Evaluation of Dysphagia-Optimising Radiotherapy Techniques in Locally Advanced Oropharyngeal Cancer

Dr Imran Abdulla Petkar

The Institute of Cancer Research and

Royal Marsden Hospital NHS Foundation Trust

A Thesis submitted to the University of London for the Degree of Doctor of Medicine (MD)

Author's Declaration

I declare as the sole author of this thesis that the data presented here represent my personal research conducted whilst a clinical research fellow at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust, UK between September 2015 and August 2017. Data generated in collaboration with other colleagues are fully acknowledged in the text.

Imran Petkar

Abstract

Radiation-associated dysphagia (RAD) following primary chemoradiation in locally advanced oropharyngeal cancers (LA-OPC) can have a devastating impact on patients' quality of life (QoL). Establishing efficient swallow-sparing radiotherapy (RT) techniques is, therefore, of paramount importance in an era where health-related QoL measures are increasingly influential determinants of curative management strategies.

Dysphagia-optimised intensity-modulated RT (Do-IMRT) is a novel planning technique that limits dose delivered to the pharyngeal constrictor muscles (PCM), a key swallowing structure implicated in RAD. A retrospective comparison planning study is presented which specifically investigated the swallow-sparing benefits of Do-IMRT over standard IMRT (S-IMRT) in LA-OPC. It is demonstrated that Do-IMRT significantly reduces the probability of persistent swallowing dysfunction, without compromising on dose to high-dose target volume or organs at risk (OAR). Prospective validation of the benefits of Do-IMRT is currently under investigation within the context of the Dysphagia-at-risk structures (DARS) trial, a national phase 3 randomised study.

Intensity-modulated proton therapy (IMPT), with its sharp dose fall-off, holds great promise as a toxicity-mitigating strategy. The benefits of IMPT in reducing RAD in LA-OPC relative to IMRT, however, remain investigational. The role of IMPT, using different beam arrangements and optimisation

techniques, in reducing the estimated risk of RAD was evaluated in a retrospective comparison planning study. It is shown that robustly optimised dysphagia-optimised IMPT (Do-IMPT$_{RO}$) could improve long-term swallowing function in selected patients with LA-OPC, compared to Do-IMRT. It is also demonstrated that the robustness of Do-IMPT$_{RO}$ plans are not affected in the presence of range and set up uncertainties, unlike conventional planning target volume (PTV) – based Do-IMPT optimisation technique.

The perceived advantage of dysphagia-optimising RT techniques is contingent on contouring accuracy of PCM, an OAR not delineated routinely in the United Kingdom (UK). Heterogeneity in delineation between oncologists may lead to differences in the reported dose-volume parameters, and this can have implications on subsequent toxicity outcomes. Inter – observer variability (IOV) in PCM contouring in the UK and subsequent potential impact on functional outcomes was studied within the context of a pre-trial RT quality assurance programme for DARS study. It is shown that there is IOV in the delineation of PCM amongst oncologists, but the impact of the variability on dose delivered to this structure was not significantly impacted upon in the pre-trial benchmark case.

Table of contents

List of tables

List of figures

Publications arising from thesis

Petkar, I., et al., *DARS: a phase III randomised multicentre study of dysphagia-optimised intensity- modulated radiotherapy (Do-IMRT) versus standard intensity- modulated radiotherapy (S-IMRT) in head and neck cancer.* BMC Cancer, 2016. **16**(1): p. 770

Petkar, I., et al., *Dysphagia-optimised Intensity-modulated Radiotherapy Techniques in Pharyngeal Cancers: Is Anyone Going to Swallow it?* Clin Oncol (R Coll Radiol), 2017. **29**(7): p. e110 -e118

Acknowledgements

The work in this thesis was carried out with the support of the Oracle Cancer Trust.

I would like to thank my supervisors, Professor Chris Nutting and Professor Kevin Harrington, together with their colleagues Dr Kate Newbold, and Dr Shree Bhide, for providing me with the opportunity to undertake this project and for their invaluable guidance throughout my time as a clinical research fellow. I owe a significant debt of gratitude to my colleagues in Radiotherapy Physics, Alex Dunlop, and Dualta McQuaid, for teaching me the technical nuances of radiotherapy planning, and for generating numerous scripts that were necessary for this thesis. This work would not have been possible without their great patience, dedication and invaluable assistance. I also wish to thank Sarah Gulliford for her advice and help with toxicity model generation.

I would particularly like to thank Dr Robert Malyapa for inviting me to the Maryland Proton Treatment Center, USA to share his proton beam therapy experience in head and neck cancers, and the dosimetry team there as well for providing invaluable tips on proton planning. Thank you also to Dr Richard Amos for his guidance on proton planning in the UK.

Finally, I would like to thank my family. I am truly grateful for the unwavering support of my wife Faizah throughout the peaks and troughs of my MD. A big

thank you to my children, Raihan and Aisha, who showed maturity far beyond their years particularly when family holidays were intertwined with sessions of thesis writing. Last but not the least, I would like to thank my parents Abdulla and Safia Petkar for the sacrifices they made to help me achieve my goals.

Abbreviations

3-D	3-dimensional
95 % CI	95 % Confidence Interval
AEF	Aryepiglottic fold
AJCC	American Joint Committee on Cancer
ANOVA	Analysis of variances
BoT	Base of tongue
BP	Bragg peak
BS	Brainstem
C-I	Concordance index
CE	Cervical oesophagus
CI	Conformity index
CL	Conformity level
CPI	Cricoid pharyngeal inlet
CRT	Chemoradiation
CRUK	Cancer Research United Kingdom
CT	Computed tomography
CTCAE	Common toxicity criteria for adverse events
CTV	Clinical target volume
DARS	Dysphagia/aspiration-related structures
DI	Discordance index
DICOM	Digital imaging and communications in medicine
Do-IMPT	Dysphagia-optimised Intensity Modulated Proton

	Therapy
Do-IMRT	Dysphagia-optimised Intensity Modulated Radiotherapy
DSC	Dice Similarity Coefficient
DTA	Distance to agreement
DVH	Dose-volume histogram
EORTC QLQ-H&N35	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Head and Neck Module
EQD2	Equivalent dose in 2 Gy fraction
FT	Feeding tube
G-tube	Gastrostomy tube
GMI	Geographical Miss Index
GS	Gold standard
GSL	Glottic-supraglottic larynx
GTV	Gross tumour volume
HI	Homogeneity index
HN-QOL	Head and neck quality of life
HNC	Head and neck cancers
HNCI	Head and neck cancer inventory
HPC	Hypopharyngeal cancers
HPV	Human papillomavirus
HR-QoL	Health-related quality of life
HU	Hounsfield units
IC	Induction chemotherapy
ICC	Intraclass correlation coefficient

IMPT	Intensity Modulated Proton Therapy
IMRT	Intensity Modulated Radiotherapy
IOV	Inter-observer variability
IPC	Inferior Pharyngeal Constrictor
IQR	Inter-quartile range
LA-OPC	Locally-advanced oropharyngeal cancer
LN	Lymph node
LRC	Locoregional control
MBS	Modified barium swallow
MDADI	MD Anderson dysphagia inventory
MFO	Multi-field optimisation
MHM	Myelo-geniohyoid complex
mo	months
MPC	Middle Pharyngeal Constrictor
MRI	Magnetic resonance imaging
NPC	Nasopharyngeal cancer
NSW-OAR	Non-swallowing organ at risk
NT	Normal tissue
NTCP	Normal tissue complication probability
OAR	Organ at risk
OC	Oral cavity
OPC	Oropharyngeal cancer
OPSE	Oropharyngeal swallowing efficiency
ORD	Observer
OS	Oesophageal sphincter

OS	Overall survival
PBS	Pencil beam scanning
PBT	Proton Beam Therapy
PC	Pharyngeal constrictor
PCM	Pharyngeal Constrictor muscle
PET	Positron emission tomography
PI	Principal Investigator
PORT	Post-operative radiotherapy
PPW	Posterior pharyngeal wall
PRD	Patient-reported dysphagia
PRV	Planning at risk volume
PRX	Patient-reported xerostomia
PSPT	Passive scanning proton therapy
PTV	Planning target volume
PW	Pharyngeal wall
RAD	Radiation-associated dysphagia
RAD6M	Radiation-associated dysphagia at six months
RCR	Royal College of Radiologists
RCT	Randomised controlled trial
RE	Robust evaluation
RMH	Royal Marsden Hospital
RMW	Range modulator wheel
RO	Robust optimisation
RPN	Retropharyngeal lymph node
RS	Raystation

RT	Radiotherapy
RTOG/EORTC	Radiotherapy Oncology Group/European Organisation for Research and Treatment of Cancer Late Radiation
LRMS	Morbidity Scale
RTTQA	Radiotherapy trials and quality assurance
RU	Range uncertainty
RVR	Remaining volume at risk
S-IMRT	Standard Intensity Modulated Radiotherapy
SC	Spinal cord
SD	Standard deviation
SE	Setup errors
SFO	Single-field optimisation
SGL	Supraglottic larynx
SMPCM	Superior and middle pharyngeal constrictor
SOBP	Spread-out Bragg peak
SPC	Superior Pharyngeal Constrictor
SPS	Swallowing performance scale
SPSS	Swallowing performance status scale
SPSSv25	Statistical Package for the Social Sciences version 25
STAPLE	Simultaneous Truth and Performance Level Estimation
SW-IMRT	Swallow-sparing Intensity Modulated Radiotherapy
SW-OAR	Swallowing organ at risk
TD25	Tolerance dose estimating 25 % probability of toxicity
TD50	Tolerance dose estimating 50 % probability of toxicity
TNS	Threshold not specified

TORS	Transoral surgery
TPS	Treatment Planning System
UK	United Kingdom
UW-QOL	University of Washington Head and neck-related quality of life
V50	Volume of structures receiving $\geq$ 50 Gy
VC	Vocal cords
VF	Videofluoroscopy
VMAT	Volumetric modulated arc therapy

1 Chapter 1: Introduction

1.1 Background

The incidence of oropharyngeal cancers (OPC) in the United Kingdom (UK) has doubled in recent years, with an age-standardised rate of 4.1 in 2011 compared to 2.1 in 2002[1]. In parallel, survival outcomes have improved substantially. Contributory factors include the emergence of good-prognosis, radiosensitive, human papillomavirus - (HPV) driven tumours, and evolving treatment paradigms for poor-prognosis tumours developed on a framework of innovations in diagnostic and therapeutic options.

For patients with American Joint Committee on Cancer (AJCC) (7th edition) stage III/IV disease, termed as locally advanced OPC (LA-OPC), primary platinum-based chemo-radiation (CRT) or RT alone is the curative treatment modality of choice, delivering equivalent or improved survival and better functional outcomes compared to surgery. Primary surgical options are limited to T1-2 N0 OPC, where minimally invasive transoral surgery (TORS) in conjunction with neck dissection might be an alternate function sparing, less morbid therapeutic option[2]. However, a significant proportion of patients treated with TORS require adjuvant RT-based treatment, thereby undermining the benefits of primary surgery[4].

Despite the routine use of intensity-modulated RT (IMRT) in head and neck cancers (HNC) to spare dose to normal tissue, RT-based therapeutic options remain toxic. On a background of rising tumour incidence in a relatively younger population and improved cure rates, an enlarging cohort of cancer survivors are therefore living with devastating long-term functional impairments - paradoxically as a result of 'organ-preserving' RT[5]. Toxicity-mitigating strategies are, therefore, an increasing focus of research in LA-OPC.

Chronic radiation - associated dysphagia (RAD) represents a substantial burden in this context, with nearly 50% of patients highlighting it as a distressing symptom a year following treatment completion[6]. Persistent swallowing dysfunction leads to an increased risk of aspiration, which is typically under-reported in most HNC trials, where assessments are undertaken only at the onset of clinical symptoms, thereby failing to identify patients who aspirate silently, and detected only after incidental objective evaluation. Potentially life-threatening aspiration pneumonia, dietary modifications, malnutrition, and prolonged feeding tube dependence are a direct consequence of RAD, leading to poor social interactions along with lifestyle alterations for both patients and their carers/family members[7-13]. A rare, but devastating, delayed manifestation of delivered RT dose to critical swallowing structures in OPC is lower cranial nerve neuropathies, which results in refractory functional pharyngeal dysphagia, recurrent aspiration pneumonia, and subsequent permanent dependence on feeding tube[14].

There is, therefore, a real risk that current therapeutic options for LA-OPC will lead to a generation of cured patients who are subsequently exposed to decades of debilitating post treatment swallowing dysfunction, resulting in an adverse impact on health-related quality of life (HR-QoL). Consequently, there is an urgent unmet need to devise efficient dysphagia-optimising RT strategies in these tumours.

1.2 Swallowing Organ at risk (SW-OAR) for swallowing function and dosimetric correlation with post-RT dysfunction

1.2.1 Physiology of swallowing

The swallowing process is complex, involving the intricate co-ordination of > 25 pairs of muscles in the oropharynx and larynx, and cartilages[15]. It begins with the voluntary oral preparatory phase, in which a bolus of food is prepared by the rhythmic activity of the intrinsic tongue muscles and genioglossus muscle, the largest extrinsic muscle. This food bolus is then pushed posteriorly to the base of tongue (BoT), to trigger the pharyngeal phase of swallowing. During this involuntary phase, contraction of the extrinsic tongue muscles facilitates the retraction of BoT, moving it posteriorly towards the posterior pharyngeal wall (PPW). Subsequently, the food bolus is propelled into the oesophagus by the downward contraction wave of the PPW, initiated by the pharyngeal

constrictor muscle (PCM). During the pharyngeal phase of swallowing, the larynx is protected by hyolaryngeal elevation, wherein there is up and forward movement of the hyoid bone and the larynx. This results in the thickening of the base of the epiglottis and subsequent tilt downwards to close the laryngeal entrance, and is facilitated by the actions of the floor of mouth and the digastric muscles[16]. The pressure built up during this phase also permits the opening of the upper oesophageal sphincter, thereby directing the food bolus into the oesophagus to initiate the oesophageal phase of swallowing.

1.2.2 Dosimetric relationship between RT dose to SW-OARs and RAD

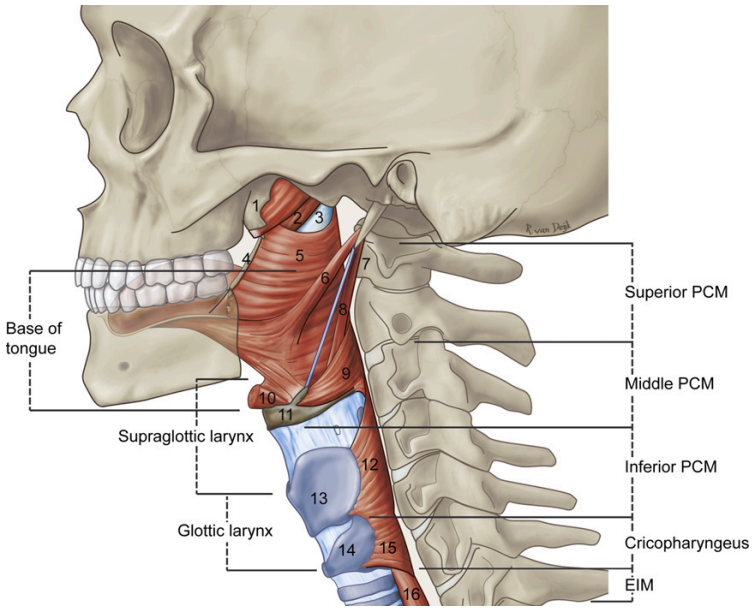


Figure 1.1 Overview of anatomical structures involved in swallowing

The PCM consists of 3 groups (superior, middle, and inferior PCM) of circular muscles that form the postero-lateral wall of the pharynx. Impaired functioning of the PCM during swallowing leads to reduced peristalsis and a non-synchronised pharyngeal constriction wave. This increases the risk of dysphagia, and aspiration due to the presence of residual food in the oro- and hypo- pharynx. (Reproduced from Christianen et al [3]).

Pioneering work by Eisbruch *et al* firmly established the strong influence of the PCM and glottis-supraglottic larynx (GSL) irradiation on persistent functional impairment following CRT in HNC[17] (Figure 1.1). Their study additionally demonstrated that 50 Gy was the lowest minimal dose delivered to a stricture volume – a surrogate for RAD, implying that it may be clinically advantageous to minimise the volumes receiving of $\geq$ 50 Gy (V_{50}) in such critical dysphagia/aspiration-related structures (DARS). Compared to 3-dimensional (3-D) conformal radiotherapy, IMRT reduced DARS V_{50} by 7-10% on average, consequently motivating a number of centres to analyse the influence of the dose delivered to DARS on various measures of late dysphagia. In recent years, additional structures such as the oral cavity, and the mylo/geniohyoid complex have also been implicated with persistent dysfunction[18, 19].

Several statistically significant dose-response constraints for key swallowing structures have been proposed as a result (Table 1.1). Strong correlations exist between both partial volume doses and mean doses with persistent dysphagia outcomes, implying that the mean dose as a solitary dosimetric variable should suffice for planning optimisation.

Despite an abundance of published literature, it is challenging to make unequivocal conclusions regarding the optimal swallow-sparing parameters. The systematic review by Duprez *et al,* which concurred that the mean dose to the PCM was a strong predictor of subsequent functional impairment, highlighted a number of crucial methodological and statistical variations

amongst the analysed studies that hindered the validity of the review's outcomes[20]. Significant heterogeneity in a number of confounding and prognostic variables, such as primary tumour location, tumour stage, use of concomitant chemotherapy, fractionation schedule and target volume definition, limit the conclusions that can be drawn. The robustness of the reported results is further diluted with the predominantly retrospective nature of most studies, together with small sample sizes and inconsistent recording of swallowing outcomes.

Nonetheless, reducing the radiation dose to DARS, without compromising on survival outcomes, is necessary for improving long-term swallowing function in OPC. Furthermore, the degree of sparing of individual DARS required to generate potential increments in function will vary, depending upon the site of the primary tumour. For instance, in nasopharyngeal (NPC) and OPC, the superior pharyngeal constrictors (SPC) are more likely to be irradiated to a radical dose compared to GSL, and consequently more likely to account for swallowing difficulties following treatment completion. Focussed efforts to reduce the SPC dose in such scenarios are more likely to be advantageous, rather than sparing the GSL. Similarly, in hypopharyngeal cancers (HPC), the inferior constrictors and GSL radiation doses are likely to play a more influential role on long-term swallowing function.

Authors	Patient nos	Tumour site	Treatment modality	Dysphagia outcome measure	Dysphagia endpoint	Timing	DARS	Dosimetric parameters correlating with late dysphagia
Feng et al[21]	36	OPC, NPC	CRT	VF UW-QOL HN-QOL RTOG/EORTC LRMS	Aspiration Stricture G2 ORD PRD liquids PRD solids	3 mo	PCM SGL - PCM Oesophagus, PCM PCM	Mean 60 Gy, V_{65} > 50% V_{50} > 50% Limited cases; statistical analysis not possible Mean dose - TNS Mean dose – TNS Mean dose - TNS
Schwartz et al[22]	31	OPC	CRT	MBS MDADI	OPSE Composite MDADI score Aspiration	6-24 mo	Anterior OC SPC - -	V30 > 65% V55 > 80% None predictive Limited cases, statistical analysis not possible
Caglar et al[23]	96	OC, OPC, NPC, HPC, Larynx, unknown primary, maxillary sinus	IC + CRT CRT RT PORT	SPS Video swallow study	Aspiration Stricture	1-2 mo	Larynx IC IC Larynx	Mean 48 Gy; V_{50} > 21% Mean 54 Gy; V_{50} >51% Mean 54 Gy; V_{50} > 51% V_{50} >21%

Authors	Patient nos	Tumour site	Treatment modality	Dysphagia outcome measure	Dysphagia endpoint	Timing	DARS	Dosimetric parameters correlating with late dysphagia
Bhide et al[24]	37	OPC, NPC, HPC, Larynx	IC + CRT	MDADI RTOG/EORTC-scored dysphagia	MDADI score ORD	12 mo	PCM Suprahyoid muscles	Nil Significant
Caudell et al[25]	83	OC, OPC, NPC, HPC, Larynx, Unknown Primary	CRT RT Cetuximab + RT	MBS	FT dependence Aspiration Stricture requiring dilatation	12 mo	Larynx IPC Larynx IPC SPC MPC	Mean 51 Gy V_{60} > 12% Mean 41 Gy, V_{60} > 24% V_{60} > 12% V_{65} > 33% V_{65} > 75%
Dornfeld et al[26]	27	OC, OPC, HPC, Larynx, Unknown/Other	CRT RT	HNCI	QoL score Weight loss FT rates Dietary modifications	12 mo	- AEF False VC False VC Upper OS Lateral PW	Nil significant Mean 50 Gy, at the level of false VC
Li et al[27]	39	OC, OPC, HPC, Larynx, Unknown Primary	CRT	CTCAE v2.0 UW-QOL	FT dependence	>192 days	IPC CPI	Mean 60 Gy; V_{65} > 30%; V_{60} > 60% Dmax 62 Gy
Dale et al [18]	34	OPC	CRT	VF	Aspiration Stricture FT dependence Aspiration pneumonia	≥12 mo	SPC MHM	MHM V_{69} and SPC V_{70} associated with late dysphagia

Authors	Patient nos	Tumour site	Treatment modality	Dysphagia outcome measure	Dysphagia endpoint	Timing	DARS	Dosimetric parameters correlating with late dysphagia
Christianen et al[28]	354*	OC, OPC, NPC, HPC, Larynx, Unknown primary, Other	RT CRT	RTOG/EORTC LRMS EORTC QLQ-H&N 35	≥ Grade 2 dysphagia (Primary Endpoint) Patient-reported symptoms (Secondary Endpoint)	6 mo	SPC SGL	Mean dose to both most predictive of late toxicity. Different predictive models were found for solid food, liquids, soft food and choking
Mortensen et al[29]	259	OC, OPC, NPC, HPC, Larynx	RT CRT Other	EORTC QLQ-H&N 35 MBS** DAHANCA dysphagia scale	PRD SPSS, Aspiration ≥ Grade 2 dysphagia	3 yrs	SGL SPC, MPC	Mean dose 55 Gy Mean dose 60 Gy
Mazzola et al[30]	56	OC, OPC, NPC, Larynx, Salivary Glands	CRT IC + CRT RT, PORT	RTOG/EORTC LRMC	≥ Grade 2 dysphagia	6 mo	SPC	V60 ≥ 70%

Table 1.1 IMRT studies investigating the correlation between radiation dose to swallowing structures and late dysphagia

Abbreviations: *AEF* Aryepiglottic fold, *CE* cervical oesophagus, *CPI* Cricoid pharyngeal inlet, *FT* Feeding tube, *HNCI* Head and Neck Cancer Inventory, *HN-QOL* Head and Neck Quality of Life, *IC* Induction chemotherapy, *IPC* Inferior pharyngeal constrictors, *MBS* Modified Barium Swallow, *MHM* Myelo-geniohyoid complex, *MDADI* MD Anderson Dysphagia Inventory, *mo* months, *MPC* middle pharyngeal constrictors, *OC* Oral Cavity, *OPSE* Oropharyngeal Swallowing Efficiency, *ORD* Observer-rated dysphagia, *OS* oesophageal sphincter, *PORT* Post-operative RT, *PRD* Patient-reported dysphagia, *PW* pharyngeal wall, *RTOG/EORTC LRMS* Radiotherapy Oncology Group/European Organisation for Research and Treatment of Cancer Late Radiation Morbidity Scale, *SPS* Swallowing performance scale, *SPSS* Swallowing performance status scale, *TNS* Threshold not specified, *UW - QOL* University of Washington Head and Neck-related Quality of Life, *VC* vocal cords, *VF* Videofluoroscopy
* 38% of patients treated with IMRT
** 65 patients only

1.3 Swallowing-sparing IMRT approaches

1.3.1 Reducing mean dose to the SPC and supraglottic larynx (SGL) using a model-based validation approach

Designed and promoted by a consortium of Dutch radiation oncologists primarily for the selective implementation of proton therapy in their country, the normal tissue complication probability (NTCP) model-based concept is described as a practical alternative to randomised controlled trials (RCT), particularly where the principal aim of a novel radiation technique is toxicity-reduction rather than survival gains. In this multi-step methodology, any potential benefit predicted during the early phases is subsequently confirmed by validating its model-based estimates in a cohort of patients who are prospectively followed-up[31, 32].

Christianen *et al* initially determined that mean doses to the SPC and SGL were most predictive of RTOG grade $\geq$ 2 dysphagia at 6 months following treatment completion in a heterogeneous group of HNC patients[28]. Their subsequent in-silico comparative planning study, in predominantly pharyngeal cancers (PC), suggested a 8.9% reduction in mean NTCP (42% v 33%) for the physician-rated toxicity scores with swallow-sparing IMRT (SW-IMRT) that was additionally optimised to reduce doses to SPC, SGL, middle constrictors and oesophageal inlet, in that order of priority, compared to standard IMRT

(S-IMRT) [33]. DARS-sparing was achieved by reducing planning target volume (PTV) coverage in its vicinity until exactly 98% of the PTV volume received 95% of the prescribed dose, together with accepting moderate shift of dose to non-specified tissues, such as the neck muscles and the oral cavity. Absolute gains in NTCP values varied considerably, depending upon the primary tumour site, nodal involvement, and tumour stage[34]. Finally, their model was clinically validated in a prospective cohort of 186 patients treated with SW-IMRT, where the mean predicted $NTCP_{SW\text{-}IMRT}$ for the entire group corresponded perfectly with the observed grade $\geq$2 dysphagia prevalence of 22.6%, and was significantly lower than the predicted $NTCP_{S\text{-}IMRT}$ of 27.5%. The predicted differences were significantly larger (24.1% v 32.2%), and, importantly, clinically relevant in approximately 50% of patients with a ΔNTCP ($NTCP_{S\text{-}IMRT} – NTCP_{SW\text{-}IMRT}$) > 5%, with observed toxicity prevalence of 25.3%[35]. Patients in this subset typically had higher T stages, primary OPC or NPC and were treated more often with conventional RT or CRT.

The group's novel SW-IMRT technique did not compromise target volume coverage, a detrimental limitation of some of the previous planning studies. Likewise, doses to the major salivary glands did not differ compared to S-IMRT, crucial as patients' perception of swallowing difficulties can often be influenced to varying degrees by co-existing xerostomia. By excluding patients with grade >1 dysphagia at baseline, the investigators ensured that any subsequently reported dysphagia was purely treatment-related.

Implementing such a model-based approach in routine clinical practice, however, will be resource- and time-intensive, with its success reliant on experienced physicists having iteratively to adjust the planning objectives for DARS till a suitable plan is achieved. Equally, treatment planning systems with a fully automated or class solution, where comparatively less effort is required to achieve a satisfactory plan, are less likely to spare DARS sufficiently to observe a clinical benefit. Crucially, as with any planning modelling exercise, the predictability of the reported benefits ultimately depends on the robustness of the primary endpoint selected to develop the particular model. In that context, the use of physician-scored, RTOG-graded dysphagia at 6 months to define post-radiotherapy long-term dysphagia is debatable, with a number of HNC studies demonstrating substantial variation in swallowing outcomes beyond 6 months. Furthermore, data recently published by the same group analysing patterns of RTOG-scored swallowing dysfunction following HNC treatment established that 23% of patients could have a clinically relevant change in the physician-reported dysphagia scores beyond 6 months, indicating that the 6-month timeline might be inconsistent at predicting future toxicity[36]. The same study additionally demonstrated the decreasing influence of radiation dose to SGL over time, leaving the SPC as the sole significant variable.

It must also be emphasised that physician-reported swallowing scores often do not correlate well with patient-reported outcomes and, as a primary endpoint, may not necessarily provide the best measure of toxicity outcomes. For instance, in the Dutch group's model, NTCP-based reductions in patient-

reported swallowing dysfunction with swallow-sparing IMRT were variable and lower than observer-rated NTCP reductions. Similar inconsistencies between observer- and subjective- reported dysphagia have been reported in other studies as well, strengthening the argument for multi-dimensional swallowing assessments[37, 38]. Nonetheless, this strategy demonstrated that swallow-sparing RT strategies are feasible, and the groups' NTCP model is the most frequently applied predictive model to evaluate late dysphagia.

1.3.2 Reducing the radiation dose to DARS outside the target volume

Situated in close proximity to the PCM, both medial and lateral groups of retropharyngeal lymph nodes (RPN) have been historically included in radiation target volumes for PC. As a result, the constrictors usually receive a substantial radiation dose, making it challenging to preserve long-term function. Feng *et al* observed that the practice of irradiating the entire uninvolved RPN compartment was inconsistent with the available evidence on patterns of nodal spread, as indicated by the paucity of metastasis to medial RPN in several surgical and radiological series[21]. The group postulated that the medial group of RPN could be safely excluded from OPC target volumes in their novel toxicity-mitigating IMRT approach, thereby potentially improving function without affecting survival outcomes.

They prospectively evaluated their hypothesis in a selective group of stage III/IV OPC patients treated with primary CRT[39]. Parts of PCM, GSL and

oesophagus in the region of the uninvolved medial RPN were spared by setting an optimal dose constraint of <50 Gy in the IMRT planning objectives, which subsequently delivered mean doses of 48, 42 and 32 Gy, respectively, to the spared regions. Corresponding mean doses to the entire structures were 58 Gy, 48 Gy and 34 Gy, respectively. With a median follow-up period of 3 years, the clinical outcomes of such dosimetric modulation in this single-arm study were no worse than standard approaches, with loco-regional recurrence-free and disease-free survival rates of 96% and 88%, respectively. Crucially, there were no failures observed within or near the spared region, thereby establishing the safety of this swallow-sparing technique.

Patient-reported swallowing outcomes from 2 established questionnaires showed worsening soon after completion of treatment, with gradual improvement through 12 months and subsequent stabilisation, whereas common toxicity criteria for adverse events (CTCAE) v2.0 – based observer-reported dysphagia scores at 12 months almost matched baseline levels. Unlike these two measures of swallowing toxicity, VF-related scores did not show longitudinal improvements, with no significant reductions observed beyond 3 months following treatment. The lack of late CTCAE-graded toxicities precluded any dosimetric analysis of observer-reported toxicities. Mean doses to PCM, GSL and oesophagus correlated significantly with worsening subjective and instrumental swallowing assessments, and different NTCPs with no particular threshold were observed with differing endpoints. The tolerance doses that estimated a 50 % (TD_{50}) and 25 % (TD_{25}) probability for VF-assessed dysphagia were 63 and 56 Gy, respectively, for PCM, and 56 Gy

and 39 Gy, respectively, for GSL[40]. The corresponding TDs for patient-reported worsened outcomes were substantially higher, reflecting to some extent the increased sensitivity of VF to detect silent aspirators.

The above study presents an innovative, practical and adaptable solution to generate a potentially beneficial toxicity-mitigating strategy in OPC; integrating existing knowledge of patterns of nodal disease spread into the IMRT planning objectives to further refine radiation delivery. Although the study only included patients with OPC, its methodology can be easily extended to other pharyngeal tumours too. A novel hypothesis at the time of study design, the concept of sparing the medial RPN from target volumes has been endorsed in the recently updated HNC nodal outlining consensus guidelines[41]. A relatively favourable patient-reported toxicity outcome, together with minimal physician-graded toxicity scores, supports the application of similar dysphagia-optimising strategies in randomised studies to better define its true benefits. Notably, the group has also reported that HR-QoL appears to remain stable with longer follow-up, with new late toxicity uncommon beyond 2 years[42].

It is clear that the above strategy cannot be extrapolated to all OPC; the study was selective by excluding any tumour with PPW or RPN involvement, explaining to a certain extent the excellent survival rates reported. The absence of a steep dose-response curve makes it difficult to establish definitive IMRT dose constraints for PCM and GSL, though the study authors have employed the VF-based TD_{25} information to guide their planning

objectives at their centre. Sharp dose fall-offs would be expected in the regions of the spared structures, and currently available routine imaging techniques do not yet possess the required sensitivity to accurately define the mucosal extent of tumours; target contouring, therefore, may need to be more generous in such scenarios to reduce the risk of marginal recurrence which eventually would prove to be counter-productive to the primary goal of toxicity reduction. Finally, attempting to spare the medial RPN with IMRT may result in a dose splash to adjacent structures linked to swallowing dysfunction such as the salivary glands and oral cavity, which could potentially worsen long-term functional outcomes.

1.4 Potential of protons for reducing chronic RAD in HNC

The use of multiple beams and fluence modulation in IMRT to improve target conformity and spare dose to OAR invariably leads to a ‘low-dose bath’ of normal tissue, and subsequent toxicity. Consequently, there has been an increased interest in alternative modalities of RT as a means of reducing treatment-related toxicity. In this context, protons offer distinct depth dose advantages that may further enhance the therapeutic ratio in HNC. Protons have a finite range, and a mono-energetic beam deposits its dose predominantly towards the end of its range, known as the Bragg peak (BP). To treat a target volume at depth, proton beams of multiple energies are combined to form a spread-out Bragg peak (SOBP)[43]. As shown in Figure 1.2, a SOBP has a lower proximal dose and almost zero dose distal to the

target when compared to a photon beam. Additionally, proton beams have a much sharper lateral penumbra. As a result, the integral dose delivered to normal tissue with the highly conformal proton therapy is decreased by a factor of approximately two, which could potentially improve local control and quality of life by allowing dose escalation to the tumour and greater sparing of non-target tissue.

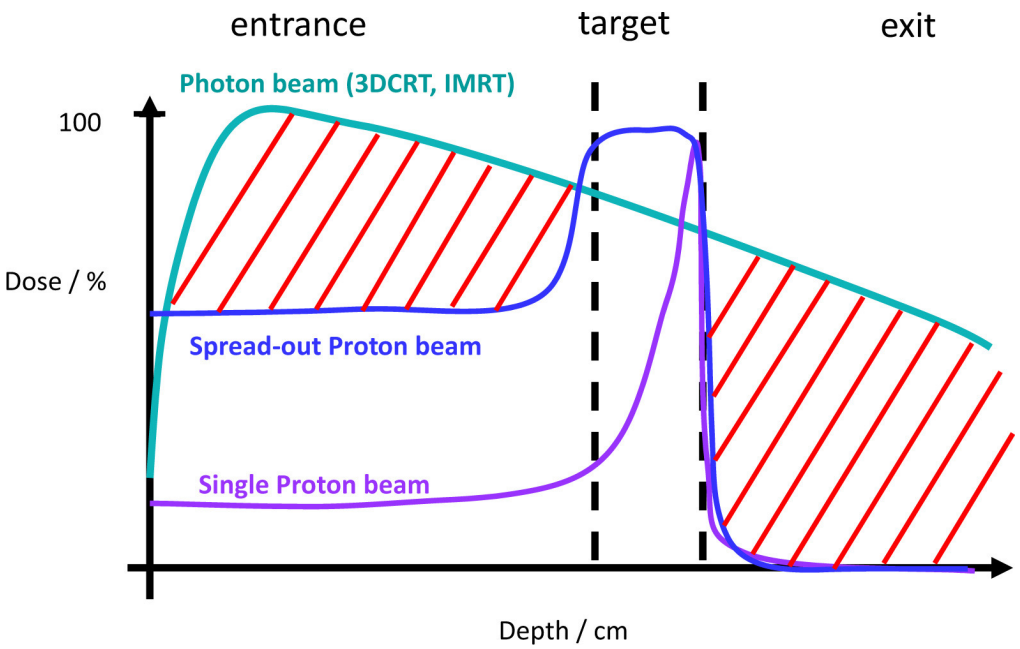


Figure 1.2 Dose depth curves of photons and protons

(Reproduced from Blanchard et al[44]). The dashed red zone represents the unnecessary dose delivered by photons that could be avoided using proton therapy.

1.4.1 Treatment delivery techniques of proton therapy

Therapeutic proton energies, ranging from 70 to 250 MeV, are generated initially as thin pencil beams by proton accelerators. To deliver treatment dose to a target at depth, proton beams are spread longitudinally and laterally to create the SOBP. This is achieved either by the use of mechanical shaping devices, known as passive scattering proton therapy, or by magnetic deflection of the thin beams, known as pencil beam scanning.

1.4.1.1 Passive scattering proton therapy (PSPT)

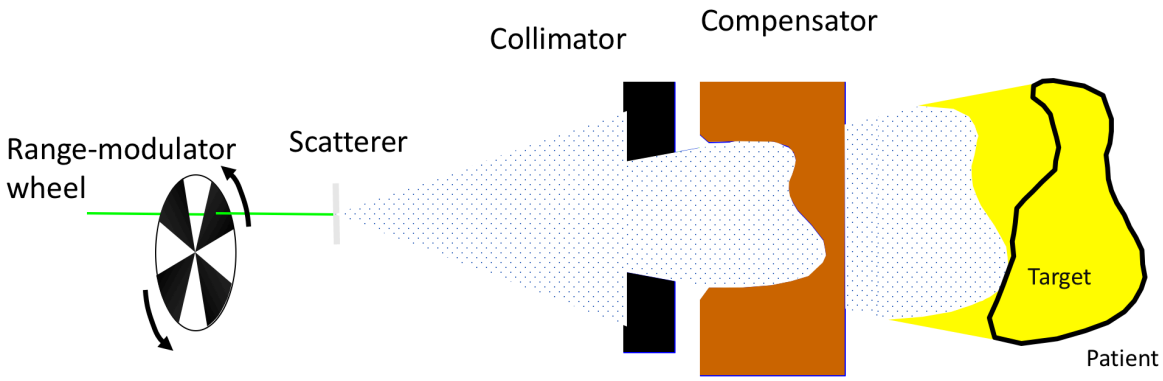


Figure 1.3 Diagrammatic representation of passive scatter proton therapy

This demonstrates deposition of extra dose proximal to the target.(Adapted from Pedroni et al[45])

In PSPT, longitudinal and lateral beam shaping is achieved with a combination of range modulator wheel (RMW), and scatterers (Figure 1.3). A RMW is a rotating wheel comprising steps of varying thickness that attenuates the proton

beam energy for creating the desired SOBP. Lateral spread is achieved by using 2 scatterers with a high atomic number. The SOBP is conformed to the target laterally by a brass aperture, and distally by a compensator made of water-equivalent material such as Lucite.

The SOBP in passive scattering is of constant modulation, designed to conform to the distal edge of the target only, thereby providing limited control over dose delivered to tissue that is proximal to the target. PSPT can, therefore, be considered analogous to 3-D conformal photon RT delivery.

1.4.1.2 Pencil beam scanning (PBS) or spot-scanning proton therapy

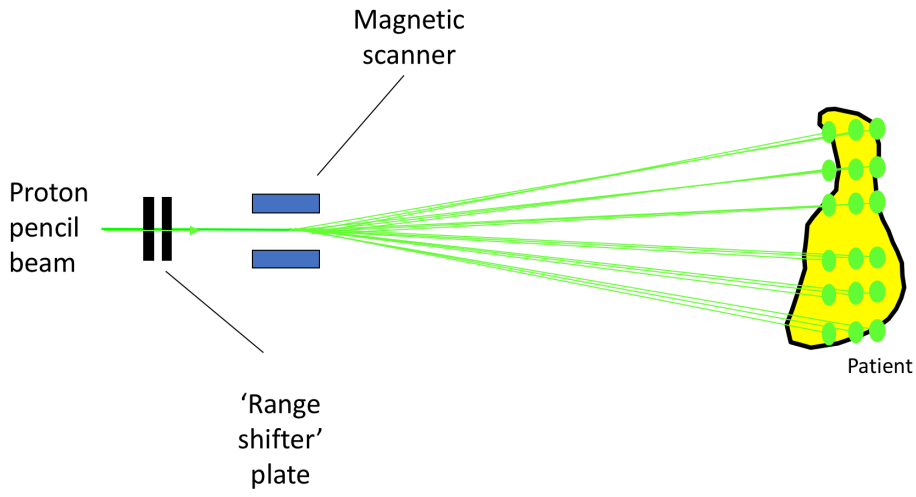


Figure 1.4 Diagrammatic representation of pencil beam proton therapy

The figure demonstrates improved target conformity compared to passive scatter proton therapy (Adapted from Pedroni et al[45])

In PBS, magnetic scanners sequentially direct thin pencil beamlets with a desired intensity to a predetermined spot within the target. The required dose distribution is achieved by depositing thousands of spots in layers in the target (Figure 1.4). PBS systems do not require RMW, apertures or compensators for beam shaping, and also provide superior conformity compared to PSPT.

Single-field optimisation (SFO) and multi-field optimisation (MFO) are the two approaches to deliver PBS treatment. In SFO, each individual proton field uniformly covers the target. This limits normal tissue sparing, as a few spots must pass through the OAR to deposit dose to parts of the target situated behind the OAR. On the other hand, uniform target coverage in MFO is provided by the optimisation of heterogeneous dose distribution from individual fields. Therefore, MFO offers more potential to spare RT dose to normal tissue as no spots are placed in the target volume lying behind the OAR from each individual field (Figure 1.5). MFO is also known as intensity-modulated proton therapy (IMPT) in ICRU78. IMPT is capable of true 3D dose painting as it provides the ability to not only modulate the dose laterally, similar to IMRT, but also at depth by varying proton energies, which IMRT cannot provide. Such dosimetric advantages could have potential beneficial clinical implications in OPC, with single-centre retrospective studies reporting reducing symptom burden following treatment[46, 47].

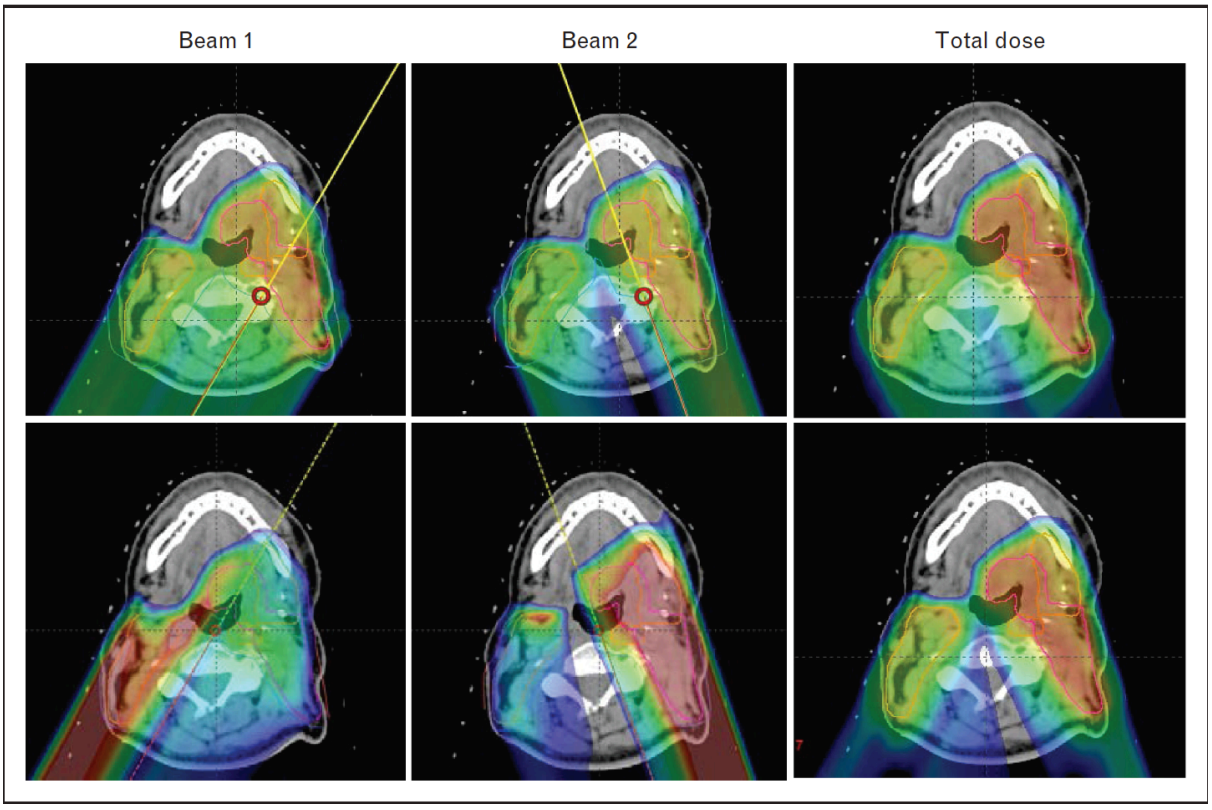


Figure 1.5 Single-field and multi-field optimisation pencil beam proton therapy

Example dose distribution of individual beams (left, centre), and the sum of the planned beams for SFO plan (top) and MFO plan (bottom). Reproduced from Lukens et al[48]

1.4.2 Clinical experience with protons in pharyngeal cancers

Few retrospective studies have compared toxicity and survival outcomes of PC patients treated with IMPT and IMRT (Table 1.2). The case-matched analysis of predominantly HPV-positive OPC requiring bilateral neck irradiation by Blanchard et al showed no significant differences in clinical outcomes between the two RT techniques after a median follow up of 32 months[49]. IMRT patients had more weight loss or gastrostomy tube insertion at 3 – and 12 – months following treatment completion compared to IMPT. Interestingly, IMPT did not reduce the rates of acute mucositis or chronic xerostomia compared to IMRT. An analysis of loco-regional failure patterns in the IMPT group reassuringly did not demonstrate any marginal failures[47]. Sio et al reported fewer patient-reported symptom burden at 3 months following treatment completion with IMPT, though no differences were observed at 12 months between the 2 RT techniques[50].

These data are hypothesis-generating studies, which require prospective validation to confirm the role of IMPT in reducing toxicities in PC patients.

Reference	Number of patients	Tumour site	Methodology	Survival rates	Toxicity endpoint	Outcomes		
						IMRT	IMPT	p value
Blanchard et al[49]; Gunn et al[47]	50 IMPT, 100 IMRT	OPC	Retrospective case-matched control comparison	3 yr OS; 94.3 % (IMPT) v 89.3 % (IMRT); p = 0.44	> Grade 2 PRX at 12 months	47.2 %	42 %	0.23
					G-tube or weight loss at 12 months	24.7 %	8 %	0.01
Holliday et al[46]	10 IMPT, 20 IMRT	NPC	Retrospective case-matched control comparison	2 yr LRC; 100 % (IMPT) v 95 % (IMRT)	G-tube during or after treatment	65 %	20 %	0.02
Sio et al[50]	35 IMPT, 46 IMRT	NPC	Retrospective cohort comparison	Not mentioned	Subacute food taste symptoms[§];	7.70	5.76	0.01
					Subacute appetite symptoms[§];	6.37	4.68	0.048
					Chronic appetite symptoms[§]	4.14	2.12	0.036
					Subacute mucous symptoms (% with moderate-severe symptoms)	84 %	62 %	0.038

Table 1.2 Direct comparisons of IMRT versus IMPT in pharyngeal cancers

G-tube, gastrostomy tube; IMPT, intensity-modulated proton therapy; IMRT, intensity-modulated radiotherapy; LRC, loco-regional control; NPC, nasopharyngeal cancer; OPC, oropharyngeal cancer; OS, overall survival; PRX, patient-reported xerostomia
[§] Mean MD Anderson Symptom Inventory – Head and Neck score

1.5 Inter-observer variability (IOV) in OAR delineation in HNC

Clinician-led definition of both target and non-target volumes remains the weakest link in the RT pathway, with several studies highlighting it as a major contributor for geometric errors during RT delivery[51]. To minimise treatment-related morbidity, accuracy in OAR delineation is necessary for optimal avoidance during inverse planning IMRT. In addition, significant inter-observer variability (IOV) in the delineation of OARs will invariably result in inconsistent toxicity reporting and interpretation of RT effects, and any subsequently modelled NTCP model is therefore unlikely to be robust.

1.5.1 Summary of studies evaluating variability in OAR delineation in HNC

Few studies have described OAR contouring variability in HNC, particularly for the constrictor muscles. Their results suggest the presence of substantial variability, though the clinical impact of such inconsistency on delivered dose and subsequent toxicity remains uncertain.

Mukesh et al observed substantial inter-observer differences in the volumes of the OARs in their study, where four HN oncologists outlined the contralateral parotid, brainstem (BS), and spinal cord (SC), together with the parotid bed, in 5 cases treated with post-operative RT for parotid tumours[52]. The mean

conformity levels (CL; ratio of intersection and union volumes for two or more structures; ideal score 100 %) for the OARs were 60 % (range 53 % - 68 %), 23 % (13 % - 28 %), and 25 % (22 % - 31 %) respectively. CL improved for the BS (45 %, range 42 % - 51 %) and SC (60 %, range 58 % - 64 %) when the analysis was restricted to the CT slices where they were contoured by all observers. Considerable variability was noted along the axial directions for the OARs, as well as for junction between the BS and SC. Such shortcomings could potentially be addressed by implementing contouring guidelines.

Brouwer et al investigated the influence of delineation guidelines to reduce IOV, by analysing the magnitude of contouring variability in OARs (SC, parotids, submandibular glands, thyroid cartilage, glottic larynx) between five HN oncologists in 6 HNC patients in their study. The group used intraclass correlation coefficient (ICC) to assess differences in volume (where a score of 0.81 – 1.0 implies substantial agreement between two volumes), concordance index (C-I; ratio of intersection volume to union volume where a score of 1.0 indicates perfect overlap) to evaluate positional changes, and a 3-D analysis to compare the local variation of each observer's contours for individual OAR with the median contour of all 5 observers' outlining. [53]. There was considerable uncertainty in identifying the superior and inferior borders for all OARs, which may be a result of poor soft tissue discrimination on the RT planning scan or the incorrect interpretation delineation guidelines. In particular, significant variability was observed for the glottic larynx for all 3 endpoints (ICC 0.27, C-I 0.37, 3D standard deviation 3.9 mm), with poor compliance to the available delineation guideline. Results of this study were supported by a separate study

by Geets et al, where a small but significant IOV was demonstrated for the parotid glands and SC[54].

1.5.2 Clinical impact of OAR delineation variability

Nelms et al evaluated the IOV in OAR contouring, and the dosimetric impact of such variability amongst 32 clinicians[55]. Centres were provided with a common CT dataset of OPC with predefined target volumes and expected to contour OARs routinely outlined in their clinical practice; the most commonly delineated OARs (BS, parotids, SC, mandible) were then compared with the study investigators' reference standard using metrics that compare the level of agreement between two contours such as Dice similarity coefficient (DSC) (ideal score 1), and distance to agreement (DTA). Significant variation in OAR delineation was observed, most pronounced for the BS, which had a DSC score of 0.66. Each centre also designed an IMRT plan using their delineated OARs during optimisation; the subsequent dose-volume histogram (DVH) data was compared with DVH of the reference OARs that was generated by superimposing the reference contours on each centres' plan. Such an approach helps to quantify the differences in OAR doses due to contouring variation alone, independent of RT optimisation. Substantial variation in RT dose was noted; for instance, -1.8 % to 22.6 % difference in maximum dose to the BS and -18.2 % to 56.1 % difference in mean dose delivered to the ipsilateral parotid gland.

In another study, Feng et al investigated the variability in the contouring of salivary glands, PCM, and larynx by three oncologists, who jointly delineated the OARs on 3 separate occasions in 10 OPC patients[56]. Substantial variation in fractional overlap (intersection volume divided by union volume) was observed for all organs (mean 0.7 $\pm$ 0.1), but particularly for PCM (0.5 $\pm$ 0.1). The group assessed the dosimetric impact of delineation uncertainty by comparing three IMRT plans: expert, first set of joint contours; union, representing maximal OAR sparing; intersection, representing minimal OAR sparing by including only the volume where all 3 sets of contours matched. Despite spatial variation in contouring, there was minimal difference in dose delivered to OARs (mean difference 0.9 Gy; range 0.6 Gy – 1.1 Gy), with the largest variation in dose observed for the submandibular glands and larynx. This suggested relative insensitivity of dose delivered to OARs to contouring variability.

1.6 Dysphagia-optimised IMRT (Do-IMRT) for OPC

Do-IMRT is a novel planning technique that optimises the dose delivered to the constrictor muscles, and is currently under evaluation in the CRUK-funded DARS (ISRCTN 25458988) trial. The study is a UK multi-centre phase III randomised clinical trial with blinded assessments of key outcome measures, in patients undergoing radical primary CRT or radiation alone, for T1-4, N0-3, M0 primary PC not involving the RPN or posterior pharyngeal wall and requiring bilateral neck irradiation[57]. Eligible patients are randomised to

either standard IMRT (S-IMRT), or Do-IMRT where the mean dose to parts of pharyngeal constrictors lying outside the radical treatment volumes will be limited to <50 Gy. The primary objective of the study is to determine whether Do-IMRT improves swallowing outcomes, which will be evaluated as a patient-reported outcome using the MD Anderson Dysphagia Inventory (MDADI) score. The difference in the mean MDADI scores at 12 months following treatment completion between the 2 arms forms the primary endpoint of the trial. Swallowing outcomes will be comprehensively assessed using a multi-dimensional, longitudinal panel of objective and subjective functional outcome measures.

1.7 Outline of thesis

This thesis investigates the potential long-term clinical benefits of dysphagia-optimising RT techniques in OPC, together with assessing the impact of variability in SW-OAR delineation on treatment-related morbidity.

Chapter 2 investigates the toxicity-mitigating benefits of Do-IMRT in OPC requiring bilateral neck irradiation compared to S-IMRT, in a comparative planning study in which the predicted probabilities of chronic RAD for Do-IMRT and S-IMRT are evaluated.

IMPT holds great promise as a toxicity-mitigating strategy due to its physical properties. In **Chapter 3**, I determine the swallow-sparing benefits of standard

– and dysphagia-optimised IMPT (Do – IMPT) over Do-IMRT in OPC in a comparative planning study. **Chapter 4** investigates the robustness of the Do-IMPT plans generated in chapter 4 in the presence of set-up and range uncertainties.

The perceived advantage of Do-RT is contingent on contouring accuracy of PCM, an organ-at-risk not delineated routinely in the UK. Heterogeneity in delineation between oncologists may lead to differences in the reported dose-volume parameters, and this can have implications on subsequent toxicity outcomes. **Chapter 5** reports on the qualitative and quantitative variability in PCM contouring in the UK and subsequent potential impact on functional outcomes, within the context of a pre-trial RT quality assurance programme for DARS, a national phase III head and neck trial.

1.8 References list

1. Schache, A.G., et al., *HPV-Related Oropharynx Cancer in the United Kingdom: An Evolution in the Understanding of Disease Etiology.* Cancer Res, 2016. **76**(22): p. 6598-6606.

2. NICE Guidelines. https://www.nice.org.uk/guidance/ng36.

3. Christianen, M.E., et al., *Delineation of organs at risk involved in swallowing for radiotherapy treatment planning.* Radiother Oncol, 2011. **101**(3): p. 394-402.

4. Evans, M. and T.M. Jones, *Transoral Surgery or Radiotherapy for Oropharyngeal Carcinoma - Is It Either Or...?* Clin Oncol (R Coll Radiol), 2016. **28**(7): p. 413-20.

5. Simcock, R. and R. Simo, *Follow-up and Survivorship in Head and Neck Cancer.* Clin Oncol (R Coll Radiol), 2016. **28**(7): p. 451-8.

6. Roe, J.W., et al., *Patient-reported outcomes following parotid-sparing intensity-modulated radiotherapy for head and neck cancer. How important is dysphagia?* Oral Oncol, 2014. **50**(12): p. 1182-7.

7. Hunter, K.U., et al., *Aspiration pneumonia after chemo-intensity-modulated radiation therapy of oropharyngeal carcinoma and its clinical and dysphagia-related predictors.* Head Neck, 2014. **36**(1): p. 120-5.

8. Patterson, J.M., et al., *Head and neck cancer and dysphagia; caring for carers.* Psychooncology, 2013. **22**(8): p. 1815-20.

9. Mortensen, H.R., K. Jensen, and C. Grau, *Aspiration pneumonia in patients treated with radiotherapy for head and neck cancer.* Acta Oncol, 2013. **52**(2): p. 270-6.

10. Xu, B., et al., *Aspiration pneumonia after concurrent chemoradiotherapy for head and neck cancer.* Cancer, 2015. **121**(8): p. 1303-11.

11. Chen, S.W., et al., *The outcome and prognostic factors in patients with aspiration pneumonia during concurrent chemoradiotherapy for head and neck cancer.* Eur J Cancer Care (Engl), 2010. **19**(5): p. 631-5.

12. Hutcheson, K.A. and J.S. Lewin, *Functional outcomes after chemoradiotherapy of laryngeal and pharyngeal cancers.* Curr Oncol Rep, 2012. **14**(2): p. 158-65.

13. Vlacich, G., et al., *Dose to the inferior pharyngeal constrictor predicts prolonged gastrostomy tube dependence with concurrent intensity-modulated radiation therapy and chemotherapy for locally-advanced head and neck cancer.* Radiother Oncol, 2014. **110**(3): p. 435-40.

14. Awan, M.J., et al., *Late radiation-associated dysphagia (late-RAD) with lower cranial neuropathy after oropharyngeal radiotherapy: a preliminary dosimetric comparison.* Oral Oncol, 2014. **50**(8): p. 746-52.

15. Murphy, B.A. and J. Gilbert, *Dysphagia in Head and Neck Cancer Patients Treated With Radiation: Assessment, Sequelae, and Rehabilitation.* Seminars in Radiation Oncology, 2009. **19**(1): p. 35-42.

16. Gawryszuk, A., et al., *Functional Swallowing Units (FSUs) as organs-at-risk for radiotherapy. PART 1: Physiology and anatomy.* Radiother Oncol, 2019. **130**: p. 62-67.

17. Eisbruch, A., et al., *Dysphagia and aspiration after chemoradiotherapy for head-and-neck cancer: which anatomic structures are affected and can they be spared by IMRT?* Int J Radiat Oncol Biol Phys, 2004. **60**(5): p. 1425-39.

18. Head, M.D.A. and G. Neck Cancer Symptom Working, *Beyond mean pharyngeal constrictor dose for beam path toxicity in non-target swallowing muscles: Dose-volume correlates of chronic radiation-associated dysphagia (RAD) after oropharyngeal intensity modulated radiotherapy.* Radiother Oncol, 2016. **118**(2): p. 304-14.

19. Gawryszuk, A., et al., *Functional Swallowing Units (FSUs) as organs-at-risk for radiotherapy. PART 2: Advanced delineation guidelines for FSUs.* Radiother Oncol, 2019. **130**: p. 68-74.

20. Duprez, F., et al., *Systematic review of dose--volume correlates for structures related to late swallowing disturbances after radiotherapy for head and neck cancer.* Dysphagia, 2013. **28**(3): p. 337-49.

21. Feng, F.Y., et al., *Intensity-modulated radiotherapy of head and neck cancer aiming to reduce dysphagia: early dose-effect relationships for*

the swallowing structures. Int J Radiat Oncol Biol Phys, 2007. **68**(5): p. 1289-98.

22. Schwartz, D.L., et al., *Candidate dosimetric predictors of long-term swallowing dysfunction after oropharyngeal intensity-modulated radiotherapy.* Int J Radiat Oncol Biol Phys, 2010. **78**(5): p. 1356-65.

23. Caglar, H.B., et al., *Dose to larynx predicts for swallowing complications after intensity-modulated radiotherapy.* Int J Radiat Oncol Biol Phys, 2008. **72**(4): p. 1110-8.

24. Bhide, S.A., et al., *Correlation between dose to the pharyngeal constrictors and patient quality of life and late dysphagia following chemo-IMRT for head and neck cancer.* Radiother Oncol, 2009. **93**(3): p. 539-44.

25. Caudell, J.J., et al., *Dosimetric factors associated with long-term dysphagia after definitive radiotherapy for squamous cell carcinoma of the head and neck.* Int J Radiat Oncol Biol Phys, 2010. **76**(2): p. 403-9.

26. Dornfeld, K., et al., *Radiation doses to structures within and adjacent to the larynx are correlated with long-term diet- and speech-related quality of life.* Int J Radiat Oncol Biol Phys, 2007. **68**(3): p. 750-7.

27. Li, B., et al., *Clinical-dosimetric analysis of measures of dysphagia including gastrostomy-tube dependence among head and neck cancer patients treated definitively by intensity-modulated radiotherapy with concurrent chemotherapy.* Radiat Oncol, 2009. **4**: p. 52.

28. Christianen, M.E., et al., *Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.* Radiother Oncol, 2012. **105**(1): p. 107-14.

29. Mortensen, H.R., et al., *Late dysphagia after IMRT for head and neck cancer and correlation with dose-volume parameters.* Radiother Oncol, 2013. **107**(3): p. 288-94.

30. Mazzola, R., et al., *Dose-volume-related dysphagia after constrictor muscles definition in head and neck cancer intensity-modulated radiation treatment.* Br J Radiol, 2014. **87**(1044): p. 20140543.

31. Langendijk, J.A., et al., *Selection of patients for radiotherapy with protons aiming at reduction of side effects: The model-based approach.* Radiotherapy and Oncology, 2013. **107**(3): p. 267-273.

32. Widder, J., et al., *The Quest for Evidence for Proton Therapy: Model-Based Approach and Precision Medicine.* International Journal of Radiation Oncology*Biology*Physics.

33. van der Laan, H.P., et al., *The potential benefit of swallowing sparing intensity modulated radiotherapy to reduce swallowing dysfunction: an in silico planning comparative study.* Radiother Oncol, 2012. **103**(1): p. 76-81.

34. van der Laan, H.P., et al., *Swallowing-sparing intensity-modulated radiotherapy for head and neck cancer patients: treatment planning optimization and clinical introduction.* Radiother Oncol, 2013. **107**(3): p. 282-7.

35. Christianen, M.E., et al., *Swallowing sparing intensity modulated radiotherapy (SW-IMRT) in head and neck cancer: Clinical validation according to the model-based approach.* Radiother Oncol, 2015.

36. Christianen, M.E., et al., *Patterns of long-term swallowing dysfunction after definitive radiotherapy or chemoradiation.* Radiother Oncol, 2015. **117**(1): p. 139-44.

37. Gluck, I., et al., *Evaluating and reporting dysphagia in trials of chemoirradiation for head-and-neck cancer.* Int J Radiat Oncol Biol Phys, 2010. **77**(3): p. 727-33.

38. Russi, E.G., et al., *Swallowing dysfunction in head and neck cancer patients treated by radiotherapy: review and recommendations of the supportive task group of the Italian Association of Radiation Oncology.* Cancer Treat Rev, 2012. **38**(8): p. 1033-49.

39. Feng, F.Y., et al., *Intensity-modulated chemoradiotherapy aiming to reduce dysphagia in patients with oropharyngeal cancer: clinical and functional results.* J Clin Oncol, 2010. **28**(16): p. 2732-8.

40. Eisbruch, A., et al., *Chemo-IMRT of oropharyngeal cancer aiming to reduce dysphagia: swallowing organs late complication probabilities and dosimetric correlates.* Int J Radiat Oncol Biol Phys, 2011. **81**(3): p. e93-9.

41. Gregoire, V., et al., *Delineation of the neck node levels for head and neck tumors: a 2013 update. DAHANCA, EORTC, HKNPCSG, NCIC CTG, NCRI, RTOG, TROG consensus guidelines.* Radiother Oncol, 2014. **110**(1): p. 172-81.

42. Vainshtein, J.M., et al., *Long-term quality of life after swallowing and salivary-sparing chemo-intensity modulated radiation therapy in survivors of human papillomavirus-related oropharyngeal cancer.* Int J Radiat Oncol Biol Phys, 2015. **91**(5): p. 925-33.

43. McGowan, S.E., N.G. Burnet, and A.J. Lomax, *Treatment planning optimisation in proton therapy.* Br J Radiol, 2013. **86**(1021): p. 20120288.

44. Blanchard, P., et al., *Proton Therapy for Head and Neck Cancers.* Semin Radiat Oncol, 2018. **28**(1): p. 53-63.

45. Pedroni, E., et al., *The 200-MeV proton therapy project at the Paul Scherrer Institute: conceptual design and practical realization.* Med Phys, 1995. **22**(1): p. 37-53.

46. Holliday, E.B., et al., *Proton Therapy Reduces Treatment-Related Toxicities for Patients with Nasopharyngeal Cancer: A Case-Match Control Study of Intensity-Modulated Proton Therapy and Intensity-Modulated Photon Therapy.* International Journal of Particle Therapy, 2015. **2**(1): p. 19-28.

47. Gunn, G.B., et al., *Clinical Outcomes and Patterns of Disease Recurrence After Intensity Modulated Proton Therapy for Oropharyngeal Squamous Carcinoma.* Int J Radiat Oncol Biol Phys, 2016. **95**(1): p. 360-7.

48. Lukens, J.N., A. Lin, and S.M. Hahn, *Proton therapy for head and neck cancer.* Curr Opin Oncol, 2015. **27**(3): p. 165-71.

49. Blanchard, P., et al., *Intensity-modulated proton beam therapy (IMPT) versus intensity-modulated photon therapy (IMRT) for patients with oropharynx cancer - A case matched analysis.* Radiother Oncol, 2016. **120**(1): p. 48-55.

50. Sio, T.T., et al., *Intensity Modulated Proton Therapy Versus Intensity Modulated Photon Radiation Therapy for Oropharyngeal Cancer: First*

Comparative Results of Patient-Reported Outcomes. Int J Radiat Oncol Biol Phys, 2016. **95**(4): p. 1107-14.

51. Vinod, S.K., et al., *Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies.* Radiotherapy and Oncology, 2016. **121**(2): p. 169-179.

52. Mukesh, M., et al., *Interobserver variation in clinical target volume and organs at risk segmentation in post-parotidectomy radiotherapy: can segmentation protocols help?* Br J Radiol, 2012. **85**(1016): p. e530-6.

53. Brouwer, C.L., et al., *Differences in delineation guidelines for head and neck cancer result in inconsistent reported dose and corresponding NTCP.* Radiother Oncol, 2014. **111**(1): p. 148-52.

54. Geets, X., et al., *Inter-observer variability in the delineation of pharyngo-laryngeal tumor, parotid glands and cervical spinal cord: comparison between CT-scan and MRI.* Radiother Oncol, 2005. **77**(1): p. 25-31.

55. Nelms, B.E., et al., *Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer.* Int J Radiat Oncol Biol Phys, 2012. **82**(1): p. 368-78.

56. Feng, M., et al., *Normal tissue anatomy for oropharyngeal cancer: contouring variability and its impact on optimization.* Int J Radiat Oncol Biol Phys, 2012. **84**(2): p. e245-9.

57. Petkar, I., et al., *DARS: a phase III randomised multicentre study of dysphagia- optimised intensity- modulated radiotherapy (Do-IMRT) versus standard intensity- modulated radiotherapy (S-IMRT) in head and neck cancer.* BMC Cancer, 2016. **16**(1): p. 770.

2 Chapter 2: The potential benefits of dysphagia-optimised intensity-modulated radiotherapy to reduce long-term dysphagia in locally-advanced oropharyngeal cancer

2.1 Introduction:

As discussed in chapter 1, optimising long-term swallowing function following RT-based treatment is an important priority in OPC. This is particularly relevant in those with stage III/IV disease where extensive RT target volumes result in significant long-term morbidity. Though encouraging toxicity-mitigating RT strategies have emerged in recent years for this group of patients with locally advanced OPC (LA-OPC), the lack of randomised evidence precludes their routine implementation in clinical practice in the UK.

For OPC, a strong correlation exists between dose delivered to the PCM, particularly the superior constrictors which usually lie in close proximity to the tumour, and the risk of developing persistent RAD following treatment completion[1-3]. The priority, therefore, in LA-OPC is to reduce dose to the PCM to improve long-term function. Do-IMRT is a novel swallow-sparing RT planning technique that reduces the mean dose delivered to the PCM by minimising dose to that part of the muscle situated outside the high-dose RT volume. The definitive role of Do-IMRT is currently under evaluation in the DARS trial, a Cancer Research UK-funded national phase III study[4].

2.2 Aims of the study

This retrospective planning study was designed to compare the long-term swallow-sparing benefits of Do-IMRT with S-IMRT in LA-OPC. The ability of Do-IMRT to reduce the dose to the PCM, and consequently lower physician-scored and patient-reported NTCPs for late swallowing complication on available predictive toxicity models was investigated, and the magnitude of the gain was quantified. A second aim was to investigate the efficiency of delivery of complex Do-IMRT by analysing the dose delivered to non-swallowing OARs (NSW-OARs) and non-target normal tissue.

2.3 Null hypothesis

- There will be no difference in calculated NTCP for physician-reported RAD_{6M} between S-IMRT and Do-IMRT for LA-OPC.

2.4 Primary objective

- To determine, if using Do-IMRT in LA-OPC to limit RT dose delivered to the PCM, results in a ΔNTCP of $\geq$ 5% for physician-reported RAD_{6M}, where ΔNTCP is the difference between the risk of RAD_{6M} with S-IMRT ($NTCP_{S\text{-}IMRT}$) and Do-IMRT ($NTCP_{Do\text{-}IMRT}$).

A mean change in NTCP of $\geq$ 5% was chosen as it has previously been demonstrated to be the minimally clinically important difference for treatment-related toxicity in HNC[5, 6].

2.5 Primary endpoint

- The difference in mean NTCP for physician-reported RAD_{6M} with S-IMRT and Do-IMRT.

2.6 Secondary objectives

- To determine the risk of patient-scored RAD_{6M} with the two RT plans.
- To compare the differences in dose delivered to SW-OARs between the two RT planning techniques.
- To investigate the impact of Do-IMRT on dose delivered to (NSW-OARs).
- To investigate the impact of Do-IMRT on target volume coverage.
- To assess the impact of Do-IMRT on dose delivered to non-target normal tissue.
- To determine the complexity of delivering Do-IMRT.
- To evaluate intra-observer variability in PCM delineation.

2.7 Secondary endpoints

- Difference in mean NTCPs for patient-scored RAD_{6M}.
- Difference in RT dose delivered to SW-OARs and NSW-OARs.
- Difference in target volume coverage between the 2 RT techniques.
- Difference in integral RT dose between the 2 plans.
- Difference in monitor units and RT delivery times.
- Determine DSC values for PCM delineation.

2.8 Materials and Methods

2.8.1 Study design

This was a retrospective comparative IMRT planning study of patients with LA-OPC treated with bilateral neck irradiation at the Royal Marsden Hospital (RMH). Patients were identified from the INSIGHT study (Committee for Clinical Research number 3926), which was a prospective functional and molecular imaging study conducted at RMH between July 2013 and February 2015 in HNC receiving definitive treatment with primary CRT, with or without induction chemotherapy[7]. All patients in this research study had a pre-treatment magnetic resonance imaging (MRI) performed in the RT planning position. This improved confidence in the accuracy of tumour delineation within the context of volumetric outlining, thereby ensuring that subsequent predicted toxicity outcomes modelled on Do-IMRT plans were robust. Eligible patients had biopsy-proven squamous cell carcinoma of the tonsil, or BoT, and required bilateral neck irradiation. Tumours involving the posterior pharyngeal wall, and/or retropharyngeal lymph node involvement were excluded, as sufficient sparing of PCM required to meet the mandatory dose constraints of Do-IMRT would not be possible, as per DARS trial protocol. A total of twenty patients with LA-OPC (ten tonsil and ten BoT), meeting the pre-defined eligibility criteria described in section, were selected for this study. This sample size was considered to be representative of this tumour type based on previous RT planning studies in HNC[8-11]. For each patient, two RT plans were generated and compared: a S-IMRT plan, and a second Do-IMRT plan that was additionally optimised to reduce dose to PCM.

2.8.2 Demographic data

Demographic data was collected and included the following:

2.8.2.1 Patient demographics

- Age
- Sex

2.8.2.2 Tumour characteristics

- Location – tonsil, BoT
- Tumour (T) and nodal (N) stage
- p16 status – positive, negative, or unknown

2.8.2.3 Treatment

- Use of chemotherapy – induction, concomitant
- Radiotherapy – dose/fractionation

2.8.3 Radiotherapy

2.8.3.1 Planning procedure

All patients had been clinically treated with volumetric modulated arc therapy (VMAT) using 2-dose volumes. Patients were immobilised in a custom-made five-point thermoplastic shell. RT planning CT scans with contrast was performed with 2 – 2.5 mm slice thickness. Anatomical T_2-weighted and T_1-weighted MRI images, together with diffusion-weighted and dynamic contrast-

enhanced functional imaging sequences were acquired on a 1.5-T scanner; patients were set up on a flat-top couch in the RT treatment position, using the same thermoplastic shell described above. Clinical RT plans were produced using the available treatment planning systems (TPS) at the two sites of RMH. I generated the RT plans for this study on the research version of Raystation (RS) TPS (version 5.9.9, RaySearch Medical Laboratories, AB Stockholm, Sweden). All my plans were reviewed and approved by an experienced HN physicist (Alex Dunlop).

2.8.3.2 Target volume delineation

The planning MRI, and the CT with the contoured target volumes and OARs of this study population were exported to RS via a secure local computer network. I used a volumetric 2-dose level approach to define my target volumes. As the anatomical approach was used clinically to define the high-dose RT volume for all patients, I edited the contoured target volumes as follows:

Gross tumour volume (GTV):

GTV was defined as the visible primary tumour and involved lymph nodes on the CT and MRI. The tumour was delineated on T_2-weighted MRI images with reference to T_1-weighted images. In patients who received induction chemotherapy, the GTV included the pre-chemotherapy primary tumour volume. If the involved LN was no longer visible following IC, the entire lymph node level was included in the high-dose RT volume.

Clinical target volume (CTV):

CTV65:

The GTV was expanded with a 1 cm isotropic margin to create the high-dose CTV65, which received a dose of 65 Gy. This CTV was edited to exclude natural barriers to disease spread such as bone, muscle and air, unless there was radiological evidence of tumour infiltration.

CTV54:

The elective dose CTV54 received a dose of 54 Gy and included the remainder of the oropharynx along with the nodal levels at risk of microscopic disease as described below.

Oropharynx:

The oropharyngeal compartment for delineation was defined as follows:

- Superior: Cranial aspect of soft palate
- Inferior: Cranial aspect of hyoid bone for tonsil tumours; caudal edge of hyoid for BoT tumours
- Lateral: To include bilateral parapharyngeal space
- Anterior: Posterior one-third of tongue
- Posterior: Pharyngeal mucosa

Lymph node (LN) levels:

LN levels were contoured as defined by the 2013 DAHANCA consensus guidelines[12]. Elective nodal volume selection for this study were as per

recommendations from recruiting UK national phase III trials[4, 13], and were defined as follows:

- Node-negative side of neck: levels II – IVa
- Node-positive side of neck: Ib-IVa, Vab and VIIa

Planning target volume (PTV)

CTV65 and CTV54 were grown isotropically by 3 mm to generate uncropped PlanPTV65 and PlanPTV54 respectively. For final dose reporting, PlanPTV65 was edited off external body contour by 5 mm to generate PTV65, which received a dose of 65 Gy. Similarly, PlanPTV54 was cropped from external body contour, and PTV65 to create PTV54, which received a dose of 54 Gy.

2.8.3.3 OAR delineation

SW-OARs:

Outlining for the PCM and larynx, the other variable that influences NTCP calculations, was based on the published contouring guidelines defined by Christianen et al.[14], in conjunction with the atlas produced for the Post-operative adjuvant treatment for HPV positive tumours (PATHOS; NCT02215265) trial[15] (Table 2.1). A thickness of 3 mm was used to contour the constrictor muscles.

Organ at risk	Anatomic borders					
	Cranial	Caudal	Anterior	Posterior	Lateral	Medial
Superior PCM	Caudal tip of pterygoid plates (hamulus)	Lower edge of C2	Hamulus of pterygoid plate, base of tongue, pharyngeal lumen	Prevertebral muscle	Medial pterygoid muscle	Pharyngeal lumen
Middle PCM	Upper edge of C3	Lower edge of hyoid bone	Base of tongue, hyoid bone	Prevertebral muscle	Greater horn of hyoid bone	Pharyngeal lumen
Inferior PCM	First slice caudal to lower edge of hyoid bone	Lower edge of arytenoid cartilages	Soft tissue of supraglottis/glottis	Prevertebral muscle	Superior horn of thyroid cartilage	
Supraglottic larynx	Tip of epiglottis	First slice cranial to upper edge of arytenoid cartilage	Hyoid bone, pre-epiglottic space, thyroid cartilage	Pharyngeal lumen, inferior PCM	Thyroid cartilage, parapharyngeal space	Pharyngeal lumen (lumen excluded)
Glottic larynx	Upper edge of arytenoid cartilage	Lower edge of cricoid	Thyroid cartilage	Pharyngeal lumen, inferior PCM, cricoid cartilage	Thyroid cartilage	Pharyngeal lumen (lumen excluded)

Table 2.1 Delineation guidelines for swallowing organs at risk (reproduced from Christianen et al)

PCM, pharyngeal constrictor muscle

For the purpose of Do-IMRT planning optimization, the superior and middle pharyngeal constrictors were combined to create superior and middle PCM (SMPCM), while the inferior pharyngeal constrictor muscle (IPCM) was defined as a separate structure. PlanSMPCM and PlanIPCM structure sets were additionally generated by cropping SMPCM and IPCM away from CTV65 (Figure 2.1).

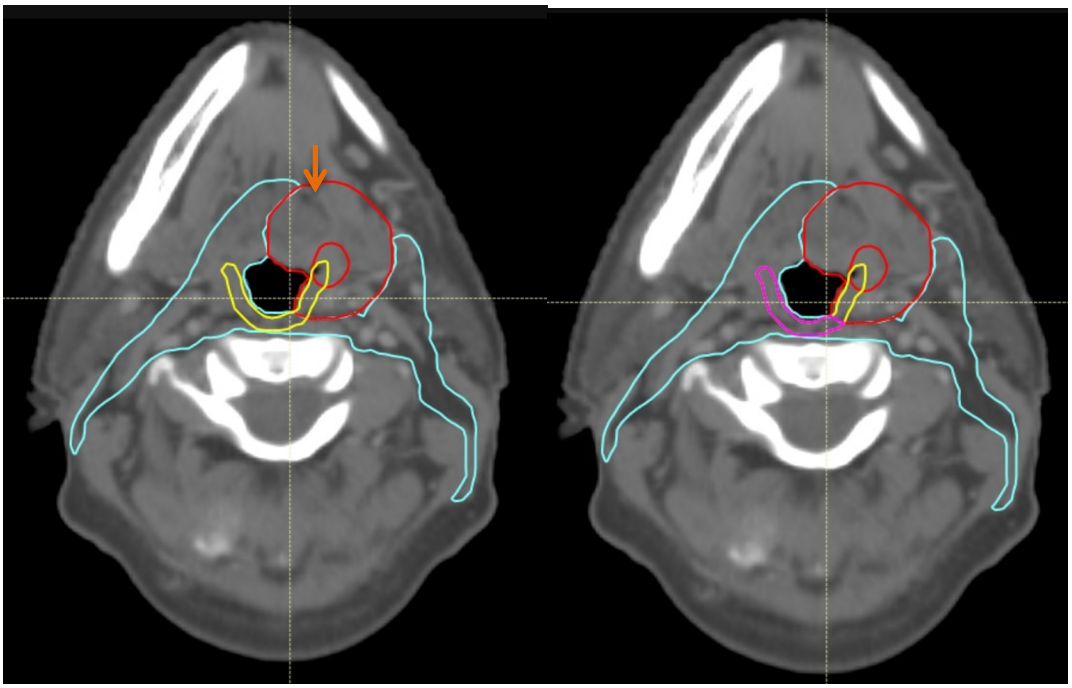


Figure 2.1 SW-OAR delineation for Do-IMRT

Left: Axial CT scan showing GTV (red), CTV65 for radical dose radiotherapy (red with arrow), CTV54 for prophylactic dose radiotherapy (cyan), superior and middle pharyngeal constrictor (SMPCM) (yellow), Right: SMPCM lying outside CTV65 edited to create PlanSMPCM (pink)
CTV, clinical target volume; GTV, gross tumour volume

The percentage volume overlap between PTV65 and various swallowing structures of interest (PTV_{OL_M}) were computed as follows:

$$PTV_{OL\_M} = \left(\frac{M \cap PTV65}{M}\right) \times 100$$; where:

M: represents volume of PCM or an individual muscle of PCM (SPC, MPC, or IPC), or volume of SGL

$M \cap PTV65$: Volume of intersection of M and PTV65

NSW-OARs

The parotids, SC, and BS had been previously delineated. These were reviewed and edited where necessary to ensure consistency with published OAR delineation guidelines[16]. For analysis, the parotids were labelled as ipsilateral or contralateral parotid depending upon the location of the primary tumour.

SC and BS planning at risk volumes (PRVs) were generated by a 3 mm isotropic expansion of the SC and BS.

2.8.3.4 IMRT plans

Plan design

Both S-IMRT and Do-IMRT plans were generated and optimised using the collapsed cone v3.4 algorithm in RS.

Beam design

VMAT was preferentially selected over static step and shoot IMRT to generate the RT plans, as an earlier study had shown that acceptable Do-IMRT plans were likely to be more difficult to achieve with static IMRT[17]. All RT plans consisted of two 6-MV 360° arcs with mirrored collimator angles of 30° and 330° respectively. Each arc consisted of 180 control points with 2° control point spacing, leaf motion constraint of 0.8 cm/degree, and a maximum delivery time of 120 seconds.

Beam optimisation

Non-anatomical volumes were added to guide the TPS to provide an optimal plan. These included ring structures around the PTVs to increase dose conformality and homogeneity, together with the application of the dose fall-off optimising function to reduce dose to non-target tissue. This function penalises dose that exceeds the specified dose fall-off away from the target and is defined by 3 parameters: the high-dose parameter representing the high-dose within the target; the low-dose parameter representing the acceptable dose within the surrounding healthy tissue; and the distance parameter representing the distance at which the high dose will have fallen to the low dose. For Do-

IMRT plans, additional PTV54 dummy volumes based on PTV54 and cropped from PTV65 by 1 cm and the constrictors by 5 mm were created to allow a dose gradient to achieve acceptable 54 Gy dose coverage. The PCM overlap usually with both PTVs and coverage was compromised to PTV54 in the region of PlanSMPCM and PlanIPCM respectively to meet planning goals, but no under-dosage was allowed to PTV65.

In order to avoid the TPS boosting target tissue within the build-up region of the patient, virtual bolus (density = 1.0 g/cm^3) was generated exterior to the body contour such that PlanPTV65 and PlanPTV54 were always at least 1 cm from either the external body contour or the virtual bolus surface[18]. Once a satisfactory plan was generated, virtual bolus was removed, the dose re-prescribed to PTV65, and the final dose was computed. The RT plan was re-evaluated to ensure optimal coverage was maintained.

Planning objectives

For each plan, baseline dose-based and dose-volume- based objectives were designed for the inverse planning optimisation process, in order to satisfy clinical planning goals (Table 2.2 and 2.3). The optimal DVH was obtained by modifying the objectives, and/or increasing the priorities for each objective during optimisation. Each optimisation sequence consisted of 60 automated iterations.

The planning goals were prioritised in the following order:

- SC and BS constraints
- PTV65 coverage
- PlanSMPCM and PlanIPCM constraints (in Do-IMRT arm only)
- PTV54 coverage
- Parotids constraints

Volume (%)	Dose (%)		
	PTV65	**S-IMRT: PTV54**	**Do-IMRT: PTV54**
99	> 90	> 90	As high as possible
98	(> 95 optimal)	(> 95 optimal)	As high as possible
95	> 95	> 95	As high as possible
50	= 100	= 100 (±1Gray)	= 100 (±1Gray)
5	< 105	As low as possible	As low as possible
2	< 107	As low as possible	As low as possible

Table 2.2 Planning target volume (PTV) constraints for standard IMRT (S-IMRT) and dysphagia-optimised IMRT (Do-IMRT)

Structure	Constraint	Mandatory Dose constraint	Optimal Dose constraint
Spinal cord	Max	< 48 Gy	
	1cm^3	< 46 Gy	
Spinal cord PRV	1cm^3	< 48 Gy	
Brainstem	Max	< 55 Gy	
	1cm^3	< 54 Gy	
Brainstem PRV	1cm^3	< 55 Gy	
Contralateral Parotid	Mean dose	as low as possible	< 24 Gy
Ipsilateral Parotid	Mean dose	as low as possible	< 24 Gy
PlanSMPCM	Mean dose	< 50Gy*	
PlanIPCM	Mean dose		< 20Gy*

Table 2.3 Planning target volume (PTV) constraints for standard IMRT (S-IMRT) and dysphagia-optimised IMRT (Do-IMRT)

IPCM, inferior pharyngeal constrictor muscle; PRV, planning organ at risk volume; SMPCM, superior and middle pharyngeal constrictor muscle
* - Do-IMRT arm only

Dose prescription

In accordance with ICRU 83, prescription was to the median dose point on the DVH such that the prescription dose of 65 Gy was received by 50% of PTV65.

Intra-observer variation

I re-contoured the SMPCM and IPCM structures one month after first delineation on 5 randomly chosen datasets from the study group. DSC was used to establish the level of agreement between the contours.

2.8.3.5 Plan evaluation

PTV coverage

DVHs for each plan were generated and clinical goals recorded. To quantify the dose conformality and homogeneity in PTV65, the conformity index (CI) and homogeneity index (HI) were defined as follows:

- CI = (cover factor) x (spill factor)

$$= \left(\frac{V_{95}(\text{PTV65})}{V_{\text{PTV65}}}\right) \text{ x } \left(\frac{V_{95}(\text{PTV65})}{V_{95}(\text{body})}\right)$$

where V_{95}(PTV65) and V_{95}(body) are the volumes of the PTV65 and body receiving at least 95 % of the prescription dose respectively, and V_{PTV} is the volume of the PTV65[19].

CI = 1 represents a perfectly conformal plan, and smaller values denote that the dose distribution conforms less to the target shape and size.

- $HI = \left(\frac{D2\ \% - D98\ \%}{Dmedian}\right)$

where D2% and D98% are the doses delivered to 2 % and 98 % of PTV65, respectively.

A perfectly homogeneous plan would correspond to HI = 0. Higher values for represent less homogeneous dose distribution within PTV65.

Dose delivered to OARs

SW- OARs

Mean RT dose delivered to the PCM, as well as to individual constrictor muscle for each patient with the 2 RT plans were recorded. In addition, mean dose to the SGL was also generated for analysis.

NSW-OARs

The clinical goals for SC, BS, and both parotids were tabulated, and analysed.

Integral dose to non-target normal tissue

The integral dose is helpful in evaluating the low dose bath for a given RT plan, and is defined as the mean dose delivered to an organ multiplied by the volume of that organ[20]. Two structures were defined to ascertain firstly, whether Do-IMRT increased the integral dose to non-target tissue compared to S-IMRT; and secondly, to establish the distribution of such dose spillage.

1. Normal Tissue (NT): This was defined as the external contour of the patient minus the PTV, and reflects the integral dose delivered to the patient outside the target volumes.

2. Remaining volume at risk (RVR): RVR was defined as the difference between the volumes enclosed by the external contour of the patient and that of the target volumes and OARs, and evaluates integral dose distribution to the soft tissues within the patient.

2.8.4 NTCP calculation for RAD_{6M}

$NTCP_{S\text{-}IMRT}$ and $NTCP_{Do\text{-}IMRT}$ for each patient was determined by applying the predictive models of Christianen *et al* discussed in chapter 1[6, 21-23]. At present, this toxicity model is the most widely used for estimating the risk of RAD[10, 24]. The four NTCP models were based on a physician-scored toxicity, **the primary endpoint of this study**, and three patient-reported endpoints at 6 months following treatment completion respectively as stated below

- Physician-scored RTOG $\geq$ grade 2 dysphagia.
- Patient-reported moderate to severe difficulties with swallowing solids.
- Patient-reported moderate to severe difficulties with swallowing soft food.
- Patient-reported moderate to severe difficulties with swallowing liquids.

The patient-reported toxicity models were developed by the Dutch group using 3 swallowing-related questions in the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Head and Neck Module (EORTC QLQ-H&N35) questionnaire.

The equation for calculating the probability of late swallowing toxicity was defined as

$$\mathbf{NTCP = (1 + e^{-S})^{-1}}$$

The factors influencing the s-value for the various NTCP models are shown in table 2.4.

RAD_{6M} endpoint	**s-value**
RTOG grade 2-4	-6.09 + (mean dose SPC × 0.057) + (mean dose supraglottic larynx × 0.037)
Liquid food	-5.98 + (mean dose supraglottic larynx × 0.074)+ (radiation technique × -1.209)
Solid food	-6.89 + (mean dose SPC × 0.049)+ (mean dose supraglottic larynx × 0.048) + (age × 0.7
Soft food	-5.83 + (mean dose MPC × 0.061)+(age ×1.203)+ (tumour site ×1.122) +(radiation technique × -0.912)

Table 2.4 Parameters that determine the s-value for the different normal tissue complication probabilities (NTCP) models

RAD_{6M}, radiation-associated dysphagia at 6 months; RTOG, radiation therapy oncology group; SPC, superior pharyngeal constrictor

To facilitate comparisons with published toxicity outcomes, the dose distributions (65 Gy/ 30 fractions) were converted to their iso-effective dose in 2 Gy fraction equivalents (EQD2), using α/β ratio of 3 for normal tissues. The

first step was to convert the cumulative DVHs exported from RS, as they were equal bin widths of volume rather than dose. Alex Dunlop, our HN physicist, wrote an excel script that changed the above DVHs into a cumulative DVH with 1 Gy dose bins. I entered the data from RS on this excel spreadsheet to generate the new cumulative DVHs for both Do-IMRT and S-IMRT in each patient. Separate folders were created for each DVH in every patient, and individual folders were successively copied on to a script created by Sarah Gulliford (see acknowledgement) on the statistical programme 'R' to generate the EQD2.

A separate excel spreadsheet for each NTCP model using the above equation was created, and the EQD2 mean doses to the variables obtained from the DVHs of S-IMRT and Do-IMRT plans of each patient was incorporated into it to obtain the corresponding NTCP values for all twenty patients. These values were subsequently analysed as described in the next section.

2.8.5 Statistical analysis

Statistical analysis was performed on Statistical Package for the Social Sciences version 25 (SPSSv25). Data collected from the DVHs of individual plans were tabulated for each patient, and NTCP calculated. The differences between the data of interests (eg mean dose to PCM, maximum dose to SC, or NTCP values) between the two planning techniques were shown to exhibit a normal distribution, assessed by the Shapiro-Wilk method. The significance of the difference between the two treatment techniques were therefore analysed using the parametric two-tailed paired Student's t-test. This was used as each patient had 2 plans, S-IMRT and Do-IMRT, and a paired t-test can analyse the mean difference for a volume of interest in such a paired sample where the individual differences are normally distributed. Differences were reported as statistically significant at the $p < 0.05$ level.

For patient sub-group analysis, the reduction in the risk of RAD_{6M} with Do-IMRT, relative to S-IMRT was determined for each group. Analysis of variances (ANOVA) or independent-samples t-test was subsequently performed on this, depending upon the number of patient sub-groups.

2.9 Results

2.9.1 Intra-observer variation

There was good agreement for both SMPCM and IPCM contours, with DICE scores of 0.76 (range 0.74 – 0.79), and 0.77 (range 0.75 – 0.80) respectively.

2.9.2 Baseline characteristics

Patient, tumour site, and target volumes characteristics are summarised in table 2.5. The two tumour sub-sites had comparable GTVs. BoT tumours had higher T-stages (50 % T3/4) compared to tonsillar tumours (30 %). The median PlanSMPCM volume was 9.1 cc^3 (range 6.2 cc^3 – 10.8 cc^3); this translated into a median of 75.4 % (range 51.7 % - 99.1 %) of SMPCM that was available in this cohort for sparing of RT dose during Do-IMRT optimisation. The overlap between the swallowing structures and high-dose PTV in the entire cohort was predominantly in the region of the SPC (median 38 %, range 5 % - 75 %), compared to MPC (13 %, 0% - 55 %). The median PTV65 – SPC overlap for BoT and tonsil tumours were 26 % (range 5 - 47) and 41 % (range 25 - 75) respectively. There was no overlap between IPCM and PTV65 in the study group. The median proportion of the combined SPC and SGL volumes, the two parameters that determine the primary endpoint of this study, overlapping with PTV65 was 19.4 % (range 7.6 % - 53.0 %).

Characteristics	Results
Age in years	62 (43 – 69)
Sex	
Male	20 (100)
Tumour site	
Base of tongue	10 (50)
Tonsil	10 (50)
Tumour stage	
T1	1 (5)
T2	11 (55)
T3	5 (25)
T4	3 (15)
Nodal stage	
N0	2 (10)
N1	2 (10)
N2a	2 (10)
N2b	12 (60)
N2c	2 (10)
AJCC stage (7^{th} edition)	
III	4 (20)
IVa	16 (80)
p16 status	
Positive	14 (70)
Negative	5 (25)
Unknown	1 (5)
Treatment modality	
CRT	13 (65)
IC + CRT	7 (35)
Gross Tumour Volume (cm^3)	16.4 (5.2 – 56.6)
PTV65 (cm^3)	160.8 (91.2 – 283.2)
PTV54 (cm^3)	394.3 (315 – 486)

Table 2.5 Baseline patient characteristics

AJCC, American Joint Committee on Cancer; CRT, chemoradiation; IC, induction chemotherapy; PTV, planning target volume
Results are number (percentage) or median (range)

2.9.3 Dose comparisons for PTVs

The mean (standard deviation (SD)) values of the PTV dose parameters, CI, and HI for S-IMRT and Do-IMRT plans are summarized in table 2.6. Both techniques produced clinically acceptable plans that achieved the PTV planning objectives in all patients. Small, but statistically significant differences were observed for all PTV65 dose objectives and HI in the Do-IMRT plans. Dose conformity for PTV65 was similar between the two plans. For Do-IMRT, the only mandatory constraint for PTV54 was the median dose, and this was achieved in all cases. An example of the dose distribution with S-IMRT and Do-IMRT for a patient is shown in fig 2.2.

Objective	S-IMRT Mean (SD)	Do-IMRT Mean (SD)	p-value
PTV65			
D99% ≥ 90%	62.1 (0.3)	62.3 (0.3)	< 0.001
D98% ≥ 95%	62.5 (0.2)	62.8 (0.3)	< 0.001
D95% ≥ 95%	63.1 (0.2)	63.6 (0.3)	< 0.001
D50% = 100%	65.1	65.1	
D5% ≤ 105%	66.9 (0.3)	66.3 (0.3)	< 0.001
D2% ≤ 107%	67.3 (0.3)	66.6 (0.3)	< 0.001
PTV54			
D99% ≥ 90%	51.3 (0.3)	43.9 (2.2)	< 0.001
D98% ≥ 95%	51.6 (0.2)	45.8 (2.1)	< 0.001
D95% ≥ 95%	52.1 (0.1)	49.2 (1.3)	< 0.001
D50% = 100%	53.8 (0.2)	54.1 (0.3)	< 0.001
D5%	61.4 (1.0)	61.5 (0.8)	0.29
D2%	63.1 (0.7)	63.3 (0.4)	0.01
CI	0.79 (0.03)	0.79(0.03)	0.81
HI	0.93 (0.01)	0.94 (0.01)	< 0.001

Table 2.6 PTV coverage with S-IMRT and Do-IMRT

CI, conformity index; Do-IMRT, dysphagia-optimised intensity modulated radiotherapy; Dx%, dose to x% volume; HI, homogeneity index, PTV, planning target volume; S-IMRT, standard IMRT; SD, standard deviation

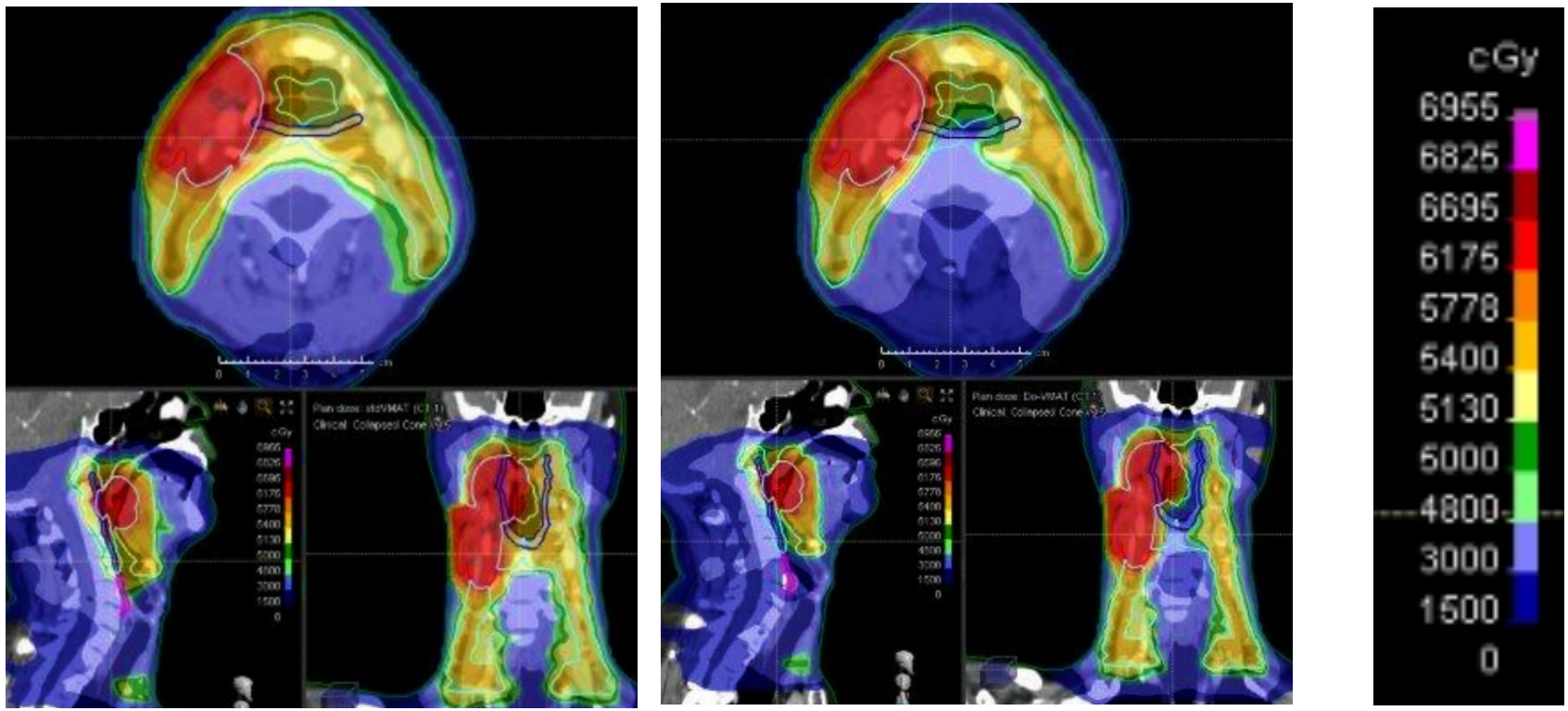


Figure 2.2 Illustration of dose distribution with S-IMRT and Do-IMRT

Dose distributions on the same computerised tomography scan slice with standard IMRT (S-IMRT) (left), and dysphagia-optimised IMRT (Do-IMRT) (right). Planning target volume PTV65, red colourwash; PTV54, brown colourwash. PTV54 coverage, but not PTV65, is compromised in the region of SMPCM (blue) to meet planning objectives of Do-IMRT

2.9.4 Dose comparisons for SW-OARs

The mean dose (SD) of the SW-OARs, including those structure sets that did not have specified RT planning objective for optimisation, are summarised in table 2.7. The mandatory PlanSMPCM dose constraint was achieved in all twenty patients, while the optimal PlanIPCM dose objective was not met in any patient. Statistically significant differences were noted for all SW-OARs, in favour of Do-IMRT plan. The mean dose delivered to the SPC and PCM with Do-IMRT was 8.1 % (SD 3.0; range 1.9 – 13.3) and 16.5 % (SD 2.5; range 12.0 – 22.2) lower than with S-IMRT respectively. The largest absolute dosimetric benefit was observed for the IPC, where a mean difference of 22.7 Gy (SD 4.1; range 8.4 – 27.7) was seen between the two planning techniques. Significantly lower doses were also delivered to the larynx in the Do-IMRT arm, despite no active efforts to spare it during the optimisation process.

Structure	S-IMRT		Do-IMRT		Reduction in dose between mean S-IMRT and Do-IMRT (Gy)	p-value
	Mean (SD)	Range	Mean (SD)	Range		
Plan SMPCM*	56.3 (1.1)	54.9 – 58.4	49.4 (0.3)	48.9 – 49.9	6.9	< 0.001
Plan IPCM*	48.6 (3.4)	43.3 – 53.9	24.9 (3.6)	20.7 – 32.1	23.7	< 0.001
PCM	56.3 (1.2)	54.5 – 58.7	47.0 (1.9)	42.4 – 50.6	9.3	< 0.001
SPC	58.7 (2.0)	55.6 – 63.8	54.0 (3.4)	48.4 – 62.6	4.7	< 0.001
MPC	57.2 (3.1)	51.7 – 62.6	49.8 (6.4)	37.9 – 59.6	7.4	< 0.001
IPCM	49.4 (3.5)	43.3 – 54.3	26.7 (6.1)	21.1 – 45.5	22.7	< 0.001
SGL	52.4 (6.1)	44.0 – 62.5	49.8 (7.8)	36.6 – 62.8	2.6	< 0.001
GSL	49.7 (5.3)	43.1 – 60.1	46.6 (7.0)	35.4 – 60.2	3.1	< 0.001

Table 2.7 Comparison swallowing organs at risk mean dose parameters (standard deviation (SD)) for the study cohort for standard intensity modulated radiotherapy (S-IMRT) and dysphagia-optimised IMRT (Do-IMRT)

GSL, glottis and supraglottic larynx (SGL); IPCM, inferior pharyngeal constrictor muscle; MPC, middle pharyngeal constrictor; PCM, pharyngeal constrictor muscle; SMPCM, superior and middle pharyngeal constrictor muscle; SPC, superior pharyngeal constrictor muscle

* structure with dose constraint in Do-IMRT

The extent of overlap of SMPCM with PTV65 did not appear to influence the reduction in dose to this structure with Do-IMRT relative to S-IMRT, as illustrated in Fig 2.3 by the lack of a linear relationship between the two.

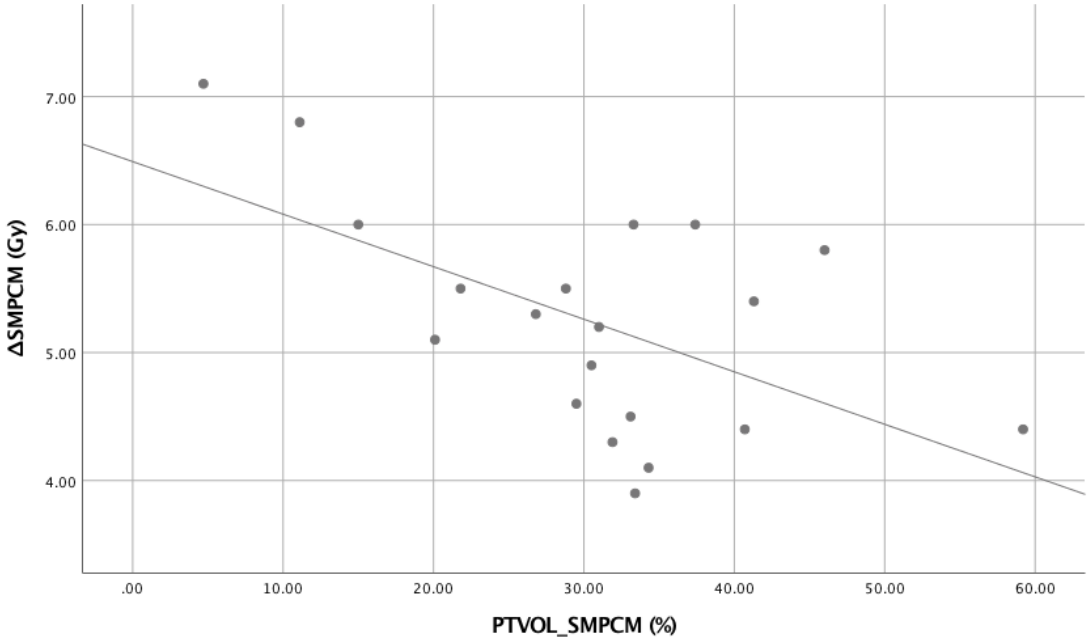


Figure 2.3 Scatter plot showing relationship between the mean dose reduction in superior and middle pharyngeal constrictor muscle (ΔSMPCM) between the two radiotherapy techniques and the percentage volume overlap of SMPCM with planning target volume (PTV) 65 (PTVOL_SMPCM)

2.9.5 Dose comparisons for NSW-OARs

The mean (SD) values of the NSW-OARs dose parameters for both plans are outlined in table 2.8. Statistically significant differences were noted in all the dose objectives for SC, BS, and mean dose to the ipsilateral parotid gland, with lower dose delivered in the Do-IMRT plan. No significant difference was seen in the contralateral parotid mean dose between the two techniques.

Structure	S-IMRT Gy (SD)	Do-IMRT Gy (SD)	p - value
Spinal cord			
Dmax	43.6 (1.7)	41.2 (2.7)	0.003
Dmax to 1 cm^3	39.2 (1.6)	35.9 (1.8)	< 0.001
Dmax to PRV 1 cm^3	42.1 (1.5)	38.8 (2.3)	< 0.001
Brainstem			
Dmax	45.0 (2.0)	41.9 (2.7)	< 0.001
Dmax to 1 cm^3	39.6 (2.7)	36.3 (2.3)	< 0.001
Dmax to PRV 1 cm^3	42.4 (2.1)	39.0 (2.5)	< 0.001
Contralateral parotid			
Mean dose	26.6 (3.4)	25.8 (1.9)	0.14
Ipsilateral parotid			
Mean dose	40.4 (4.5)	37.4 (2.8)	< 0.001
Contralateral submandibular gland			
Mean dose	51.9 (4.8)	50.7 (5.8)	0.004
Ipsilateral submandibular gland			
Mean dose	63.3 (2.8)	63.2 (3.0)	0.56

Table 2.8 Comparison swallowing organs at risk mean dose parameters (standard deviation (SD)) for the study cohort for standard intensity modulated radiotherapy (S-IMRT) and dysphagia-optimised IMRT (Do-IMRT)

PRV, planning organ at risk volume; SD, standard deviation

2.9.6 Integral RT dose

Both mean $NT_{Do\text{-}IMRT}$ (130.5 Gy-litre (Gy-L), 95 % CI 122.8 – 138.3), and $RVR_{Do\text{-}IMRT}$ (101.7 Gy-L, 95 % CI 94.6 – 108.9) were significantly lower than the corresponding mean $NT_{S\text{-}IMRT}$ (138.8 Gy-L, 95 % CI 130.1 – 147.5; p = 0.00) and $RVR_{S\text{-}IMRT}$ (108.9 Gy-L, 95 % CI 101.0 – 116.8; p = 0.00) doses respectively. Compared to S-IMRT, the integral doses to NT, and RVR were 5.9 % and 6.7 % lower with Do-IMRT. An example of the integral dose deposition for the 2 RT plans is shown in Fig 2.4.

There was a statistically significant increase in the number of monitor units required to deliver Do-IMRT plans; delivery time between the 2 plans were similar.

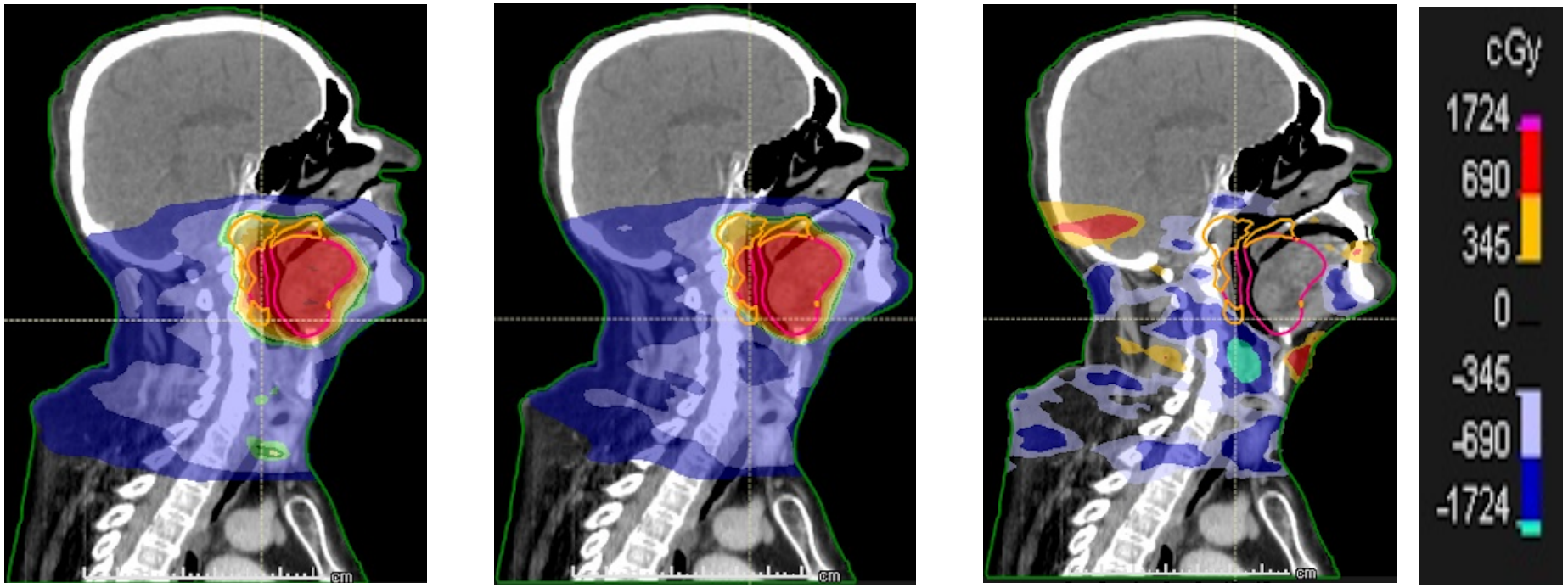


Figure 2.4 Integral dose deposition with S-IMRT and Do-IMRT

Representative saggital plans for the same patient with LA-OPC. (left) The dose distribution of standard IMRT (S-IMRT) plan; (middle) shows dose distribution of the comparative dysphagia-optimised IMRT (Do-IMRT) plan; (right) The result of a dose subtraction to demonstrate where unnecessary dose is avoided with Do-IMRT, relative to S-IMRT.

2.9.7 NTCP of late swallowing dysfunction

2.9.7.1 Physician-scored RTOG grade 2-4 RAD_{6M}

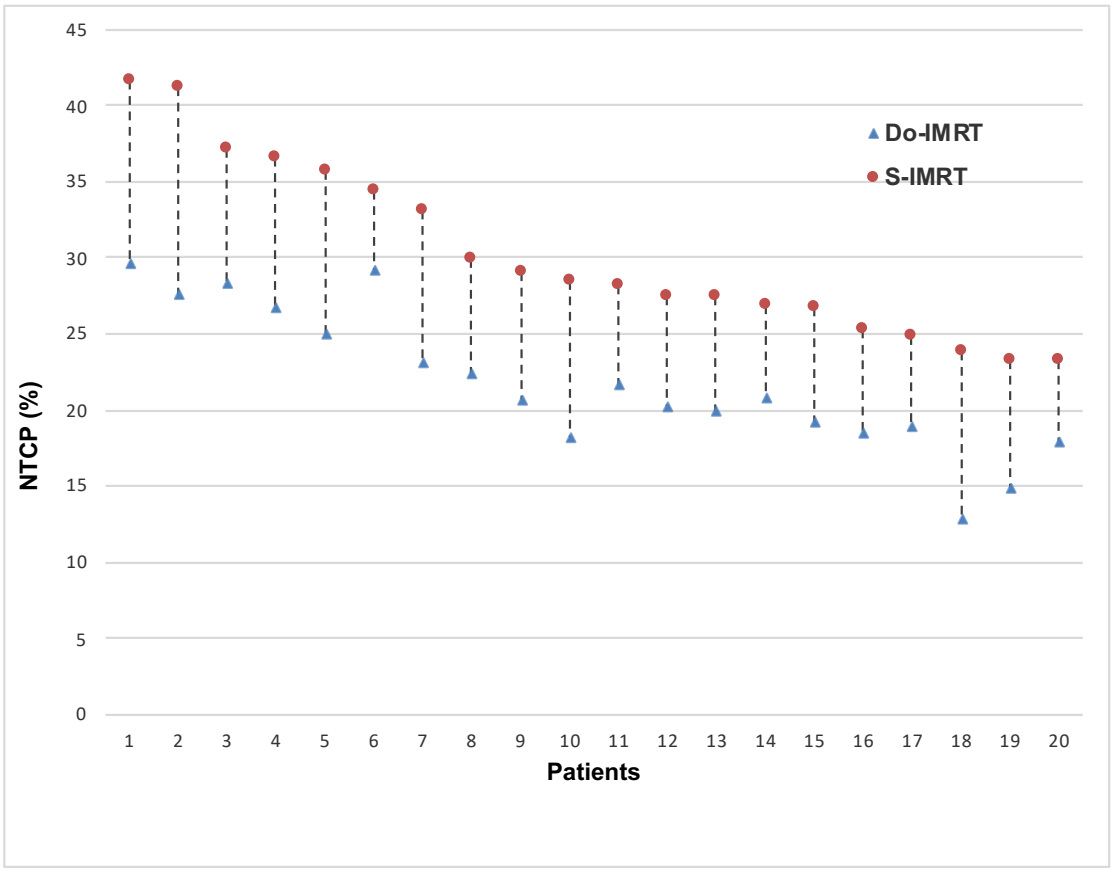


Figure 2.5 Normal tissue complication probabilities (NTCPs) of physician-scored Radiation Therapy Oncology Group (RTOG) grade 2-4 swallowing dysfunction with standard IMRT (S-IMRT) and dysphagia-optimised IMRT (Do-IMRT) (patients re-sorted as per the S-IMRT NTCP values)

The NTCP results for the study primary endpoint with both RT planning techniques for each patient are provided in figure 2.5. Mean ΔNTCP was 8.4 % (SD 2.2; range 5.2 – 13.6), and achieved the pre-defined threshold of $\geq$ 5 % reduction in toxicity with Do-IMRT in all patients. The mean $NTCP_{Do\text{-}IMRT}$ was 21.8 % (SD 4.6; range 12.9 – 29.6), which was significantly lower than the

mean $NTCP_{S\text{-}IMRT}$ of 30.2 % (SD 5.6; range 23.2 – 41.6); $p < 0.001$. The absolute NTCP values for the two planning techniques, together with the absolute NTCP gain with Do-IMRT, varied between individual patients.

Fig 2.6 illustrates the influence of ΔSPC (mean $SPC_{S\text{-}IMRT}$ – mean $SPC_{Do\text{-}IMRT}$) and ΔSGL (mean $SGL_{S\text{-}IMRT}$ – mean $SGL_{Do\text{-}IMRT}$) on ΔNTCP in the study group. ΔSPC was larger than ΔSGL in 75 % of patients.

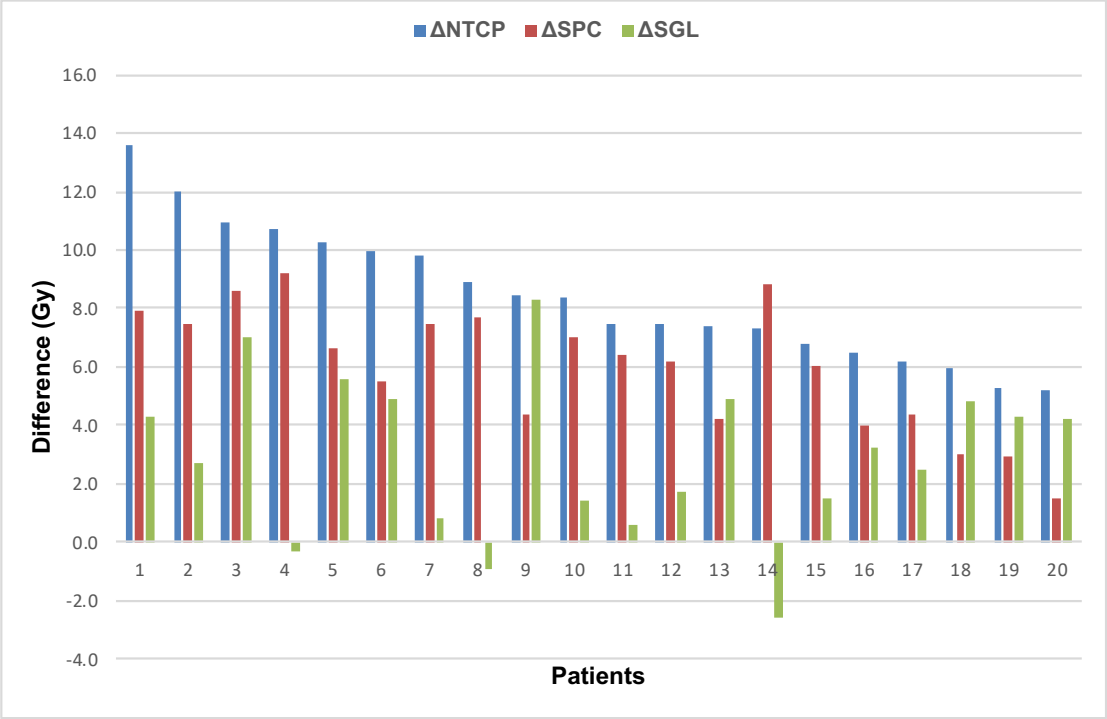


Figure 2.6 ΔNTCP for physician-rated radiation therapy oncology group (RTOG) grade 2-4 dysphagia at 6 months, ΔSPC, and ΔSGL for each patient (patients resorted as per ΔNTCP values)

NTCP, normal tissue complication probability; SGL, supraglottic larynx; SPC, superior pharyngeal constrictor

On sub-group analysis (Table 2.9), larger ΔNTCP were observed for BoT tumours, though this was not statistically significant. Likewise, advanced T stage, higher nodal staging, or the use of induction chemotherapy did not appear to significantly impact on the toxicity-mitigating benefits of Do-IMRT in this study group.

		ΔNTCP for physician-rated RTOG grade ≥ 2 dysphagia	Independent t-test
N = 20	N	Mean (Standard deviation)	p-value
Tumour site			0.19
BoT	10	9.1 (1.9)	
Tonsil	10	7.8 (2.5)	
T stage			0.20
T1/2	12	7.9 (2.0)	
T3/4	8	9.3 (2.6)	
Nodal stage			0.92
N0-2a	6	8.4 (3.2)	
N2b/c	14	8.5 (1.9)	
Treatment modality			0.52
IC + CRT	7	8.0 (1.6)	
CRT	13	8.7 (2.6)	

Table 2.9 Normal tissue complication probabilities (NTCP) for physician-scored dysphagia according to tumour site, stage, nodal stage and treatment modality

CRT, chemoradiation; IC, induction chemotherapy

2.9.7.2 Patient-reported moderate-severe RAD_{6M}

The NTCPs for the three patient-reported swallowing toxicities are presented in Figures 2.7 – 2.9. The mean NTCP reductions with Do-IMRT for patient-reported moderate to severe swallowing issues were: 6.4 % (range 3.1 % - 11.6 %) for solid food, 4.4 % (range 1.8 % - 9.5 %) for soft food, and 0.5 % (range – 0.8 % - 1.6 %) for liquid food. 60 % of patients had at least 5% reduction in predicted toxicities with Do-IMRT compared to S-IMRT, with respect to difficulties in swallowing solid food.

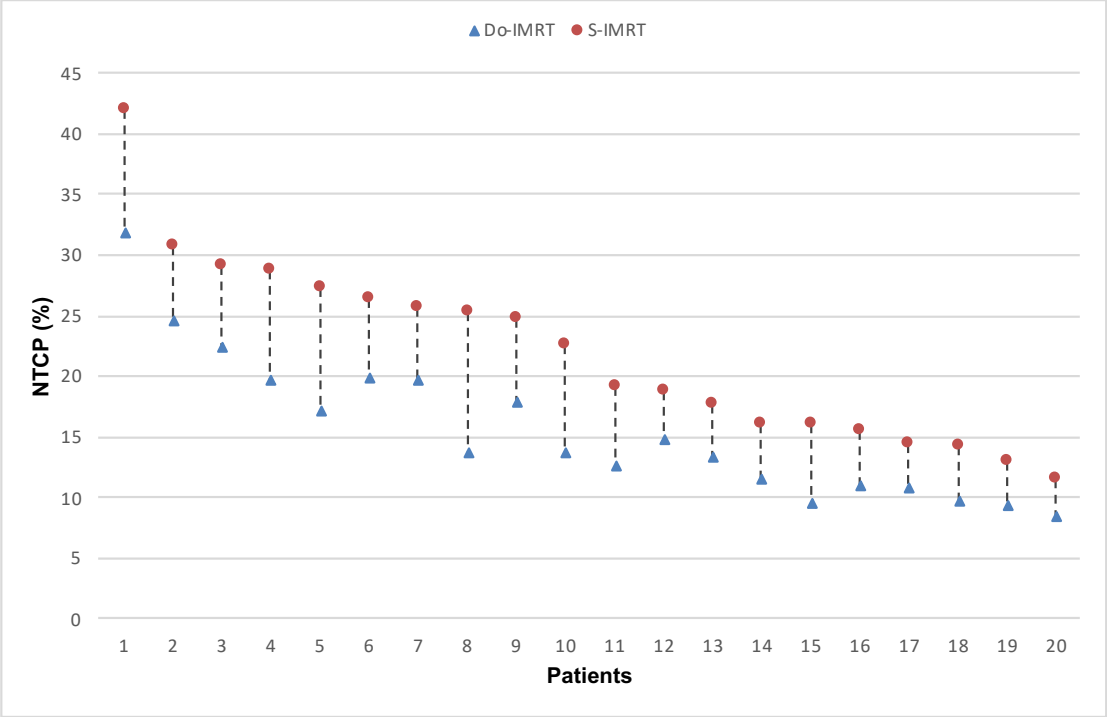


Figure 2.7 Normal tissue complication probabilities (NTCPs) of patient-reported moderate to severe problems with swallowing solid food with standard IMRT (S-IMRT) and dysphagia-optimised IMRT (Do-IMRT)

Patients re-sorted as per the S-IMRT NTCP values

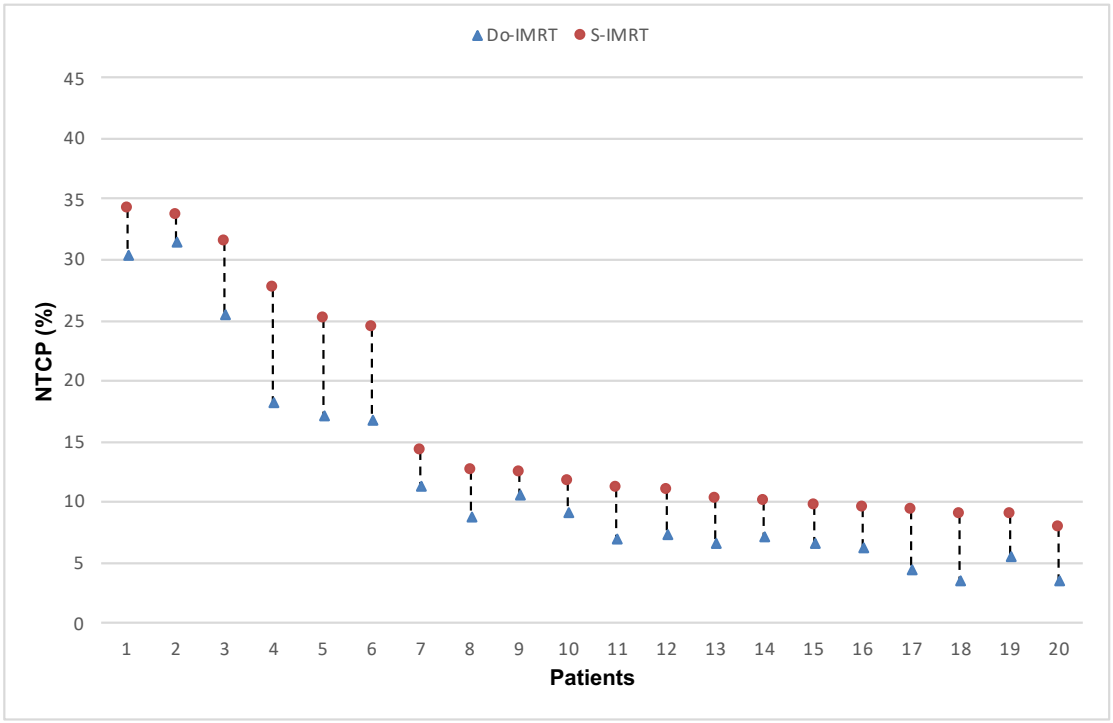


Figure 2.8 NTCPs of patient-reported moderate to severe problems with swallowing soft food with S-IMRT and Do-IMRT (patients re-sorted as per the S-IMRT NTCP values)

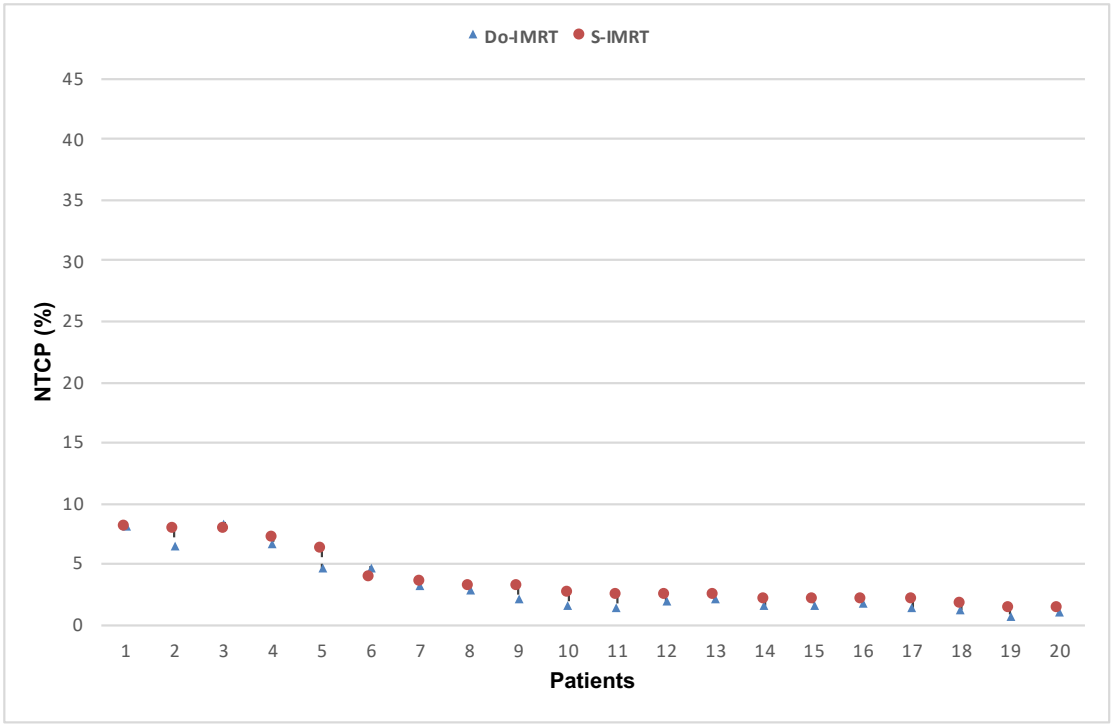


Figure 2.9 NTCPs of patient-reported moderate to severe problems with swallowing liquid food with S-IMRT and Do-IMRT (patients re-sorted as per the S-IMRT NTCP values)

The result of the predicted RAD_{6M} for the four endpoints for each patient in this study is illustrated in Fig 2.10 below.

Figure 2.10 Physician- and patient-scored ΔNTCPs ($NTCP_{S\text{-}IMRT}$ – $NTCP_{Do\text{-}IMRT}$) for each patient. The horizontal dotted line represents the clinically relevant 5 % difference in predicted toxicities between standard IMRT (S-IMRT) and dysphagia-optimised IMRT (Do-IMRT)

2.10 Discussion

This study has demonstrated that clinically relevant reductions in estimated physician-scored RAD_{6M} in patients with LA-OPC can be achieved by minimising dose to the constrictor muscles with Do-IMRT, compared to S-IMRT. On average, the probability of significant swallowing dysfunction was reduced with Do-IMRT from 30.2 % to 21.8 %, without compromising PTV65 coverage or dose conformity.

The toxicity-mitigating benefit of Do-IMRT varied across the individual patients in this study cohort. Primary and nodal tumour location and volume, along with achieving the mandatory median PTV 54 dose objective are important factors that determine the extent of sparing of RT dose to the constrictors, and therefore subsequent dysphagia. Tonsillar primaries with bilateral nodal disease in level 2 will limit sparing of dose to PlanSMPCM, and therefore the predicted benefits with Do-IMRT will be small. On the other hand, substantial reduction in dose delivered to the constrictors is possible with a centrally placed BoT tumour with nodal disease; subsequent gains in function will be much larger with our swallow-sparing technique. Nevertheless, we have demonstrated that $\geq$ 5 % reduction in predicted physician-scored toxicity was achieved in all patients.

Improvements in patient-reported toxicities were comparatively small, and mainly observed with difficulties in swallowing solid foods. For many, the predicted RAD_{6M} with S-IMRT were small, and therefore further clinically relevant reductions with Do-IMRT could not be achieved. Additionally, the NTCP model for patient-reported outcomes for swallowing liquids relied on dose delivered to SGL, an OAR that has no objective function in Do-IMRT planning optimisation. Consequently, substantial benefits with this swallow-sparing strategy would not be expected using the above model. Nevertheless, the predicted toxicities were similar to the outcomes in the Dutch groups' SW-IMRT in-silico planning comparative study.

2.10.1 Comparison of Do-IMRT with other swallow-sparing RT strategies

The 8.4 % improvement in physician-scored RAD_{6M} observed with Do-IMRT in our study is consistent with the 9 % toxicity reduction predicted in the Dutch groups' SW-IMRT in-silico planning study[22]. The two RT techniques are quite different in their methodology, and the merits of implementing Do-IMRT to maintain long-term swallow following treatment completion are best established by considering the differences in the approach to SW-OAR sparing, planning optimisation, and RT dose to non-target tissue between the 2 toxicity-mitigating strategies as discussed below.

2.10.1.1 Dose delivered to SW-OARs

The pivotal assumption of Do-IMRT is that a significant reduction in dose to PCM, in particular the superior and middle constrictors which are often near the tumour in LA-OPC, is more likely to result in reduced symptom burden in this selective cohort of patients. This rationale is based on the evidence presented by Eisbruch's group in a similar patient population previously discussed in chapter 1, where an average reduction in mean dose to PCM to 58 Gy demonstrated favourable long-term physician-reported, observer-rated, and objective measures of swallowing function in a prospective non-randomised study[2]. This hypothesis underpins the planning optimisation of Do-IMRT, which reduced the mean dose to PCM to 47 Gy, with no active attempt made to spare the SGL. In contrast, SW-IMRT was introduced as a SPC – and SGL – sparing RT strategy in keeping with their modelling study that identified the above 2 SW-OARs as the key determinants of subsequent toxicity across a non-selective population of patients with HNC. Their reported swallow-sparing benefits were therefore a direct outcome of attempting to spare dose to both these structures during planning optimisation. As a result, the predicted toxicities with our dysphagia-optimising technique would be expected to be smaller, and not comparable, to that of the Dutch groups' swallow-sparing RT. The lower risk of RAD with SW-IMRT in LA-OPC is predominantly due to SGL-sparing; the technique reduced the mean dose to SPC and SGL by 2.7 Gy and 7.3 Gy respectively. Corresponding reductions with Do-IMRT were 4.7 Gy and 2.6 Gy respectively. The possibility of limiting dose to SGL with our technique was an interesting observation, and in cases

where reductions in dose to SPC were small as shown earlier in Fig 2.6, would have played a more dominant role in improving swallowing function. This dual SPC- and SGL – sparing benefits of Do-IMRT can be explained by the fact that the superior aspect of the SGL is closely related to the MPC, and the lower RT dose to SGL is a consequence of the optimiser attempting to decrease the dose to PlanSMPCM (Fig 2.11). There may, therefore, be potential to further augment the swallow-sparing benefits of Do-IMRT, compared to SW-IMRT, by defining the SGL as an additional OAR during the optimisation process. Further assessment is required though to ensure that the adoption of such a strategy does not significantly compromise target coverage or impact on dose delivered to the pharyngeal constrictors.

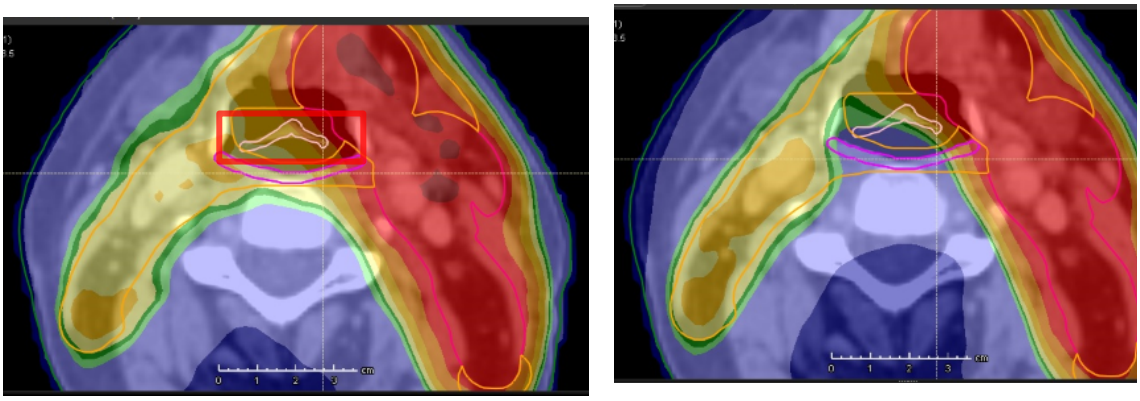


Figure 2.11 Supraglottic larynx (SGL) sparing effect of dysphagia-optimised IMRT (Do-IMRT). (Left) Standard IMRT, where there is no dose sparing of SGL (red box); (Right) Do-IMRT where limiting dose to PlanSMPCM results in sparing of RT dose to SGL.

2.10.1.2 Delivered dose to the salivary glands

Radiation-induced xerostomia is a well-recognised toxicity following HNC treatment, and is a contributory factor for patient-reported persistent

swallowing dysfunction following completion of treatment. The S-IMRT technique described earlier is synonymous with parotid-sparing IMRT, and has been previously shown to reduce the incidence of ≥ grade 2 subjective xerostomia at 12 months when compared with conformal radiotherapy. During Do-IMRT optimisation, a higher priority is placed on constrictor sparing compared to parotid sparing. Such preferential swallow-sparing IMRT planning approach could potentially result in redistribution of dose from the SW-OAR to the salivary glands, thereby nullifying the toxicity-mitigating gain achievable with this novel RT strategy. This comparative planning study demonstrated that there was no difference in mean dose delivered to either parotids or submandibular glands between the two RT techniques, implying that Do-IMRT functions as a dual salivary gland – and swallowing structure – sparing RT strategy, similar to SW-IMRT.

2.10.1.3 Integral dose

The application of complex IMRT techniques has been shown to increase the integral dose delivered to normal tissue compared to 3D conformal RT, as it optimises the conformality of the target volume at the expense of low-dose spillage to surrounding tissue. Such low-dose bath to non-target tissue is a potential concern for developing iatrogenic secondary malignancies. Different mathematical models exist to predict this risk, relying not just on total body exposure to RT, but also on the organ irradiated, patient age, genetic predisposition, and life-style factors. An interesting, and unexpected, finding of this study was that Do-IMRT reduces the integral dose delivered to non-target

soft tissue, relative to S-IMRT. A likely explanation for this favourable outcome is that Do-IMRT optimisation functions by relaxing most of the standard mandatory dose constraints for PTV54 (D99 %, D98 %, D95 %) in order to meet the planning constraints for PlanSMPCM and PlanIPCM respectively. The optimiser, therefore, is permitted to deposit a reduced dose in the region of the constrictors compared to S-IMRT, and consequently no dumping of dose to other OARs or soft tissue is necessary to achieve it. On the other hand, SW-IMRT permits increased spill, and indeed this is necessary to achieve the gain in swallowing function. Improved SW-OAR sparing with that technique was obtained by accepting reduced target conformity, along with a moderate shift of dose to the oral cavity and the neck muscles[23].

2.10.1.4 RT planning technique

The setting of dose-based objective function constraints at the outset with Do-IMRT facilitates the generation of good quality plans with minimal iteration rounds and decision making from the planner. Consequently, the planning methodology should be compatible with centres using fully automated or class solutions for optimisation. In contrast, the SW-IMRT optimisation can be resource- and time- consuming, wherein the objective functions and weights are iteratively adjusted to obtain a satisfactory RT plan, and therefore may be difficult to reproduce in routine clinical practice. Langendijk et al found that that the mean dose delivered to SPC and SGL was reduced by only 1.6 and 2.2 Gy respectively with SW-IMRT, compared to S-IMRT, in the first 20 patients treated clinically with SW-IMRT at their centre. This translated into minimal,

clinically irrelevant NTCP gain with SW-IMRT for this cohort. Further reductions in dose to the SW-OARs, and swallowing complication probabilities, were only observed when they were replanned by more experienced dosimetrists with time available to maximise sparing of RT dose to these swallowing structures.

2.10.2 Risk of loco-regional recurrence with Do-IMRT

Under-dosage to a small volume of mucosal and nodal compartment within the PTV54 volume and situated in close proximity to PlanSMPCM and PlanIPCM is accepted to meet the mandatory planning constraints of Do-IMRT. In order to ensure that this region received an acceptable dose, minimum DVH objectives were placed on PTV54 during optimisation to cover 99 % with 30 Gy and 95 % with 48.6 Gy. Additionally, setting a highly weighted minimum dose objective of 51.3 Gy and a uniform objective of 54 Gy to that part of PTV54 away from the constrictor ensured that the mandatory median dose constraint for PTV54 was achieved. There is, however, a theoretical increased risk of loco-regional recurrence around the spared muscle with Do-IMRT, which could potentially compromise long-term cure rates with primary CRT. In this regard, studies that have analysed failure patterns in HNC following IMRT with a volumetric approach to CTV delineation have consistently demonstrated that a majority of relapses occur within the central high-dose volumes, implying poor biology rather than inaccurate risk assessment as a primary reason for treatment failure[25, 26]. For instance, the MD Anderson group mapped the

spatial location of recurrence in 47 patients to a pre-treatment PET scan using deformable image registration and identified that 90 % of failures occurred in regions that had received $\geq$ 95 % dose prescribed to the high dose PTV[26]. Compromising PTV54 dosimetric coverage, but not PTV65, to that part of the oropharynx near the constrictors with Do-IMRT is therefore unlikely to significantly increase the risk of mucosal relapse. Further corroboration about the low risk of local failure in the spared region is provided by the recently published international consensus delineation guidelines, which recommends the inclusion of primary tumour with a 1 cm margin only in the prophylactic dose rather than the entire compartment as in our study[27]. It can therefore be argued that our target volume delineation is conservative. Evidence from the above studies have additionally shown that tumour recurrence in the central elective nodal volumes, where under-dosage is accepted with Do-IMRT, is uncommon. Consequently, we feel that the perceived risk of loco-regional recurrence in the spared region is low.

2.10.3 Study limitation

A limitation of this study is the application of NTCP models that were not validated for the patient cohort and RT treatment technique. Consequently, this may introduce uncertainties to the results and it is therefore plausible that the clinical use of Do-IMRT may lead to different toxicity outcomes than predicted in this study[28, 29]. However, NTCP models for RAD are limited, and we elected to use the most widely quoted toxicity model with a relatively comparable patient population. The NTCP values for S-IMRT in this study are

similar to published outcomes, and the use of this model for predicting toxicity therefore appears reasonable. In addition, while the absolute NTCP values may differ with an alternative model, it is likely that the relative difference between S-IMRT and Do-IMRT will be similar.

2.11 Conclusion

This hypothesis-driven planning study met its primary endpoint of demonstrating clinically relevant reduction in the predicted risk of persistent swallowing dysfunction with Do-IMRT in LA-OPC, relative to S-IMRT. The predicted benefits varied, but met the pre-defined threshold of $\geq$ 5 % improvement in physician-scored RAD_{6M} for each patient in this study. These improvements were achieved without compromising dose to the high-dose target volume and OARs.

2.12 References list

1. Petkar, I., et al., *Dysphagia-optimised Intensity-modulated Radiotherapy Techniques in Pharyngeal Cancers: Is Anyone Going to Swallow it?* Clin Oncol (R Coll Radiol), 2017. **29**(7): p. e110-e118.

2. Feng, F.Y., et al., *Intensity-modulated chemoradiotherapy aiming to reduce dysphagia in patients with oropharyngeal cancer: clinical and functional results.* J Clin Oncol, 2010. **28**(16): p. 2732-8.

3. Eisbruch, A., et al., *Chemo-IMRT of oropharyngeal cancer aiming to reduce dysphagia: swallowing organs late complication probabilities and dosimetric correlates.* Int J Radiat Oncol Biol Phys, 2011. **81**(3): p. e93-9.

4. Petkar, I., et al., *DARS: a phase III randomised multicentre study of dysphagia- optimised intensity- modulated radiotherapy (Do-IMRT) versus standard intensity- modulated radiotherapy (S-IMRT) in head and neck cancer.* BMC Cancer, 2016. **16**(1): p. 770.

5. Samuels, S.E., et al., *Methods for Reducing Normal Tissue Complication Probabilities in Oropharyngeal Cancer: Dose Reduction or Planning Target Volume Elimination.* Int J Radiat Oncol Biol Phys, 2016. **96**(3): p. 645-52.

6. Christianen, M.E., et al., *Swallowing sparing intensity modulated radiotherapy (SW-IMRT) in head and neck cancer: Clinical validation according to the model-based approach.* Radiother Oncol, 2015.

7. Welsh, L., et al., *Prospective, longitudinal, multi-modal functional imaging for radical chemo-IMRT treatment of locally advanced head and neck cancer: the INSIGHT study.* Radiat Oncol, 2015. **10**: p. 112.

8. Wiezorek, T., et al., *Rotational IMRT techniques compared to fixed gantry IMRT and tomotherapy: multi-institutional planning study for head-and-neck cases.* Radiat Oncol, 2011. **6**: p. 20.

9. Teoh, M., et al., *Volumetric-modulated arc therapy (RapidArc) vs. conventional fixed-field intensity-modulated radiotherapy for (1)(8)F-FDG-PET-guided dose escalation in oropharyngeal cancer: a planning study.* Med Dosim, 2013. **38**(1): p. 18-24.

10. Cilla, S., et al., *Volumetric modulated arc therapy (VMAT) and simultaneous integrated boost in head-and-neck cancer: is there a place for critical swallowing structures dose sparing?* Br J Radiol, 2016. **89**(1059): p. 20150764.

11. van de Water, T.A., et al., *Potential benefits of scanned intensity-modulated proton therapy versus advanced photon therapy with regard to sparing of the salivary glands in oropharyngeal cancer.* Int J Radiat Oncol Biol Phys, 2011. **79**(4): p. 1216-24.

12. Gregoire, V., et al., *Delineation of the neck node levels for head and neck tumors: a 2013 update. DAHANCA, EORTC, HKNPCSG, NCIC CTG, NCRI, RTOG, TROG consensus guidelines.* Radiother Oncol, 2014. **110**(1): p. 172-81.

13. Thomson, D., et al., *NIMRAD - a phase III trial to investigate the use of nimorazole hypoxia modification with intensity-modulated radiotherapy in head and neck cancer.* Clin Oncol (R Coll Radiol), 2014. **26**(6): p. 344-7.

14. Christianen, M.E., et al., *Delineation of organs at risk involved in swallowing for radiotherapy treatment planning.* Radiother Oncol, 2011. **101**(3): p. 394-402.

15. Owadally, W., et al., *PATHOS: a phase II/III trial of risk-stratified, reduced intensity adjuvant treatment in patients undergoing transoral surgery for Human papillomavirus (HPV) positive oropharyngeal cancer.* BMC Cancer, 2015. **15**: p. 602.

16. Brouwer, C.L., et al., *CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines.* Radiother Oncol, 2015. **117**(1): p. 83-90.

17. *Tyler J BD, Rooney K, Nutting C. Development of dysphagia-optimised IMRT for head and neck cancer treatment in the DARS trial. in Proceedings of the 35th Annual Meeting of Estro 2016, Turin, Italy. 2016*.

18. Thilmann, C., et al., *[Virtual bolus for inversion radiotherapy planning in intensity-modulated radiotherapy of breast carcinoma within the scope of adjuvant therapy].* Strahlenther Onkol, 2002. **178**(3): p. 139-46.

19. MacFarlane, M., et al., *Evaluation of unified intensity-modulated arc therapy for the radiotherapy of head-and-neck cancer.* Radiother Oncol, 2016. **119**(2): p. 331-6.

20. Yang, R., et al., *Integral Dose in Three-dimensional Conformal Radiotherapy, Intensity-modulated Radiotherapy and Helical Tomotherapy.* Clinical Oncology, 2009. **21**(9): p. 706-712.

21. Christianen, M.E., et al., *Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.* Radiother Oncol, 2012. **105**(1): p. 107-14.

22. van der Laan, H.P., et al., *The potential benefit of swallowing sparing intensity modulated radiotherapy to reduce swallowing dysfunction: an in silico planning comparative study.* Radiother Oncol, 2012. **103**(1): p. 76-81.

23. van der Laan, H.P., et al., *Swallowing-sparing intensity-modulated radiotherapy for head and neck cancer patients: treatment planning optimization and clinical introduction.* Radiother Oncol, 2013. **107**(3): p. 282-7.

24. Mazzola, R., et al., *Dose-volume-related dysphagia after constrictor muscles definition in head and neck cancer intensity-modulated radiation treatment.* Br J Radiol, 2014. **87**(1044): p. 20140543.

25. Tandon, S., et al., *Failure patterns of head and neck squamous cell carcinoma treated with radical radiotherapy by intensity modulated radiotherapy technique using focal volume and dosimetric method.* Head Neck, 2018.

26. Mohamed, A.S.R., et al., *Patterns-of-failure guided biological target volume definition for head and neck cancer patients: FDG-PET and dosimetric analysis of dose escalation candidate subregions.* Radiother Oncol, 2017. **124**(2): p. 248-255.

27. Gregoire, V., et al., *Delineation of the primary tumour Clinical Target Volumes (CTV-P) in laryngeal, hypopharyngeal, oropharyngeal and oral cavity squamous cell carcinoma: AIRO, CACA, DAHANCA, EORTC, GEORCC, GORTEC, HKNPCSG, HNCIG, IAG-KHT, LPRHHT, NCIC CTG, NCRI, NRG Oncology, PHNS, SBRT, SOMERA, SRO, SSHNO, TROG consensus guidelines.* Radiother Oncol, 2018. **126**(1): p. 3-24.

28. Jakobi, A., et al., *NTCP reduction for advanced head and neck cancer patients using proton therapy for complete or sequential boost treatment versus photon therapy.* Acta Oncol, 2015. **54**(9): p. 1658-64.

29. Bijman, R.G., et al., *Impact of model and dose uncertainty on model-based selection of oropharyngeal cancer patients for proton therapy.* Acta Oncol, 2017. **56**(11): p. 1444-1450.

3 Chapter 3: A comparative planning study of IMRT and IMPT techniques to reduce swallowing dysfunction in LA-OPC

3.1 Introduction

In the previous chapter, we have demonstrated the potential toxicity-mitigating benefits of Do-IMRT over S-IMRT in LA-OPC. This chapter investigates whether the dosimetric advantage provided by the finite range of protons, as discussed in chapter 1, can be exploited to further reduce dose to critical SW-OARs compared to photons and consequently improve long-term swallowing function.

3.1.1 UK experience in proton beam therapy (PBT)

IMPT planning is complex, and highly sensitive to treatment uncertainties compared to IMRT[1-3]. An accurate understanding of the sources and impact of such uncertainties on the proton beam is necessary, to ensure their incorporation during the RT optimisation process to produce robust plans. In this context, it is pertinent to emphasise that knowledge and expertise in PBT amongst UK oncologists and physicists was limited when this thesis was undertaken. There were, therefore, initial concerns about the feasibility of this project.

To address these issues, I visited the Head and Neck Unit at Maryland Proton Treatment Centre, Baltimore, USA, courtesy of the Keith Durrant Memorial Fellowship that was awarded by the Royal College of Radiologists (RCR). In addition, our physicist Alex Dunlop (see acknowledgements) visited the MD Anderson Proton Therapy Center, USA on a separate RCR fellowship. These visits provided valuable insight into the proton treatment-planning pathway including RT optimisation techniques, assessment of plan robustness, and treatment verification process. In addition, I attended the proton symposium at University College London (UCL), and developed collaboration with Dr Richard Amos, an experienced proton physicist and operational lead for PBT physics at University College London. A number of planning workshops were also held at RMH with Dr Amos. There was, therefore, sufficient confidence in proton optimisation techniques in HNC during this study.

3.1.2 Treatment uncertainties in proton delivery

Protons lose energy but not intensity as they travel through tissue; in contrast photons lose intensity but no energy. Therefore, any uncertainty in tissue density has a direct impact on the proton range and subsequent spread-out Bragg peak, while there is minimal change in the photon dose distribution (Figure 3.1). In addition to accounting for geometrical uncertainties due to setup errors (SE) and anatomical changes that are also observed with IMRT, range uncertainty (RU) due to systematic and random errors in calculating the stopping power of protons needs to be considered in proton planning. RU arise due to anatomical changes (target volume shrinkage, weight loss, sinus filling,

organ motion), uncertainties in relative biological effectiveness (RBE), computed tomography (CT) Hounsfield units (HU) calculation, conversion of HU to proton stopping power, metal/dental amalgam artefacts, and beam hardening[4]. The distinct physical properties of protons that make it an attractive alternative to photons, therefore, also makes it highly sensitive to changes in tissue density during the course of a treatment (Figure 3.2).

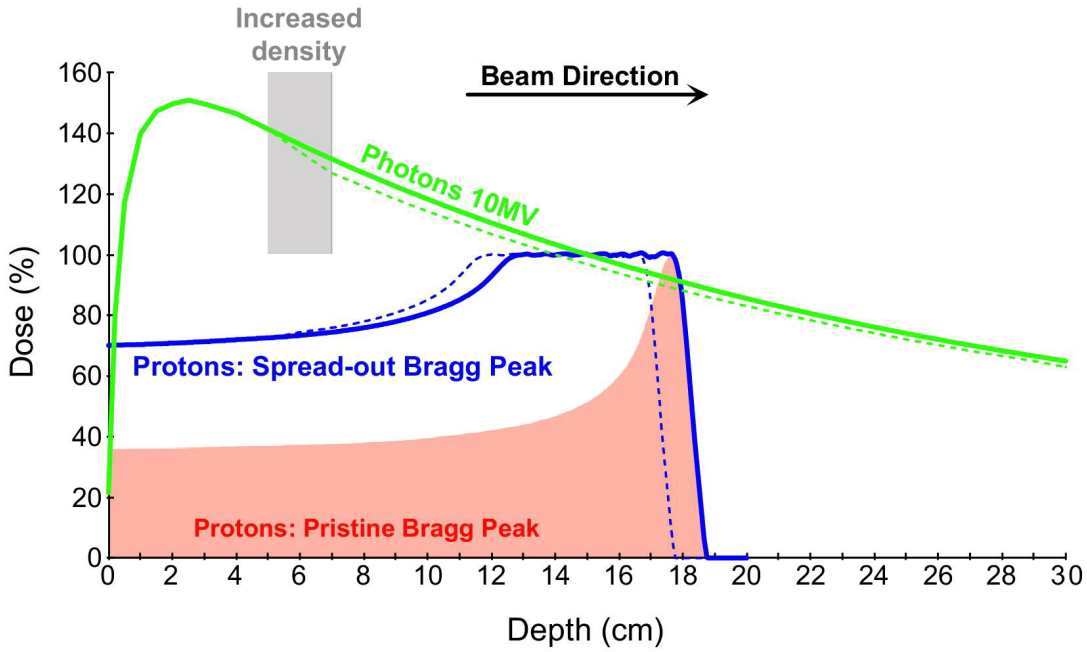


Figure 3.1 Depth-dose curves without (solid line) and with (dashed) an anatomical density variation in the beam entrance region

(Reproduced from Engelsman et al[5])

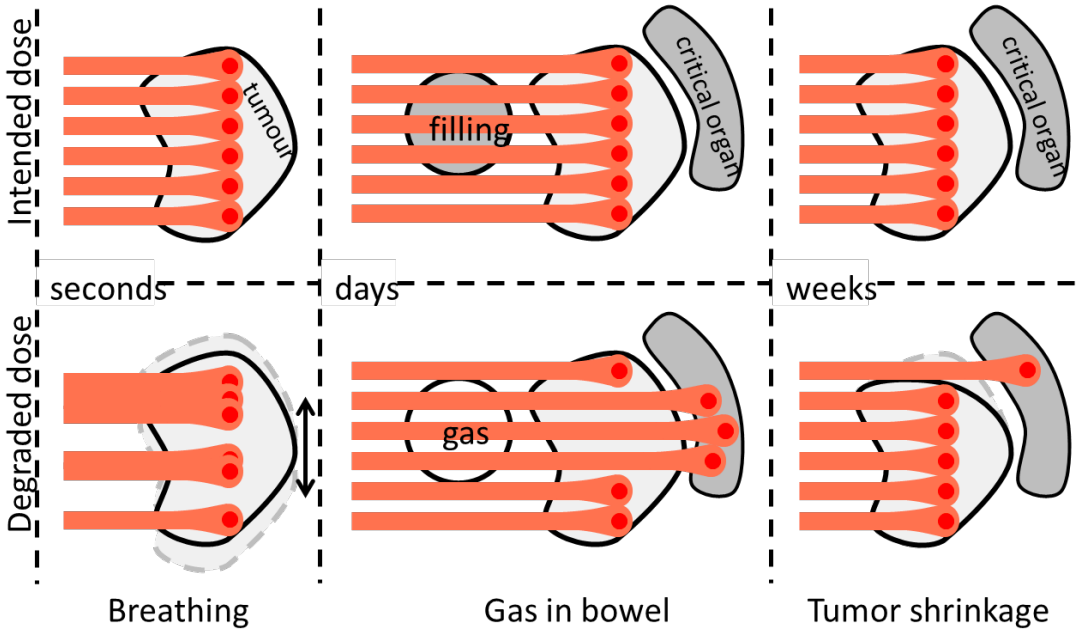


Figure 3.2 Impact of treatment uncertainties on proton RT plans, illustrating significant degradation of the delivered dose in the presence of such errors, and consequently risking under-dosage of target volumes and/or overdosage of organs at risk (OAR). PTV margins do not account for range uncertainties, a major source of uncertainty in proton planning

Reproduced from Hoogeman presentation[6]

3.1.3 Management of uncertainties in IMPT

In IMRT, the PTV provides a safety margin for positional uncertainties in the CTV. An underlying principle behind the isotropic expansion of CTV to generate the PTV is that the photon dose distribution is not perturbed by changes in tissue density. Therefore, as long as the CTV moves within the PTV that is irradiated to the prescribed dose, it is assumed that there will be adequate coverage of the CTV. This traditional concept of CTV – PTV expansion does not readily apply for MFO-IMPT planning, as a major source of uncertainty in the direction of the incident beam is RU, which has been estimated to be approximately 2.5 – 3.5 % of the path length [7]. Range errors during treatment can result in relative shift of the composite and highly complex MFO-IMPT planned dose distribution, resulting in undesirable hot or cold spots within the CTV; the addition of a margin to the CTV is unlikely to mitigate the impact of RU on delivered dose in such scenarios[8](Fig 3.3). Incorporation of RU, in addition to geometrical errors, in the PBT planning process is therefore integral to ensure that there is minimal degradation of plan quality between the planned and delivered dose distribution. The two approaches to account for the above uncertainties in MFO-IMPT are robustness in plan optimisation and robustness in plan evaluation, which are discussed in the following section.

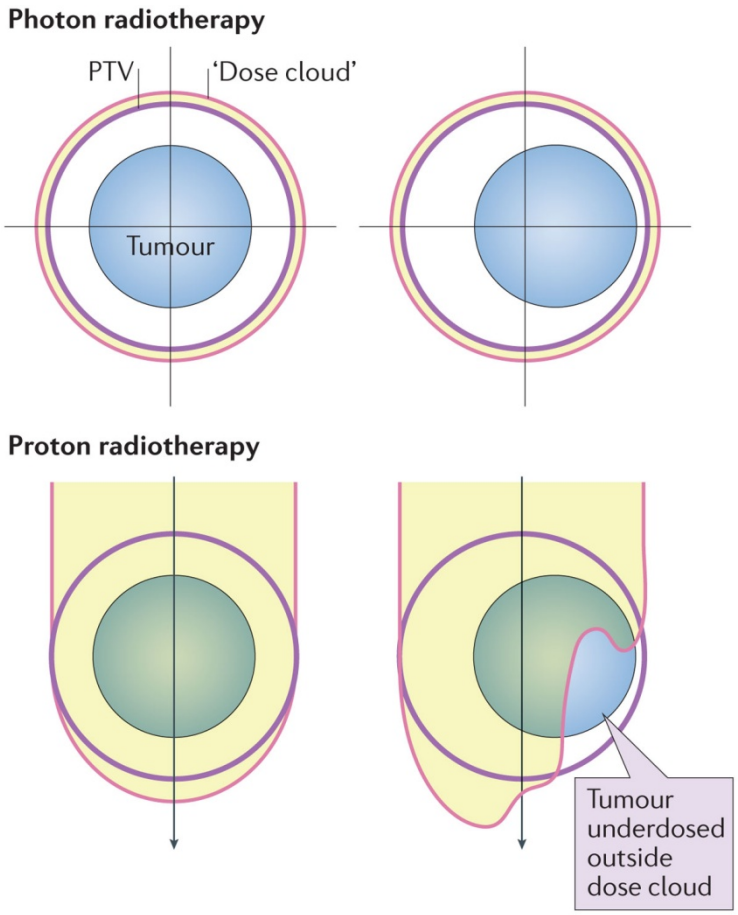


Figure3.3 PTV for proton beam therapy

The standard way of dealing with positional uncertainties is to add a planning target volume (PTV) margin and to shape the radiation dose distribution ('dose cloud') to conform to the PTV. If in the case of photon RT the tumour moves within the PTV most of the time (eg, to the right, as shown here), tumour dose coverage is guaranteed because the shape of the dose cloud is largely unaffected by the tumour motion (left panel). In the case of proton radiotherapy, the simple margin approach is not sufficient because here the dose cloud is affected by tumour motion as well as motion of organs far away from the tumour (right panel)(Reproduced from Baumann et al[9].

3.1.3.1 Robustness in plan optimisation

In robust optimisation (RO), a set of range and setup error scenarios is provided to the proton TPS. The RO algorithm incorporates these uncertainties when determining spot weights during optimisation, and generates an IMPT plan where the CTV is robust against the effects of the above possible errors. The two approaches for implementing the errors into the optimisation cost function are the worst-case approach and the probabilistic approach[4, 5, 8].

The voxel-wise worst-case RO algorithm determines the worst-case dose distribution by simulation of errors for each voxel (minimum dose in target and maximum dose in OAR), and subsequently including the dose distribution of all worst-case voxels in the optimiser cost function to generate a plan that is robust for such scenario[10]. A potential limitation of this approach is that the worst-case dose distribution is unphysical, as the worst-case dose in 2 voxels within the target may occur under different scenarios. Therefore, plans generated by this method may be conservative. An alternative is the minimax or composite worst-case optimisation method, where only physically realisable dose distributions are considered[11].

The probabilistic approach generates an IMPT plan that produces an acceptable dose distribution for all error scenarios. This approach optimises the average plan quality, though it remains possible that the dose distribution may be inferior for the worst scenario[12].

Currently, there is no evidence to demonstrate the superiority of one approach over another. Either way, RO requires significant computational time, and is an evolving strategy with a lack of universal agreement regarding the most appropriate methodology to adopt for PBT planning.

3.1.3.2 Robustness in plan evaluation (RE)

RO software does not exist on all TPS, and therefore the default strategy in this situation involves the use of PTVs and PRVs to direct plan optimisation, similar to photon planning. Dose distributions for the CTVs (not PTVs) and OARs (not PRVs) are computed under different scenarios of SE and RU, and the corresponding DVHs plotted. The band of DVHs generated for a structure then represents the possible variations in delivered dose, and a robust plan will achieve target coverage in each of the simulated scenarios. On the other hand, considerable variation will exist in CTV coverage in a non-robust plan, and further plan optimisation is carried out if it is deemed as unacceptable. For OARs, bandwidths at critical points (such as maximum dose to spinal cord) would inform the need for re-optimisation. The number, or nature of scenarios that need to be included in the evaluation model remains investigational.

The proton centres due to open in the UK will be employing this technique to verify the robustness of the nominal proton plan in HNC (personal communication with Dr David Thompson, Christie Hospital).

3.2 Aims of this study

This retrospective planning study was designed to investigate whether the dose delivered to PCM in LA-OPC can be reduced using MFO-IMPT, planned with RE and RO, compared to IMRT. The expected toxicity-mitigating benefits of SW-OAR sparing with IMPT, compared to IMRT, was evaluated by applying the predictive models for RAD_{6M} previously described in chapters 1 and 2 respectively.

3.3 Null hypothesis

There will be no difference in the predicted risk of physician-scored RAD_{6M} between dysphagia-optimised IMPT (Do-IMPT) and Do-IMRT following CRT in LA-OPC.

3.4 Primary objective

To determine, if using Do-IMPT in LA-OPC results in a $\Delta NTCP_{DO}$ of $\geq$ 5 % for physician-scored RAD_{6M} respectively, where

- $\Delta NTCP_{DO}$ is the difference between the risk of RAD_{6M} with Do-IMRT ($NTCP_{Do\text{-}IMRT}$) and Do-IMPT ($NTCP_{Do\text{-}IMPT}$).

3.5 Primary endpoint

The differences in mean estimated risks of physician-scored RAD_{6M} with Do-IMRT and Do-IMPT.

3.6 Secondary objectives

- To determine $\Delta NTCP_{standard}$, which is the difference between the risk of RAD_{6M} with S-IMRT and standard IMPT (S-IMPT).
- To compare the differences in dose delivered to SW-OARs and NSW-OARs between IMRT and IMPT.
- To assess the dosimetric differences in target volume coverage between photons and PBT.
- To determine the difference in integral dose delivered with photons and PBT.
- To evaluate the differences in delivered dose to target volumes and OARs between RO and PTV-based Do-IMPT plans.
- To analyse the influence of 3-beam and 7-beam PBT on OAR sparing and target volume coverage.

3.7 Materials and methods

Ten LA-OPC patients from the study group described in chapter 2 were chosen for this retrospective comparative planning study. Selected patients had minimal dental artefacts, which reduced uncertainties in the calculation of the Hounsfield units of nearby tissues on the planning CT scan, thereby permitting reliable dose calculations. Clinically, the impact of dental artefacts is usually overcome by using either complex processing algorithms or replacing the metal dental amalgam with low-density composite ceramic filling.

For each patient, four proton plans were generated and compared with corresponding S-IMRT and Do-IMRT plans created in chapter 2:

1. PTV-based S-IMPT plan (S-IMPT)
2. PTV-based Do-IMPT plan using 3 beams (Do-IMPT$_{3B}$)
3. PTV-based Do-IMPT using 7 beams (Do-IMPT$_{7B}$)
4. CTV-based robustly optimised Do – IMPT plan using 3 beams (Do-IMPT$_{RO}$)

3.7.1 Demographic data

Demographic data was collected as described in chapter 2

3.7.2 Radiotherapy

Planning procedure, target volume and OAR delineation are as described in chapter 2. PTV was not created for the robustly optimised Do-IMPT$_{RO}$ plan.

3.7.2.1 Plan design for PBT

IMPT plans were generated and optimised on the research version of RayStation v6 TPS (RaySearch Laboratories, Stockholm, Sweden). Proton beam data for the Probeam machine was provided by Raysearch Laboratories.

Monte Carlo dose engine (MC) was used to produce RT plans based on physical dose. Compared to analysis-based algorithms such as the pencil beam model, this computer-generated algorithm provides a more accurate proton dose distribution by simulating the path of individual particles through the patient, thereby incorporating the different densities of the different tissues the beam traverses through[13, 14]. MC was used for calculating spot dose distributions during the optimisation process, as well as for calculating the final dose distribution.

Beam design:

3 beams were used to generate the S-IMPT, Do-IMPT$_{3B}$, and Do-IMPT$_{RO}$ plans: right anterior oblique (RAO), left anterior oblique (LAO), and a posterior beam. Such a beam arrangement for proton treatment is used clinically at MD Anderson Proton Therapy Center and minimises uncertainties along the proton path by avoiding beams going through the mouth, teeth, and maxillary

sinuses[15]. The oblique beams place its distal spots lateral to SC, and the posterior beam traverses the cord placing its distal spots just posterior to the parotid gland. The gantry angle for each beam between patients varied, avoiding inhomogeneity regions, depending upon target volume shape. Beam angles of 30°, 110°, 150°, 180°, 230°, 260°, and 330° respectively were used for Do-IMPT$_{7B}$ optimisation.

Spot spacing and energy layers were automatically selected by the MC dose engine, depending upon the spot size at a particular depth and the shape of the target volume in the beam direction respectively. The final average pencil beam spot size sigma varied from 3.5 mm to 6 mm for energies from 70 MeV to 245 MeV, where sigma represents the standard deviation of the lateral intensity distribution of the proton beamlet. An air gap of 5 cm between the snout and patient surface was used for all plans.

Beam optimisation*:*

MFO, which involves the use of an inverse optimisation method to adjust spot weights from all fields simultaneously to achieve the planning objectives previously discussed in chapter 1, was used to generate a conformal dose distribution for all proton plans. The use of non-anatomical volumes to guide the TPS during optimisation was similar to IMRT, as described in chapter 2.

Robust optimisation for Do-IMPT$_{RO}$:

The RO tool on RS takes into account the impact of possible SE and RU on the proton dose distribution during the optimisation process to create a plan that is sufficiently robust to these uncertainties. With a SE of 3 mm (CTV – PTV margin at our centre) and RU of 3.5 %, a total of 21 different uncertainty scenarios using the worst-case approach were considered during the optimisation process for both target volumes and OARs. Patient position errors were not considered independently for all beams during RO, as this would have exponentially increased the number of scenarios with the number of beams, resulting in significantly longer computation time.

Planning objectives:

Baseline dose-based and dose-volume – based objectives were as described in chapter 2. For S-IMPT, objectives were similar to S-IMRT plans described. For Do-IMPT$_{RO}$, the DVH objectives for PTV65 and PTV54 were replaced by CTV65 and CTV54 respectively.

Dose prescription:

Prescription was to the median dose point on the DVH such that the prescription dose of 65 Gy was received by 50 % of PTV65 (for PTV-based plans), and CTV65 (robustly optimised plan). Physical dose was considered for this comparative planning study and is not corrected for RBE, as this was not available on the TPS at the time of this thesis.

3.7.2.2 Plan evaluation

Plan evaluation was as described in chapter 2. As the target volumes were different for PTV-based plans and RO PBT, the approach to determine the integral dose used in chapter 2 may not provide a fair comparison. For this study, the following structures were defined:

1. Body tissue: This was defined as the external contour of the patient and reflects the integral dose delivered to the patient.
2. Percentage volume of external contour receiving 5 Gy (V_5) and 30 Gy (V_{30}) respectively.

3.8 Statistical analysis

In total, 60 plans were generated and evaluated. Statistical analysis was carried out on SPSS version 25, using the one-way repeated measures analysis of variance (ANOVA) where data was normally distributed, with all pairwise comparisons tested using Bonferroni's multiple comparison test wherein p-values $\leq$ 0.05 indicated statistical significance. The one-way repeated measures ANOVA is an extension of the paired samples t-test and is an appropriate statistical test when the same group of patients are tested on three or more different occasions on the same dependant variable[16, 17]. The Bonferroni correction was used to reduce the risk of type 1 error that is otherwise possible with multiple comparisons. For data that was not normally distributed, the non-parametric Friedman test was applied to carry out the

statistical analysis[18]. Target volume coverage and OAR doses were compared between each of the six techniques.

3.9 Results

Patient, tumour site, and target volume characteristics are summarised in table 3.1. Fig 3.4 illustrates dose distributions on the same sagittal CT slice in the same patient for the six RT techniques.

Patient nos	Age	Sex	Tumour site	TNM stage	AJCC stage	p16 status	Treatment modality	GTV (cm^3)	PTV65 (cm^3)	PTV54 (cm^3)
1	43	M	BoT	T2N1	III	Positive	CRT	7.2	91.2	375
2	68	M	BoT	T1N2c	IVa	Negative	IC + CRT	5.2	88.1	423
3	47	M	BoT	T2N2b	IVa	Positive	CRT	35.3	203	347
4	67	M	BoT	T2N2a	IVa	Positive	CRT	14.2	122.3	435
5	61	M	BoT	T2N2b	IVa	Positive	CRT	26.9	189	433
6	52	M	T	T2N2b	IVa	Negative	CRT	19.3	143	486
7	58	M	T	T3N0	III	Positive	CRT	27.3	142	371
8	63	M	T	T2N1	III	Positive	CRT	15.2	143	473
9	64	M	T	T2N2b	IVa	Negative	IC + CRT	12.8	187	315
10	61	M	T	T3N2b	IVa	Positive	CRT	35.7	275	439

Table 3.1 Baseline patient characteristics

AJCC, American Joint Commission on Cancer; BoT, base of tongue; CRT, chemoradiation; GTV, gross tumour volume; IC, induction chemotherapy; M, male; PTV65, planning target volume receiving 65 Gy; PTV54, planning target volume receiving 54 Gy

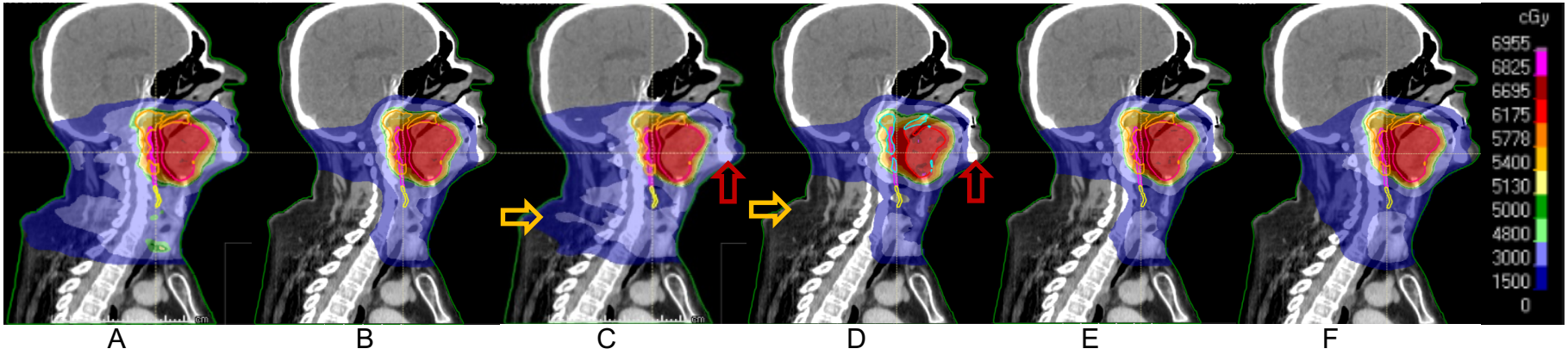


Figure 3.4 Example of dose distribution with the different IMRT and IMPT techniques for LA-OPC

Sagittal images of the same computed tomography slice in the same patient showing the dose distribution achieved by each of the six radiotherapy (RT) planning techniques. A, standard intensity-modulated RT (S-IMRT); B, S-intensity modulated proton therapy (S-IMPT); C, dysphagia-optimised IMRT (Do-IMRT); D, robust optimised Do-IMPT (Do-IMPT$_{RO}$); E, three field PTV-based Do-IMPT (Do-IMPT$_{3B}$); F, seven field PTV-based Do-IMPT (Do-IMPT$_{7B}$).
Red arrow indicates the relative sparing of dose to the anterior oral cavity with Do-IMPT$_{RO}$ relative to Do-IMRT; orange arrow demonstrates the sparing of healthy tissue and spinal cord between the two RT techniques

3.9.1 Dose comparisons for PTVs:

The mean (SD) and statistical results for the various PTV (CTV for robustly optimised Do-IMPT) dose constraints and objectives with the RT techniques are tabulated in table 3.2 and table 3.3 respectively. The planning objectives for the target volumes were achieved with all proton and photon plans. There was no significant difference for the mandatory target volume objectives between S-IMRT and S-IMPT, apart from dose to 99% of PTV65 that was higher with S-IMRT. PTV65 coverage was statistically superior with Do-IMRT compared with the PTV-based dysphagia-optimised proton plans; likewise Do-$IMPT_{7B}$ significantly improved PTV65 coverage compared to Do-$IMPT_{3B}$. On the other hand, the PTV54 coverage was superior with PBT compared to Do-IMRT, and met statistical significance for some of the dose objectives. Do-$IMPT_{RO}$ target volumes, which had CTV optimisation objectives, were not included in this statistical analysis for comparison of the PTVs.

Structure	Standard RT treatment		Dysphagia-optimised RT treatment			
	S-IMRT Gy (SD)	S-IMPT Gy (SD)	Do-IMRT Gy (SD)	Do-IMPT$_{RO}$ * Gy (SD)	Do-IMPT$_{3B}$ Gy (SD)	Do-IMPT$_{7B}$ Gy (SD)
PTV65						
D99% ≥ 90% (58.5 Gy)	62.1 (0.3)	61.7 (0.2)	62.3 (0.3)	62.1 (0.1)	61.8 (0.2)	62.1 (0.1)
D98% ≥ 95% (61.75 Gy)	62.5 (0.3)	62.2 (0.2)	62.8 (0.3)	62.5 (0.1)	62.2 (0.1)	62.6 (0.1)
D95% ≥ 95% (61.75 Gy)	63.1 (0.2)	63.0 (0.1)	63.5 (0.3)	63.1 (0.1)	63.0 (0.1)	63.2 (0.1)
D50% = 100% (65.1 Gy)	65.1 (0.0)	65.1 (0.0)	65.1 (0.0)	65.1 (0.0)	65.1 (0.0)	65.1 (0.0)
D5% ≤ 105% (68.25 Gy)	66.9 (0.3)	66.7 (0.2)	66.2 (0.2)	67.2 (0.1)	67.1 (0.1)	66.4 (0.1)
D2% ≤ 107% (69.55 Gy)	67.3 (0.3)	67.3 (0.2)	66.6 (0.2)	67.8 (0.2)	67.6 (0.2)	66.8 (0.1)
PTV54						
D99% ≥ 90% (48.6 Gy)	51.2 (0.2)	51.2 (0.2)	43.8 (2.5)	44.7 (2.8)	44.3 (2.9)	44.2 (2.9)
D98% ≥ 95% (51.3 Gy)	51.6 (0.2)	51.6 (0.2)	45.9 (2.4)	46.7 (2.5)	46.6 (2.6)	46.4 (2.7)
D95% ≥ 95% (51.3 Gy)	52.1 (0.1)	52.1 (0.1)	49.4 (1.4)	49.3 (1.7)	50.0 (1.3)	50.1 (1.4)
D50% = 100% (54.0 Gy)[$]	53.8 (0.2)	54.1 (0.1)	54.0 (0.3)	54.5 (0.1)	54.3 (0.1)	54.2 (0.1)
D5% ≤ 105%	61.6 (0.8)	61.0 (0.4)	61.6 (0.6)	61.2 (0.6)	61.5 (1.3)	61.0 (0.4)
D2% ≤ 107%	63.1 (0.6)	62.7 (0.3)	63.3 (0.3)	62.8 (0.5)	63.2 (1.1)	62.6 (0.3)

Table 3.2 PTV coverage with the six RT techniques

Do-IMPT$_{xB}$, dysphagia-optimised intensity modulated proton therapy with x number of beams; Do-IMPT$_{RO}$, robustly optimised Do-IMPT; Do-IMRT, dysphagia-optimised intensity modulated radiotherapy; Dx%, dose to x % volume; PTV, planning target volume; S-IMRT, standard IMRT

***** tabulated results are for clinical target volume (CTV) 65 and CTV54 respectively; [$] only mandatory PTV54 objective for Do plans

Structure	S-IMRT v S-IMPT	S-IMRT v Do-IMRT	S-IMRT v Do-IMPT$_{3B}$	S-IMRT v Do-IMPT$_{7B}$	Do-IMRT v S-IMPT	Do-IMRT v Do-IMPT$_{3B}$	Do-IMRT v Do-IMPT$_{7B}$	S-IMPT v Do-IMPT$_{3B}$	S-IMPT v Do-IMPT$_{7B}$	Do-IMPT$_{3B}$ v Do-IMPT$_{7B}$
	p-values, shaded boxes indicate statistically significant difference									
PTV65										
D99%	0.01	0.12	0.16	1.00	<0.01	<0.01	0.41	1.00	<0.01	<0.01
D98%	0.01	0.02	0.13	1.00	<0.01	<0.01	0.25	1.00	<0.01	<0.01
D95%	0.09	<0.01	0.22	1.00	<0.01	<0.01	0.04	1.00	<0.01	<0.01
D5%	1.00	<0.01	1.00	<0.01	<0.01	<0.01	0.04	0.19	<0.01	<0.01
D2%	1.00	<0.01	0.08	0.01	<0.01	<0.01	0.10	0.08	<0.01	<0.01
PTV54										
D99%	1.00	<0.01	<0.01	<0.01	<0.01	0.72	1.00	<0.01	<0.01	1.00
D98%	1.00	<0.01	<0.01	<0.01	<0.01	0.02	0.12	<0.01	<0.01	<0.01
D95%	1.00	<0.01	<0.01	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	1.00
D50%	0.06	<0.01	<0.01	<0.01	1.00	0.04	0.17	<0.01	0.02	0.64
D5%	0.07	1.00	1.00	0.05	<0.01	1.00	<0.01	1.00	1.00	1.00
D2%	0.05	1.00	1.00	0.02	<0.01	1.00	<0.01	1.00	0.44	1.00

Table 3.3 Statistical comparisons for PTV coverage between the RT techniques

3.9.2 Dose comparisons for SW-OARs

Results of doses to the SW-OARs are provided in table 3.4 and fig 3.5 respectively. The mandatory PlanSMPCM dose constraint was achieved with the dysphagia-optimised RT techniques in all ten patients and was lowest with the robustly optimised proton plan. The optimal PlanIPCM dose objective was achieved with Do-IMPT$_{RO}$ in six patients; in contrast to Do-IMRT and PTV-based proton plans where it was not achieved in any patient. There was no significant difference in mean dose to SPC between S-IMRT and S-IMPT, or between Do-IMRT and Do-IMPT$_{RO.}$ PTV-based Do-IMPT plans resulted in a significant increase in SPC dose compared to Do-IMRT. Mean PCM doses was significantly reduced using PBT and Do-IMRT compared to S-IMRT; statistically significant improvement in dose was observed with Do-IMPT$_{RO}$ as compared to PTV-based dysphagia-optimised proton and photon plans. Mean dose to SGL was significantly reduced with PBT and Do-IMRT compared to S-IMRT. The largest reduction was observed with Do-IMPT$_{RO}$, where the mean SGL dose was significantly less compared with all other techniques. There were no clinically significant reductions in mean dose to SPC, PCM and SGL when PBT was planned with 7 beams as compared with 3 beams.

Structure	Standard RT treatment		Dysphagia-optimised RT treatment			
	S-IMRT Gy (SD)	S-IMPT Gy (SD)	Do-IMRT Gy (SD)	Do-IMPT$_{RO}$ Gy (SD)	Do-IMPT$_{3B}$ Gy (SD)	Do-IMPT$_{7B}$ Gy (SD)
PlanSMPCM	56.5 (1.1)	56.2 (0.9)	49.3 (0.2)	48.5 (0.5)	49.7 (0.1)	49.5 (0.2)
PlanIPCM	48.3 (3.2)	41.9 (7.2)	24.5 (3.5)	20.5 (1.9)	23.1 (2.1)	22.8 (2.1)
SPC	59.1 (2.4)	59.0 (2.1)	54.5 (4.1)	54.6 (3.9)	55.4 (4.0)	55.2 (4.0)
MPC	57.4 (3.3)	56.6 (4.0)	49.7 (7.6)	46.7 (8.1)	46.4 (7.9)	48.3 (7.6)
IPC	48.4 (3.2)	43.1 (5.2)	24.8 (3.6)	20.8 (2.0)	23.4 (2.3)	23.1 (2.2)
PCM	56.6 (1.3)	55.4 (1.3)	47.4 (2.2)	46.0 (2.0)	47.3 (2.0)	47.2 (2.0)
SGL	52.0 (5.2)	49.1 (6.2)	48.8 (6.6)	43.8 (7.6)	47.1 (7.1)	46.3 (7.3)

Table 3.4 Comparison swallowing organs at risk mean dose parameters (standard deviation (SD)) for the 6 radiotherapy (RT) techniques

Do-IMPT$_{xB}$, dysphagia-optimised intensity modulated proton therapy with x number of beams; Do-IMPT$_{RO}$, robustly optimised Do-IMPT; Do-IMRT, dysphagia-optimised intensity modulated radiotherapy IPCM, inferior pharyngeal constrictor muscle; MPC, middle pharyngeal constrictor; PCM, pharyngeal constrictor muscle; SMPCM, superior and middle pharyngeal constrictor muscle; SPC, superior pharyngeal constrictor muscle; S-IMRT, standard intensity-modulated RT

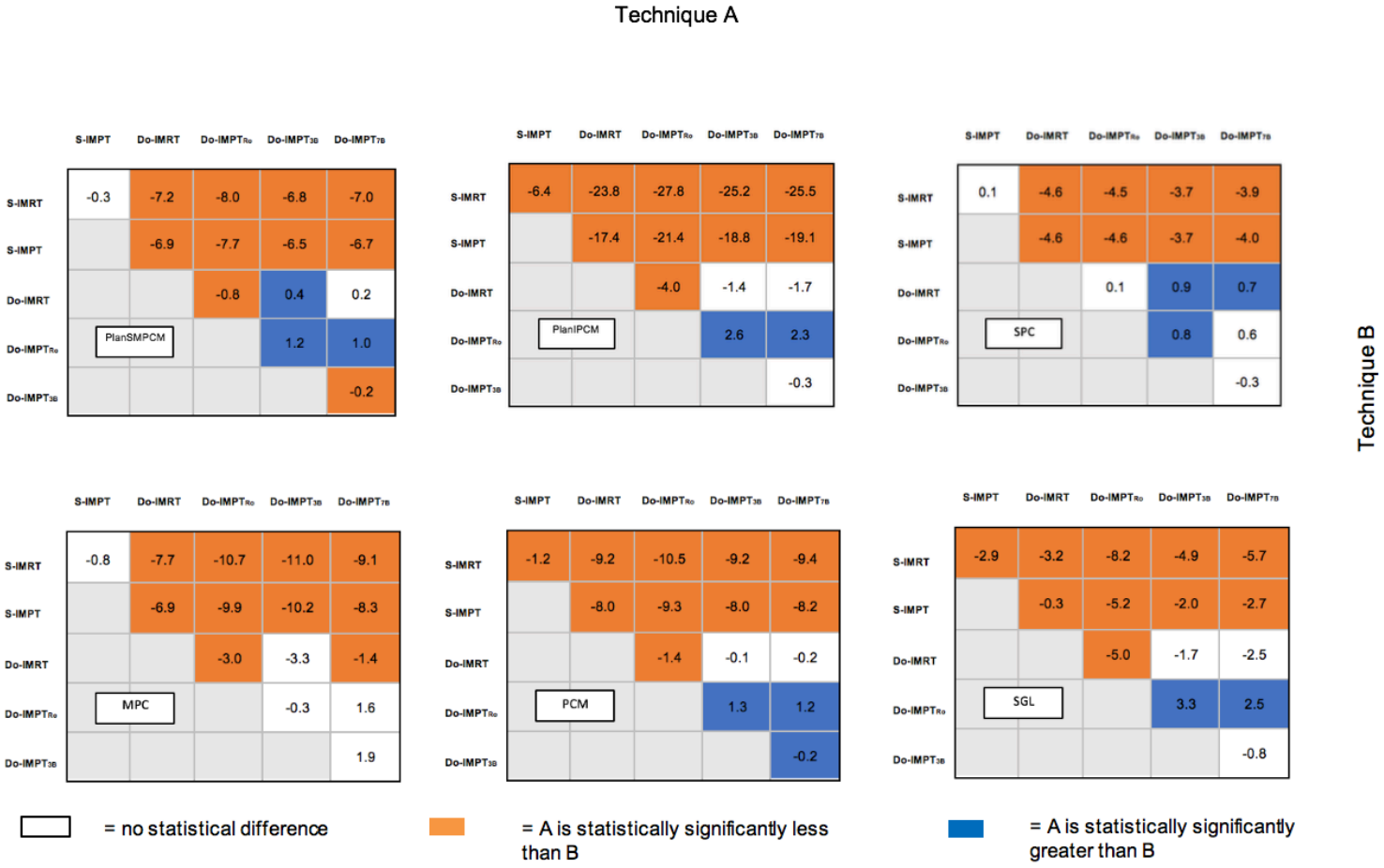


Figure 3.5 Differences in mean swallowing organ at risk doses between radiotherapy (RT) techniques (technique A, technique B)

3.9.3 Dose comparisons for NSW-OARs

The mean (SD) dose and statistical results for SC, BS, and the parotids for the six RT techniques are tabulated in table 3.5 and 3.6 respectively. The delivered dose to the SC and BS were lower with PBT compared to IMRT, some of which were statistically significant. The three dysphagia-optimised proton plans delivered lower mean dose to the contralateral parotid gland compared to the corresponding photon plan, with the largest difference observed with the RO plan which was statistically significant. No significant difference was seen in the ipsilateral parotid mean dose between the photon and proton plans. Increasing the number of fields for the PTV-based PBT did not improve OAR sparing.

Structure	Standard RT treatment		Dysphagia-optimised RT treatment			
	S-IMRT Gy (SD)	S-IMPT Gy (SD)	Do-IMRT Gy (SD)	Do-IMPT$_{RO}$ Gy (SD)	Do-IMPT$_{3B}$ Gy (SD)	Do-IMPT$_{7B}$ Gy (SD)
SC						
Dmax	43.1 (2.1)	36.1 (5.1)	40.3 (2.9)	39.5 (3.6)	36.4 (4.9)	39.4 (3.3)
Dmax$_{1cc}$	38.7 (1.8)	25.0 (5.2)	35.2 (1.7)	29.6 (2.8)	26.0 (5.5)	32.3 (1.7)
BS						
Dmax	44.7 (2.2)	38.2 (4.4)	41.6 (3.5)	38.9 (3.0)	38.1 (4.8)	40.5 (2.7)
Dmax$_{1cc}$	39.0 (2.7)	26.1 (5.6)	35.8 (2.4)	28.2 (5.8)	26.1 (5.9)	31.6 (3.2)
CL parotid						
Mean dose	26.7 (3.9)	25.5 (2.8)	25.9 (2.4)	23.6 (3.3)	24.8 (2.8)	24.5 (3.5)
IL parotid						
Mean dose	41.3 (4.6)	38.4 (5.8)	38.2 (3.2)	37.5 (5.8)	37.5 (5.8)	37.9 (5.3)

Table 3.5 Comparison non - swallowing organs at risk mean dose parameters (standard deviation (SD)) for the 6 radiotherapy (RT) techniques

Structure	S-IMRT v S-IMPT	S-IMRT v Do-IMRT	S-IMRT v Do-IMPT$_{3B}$	S-IMRT v Do-IMPT$_{7B}$	Do-IMRT v S-IMPT	Do-IMRT v Do-IMPT$_{3B}$	Do-IMRT v Do-IMPT$_{7B}$	S-IMPT v Do-IMPT$_{3B}$	S-IMPT v Do-IMPT$_{7B}$	Do-IMPT$_{3B}$ v Do-IMPT$_{7B}$	S-IMPT v Do-IMPT$_{RO}$	S-IMRT v Do-IMPT$_{RO}$	Do-IMRT v Do-IMPT$_{RO}$	Do-IMPT$_{RO}$ v Do-IMPT$_{3B}$	Do-IMPT$_{RO}$ v Do-IMPT$_{7B}$
	p-values, shaded boxes indicate statistically significant difference														
Spinal Cord															
Dmax	0.04	0.43	0.04	0.05	0.051	0.09	1.00	1.00	0.77	0.91	0.30	0.20	1.00	0.35	1.00
Dmax$_{1cc}$	<0.01	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.052	0.04	0.15	0.71	<0.01	<0.01	1.00	0.09
Brainstem															
Dmax	<0.01	0.03	<0.01	<0.01	0.12	0.06	1.00	1.00	0.34	0.50	1.00	<0.01	0.03	1.00	0.09
Dmax$_{1cc}$	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	1.00	0.02	0.03	1.00	<0.01	<0.01	1.00	0.43
CL parotid															
Mean	0.35	1.00	0.08	0.03	1.00	0.07	0.41	0.10	0.23	1.00	<0.01	<0.01	<0.01	0.03	0.41
IL parotid															
Mean	0.07	0.01	0.01	<0.01	1.00	1.00	1.00	0.40	1.00	1.00	0.12	<0.01	1.00	1.00	1.00

Table 3.6 Comparison non - swallowing organs at risk mean dose parameters (standard deviation (SD)) for the 6 radiotherapy (RT) techniques

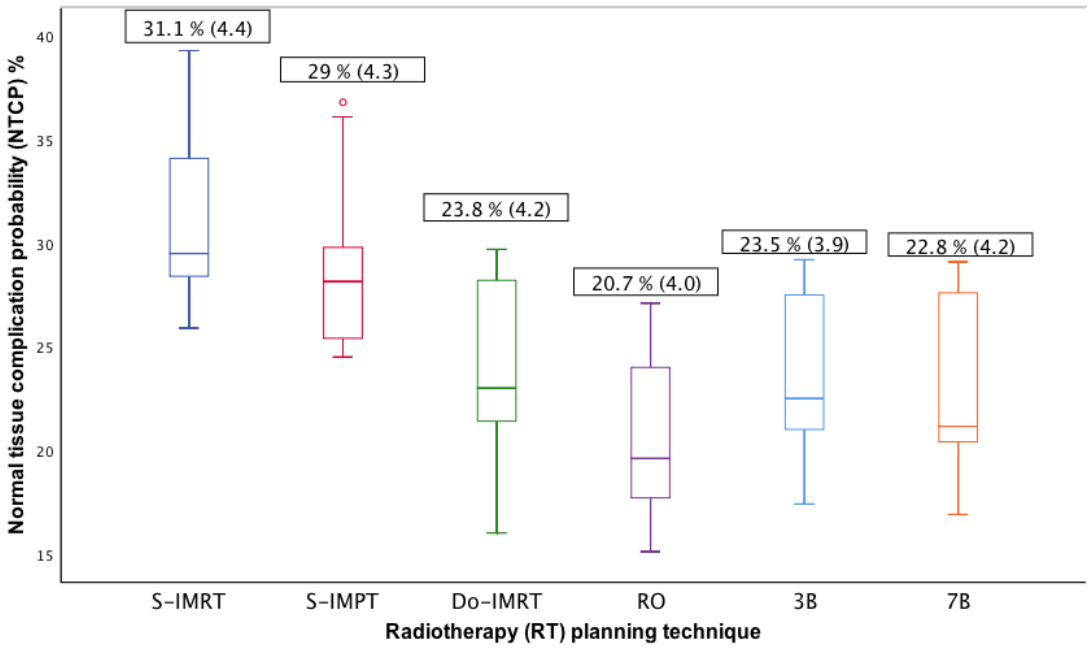
3.9.4 Comparisons of NTCP for RAD_{6M}

The NTCP results for the physician - and patient - scored RAD_{6M} with the different planning techniques are illustrated in figures 3.6 – 3.9.

The estimated risk of RAD_{6M} was lowest with Do-$IMPT_{RO}$ compared to the other 5 RT techniques for all predictive models. Compared with S-IMRT, the average probability of physician-scored RAD_{6M} was reduced by 2.1 % (95 % CI 0.32 % – 4.0 %) with S-IMPT, 7.3 % (4.6 % – 10.1 %) with Do-IMRT, 7.6 % (5.7 % – 9.5 %) with Do-$IMPT_{3B}$, 8.3 % (6.4 % - 10.3%) with Do-$IMPT_{7B}$ and 10.4 % (8.6 % - 12.2 %) with Do-$IMPT_{RO}$ respectively.

Mean $\Delta NTCP_{DO\_RO}$ was 3.1 % (range 0.8 % - 5.7 %) and did not reach our pre-defined threshold of $\geq$ 5 % reduction in toxicity with dysphagia-optimised PBT across the entire cohort. Likewise the mean $\Delta NTCP_{standard}$ of 2.1 % (- 0.1 % - 4.3 %) was not clinically significant. Mean $\Delta NTCP_{DO\_PTV}$ was negligible with both PTV-based proton plans, and the addition of more fields did not reduce the risk of persistent swallowing dysfunction.

For patient-scored toxicities, while the predicted risk of toxicities was again lowest with Do-$IMPT_{RO}$, the benefits compared to the corresponding Do-IMRT plan was less than < 5 % and therefore not clinically meaningful.



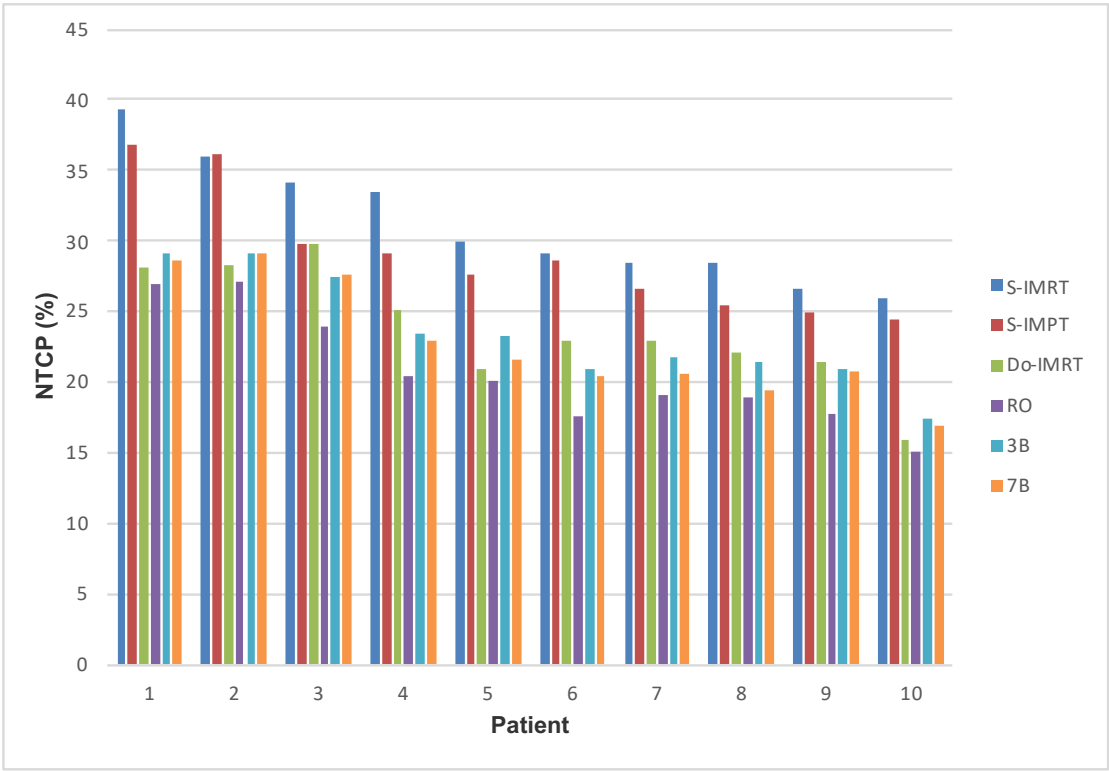
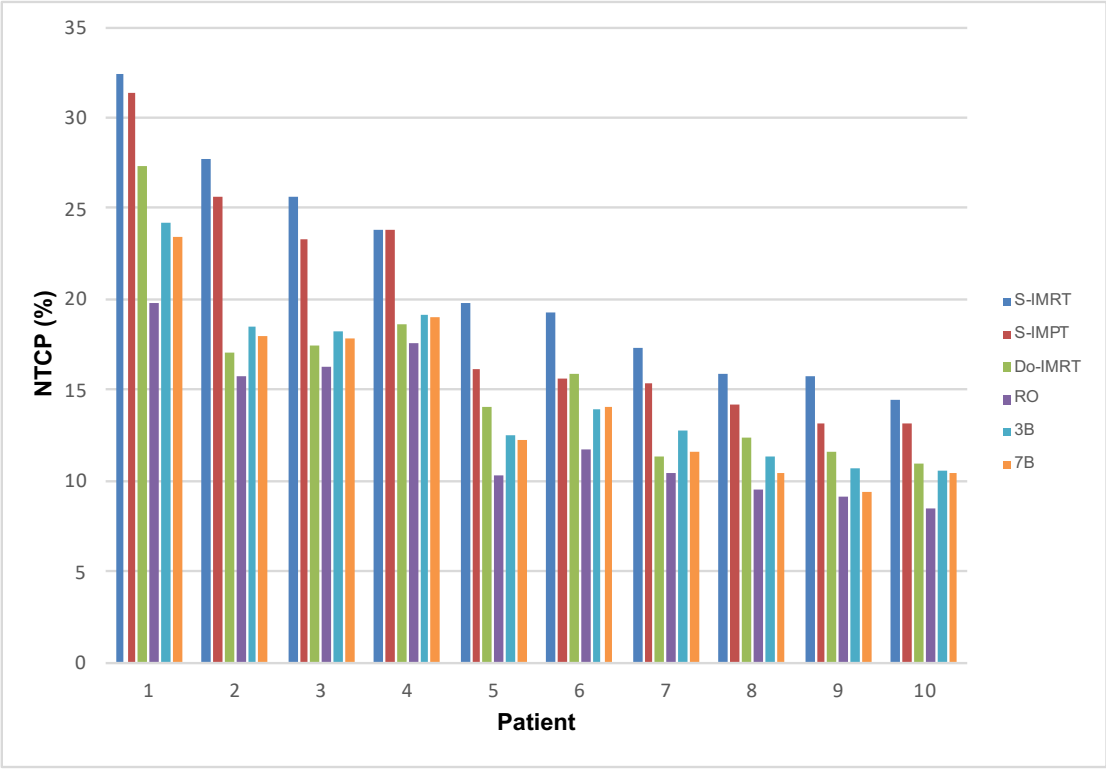


Figure 3.6(A) Boxplots of the normal tissue complication probabilities (NTCP) of physician-scored RTOG ≥ 2 dysphagia at 6 months for the various radiotherapy (RT) techniques. Mean (standard deviation) NTCP for each technique is provided above the respective boxplot, and (B) NTCPs for individual patients with the 6 RT plans

RO, robustly optimised dysphagia-optimised intensity modulated proton plan (DO-IMPT); RTOG, RT oncology group; S-IMPT, standard IMPT; S-IMRT, standard intensity modulated RT; Do-IMRT, dysphagia-optimised IMRT; 3B, planning target volume (PTV)- based Do-IMPT with 3 beams; 7B, PTV-based Do-IMPT with 7 beams



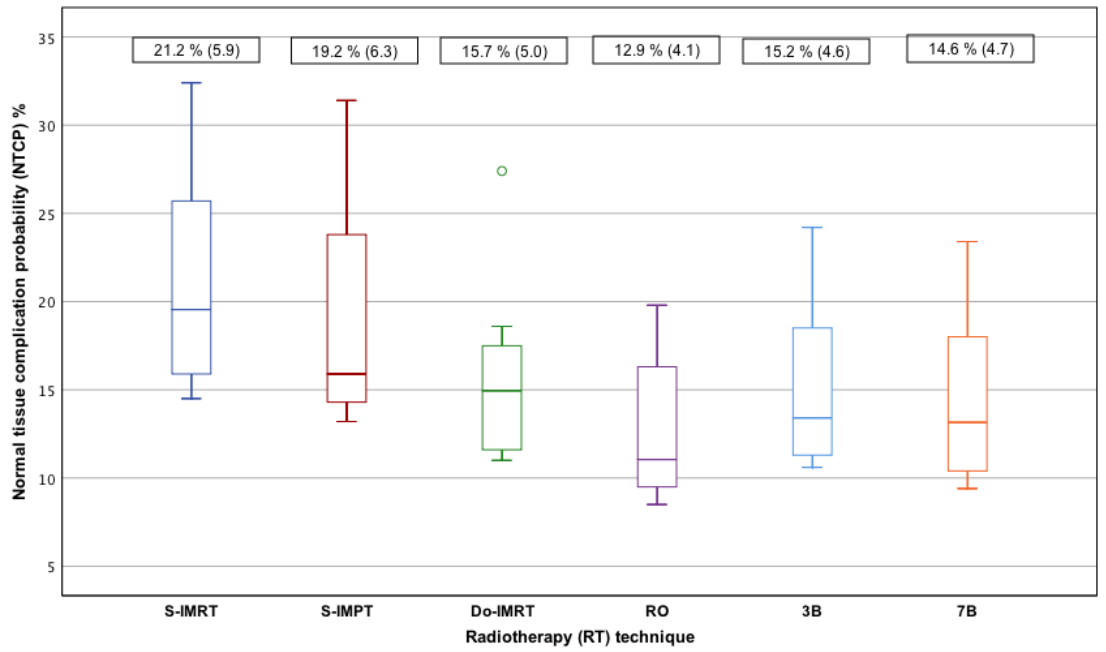
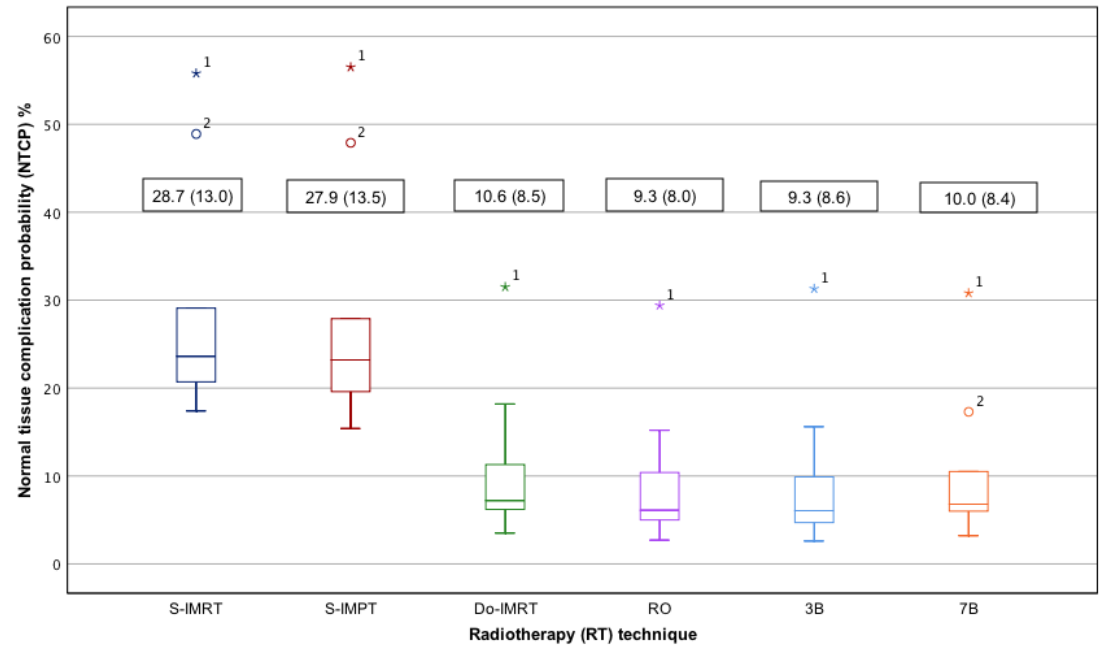


Figure 3.7 A) Boxplots of NTCPs of patient-rated moderate to severe difficulties in swallowing solids at 6 months for the various RT techniques. Mean (standard deviation) NTCP for each technique is provided above the respective boxplot, and (B) NTCPs for individual patients with the 6 RT plans



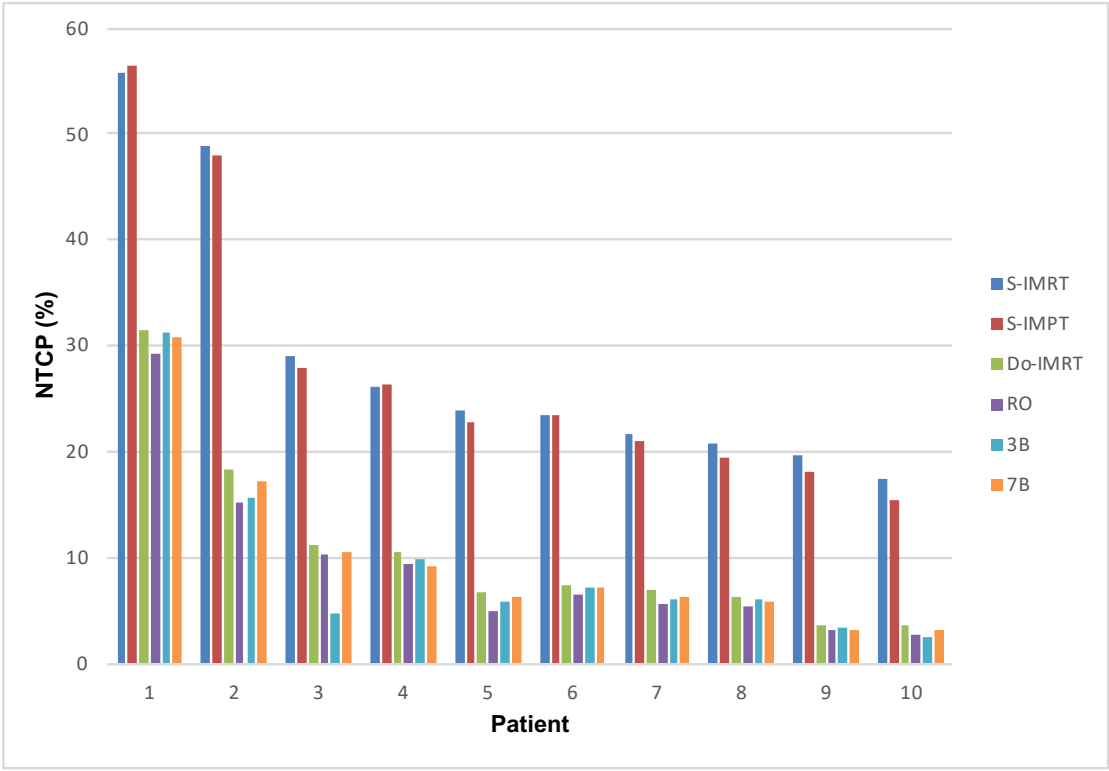
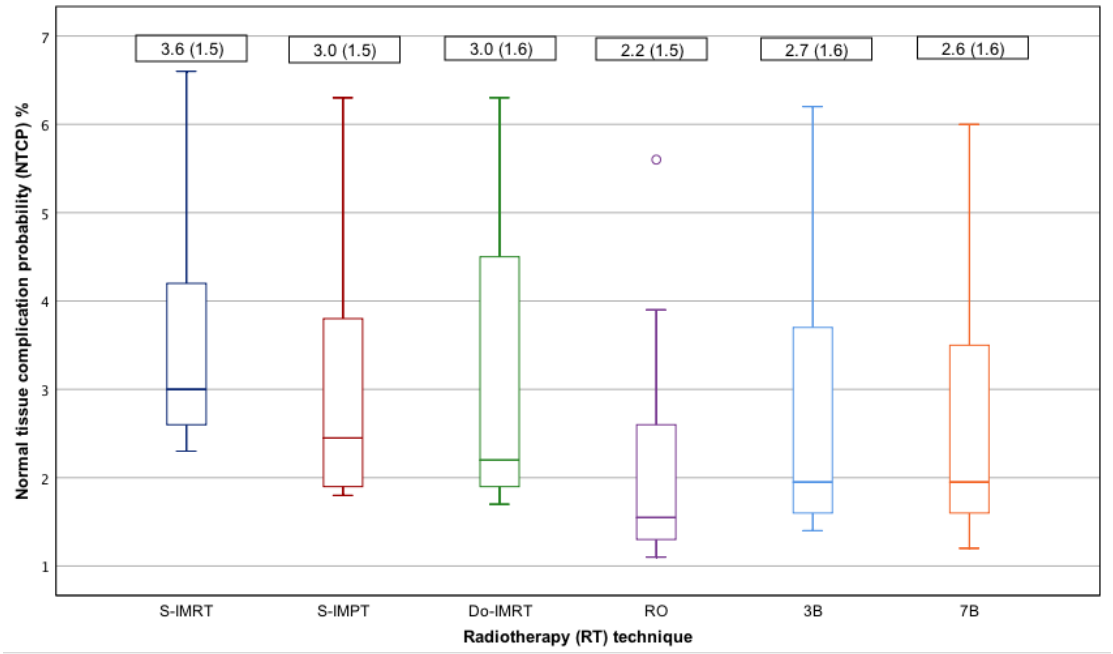


Figure 3.8(A) Boxplots of NTCPs of patient-rated moderate to severe difficulties in swallowing soft food at 6 months for the various RT techniques. Mean (standard deviation) NTCP for each technique is provided above the respective boxplot. ° and * represent outliers and extreme data points.(B) NTCPs for individual patients with the 6 RT plans



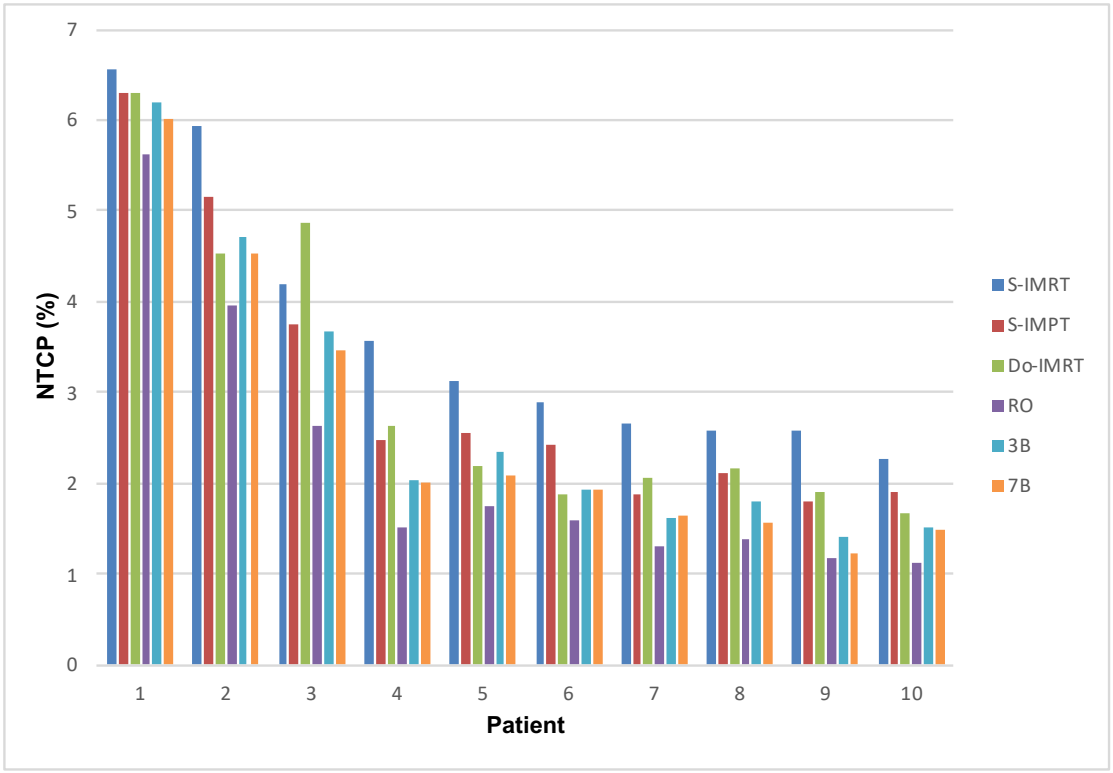


Figure 3.9(A) Boxplots of NTCPs of patient-rated moderate to severe difficulties in swallowing liquid food at 6 months for the various RT techniques. Mean (standard deviation) NTCP for each technique is provided above the respective boxplot. ° represent outliers and extreme data points.(B) NTCPs for individual patients with the 6 RT plans

3.9.5 Integral dose:

The results of the mean dose to the body tissue, together with percentage volume of the body receiving 5 Gy and 30 Gy respectively with the different RT techniques are tabulated in table 3.7.

For all three studied parameters, the integral dose was significantly less with PBT compared to the IMRT plans ($p < 0.001$). The lowest deposited dose was observed with Do-IMPT$_{RO}$. Increasing the number of fields for the PTV-based PBT resulted in larger amount of total dose to the patient.

Structure	Body tissue		V5		V30	
	Mean, Gy-L	Range, Gy-L	%	Range	%	Range
S-IMRT	163.8	139.6 – 213.8	37.3	28.9 – 44.0	15.2	12.6 – 18.3
S-IMPT	115.2	101.1 – 139.1	26.8	21.8 – 30.5	10.9	9.0 – 13.0
Do-IMRT	156.9	130.2 – 201.8	37.2	28.9 – 43.7	13.7	11.1 – 16.6
Do-IMPT$_{RO}$	111.2	96.7 – 135.4	26.4	21.9 – 31.2	10.5	8.8 – 12.6
Do-IMPT$_{3B}$	113.8	99.7 – 138.5	26.5	21.6 – 30.3	10.7	8.9 – 12.8
Do-IMPT$_{7B}$	125.5	108.6 – 153.6	32.7	25.4 – 30.2	11.0	9.1 – 13.1

Table 3.7 Integral doses delivered with the six radiotherapy (RT) techniques.

Gy-L, Gray-litre; VX – percentage volume of body receiving X Gy

3.10 Discussion:

The use of PBT as a toxicity-mitigating RT modality in LA-HNC has been of significant interest in recent years[19, 20]. To date, there is no prospective randomised evidence confirming the superiority of PBT over IMRT in this context. The primary aim of this planning study, comparing six photon and proton RT techniques, was to evaluate the potential swallow-sparing benefits of Do-IMPT over Do-IMRT in patients with LA-OPC. Our results demonstrate that the risk of physician – scored and patient – reported RAD_{6M} was lowest with Do-$IMPT_{RO,}$ with comparable target volume coverage. However, the magnitude of this benefit relative to Do-IMRT varied among individual patients in this cohort, and did not meet our predetermined minimal clinically important difference threshold of > 5 % mean change in predicted treatment-related toxicity across the entire cohort between dysphagia-optimised photon and proton plans. Two patients had a mean $\Delta NTCP_{DO\_RO}$ of > 5 %, suggesting careful selection may help identify patients likely to benefit from PBT. In contrast, PTV-based Do-IMPT plans did not show any meaningful improvement in long-term swallowing function compared to Do-IMRT.

To our knowledge, there have been no published papers comparing both RO- and PTV- based swallow-sparing MFO-IMPT with the corresponding VMAT plan in reducing RAD in a homogenous cohort of LA-OPC. Previous planning studies have investigated the swallow-sparing benefits of PBT relative to fixed-field IMRT in a heterogeneous population of HNC.

Van der Laan et al showed an incremental sparing of mean RT dose to SGL with 7-beam swallow-sparing IMPT (SW-IMPT$_{7B}$) (40 Gy) relative to 3-beam SW-IMPT (SW-IMPT$_{3B}$) (46 Gy) and SW-IMRT (40 Gy) in a comparison study of twenty-five patients with LA-OPC (21/25) and HPC (4/25)[21]. This reduction in mean dose to SGL with the PTV-based proton plans translated into reduced estimated risk of RAD$_{6M}$. Compared to SW-IMRT, the probability of physician-scored swallowing dysfunction was reduced on average by 8 % and 3 % with SW-IMPT$_{7B}$ and SW-IMPT$_{3B}$ respectively. However, this clinical benefit did not extend to all patients; the respective ranges for the NTCPs were wide (SW-IMRT, 17 % - 62 % versus SW-IMPT$_{3B}$, 6 % - 62 % versus SW-IMPT$_{7B}$, 4 % - 62 %). It is also noteworthy that their improvements in predicted swallowing function were driven by reductions in mean dose to SGL alone, as there was no difference in mean dose to SPC with the three RT techniques. This lack of sparing of dose to the constrictor with SW-PBT relative to the corresponding photon technique was also observed in our study. On another note, the relative reduction in mean dose to SGL noted with Do-IMPT$_{RO}$ in our study may be improved further by setting a planning objective function to this SW-OAR during optimisation; this could translate into substantial reductions in predicted swallowing dysfunction with our robustly optimised proton technique.

Similarly, Van Dijk et al showed that robustly optimised SW-IMPT reduced NTCP values for RAD$_{6M}$ by 8% compared to fixed-field IMRT in a heterogeneous cohort of 10 patients with HNC[22]. Baseline characteristics

varied widely, including patients with T1N0 oropharynx and early laryngeal cancer.

A significant finding in our study was the lower mean dose to the contralateral parotid with Do-IMPT$_{RO}$ compared to Do-IMRT; dose reductions with PTV-based proton plans were not statistically significant. Xerostomia has been shown to contribute significantly to patients' perception of dysphagia[23], and therefore the relative parotid-sparing effect of Do-IMPT$_{RO}$ is likely to further augment the swallow-sparing benefits of PBT compared to Do-IMRT. Our contralateral parotid mean doses with the robustly optimised proton plans were, however, higher than those observed in the study by Van Der Laan's group. Possible explanations for this discrepancy is that the Dutch study also included primary hypopharyngeal tumours that are likely to be situated away from the salivary glands, and the higher priority placed on sparing dose to the constrictor muscle with our dysphagia-optimised plans.

Another significant finding in our study was the larger integral dose delivered with photons, in particular with Do-IMRT compared to robustly optimised proton plans; this is consistent with published outcomes[24]. Such increase in the low dose radiation bath may be associated with a higher risk of secondary radiation-induced malignancies[25, 26], which is of particular relevance in the relatively younger population with good prognosis HPV-related tumours. Additionally, as can be seen from Fig, Do-IMPT$_{RO}$ can minimise dose delivered to the anterior oral cavity, potentially reducing the severity of acute toxicity and

thereby influence the prevalence of late RAD. The oral cavity was not delineated in our study, and therefore accurate dose delivered to this structure was not available for detailed analysis.

This study has shown that robustly optimised Do-IMPT is dosimetrically superior to PTV-based Do-IMPT, consistent with published results. It could be argued that the improved normal tissue sparing is a consequence of smaller target volumes due to omission of the CTV-PTV margins. There is, therefore, a need to determine whether the RO plan is able to maintain plan quality in the presence of uncertainties and compare it to PTV-based PBT. This has been investigated in the next chapter. In addition, we have shown that increasing the number of fields for PTV-based optimisation does not improve sparing of dose to OARs. Previous studies have reported conflicting results regarding the benefits of more fields to limit dose to OARs, probably a consequence of different planning techniques.[22, 27]

The limitations of the planning study of chapter 2 apply to this study as well. A further limitation of this study is that proton plans were based on physical dose, and not corrected for RBE. A particular concern with RBE uncertainty is the potential increase in dose to OARs situated distal to the target volume, and its incorporation is therefore critical during delivery of a clinical plan. However, it is less relevant for a hypothesis-testing comparative planning study such as ours. Furthermore, the gantry angles for the 3-beam PBT meant that no critical

structure was distal to the target volume, and therefore the perceived risk of over-dosage to critical OARs is unlikely.

3.11 Conclusions

Do-IMPT$_{RO}$ reduced the risk of RAD$_{6M}$ compared to Do-IMRT in LA-OPC, though this did not lead to clinically relevant improvements for all patients in this cohort. The additional benefits provided by the robustly optimised PBT over Do-IMRT in the form of improved parotid sparing could further reduce the risk of xerostomia, and potentially improve long-term swallowing function and long-term HR-QoL. The definitive role of PBT as a toxicity-mitigating strategy needs to be investigated in the setting of a prospective randomised study.

3.12 References list

1. Kraan, A.C., et al., *Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range, and setup errors.* Int J Radiat Oncol Biol Phys, 2013. **87**(5): p. 888-96.

2. Thomson, D.J., et al., *The Impact of Anatomic Change on Pencil Beam Scanning in the Treatment of Oropharynx Cancer.* International Journal of Particle Therapy, 2015. **2**(2): p. 394-403.

3. Tangsriwong, K., et al., *Potential Impact of Daily Setup Variation on Pencil-Beam Scanning for Head and Neck Cancer.* International Journal of Particle Therapy, 2015. **2**(1): p. 44-49.

4. McGowan, S.E., N.G. Burnet, and A.J. Lomax, *Treatment planning optimisation in proton therapy.* Br J Radiol, 2013. **86**(1021): p. 20120288.

5. Engelsman, M., M. Schwarz, and L. Dong, *Physics Controversies in Proton Therapy.* Seminars in Radiation Oncology, 2013. **23**(2): p. 88-96.

6. Hoogeman, M. *Robust planning and margin concepts : What is the way to go*. [PowerPoint Slides] 2015 27/05/2019; Available from: https://www.oncoray.de/fileadmin/files/Veranstaltungen/4DTPWS/talks/Session2_talk3_MSHoogeman_publish.pdf.

7. Mitin, T. and A.L. Zietman, *Promise and pitfalls of heavy-particle therapy.* J Clin Oncol, 2014. **32**(26): p. 2855-63.

8. Unkelbach, J. and H. Paganetti, *Robust Proton Treatment Planning: Physical and Biological Optimization.* Seminars in Radiation Oncology, 2018. **28**(2): p. 88-96.

9. Baumann, M., et al., *Radiation oncology in the era of precision medicine.* Nature Reviews Cancer, 2016. **16**: p. 234.

10. Pflugfelder, D., J.J. Wilkens, and U. Oelfke, *Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy.* Physics in Medicine and Biology, 2008. **53**(6): p. 1689-1700.

11. Fredriksson, A., A. Forsgren, and B. Hardemark, *Minimax optimization for handling range and setup uncertainties in proton therapy.* Med Phys, 2011. **38**(3): p. 1672-84.

12. Unkelbach, J., et al., *Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning.* Med Phys, 2009. **36**(1): p. 149-63.

13. Kooy, H.M. and C. Grassberger, *Intensity modulated proton therapy.* Br J Radiol, 2015. **88**(1051): p. 20150195.

14. Paganetti, H., *Range uncertainties in proton therapy and the role of Monte Carlo simulations.* Phys Med Biol, 2012. **57**(11): p. R99-117.

15. Frank, S.J., et al., *Multifield Optimization Intensity Modulated Proton Therapy for Head and Neck Tumors: A Translation to Practice.* International Journal of Radiation Oncology • Biology • Physics, 2014. **89**(4): p. 846-853.

16. Statistics, L. *Repeated Measures ANOVA*. 25/05/2019]; Available from: https://statistics.laerd.com/statistical-guides/repeated-measures-anova-statistical-guide.php.

17. Marshall, E. *Repeated measures ANOVA in SPSS*. Available from: https://www.sheffield.ac.uk/polopoly_fs/1.531222!/file/MASH_repeated_measures_ANOVA_SPSS.pdf.

18. Statistics, L. *Friedman test in SPSS Statistics*. 25/05/2019]; Available from: https://statistics.laerd.com/spss-tutorials/friedman-test-using-spss-statistics.php.

19. Blanchard, P., et al., *Proton Therapy for Head and Neck Cancers.* Semin Radiat Oncol, 2018. **28**(1): p. 53-63.

20. Sio, T.T., et al., *Intensity Modulated Proton Therapy Versus Intensity Modulated Photon Radiation Therapy for Oropharyngeal Cancer: First Comparative Results of Patient-Reported Outcomes.* Int J Radiat Oncol Biol Phys, 2016. **95**(4): p. 1107-14.

21. van der Laan, H.P., et al., *The potential benefit of swallowing sparing intensity modulated radiotherapy to reduce swallowing dysfunction: an*

in silico planning comparative study. Radiother Oncol, 2012. **103**(1): p. 76-81.

22. van Dijk, L.V., et al., *Robust Intensity Modulated Proton Therapy (IMPT) Increases Estimated Clinical Benefit in Head and Neck Cancer Patients.* PLoS One, 2016. **11**(3): p. e0152477.

23. Vainshtein, J.M., et al., *Impact of xerostomia on dysphagia after chemotherapy-intensity-modulated radiotherapy for oropharyngeal cancer: Prospective longitudinal study.* Head Neck, 2016. **38 Suppl 1**: p. E1605-12.

24. van der Laan, H.P., et al., *The potential of intensity-modulated proton radiotherapy to reduce swallowing dysfunction in the treatment of head and neck cancer: A planning comparative study.* Acta Oncol, 2013. **52**(3): p. 561-9.

25. Hall, E.J. and C.S. Wuu, *Radiation-induced second cancers: the impact of 3D-CRT and IMRT.* Int J Radiat Oncol Biol Phys, 2003. **56**(1): p. 83-8.

26. Ruben, J.D., et al., *The effect of intensity-modulated radiotherapy on radiation-induced second malignancies.* Int J Radiat Oncol Biol Phys, 2008. **70**(5): p. 1530-6.

27. Barten, D.L., et al., *Comparison of organ-at-risk sparing and plan robustness for spot-scanning proton therapy and volumetric modulated arc photon therapy in head-and-neck cancer.* Med Phys, 2015. **42**(11): p. 6589.

4 Chapter 4: An analysis of proton plan robustness for LA-OPC: Comparing robust optimised Do-IMPT and PTV-based Do-IMPT planning

4.1 Introduction

CTV-based RO and PTV-based RE are the two approaches used to account for treatment uncertainties in MFO-IMPT planning[1-3]. We previously demonstrated that RO-based Do-IMPT improves plan optimality compared to the PTV approach for LA-OPC by reducing dose to OARs with comparable target volume coverage. It may be argued that the beneficial effect may be predominantly due to optimising the plan to the smaller CTV, wherein the larger dose gradient between target volume and OAR allows for improved OAR sparing, rather than due to superior optimisation technique. The main difference between the 2 techniques is that RO optimises the CTV as it takes into account RU and SE during optimisation, while RE considers only SE similar to IMRT. RO is therefore considered to be more resilient to proton beam treatment errors[4-7], which is particularly relevant for MFO-IMPT where deviations in the position of each spot in the target during treatment from the nominal plan could potentially result in under – or over-dosage of target and/or increased toxicity to non-target normal tissue. RO is comparatively more expensive, and requires significant computational time relative to the traditional

planning technique with PTV. It is, therefore, pertinent to quantify the expected difference in delivered dose from the planned dose distribution with both dysphagia-optimising PBT planning techniques, if uncertainties are introduced during the course of a treatment.

This chapter investigates the resilience of the nominal dose distributions for the different dysphagia-optimised proton plans, assessing the effects of SE and RU on the dose distribution for target and normal tissues.

4.2 Aims of this study

The first aim of this study was to investigate the sensitivity of robustly optimised and PTV-based Do-IMPT plans to SE and RU. A second aim was to evaluate whether increasing the number of fields with the PTV-based approach contributed to increased resilience of the nominal dose distribution.

4.3 Null hypothesis

There is no difference in the robustness of RO-based and PTV-based Do-IMPT plans. Increasing the number of fields do not improve robustness for margin-based PBT.

4.4 Materials and methods

The nominal Do-IMPT$_{3B}$, Do-IMPT$_{7B}$, and Do-IMPT$_{RO}$ plans generated for the 10 LA-OPC patients in chapter 3 were evaluated for robustness to SE and RU in this study. All simulated plans were created on RS, using the provided 'compute perturbed dose' tool. Density changes and isocentre shifts were entered into this tool, which then recalculated a new dose distribution for each perturbed plan.

4.4.1 Simulation of dose uncertainties

The main purpose of this work was a relative comparison of robustness of swallow-sparing proton plans with different optimisation methods and beam designs. In theory, it is possible to assess the influence of all possible uncertainty scenarios on the planned dose distribution. However, such evaluation would require substantial computation and data processing time, which would not be possible within the available timeframe for this thesis. Instead, 11 different dose uncertainties incorporating either SE or RU and a combination of both were investigated for each proton plan, shown in table 4.1. SE were simulated by shifting the isocentre by 3 mm, which is the CTV-PTV margin at our centre, in three directions X (left shift), Y (superior shift), and Z (anterior shift). The influence of RU on proton plans was evaluated by introducing $\pm$ 3.5 % errors.

Perturbation	Range uncertainty error	Setup error, mm		
		X	Y	Z
1	0	3	0	0
2	0	0	3	0
3	0	0	0	3
4	3.5 %	0	0	0
5	3.5 %	3	0	0
6	3.5 %	0	3	0
7	3.5 %	0	0	3
8	-3.5 %	0	0	0
9	-3.5 %	3	0	0
10	-3.5 %	0	3	0
11	-3.5 %	0	0	3

Table 4.1 Table showing the different uncertainty scenarios evaluated for each nominal plan.

4.4.2 Evaluation of plan robustness

Dose distributions of CTVs and OARs were assessed, rather than the PTVs and PRVs. For each perturbed plan, the CTV65 $D_{99\%}$, $D_{95\%}$, $D_{50\%}$, $D_{5\%}$, and $D_{2\%}$ and CTV54 $D_{50\%}$ were recorded. In addition, maximum dose to critical OARs, along with mean doses to the parotids, planSMPCM, planIPCM, SPC, SGL and PCM were also determined. The above dataset was subsequently exported from RS to an excel spreadsheet for evaluation.

4.4.2.1 Qualitative assessment of plan robustness

For visual assessment, the 11 perturbed DVHs, together with the nominal plan, were plotted as DVH bands. The DVH band for an organ or volume of interest would demonstrate the range of dose distribution possible for that structure in relation to the nominal plan in the presence of uncertainties, where a narrow band would indicate high robustness and a wide band implying increased sensitivity to uncertainties. In addition, such bands guide plan evaluation by providing a quick, qualitative assessment of the impact of the errors on dose distribution. For instance, a DVH band that lies predominantly to the left of the nominal CTV65 DVH would indicate under-dosage to the high-dose target volume under most uncertainties, and therefore prompt further re-optimisation of the nominal plan to improve robustness. I generated the DVH bands on a python script created by our physicist Alex Dunlop.

4.4.2.2 Quantitative assessment of plan robustness

DVH bandwidth:

For each nominal proton plan, the difference between the minimum and maximum value of a point of interest between the perturbed plans for a given target volume or organ, such as $D_{x\%}$ or mean dose, was defined as the bandwidth at that point. Three bandwidths were generated for a given

structure, one for the robustly optimised and two for the PTV-based optimisation plans; these were then compared to give a relative assessment of each plan's robustness. A broad bandwidth indicates that the dose distribution varies widely under uncertainty at that point, and the nominal plan is therefore less robust.

CTV65$_{D95\%}$ coverage:

The CTV – PTV margin is expected to deliver atleast 95 % of the prescribed dose to the CTV65 $_{D95\%}$ in 90 % of the population, as per the methodology described by van Herk et al[8]. Adopting this concept as a key determinant of plan robustness, the perturbed doses at CTV65$_{D95\%}$ were analysed for the different proton plans, and the percentage of simulations that complied with the above mandatory dose constraint were recorded as a measure of the success of the planning technique robustness.

CTV54$_{D50\%}$ coverage:

The dose to this mandatory objective function under different simulations was determined, and the percentage of perturbations for each nominal plan meeting the constraint was determined.

RU-related uncertainty:

The impact of RU-related errors on the delivered dose is a major parameter taken into the account in the RO algorithm, and is considered not to be adequately covered by the PTV concept. We therefore analysed the effect of RU alone, in the absence of SE, on the dose distributions for target volumes in our simulated plans for the RO-based and PTV- based plans.

NSW-OARs:

The maximum dose to the SC and BS, and the mean dose to the parotids under different perturbations for each nominal plan were analysed. For the SC and BS, the percentages of simulations not complying with the mandatory dose constraint was evaluated.

SW-OARs:

Limiting the mean dose to planSMPCM to < 50 Gy was a mandatory dose objective for our dysphagia-optimised plans and achieved in all three nominal proton plans. The variation in dose delivered to this structure under different scenarios of uncertainties was compared between the plans, and the percentage of simulations which did not meet this planning objective was determined. Similar analysis was performed for PlanIPCM, though meeting its dose constraint was not obligatory.

Variation in NTCPs for physician-scored RAD_{6M} ***:***

Differences in late swallowing dysfunction under various conditions of uncertainties for the three plans were determined and compared.

4.5 Statistical analysis

Statistical analysis was performed on SPSS version 25, using the one-way repeated measures ANOVA, with all pairwise comparisons tested using Bonferroni's multiple comparison test wherein p-values < 0.05 indicated statistical significance.

4.6 Results

4.6.1 Qualitative assessments with DVH bands

Representative illustrations of the DVH bands for the target volumes and OARs for one LA-OPC case are shown in Figure 4.1 – 4.4 respectively. For all ten patients, the DVH bands for CTV65 and CTV54 were narrower with the RO plan relative to the PTV-based plans, indicating its superior robustness. Increasing the number of fields for PTV-based PBT appeared to make the plans more robust, evidenced by the wider bands for Do-IMPT$_{3B}$ relative to Do-IMPT$_{7B}$. The DVH bands for SW-OARs and NSW-OARs were comparable between Do-IMPT$_{RO}$ and Do-IMPT$_{7B}$, and appeared perceptibly better than Do-IMPT$_{3B}$.

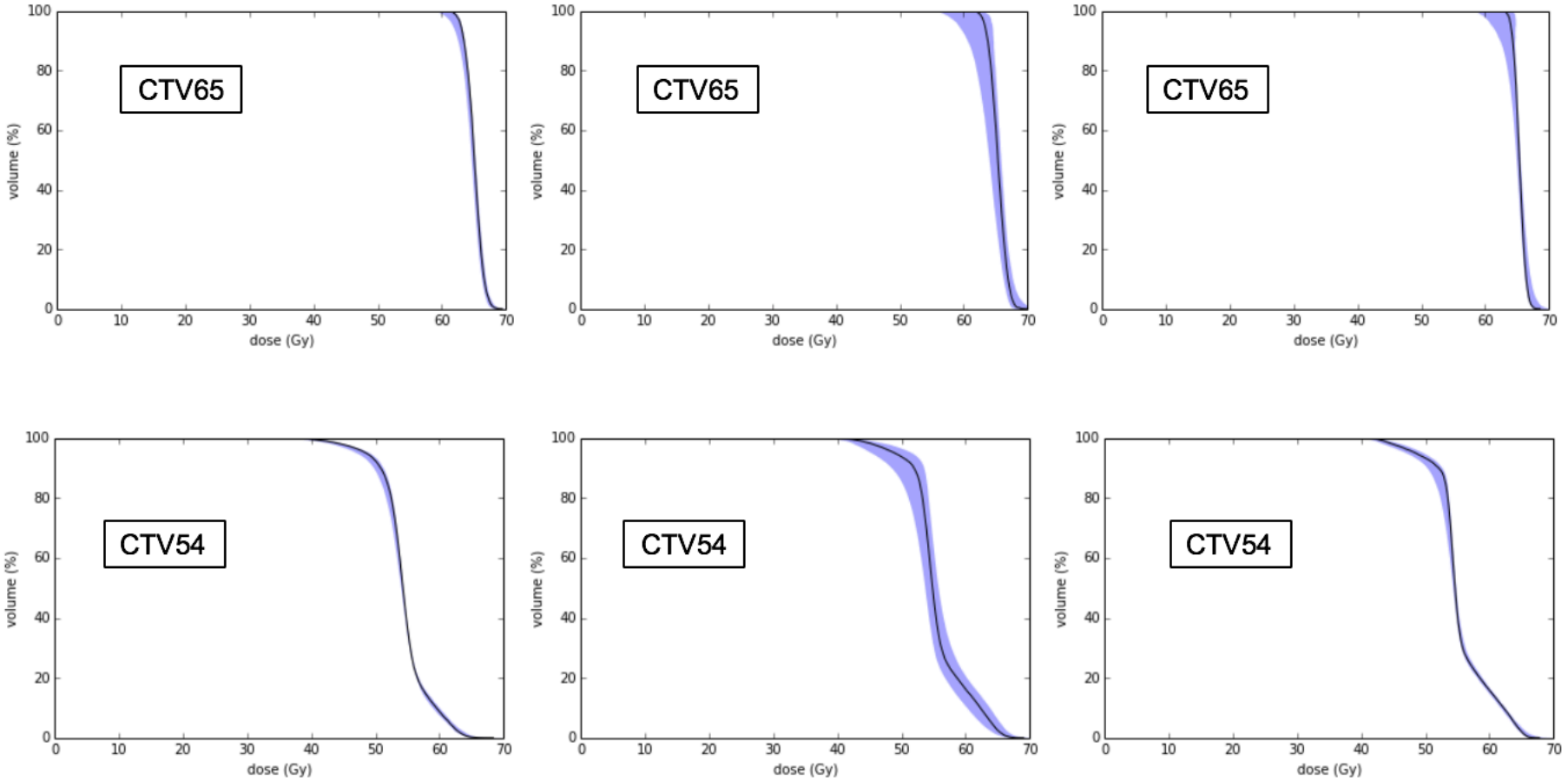


Figure 4.1Colour wash represents the DVH bands for dose distributions from the 11 perturbed plans for the CTVs in the RO plan (left column), 3-beam PTV-based plan (middle), and 7-beam PTV-based plan (right) for one LA-OPC case

The solid lines are the DVHs for the nominal dose distribution. The narrow CTV bands for the RO plan, relative to the other 2 plans, indicates its superior robustness. Likewise, the addition of more beams confers more robustness for PTV-based Do-IMPT plans.

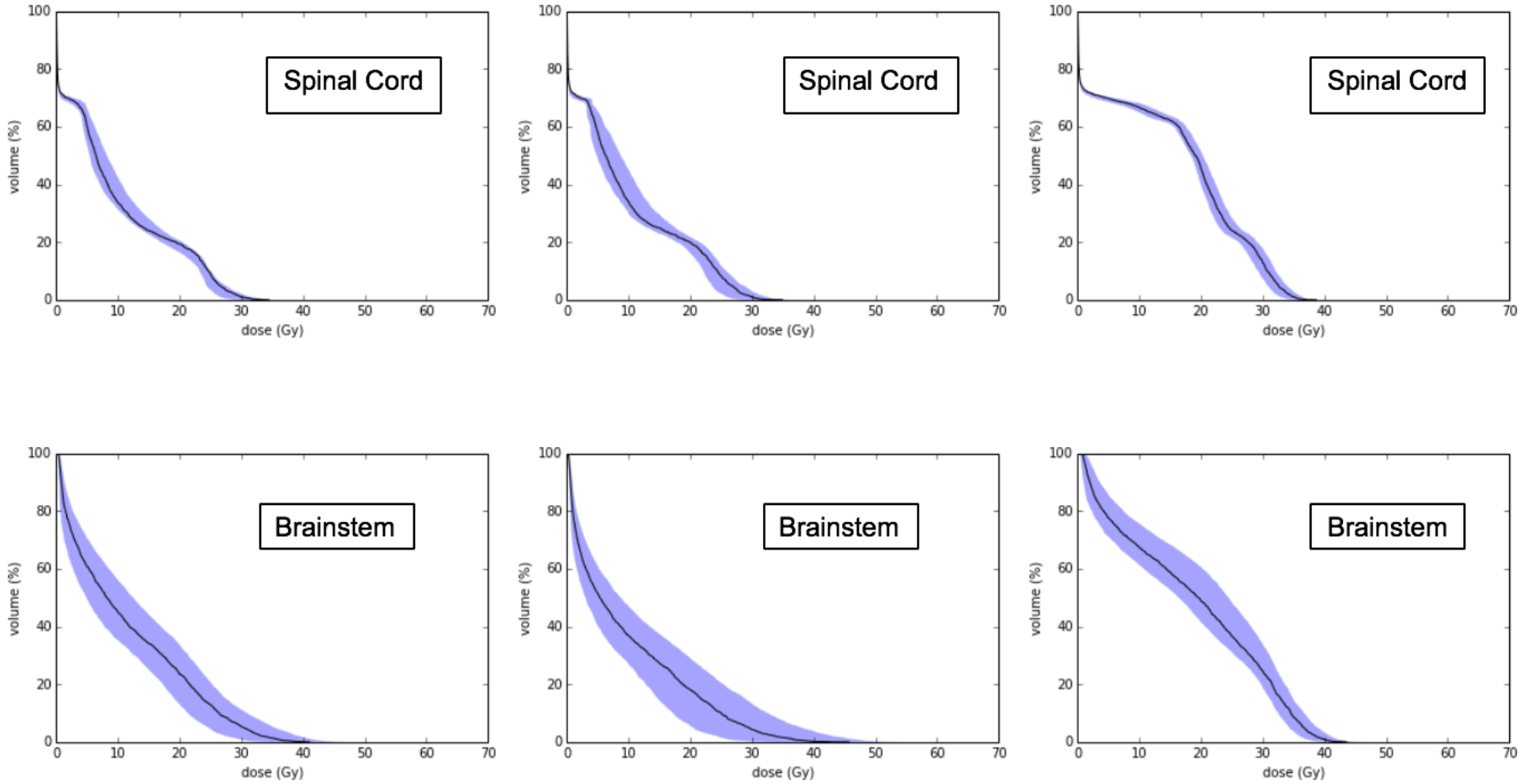


Figure 4.2 Colour wash represents the DVH bands for dose distributions from the 11 perturbed plans for the spinal cord (first row) and brainstem (second row) in the RO plan (left column), 3-beam PTV-based plan (middle), and 7-beam PTV-based plan (right) for one LA-OPC case.

The solid lines are the DVHs for the nominal dose distribution.

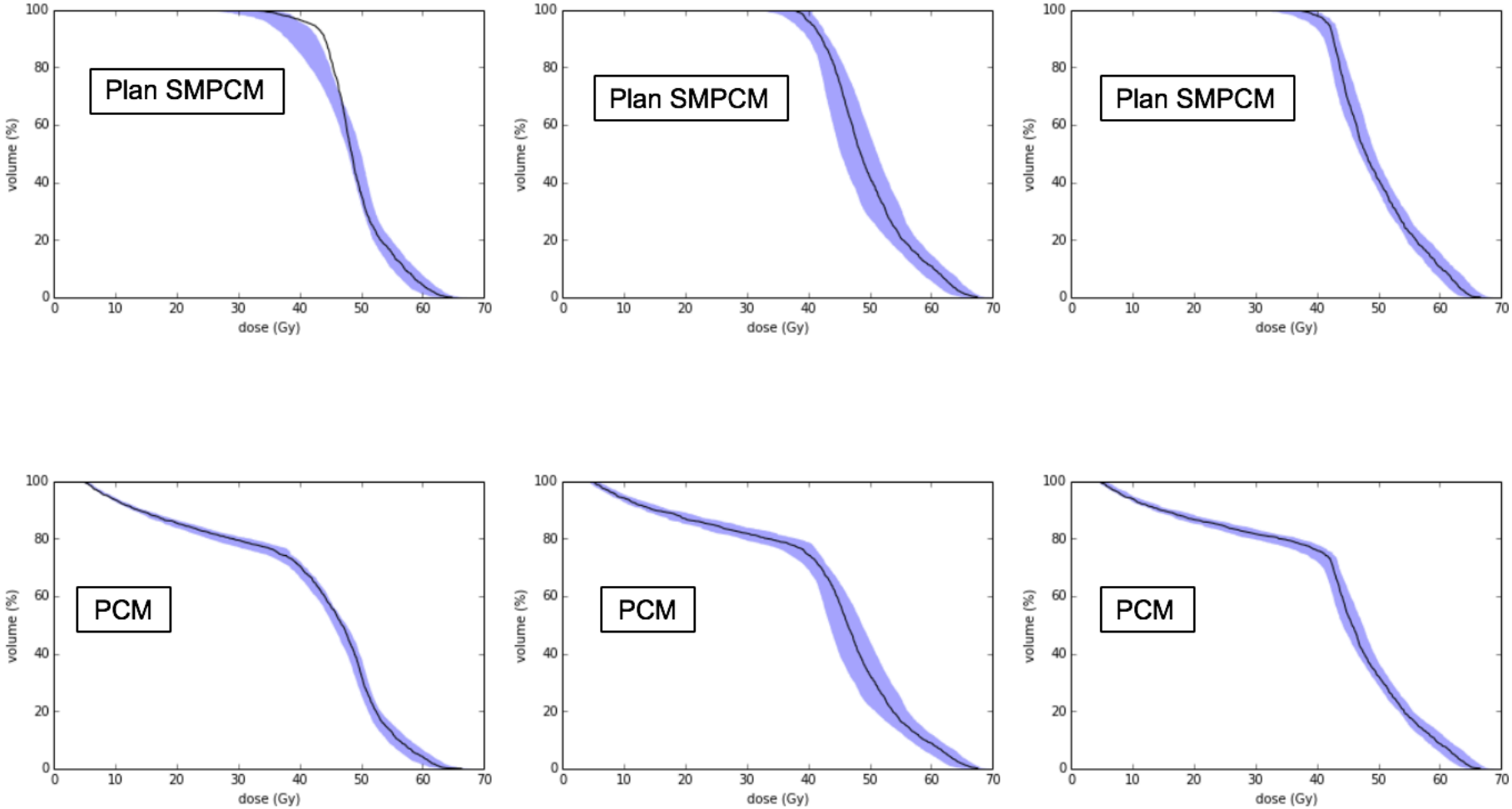


Figure 4.3 Colour wash represents the DVH bands for dose distributions from the 11 perturbed plans for Plan SMPCM (first row) and PCM (second row) in the RO plan (left column), 3-beam PTV-based plan (middle), and 7-beam PTV-based plan (right) for one LA-OPC case.

The solid lines are the DVHs for the nominal dose distribution.

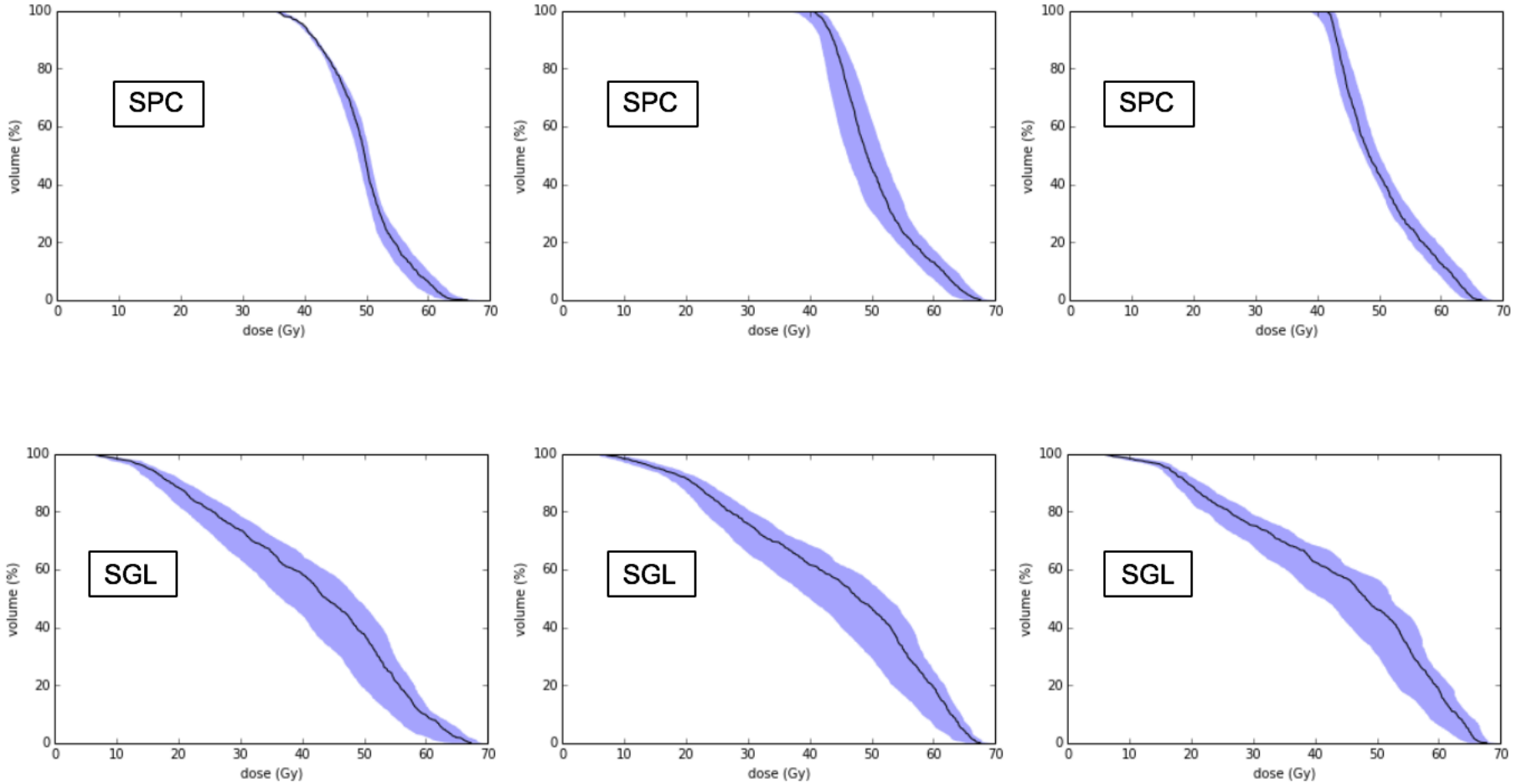


Figure 4.4 Colour wash represents the DVH bands for dose distributions from the 11 perturbed plans for the SPC (first row) and SGL (second row) in the RO plan (left column), 3-beam PTV-based plan (middle), and 7-beam PTV-based plan (right) for one LA-OPC case.

The solid lines are the DVHs for the nominal dose distribution.

4.6.2 Quantification of differences in perturbed dose in presence of treatment uncertainties

4.6.2.1 Target volumes

The mean DVH bandwidth for the mandatory objective functions of both CTVs were significantly narrower with the RO plan compared to PTV-based plans, as shown in table 4.2. In conjunction with the DVH bands, this indicated that the nominal RO target volume coverage was affected less in the presence of SE and RU with Do-IMPT$_{RO}$. Between the 2 margin-based plans, the addition of more beams improved robustness for CTV65$_{D99\%}$ and CTV65$_{D95\%}$ respectively in the presence of uncertainties.

Figure 4.5 shows the influence of the 11 treatment uncertainties on CTV65$_{D95\%}$ coverage for each patient. The RO plan was the most robust amongst the 3 proton plans: for all patients across each simulation, the CTV65 $D_{95\%}$ was > 95 % of the prescribed dose. In contrast, only 85 % (range 90 % - 97 %) and 92 % (range 93 % - 97 %) of all simulations met this criterion for Do-IMPT$_{3B}$ and Do-IMPT$_{7B}$ plans respectively. The minimum doses were observed for the combined errors of 3.5 % RU and 3 mm displacement in the X direction across both PTV-based plans for all but 1 patient (y direction, 7-beam plan). The mandatory CTV54 dose objective was achieved across all simulations for the 3 proton plans.

Volume	Plan	Bandwidth Mean (SD)	Range
CTV65			
D99%	Do-IMPT$_{RO}$	1.2 (0.3) *	0.7 – 1.6
	Do-IMPT$_{3B}$	4.5 (1.9)	1.6 – 6.6
	Do-IMPT$_{7B}$	3.5 (1.4) †	1.5 – 5.5
D95%	Do-IMPT$_{RO}$	0.6 (0.2) *	0.4 – 1.0
	Do-IMPT$_{3B}$	3.1 (1.5)	0.8 – 5.0
	Do-IMPT$_{7B}$	2.3 (1.1) †	1.0 – 4.1
D50	Do-IMPT$_{RO}$	0.3 (0.1) *	0.2 – 0.5
	Do-IMPT$_{3B}$	0.9 (0.3)	0.6 – 1.7
	Do-IMPT$_{7B}$	1.0 (0.3)	0.5 – 1.5
D5	Do-IMPT$_{RO}$	0.3 (0.1) *	0.1 – 0.5
	Do-IMPT$_{3B}$	1.2 (0.5)	0.5 – 2.3
	Do-IMPT$_{7B}$	1.1 (0.2)	0.9 – 1.4
D2	Do-IMPT$_{RO}$	0.4 (0.1) *	0.2 – 0.6
	Do-IMPT$_{3B}$	1.6 (0.6)	0.7 – 2.8
	Do-IMPT$_{7B}$	1.2 (0.2)	0.9 – 1.5
CTV54			
D50	Do-IMPT$_{RO}$	0.3 (0.1) *	0.2 – 0.4
	Do-IMPT$_{3B}$	0.8 (0.4)	0.4 – 1.5
	Do-IMPT$_{7B}$	0.5 (0.2)	0.2 – 0.7

Table 4.2 Mean CTV bandwidth values for the 11 perturbed plans for each of the 3 nominal proton plans.

* Significantly better values ($p < 0.05$) for RO plan compared to PTV-based plans
† Significantly better values ($p < 0.05$) for Do-IMPT$_{7B}$ compared to Do-IMPT$_{3B}$

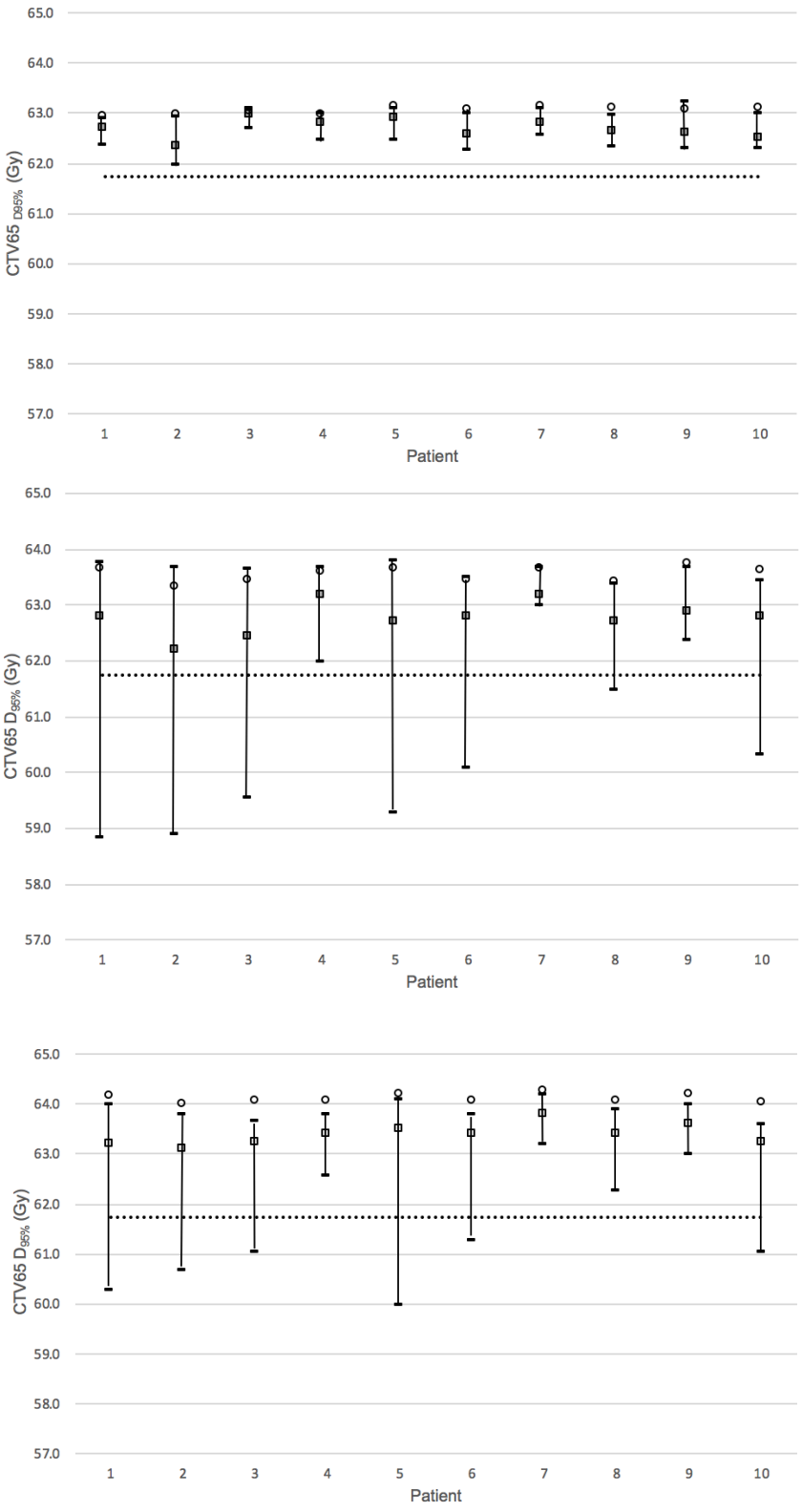


Figure 4.5 CTV65$_{D95\%}$ dose for each patient for the nominal plan (circles), with perturbed plans' median dose (squares) and maximum/minimum values (error bars). Top - Do-IMPT$_{RO}$, Middle - Do-IMPT$_{3B}$, Bottom - Do-IMPT$_{7B}$. The horizontal dotted line represents the 95 % dose (61.75 Gy) of the prescribed dose of 65.1 Gy.

The influence of RU alone on $CTV65_{95\%}$ with the three proton plans is shown in Figure 4.6. All three plans met the minimum dose constraint for each patient. The dose deterioration with the simulated RU relative to the nominal plan was least with the RO plan.

Similar findings were observed for the other CTV65 dose constraints, as shown in Figure 4.7.

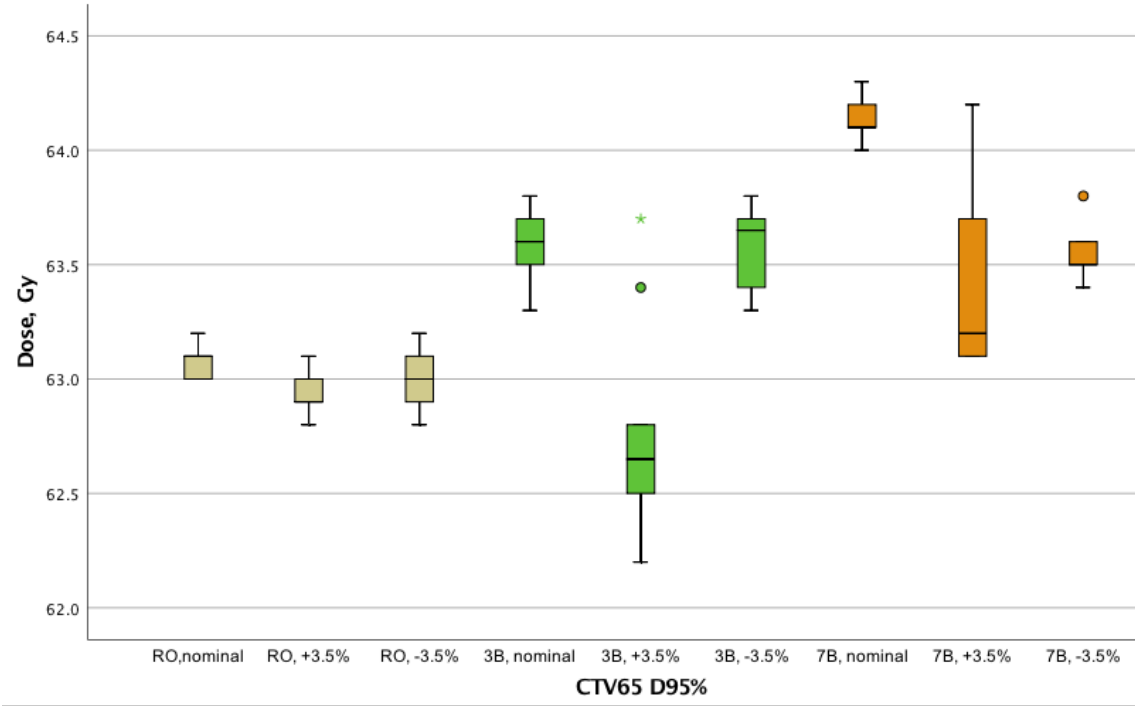


Figure 4.6 Boxplot demonstrating the impact of RU alone on $CTV65_{D95\%}$ coverage for the different proton plans.

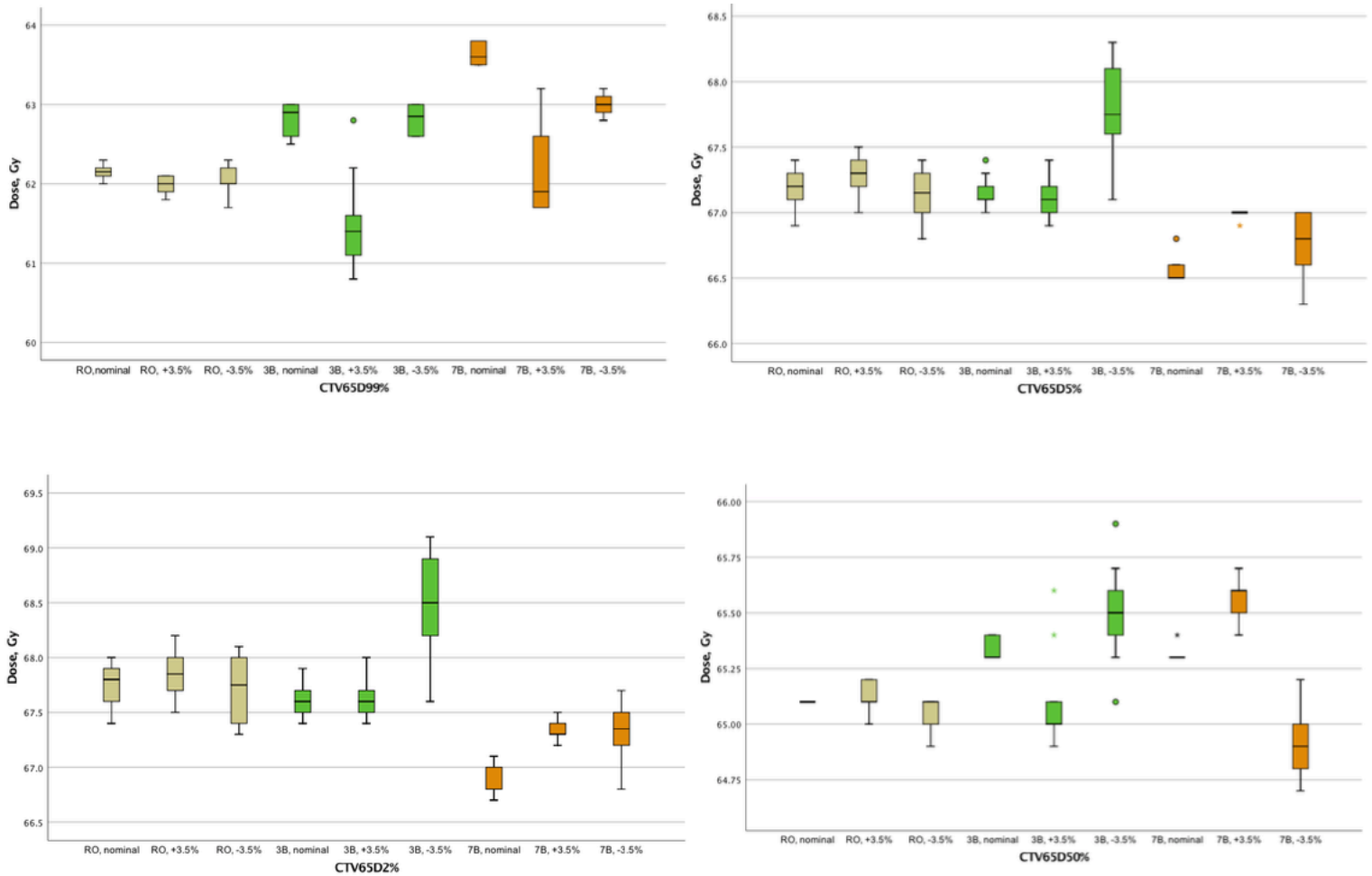


Figure 4.7 Boxplot demonstrating the impact of RU alone on CTV65 coverage for the different proton plans.

4.6.2.2 SW-OARs

The mean bandwidths for PlanSMPCM and PlanIPCM were statistically significantly worse with the PTV-optimised PBT compared to the RO plan (table 4.3). Although SPC mean bandwidth was smaller with Do-IMPT$_{RO}$, this reached statistical significance in comparison to Do-IMPT$_{3B}$ only. The delivered dose to the PCM and SGL were more sensitive to uncertainties with the PTV-based proton plans, evidenced by the significantly larger bandwidths compared to the RO plan. Differences for most endpoints were inferior for Do-IMPT$_{3B}$ compared to Do-IMPT$_{7B}$.

Structure	Plan	Bandwidth Mean (SD)	Range
Plan SMPCM			
Mean	Do-IMPT$_{RO}$	1.8 (0.8) *	1.0 – 3.7
	Do-IMPT$_{3B}$	4.3 (0.9)	2.4 – 5.8
	Do-IMPT$_{7B}$	3.1 (1.1) †	1.6 – 4.6
Plan IPCM			
Mean	Do-IMPT$_{RO}$	3.0 (1.3) *	0.9 – 5.2
	Do-IMPT$_{3B}$	5.3 (2.5)	1.9 – 10.2
	Do-IMPT$_{7B}$	4.8 (2.5)	2.1 – 9.8
SPC			
Mean	Do-IMPT$_{RO}$	1.3 (0.3) ¥	0.9 – 1.7
	Do-IMPT$_{3B}$	2.9 (1.1)	1.1 – 4.6
	Do-IMPT$_{7B}$	1.7 (0.6) †	0.5 – 2.3
PCM			
Mean	Do-IMPT$_{RO}$	1.6 (0.4) *	1.0 – 2.3
	Do-IMPT$_{3B}$	3.3 (1.1)	1.6 – 4.8
	Do-IMPT$_{7B}$	2.5 (1.0)	0.8 – 4.3
SGL			
Mean	Do-IMPT$_{RO}$	4.2 (1.3) §	2.4 – 6.5
	Do-IMPT$_{3B}$	5.0 (1.4)	2.3 – 7.8
	Do-IMPT$_{7B}$	5.3 (1.2)	3.5 – 7.3

Table 4.3 Mean SW-OAR bandwidth values for the 11 perturbed plans for each of the 3 nominal proton plans.

* Significantly better values ($p < 0.05$) for RO plan compared to PTV-based plans
† Significantly better values ($p < 0.05$) for Do-IMPT$_{7B}$ compared to Do-IMPT$_{3B}$
¥ Significantly better values ($p < 0.05$) for RO plan compared to Do-IMPT$_{3B}$
§ Significantly better values ($p < 0.05$) for RO plan compared to Do-IMPT$_{7B}$

The mandatory PlanSMPCM mean dose objective of < 50 Gy was achieved for all patients with all perturbations for the robustly optimised PBT. In comparison, only 65.5 % and 69.1 % of all simulations for Do-IMPT$_{3B}$ and Do-IMPT$_{7B}$ respectively met this planning objective (Figure 4.8). On the other hand, the variation in the mean dose to PlanIPCM under different simulations for each patient was comparable for all three planning techniques.

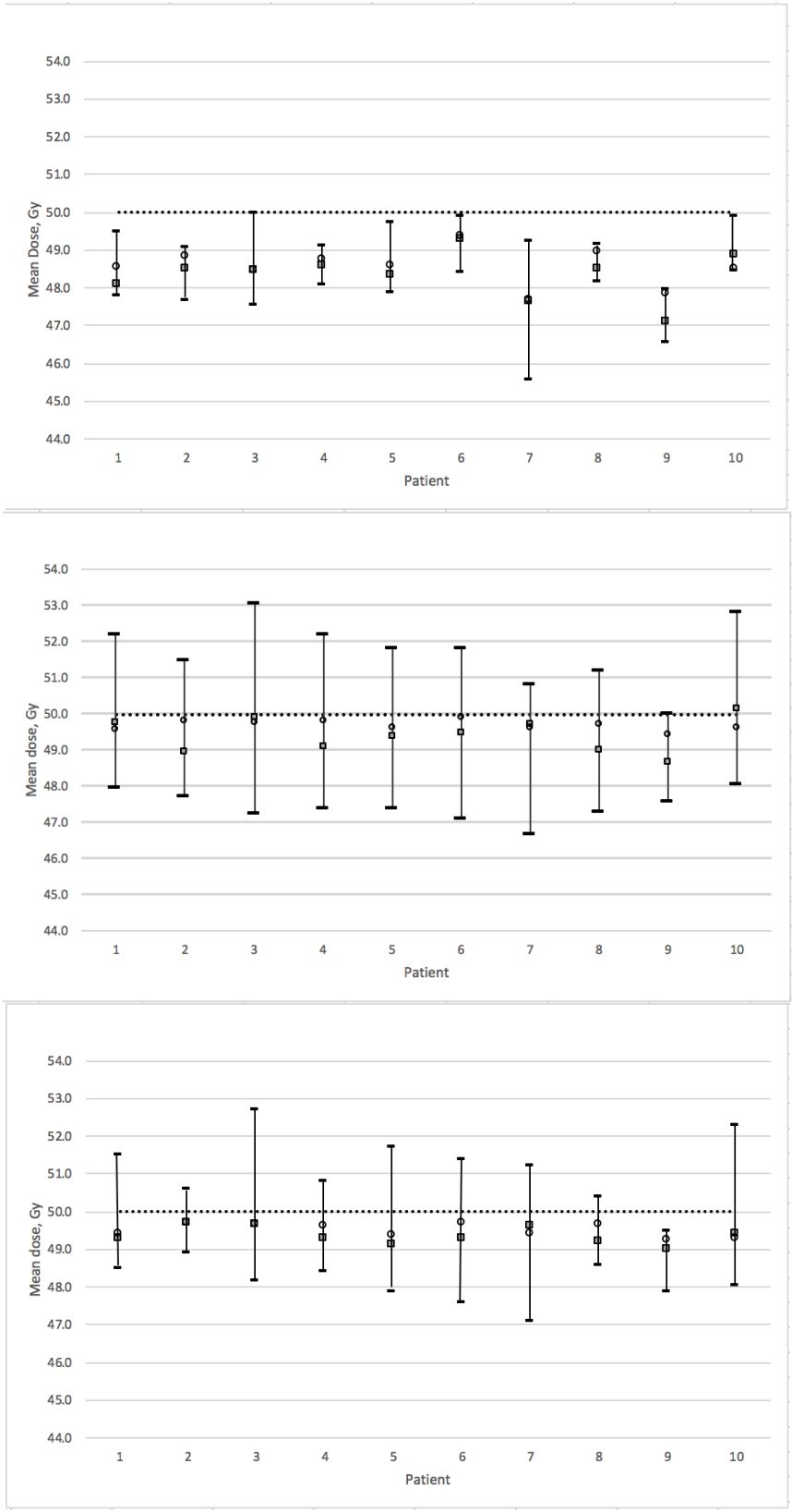


Figure 4.8 Mean PlanSMPCM dose for each patient for the nominal plan (circles), with perturbed plans' median dose (squares) and maximum/minimum values (error bars). Top - Do-IMPT$_{RO}$, Middle - Do-IMPT$_{3B}$, Bottom - Do-IMPT$_{7B}$. The horizontal dotted line represents the mandatory planning objective of < 50 Gy for this structure.

4.6.2.3 NSW-OARs

Bandwidths were smallest with the 7-beam PTV plan (Table 4.4). All perturbed plans met the dose objective for the critical OARs.

Structure	Plan	Bandwidth Mean (SD)	Range
Spinal cord			
Dmax	Do-IMPT$_{RO}$	5.4 (1.7)	2.2 – 7.9
	Do-IMPT$_{3B}$	5.0 (2.6)	2.5 – 10.0
	Do-IMPT$_{7B}$	2.1 (1.2) †	0.7 – 4.3
D1cc	Do-IMPT$_{RO}$	6.7 (1.8)	4.9 – 10.2
	Do-IMPT$_{3B}$	8.6 (3.5)	4.3 – 14.7
	Do-IMPT$_{7B}$	5.2 (1.9) †	2.9 – 8.3
Brainstem			
Dmax	Do-IMPT$_{RO}$	9.2 (2.8)	4.3 – 14.0
	Do-IMPT$_{3B}$	9.1 (3.5)	5.2 – 15.5
	Do-IMPT$_{7B}$	4.8 (1.6) †	3.1 – 8.3
D1cc	Do-IMPT$_{RO}$	8.2 (2.2)	4.8 – 12.4
	Do-IMPT$_{3B}$	10.3 (4.2)	4.4 – 17.9
	Do-IMPT$_{7B}$	6.3 (1.5) §	4.3 – 8.8
CL parotid			
Mean	Do-IMPT$_{RO}$	4.8 (0.6) *	4.0 – 6.0
	Do-IMPT$_{3B}$	5.9 (0.8)	4.9 – 7.0
	Do-IMPT$_{7B}$	6.1 (0.7)	4.5 – 7.2
IL parotid			
Mean	Do-IMPT$_{RO}$	6.0 (1.3) ¥	4.2 – 7.8
	Do-IMPT$_{3B}$	7.4 (1.5)	4.9 – 9.3
	Do-IMPT$_{7B}$	6.6 (1.1)	5.3 – 8.0

Table 4.4 Mean SW-OAR bandwidth values for the 11 perturbed plans for each of the 3 nominal proton plans.

* Significantly better values ($p < 0.05$) for RO plan compared to PTV-based plans
† Significantly better values ($p < 0.05$) for Do-IMPT$_{7B}$ compared to RO & Do-IMPT$_{3B}$
¥ Significantly better values ($p < 0.05$) for RO plan compared to Do-IMPT$_{3B}$

4.6.2.4 Perturbed NTCP

The absolute difference between the median perturbed NTCP and the nominal NTCP value for physician-scored RAD_{6M} with each dysphagia- optimised proton plan was < 2 %, and therefore modest as shown in Fig 4.9. There were no instances of deterioration in the predicted toxicity of > 5 % relative to the corresponding nominal proton plan.

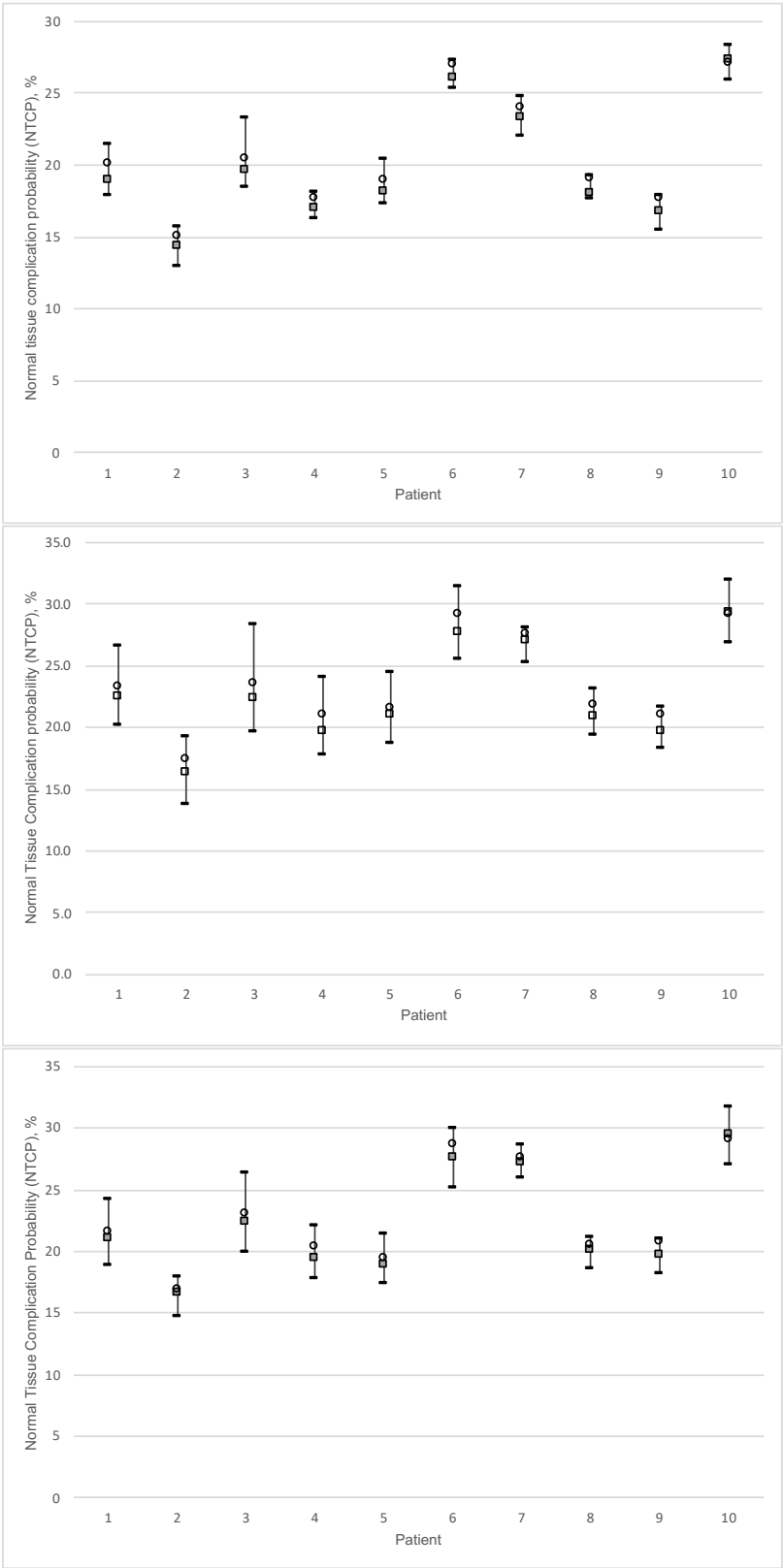


Figure 4.9 NTCP values for each patient for the nominal plan (circles), with perturbed plans' median dose (squares) and maximum/minimum values (error bars). Top - Do-IMPT$_{RO}$, Middle - Do-IMPT$_{3B}$, Bottom - Do-IMPT$_{7B}$.

4.7 Discussion

This chapter studied the influence of treatment uncertainties on the nominal dose distributions for CTV-based and PTV-based Do-IMPT plans in LA-OPC. The results showed that target volume coverage was significantly less sensitive to SE and RU with RO plans compared to PTV-based plans, evidenced by the narrower DVH bands and DVH bandwidths. Median simulated dose deviations from nominal Do-IMPT$_{RO}$ were small; in comparison this was larger with both Do-IMPT$_{3B}$ and Do-IMPT$_{7B}$ respectively. Crucially, the effects of the combined errors meant that the margin-based proton plans did not comply with the treatment intent that 90 % of the population receive CTV$_{D95\%}$ > 95 %, consequently risking target under-dosage (Figure 4.10). Likewise, we found that the RO plan was more resilient to RU and SE for SW-OARs compared to the PTV-based plans. These results demonstrate that the benefit of comparable target coverage and improved SW-OAR sparing by the nominal Do-IMPT$_{RO}$ relative to PTV-based Do-IMPT plans remains intact under considerations of treatment uncertainties. On the other hand, further planning optimisation to ensure adequate robustness of the target dose with margin-based plans might be necessary for some patients, particularly if the errors that were simulated in this study were reproduced at every fraction.

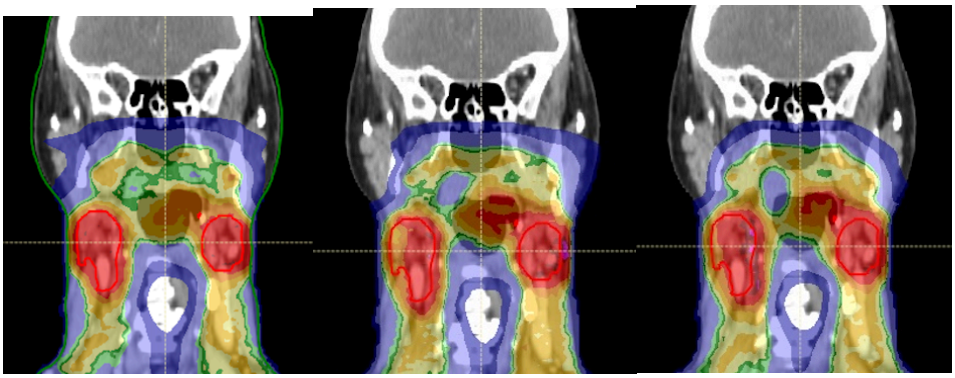


Figure 4.10 Dose distributions in the coronal plane for a representative patient illustrate the insensitivity of the robustly optimized plan (left panel) to range (+3.5%) and set up uncertainty (+ 3 mm) compared with the conventional PTV-based 3-beam (middle) and 7-beam plan (right panel).

The figure shows that the CTV (red) is inadequately covered with both PTV-based plans compared to the RO plan.

Our findings are consistent with previous studies that have indicated that PTV-based PBT for HNC produce treatment plans that are less robust in the presence of treatment uncertainties compared to RO-based PBT[5, 9-12].

Liu et al analysed the effectiveness of worst-case RO and the PTV approach to account for RU and SE in IMPT planning in 14 HNC patients. Three fields were used to generate the plans. Plan robustness was quantified by determining the area under the curve (AUC) of the root-mean-square DVH, in which the dose spread per voxel was recorded by the root-mean-square for 21 scenarios of treatment uncertainties. Smaller AUC values indicated that the plan was less sensitive to errors. Statistically significant differences in AUC values in favour

of the RO plan were found for all endpoints (high-dose CTV, 10.8 v 15.4; low-dose CTV, 11.4 v 16.7; BS, 16.3 v 16.5), apart from SC.

Another study by the same group investigated the dose distributions between robustly optimised and PTV- based IMPT in 5 patients with HNC and additionally investigated the sensitivity of the nominal plans to 21 treatment uncertainties. Plan robustness was studied by quantifying the CTV DVH bandwidth at the median dose, along with determining the target dose coverage and homogeneity in the worst-case scenario. The study showed that target volume coverage and homogeneity, along with OAR sparing, was superior with robust optimisation. Furthermore, the nominal RO plan was less sensitive to treatment uncertainties. Their findings showed that the CTV bandwidth was 0.59 with RO plan, and substantially narrower than the 3.53 of the PTV-based plan.

A finding from our study was that increasing the number of fields for PTV-based optimisation did not contribute to increased resilience to uncertainties compared to RO plan. We also demonstrated that certain dose volume parameters, such as $CTV65_{D95\%,}$ and maximum doses to the SC and BC, were more robust to SE and RU with Do-$IMPT_{7B}$ relative to Do-$IMPT_{3B}$. However, the dose to $CTV65_{D95\%}$ was found to be unsatisfactory with both margin-based plans in the presence of treatment uncertainties, thereby triggering re-optimisation for both to achieve a more robust plan. Likewise, the dose constraints for the critical OARs were achieved for all patients in the presence

of the various uncertainty scenarios with the 3-field PTV plan. Therefore, increasing the number of fields for PTV-based Do-IMPT does not translate into clinically relevant improved robustness. The above findings for the Do-IMPT plans are consistent with previous studies that evaluated PBT robustness in HNC. In the study by Van Dijk et al, SW-IMRT (7 fields), RO – based and PTV – based SW-IMPT (2,3,5 and 7 fields) were generated for 10 patients with HNC. RU was simulated by altering the CT intensity values by ± 3 %, while SE was analysed by shifting the isocentre isotropically in 26 different directions. All perturbed IMRT and RO-IMPT plans showed satisfactory coverage of both CTVs. In contrast, only 1 of 40 simulated plans demonstrated adequate coverage of both target volumes. In addition, increasing the number of fields did not make any meaningful difference on plan resilience for RO plans, while appeared to have a detrimental impact on the low-dose CTV for the PTV approach.

Another key finding of this study was that RU alone did not result in any significant deterioration in target volume coverage between the different proton plans. This is an important observation, as the adoption of the conventional CTV-PTV margin for PBT is traditionally considered insufficient to account for RU.

4.8 Conclusions

Do-IMPT$_{RO}$ was more robust compared to the PTV-based Do-IMPT in LA-OPC, particularly for target volumes. Increasing the number of fields did not result in improved plan robustness for the PTV-based approach.

4.9 References list

1. Mohan, R. and D. Grosshans, *Proton therapy - Present and future.* Adv Drug Deliv Rev, 2017. **109**: p. 26-44.

2. Engelsman, M., M. Schwarz, and L. Dong, *Physics Controversies in Proton Therapy.* Seminars in Radiation Oncology, 2013. **23**(2): p. 88-96.

3. Kraan, A.C., et al., *Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range, and setup errors.* Int J Radiat Oncol Biol Phys, 2013. **87**(5): p. 888-96.

4. Unkelbach, J., et al., *Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning.* Med Phys, 2009. **36**(1): p. 149-63.

5. Liu, W., et al., *Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers.* Med Phys, 2013. **40**(5): p. 051711.

6. Liu, W., et al., *Robust optimization of intensity modulated proton therapy.* Med Phys, 2012. **39**(2): p. 1079-91.

7. Chen, W., et al., *Including robustness in multi-criteria optimization for intensity-modulated proton therapy.* Phys Med Biol, 2012. **57**(3): p. 591-608.

8. van Herk, M., et al., *The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy.* Int J Radiat Oncol Biol Phys, 2000. **47**(4): p. 1121-35.

9. Bai, X., et al., *Robust optimization to reduce the impact of biological effect variation from physical uncertainties in intensity-modulated proton therapy.* Phys Med Biol, 2019. **64**(2): p. 025004.

10. Liu, W., et al., *PTV-based IMPT optimization incorporating planning risk volumes vs robust optimization.* Med Phys, 2013. **40**(2): p. 021709.

11. Cubillos-Mesias, M., et al., *Impact of robust treatment planning on single- and multi-field optimized plans for proton beam therapy of*

unilateral head and neck target volumes. Radiat Oncol, 2017. **12**(1): p. 190.

12. van Dijk, L.V., et al., *Robust Intensity Modulated Proton Therapy (IMPT) Increases Estimated Clinical Benefit in Head and Neck Cancer Patients.* PLoS One, 2016. **11**(3): p. e0152477.

5 Chapter 5: Inter-observer variation in pharyngeal constrictor muscle delineation and subsequent impact on predicted radiation-associated dysphagia

5.1 Introduction

IMRT creates steep dose gradients between PTVs and non-target normal tissue in order to deliver highly conformal RT, and consequently accuracy in defining RT volumes and planning is required to fully exploit its benefits in HNC[1, 2]. Variability in the delineation of either target volume or OAR could result in under-dosage of tumour, over-dosage of critical normal structures or both. In this context, inconsistent contouring of the PCM could offset the potential toxicity-mitigating benefits of Do-IMRT presented in chapter 2. Likewise, heterogeneity in the definition of this SW-OAR amongst clinicians recruiting to the DARS study might also lead to erroneous interpretation of RT-related morbidity, and consequently affect the primary endpoint of this randomised study. In addition, variable contouring of PCM may lead to inaccurate correlation between PCM DVH and RAD, and subsequent parameters generated for predicting swallowing toxicity may be misleading.

As part of the pre-trial RTTQA programme for DARS trial, participating centres were expected to successfully complete a pre-trial contouring case before

enrolling patients in the study. An important component of this quality assurance was the outlining of the PCM, as delineation of this structure in UK is not routine clinical practice. In this chapter, I investigate the IOV in delineation of this structure within the context of the pre-trial RTTQA for DARS trial, and study its impact on Do-IMRT plans.

5.2 Aims

The first aim of this study was to quantify the differences in PCM delineation between UK head and neck oncologists within the context of a pre-trial RTTQA programme. The second aim was to evaluate the impact of IOV on delivered dose to the constrictor muscle for Do-IMRT. The third aim was to analyse the influence of outlining variability on complication probability for physician-scored RAD_{6M}.

5.3 Null Hypothesis

There will be no IOV in the delineation of the constrictors, and therefore, no difference in NTCP for physician-scored RAD_{6M} will be observed.

5.4 Materials and methods

5.4.1 DARS pre-trial RTTQA programme

The pre-trial RTTQA required principal investigators' (PI) from participating centres to outline target volumes and OARs, including SMPCM and IPCM, on a benchmark case of a patient with T2N2c tongue base tumour. Centres downloaded the planning CT scan dataset, with GTV pre-outlined, in digital imaging and communications in medicine (DICOM) – RT from the RTTQA website. The completed DICOM-RT dataset was submitted for central review at the Royal Marsden Hospital. Detailed feedback was provided to each PI, and if there was substantial contouring deviation from the trial protocol, a resubmission was requested. Pre-trial contouring approval was given only when compliant with the protocol. For the purpose of my study, I evaluated the constrictor muscle delineation from the initial submission of each PI (n=15).

5.4.2 Delineation process for PI - PCM

The trial RT protocol provided a comprehensive instruction manual on delineation guidelines for PCM. In addition, participating centres had access to a detailed slice-by-slice CT atlas of delineated PCM from the PATHOS trial. PIs were expected to follow the trial protocol and contour the SMPCM (PI-SMPCM) as one structure, and the IPCM (PI-IPCM) as a separate structure.

5.4.3 Gold standard (GS) delineation

To quantify the IOV in volume delineation, a reference volume is required. This can be defined by using either a GS contour, delineated by an expert or a panel of experts, or the use of a probabilistic algorithm such as Simultaneous Truth and Performance Level Estimation (STAPLE) to generate an 'ideal' contour from a number of contours defined by different experts[3, 4].

For the purpose of this study, a GS set of reference structures (GS-SMPCM, GS-IPCM) was created by the chief investigator (Professor Chris Nutting), and approved by the trial management group.

5.4.4 Evaluation of PI – PCM delineation

The DICOM-RT data from the 15 participating centres, together with the GS dataset were exported to the research version of RS for qualitative and quantitative analysis. The outlining accuracy of each PI – SMPCM and PI – IPCM were compared against GS contours, using different metrics described below.

5.4.4.1 Volumetric assessment

Volumes of each PI - SMPCM and PI - IPCM in cm^3 were generated and compared against GS – SMPCM and GS – IPCM contours respectively.

5.4.4.2 Overlap assessment

5.4.4.2.1 Whole- volume overlap assessment

Conformity Indices:

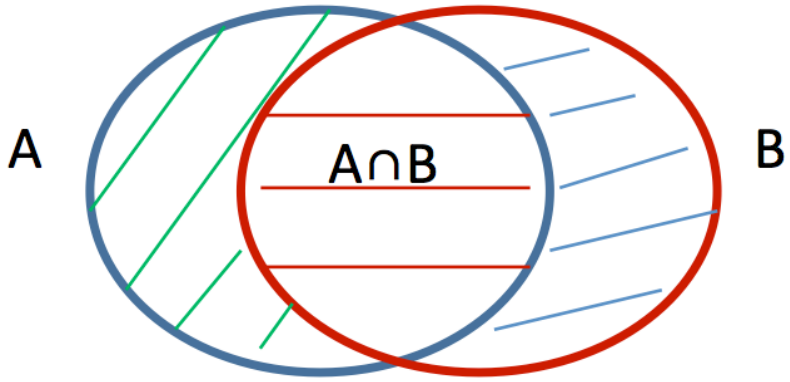
Volume-based conformity indices (CI) were recorded to determine the concordance between the investigators' and GS contours of SMPCM and IPCM respectively, where the CI represent a ratio of different volumes. In addition to the constrictor muscle analysis, CI for the BS and parotid glands were also determined. These structures are routinely delineated as OARs in HNC. Good conformity between clinicians can therefore be expected, and would serve as a useful comparator for the constrictor muscle.

Dualta McQuaid, one of the physicists in our team, created a python script in RS that I used to determine the above CI. The script was run for each PI contour to generate the results, which was subsequently exported to an excel spreadsheet for analysis. The CI used for analysing contouring variation in this study are summarised below (Figure 5.1):

- Dice similarity coefficient (DSC): reflects the overall agreement between the volumes of two contours. An ideal score is 1, indicating perfect overlap with the GS contour[5]. A score of > 0.7 is considered to represent good agreement between 2 contours[6-8].

- Geographical Miss Index (GMI): indicates the amount of GS contour not included in the PI contour. An ideal score is 0, implying no ‘under-contouring’[9].

- Discordance Index (DI): indicates the amount of PI outlining not included in the GS contour. An ideal score is 0, indicating no ‘over-contouring’[10].



Conformity Index	Equation	Ideal score
Dice Similarity Coefficient	$\frac{2\,(A \cap B)}{A + B}$	1
Geographical Miss Index	$\frac{B - (A \cap B)}{B}$	0
Discordance Index	$\frac{1 - (A \cap B)}{A}$	0

A = Principal investigator contour; B = Gold standard contour; A∩B = volume of intersection of A and B

Figure 5.1Conformity indices for volume overlap assessment

Surface-based distance to agreement:

Distance to agreement (DTA) is generated by computing the minimum distance from a point on the GS contour surface to a point on the investigator contour surface, which is then repeated for all points on the GS surface[11]. This parameter is dependent on the direction the measurements are taken and is usually accounted for by reporting the mean- and maximum DTA. An ideal score is 0.

5.4.4.2.2 Slice-by-slice overlap assessment

In order to determine the location of maximal variation between investigators' and GS contours, a slice-by-slice conformity analysis [6-9] was carried out as described below (Figure 5.2).

There were 26 slices of GS – SMPCM delineated on the planning CT dataset. This structure set was duplicated. On the duplicate set, all SMPCM slices below the cranial- most contour were deleted and a new structure set, GS – SMPCM_1 was created. This method was successively repeated on the duplicate set to generate 26 individual structure sets, representing each slice of SMPCM (ie GS – SMPCM_1, GS – SMPCM_2,……, GS – SMPCM_26). The same methodology was applied to generate 6 individual GS – IPCM structure sets (GS – IPCM_1,…….., GS – IPCM_6).

Corresponding 26 and 6 structure sets were created for each PI – SMPCM and PI – IPCM respectively. The slice DSC (s-DSC), slice GMI (s-GMI), slice DI (s-DI), slice mean DTA (s-mean DTA), and slice maximum DTA (s-maximum DTA) were subsequently determined to identify volume variation on a slice-by-slice basis, using the equation described in Fig 5.1. Positional variation on each slice was additionally established by evaluating the maximum distance from the surface of GS delineation to the PI contour in four directions (anterior, posterior, right lateral, and left lateral) on each slice.

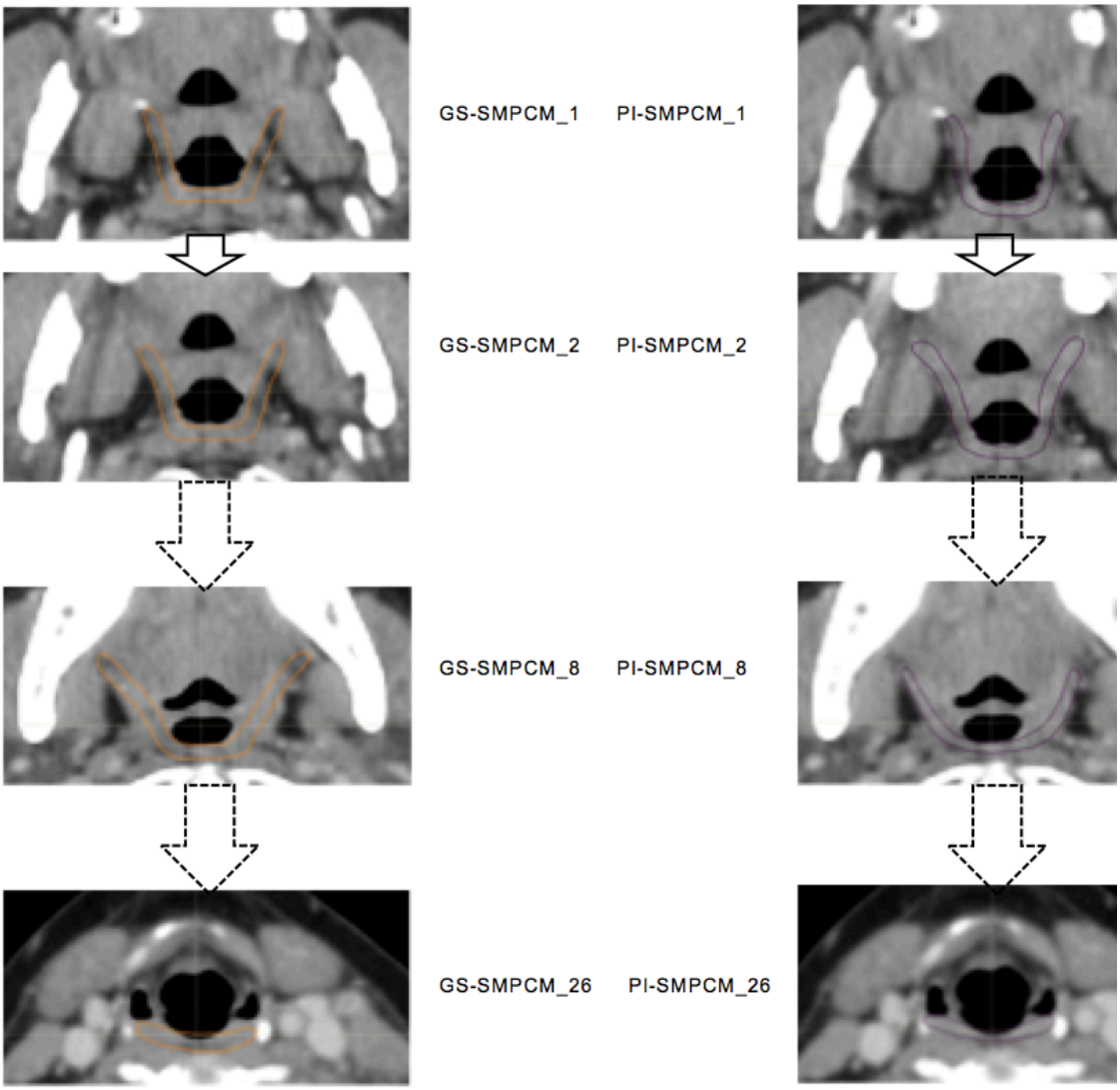


Figure 5.2 Gold standard (GS), and each principal investigator's (PI) contouring of superior and middle pharyngeal constrictor muscle (SMPCM) was split into individual slices. Conformity indices were subsequently calculated on a slice-by-slice basis. The same process was repeated for the inferior constrictor delineation.

5.5 Comparison of dose delivered to the swallowing OARs

This was analysed as follows (Figure 5.3):

1. The mean RT dose delivered to the SW-OARs using the GS contours was recorded. This was achieved by creating a GS Do-IMRT plan using the GS target volumes and OARs including SMPCM and IPCM, as per the methodology discussed in chapter 1. Mean dose delivered to the superior PCM (GS - SPC), IPCM (GS - IPCM), PCM (GS - PCM), and SGL (GS - SGL) were generated.

2. Individual PI Do-IMRT (PI DO-IMRT) plans based on their delineation of SMPCM and IPCM were created, and corresponding dose to the swallowing structures (PI-SPC, PI-IPCM etc) were tabulated. For these plans, GS target volumes and NSW-OARs were used for RT optimisation. This ensured that any difference in mean dose delivered to constrictors was solely due to variation in its contouring by the PIs.

3. GS - SMPCM and GS - IPCM structure sets were superimposed on PI Do-IMRT plans constructed in step 2, and the mean dose delivered to the GS contours on PI Do-IMRT plans was derived. The reference dose (R-SPC, R-IPC etc) thus generated would reflect the true dose delivered to the constrictors, and would be useful to highlight any clinically relevant dosimetric differences if there were to be substantial PI contouring deviation from the benchmark delineation.

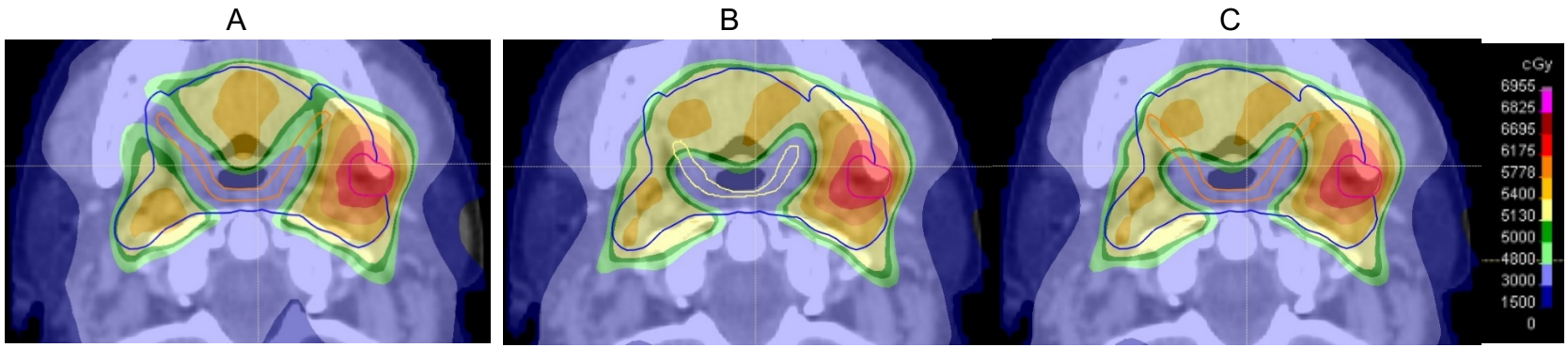


Figure 5.3 Example of evaluation of dose delivered to PCM based on PI contours on an axial CT slice.

A) The GS Do-IMRT plan based upon the GS SMPCM (orange) and GS IPCM contour was created to record the DVH for this SW-OAR; B) shows a PI Do-IMRT plan that was generated using the PI SMPCM (yellow) and PI IPCM delineation to derive the relevant dose metrics; C) GS SMPCM and GS IPCM contour was superimposed on the PI Do-IMRT plan in (B) to allow their DVHs to be derived. This was then compared to the original DVH obtained in (A). The presence of variation between the GS and PI contour, as in this slice, would highlight differences in dose delivered. In this example, it can be seen that there was less sparing of GS SMPCM laterally on PI Do-IMRT plan compared to GS plan.

5.6 Clinical impact of dosimetric variation

Three NTCPs for RAD_{6M} were created and compared for each PI as follows:

1. Gold standard NTCP (GS-NCTP), based on GS Do-IMRT plan

2. Investigator NTCP (PI-NTCP), based on PI Do-IMRT plans created in section step 2.

3. Reference NTCP (R-NTCP), based on GS contours that were superimposed on PI Do-IMRT plans created in section step 3.

5.7 Statistical analysis

SPSS v25 was used to conduct the statistical analysis. Data was assessed for their normality using Shapiro-Wilk before summary statistics and correlations were analysed. For variables that were normally distributed, the mean and 95% confidence interval (95% CI) are reported. For those variables that were not normally distributed, the median and interquartile range (IQR) are reported.

5.8 Results

An example of the heterogeneity in SMPCM definition between the PIs is illustrated in Fig 5.4.

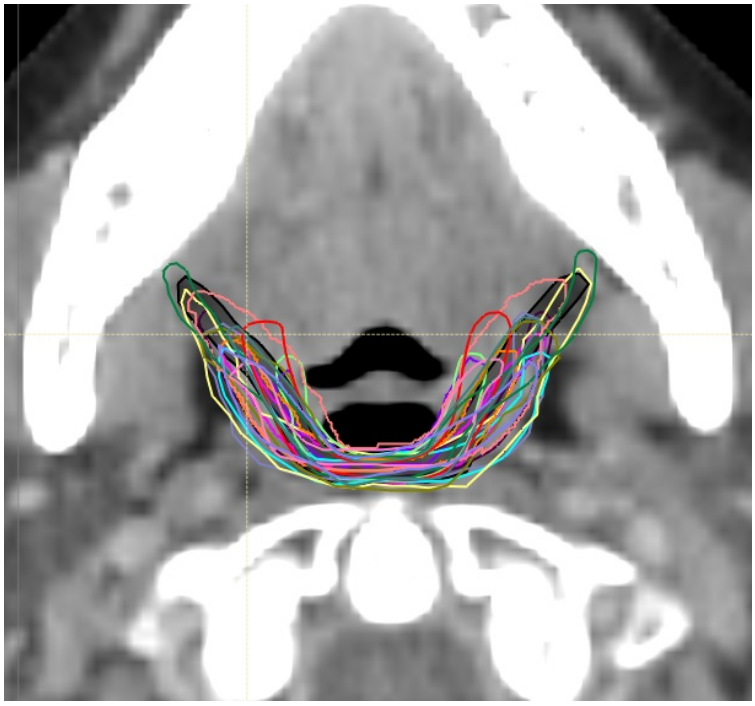


Figure 5.4 Example of PI-SMPCM contours on a single axial CT slice. The gold standard contour is shown in black

5.8.1 Volume assessment

5.8.1.1 Simple whole - volume assessments

GS-SMPCM and GS-IPCM volumes were 13.51 cm^3 and 1.67 cm^3 respectively. Mean PI-SMPCM and PI-IPCM volumes were 12.18 cm^3 (95 % CI 10.53 – 13.83 cm^3, range 8.53 – 18.19 cm^3) and 2.40 cm^3 (95% CI 1.88 – 2.93, range 1.34 – 4.43 cm^3) respectively (Figure 5.5). PI-SMPCM volumes were predominantly smaller than GS volume (11/15), while the converse was true for PI-IPCM volumes (11/15).

For GS contour, the volume of Plan SMPCM, which is that part of SMPCM lying outside CTV65 and is actively spared during the optimisation process for Do-IMRT, was 13.47 cc (99.71 % of SMPCM volume). For PI contours, mean PlanSMPCM volume was 11.74 cc (96.5 % of SMPCM volume, range 93.46 % - 99.85 %).

The mean PI contralateral parotid volume was 37.1 cc (95 % CI 35.4 – 38.9), and similar to the GS volume. Corresponding values for ipsilateral parotid and brainstem volumes were 32.5 cc (95 % CI 30.7 – 34.4; GS 35.2 cc) and 22.6 cc (95% CI 19.9 – 25.3; GS 25.7 cc) respectively.

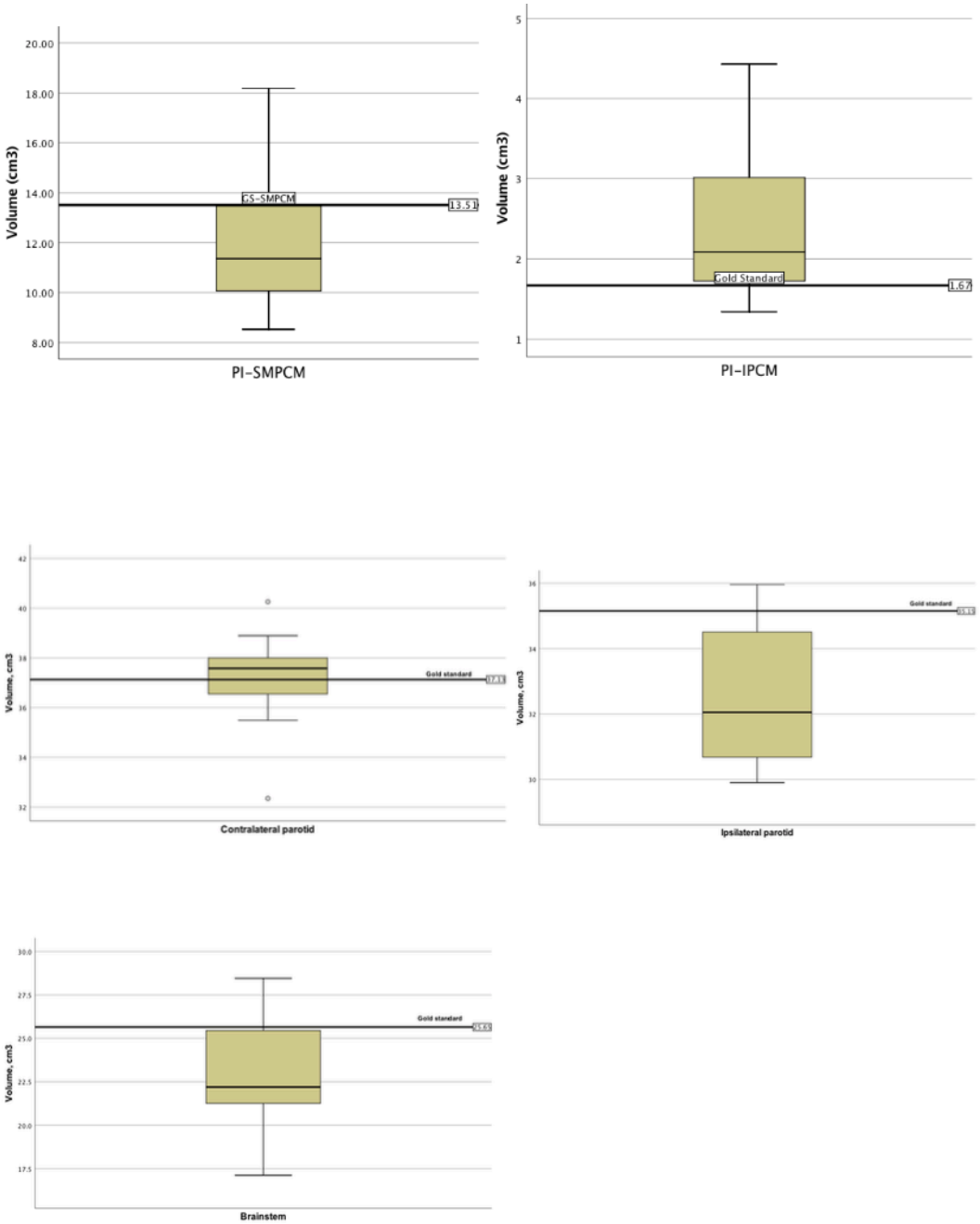


Figure 5.5 Simple whole-volume analysis of principal investigators' (PI) contouring of the constrictor muscles.

GS, gold standard; IPCM, inferior pharyngeal constrictor muscle; SMPCM, superior and middle pharyngeal constrictor muscle

5.8.1.2 Overlap assessments

There was substantial discordance in conformality between PI and GS contours for both SMPCM and IPCM contours, with low DSC and high GMI and DI values for both structure sets. 2 of the 15 PIs achieved a DSC > 0.70 for their IPCM delineation, and none for SMPCM contouring (Figure 5.6 and 5.7 respectively). The results of the whole-volume CI and DTA for the constrictor muscles and NSW-OARs are tabulated in Table 5.1. In comparison, there was good agreement for the NSW-OARs, with DSC of > 0.80 for both parotids and BS (Table 5.2). The GMI values indicated that a mean of 6.3 cm^3 (range 3.2 – 8.0 cm^3) and 0.5 cm^3 (range 0.2 – 0.9 cm^3) of the GS – SMPCM and – IPCM contours were outside the investigator outlinings respectively. In other words, on average 46.6 % and 30.0 % of GS – SMPCM and – IPCM volumes were not included in the PIs delineation. The DI values, particularly for IPCM, imply substantial over-contouring by the PIs. For 11 (73%) PI-SMPCM and 3 (20%) PI-IPCM contours, the maximum DTA was > 1 cm relative to the corresponding GS contours.

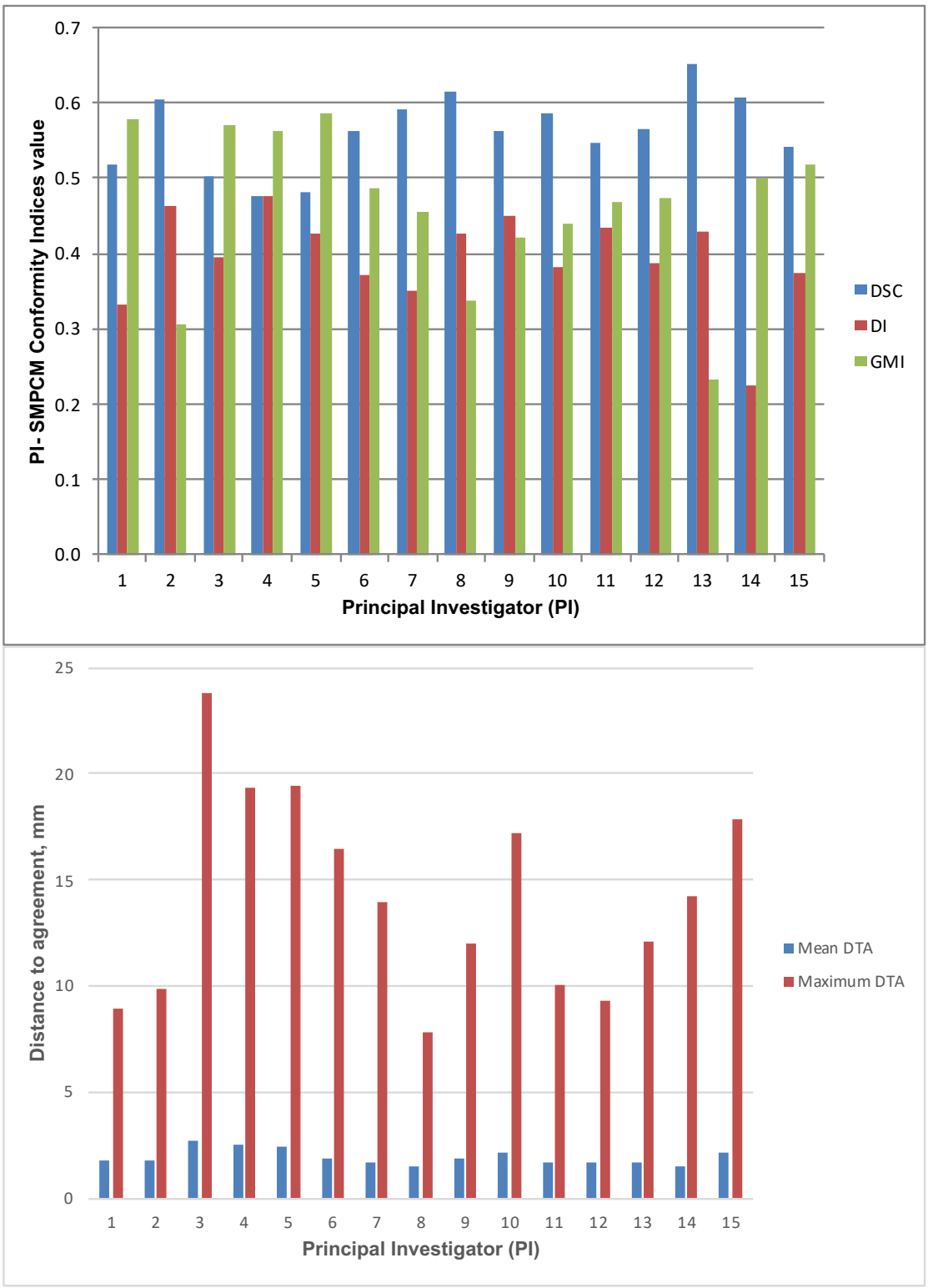


Figure 5.6 Conformity indices (top) and DTA (bottom) results for PI-SMPCM contours

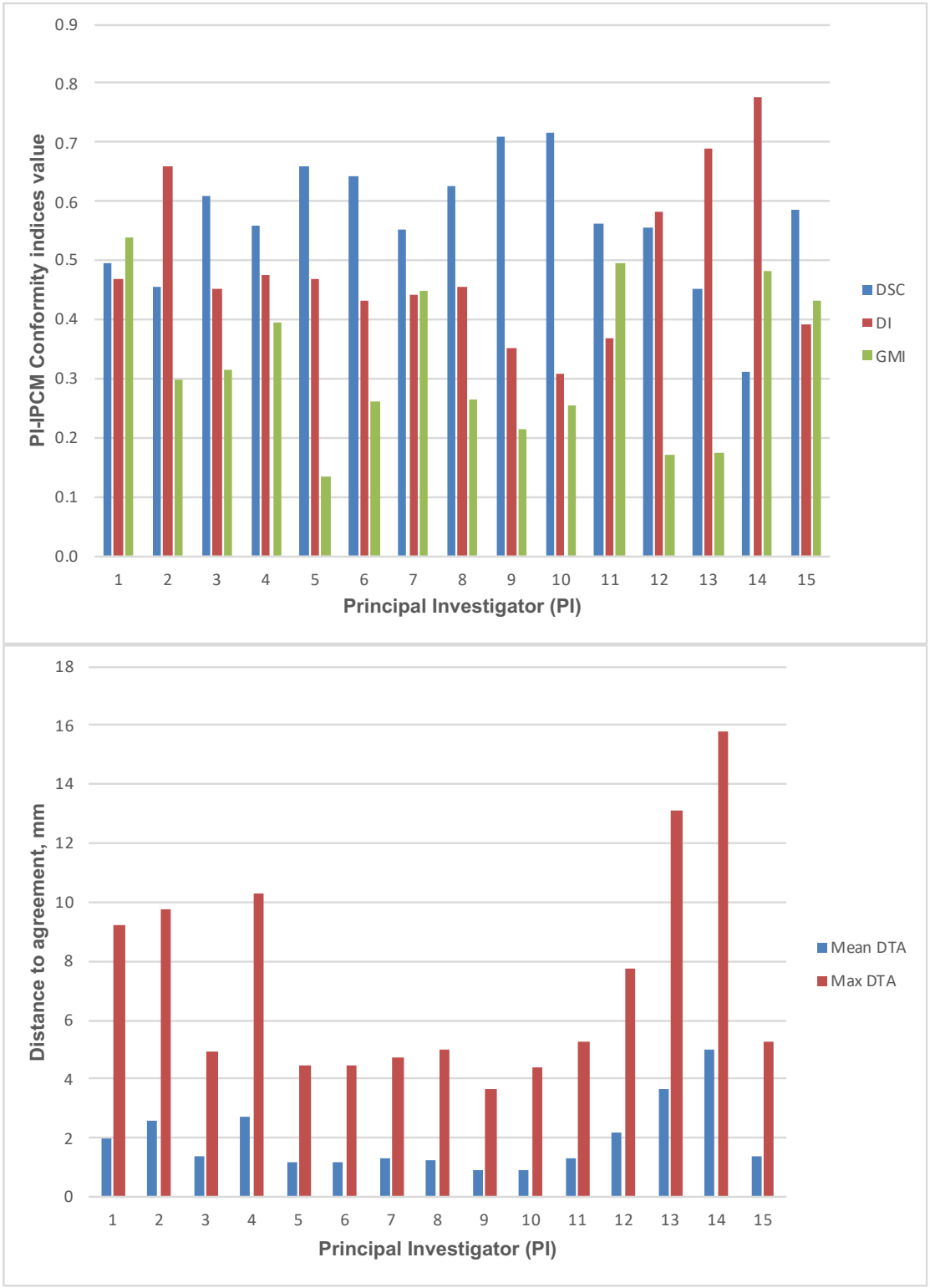


Figure 5.7 Conformity indices (top) and DTA (bottom) results for PI-IPCM contours

Structure	SMPCM					IPCM				
	DSC	DI	GMI	Mean DTA (mm)	Max DTA (mm)	DSC	DI	GMI	Mean DTA (mm)	Max DTA (mm)
Range	0.48-0.65	0.23-0.48	0.23-0.59	1.5-2.8	7.8-23.8	0.31-0.72	0.31-0.78	0.14-0.54	0.9-5.0	3.6-15.8
Median	-	-	-	1.8	13.9	-	-	-	1.4	5.3
Mean	0.56	0.40	0.46	2.0	14.2	0.57	0.49	0.33	-	-
SD	0.05	0.06	0.10	0.04	0.47	0.11	0.13	0.13	0.11	0.37
95 % CI	0.53-0.59	0.36-0.43	0.40-0.52	-	11.5-16.8	0.51-0.63	0.41-0.56	0.26-0.40	-	-
IQR	-	-	-	1.7-2.2	-	-	-	-	1.2-2.6	4.5-9.8
GS	1	0	0	0	0	1	0	0	0	0

Table 5.1 Results for conformity indices and distance to agreement for SMPCM and IPCM

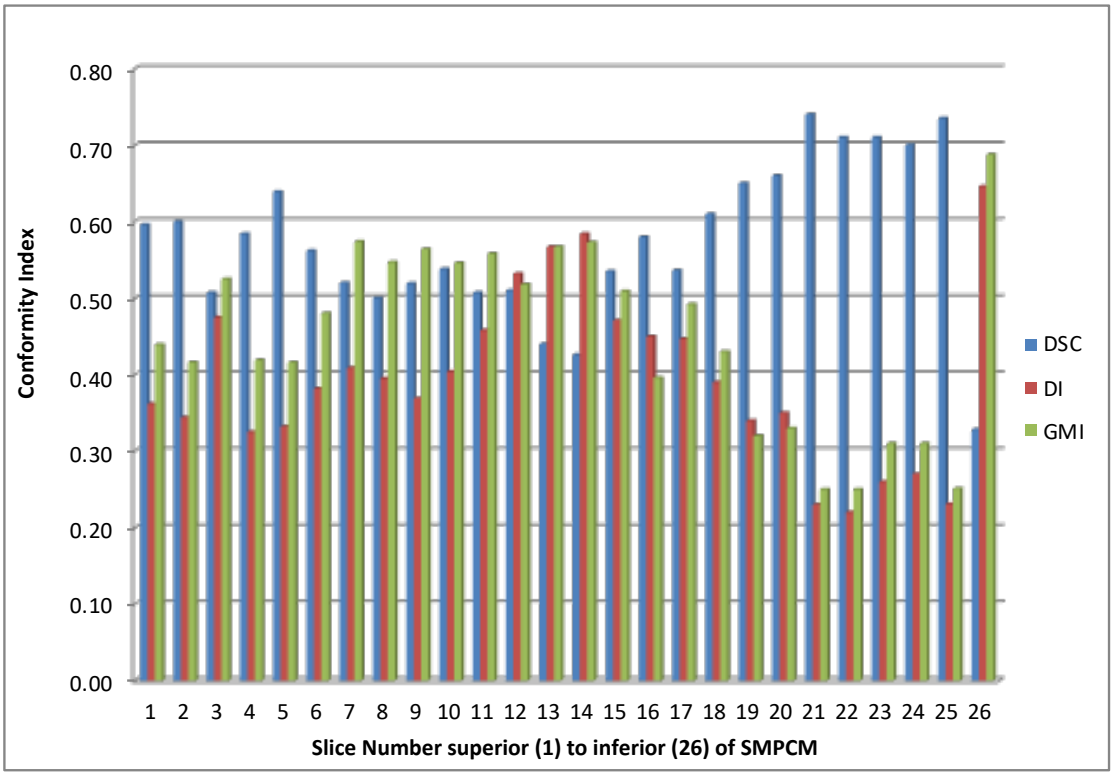
Table 5.2 Results for conformity indices and distance to agreement for ipsilateral and contralateral parotid gland and brainstem

Structure	Ipsilateral parotid gland					Contralateral Parotid gland				
	DSC	DI	GMI	Mean DTA (mm)	Max DTA (mm)	DSC	DI	GMI	Mean DTA (mm)	Max DTA (mm)
Range	0.85-0.89	0.05-0.15	0.12-0.2	1.3-2.1	7.2-15.5	0.82-0.90	0.08-0.17	0.08-0.20	1.1-1.6	6.0-12.2
Median	0.87	-	-	-	-	-	-	0.10	-	-
Mean	-	0.09	0.16	1.6	1.18	0.87	0.11	-	1.4	9.3
SD	0.01	0.03	0.03	0.02	0.33	0.02	0.03	0.04	0.02	0.2
95 % CI	-	0.07-0.12	0.14-0.19	1.4-1.8	9.1-14.5	0.85-0.89	0.09-0.14	-	1.2-1.6	7.8-1.1
IQR	0.87-0.87	-	-	-	-	-	-	0.10-0.13	-	-

Structure	Brainstem				
	DSC	DI	GMI	Mean DTA (mm)	Max DTA (mm)
Range	0.74-0.88	0.05-0.25	0.15-0.38	1.2-3.3	4.1-11.9
Median	-	-	-	-	-
Mean	0.82	0.12	0.23	2.0	0.73
SD	0.05	0.07	0.08	0.07	0.23
95 % CI	0.78-0.86	0.06-0.17	0.17-0.29	1.5-2.6	5.5-9.1
IQR	-	-	-	-	-

5.8.1.3 Slice by slice analysis

For slice-by-slice analysis of PI-SMPCM contours, the median s-DSC was 0.57 (IQR 0.51 – 0.65); s-GMI, 0.46 (IQR 0.33 – 0.55); and s-DI 0.39 (IQR 0.33 – 0.46). Corresponding values for PI-IPCM were 0.70 (IQR 0.50 – 0.76); 0.22 (IQR 0.16 – 0.46); and 0.34 (IQR 0.23 – 0.59) respectively. There was considerable variation in defining the superior-inferior extents of both SMPCM and IPCM relative to GS, with perfect concordance observed in only one PI-IPCM and three PI-SMPCM delineations respectively. Apart from the caudal-most slice, the highest agreement with the GS-SMPCM contours was observed inferiorly for slices 21 – 25, with median s-DSC > 0.7 and low values of s-GMI (0.25) and s-DI (0.23) respectively. Positional analysis showed that the largest variation was noted mid-way between the superior and inferior slices (Figure 5.8 and 5.9 respectively). Visual assessment of the contours showed that this was a consequence of PIs not extending the SMPCM contours laterally to encompass the pterygoid muscle, as recommended. The s-CI values for PI-IPCM imply that the relatively poor corresponding whole volume CI values were largely due to uncertainty in defining the superior and inferior extent of this structure.



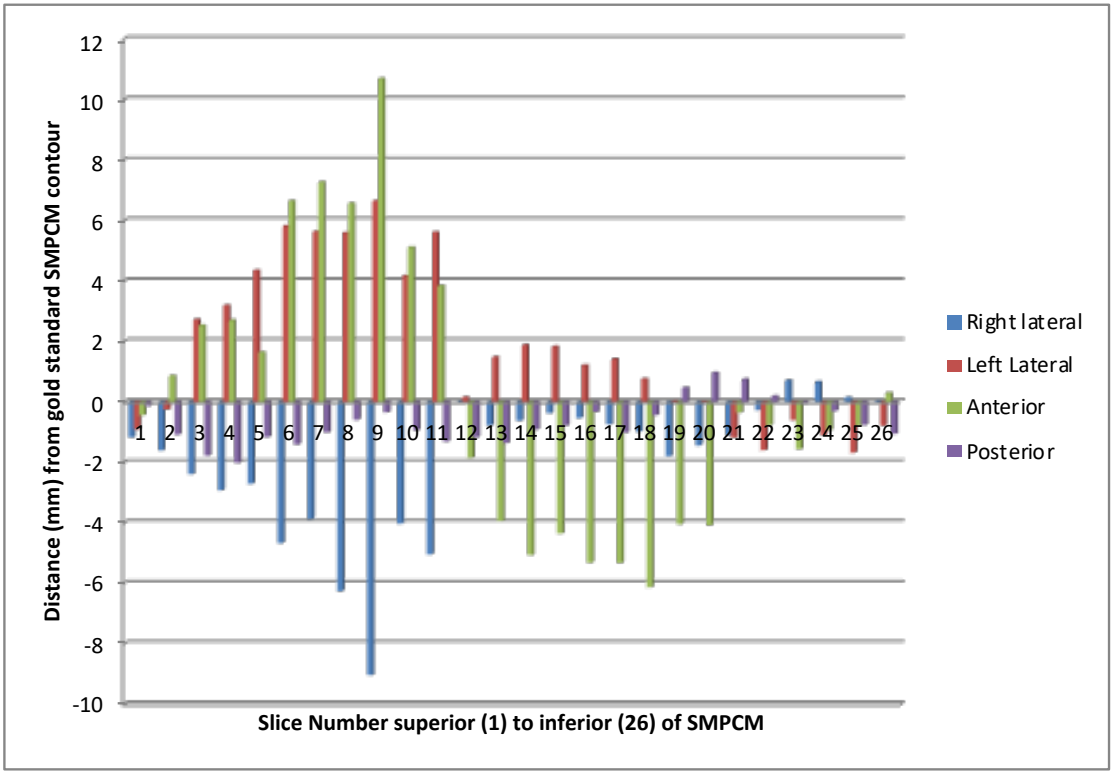
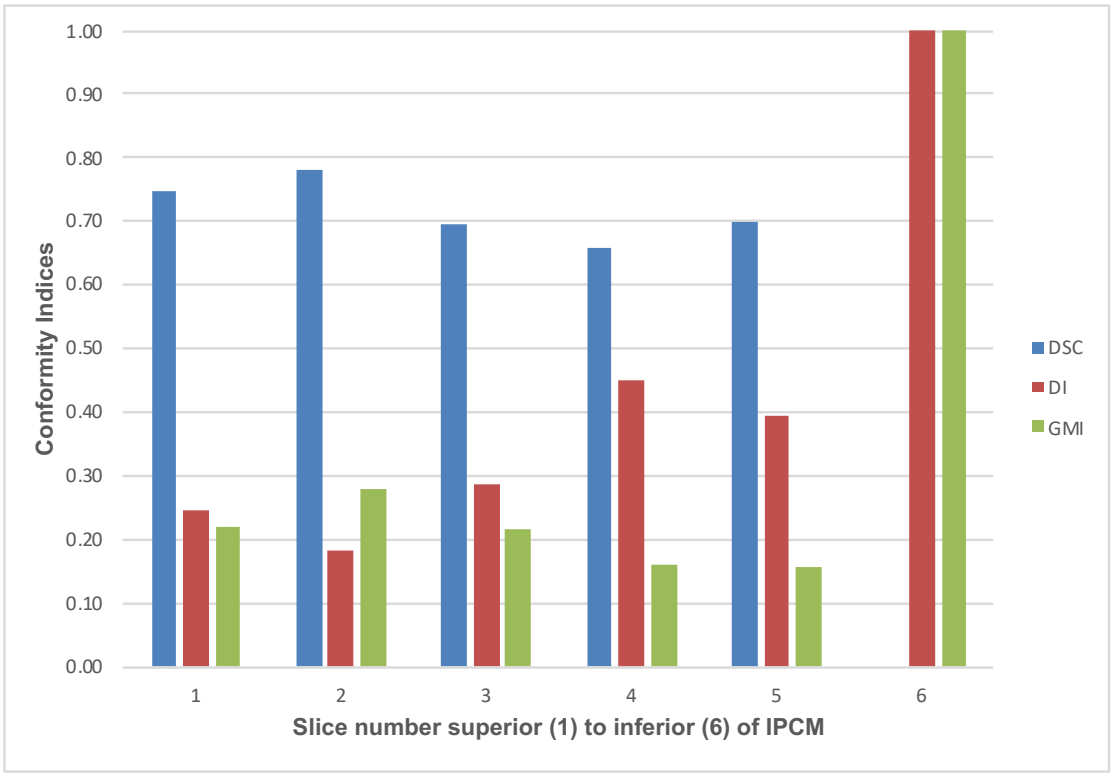


Figure 5.8 Slice-by-slice conformity (top) and positional (bottom) analysis of PI-SMPCM contours



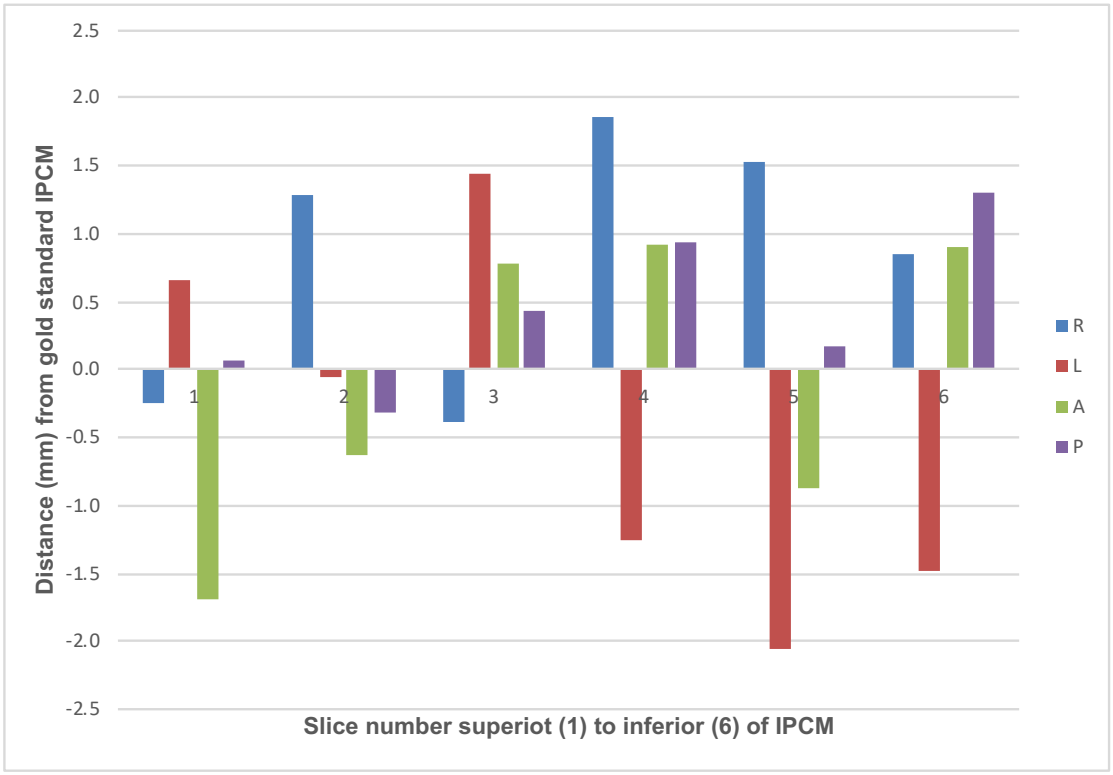


Figure 5.9 Slice-by-slice conformity (top) and positional (bottom) analysis of PI-IPCM contours

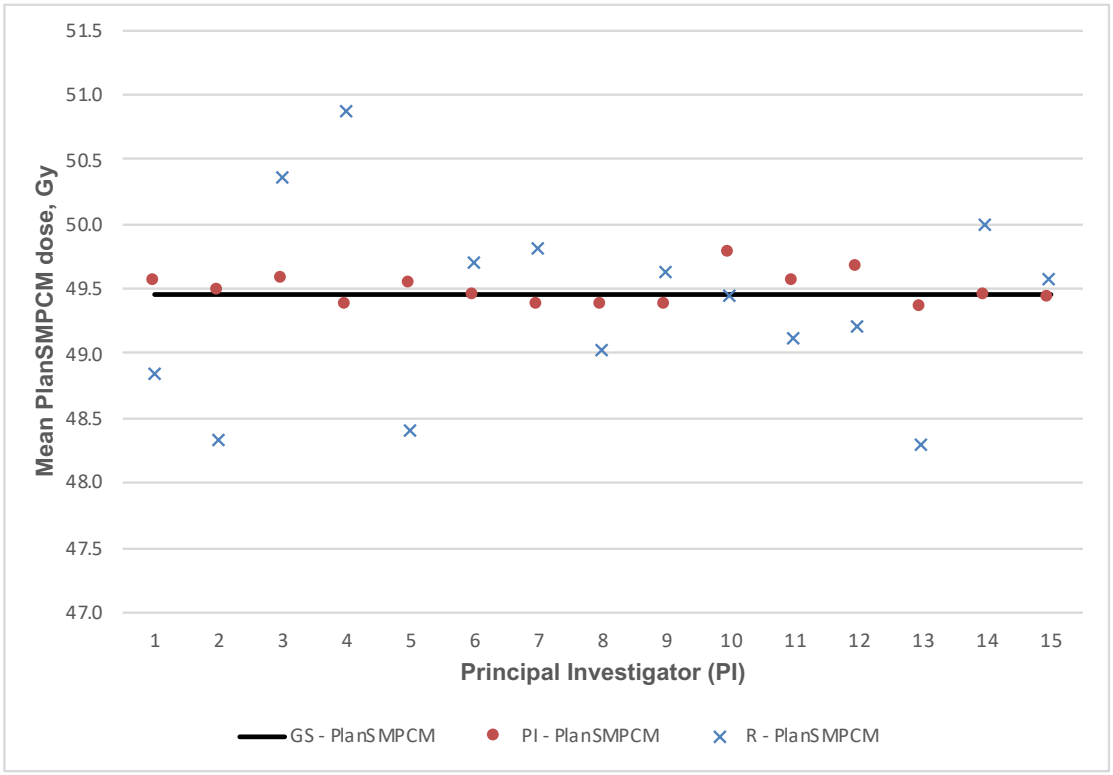
5.8.1.4 Dose delivered

Dose delivered to the SW-OAR is presented in figures 5.10 and 5.11 respectively.

Mean GS – PlanSMPCM was 49.5 Gy. This was similar to the mean of the mean dose to PlanSMPCM with PI Do-IMRT plans (49.5 Gy; range 49.4 Gy to 49.8 Gy; SD 0.1). The average difference between PI-PlanSMPCM and R-PlanSMPCM mean dose (49.4 Gy; range 48.3 Gy to 50.9 Gy; SD 0.7) was small at 0.1 Gy. For 3 PI-PlanSMPCM contours, the true dose delivered to that structure, ie R-PlanSMPCM, would be greater than the dose constraint of < 50 Gy.

For the SPC, the mean GS dose of 48.3 Gy was, on average, 0.5 Gy lower than the mean PI-SPC dose (48.8 Gy; range 48.2 Gy to 49.4 Gy; SD 0.3). There was a 0.3 Gy mean difference between PI-SPC and R-SPC mean doses (48.5 Gy; range 47.2 Gy to 49.9 Gy; SD 0.7).

Mean GS-PlanIPCM dose was 20.2 Gy. The mean of the mean PI-PlanIPCM dose was 20.6 Gy (range 19.6 Gy – 22.3 Gy, SD 0.8); corresponding value for R-PlanIPCM was 19.4 Gy (range 15.6 Gy to 24.9 Gy, SD 2.4).



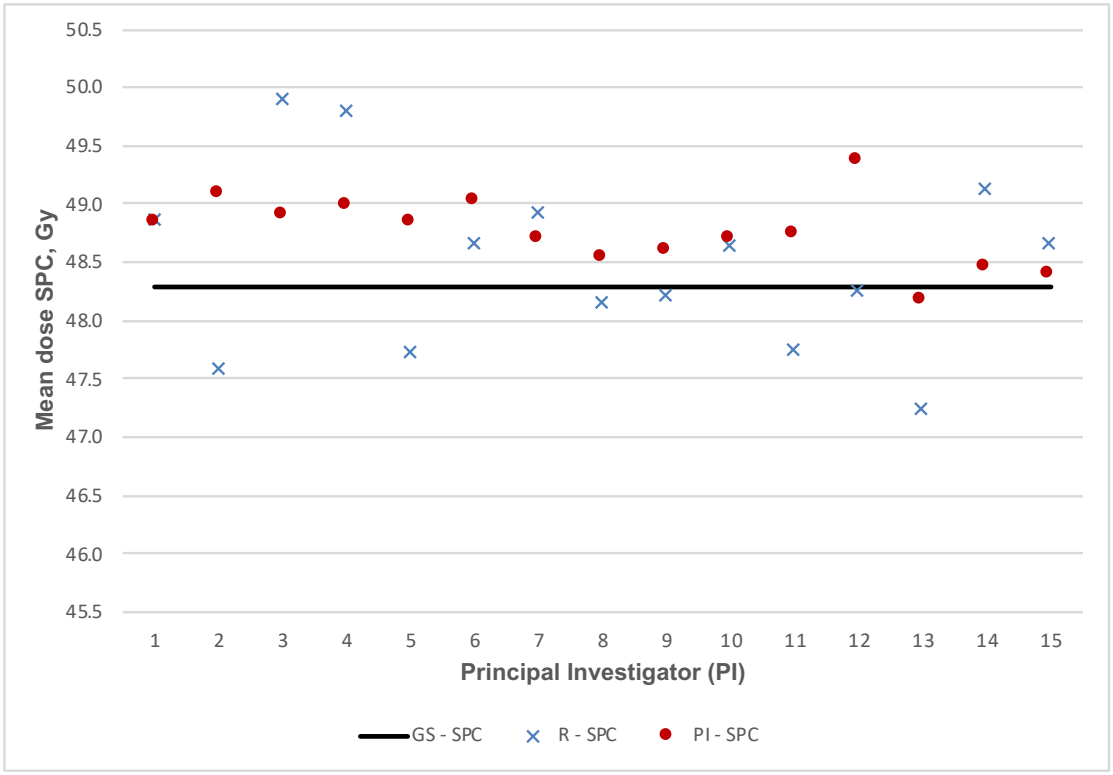


Figure 5.10 Mean dose delivered to PlanSMPCM (top) and SPC (bottom) with PI Do-IMRT plan (PI-PlanSMPCM, PI-SPC), and the GS contour superimposed on the PI Do-IMRT plan(R-PlanSMPCM, R-SPC).

The horizontal line represents the mean dose delivered to the structures on the GS plan, based on GS contours.

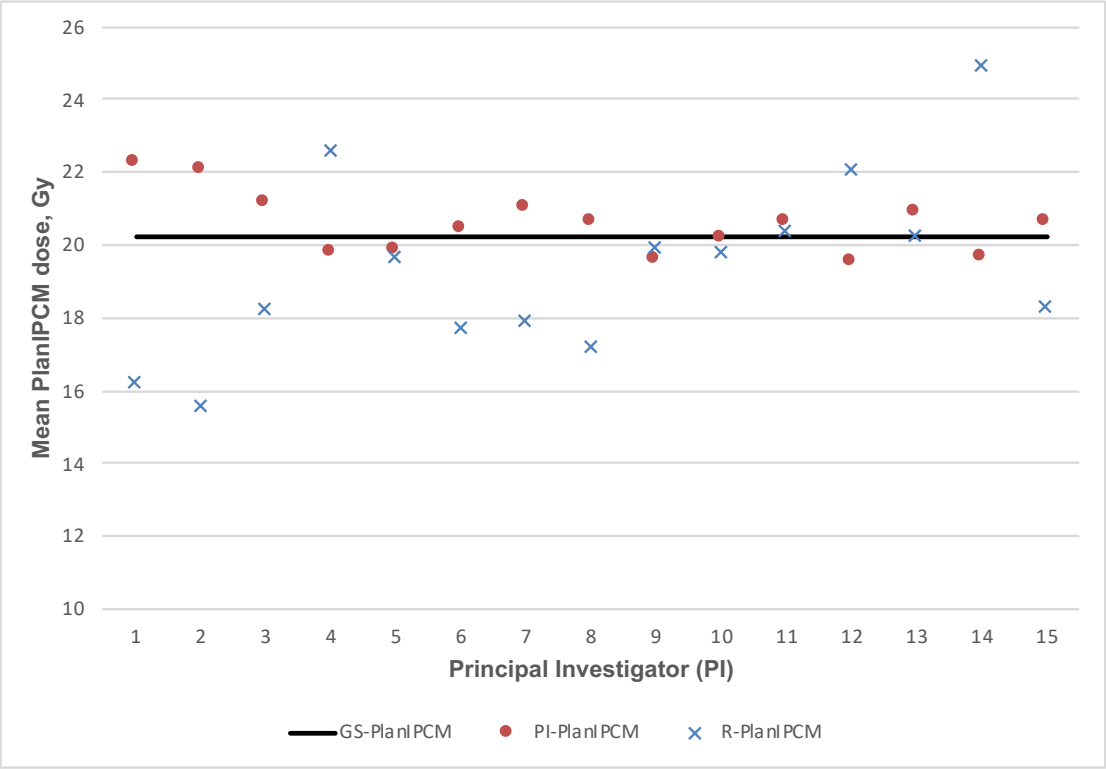


Figure 5.11 Mean dose delivered to Plan IPCM (top) with PI Do-IMRT plan (PI-PlanIPCM), and the GS contour superimposed on the PI Do-IMRT plan(R-PlanIPCM).

The horizontal line represents the mean dose delivered to the structures on the GS plan, based on GS contours.

5.8.1.5 NTCP

The estimated risk of swallowing dysfunction at 6 months is shown in figure 5.12. There was no significant difference in either the PI-NTCP or R-NTCP, compared to the GS-NTCP of 24.7%.

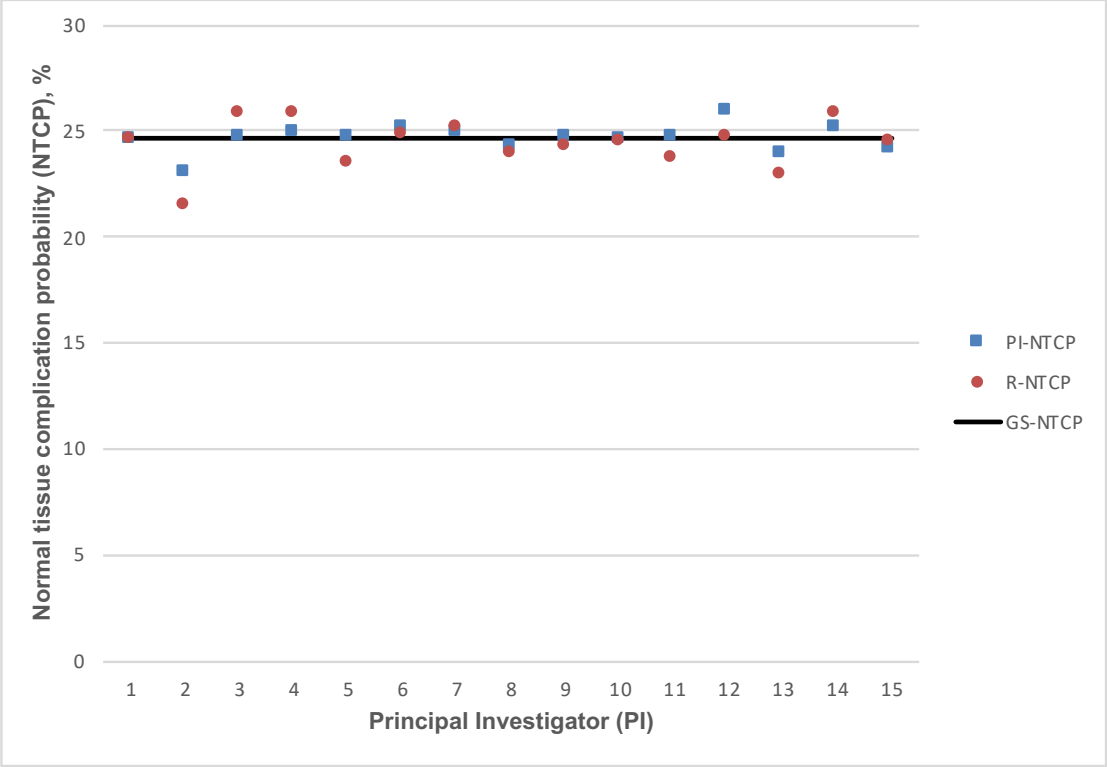


Figure 5.12 NTCP values for physician-scored RAD_{6M} based on PI Do-IMRT plan.

The horizontal line represents the NTCP value for the GS Do-IMRT plan, based on GS PCM contours

5.9 Discussion

IOV in target volume contouring in HNC has been extensively evaluated, in contrast to OAR variability where there have been few studies only[12-15]. The purpose of this study was to quantify the inter-clinician variability in PCM contouring within the context of a pre-trial contouring QA, and more importantly, to determine whether any inconsistency in its delineation affected the planned dose metrics and subsequent probability of RAD.

We have shown that PI-SMPCM volumes were predominantly smaller than the GS volume, with substantial variance between different PIs. We have also shown that conformity to the GS volume for both SMPCM and IPCM was poor, as evidenced by the low DSC and high DI and GMI scores. In contrast, there was good inter-clinician agreement for defining the parotids and brainstem. An explanation for the difference in observed CI values is that CT provides good spatial resolution and soft tissue contrast to help identify the standard OARs, thereby permitting consistent contouring. On the other hand, the PCM is not readily visualised on CT and its delineation is therefore reliant on accurate interpretation of the contouring guidelines based on different anatomical landmarks, which is likely to have contributed to the higher degree of variation observed in this study. For instance, the cranial and caudal extent of PCM was subject to substantial IOV implying uncertainty in identifying the tip of the pterygoid plates and the lower edge of the arytenoid cartilages, which may be due to unfamiliarity with identifying these on CT. Spatial assessment for PI-SMPCM delineation additionally demonstrated that concordance with the GS contour was poor in the middle section of this structure, where the lower s-GMI and s-DSC compared to the mean overall GMI and DSC suggested under-outlining as the contouring error. Visual assessment of the discordant slices identified that under-outlining was often due to failure to extend the delineation of SMPCM laterally to encompass the pterygoid muscle. It is also pertinent to consider the relatively smaller volume of the constrictors relative to the standard OARs when interpreting the differential CI values. CIs are more sensitive to small volumes, as a few missing or extra voxels on one contour is sufficient to skew their values; on the other hand they are more forgiving for

larger volumes such as the parotids where a relatively larger variation is required to demonstrate a comparable CI result.

CIs are the most commonly used metric for reporting IOV in RT studies and are helpful in determining similarity in volumes, particularly in comparison to a reference volume. In this study, an 'expert-defined' gold standard was used as the benchmark contour against which all PI contours were compared, rather than a mathematically-derived consensus contour from a number of contours defined by different oncologists such as STAPLE. Both approaches have been adopted in various quality assurance studies for different tumour sites, dependant upon the study endpoint. It must be emphasised that the focus of our study was to understand the variability in PCM delineation and subsequent dosimetric impact, and therefore whether the GS contour accurately represented the structure is less relevant. The use of an expert-defined reference has helped us to identify a systematic pattern of under-contouring of the constrictor muscle near the pterygoid muscle, which would have been obscured by the use of a consensus contour.

An inherent limitation of whole-volume CIs is that it does not provide sufficient information about differences in size, shape or location that may exist between 2 volumes. Similar CI values for different contours, therefore, do not necessarily indicate that the contours are identical. For instance, one PI achieved a DSC of 0.65 (ranked 1st of 15), GMI of 0.23 (ranked 1st of 15), but a DI of 0.43 (ranked 11th of 15) for SMPCM delineation. Visual assessment of the PI contours, however, showed the contour did not extend to the pterygoid

muscle, thereby inadvertently reducing the volume of PlanSMPCM. This could have detrimental dosimetric implications for the mean dose delivered to this structure with Do-IMRT, and the PI was requested to re-submit. On the other hand, no clinically relevant SMPCM contouring error was identified for another PI who scored a DSC of 0.62 (2^{nd} of 15), GMI of 0.34 (3^{rd} of 15) and DI of 0.43 (10^{th} of 15). Clinically relevant errors may therefore be missed if whole-volume CI alone were used to establish levels of agreement between contours.

The observed discrepancies in defining this SW-OAR highlight the importance of pre-trial contouring QA. Systematic errors can be identified early and rectified. For instance, three PIs wrongly assumed the caudal edge of cricoid cartilage as the inferior border of the IPCM, despite the presence of a detailed RT protocol and a CT atlas. Our contouring variability results are consistent with previously published studies that have been described in chapter 1. Alterio et al additionally showed that there was increased intra- and inter-observer variability in SPC delineation, along with lower adherence to the corresponding MRI-contoured muscle, amongst 34 HN oncologists[16].

Despite volumetric, overlap, and spatial variability in contouring of the PCM, there was no difference in either the mean dose to this structure or the risk of persistent swallowing dysfunction compared to GS, even for PIs (n = 7) who were requested to re-submit due to unacceptable delineation of the muscle. Such an outcome suggests that variability in the delineation of this SW-OAR does not impact on the dose delivered with Do-IMRT. Before drawing firm conclusions to that effect, it is pertinent to consider certain factors that may

influence the dosimetry and toxicity outcomes in clinical practice. This analysis was conducted on a single benchmark case with minimal PTV65-PCM overlap, and it is possible that the clinical outcomes with similar PCM contouring variability could differ with increasing number of cases and/or greater overlap. Furthermore, the ball diameter used to contour the PCM with few PIs was wider than the 3 mm used for the GS contour. Consequently, there was a larger dose gradient on the PI plans relative to the GS plan, explaining why the mean doses to the GS on PI plans was smaller. Finally, variability in SGL delineation was not assessed in this study and there is a probability that outlining uncertainties for this structure may lead to different toxicity outcomes than the one presented in this study.

5.10 Conclusion

There was considerable variation in the contouring of the PCM, based on CI values and DTA measurements. Despite this, there was no significant difference in doses delivered to this structure, relative to the GS contour.

5.11 References list

1. Stapleford, L.J., et al., *Evaluation of automatic atlas-based lymph node segmentation for head-and-neck cancer.* Int J Radiat Oncol Biol Phys, 2010. **77**(3): p. 959-66.

2. Peters, L.J., et al., *Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: results from TROG 02.02.* J Clin Oncol, 2010. **28**(18): p. 2996-3001.

3. Vinod, S.K., et al., *Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies.* Radiotherapy and Oncology, 2016. **121**(2): p. 169-179.

4. Warfield, S.K., K.H. Zou, and W.M. Wells, *Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation.* IEEE Trans Med Imaging, 2004. **23**(7): p. 903-21.

5. Hanna, G.G., A.R. Hounsell, and J.M. O'Sullivan, *Geometrical analysis of radiotherapy target volume delineation: a systematic review of reported comparison methods.* Clin Oncol (R Coll Radiol), 2010. **22**(7): p. 515-25.

6. Mattiucci, G.C., et al., *Automatic delineation for replanning in nasopharynx radiotherapy: what is the agreement among experts to be considered as benchmark?* Acta Oncol, 2013. **52**(7): p. 1417-22.

7. Zijdenbos, A.P., et al., *Morphometric analysis of white matter lesions in MR images: method and validation.* IEEE Trans Med Imaging, 1994. **13**(4): p. 716-24.

8. Thomson, D., et al., *Evaluation of an automatic segmentation algorithm for definition of head and neck organs at risk.* Radiat Oncol, 2014. **9**: p. 173.

9. Gwynne, S., et al., *Toward semi-automated assessment of target volume delineation in radiotherapy trials: the SCOPE 1 pretrial test case.* Int J Radiat Oncol Biol Phys, 2012. **84**(4): p. 1037-42.

10. Gwynne, S., et al., *Improving radiotherapy quality assurance in clinical trials: assessment of target volume delineation of the pre-accrual benchmark case.* Br J Radiol, 2013. **86**(1024): p. 20120398.

11. Nelms, B.E., et al., *Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer.* Int J Radiat Oncol Biol Phys, 2012. **82**(1): p. 368-78.

12. van der Veen, J., A. Gulyban, and S. Nuyts, *Interobserver variability in delineation of target volumes in head and neck cancer.* Radiother Oncol, 2019. **137**: p. 9-15.

13. Hermans, R., et al., *Laryngeal tumor volume measurements determined with CT: a study on intra- and interobserver variability.* Int J Radiat Oncol Biol Phys, 1998. **40**(3): p. 553-7.

14. Christiaens, M., et al., *Quality assurance of radiotherapy in the ongoing EORTC 1219-DAHANCA-29 trial for HPV/p16 negative squamous cell carcinoma of the head and neck: Results of the benchmark case procedure.* Radiother Oncol, 2017. **123**(3): p. 424-430.

15. Loo, S.W., et al., *Interobserver variation in parotid gland delineation: a study of its impact on intensity-modulated radiotherapy solutions with a systematic review of the literature.* Br J Radiol, 2012. **85**(1016): p. 1070-7.

16. Alterio, D., et al., *Contouring of the Pharyngeal Superior Constrictor Muscle (PSCM). A cooperative study of the Italian Association of Radiation Oncology (AIRO) Head and Neck Group.* Radiotherapy and Oncology, 2014. **112**(3): p. 337-342.

6 Chapter 6: Thesis summary and future directions

6.1 Chapter 2: Do-IMRT to reduce RAD in LA-OPC

Previous non-randomised studies have identified a strong correlation between irradiation of PCM, in particular SPC, and risk of persistent swallowing dysfunction following RT-based treatment in LA-OPC. The comparative planning study results from this chapter demonstrated that it was possible to significantly reduce the mean dose delivered to the constrictor muscle with the novel PCM-sparing Do-IMRT planning technique, relative to S-IMRT, in patients receiving bilateral neck irradiation for LA-OPC. The DVH parameters for reducing dose to the PCM with Do-IMRT were derived from published hypothesis-generating studies. An additional important finding was that Do-IMRT reduced the mean dose to the SGL, another SW-OAR implicated in RAD. We demonstrated that these reductions in delivered dose to the SW-OARs lowered the estimated risk of physician-scored RAD_{6M} by > 5% compared to S-IMRT in each patient in the study cohort. This was achieved without compromising PTV65 coverage, spillage of dose to the salivary glands, or increasing integral dose. It was possible to generate good quality PCM-sparing RT plans with minimal iteration by setting dose-based objective function; in contrast the Dutch groups' SW-IMRT technique has been shown to be time-consuming and difficult to replicate in clinical practice. These promising results suggest that the relatively easy to implement Do-IMRT, compared to

SW-IMRT, may be an effective swallow – and salivary gland – sparing RT strategy for LA-OPC, and justifies the need to undertake a prospective randomised study to validate the toxicity-mitigating benefits of Do-IMRT.

6.2 Chapter 3: PBT to reduce RAD in LA-OPC

The introduction of the PBT in the UK necessitates the identification of patients who will be expected to benefit most from protons compared to IMRT. This chapter tested the hypothesis that the implementation of IMPT would result in significant reductions in the delivered dose to the constrictor muscle, and consequently lower toxicity, relative to IMRT in LA-OPC. Four PBT plans were generated, and compared to the corresponding IMRT plans from chapter 1, for each patient in this retrospective planning study; a S-IMPT plan, and 3 Do-IMPT plans using different optimisation techniques and beam arrangements. We have shown that there was a significant reduction in the mean dose to MPC (3 Gy), IPC (4 Gy), PCM (1.4 Gy), and SGL (5.0 Gy) with Do-IMPT$_{RO}$ compared to Do-IMRT. However, there was no difference in the delivered dose to the SPC, and the study failed to meet its primary endpoint of demonstrating a mean reduction of > 5 % in the estimated risk of persistent swallowing toxicity for the entire cohort with Do-IMPT, compared to Do-IMRT. This threshold was demonstrated in two patients, suggesting that careful selection of patients suitable for PBT may be required. It may be possible that the incremental benefits demonstrated in this study with Do-IMPT$_{RO}$ could be improved by setting a dose constraint on SGL, and in combination with the significantly reduced mean dose delivered to the contralateral parotid and

integral dose this may provide significant improvements in patients' HR-QoL with PBT compared with IMRT. This study has also demonstrated that the nominal robustly optimised proton plan was superior to the corresponding PTV-based proton plans for SW-OAR sparing.

6.3 Chapter 4: Robustness of PBT

This chapter addressed two key aspects pertinent to PBT planning: 1) resilience of the nominal Do-IMPT$_{RO}$ and Do-IMPT$_{PTV}$ plans to treatment uncertainties; 2) sensitivity of PTV-based proton plans as a function of number of beams used.

We demonstrated that the nominal Do-IMPT$_{RO}$ was more robust to treatment uncertainties compared to PTV-based PBT, particularly for target coverage. This was observed if there were to be a combination of SE and RU, while RU on its own did not appear to have a detrimental impact on plan quality for PTV-based IMPT. RO-based plans require significant computational time compared to traditional margin-based ones, though this could be offset by the time required to replan due to suboptimal coverage during robust evaluation of PTV-based IMPT. This was not the aim of this study, and therefore not evaluated. Furthermore, it was also shown that increasing the number of beams for Do-IMPT$_{PTV}$ did not improve the resilience of the nominal PTV plan to treatment uncertainties.

6.4 Chapter 5: Inter-observer variability in PCM delineation

This chapter was a comprehensive analysis of variability in the delineation of the PCM, and consequent impact on predicted toxicity outcomes, amongst UK HN oncologists within the context of a pre-trial quality assurance programme. We have shown that there were substantial volumetric, overlap, and spatial differences in PCM outlining, in comparison to the gold standard contour. Significant delineation errors were due to a combination of inaccurate interpretation of the contouring guidelines, and unfamiliarity with CT-based anatomical landmarks. In particular, we identified a systematic error in the definition of the SMPCM in the region of pterygoid muscle. This inconsistent definition, however, did not have a detrimental impact on dose delivered to it or estimated risk of RAD_{6M} in comparison to the GS. As this analysis was conducted on a single case with minimal PTV65-PCM overlap, it is premature to conclude that inter-observer variability in delineation of PCM is not associated with any dosimetric differences to this structure.

6.5 Future directions

This thesis has demonstrated a need for conclusive evidence to establish the benefits of swallow-sparing RT techniques for LA-OPC on long-term function. Prospective validation of the predicted benefits of Do-IMRT presented in chapter 2 is required before adopting it into routine clinical practice. In this context, results from the randomised phase III DARS trial should establish the

definitive role of this novel RT in patients with LA-OPC. Unlike current evidence that predominantly relies on physician-scored swallowing dysfunction at 6 months as a benchmark for persistent toxicity, the availability of both patient-reported and physician-scored swallowing outcomes, along with VF-based data at 12 months in DARS will provide robust information about the clinical benefits of Do-IMRT on long-term dysphagia. In parallel to establishing clinical outcomes from the study, analysis of data from the above prospective study would help to determine DVH parameters for the constrictors and pre-treatment factors that predict an increased risk of patient-reported RAD following arc-based RT in LA-OPC, that could then be used to develop a predictive model. Further analysis would involve contouring different swallowing structures implicated with dysphagia, such as the oral cavity and floor of mouth muscles, and studying the relationship between doses delivered to them and toxicity outcomes.

This thesis has also demonstrated the presence of considerable variation in PCM contouring amongst the participating PIs from the DARS pre-trial RTTQA programme. As part of the DARS on-trial quality assurance programme, the target volumes and OARs were prospectively reviewed for atleast the first two recruited patients at each centre, to ensure adherence to the protocol. Further analysis would involve determining the number of cases that were deemed to have significant protocol deviation with regards to PCM definition. It would also be of interest to perform a retrospective review of PCM outline for all patients recruited into DARS; this would involve comparison of constrictor delineation by PIs with gold standard, re-outlined by the central team. This would provide

more exhaustive information regarding the utility of CI and DTA in predicting toxicity outcomes, along with assessing whether the variation in contouring decreased with time as PIs became more familiar with the delineation guidelines. Such a study would also guide whether future trial with novel targets/OARs should involve prospective central review for all recruited patients.

Finally, from the proton data presented in this thesis that demonstrated the dosimetric advantages of PBT compared to IMRT in OPC, a phase III multi-centre randomised controlled trial looking at IMPT versus IMRT as a toxicity-mitigating strategy in HPV-associated oropharyngeal cancer is proposed. The primary endpoint would be the difference in the mean UW-QoL physical composite score at 12 months following treatment completion. The trial proposal was received positively at the national Clinical and Translational Radiotherapy Research working group meeting, and is likely to be amongst the first PBT trials in the UK.

6.6 Conclusion

The work presented in this thesis demonstrated the feasibility of generating clinically deliverable dysphagia-optimising RT plans that resulted in significant reductions in doses to the PCM and, therefore, reduced the probability of subsequent treatment-induced long-term swallowing impairment.