-48.8)	23.1 (5.9-43.1)	0.313	0.313
<b>Left Parotid</b>					
D <sub>mean</sub> (Gy)	18.5 (8.0-28.1)	16.9 (9.8-25.2)	14.2 (8.1-22.0)	0.188	0.125
<b>Right Parotid</b>					
D <sub>mean</sub> (Gy)	17.3 (12.0-26.6)	16.9 (11.4-26.6)	14.5 (10.4-21.1)	0.100	0.058

The dose given to non-prioritised OARs was not increased as a consequence of the focus put on reducing the prioritised OAR doses. C2 and C3 plans produced non-prioritised OAR doses that were at least as low as what was achieved with the C1 plans. Mean oral cavity dose (calculated for five cases) was similar for all complexity levels ( $22 \pm 9$  Gy). The mandible dose was calculated for two cases. For case #4 the mandible dose was similar for all complexity levels while for case #5 the C3 plan reduced the mandible dose by 5 Gy.

In specific cases where sparing was possible, there were noteworthy improvements in sparing of the larynx and pharyngeal constrictors. The C3 plan for case #5 decreased the larynx dose from 50.7 Gy (C1 plan) to 35.0 Gy. A smaller 5.8 Gy decrease was achieved by the C3 plan for case #1. Finally, the C3 plan for Case #2 reduced the pharyngeal constrictor dose by 5.8 Gy.

### 3.1.3. Case Study

In this section, case #1 is presented to illustrate the effect plan complexity typically had on the dose distributions of an individual patient. Cumulative dose volume

histograms comparing C1, C2 and C3 plans are plotted in figure 3.1 and DVH metrics are displayed in table 3.3.

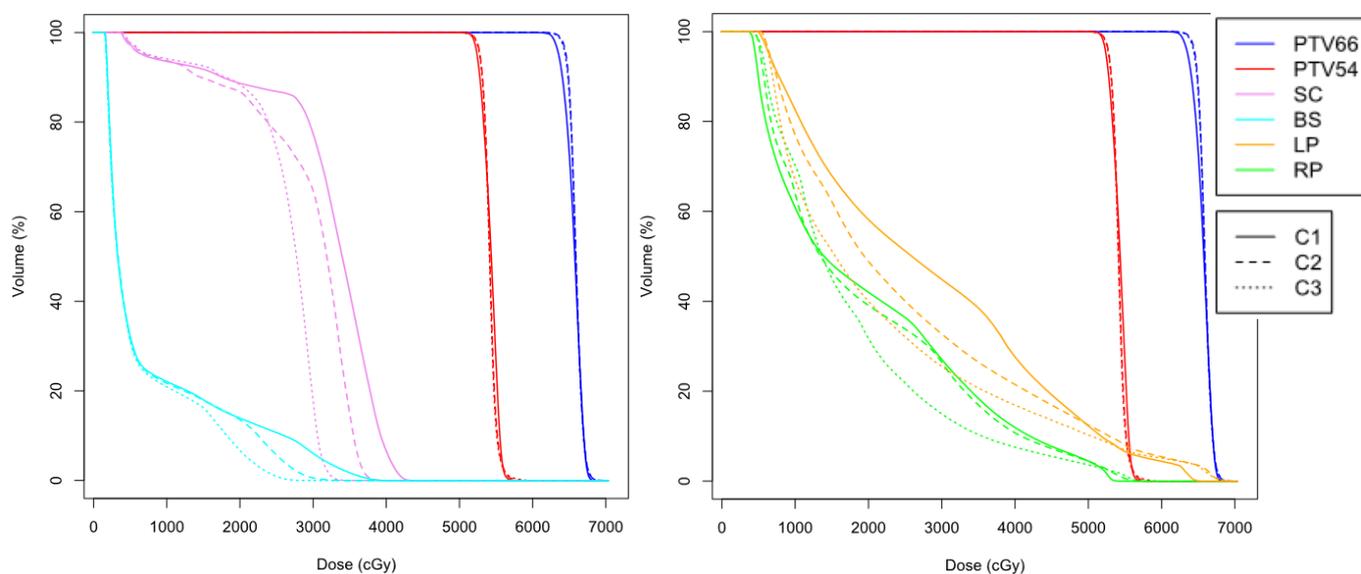


Figure 3.1: Cumulative dose volume histograms for case #1, comparing C1 (solid lines), C2 (dashed lines) and C3 (dotted lines) plans. DVHs for PTV66, PTV54, Spinal Cord (SC), Brainstem (BS), left Parotid Gland (LP) and right Parotid Gland (RP) are shown in their respective colours.

For this case, the C2 and C3 plans showed improvement in both target coverage and OAR sparing. The C2 and C3 plans increased the PTV66  $D_{98}$  by 1.1 Gy and 1.0 Gy respectively. The C2 and C3 PTV66 DVH dropped off more steeply without compromising the  $D_{50}$  or increasing  $D_2$  indicating a more homogeneous dose distribution. Similarly, C2 and C3 plans increased the PTV54  $D_{98}$  by 0.4 Gy and 0.3 Gy respectively. Similar to the PTV66 DVH, The C2 and C3 PTV54 DVH also dropped off more steeply than the C1 plan.

The C2 and C3 plans decreased the SC  $D_{1cc}$  by 4.5 Gy and 9.4 Gy respectively. Similarly, the C2 and C3 plans reduced the BS  $D_1$  by 7.1 Gy and 11.4 Gy respectively. The C2 and C3 plans reduced the mean left parotid gland dose by 2.9 Gy and 6.1 Gy respectively. The C3 plan decreased the mean right parotid gland dose by 2.0 Gy while the C2 plan produced a negligible 0.2 Gy decrease.

Table 3.3: Target volume and OAR DVH metrics for the C1, C2 and C3 plans of case #1. For definition of DVH metrics see section 2.1.2.

<b>Complexity level</b>	C1	C2	C3
<b>PTV66</b>			
D <sub>98</sub> (Gy)	63.0	64.1	64.0
D <sub>50</sub> (Gy)	65.8	65.9	65.9
D <sub>2</sub> (Gy)	67.5	67.6	67.7
<b>PTV54</b>			
D <sub>98</sub> (Gy)	52.0	52.4	52.3
D <sub>50</sub> (Gy)	54.3	54.1	54.1
D <sub>2</sub> (Gy)	56.1	56.2	56.2
<b>Spinal Cord PRV</b>			
D <sub>1cc</sub> (Gy)	41.5	37.0	32.1
Min. Distance to 50 Gy isodose (mm)	3.0	4.0	6.0
<b>Brainstem PRV</b>			
D <sub>1</sub> (Gy)	36.2	29.1	24.8
<b>Left Parotid</b>			
D <sub>mean</sub> (Gy)	28.1	25.2	22.0
<b>Right Parotid</b>			
D <sub>mean</sub> (Gy)	19.8	19.6	17.8

The dose reduction to OAR volumes was a result of increased dose fall off between the target volume and the respective OAR. Sagittal plane dose distributions through the centre of the SC<sub>PRV</sub> volume illustrate the dose fall off achieved by the three complexity levels (figure 3.2). It can be seen that along the entire length of the spinal cord, the steepness of the dose fall off from PTVs was increased with higher complexity plans. Additionally, with increasing plan complexity the minimum distance from the proximal edge of the SC<sub>PRV</sub> to the 50 Gy isodose increased from 3mm (C1) to 4mm (C2) and 6mm (C3) (figure 3.3). The dose reduction to the spinal cord and region posterior to the spinal cord is also visible on the transverse plane (figure 3.4). Improved sparing of larynx is visible on the sagittal plane where the C2 and C3 plans decreased the mean larynx dose from 49.3 Gy (C1 plan) to 45.2 Gy and 43.5 Gy respectively. The transverse plane dose distributions (figure 3.4) show that the reduction in parotid gland dose was achieved with increased dose fall off localised in the regions between the target volumes and the parotid glands.

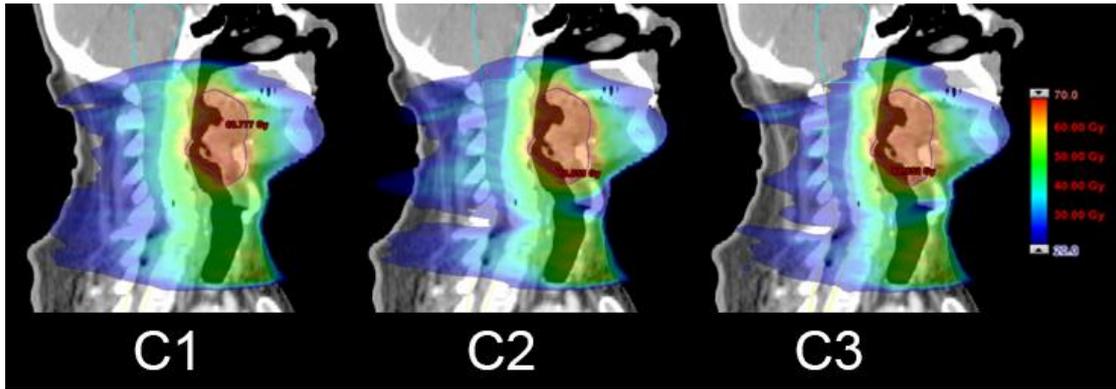


Figure 3.2: Sagittal plane dose distributions through the centre of the spinal cord PRV volume (yellow contour). Increased plan complexity produced steeper dose drop off from the PTV66 volume (purple contour), resulting in a reduction to the SC PRV  $D_{1cc}$ .

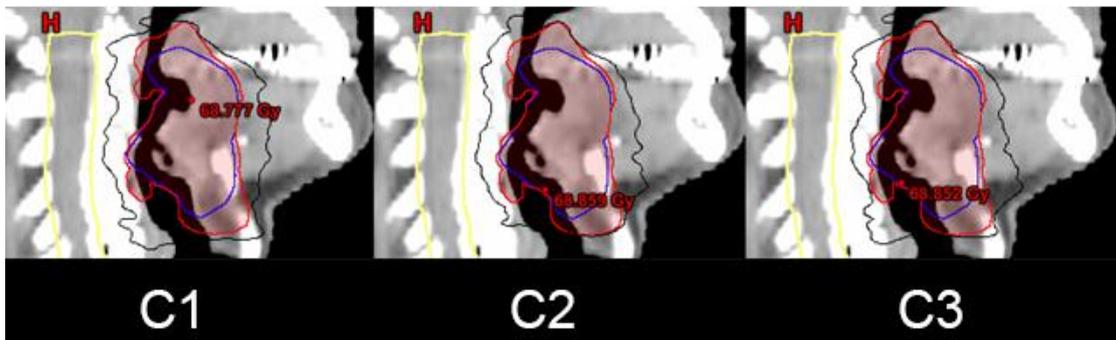


Figure 3.3: Sagittal plane dose distributions through the centre of the spinal cord PRV volume (yellow contour). Higher plan complexity increased the minimum distance from the proximal edge of the SC PRV to the 50 Gy isodose (black isodose) from 3mm (C1) to 4mm (C2) and 6mm (C3).

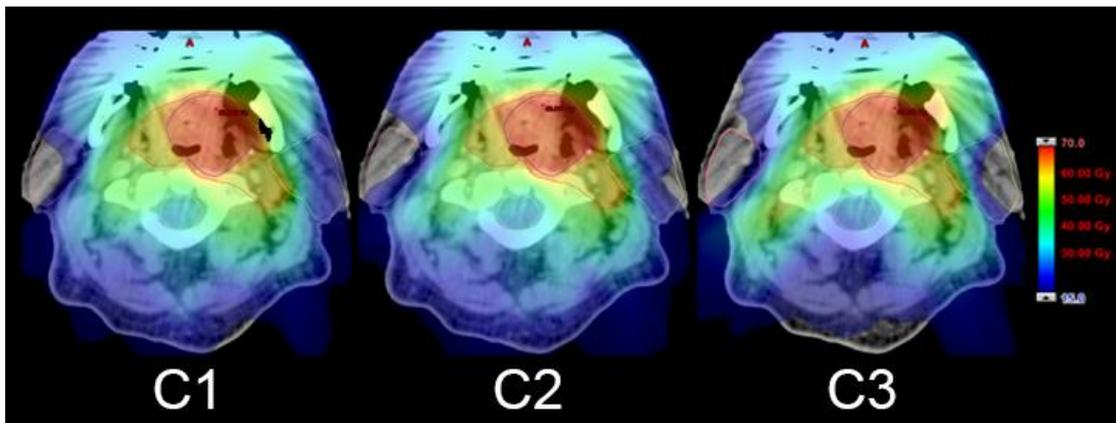


Figure 3.4: Transverse plane dose distributions through the left (Yellow contour) and right (Coral contour) parotid glands. Increased plan complexity produced steeper dose drop off from the PTV66 (blue contour) and PTV54 volumes (red), reducing the mean dose to the left and right parotid glands.

## 3.2. Plan complexity

The planning strategy using complexity levels produced three distinct categories of plans defined by plan quality. In order to study complexity in relation to verification measurement results, the fundamental parameters of the plans were quantified with complexity metrics. Complexity metrics calculated for each complexity level were used to quantify MLC motion and the field sizes defined by MLC apertures. Dose rate and gantry speed modulation were analysed by plotting both parameters for each segment. The calculated parameters were related back to the MU used by the beam so that MU may be used as an informed guide of complexity during optimisation.

### 3.2.1. Complexity metrics

The complexity metrics for the counter clockwise (CCW) and clockwise (CW) beams are plotted separately in figure 3.5. For all cases, the complexity metrics displayed clear trends with complexity level. All metrics also displayed considerable variation from case to case, indicating that patient specific factors such as the relative location and size of planning volumes may also play a role in defining the complexity of the beams.

#### Monitor Units

The mean MU utilised by C1, C2 and C3 plans was  $444 \pm 36$  (1 S.D.),  $578 \pm 41$  and  $683 \pm 49$  respectively. Increased MU utilisation with higher complexity level plans was consistently observed with both CCW and CW beams for all cases. The variation in MU utilisation between cases increased with complexity level for CW beams but not for CCW beams. On average, the CW beam used  $19 \pm 26$  (1 S.D.) more MU than the CCW beam (figure 3.6), indicating that the complexity of each pair of beams may not be equal

## ALPO & CLPO

For both beams, ALPO decreased, and CLPO increased with complexity level, indicating that smaller fields were utilised in the higher complexity beams. The mean ALPO for C1, C2 and C3 plans was  $41 \pm 5$  mm,  $32 \pm 6$  mm and  $29 \pm 6$  mm respectively. Decreased ALPO with higher complexity level plans was consistently observed with both CCW and CW beams for all cases. The mean CLPO for C1, C2 and C3 plans was  $16 \pm 5$  %,  $24 \pm 9$  % and  $29 \pm 9$  % respectively.

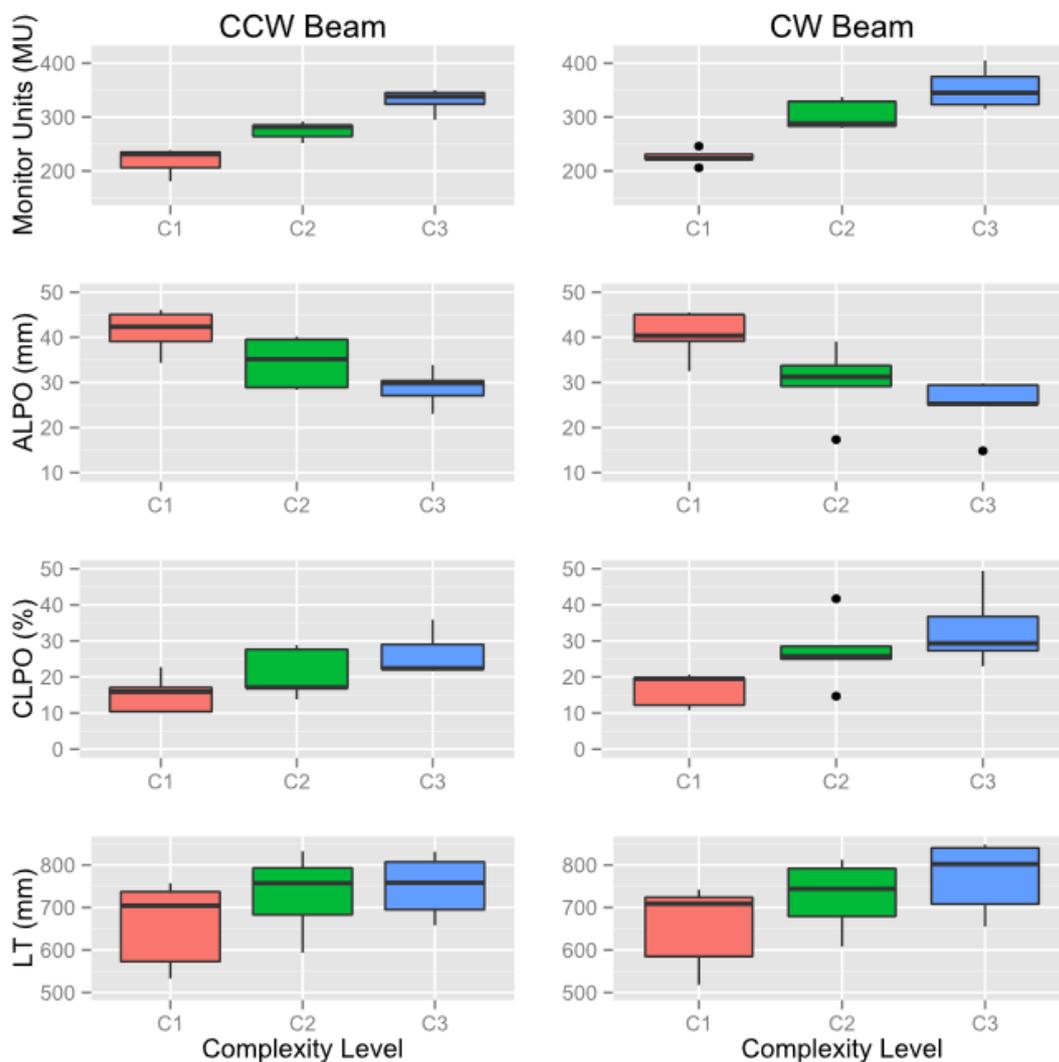


Figure 3.5: Box plots of complexity metrics, MU, ALPO, CLPO and LT against complexity level. Median values are plotted with first and third quartiles. Whiskers represent the range of data within 1.5 times the inter-quartile range. Outliers, plotted as black dots are defined as data outside 1.5 times the inter-quartile range.

Increased CLPO with higher complexity level plans was consistently observed with both CCW and CW beams for all cases. Like beam MU, there were differences between the CW and CCW beams for the ALPO and  $CLPO_{10mm}$  metrics. On average, the ALPO was  $3 \pm 5$  mm lower for CW beams for CCW beams and the CLPO was  $5 \pm 5\%$  higher for CW beams than for CCW beams (figure 3.6).

### Leaf travel

LT primarily increased from levels C1 to C2. A smaller increase was seen from C2 to C3 for the CW beams while no increase was observed for the CCW beams. The mean LT for C1, C2 and C3 plans was  $658 \pm 94$  mm,  $742 \pm 91$  mm and  $760 \pm 76$  mm respectively. On average, LT was higher by  $12 \pm 18$  mm for CW beams than for CCW beams

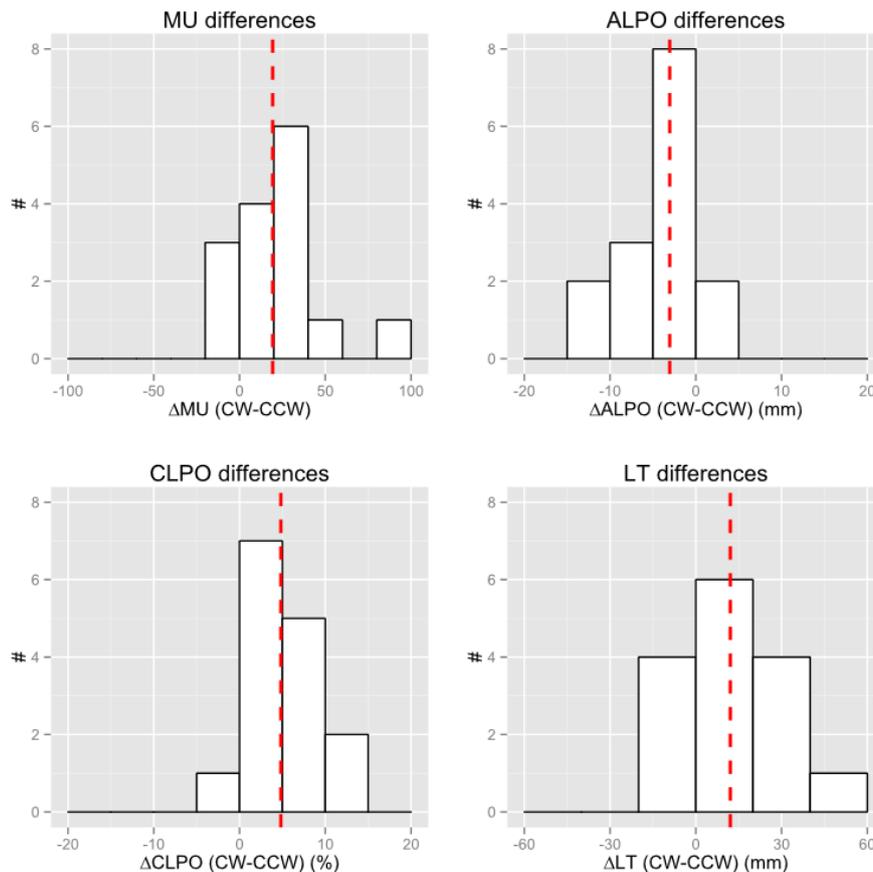


Figure 3.6: Histograms illustrating differences between CW and CCW beams for MU, ALPO, CLPO and LT complexity metrics. The red dashed line indicates the mean difference between CW and CCW beams.

### 3.2.2. ALPO, CLPO, LT relationships with MU

ALPO, CLPO and LT were correlated with beam MU (figures 3.7, 3.8, 3.9). ALPO displayed a negative linear trend with beam MU ( $r^2 = 0.62$ ), indicating that an increase in beam MU corresponds with a decrease in average field size. The CLPO metric displayed a positive linear trend with beam MU ( $r^2 = 0.53$ ) indicating that there was also an increase in the proportion of LPOs smaller than 10 mm with increased beam MU. A linear trend also existed between LT and beam MU ( $r^2 = 0.27$ ). This correlation was much weaker than the two other metrics, but on average, beams with higher MU did show higher LT.

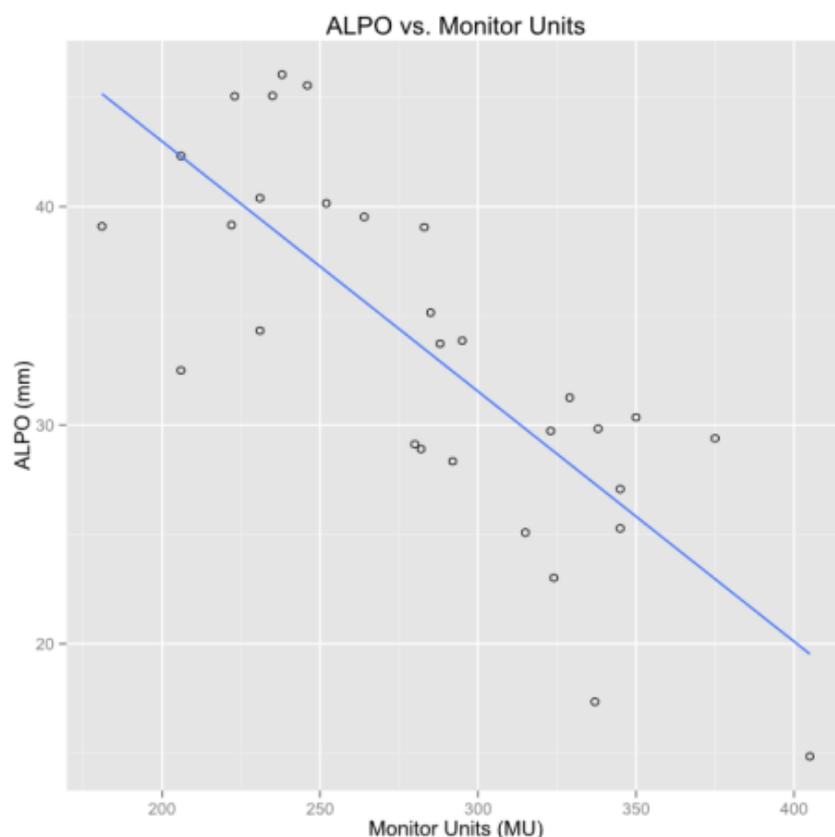


Figure 3.7: ALPO plotted against MU for all beams.  $r^2$  was 0.62 for the fitted linear model.

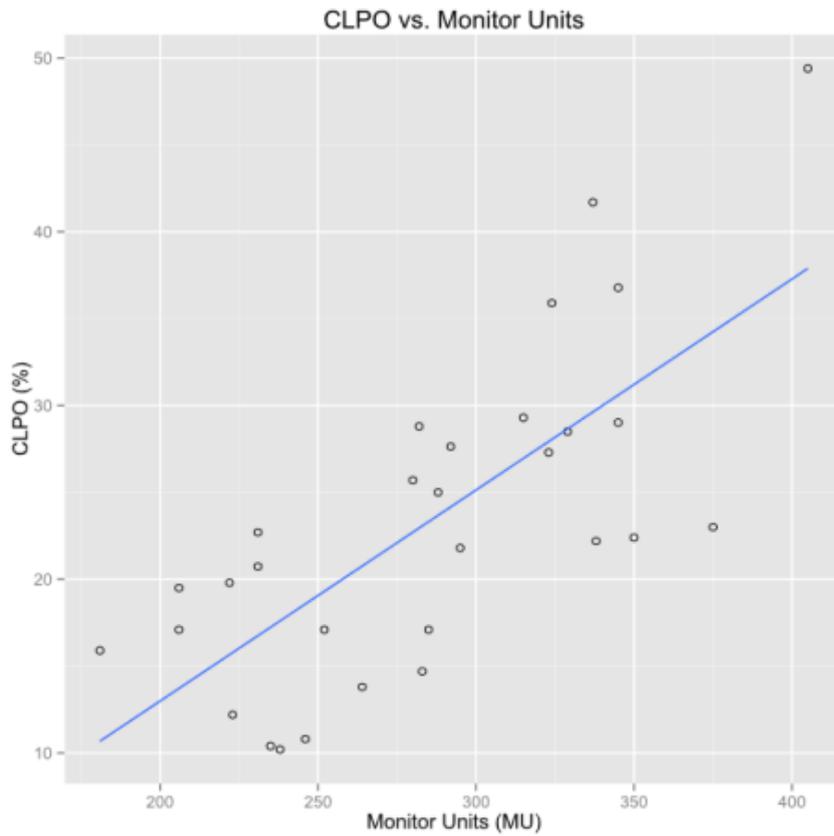


Figure 3.8: CLPO plotted against MU for all beams.  $r^2$  was 0.53 for the fitted linear model

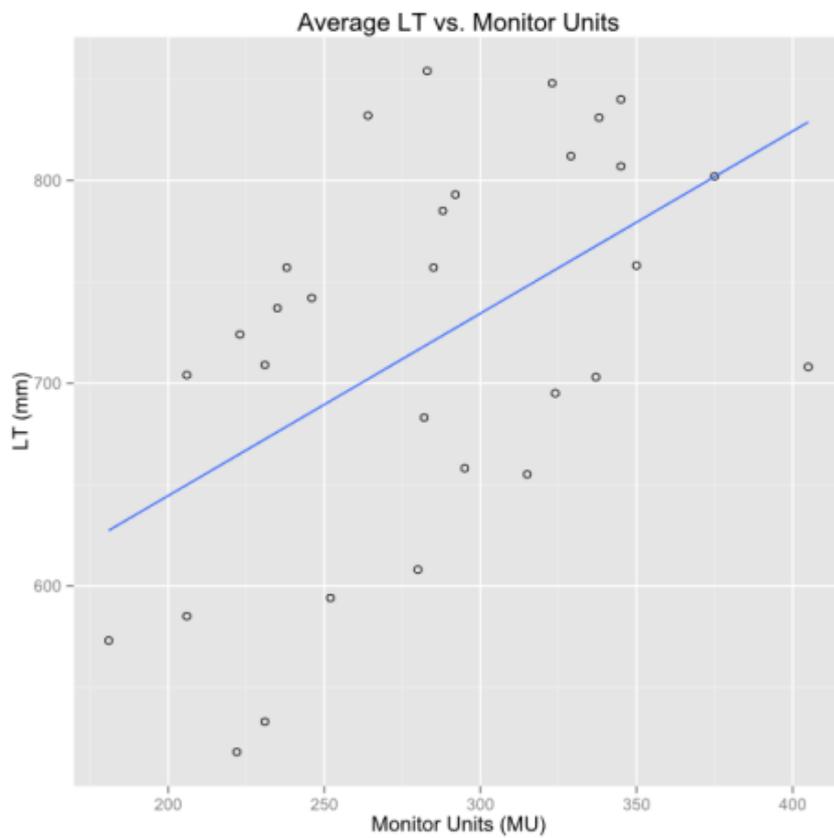


Figure 3.9: LT plotted against MU for all beams.  $r^2$  was 0.27 for the fitted linear model.

### 3.2.3. Case study

Continuing the case study started in section 3.1.3, the complexity metrics for case #1's C1, C2 and C3 plans are listed in table 3.4.

Table 3.4: Complexity metrics for case #1's C1, C2 and C3 plans.

Complexity Level	C1	C2	C2
Plan MU	437	621	690
ALPO (mm)	41	30	26
CLPO (%)	19%	28%	33%
LT (mm)	707	803	824

### 3.2.4. TPS version 11 upgrade

The eclipse TPS was unavoidably upgraded from version 10.0.28 (v10) to version 11.0.31 (v11) before the majority of the C2 and C3 plans were generated. As a result, all C1 plans were optimised using v10 while the majority of C2 and C3 plans were optimised in v11. The exception was case #5 where all plans were optimised using version 10. The major differences between the two optimiser versions appeared to be that the v10 optimiser produced plans with a more continuous style dose rate while v11 optimiser generated plans with a more step-like style dose rate (figure 3.10). It was important to ensure that the new optimiser version also did not fundamentally change the other complexity metrics: MU, ALPO, CLPO and LT. To investigate whether the new TPS version introduced new optimisation behaviour, three similar quality plans from the same case were compared. Plan “v10<sub>orig</sub>” was the original v10 clinical plan, plan “v10<sub>reopt</sub>” was a re-optimised plan using v10 and plan “v11<sub>reopt</sub>” was a re-optimised plan using v11 and the same optimisation objectives as v10<sub>reopt</sub>. By performing a re-optimisation of the plan using v10 (v10<sub>reopt</sub>), the effect of a re-optimisation using similar but potentially different optimisation objectives could be observed for the same optimiser version. The re-optimisation of v11<sub>reopt</sub> was performed to observe the differences between optimiser versions with identical

optimisation objectives. This comparison of three plans was made for three different H&N cases (independent of the five included in the complexity study).

It was found that the step-like dose rate could only be found in the plans generated with the v11 optimiser. Additionally, Case #5's C2 and C3 plans (optimised in v10) had a continuous dose rate like the other v10 plans indicating that the new optimiser version was the primary cause of this effect. The differences between the optimiser versions were negligible for the MU, ALPO and CLPO metrics. Larger differences were observed between v10, v10<sub>reopt</sub> and v11<sub>reopt</sub> for leaf travel. However, LT was also changed for the v10<sub>reopt</sub> plans implying that LT can change as a result of re-optimisation alone.

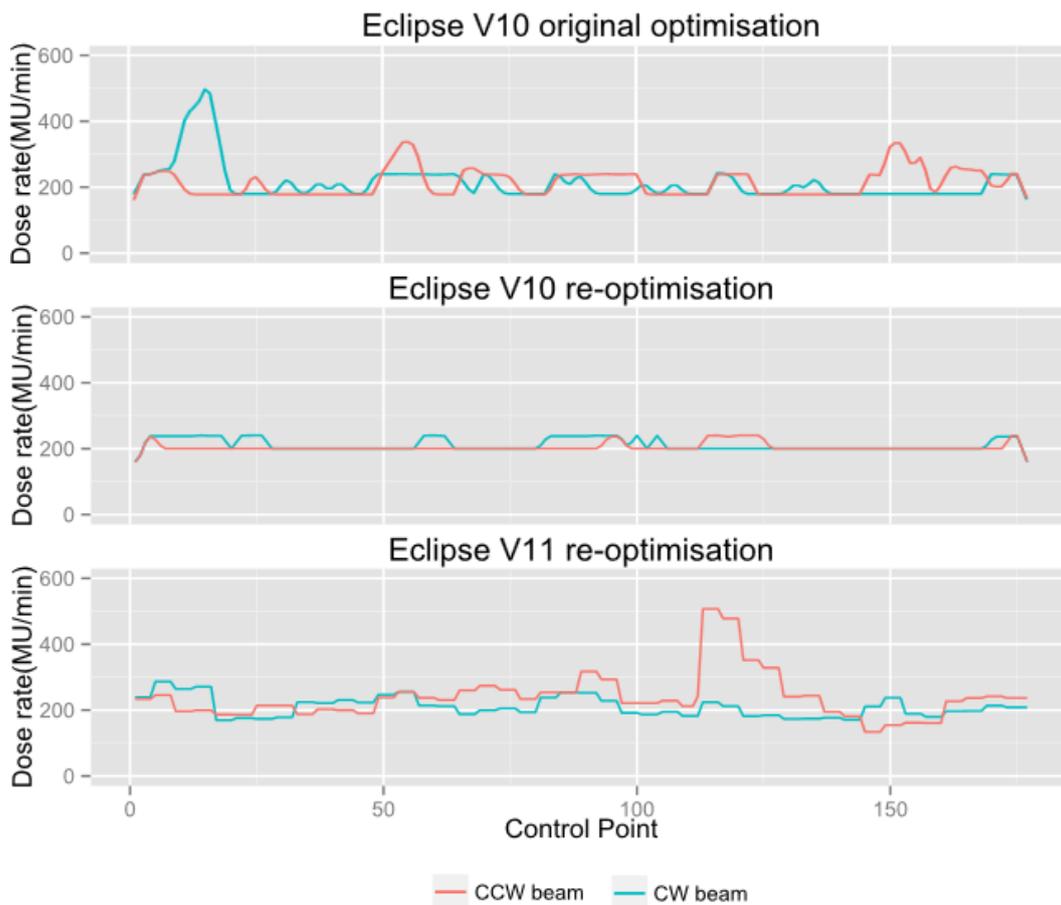


Figure 3.10: Dose rate modulation for CCW and CW beams optimised with: eclipse version 10 original optimisation (top), version 10 re-optimisation (middle) and version 11 re-optimisation (bottom). V10 optimisations displayed a continuous style dose rate modulation while the V11 optimisation displayed a “step-like” dose rate modulation. Re-optimisation of the plan was seen to change the magnitude, frequency and location of dose rate modulation.

## 3.3. Plan dosimetry verification

### 3.3.1. VMAT delivery repeatability

Ten consecutive time resolved point dose measurements were used to assess the repeatability of a C1 plan delivery while nine measurements (data for one run was corrupted) were used for a C3 plan delivery (using case #5 plans). The measured integral dose for both sets of measurements is shown in table 3.5.

Table 3.5: Results of the repeatability experiments for the integral measured dose. Mean integral measured dose and standard deviations for the C1 and C3 plans are displayed.

Plan	C1	C3
Average measured integral dose (cGy)	196.0	194.0
Absolute standard deviation (cGy)	0.06	0.05
Relative standard deviation	< 0.1%	< 0.1%

The relative standard deviation of the integral dose was less than 0.1% for both sets of measurements. The average measured segment dose is plotted against control point for the C1 plan (figure 3.11) and C3 plan (figure 3.12). Control points 1-177 represent the first beam while control points 178-354 make up the second beam. The standard deviation of segment resolved dose measurements was calculated for all 354 segments. Red error bars (1 S.D.) are plotted with the average segment dose (figures 3.11 & 3.12, top), but are not easily seen on this scale. The average relative standard deviation per segment was 0.6% for the C1 plan and 0.5% for the C3 plan. Log file analysis shows that MLC position deviations were within  $\pm 0.6\text{mm}$  for both plans. Log files reported that gantry rotation errors were  $\pm 0.15^\circ$  at the start and end of the arc and less than  $0.05^\circ$  while the rotation speed was constant. It is important to note that deviations calculated from LINAC log files represent deviations measured by the LINAC control systems, but do not take into account deviations due to error in calibration. These results indicate that the Truebeam LINAC can deliver VMAT with excellent repeatability in terms of both the segment dose and the integral dose.

Further, no difference in dose delivery repeatability was found between the low complexity and high complexity plans.

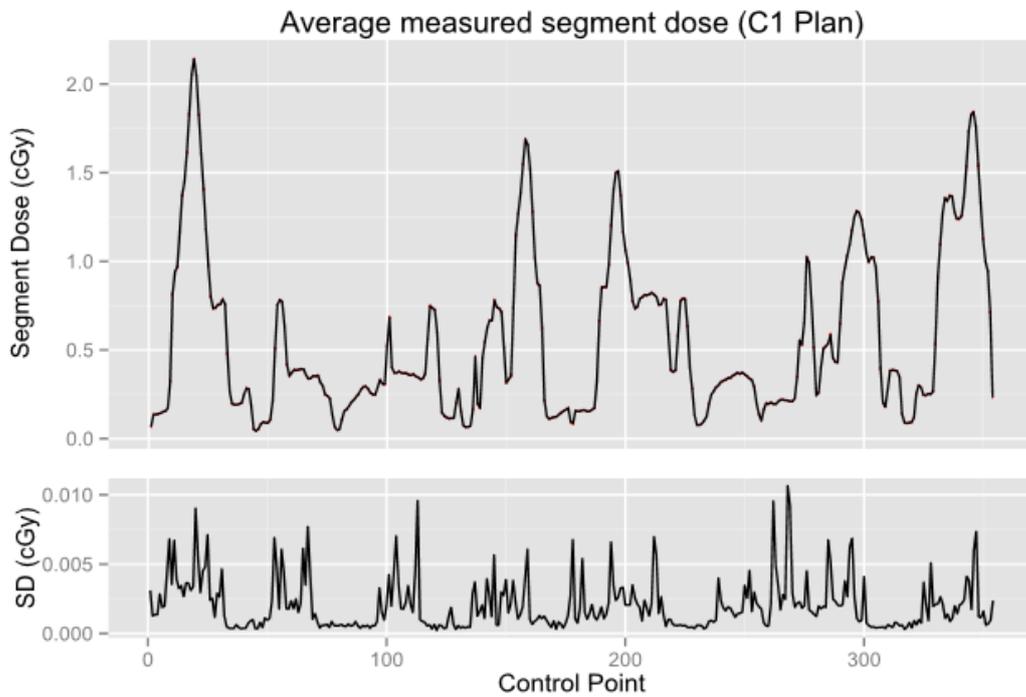


Figure 3.11: C1 plan. The average (of ten measurements) measured segment dose plotted against control point (top). The standard deviation plotted against control point (bottom). The average segment relative standard deviation was 0.6%

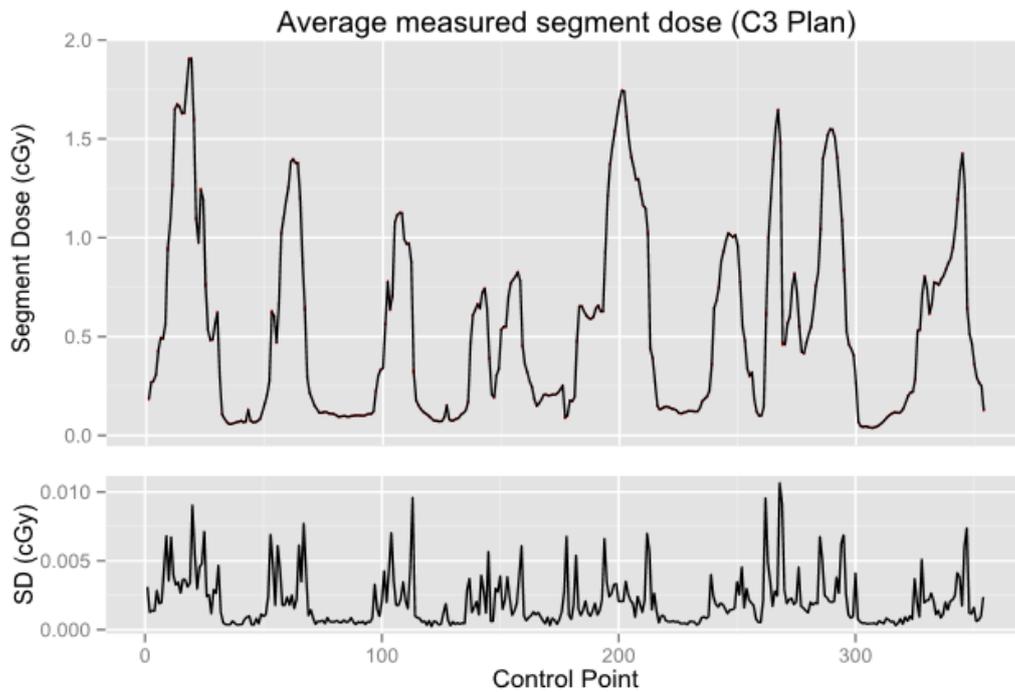


Figure 3.12: C3 plan. The average (of nine measurements) measured segment dose plotted against control point (top). The standard deviation plotted against control point (bottom). The average segment relative standard deviation was 0.5%

### **3.3.2. Impact of dose grid resolution**

The impact of dose grid resolution on the calculated dose was determined through comparison of point doses for 1.0 mm and 1.5 mm dose grid resolutions.

When the dose calculation grid resolution was reduced from 1.5 mm to 1.0 mm, TPS calculated point doses were reduced by  $0.9\% \pm 0.4\%$  for C1 plans,  $1.0\% \pm 0.3\%$  for C2 plans and  $1.2\% \pm 0.4\%$  for C3 plans. Qualitative comparison of the 1.5 mm and 1.0 mm calculated dose using dose profiles showed that the decrease in dose was generally a systematic drop for all areas of the dose distribution. Dose calculation using 1.0 mm dose grid resolution gave the best agreement between measured and calculated dose for all end to end verification methods and therefore was used for analysis of all end-to-end dose verification measurements.

### **3.3.3. Point dose verification**

#### **Sensitivity to phantom positioning**

To assess the sensitivity of point dose measurements to small phantom positioning errors, multiple measurements were made with artificially induced setup errors in three directions. It was found that measurements made in large dose gradients were more sensitive to positioning errors than measurements made in small dose gradients. Effectively, a position offset from the nominal measurement point would result in the measurement of a different part of the dose distribution. From these observations, it was expected that measurements made in large dose gradients (e.g. OAR regions) would have higher uncertainty than measurements made in small dose gradients (e.g. PTV regions). Therefore, measurements performed in OAR and PTV regions were analysed separately.

### Integral point dose analysis

Average point dose deviations for PTV points, OAR points and all points combined, for all three complexity levels are summarised in table 3.6. All point dose results were within clinically acceptable limits. No significant differences in the average point dose deviations were observed between complexity levels.

### Segment point dose analysis

Analysis of the dose deviation per segment using the methodology described by Louwe *et al.* [48] did not reveal significant systematic differences between the three complexity levels. As an example, segment deviations (PTV points) expressed as a percentage of the fraction dose for are displayed as a function of the detector distance to the closest field edge (figure 3.13). The only deviations that could be observed above the background noise level were for segments where the detector was located near the beam penumbra.

Table 3.6: Average point dose deviations for PTV points, OAR points and all points combined for C1, C2 and C3 plans. Matched-pair t-test (two tailed) p-values are listed for C1 vs. C2, C1 vs. C3 and C2 vs. C3

Complexity level	C1	C2	C3	C1/C2 p	C1/C3 p	C2/C3 p
<b>PTV points</b>						
Mean deviation $\pm$ 1 S.D.	-0.4% $\pm$ 0.8%	0.4% $\pm$ 0.7%	0.3% $\pm$ 0.8%	0.170	0.241	0.802
<b>OAR points</b>						
Mean deviation $\pm$ 1 S.D.	0.2% $\pm$ 1.1%	-0.6% $\pm$ 1.5%	-0.3% $\pm$ 1.6%	0.136	0.299	0.404
<b>All points</b>						
Mean deviation $\pm$ 1 S.D.	-0.1% $\pm$ 1.0%	-0.1% $\pm$ 1.2%	0.0% $\pm$ 1.2%	0.933	0.769	0.758

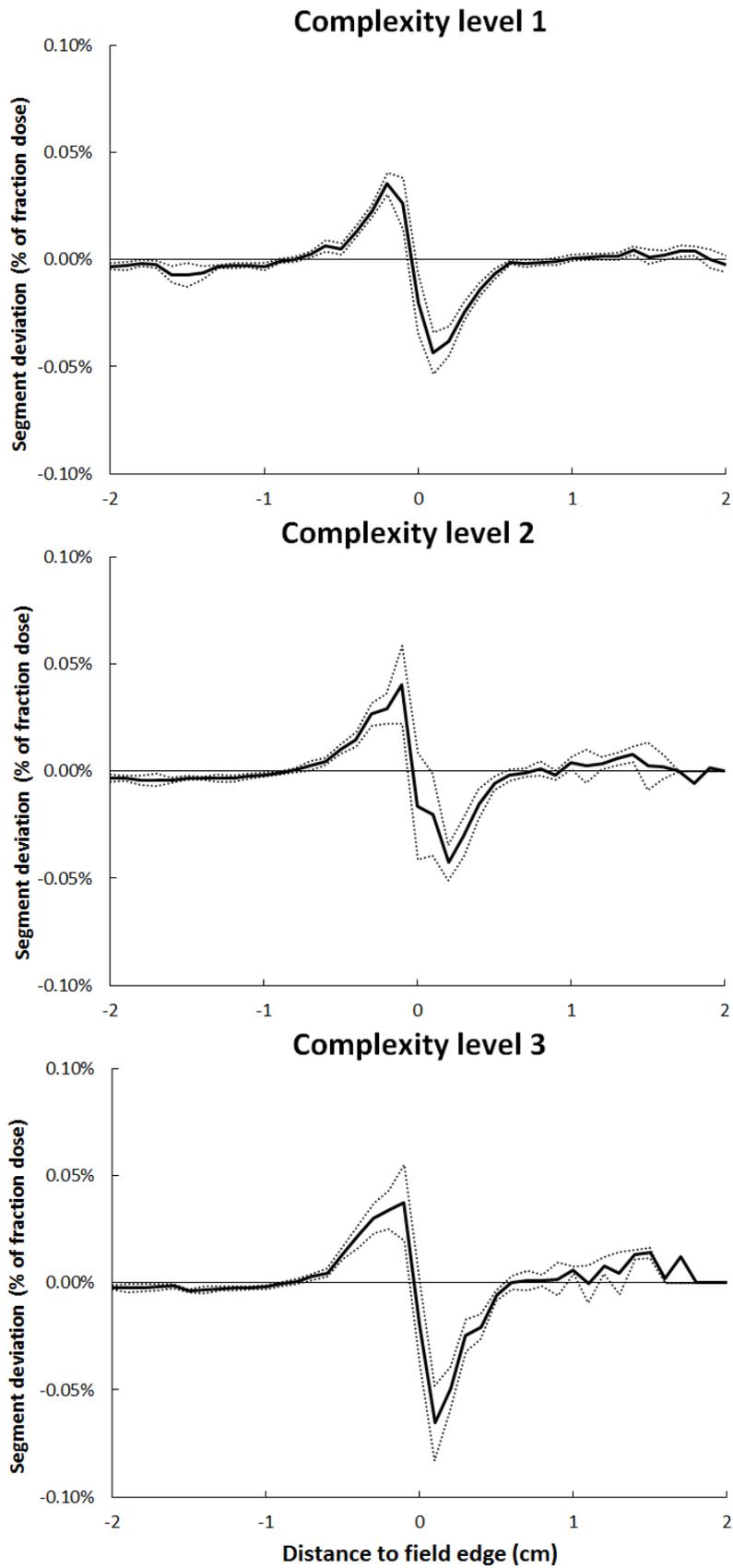


Figure 3.13: Segment point dose deviations (PTV points) expressed as a percentage of the fraction dose are displayed as a function of the detector distance to the closest field edge. The largest deviations occurred where the detector was measuring in the beam penumbra.

### 3.3.4. Film dose verification

#### Coronal plane film dose analysis

Gamma analysis (2%, 2mm criterion) was performed between the measured film dose and calculated TPS dose for each measurement plane. Average gamma pass rates and mean gamma values calculated for PTV planes, OAR plans, and all planes combined, for all three complexity levels are summarised in table 3.7. All film results were within clinically acceptable limits. No significant differences in gamma pass rate and mean gamma was observed between the three complexity levels for all measurement planes. OAR planes tended to have slightly lower gamma pass rates and higher mean gamma values than the corresponding PTV planes for a given plan. Point doses (voxels) extracted from the film were within 2% of the corresponding ionisation chamber point dose measurement

Table 3.7: Average film gamma pass rates (2%, 2mm criterion) and mean gamma values for PTV planes, OAR planes and all planes combined for C1, C2 and C3 plans. Wilcoxon matched-pair signed rank test (two tailed) p-values are listed for C1 vs. C2, C1 vs. C3 and C2 vs. C3

Complexity level	C1	C2	C3	C1/C2 p	C1/C3 p	C2/C3 p
<b>PTV planes</b>						
Mean $\gamma$ pass rate $\pm$ 1S.D.	96% $\pm$ 2%	96% $\pm$ 2%	97% $\pm$ 1%	0.586	0.410	0.572
Mean $\gamma \pm$ 1S.D.	0.25 $\pm$ 0.05	0.24 $\pm$ 0.04	0.23 $\pm$ 0.01	0.625	0.343	0.784
<b>OAR planes</b>						
Mean $\gamma$ pass rate $\pm$ 1S.D.	94% $\pm$ 3%	94% $\pm$ 2%	94% $\pm$ 1%	0.789	1.000	0.490
Mean $\gamma \pm$ 1S.D.	0.30 $\pm$ 0.07	0.32 $\pm$ 0.06	0.31 $\pm$ 0.04	0.458	0.625	1.000
<b>All planes</b>						
Mean $\gamma$ pass rate $\pm$ 1S.D.	95% $\pm$ 3%	95% $\pm$ 2%	96% $\pm$ 2%	1.000	0.606	0.470
Mean $\gamma \pm$ 1S.D.	0.27 $\pm$ 0.06	0.28 $\pm$ 0.07	0.27 $\pm$ 0.05	0.357	0.123	0.601

### 3.3.5. ArcCheck dose verification

#### Sensitivity to phantom positioning

Systematic phantom shifts were applied in three directions to assess the impact of positioning errors on ArcCheck results. Lateral phantom shifts from the nominal position at isocentre decreased gamma pass-rates (2%, 2mm) by 4% per mm shift.

Anterior/Posterior shifts from the nominal position at isocentre decreased gamma pass-rates by 2% per mm shift. Superior/Inferior shifts from the nominal position at isocentre decreased gamma pass-rates by 13% per mm shift.

### ArcCheck analysis

Gamma analysis (2%, 2mm criterion) was performed between the measured diode dose and the calculated diode dose for each plan. Average gamma pass rates for all three complexity levels are summarised in table 3.8. All ArcCheck results were within clinically acceptable limits. No significant differences in the gamma pass rates were observed between the three complexity levels.

Table 3.8: Average ArcCheck gamma pass rates (2%, 2mm criterion) for C1, C2 and C3 plans. Wilcoxon matched-pair signed rank test (two tailed) p-values are listed for C1 vs. C2, C1 vs. C3 and C2 vs. C3

Complexity level	C1	C2	C3	C1/C2 p	C1/C3 p	C2/C3 p
Mean $\gamma$ pass rate $\pm$ 1S.D.	96% $\pm$ 4%	95% $\pm$ 3%	94% $\pm$ 5%	0.125	0.188	1.000

### 3.3.6. Case study

As a continuation of the case study presented in sections 3.1.3 and 3.2.3, case #1's dose verification results are presented in this section to illustrate the typical dose verification results for C1, C2 and C3 plans.

Time resolved point dose results for the PTV measurement points are displayed in figure 3.14. Integral point dose deviations for C1, C2 and C3 plans were 0.5%, -0.5% and 0.4% respectively. Dose deviations are visible for various control points for all complexity levels but the magnitude or frequency of these deviations do not increase with complexity. These deviations primarily originated from measurements made in the beam penumbra (section 3.3.3 : segment dose analysis).

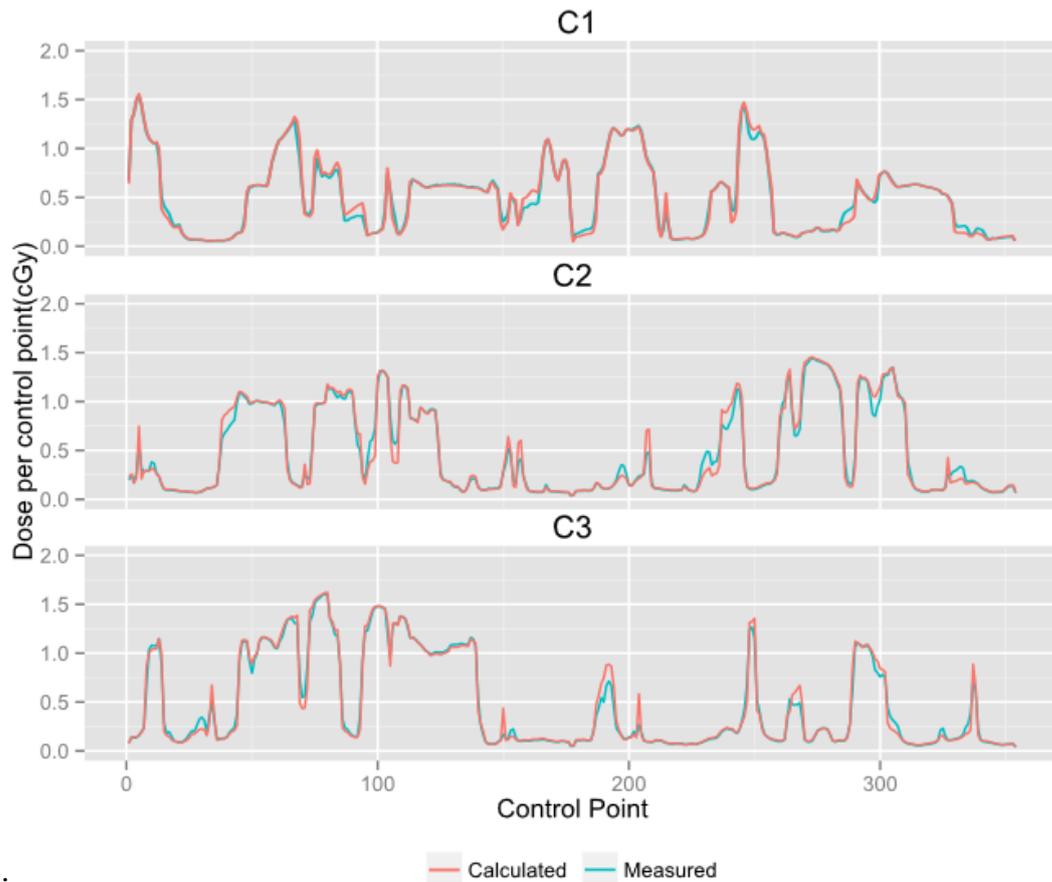


Figure 3.14: Time resolved point dose results (PTV measurement point) for case #1's C1, C2 and C3 plans. Integral point dose deviations for C1, C2 and C3 plans were 0.5%, -0.5% and 0.4% respectively. The deviations between measured and calculated dose primarily originated from measurements made in the beam penumbra.

Film gamma maps for the PTV plane of the C1, C2 and C3 plans are displayed in figure 3.15. Film gamma pass rates (2%, 2mm) for C1, C2 and C3 plans were 95%, 96% and 98% respectively.

ArcCheck gamma maps for C1, C2 and C3 plans are displayed in figure 3.16.

ArcCheck gamma pass rates (2%, 2 mm) for C1, C2 and C3 plans were 99%, 96% and 96% respectively. Overall, no clear differences in dosimetric accuracy were observed between the three complexity levels for case #1.

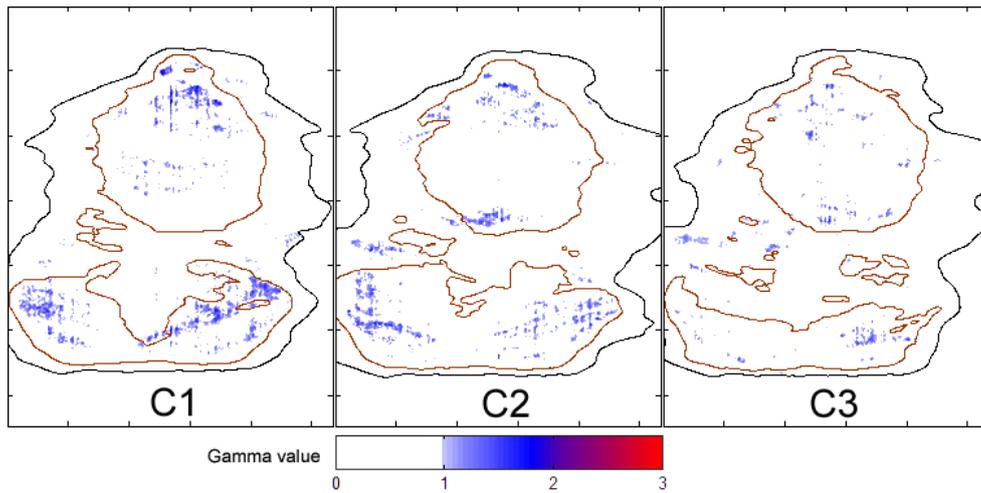


Figure 3.15: Film gamma maps (PTV planes) for case #1's C1, C2 and C3 plans. Film gamma pass rates (2%, 2mm) for C1, C2 and C3 plans were 95%, 96% and 98% respectively. Gamma values are indicated with the colour scale from white to blue to red. Gamma maps were generated using 2%, 2mm criterion.

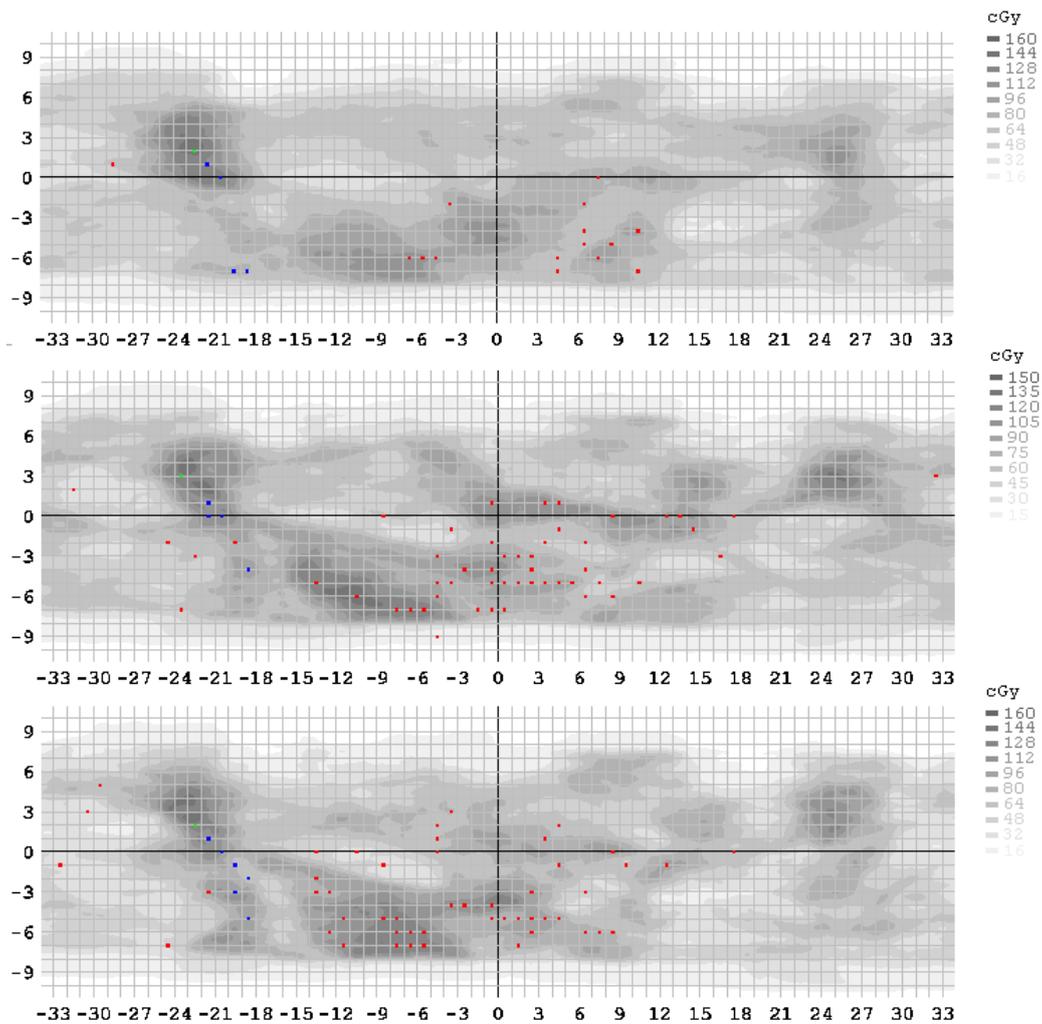


Figure 3.16: ArcCheck gamma maps for case #1's C1, C2 and C3 plans. ArcCheck gamma pass rates (2%, 2mm) for C1, C2 and C3 plans were 99%, 96% and 96% respectively. Red dots represent diodes failing gamma analysis with a high dose and blue dots represent diodes failing gamma analysis with a low dose. Gamma maps were generated using 2%, 2mm criterion.

## 4. Discussion

Ultimately, the motivation for exploring complex VMAT beam arrangements is to improve the quality of H&N RT treatment. Two major components of the overall quality of the treatment were investigated in this thesis: plan quality and plan dosimetric accuracy. Possible relationships between the two components and plan complexity were investigated by creating VMAT plans defined by a complexity level system, quantifying the complexity of those plans with metrics and performing dose verification measurements using three independent methods. Plan complexity is related to the fundamental parameters of VMAT plans such as the MLC motion, field size, dose rate modulation and gantry speed modulation. Plan complexity can be increased as a result of improving plan quality during optimisation. Additionally, plan complexity can also be related to the patient's anatomy and the geometry of planning volumes. While higher complexity plans provide an opportunity to improve the quality of H&N RT treatment, they may also deteriorate the accuracy of VMAT delivery and TPS modelling. Whether this happens or not depends on several factors including the characteristics of the LINAC and the performance of the TPS dose calculation algorithm. Furthermore, the characteristics of the applied dose verification methods determine whether specific dosimetric deviations and therefore changes in dosimetric accuracy can be detected. Each of these aspects will be discussed in this chapter.

## 4.1. Plan Quality

In this section, the plan quality characteristics produced by C1, C2 and C3 plans will be discussed in relation to the potential clinical impact they may have. The advantages and limitations of the planning methods used in this thesis will also be discussed.

The objective of VMAT is to fully cover the target volumes with the prescribed dose while limiting the dose delivered to critical OARs that are not the target of the therapy. Failing to adequately spare OARs may result in toxicity to sensitive organs, potentially lowering the patient's quality of life. Equally important, the target volumes must be covered with the prescribed dose to ensure that the effectiveness of the therapy is not compromised. Plan quality is therefore measured in terms of the ability of a VMAT plan to achieve both of these aspects simultaneously. In order to study the relationship between plan quality and plan complexity, plans categorised by three complexity levels were generated for five H&N cancer cases. It was found that by tightening the planning objectives, the dose delivered to the prioritised OARs (spinal cord, brain stem and parotid glands) could be dramatically decreased as indicated by the DVH metrics (section 3.1.2). While C2 and C3 plans were able to effectively decrease OAR doses, target coverage was not compromised and, in fact, improvements were made to the PTV66 D<sub>98</sub> and PTV54 D<sub>50</sub> on average (section 3.1.1.). Additionally, the delivered dose to the non-prioritised OARs (oral cavity, mandible, larynx and pharynx) was not increased as a consequence of the focused sparing of the prioritised OARs (section 3.1.2).

## **Statistical significance**

The statistical significance of plan quality differences observed between the three complexity levels could not be proven for the sample size used in this study.

However, some of the plan quality trends with complexity level were large in magnitude and consistently observed for all cases. It is estimated that statistical significance can be proven for the larger trends with the addition of three more cases that display similar results to the five cases presented in this thesis.

### **4.1.1 Spinal cord dose**

Relative to C1 plans, C2 and C3 plans produced steeper dose drop off along the posterior/inferior axis (figures 3.2, 3.4), reducing the  $D_{1cc}$  to the  $SC_{PRV}$  (tables 3.2, 3.3). The SC dose reduction achieved by the C2 and C3 plans would not likely reduce the NTCP for myelopathy [62] as the C1 plans were already achieving sufficiently low doses resulting in a negligible NTCP. However, reducing the planned SC dose during the first course of radiotherapy may be beneficial for possible future re-treatment. This is because the previously accumulated SC dose may often be the limiting factor for the target volume prescription of the re-treatment. More importantly for patient positioning, the minimum distance from the  $SC_{PRV}$  to the 50 Gy isodose was increased with the C2 and C3 plans (tables 3.2, figure 3.3). With day to day patient positioning setup variation including patient deformation, the SC may be moved into a 50 Gy dose region. Where the SC is at risk due to the proximity to high dose areas, the clinical instruction is to position the patient with focus on positioning the SC in the correct position. However, this compromise in patient positioning can result in an under-dosage of the target volume(s). By creating a larger distance between the SC and the 50 Gy isodose, the impact of this kind of compromise can be reduced or avoided completely.

### **4.1.2. Brain stem dose**

For cases where the brain stem was close to the target volumes, C2 and C3 plans were able to greatly reduce the BS  $D_1$  relative to the C1 plans (tables 3.2, 3.3). Discussion of the brain stem dose is similar to the spinal cord. The dose reductions achieved by the C2 and C3 plans would not likely reduce the NTCP for brain stem injury [63] because the C1 plans were already achieving sufficiently low doses resulting in negligible NTCP. However, reducing the planned BS dose during the first course of radiotherapy may be beneficial for possible future re-treatment. Additionally with reduced dose, the critical 54 Gy isodose is more likely to be positioned further away from the BS.

### **4.1.3. Parotid gland dose**

Relative to C1 plans, C2 and C3 plans reduced the mean dose delivered to the parotid glands (tables 3.2, 3.3). Large mean dose reductions of 6 to 8 Gy were achievable by C3 plans where the corresponding C1 plan produced mean doses above 20 Gy.

Smaller dose reductions of 2 to 4 Gy were obtained by C3 plans where the corresponding C1 plan produced mean doses below 15 Gy. Several groups have investigated the NTCP dose response for parotid gland xerostomia [64, 65, 66]. Their work suggests that the incidence of xerostomia can be significantly decreased if the mean dose delivered to at least one parotid gland is kept below 26 Gy. Additionally, further dose reductions below 26 Gy also imply decreased xerostomia NTCP.

Therefore, all parotid gland dose reductions achieved by C2 and C3 plans may be expected to improve post radiotherapy salivary function to some extent. Further, there is likely heterogeneity between individual's dose responses due to genetic make-up, whether the patient also has adjuvant/concurrent chemotherapy and other factors.

Additionally, uncertainty in the NTCP models adds to the uncertainty of an individual

patient response. These are all further reasons to reduce parotid dose to the greatest feasible extent for each patient.

#### **4.1.4. Advantages and disadvantages of the complexity level system**

One of the motivating factors behind this study was to determine whether the plan quality currently delivered clinically at the WBCC could be improved upon. For this reason, it was necessary to compare newly generated VMAT plans to the clinical plans. The three tier complexity level system used to generate the plans allowed for plan quality improvements to be studied for individual patients. This was important as the purpose for creating further plans remained clinically relevant (i.e. decreasing NTCP for individual patients). Furthermore, by creating multiple plans for individual patients, the effect of plan complexity on the quality of the plan could be identified without the added variability due to differences in patient anatomy and the definition of planning volumes. A disadvantage of the planning approach was that the long time required to generate new plans essentially limited the number of cases that were able to be studied within the allocated time frame. Additionally, the method used to generate C3 plans did not allow for true pareto-optimal plans to be found. With more time, multiple plans with varied sparing of each individual OAR could have been created so the trade-offs of sparing different OARs could be observed. By more fully exploring the trade-offs, it could have been stated with greater certainty that a given C3 plan was pareto-optimal.

## **4.2. Plan complexity**

In this section each aspect of plan complexity studied in this thesis will be related to the possible impact they may have on the dosimetric accuracy of the TPS dose calculation and delivery on the LINAC. Additionally, the relationships between the

three calculated complexity metrics and the plan MU will be discussed as a potential tool to be used during optimisation. The limitations of the complexity level planning system used to generate the plans will be discussed.

Plan MU was increased when optimisation constraints were tightened to improve plan quality (section 3.2.1). If plan quality is related to plan complexity then the plan MU seems to give some indication of complexity. However, the plan MU by itself does not directly provide information about the plan's fundamental VMAT parameters such as MLC motion, field size, dose rate and gantry speed. Furthermore, the three complexity levels served only as a qualitative complexity measure, but continuous quantitative complexity measures would be required to eventually relate plan complexity with plan dosimetric accuracy. Therefore, complexity metrics were calculated to quantify plan complexity in relation to the plan parameters. Additionally, each plan's gantry rotation speed and dose rate delivery were characterised by plotting both for all segments for each beam.

### **Statistical significance**

The statistical significance of the differences observed in plan complexity between the three complexity levels could not be proven for the sample size used in this study. However, some of the complexity metrics trends with complexity level were large in magnitude and consistently observed for all cases. It is estimated that statistical significance can be proven for the larger trends with three more cases that display similar results to the five cases presented in this thesis.

#### **4.2.1 MU & Gantry speed**

The results (figure 3.5) show that for the five cases in this study, improvements to plan quality made with C2 and C3 plans also consistently increased the MU utilised

by the plans. This result generally agrees with a previous study that looked at the trade-offs between achievable plan quality and required plan MU for static gantry IMRT [38]. A consequence of high total plan MU was that the MU delivered for each segment was also high. This meant that on average, the C2 and C3 plans required higher dose rates than the C1 plans during delivery on the LINAC. However, none of the segments in any of the plans required delivered dose rates that were greater than the maximum LINAC limit of 600 MU/min. Therefore, gantry speed modulation was not required to deliver the planned MU weight for any of the segments of any of the plans. The advantage of not requiring gantry speed modulation was that the treatment delivery time was identical for all plans. Additionally, delivery error potentially arising from modulating the speed of the heavy LINAC head was avoided.

#### **4.2.2. Field size**

For the five cases in this study, the results (figure 3.5) show that improvements to plan quality made with C2 and C3 plans required the use of smaller MLC defined field sizes. The ALPO metric showed that leaf pair openings were on smaller on average for C2 and C3 plans. More importantly, C2 and C3 plans used a higher proportion of leaf pair openings less than 1.0 cm in size indicated by the CLPO metric. This is important for the accuracy of TPS dose calculation because it has previously been shown that the largest dose calculation errors using AAA occur for MLC defined field sizes that are smaller than 1.0x1.0 cm. Kron *et al.* [67] showed that for 1.0x1.0 cm fields, the AAA algorithm could under predict the delivered dose by anywhere from ~1% to ~7% depending on the selection of MLC transmission, dosimetric leaf gap and focal spot beam model parameters. The WBCC TrueBeam LINAC beam model has not yet been optimised for the calculation of small fields. Therefore, it is likely that for segments containing field sizes 1.0x1.0 cm or less there

are systematic errors ranging from ~1% to ~7% introduced into the AAA dose calculation. Improvements to the WBCC beam model for small fields could be made by adjusting the beam model parameters to more closely match the dose calculation to measurements made with a high spatial resolution detector.

### **4.2.3. Leaf travel**

For the five cases in this study, the results (figure 3.5) show that the C2 and C3 plans on average had higher LT than the corresponding C1 plans. For the same gantry rotation speed, higher LT implies higher leaf speeds on average during delivery. It has been previously reported that the magnitude of random dynamic MLC position errors can increase with leaf speed [68]. Data from LINAC delivery log files showed that the maximal MLC leaf positioning error was less than 0.6 mm for all C1, C2 and C3 plans regardless of the differences in LT between complexity levels. As the magnitude of random MLC positioning errors were similar for all complexity levels, it is likely that the dosimetric impact of random MLC errors was also similar for all complexity levels. Regardless, it has previously been shown that random MLC position errors of up to 1.2 mm have negligible impact on the dose distribution [69]. Masi *et al.* previously showed that increased LT can correlate with poor dose verification results for plans calculated with 4° control point spacing [70]. However, the strength of the observed correlation was reduced when they recalculated the plans with a higher resolution 2°/3° control point spacing. In the RapidArc implementation of VMAT, plans are calculated with a control point spacing of approximately 2°, therefore the dose calculation inaccuracies observed by Masi *et al.* were unlikely to be replicated by the WBCC VMAT setup.

#### **4.2.4. Complexity variation with case**

Within a complexity level, all complexity metrics displayed considerable variation from case to case (figure 3.5). The variation possibly indicates that the complexity of a plan is related to the patient's anatomy and the geometry of the planning volumes. This demonstrates that using a MU constraint may not be ideal because some patients may require increased plan complexity to achieve the same level of plan quality as other patients. Plan complexity variation for different anatomy and planning volume geometry was similarly demonstrated for static gantry IMRT by McNiven *et al.* [36] where large variation in utilised MU and their calculated complexity metric was reported for various treatment sites. Plan complexity variation was also likely caused by the limitations of the complexity level planning system which is further explained in section 4.2.8.

#### **4.2.5. Differences in complexity between CCW & CW beams**

Differences were observed in the complexity metrics between the CCW and CW beams. CW beams were generally more complex than the corresponding CCW beam (fig 3.6). This may indicate that each beam delivers dose in different ways. For example, for one case it was observed that the CCW beam was more conformal covering just the PTV while the CW was more modulated completing the PTV dose delivery. In another case, one beam mostly delivered dose to one half of the PTV66 while the second beam covered the remaining half. It is currently uncertain how the complexity of each beam is determined and more work is required to further understand these observations.

#### **4.2.6. Relationships between ALPO/CLPO/LT and MU**

Complexity metrics have previously been related to the degradation of the agreement between TPS calculated dose and dose measured with QA verification methods [36, 70, 71]. Hypothetically, if the dosimetric agreement between calculated and measured dose was found to degrade with increased plan complexity, it would be beneficial to relate poor dose verification results with one or more particular treatment plan parameters quantified by complexity metrics. With knowledge of what plan parameters result in degradation of deliverability, steps could be taken to ensure that plans with those kinds of complexities are avoided during optimisation. Ideally, the important complexity metrics would be available to the planner in real time so the plan complexity could be effectively monitored and controlled during optimisation. However, with the current version of the optimiser (PRO version 11.0.31), only the MU used by the current iteration of the plan is available to the planner. Therefore, for the current version of the optimiser, plan complexity can only be inferred from the reported plan MU. Trends were observed between the three calculated metrics (ALPO, CLPO, LT) and plan MU, but none were strongly correlated (fig 3.7-3.9). As none of the correlations were strong, the application of such correlations to control plan complexity during optimisation is limited. However, in the case of the WBCC VMAT system, no degradation of dosimetric accuracy with increased plan complexity was observed for the five cases in this study (section 3.3). Therefore, it appears that an upper limit on plan complexity is not required to ensure dosimetric accuracy for the WBCC VMAT system. This will be discussed in more detail in section 4.3.

#### **4.2.7. TPS version 11 upgrade**

The eclipse TPS was unavoidably upgraded from version 10.0.28 to 10.0.31 half way through this study. As such, the impact of the eclipse v11 upgrade on plan

complexity was assessed (section 3.2.4). It was important to assess how the upgrade had changed the VMAT optimiser because all C1 plans were generated with the old v10 optimiser while the majority of C2 and C3 plans were generated with the v11 optimiser. Differences between the optimiser versions may have changed the complexity of the plans irrespective of plan quality. This would potentially increase the uncertainty of the observed trends between the complexity metrics and complexity level (figure 3.5). The primary difference between optimiser versions was observed for the LT metric where the v11 optimiser sometimes would generate plans with higher LT. However, changes in LT were also observed with re-optimisations of v10 plans using the same v10 optimiser. Therefore, it is not possible to attribute this change to the optimiser version alone because the process of re-optimisation also changes the LT. However, as a result, there may be increased uncertainty surrounding the change in LT from C1 to C2 plans. The only consistently observed difference between optimisation versions was that the v11 plans had more discrete changes in dose rate as compared to the continuous dose rate changes exhibited by v10 plans (figure 3.10). However, it is not likely that this difference in planned dose rate would have changed the accuracy of dose delivery by the LINAC. Differences between optimiser versions for ALPO, CLPO and MU were negligible and would not likely influence the observed trends between complexity levels.

#### **4.2.8. Limitations of the complexity level planning system**

The complexity level planning system allowed for a range of plan complexity to be generated and related with plan quality. However, the system also created overlap in the complexity metrics between the complexity levels. For example, the highest complexity C1 plans were more complex than the lowest complexity C2 plans in terms of the complexity metrics. The overlap was partially a result of the variability in

plan complexity for different patient anatomy and planning volume geometry, but also partially due to the complexity level system itself. The primary limitation of the system was that it assumed that the C3 plans were pareto-optimal. As the quality of the C2 plans were defined to be half way between the C1 and C3 plans, it was important that the C3 plans were generated in a consistent way. By not adequately controlling the plan quality in a consistent way for each case, plan complexity variability may have been increased within each complexity level.

#### **4.2.9. Other considerations**

A potential downside to increased MU utilisation with higher complexity plans is that the patient can be exposed to higher out of field dose. A study by Hall, looking at the differences between 3DCRT and IMRT has suggested that increased out of field patient exposure due to LINAC head leakage and MLC transmission radiation are factors that may increase the probability of inducing secondary cancers [27]. Hall concluded that IMRT may double the probability of induced secondary cancers relative to 3DCRT (1.5 % to 3.0 % 10 years after radiotherapy). In this study, relative to the C1 plans, C3 plans increased plan MU by 55% on average. An increase in plan MU will increase the out of field dose from radiation transmitted through the MLC leaves. However, the C3 plans would not increase patient exposure to leakage radiation from the LINAC head because the delivery times for all complexity plans were identical (4.2.1.). The differences between 3DCRT and IMRT delivery times and MU usage considered by Hall are much greater than differences between high and low complexity plans in this study. Therefore, the probability of inducing secondary cancer with higher complexity VMAT plans is not likely to be much different than with lower complexity VMAT plans. Even so, this should be kept under consideration

and steps to reduce transmission through the MLC and leakage from the LINAC head should be taken where possible to minimise the patient's out of field exposure.

### **4.3. Plan dosimetry verification**

In this section, the results of the dose verification measurements will be discussed and the strengths and weaknesses of each verification method will be highlighted. The results of the dose verification methods will also be discussed in relation to plan complexity and plan quality.

The overall accuracy of VMAT treatment depends on multiple factors including patient positioning, accuracy of the TPS dose calculation and the accuracy of dose delivery by the LINAC. In this project I have studied the agreement between the TPS calculated dose and LINAC delivered dose (dosimetric accuracy) for a variety of plan complexity. Dosimetric accuracy is important because planning decisions are made using the calculated dose and for the most part, the calculated dose is expected to be reproduced in reality. With increased plan complexity, certain plan parameters may degrade the accuracy of the dose calculation and/or the dose delivery. For this reason, end-to-end dose verification measurements were performed to assess the dosimetric accuracy for C1, C2 and C3 plans. The verification results were averaged for each complexity level to identify whether dosimetric accuracy was changed with plan complexity. Three independent verification methods were used to form a more complete picture of dosimetric accuracy. Each method has its strengths and weakness and only by considering the results of all three methods together can the question of a plan's dosimetric accuracy be answered with confidence.

### **4.3.1. LINAC dose delivery repeatability**

The Truebeam LINAC at the WBCC was able to deliver integral dose to a small volume with repeatability better than 0.1% for both C1 and C3 plans (table 3.5). Additionally, segment dose delivery repeatability was better than 0.6% for both C1 and C3 plans (table 3.5). These results suggest that the increased plan complexity of the C3 plan did not negatively impact on the repeatability of the dose delivery. It is noteworthy that the stated repeatability figures also include uncertainty of the measurement process including  $K_{tp}$  correction of each run and leakage correction. The advantage of using time resolved point dose measurements was that the dose delivery could be assessed for all segments rather than just the integral dose. Additionally, the point dose method gave adequately stable measurements so the repeatability of the LINAC dose delivery itself was resolved. This type of analysis gave confidence that the measured integral dose was not a result of large random errors cancelling out. Rather, the dose delivery at every segment was found to be highly repeatable, resulting in a highly repeatable integral dose delivery (fig 3.11 & 3.12). The high repeatability of segment dose delivery may indicate that the LINAC was operating within its limitations for both high and low plan complexity plans.

A disadvantage of the point dose method was that the repeatability was only assessed at one small volume in the phantom. It is possible that the dose delivery repeatability was different in other locations and missed by this experiment. Film or ArcCheck also could have been used for this experiment, but the uncertainties associated with these two methods were too large and LINAC dose delivery repeatability would not be resolved beneath the uncertainty of the measurement method itself.

### 4.3.2. Impact of dose grid resolution

Reducing the dose grid resolution from 1.5 mm to 1.0 mm was found to decrease the calculated volume doses by around 1 % (section 3.3.2). The calculated dose was reduced by slightly more for C3 plans than for C1 plans indicating that this effect may be larger for more highly modulated beams. A 1.0 mm dose grid resolution was selected for dose verification analysis as it provided the most accurate dose calculation. This has similarly been observed by Ong *et al.* [40] for AAA v10, where it was recommended that a 1.0 mm dose grid resolution should be used for higher complexity plans due to dose calculation inaccuracies when calculating with lower dose grid resolutions. Additionally, Ong *et al.* reported that the dosimetric inaccuracies with were larger in low density media. Therefore for high complexity plans, it is recommended that a 1.0 mm dose grid resolution is also used for the dose calculation on the patient's anatomy to maximise the dosimetric accuracy for the assessment of plan quality. A practical downside to calculating at a 1.0 mm dose grid resolution was that calculation times were doubled to 60-90 minutes compared to 45 minutes with a 1.5 mm dose grid resolution due to the increased computational overhead. This issue may be resolved with the use of higher specification computer hardware and by porting the dose calculation from the central processing unit to the graphics processing unit [72].

### **4.3.3. Point dose verification**

#### **Segment point dose analysis**

Analysis of the dose deviation per segment as described by Louwe *et al.* [48] did not reveal significant systematic differences between the three complexity levels.

The only deviations that were observed above the background noise level were for segments where the detector was located near the beam penumbra (figure 3.13). For our department, the origin and magnitude of these deviations have been well characterised [48]. Louwe *et al.* found that these deviations were primarily caused by inaccuracies in the modelling of the beam penumbra. This project's data set was too small to make further analysis of these deviations meaningful. Previous experience within the department showed that at least 50-100 time-resolved point dose measurements are required to resolve the smallest systematic deviations from the random noise resulting from inaccuracies in detector positioning.

#### **Integral point dose analysis**

No significant trends were observed in the integral point dose deviations between the three complexity levels. Additionally, averaged point dose deviations for all three complexity levels were very close to 0.0% (table 3.6) suggesting that for the five cases in this study there were no large systematic differences between the delivered and calculated doses for any of the complexity levels.

From the three methods used in this study, the point dose method provided the most accurate measurement of absolute dose. This is because for VMAT dose measurement, the characteristics of the pinpoint ionisation chambers were well understood and the dose calibration factor was traceable back to the Australasian primary standard (Australian Radiation Protection and Nuclear Safety Agency,

ARPANSA). However, individual point dose measurements were very sensitive to detector positioning where the sensitivity was mostly dependent on the steepness of dose gradients surrounding the detector. The effect of this can be seen in the results where the measurement variability for OAR points measured on steep dose gradients was higher than for PTV points measured on relatively shallow dose gradients.

Additionally, the measurement uncertainty for OAR points was higher for C2 and C3 plans than for C1 plans (table 3.6) possibly as a result of the steeper dose gradients generated by higher complexity plans. Where possible, measurement points were selected on relatively shallow dose gradients to minimise the measurement's sensitivity to detector positioning. Additionally, multiple points located at the same position were measured sequentially without moving the phantom to reduce measurement variation introduced from resetting the phantom.

Dose gradients also had an effect on the chamber measurement itself through volume averaging [53, 54]. As the differences in dosimetric accuracy between complexity levels may have been very small it was important to minimise uncertainty in the analysis method where possible. The active volume of the chamber was 5 mm long on the longitudinal axis of the chamber (table 2.5). As a result, dose gradients along this axis of the chamber introduced noticeable volume averaging into the measurements. Initially, a single voxel was used to extract the TPS calculated dose at each measurement point. However, it was found with dose gradients crossing the measurement point, the selection of the voxel had a large impact on the calculation of the dose deviation. It was realised that the measurement of dose with the pinpoint chamber was made in a small volume. Therefore, to account for volume averaging, the calculated point dose was taken as the mean dose delivered to voxels contained inside a contour drawn with the same dimensions as the chamber's active volume.

Essentially this selection of the calculated dose modelled how the chamber was actually measuring dose and the uncertainty of the point dose deviation calculation was reduced because a single voxel was no longer used as the comparison dose. This method was selected over gamma analysis so dose deviations could be averaged for each complexity level.

#### **4.3.4. Film dose verification**

For the five cases in this study, no significant trends were observed in the film gamma pass-rates and mean gamma values between the three complexity levels (table 3.7). The results indicate that the higher complexity plans did not significantly degrade dosimetric accuracy for the coronal dose planes selected in the PTV and OAR regions. Additionally, qualitative analysis using dose profiles showed that the expected plan quality improvements observed with the TPS calculated dose were accurately replicated with LINAC delivery.

Film measurements extended dose measurements to coronal planes so the spatial accuracy of the delivered dose distribution could be verified in addition to the absolute dose of the delivered dose. The absolute dose produced by film was verified by selecting measurement planes that also contained point dose measurements.

Absolute dose between film and point dose dosimetry agreed within 2% for all planes.

It is important to note the comparison between film and point dose was made using single film voxel doses rather than a volume similar to the method explained in section 2.3.1. Small dose deviations existed at various locations on all films but did not appear to increase in frequency or magnitude with increased plan complexity (figure 3.15). The overall dosimetric accuracy of each film plane was quantified with gamma analysis. For gamma analysis, the  $\Delta D = 2\%$  criterion was selected to be consistent with point dose analysis where absolute dose deviations of up to 2% was

within clinically acceptable levels. DTA = 2 mm was selected because generally a positional deviation of up to 2 mm was also considered clinically acceptable. However, for plans where the 50 Gy isodose was within 2 mm of the SC or BS, this DTA criterion may have been too lenient because a positional deviation of 1 mm to 2 mm of the 50 Gy isodose near the SC/BS would be considered important.

#### **4.3.5. ArcCheck dose verification**

For the five cases in this study, no significant trends were observed in the ArcCheck gamma pass-rates between the three complexity levels (table 3.8). The results indicate that the higher complexity plans did not significantly degrade dosimetric accuracy. The ArcCheck phantom enables analysis of VMAT dosimetric accuracy through measurement of the entrance and exit dose with diodes spaced around the cylindrical phantom. The integral measured dose is a sum of all entrance and exit dose contributions from each delivered segment. Gamma analysis comparison is performed between the integral dose measured by each diode and the TPS calculated dose. The ArcCheck is a practical method to quickly perform dose verification of the entire dose delivery. Additionally, the ArcCheck provides a test of the TPS dose calculation accuracy for a phantom material other than water (PMMA) which may be important for the dose calculation on the patient anatomy because structures of various densities exist in the H&N. However due to the ArcCheck's poor spatial resolution between diodes (1.0 cm spacing), its utility to detect differences in the dosimetric accuracy between complexity levels was limited to large dosimetric deviations. In this sense, the ArcCheck has provided data that shows that there is no large degradation of the dosimetric accuracy for higher complexity plans but cannot resolve smaller differences between complexity levels that may also exist. This lack of sensitivity to

dosimetric deviations has also been demonstrated for other commonly used detector array QC phantoms [73].

#### **4.3.6. VMAT insensitivity to errors**

Three independent dose verification methods have indicated that dosimetric accuracy was not degraded with increased plan complexity. This result initially seems counterintuitive given that large systematic dosimetric deviations originating from inaccuracies in AAA were expected to exist for segments containing LPOs smaller than 1.0x1.0 cm as discussed in section 4.2.2. Additionally, other systematic deviations may also exist that have not yet been identified. Because these deviations have not had a large effect on the integral dose, it is implied that for VMAT, the integral dose may be insensitive to certain types of dosimetric deviations. This insensitivity may be because the systematic segment deviations become diluted due to geometry of VMAT delivery itself. For example, consider a systematic deviation between the LINAC delivered and TPS calculated dose resulting from the inadequate TPS modelling of the MLC inter-leaf leakage. In this scenario, the voxels that are irradiated with inter-leaf leakage during delivery will accumulate systematic deviations. For a collimator rotation of 0 degrees, voxels remain stationary in the beam's eye view during gantry rotation and the same voxels are always irradiated with inter-leaf leakage. Therefore, the systematic deviations from each of the segments add up in the integral dose. However for a rotated collimator angle, voxels rotate in ellipses in the beam's eye view during gantry rotation. Therefore, for a rotated collimator angle many different voxels are exposed to inter-leaf leakage radiation and the systematic deviations from each segment no longer adds up to become significant in the integral dose. With this kind of mechanism, the segment systematic deviations can be averaged out of the integral dose.

A similar argument can be applied to the systematic deviations originating from the TPS calculation of small fields. In this case, the systematic deviations only become large for fields below a certain threshold size, so these deviations are transient in nature. Further, small fields can be located in many locations in the beams eye view and affect many different regions of voxels. Finally, as described previously, the collimator rotation results in voxels moving in ellipses in the beam's eye view as the gantry is rotated. As such, the systematic deviations arising from small fields would not have the opportunity to add up for any particular group of voxels during the delivery of the arc. The remaining fluence delivered through large fields is correctly calculated by AAA and do not contribute large systematic deviations. For plans with a combination of large and small fields, the systematic deviations arising from small fields make up a negligible contribution to the overall delivered fluence. This effect would not hold for regions of voxels that acquire the majority of their dose from fluence delivered through small fields.

VMAT's insensitivity to various systematic and random delivery errors has also been demonstrated by Betzel *et al.* [74].

## 5. Conclusion

VMAT is the latest technological advance in LINAC based radiotherapy that improves the quality of treatment plans used to treat H&N cancer patients. While the VMAT delivery modality itself has improved the quality of H&N treatment plans, further steps can be taken to ensure that the highest quality treatment plans are generated for each patient. This study has shown that the quality of a given VMAT plan is highly dependent on how far a plan is pushed towards a pareto-optimal solution during optimisation. It has been demonstrated that the quality of VMAT plans can be improved upon by not using a MU objective that restricts plan complexity and by aggressively applying DVOs to planning volumes during optimisation. For five cases, higher complexity plans have been generated that improve plan quality beyond what is achievable with the planning protocol currently used in WBCC clinical practice. Characteristics of plan complexity have been quantified with metrics and show that higher complexity plans on average utilise smaller field sizes, higher leaf travel and higher monitor units. The results from three independent dose verification methods have indicated that dosimetric accuracy was not degraded by the higher complexity plans compared to the clinical WBCC plans. These results give confidence that VMAT plans more complex than those currently used clinically at the WBCC can be accurately calculated and delivered within clinically acceptable limits. H&N VMAT at the WBCC can now be developed further with greater confidence in the dosimetric accuracy of higher complexity plans.

# Bibliography

- [1] Argiris, Athanassios et al. 'Head And Neck Cancer'. *The Lancet* 371.9625 (2008): 1695-1709.
- [2] Howlader, N et al. *SEER Cancer Statistics Review, 1975-2011*. Bethesda, MD: National Cancer Institute, 2015.
- [3] Argiris, Athanassios, and Cathy Eng. 'Epidemiology, Staging, And Screening Of Head And Neck Cancer'. *Cancer Treatment and Research* 114 (2004): 15-60.
- [4] Saman, Daniel M. 'A Review Of The Epidemiology Of Oral And Pharyngeal Carcinoma: Update'. *Head Neck Oncol* 4.1 (2012): 1.
- [5] Agudo, A. et al. 'Impact Of Cigarette Smoking On Cancer Risk In The European Prospective Investigation Into Cancer And Nutrition Study'. *Journal of Clinical Oncology* 30.36 (2012): 4550-4557.
- [6] D'Souza, Gypsyamber et al. 'Case–Control Study Of Human Papillomavirus And Oropharyngeal Cancer'. *New England Journal of Medicine* 356.19 (2007): 1944-1956.
- [7] Curado, Maria Paula, and Mia Hashibe. 'Recent Changes In The Epidemiology Of Head And Neck Cancer'. *Current Opinion in Oncology* 21.3 (2009): 194-200.
- [8] Marur, Shanthi, and Arlene A. Forastiere. 'Head And Neck Cancer: Changing Epidemiology, Diagnosis, And Treatment'. *Mayo Clinic Proceedings* 83.4 (2008): 489-501.
- [9] National Head and Neck Cancer Tumour Standards Working Group,. *Standards Of Service Provision For Head And Neck Cancer Patients In New Zealand - Provisional..* Wellington: Ministry of Health, 2013.
- [10] Peters, Lester J et al. 'Evaluation Of The Dose For Postoperative Radiation Therapy Of Head And Neck Cancer: First Report Of A Prospective Randomized Trial'. *International Journal of Radiation Oncology\*Biophysics* 26.1 (1993): 3-11.
- [11] Fiorica, F. 'Preoperative Chemoradiotherapy For Oesophageal Cancer: A Systematic Review And Meta-Analysis'. *Gut* 53.7 (2004): 925-930.
- [12] Webb, S, and A E Nahum. 'A Model For Calculating Tumour Control Probability In Radiotherapy Including The Effects Of Inhomogeneous Distributions Of Dose And Clonogenic Cell Density'. *Physics in Medicine and Biology* 38.6 (1993): 653-666.
- [13] Marks, Lawrence B. et al. 'Use Of Normal Tissue Complication Probability Models In The Clinic'. *International Journal of Radiation Oncology\*Biophysics* 76.3 (2010): S10-S19.
- [14] Cooper, Jay S. et al. 'Late Effects Of Radiation Therapy In The Head And Neck Region'. *International Journal of Radiation Oncology\*Biophysics* 31.5 (1995): 1141-1164.

- [15] Schultheiss, T.E. et al. 'Radiation Response Of The Central Nervous System'. *International Journal of Radiation Oncology\*Biology\*Physics* 31.5 (1995): 1093-1112.
- [16] Withers, H. Rodney. 'Biologic Basis For Altered Fractionation Schemes'. *Cancer* 55.S9 (1985): 2086-2095.
- [17] Leibel, Steven A. et al. 'Improved Dose Distributions For 3D Conformal Boost Treatments In Carcinoma Of The Nasopharynx'. *International Journal of Radiation Oncology\*Biology\*Physics* 20.4 (1991): 823-833.
- [18] Webb, S. 'The Physical Basis Of IMRT And Inverse Planning'. *BJR* 76.910 (2003): 678-689.
- [19] Spirou, Spiridon V., and Chen-Shou Chui. 'A Gradient Inverse Planning Algorithm With Dose-Volume Constraints'. *Medical Physics* 25.3 (1998): 321.
- [20] Webb, S. 'Optimizing The Planning Of Intensity-Modulated Radiotherapy'. *Physics in Medicine and Biology* 39.12 (1994): 2229-2246.
- [21] Shepard, D. M. et al. 'Direct Aperture Optimization: A Turnkey Solution For Step-And-Shoot IMRT'. *Medical Physics* 29.6 (2002): 1007.
- [22] Longobardi, Barbara et al. 'Comparing 3DCRT And Inversely Optimized IMRT Planning For Head And Neck Cancer: Equivalence Between Step-And-Shoot And Sliding Window Techniques'. *Radiotherapy and Oncology* 77.2 (2005): 148-156.
- [23] Cozzi, Luca et al. 'Three-Dimensional Conformal Vs. Intensity-Modulated Radiotherapy In Head-And-Neck Cancer Patients: Comparative Analysis Of Dosimetric And Technical Parameters'. *International Journal of Radiation Oncology\*Biology\*Physics* 58.2 (2004): 617-624.
- [24] Vergeer, Marije R. et al. 'Intensity-Modulated Radiotherapy Reduces Radiation-Induced Morbidity And Improves Health-Related Quality Of Life: Results Of A Nonrandomized Prospective Study Using A Standardized Follow-Up Program'. *International Journal of Radiation Oncology\*Biology\*Physics* 74.1 (2009): 1-8.
- [25] Van Rij, CM et al. 'Parotid Gland Sparing IMRT For Head And Neck Cancer Improves Xerostomia Related Quality Of Life'. *Radiat Oncol* 3.1 (2008): 41.
- [26] Zhu, X. Ronald, Christopher J. Schultz, and Michael T. Gillin. 'Planning Quality And Delivery Efficiency Of Smlc Delivered IMRT Treatment Of Oropharyngeal Cancers Evaluated By RTOG H-0022 Dosimetric Criteria'. *Journal of Applied Clinical Medical Physics* 5.4 (2004): 80-95.
- [27] Hall, Eric J. 'Intensity-Modulated Radiation Therapy, Protons, And The Risk Of Second Cancers'. *International Journal of Radiation Oncology\*Biology\*Physics* 65.1 (2006): 1-7.
- [28] Otto, Karl. 'Volumetric Modulated Arc Therapy: IMRT In A Single Gantry Arc'. *Medical Physics* 35.1 (2008): 310.
- [29] Verbakel, Wilko F.A.R. et al. 'Volumetric Intensity-Modulated Arc Therapy Vs. Conventional IMRT In Head-And-Neck Cancer: A Comparative Planning And Dosimetric Study'. *International Journal of Radiation Oncology\*Biology\*Physics* 74.1 (2009): 252-259.

- [30] Holt, Andrea et al. 'Multi-Institutional Comparison Of Volumetric Modulated Arc Therapy Vs. Intensity-Modulated Radiation Therapy For Head-And-Neck Cancer: A Planning Study'. *Radiat Oncol* 8.1 (2013): 26.
- [31] Bertelsen, Anders et al. 'Single Arc Volumetric Modulated Arc Therapy Of Head And Neck Cancer'. *Radiotherapy and Oncology* 95.2 (2010): 142-148.
- [32] Hong, Theodore S. et al. 'The Impact Of Daily Setup Variations On Head-And-Neck Intensity-Modulated Radiation Therapy'. *International Journal of Radiation Oncology\*Biology\*Physics* 61.3 (2005): 779-788.
- [33] Nelms, B et al. 'MO-EE-A2-01: On The Dosimetric/DVH Impact Of Variation In Organ Delineation: A Multi-Institutional Study And Proposed Quality System'. *Medical Physics* 37.6 (2010): 3348-3348.
- [34] Mohan, Radhe et al. 'The Impact Of Fluctuations In Intensity Patterns On The Number Of Monitor Units And The Quality And Accuracy Of Intensity Modulated Radiotherapy'. *Medical Physics* 27.6 (2000): 1226.
- [35] Dai, Jianrong, and Yunping Zhu. 'Minimizing The Number Of Segments In A Delivery Sequence For Intensity-Modulated Radiation Therapy With A Multileaf Collimator'. *Medical Physics* 28.10 (2001): 2113.
- [36] McNiven, Andrea L., Michael B. Sharpe, and Thomas G. Purdie. 'A New Metric For Assessing IMRT Modulation Complexity And Plan Deliverability'. *Medical Physics* 37.2 (2010): 505.
- [37] Webb, S. 'Use Of A Quantitative Index Of Beam Modulation To Characterize Dose Conformality: Illustration By A Comparison Of Full Beamlet IMRT, Few-Segment IMRT (Fsimrt) And Conformal Unmodulated Radiotherapy'. *Physics in Medicine and Biology* 48.14 (2003): 2051-2062.
- [38] Craft, David, Philipp Süß, and Thomas Bortfeld. 'The Tradeoff Between Treatment Plan Quality And Required Number Of Monitor Units In Intensity-Modulated Radiotherapy'. *International Journal of Radiation Oncology\*Biology\*Physics* 67.5 (2007): 1596-1605.
- [39] Masi, Laura et al. 'Impact Of Plan Parameters On The Dosimetric Accuracy Of Volumetric Modulated Arc Therapy'. *Medical Physics* 40.7 (2013): 071718.
- [40] Ong, Chin Loon et al. 'Impact Of The Calculation Resolution Of AAA For Small Fields And Rapidarc Treatment Plans'. *Medical Physics* 38.8 (2011): 4471.
- [41] Sobin, Leslie H., Mary K. Gospodarowicz, and Christian Wittekind, eds. TNM classification of malignant tumours. John Wiley & Sons, 2011.
- [42] Grégoire, Vincent et al. 'CT-Based Delineation Of Lymph Node Levels And Related Ctvs In The Node-Negative Neck: DAHANCA, EORTC, GORTEC, NCIC, RTOG Consensus Guidelines'. *Radiotherapy and Oncology* 69.3 (2003): 227-236.
- [43] 'Prescribing, Recording, And Reporting Photon-Beam Intensity-Modulated Radiation Therapy (IMRT): Contents'. *Journal of the ICRU* 10.1 (2010): NP-NP.

- [44] Hoffman, Aswin. 'Treatment Planning Optimisation Methods For Individualised Dose Prescription In Intensity-Modulated Radiation Therapy'. PhD thesis, Radboud Universiteit Nijmegen, The Netherlands, 2013.
- [45] Van Esch, Ann et al. 'Testing Of The Analytical Anisotropic Algorithm For Photon Dose Calculation'. *Medical Physics* 33.11 (2006): 4130.
- [46] Fogliata, Antonella et al. 'Dosimetric Validation Of The Anisotropic Analytical Algorithm For Photon Dose Calculation: Fundamental Characterization In Water'. *Physics in Medicine and Biology* 51.6 (2006): 1421-1438.
- [47] Zygmanski, P., and J. H. Kung. 'Method Of Identifying Dynamic Multileaf Collimator Irradiation That Is Highly Sensitive To A Systematic MLC Calibration Error'. *Medical Physics* 28.11 (2001): 2220.
- [48] R.J.W. Louwe et al. 'Time-resolved dosimetry using a pinpoint ionization chamber as quality assurance for IMRT and VMAT'. *Medical Physics* (2015), in press.
- [49] PTW-Freiburg,. *Ionizing Radiation Detector Catalog*. Freiburg: PTW-Freiburg, 2006.
- [50] International Atomic Energy Agency,. *Absorbed Dose Determination In External Beam Radiotherapy*. Vienna: International Atomic Energy Agency, 2015. Print. TECHNICAL REPORTS SERIES No. 398.
- [51] Satherley, Thomas. *TANDEM Software User Guide*. 2013. Print. WBCC Internal Report.
- [52] International Organization for Standardization,. *International Vocabulary Of Metrology -- Basic And General Concepts And Associated Terms (VIM)*. Geneva: International Organization for Standardization, 2007.
- [53] Low, Daniel A. et al. 'Ionization Chamber Volume Averaging Effects In Dynamic Intensity Modulated Radiation Therapy Beams'. *Medical Physics* 30.7 (2003): 1706.
- [54] Laub, Wolfram U., and Tony Wong. 'The Volume Effect Of Detectors In The Dosimetry Of Small Fields Used In IMRT'. *Medical Physics* 30.3 (2003): 341.
- [55] Devic, Slobodan. 'Radiochromic Film Dosimetry: Past, Present, And Future'. *Physica Medica* 27.3 (2011): 122-134.
- [56] Del Moral, F. et al. 'From The Limits Of The Classical Model Of Sensitometric Curves To A Realistic Model Based On The Percolation Theory For Gafchromic™ EBT Films'. *Medical Physics* 36.9 (2009): 4015.
- [57] Marquardt, Donald W. 'An Algorithm For Least-Squares Estimation Of Nonlinear Parameters'. *Journal of the Society for Industrial and Applied Mathematics* 11.2 (1963): 431-441.
- [58] Levenberg, K. 'A Method For The Solution Of Certain Problems In Least Squares'. *Quart. Applied Math* 2 (1944): 164-168.
- [59] Low, Daniel A. et al. 'A Technique For The Quantitative Evaluation Of Dose Distributions'. *Medical Physics* 25.5 (1998): 656.

- [60] Létourneau, Daniel et al. 'Novel Dosimetric Phantom For Quality Assurance Of Volumetric Modulated Arc Therapy'. *Medical Physics* 36.5 (2009): 1813.
- [61] Sun Nuclear Corporation,. *Arccheck Reference Guide*. Melbourne, FL: Sun Nuclear Corporation, 2015.
- [62] Kirkpatrick, John P., Albert J. van der Kogel, and Timothy E. Schultheiss. 'Radiation Dose–Volume Effects In The Spinal Cord'. *International Journal of Radiation Oncology\*Biology\*Physics* 76.3 (2010): S42-S49.
- [63] Mayo, Charles, Ellen Yorke, and Thomas E. Merchant. 'Radiation Associated Brainstem Injury'. *International Journal of Radiation Oncology\*Biology\*Physics* 76.3 (2010): S36-S41.
- [64] Dijkema, Tim. 'Salivary Gland Sparing Radiotherapy'. PhD thesis, Universiteit Utrecht, The Netherlands 2013.
- [65] Eisbruch, Avraham et al. 'Dose, Volume, And Function Relationships In Parotid Salivary Glands Following Conformal And Intensity-Modulated Irradiation Of Head And Neck Cancer'. *International Journal of Radiation Oncology\*Biology\*Physics* 45.3 (1999): 577-587.
- [66] BLANCO, A et al. 'Dose-Volume Modeling Of Salivary Function In Patients With Head And Neck Cancer Receiving Radiation Therapy'. *International Journal of Radiation OncologyBiologyPhysics* 60 (2004): S189-S189.
- [67] Kron, T. et al. 'Small Field Segments Surrounded By Large Areas Only Shielded By A Multileaf Collimator: Comparison Of Experiments And Dose Calculation'. *Medical Physics* 39.12 (2012): 7480.
- [68] Ling, C. Clifton et al. 'Commissioning And Quality Assurance Of Rapidarc Radiotherapy Delivery System'. *International Journal of Radiation Oncology\*Biology\*Physics* 72.2 (2008): 575-581.
- [69] Wijesooriya, Krishni et al. 'Rapidarc Patient Specific Mechanical Delivery Accuracy Under Extreme Mechanical Limits Using Linac Log Files'. *Medical Physics* 39.4 (2012): 1846.
- [70] Masi, L. et al. 'Impact Of Control Point Angular Separation On The Dosimetric Accuracy Of VMAT Plans'. *International Journal of Radiation Oncology\*Biology\*Physics* 84.3 (2012): S862.
- [71] Miura, Hideharu et al. 'DICOM-RT Plan Complexity Verification For Volumetric Modulated Arc Therapy'. *IJMPCCERO* 03.03 (2014): 117-124.
- [72] Hissoiny, Sami, Benoît Ozell, and Philippe Després. 'A Convolution-Superposition Dose Calculation Engine For Gpus'. *Medical Physics* 37.3 (2010): 1029.
- [73] Fredh, Anna et al. 'Patient QA Systems For Rotational Radiation Therapy: A Comparative Experimental Study With Intentional Errors'. *Medical Physics* 40.3 (2013): 031716.
- [74] Betzel, Gregory T. et al. 'Is Rapidarc More Susceptible To Delivery Uncertainties Than Dynamic IMRT?'. *Medical Physics* 39.10 (2012): 5882.