QUANTIFYING, UNDERSTANDING AND PREDICTING DIFFERENCES BETWEEN PLANNED AND DELIVERED DOSE TO ORGANS AT RISK IN HEAD & NECK CANCER PATIENTS UNDERGOING RADICAL RADIOTHERAPY TO PROMOTE INTELLIGENTLY TARGETED ADAPTIVE RADIOTHERAPY

David Jonathan Noble

Department of Oncology

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Emmanuel College

September 2019

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

All work done in collaboration is specified in the Acknowledgements section of each data chapter.

It does not exceed the prescribed word limit of 60,000 words for the Clinical Medicine and Clinical Veterinary Medicine Degree Committee.

ABSTRACT - Quantifying, Understanding and Predicting Differences Between Planned and Delivered Dose to Organs at Risk in Head & Neck Cancer Patients Undergoing Radical Radiotherapy to Promote Intelligently Targeted Adaptive Radiotherapy.

Dr David Jonathan Noble I Emmanuel College I Department of Oncology

Introduction: Radical radiotherapy (RT) is an effective but toxic treatment for head and neck cancer (HNC). Contemporary radiotherapy techniques sculpt dose to target disease and avoid organs at risk (OARs), but anatomical change during treatment mean that the radiation dose delivered to the patient – *delivered* dose (D_A), is different to that anticipated at planning – *planned* dose (D_P). Modifying the RT plan during treatment – Adaptive Radiotherapy (ART) – could mitigate these risks by reducing dose to OARs. However, clinical data to guide patient selection for, and timing of ART, are for lacking.

Methods: 337 patients with HNC were recruited to the Cancer Research UK VoxTox study. Demographic, disease and treatment data were collated, and both D_P and D_A to organs at risk (OARs) were computed from daily megavoltage CT image guidance scans, using an open-source deformable image registration package (Elastix). Toxicity data were prospectively collected. Relationships between D_P, D_A and late toxicities were investigated with univariate, and logistic regression normal tissue complication probability (NTCP) modelling approaches. A sub-study of VoxTox recruited 18 patients who had MRI scans before RT fractions 1, 6, 16, and 26. Changes in salivary gland volumes and relative apparent diffusion coefficient (ADC) values were measured and related to toxicity events.

Results: Spinal cord dose differences were small, and not predicted by weight loss or shape change. Mean D_A to all other OARs was higher than D_P; factors predicting higher D_A included primary disease site, concomitant therapy, shape change and advanced neck disease. Nine patients (3.7%) saw $D_A > D_P$ by ≥2Gy to more than half of the OARs assessed. These patients all had received bilateral neck RT for N-stage ≥2b oropharyngeal cancer. Strong uni- and multivariate relationships between OAR dose and toxicity were seen. Differences between D_A and D_P-based dose-toxicity models were minimal, and not statistically significant. On MRI, both parotid and submandibular glands shrank during treatment, whilst relative ADC rose. Relationships with toxicity were inconclusive.

Conclusions: Small differences between OAR D_P and D_A mean that D_A-based toxicity prediction models confer negligible additional benefit at the population level. Factors such as primary disease sub-site, concomitant systemic therapy, staging and shape change may help to select the patients that do develop clinically significant dose differences, and would benefit most from ART for toxicity reduction.

Acknowledgements

First and foremost I would like to thank my 2 principle supervisors Professor Neil Burnet, and Dr Raj Jena. Throughout the course of my fellowship, both have consistently provided intellectual and practical assistance with my work as well as mentorship, guidance and support as supervisors, colleagues and friends, and I am enormously grateful for their help. My second supervisor, Dr Ferdia Gallagher, has also been extremely supportive, and provided significant practical help in setting up the Minot-OAR project. My sincere gratitude also goes to the whole head & neck team especially the consultants Dr Richard Benson, Dr Gill Barnett, Dr Sarah Jefferies, and Dr Rashmi Jadon who patiently acquiesced with my persistent demands for contours, provided a useful sounding board for ideas, and made me feel a valued member of the team.

This project has only been possible through close collaborative working with a big team, and I owe a deep debt of gratitude to a number of talented, dynamic and supportive individuals, too numerous to list here. However, the roles of certain people do merit specific mention. Amy Bates has contributed so much to the VoxTox project, and deserves enormous credit for recruiting so many patients to the study. Karl Harrison has been an excellent colleague, and has taught me a great deal about image processing, analysis, and data science. Leila Shelley has been my 'PhD buddy' and I have been very lucky to have her company in a fascinating and fun journey into academic radiation oncology.

I have been lucky to have the chance to play competitive league cricket again whilst doing my research, and I am very grateful to the men and women of Histon Cricket Club for keeping me sane, keeping my feet on the ground, and for not being remotely interested in my day job. Finally I would like to thank my close friends, and my family, especially my wife Anna, and my children Alexander and Laura for their love and support, and the joy they bring to my life.

THIS THESIS IS DEDICATED TO THE MEMORY OF
DR. NADEERA DE SILVA

Table of contents

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Abbreviations/Glossary of terms

ABAS – Atlas-based automated segmentation tool. Software packages that generate automated segmentation based on a library of examples and deformable image registration.

Accelerated RT – Delivering radiotherapy in a shorter overall treatment time, typically by treating more than one time a day, or on more days during the week.

ADC – Apparent diffusion coefficient. A quantification of the Brownian motion of water on an MRI image.

ADL – Activities of daily living.

ART – Adaptive radiotherapy.

AUC – Area under the curve – often used to quantify ROC curves.

BNI – Bilateral neck irradiation. Radiotherapy treatment for head and neck cancer where lymph node regions in both sides of the neck are treated.

B-spline – A mathematical function which underlies the deformable component of some image registration algorithms.

CBCT – Cone-beam CT. A type of Image guidance using kV energy X-Rays produced by the linear accelerator head.

CRF – Clinical reporting form – toxicity assessment questionnaires specifically designed for the VoxTox project, that allow population of a number of recognised toxicity scoring systems.

CRT – Chemo-radiotherapy. Chemotherapy delivered concomitantly alongside radiotherapy, often to act as a radio-sensitiser**.**

CSF – Cerebrospinal fluid

CTCAE – Common Terminology Criteria for Adverse Events. A set of criteria for quantifying the severity of adverse events following cancer treatment.

CTV – Clinical target volume. A volume in radiotherapy planning that takes into account microscopic spread of disease beyond what is known or visible.

$\mathbf{D}_{A}$ – Delivered [radiation] dose.

$\mathbf{D}_{P}$ – Planned [radiation] dose.

DARS – Dysphagia/aspiration related structures. A concept that includes anatomical structures involved in the process of swallowing that can be seen and contoured on imaging.

DBC – Distance between centres. A measure of contour agreement.

DDA – Dose difference assessment. The standard technique in the Addenbrooke's hospital head and neck radiotherapy unit for estimating differences between planned and delivered radiation dose.

DICOM – Digital Imaging and Communications in Medicine. The standard file format for the storing, management, and transfer of medical images.

DIPC – Daily image guidance with positional correction.

DIR – Deformable image registration.

DSC – Dice similarity coefficient. A measure of contour agreement.

DTC – Distance to conformity. A measure of contour agreement.

DVH – Dose volume histogram.

DWI – Diffusion weighted imaging. Using specific MRI sequences to generate contrast from differences in the relative diffusion of water molecules.

$D_{2\%}$ - The minimum radiation dose to the 2% volume of a given structure that receives the highest dose in that structure. A pragmatic surrogate for maximum dose to that structure.

Elastix – Open-source deformable image registration package. Used for contour propagation from kVCT to MVCT images, and dose accumulation in this work.

EORTC QLQ H&N35 – A questionnaire designed by the European Organisation for Research and Treatment of Cancer to assess the quality of life of patients who have undergone radiotherapy for head and neck cancer.

Estro – European Society for Radiation Oncology.

FALCON – Fellowship in Anatomic DeLineation and CONtouring. A multifunctional platform for contouring and delineation workshops with Estro.

FDG – 18F-fluorodeoxyglucose. A radiolabelled tracer for positron emission tomography, PET, that highlights tissues with high glucose uptake.

FoV – Field of view.

FMISO – 18F-fluoromisonidazole. A radiolabelled tracer for positron emission tomography, PET, preferentially taken up into hypoxic tissues.

Gray (Gy) – The unit of radiotherapy dose, defined as 1 Joule of energy of radiation energy per kilogram of matter.

GTV – Gross tumour volume. Macroscopic tumour that can be seen palpated or imaged.

HD – Hausdorff Distance - a measure of contour agreement.

H&N – Head and neck.

HNC – Head and neck cancer.

HNSCC – Squamous cell carcinoma of the head and neck region.

HPV – Human papillomavirus. Now known to cause a substantial proportion of HNC, especially oropharyngeal cancer.

HRA – Health Research Authority

HU – Hounsfield unit. A quantitative scale for describing radiodensity, where the linear attenuation coefficient of a voxel is normalised to known values for air and water.

Hyperfractionation – Where radiotherapy treatment is delivered in fraction sizes smaller than 1.8 – 2Gy.

I_F: Fixed image. In deformable image registration, the image that provides the geometrical frame of reference, to which the moving image is optimally matched.

IG – Image guidance. The process whereby an image of the patient is taken in the treatment position (often by the linear accelerator itself) to ensure setup matches the plan as accurately as possible.

I_M: Moving image. In deformable image registration, the image that is moved and changed (the complexity of which changes depends on the details of the algorithm) in order to match the fixed image as closely as possible.

IMB – Imaging biomarker.

IMRT – Intensity modulated radiotherapy. A complex arrangement of radiotherapy beams with heterogeneous fluence, capable of delivering non-uniform levels of dose that may be accurately conformed to structures of interest.

IPC – Inferior pharyngeal constrictor muscle.

Jacobian determinant – A matrix that takes partial derivatives of each element of a transformation. In deformable image image registration, it gives an indication of whether a voxel in the moving image has increased or decreased its size, relative to baseline, in the process of optimising registration with the fixed image.

JCI – Jaccard conformity index. A measure of contour agreement (also referred to as CI).

kVCT – Computed tomography using X-Rays with kilo-voltage range energy. These scans may be used for diagnostics, radiotherapy planning, or image guidance.

LENT/SOM(A) – Late Effects Normal Tissue – Subjective, Objective, Management, Analytic. A system for scoring radiotherapy toxicity.

LND – Lateral neck dimension

LS – L'Hermitte's syndrome. A condition in which patients described an unpleasant sensation of electric shocks travelling down their back and into limbs.

MATLAB – A software package and programming language for technical computing.

MDA – Mean distance to agreement. A measure of contour agreement.

MDADI – MD Anderson Cancer Centre Dysphagia Inventory – a tool for scoring the severity of swallowing dysfunction after cancer treatment (specifically radiotherapy for head and neck cancer).

MI – Mutual information - a mathematical function to produce a similarity metric between fixed and moving images in image registration.

MPC – Middle pharyngeal constrictor muscle.

MRI – Magnetic resonance imaging.

MR-Linac – An MRI scanner and radiotherapy delivery machine (linear accelerator) combined into one unit. Image guidance is done with MRI, providing superior soft tissue definition compared with CT.

MSD – Mean squared difference – a mathematical function to produce a similarity metric between fixed and moving images in image registration.

MVCT - Computed tomography using X-Rays with mega-voltage range energy. Used exclusively for image guidance.

NCC – Normalised cross correlation - a mathematical function to produce a similarity metric between fixed and moving images in image registration.

NMI – Normalised mutual information - a mathematical function to produce a similarity metric between fixed and moving images in image registration.

NPC – Nasopharyngeal carcinoma.

NTCP – Normal tissue complication probability.

OAR – Organ at risk.

OC – Oral cavity

OR – Odds ratio

OS – Overall survival

PBT – Proton beam therapy.

pCT – Planning CT. The CT scan done specifically for radiotherapy planning.

PEG – A percutaneous endoscopic gastrostomy.

PET – Positron emission tomography.

PFS – Progression free survival.

PRV – Planning organs at risk volume. A margin added to a segmented OAR that accounts for the positional uncertainty of the structure, relative to the daily dose.

PTV – Planning target volume. A radiotherapy planning volume that takes into account differences in setup and, to some extent organ position.

PVD – Proportional volume difference. A measure of volume differences between 2 ROIs that makes no inference as to which is 'gold standard'.

Python – A computational programming language, very widely used in technically and mathematically orientated work.

QoL – Quality of life.

R – A powerful open-source software package for data analysis**.**

ROC curve – Receiver operator characteristic curve. A graph that plots sensitivity against 1-specificity, thereby quantifying the discriminative capabilities of a test.

ROI – Region of interest**.**

RT – Radiotherapy.

RTOG – Radiation Therapy Oncology Group. An organisation that published a scoring system for objectively measuring radiotherapy toxicity.

SC – Spinal cord.

SCC – Squamous cell carcinoma. 90% of primary head and neck cancers are SCC's.

SCT/SACT – Systemic anticancer therapy – drugs, including conventional cytotoxic, targeted therapies, immunotherapy and hormonal therapies used to treat cancer.

SDI – Salivary duct inflammation. A toxicity endpoint described in the CTACE scoring system, which describes a phenotype of thick, sticky, stringy saliva and oro-pharyngeal secretions, often with associated taste disturbance.

SGL – Supraglottic larynx.

SMG – Submandibular [salivary] gland**.**

SPC – Superior pharyngeal constrictor muscle.

SPSS – A statistical software package.

SSA – Slice surface area

SWOARS – Organs at risk involved in some way in the process of swallowing. In this work this is taken to mean both parotid and submandibular glands, the superior and middle pharyngeal constrictor muscles, the supraglottic larynx, and the oral cavity.

TCP – Tumour control probability.

TG 132 – Report of the AAPM Radiation Therapy Committee Task Group No. 132 (Brock et al – reference 20, Chapter 4). Consensus guidelines on how image registration tools should be trained, validated and assessed.

TN – Thyroid notch (anatomical landmark).

TPS – Treatment planning system – software for calculating a radiotherapy delivery plan

TRE – Target registration error. A metric for interrogating the quality of image registration, using relative positions of known anatomical landmarks.

TRIPOD – Transparent Reporting of Multivariable Prediction Model for Individual Prognosis or Diagnosis – a consensus report aimed at improving the quality and transparency of predictive models in medicine.

UMCG – University Medical Centre Groningen.

UNI - Unilateral neck irradiation. Radiotherapy treatment for head and neck cancer where lymph node regions on one side of the neck only are treated.

VMAT – Volumetric modulated arc therapy. A technique for delivering IMRT by continuous delivery of radiation dose as the linear accelerator rotates.

WL – Weight loss

Xerostomia – Persistent dry mouth, a common (and often severe) side effect of radiotherapy for HNC.

95% Surface Distance - A measure of contour agreement.

Chapter 1 – Introduction and literature review

1.1 Background – head & neck cancer

The term Head and Neck Cancer (HNC) describes a group of biologically similar neoplasms arising from the epithelium and salivary gland parenchyma of the upper aero-digestive tract. Over 90% of the tumours encompassed by this definition are squamous cell carcinomas (SCC's) arising from the epithelium of the lip, oral cavity, nasal cavity, pharynx and larynx [1]. According to data from the Office of National Statistics, combined HNC incidence in England in 2014 (excluding neoplasms of the nasopharynx and nasal cavity which were not included) was just under 9000 [2]. Depending on the year and the source, HNC is between the 6th and 8th commonest cancer in the UK, and the 6th commonest worldwide [2-5].

Overall, HNC incidence in the UK is rising, although this trend is not reflected across all sub-sites within the umbrella term of HNC. Based on data from 2013, age-standardised incidence rates for laryngeal cancer have in fact dropped by 7% since 2003 and 20% since 1993 [6]. In contrast, data for oral cancer – used in this context to include those of the oral cavity, hypo and oropharynx – show a 28% and 42% respective rise over the same period [7]. This difference is due to the fact that the underlying aetiology, biology, demographics and natural history of HNC are evolving.

It is well known that the consumption of both tobacco and alcohol, independently and in combination, are risk factors for the development of HNC, and this effect has been reflected in the demographic spread of the condition [8-10]. The decline in smoking may explain why the incidence of some sub-types of HNC (laryngeal carcinoma for example) is also falling [11]. However, it has become clear over recent decades that some HNC is caused by Human Papillomavirus (HPV) [12-14]. These tumours are predominantly found in the oropharynx and their incidence is rising, explaining the overall increase in HNC cases [15, 16]. It is also clear that HPV driven cancers are biologically separate entities from non-HPV tumours, with different biological drivers and patterns of gene dysfunction/mutation [5, 17, 18]. The demographics and natural history of HPV driven disease are different; on average these patients are younger and more affluent, and they have substantially better outcomes [16, 19-22].

This clear difference in biology and prognosis is reflected in the most recent TNM staging system for HNC (version 8), in which HPV positive and negative disease is staged differently [23]. Furthermore, this biomarker, along with the patients smoking status, drive a robust and extensively validated risk stratification system [22], which classifies patients as having low, intermediate or high risk disease. For low risk patients, a clear priority in the H&N oncology

community is the reduction of long-term toxicity, and this is reflected in many recent and on-going clinical trials [24-26].

This observation has important implications for cancer survivorship. HNC is a loco-regional disease; approximately half of patients will present with advanced local disease but it is unusual to find distant metastasis at presentation [27]. Overall 5-year survival for HNC in the US is 64%, whilst English data from a similar timeframe reports 5-year survival between 50 and 70% depending on anatomical sub-site [28]. Both sources suggest that survival is slowly improving; a possible reason being the changing biology already discussed. Partly because disease is often loco-regional at presentation, radiotherapy (RT) and surgery are key treatment modalities for this condition, and whilst there have been major advances with both; significant long-term morbidity remains a problem. The SEER data report that there were over 300,000 living with oral cavity and pharynx cancer in the US in 2013 [27]. The concern therefore is that as the biology and natural history of the condition evolve, high cure rates from intensive therapy will result in growing numbers of cancer survivors with lasting morbidity and impaired quality of life.

1.2 Radiotherapy for head & neck cancer

Whilst RT had been used to treat HNC for decades prior to this, the seminal 1991 Veterans Affairs study of induction chemotherapy followed by definitive radiotherapy or surgery for advanced laryngeal cancer caused a paradigm shift in the approach to HNC [29]. SCCs, including those of the head and neck, were known to be relatively radiosensitive [30], but this study provided evidence to support 'organ-sparing' RT for more advanced cases in addition to smaller lesions. Radiotherapy, often delivered with concomitant chemotherapy, is now established as the treatment of choice for SCC of all areas of the pharynx and is widely used, as a single modality or in combination with surgery for disease of the larynx, oral cavity and para-nasal sinuses.

Although cure rates are improving, when disease does relapse, it is often loco-regional, and is catastrophic for patients. Despite promising recent advances such as immunotherapy [31], treatment options for relapsed disease are often limited, and symptoms can be very difficult to palliate. Conversely, lifelong late side effects including xerostomia, dysphagia and osteonecrosis following radical RT for HNC are common [32-34]. Thus the concept of therapeutic ratio is critical in the approach to RT for HNC. Much research is focused on trying to improve this ratio, and many treatment strategies in clinical use aim to exploit benefits where

they exist [35]. One approach is to reduce radiotherapy fraction size, in order to increase total dose – hyperfractionating – and this has shown considerable clinical benefit [36, 37]. HNSCCs are often rapidly proliferative, thus accelerating RT by treating twice daily or over weekends has also been tested in clinical trials, again with some success [37-40].

However, the most widespread strategy for augmenting the efficacy of radiation to enhance the therapeutic ratio is to use concomitant cytotoxic chemotherapy, and this approach is supported by high quality meta-analysis [41, 42]. A number of agents have been tried and are still used, including 5FU [43-50], taxanes [49-52] and carboplatin [45-47]. However, the agent with the broadest evidence base, and which is standard of care in the UK and elsewhere, is cisplatin [24, 25, 48, 53-56]. Another drug that has been successfully used as a radiosensitiser alongside radiation protocols for HNC is the anti-egfr mono-clonal antibody cetuximab [57].

1.3 Radiotherapy toxicity

Whilst all of these approaches improve disease related outcomes, they also increase toxicity. Toxicity occurs because normal tissue structures in the vicinity of the tumour also receive radiation dose, and the 'tolerance' of these normal tissues limits the dose that can be prescribed and delivered [58]. Partly because of this, radiotherapy planning for HNC is complex [59-62]. There are multiple anatomical locations in which primary disease may arise and there are subtle differences in approach and technique for each of these sub-sites.

Loco-regional lymph node involvement at presentation is common; indeed 'neck lump' is often the presenting complaint for many patients with HNC [63]. But for those patients with a clinical and radiological N0 neck at presentation, there is still a risk of occult metastasis. This risk depends on factors such as disease sub-site, T-stage and biology, and has been well documented by surgical series [64-66]. Radiotherapy treatment protocols have evolved to reflect this risk [59], and once workup and staging is complete, the treating oncologist will decide whether one or both sides of the neck will require treatment, which lymph node levels and to what dose. However, the concept of therapeutic ratio still applies, as important organs at risk lie adjacent to these lymph node levels.

Permanent xerostomia is a common consequence of RT for HNC, and is caused by radiation damage to salivary glands [67]. There are 3 'major' human salivary glands – the parotid, submandibular and sublingual – and together they produce around 1 litre of saliva per day [68].

The parotids contribute 60-65% of total saliva production, with 20-30% coming from the submandibular and 2-5% from the sublingual glands. There are also minor salivary glands distributed throughout the oral cavity and pharynx, which vary between individuals. Parotid secretions are watery and serous, those of the submandibular and sublingual glands are mixed serous and mucous and those of the minor glands are predominantly mucous resulting in thick, sticky secretions.

The parotid gland is the largest salivary gland and is a paired structure that lies wedged between the sternocleidomastoid and masseter muscles. Each gland is typically 20-30cm^3, and they are largely responsible for stimulated saliva production [69, 70]. Importantly, they lie immediately adjacent to level II lymph nodes, which are frequently included in the Clinical Target Volume (CTV) [59]. There is clear evidence that radiation dose to this structure is associated with persistent xerostomia, and published guidance suggests that limiting one parotid gland to 20Gy, or keeping both below 25Gy, keeps the risk of severe xerostomia acceptably low [71].

The submandibular glands are found in level IB in the anterior triangle of the neck [59]. They are smaller than the parotid glands - typically 10-15 cm^3 - and contribute up to 90% of resting saliva production [69, 70]. They are often very close to tumours arising from the oral cavity and oropharynx, and effective sparing of this organ can therefore be difficult. However, data suggests that doing so has an effect on both un-stimulated and stimulated saliva production [70, 72], and dose to the submandibular gland appears to be the critical factor in the development of permanently thick, sticky saliva [67]. Tolerance doses of 35-40Gy have been quoted as RT planning objectives [70-72].

The pharyngeal constrictors are a group of more than 30 small muscles that surround the posterior and lateral aspects of the pharynx [73]. They coordinate the complex task of safe and efficient swallowing. The consequences of damage can be severe, with dysphagia, persistent aspiration, pneumonia, permanent feeding tubes and resultant reduction in quality of life as possible sequelae [74-79]. Their role as an important OAR was clearly articulated a little over a decade ago, in a study that also coined the term dysphagia/aspiration-related structure or ‘DARS’ [80]. This term encompasses the concept of anatomical structures involved in the process of swallowing that can be identified by and contoured on radiotherapy planning imaging, and it has evolved as studies have shown how dose to individual anatomical subunits can lead to loss of function. Dose to individual constrictor muscles (superior, middle and inferior) has been found to be an important predictor of outcome by a number of authors [81-

85]. The same is true for the supraglottic larynx, cricopharyngeus muscle, and cervical oesophagus [85-88].

Following this work, more recent guidelines have encouraged individual segmentation of the different substructures involved in swallowing [60, 89], to encourage contouring concordance within the community, and permit calculation of dose to each substructure. This includes the superior, middle and inferior pharyngeal constrictor muscles, cricopharyngeus muscle, cervical oesophagus, and supraglottic larynx [90]. Efforts have also been made to investigate whether reducing dose to these structures can reduce toxicity. Two UK-based clinical trials, PATHOS and DARS, were conceived to address this hypothesis, albeit in slightly different patient cohorts [26, 91].

.

Another crucial OAR when planning radiotherapy for HNC is the spinal cord. High doses of radiation to the spinal cord may lead to catastrophic transverse myelitis, whereby patients can lose motor function, sensory discrimination and bowel or bladder control [92]. Historical series reviewing patients treated with old-fashioned radiotherapy techniques do report a low risk of this toxicity in patients undergoing RT for HNC [93, 94], although the risk rises sharply above doses of 54-55Gy [94, 95]. In contemporary practice, the incidence is vanishingly low as radiotherapy teams are so careful to spare the spinal cord, both in the planning and delivery stages of the process. Another side effect of radiation dose to the spinal cord is L'Hermitte's syndrome (LS), characterised by a sensation of electric shocks travelling down the cervical spine and into limbs. The condition is transient and self-limiting, typically lasting 3-6 months [96]. No clear association between LS and progressive irreversible myelitis has been demonstrated, but there are isolated reports of LS preceding delayed radiation myelopathy and paralysis [97]. Furthermore, this condition can be both unpleasant and distressing for patients.

1.4 Intensity-modulated radiotherapy – IMRT

It has been shown that lasting morbidity secondary to radical RT for HNC is a direct consequence of radiation dose to important OAR's [75]. Therefore reducing dose to these structures should in turn reduce the probability of toxicity and improve the therapeutic ratio. Intensity modulated radiotherapy (IMRT) uses an inverse planning algorithm to generate complex arrangements of beams with heterogeneous fluence that deliver non-uniform levels of dose, and concave dose distributions that may be accurately conformed to structures of

interest [98, 99]. In this way, a high dose of radiation may be delivered to the tumour, whilst nearby OAR's are relatively spared [98, 100].

A simple approach to IMRT is 'forward planning', where iterations to the patterns of beam direction and fluence are dictated by the human operator. Such techniques are often used for planning breast and prostate cancer treatments [101, 102]. However, most contemporary IMRT, and the vast majority used in radiotherapy for HNC, uses an 'inverse planning' technique. Here, iterations are undertaken by the treatment planning software itself, guided by human defined constraints and objectives and an optimisation function. IMRT can be delivered with a set, odd number of fixed fields – usually 5 or 7 – in which multi-leaf collimators in the head of the linear-accelerator create heterogeneous fluence patterns whilst the beam is on [103]. More common in contemporary practice is to use a rotational technique in which the beam can be on and delivering treatment from a full 360 degrees as the gantry rotates around the patient [104]. Rotational IMRT delivered by standard linear accelerators is referred to as volumetric arc therapy (VMAT) [105], but is can also be delivered with alternative platforms such as TomoTherapy [106].

The uptake of IMRT in the UK was somewhat sluggish compared to both Europe and the US, although this improved following the introduction of a national strategy, and a change in funding structures [107]. There is now an evidence base showing that IMRT can reduce dose and permit dose escalation across a range of tumour sites [108, 109]. In the context of HNC, the landmark PARSPORT trial showed that IMRT dramatically improves a key clinical outcome [110]. In this study, the incidence of Grade 2 or worse xerostomia at 2 years was 29% in patients who underwent IMRT, compared with 83% in those who received conventional radiotherapy. In addition, there are further dosimetric studies demonstrating that IMRT can reduce dose to important OAR's [70, 80, 111].

1.5 PTV margins and image-guided radiotherapy – IGRT

A key principle of safe and effective radiotherapy delivery is that the patient is in the same position for treatment delivery as they were when that treatment was planned, and that this position is the same every day of treatment. This is even more critical for IMRT. By definition, these treatment plans have heterogeneous dose distributions, with high doses (to the tumour) and steep dose gradients towards lower OAR doses. Inaccuracies in patient setup may

therefore have 2 concerning consequences; namely low dose to the tumour, or excessive dose to normal tissue structures.

A number of strategies exist to mitigate and reduce these risks. To minimise setup inaccuracies, patients undergo their planning scan and subsequent treatment in a thermoplastic mask (or shell), which is comfortably but tightly fitted to minimise movement in all 6 Cartesian planes. Shells that come down over both shoulders, with 5 points of fixation, are standard practice in head and neck radiotherapy. Another important tool is the use of a margin from clinical to planning target volumes (CTV to PTV). This margin, typically 3-5mm for HNC cases [112], allows for the fact that the clinically defined target volume may be in a slightly different position on the treatment machine from day to day. There is an important cost to consider when choosing a PTV margin however, as the volume of the pertinent target will increase by the 3rd power of the margin size chosen. It is therefore a subject of much consideration and debate, and crucial work has been done to agree upon a standardised, intelligent approach to choosing PTV margins [113].

In recent years, another efficacious strategy has become increasingly widely used – image guidance. Modern systems use the linear accelerator itself to generate computed tomography (CT) imaging that shows soft tissue structures in addition to bone [114]. On most standard machines, this is achieved via kV energy X-Rays deployed as a cone beam – cone beam CT (CBCT) [115]. On the TomoTherapy system, the energy of the treatment beam is turned down, and a fan-beam of MV energy X-Rays generates the image [106]. Many patients with HNC at Addenbrookes hospital are treated with the TomoTherapy system, the technology is integral to the work of this dissertation, and is described in more detail in Chapter 2.

Image-guided radiotherapy (IGRT) is now in widespread use, and recent NHS guidelines suggest that every patient should have a form of IGRT as part of their RT treatment [116]. Although the image quality of IG-CT is substantially inferior to standard diagnostic CT, they do permit superior registration between the RT planning scan and on-set imaging than that possible with bone matching only [117]. Once setup is complete, images of the patient in the treatment position are reviewed by treating radiographers before delivering that day's fraction, and subtle adjustments of the couch position can be made if necessary. This technique confers greater certainty of the accuracy of patient's set-up, which in turn gives scope for reduced PTV margins and reduced toxicity [118].

A further advance in this field is the development of technology to combine MRI and a linear accelerator into one machine – MR-linac [119]. Although the physics of this combination

present significant technical challenges [120-123], MR-linac has a number of significant advantages over standard IG, and presents a range of opportunities for the research community [124, 125]. The lack of any ionising radiation with MRI is one clear advantage, as is the superior image quality and soft tissue resolution that may be achieved [126]. There is considerable interest in using the range of image sequences that MR-linac can produce to search for predictive biomarkers of outcome [126, 127], and optimism that these machines will facilitate superior gating and tracking of intra-fraction motion [128, 129], as well as on and off-line adaptive radiotherapy workflows [130, 131].

1.6 Differences between planned and delivered dose

Image guidance can account for some of the differences that may arise between daily setup positions, but is unable to account for the random and systematic daily variations in soft tissue anatomy that occur. Many such changes can and do occur in patients undergoing RT for HNC. SCCs often respond quickly to treatment, and bulky disease may shrink rapidly during treatment [132, 133]. HNC patients often lose a substantial amount of weight during treatment, with possible consequences for internal anatomy [134]. Perhaps most relevant of all, important OAR's may change in size, shape and even position over a course of treatment. A number of studies show that the parotid glands shrink, but also that the centre of gravity translates medially and inferiorly [132, 135-138]. The submandibular glands also shrink, and move superiorly [137].

Given that the radiation fluence delivered by the treatment machine does not significantly change from day-to-day, internal anatomical changes mean that the dose of radiation delivered to each voxel within a given structure is likely to be different from that predicted by the RT plan. In theory therefore there is a difference between planned dose (D_P), and delivered dose (D_A). Preliminary data suggests that this hypothesis is correct [136, 138-140], although the effect appears to be independent of radiotherapy delivery technique [141]. These studies have been small, and have generally used manual contouring to define structures of interest on IG scans to calculate D_A. However, such differences may be highly relevant, as tumour control probability (TCP) and normal tissue complication probability (NTCP) curves are sigmoidal, steep and have all been generated from planned dose data [67, 84, 142, 143].

1.7 Adaptive radiotherapy (ART) for head & neck cancer

1.7.1 ART concepts

A greater understanding of differences between planned and delivered dose to amend and improve RT plans *during a course of treatment*, to improve outcomes for patients, is a clear priority for the radiotherapy community [144]. This field of research is generally referred to as Adaptive Radiotherapy (ART) [145]. A recent review neatly summarises ART as "a formal approach to correct for daily tumour and normal tissue variations through streamlined online or offline modification of original target volumes and plans" [146]. The problem is that generating radiotherapy plans is a laborious, time-consuming process, requiring detailed input from doctors, radiographers and physicists. Furthermore, most current workflows require the patient to have a new CT scan and a new treatment plan, with resultant problems of time, inconvenience and radiation exposure for the patient, and resource implications for departments. Thus implementing ART is likely to be challenging with regards to resources, logistics and timing and it is becoming increasingly apparent that not every patient will benefit from ART [146].

This problem is compounded by the fact that there are a number of subtly different reasons for undertaking ART, and techniques for delivering it. A recent overview has summarised these differences, defining 5 discrete subtypes of ART with attendant nomenclature [147]. These subtypes are described in Table 1, which is adapted from Table 3 in the publication of Heukelom and Fuller [147]. However, in all 5 subtypes, there is either an explicit or implicit *primary* objective. Therefore, from the patient's perspective, and potentially also when considering primary endpoints for clinical trials, it could be argued to further concatenate these 5 subtypes into 2 discrete approaches:

1. Aiming to improve cure rates without increasing treatment toxicity; toxicity reduction would be a bonus – the iso-toxic approach.
2. Aiming to reduce toxicity whilst maintaining cure rates; improving cure rates would be a bonus – the iso-effective approach.

Subtype	Rationale & technique	Tumour dose	OAR dose
ART_{ex_aequo}	Dose accumulation to ensure adequate coverage of target volumes. Replanning where necessary to ensure planning dose parameters delivered	Maintained	Maintained (or reduced)
ART_{OAR}	Similar to ART_{ex_aequo}, but efforts to improve OAR sparing at treatment replan	Maintained	Reduced
ART_{amplio}	Dose to the tumour is increased at replanning. Dose to OARs is maintained, or improved if possible	Escalated	Maintained
ART_{reduco}	Targeted at cases with (imaging) evidence of response: for these patients target volumes shrink with regressing disease, permitting superior OAR sparing.	Maintained	Reduced
ART_{totale}	As ART_{reduc0}, but dose to residual disease is escalated	Escalated	Reduced

Table 1.1: Five Subtypes of, and approaches to Adaptive Radiotherapy, as defined by Heukelom and Fuller [147].

1.7.2 Iso-toxic Adaptive Radiotherapy

For the iso-toxic approach, it is important to understand patterns of relapse in HNC, in order to design better treatment strategies. The risk of distant metastases in HNC patients was classically thought to be low relative to many other carcinomas [27, 148, 149], a pattern that has been maintained in the era of HPV-driven disease [150]. Thus the prime importance of loco-regional control is a crucial paradigm in the management of HNC [39, 40, 151-154], and the point is magnified by the theoretical benefit that even a small increase in tumour dose can bring. Like NTCP curves, TCP curves are steep [142]. They are steepest at a TCP of 50%, and the gradient may be well approximated by a straight-line tangent at this point. From the tangent, an estimate of the predicted increase in TCP resulting from a 1% dose increase can be made – the Gamma-50 (γ_{50}) factor [155]. Trial data in HNC suggested a γ_{50} of 1.33 in HNC [36, 155], meaning that a 10% increase in dose would theoretically result in a 13.3% increase in cure rate.

Studies suggest that long-term primary site control is 70-92%, whilst relapse in involved regional nodal regions occurs in 5-15% of patients [54, 151, 152, 156-162]. Marginal recurrence is rare, as is relapse in nodal regions not included in the original treatment plan

[156-158, 160]. The data from these studies show that the majority of loco-regional relapse following RT for HNC occurs within the high-dose region of the radiotherapy treatment plan. The findings are based upon planned imaging and dose data. Therefore, they may not fully account for anatomical changes that occur during treatment. However, a study using a deformable image registration (DIR) algorithm to try and account for anatomical changes appear to confirm these findings and reinforce the conclusion that the majority of relapses occur within the high dose volume [163].

These findings imply that loco-regional relapse is predominantly a biological issue, rather than a physical, computational or technical one. Problems such as inherently radio-resistant cancer stem cells, and the hypoxic conditions that are often found within primary tumours and involved nodes are well documented [164-168], and may underlie many of the loco-regional recurrences seen in the clinic. Novel imaging techniques such as Positron Emission Tomography (PET) and functional MRI have therefore generated significant interest in this space [169-176], two recent overviews summarise much of the recent research in this field [177, 178], and further studies to establish how these images may guide safe and effective dose adaptation (both escalation and reduction) during treatment are well underway [179-182].

The notion of iso-toxic ART is slightly complicated by the fact that early iterations of the approach postulated that anatomical change during treatment might lead to systematic under-dosing of target volumes in a cohort of patients, thus jeopardising local control [183, 184]. Therefore, ART in this context aimed to *maintain* dose to the target, with the implicit assumption that not doing so would reduce dose – this is the concept of ART_{ex_aequo} as defined by Heukelom and Fuller [147].

A recent study used weekly re-CT to guide adaptation using exactly this principle, and found that failure to undertake ART would have resulted in target volume under-dose in 76% of patients [185]. Another paper assessed the same question in the context of different PTV margins [186]. By accumulating dose using daily IG scans, this study specifically addressed whether reducing PTV margins would increase the risk of clinically significant under-dosing of individual target volumes, and whether ART could reduce the risk in those patients where this occurred. In contrast to the work of Castelli and colleagues [185], the authors found that under-dosing of CTV's was surprisingly unusual with reduced margins, and that in cases in which it did occur, this could be managed by adaptive replanning [186]. On the basis of this work, this centre reduced its PTV margins from 5mm to 3mm [187]. It is worth noting however, that this group have also shown that adapting GTV volumes mid-treatment is a potentially risky enterprise, because of the ways in which tumours can shrink or 'dissolve' [188].

1.7.3 Iso-effective Adaptive Radiotherapy

However, the majority of recent and current work on ART for HNC patients has focussed on reducing OAR dose and side effects – the iso-effective approach [147]. The results of many of these studies have been summarised in a systematic review [146]. Two more recent overviews cover many of these papers, as well as newer studies [147, 189]. Many studies cited in these reviews looked at a range of OARs, including the parotid and submandibular glands, spinal cord and pharyngeal constrictor muscles. Many also made measurements of anatomical change, and endeavoured to link these changes to both dosimetric endpoints, and adaptive practices [190].

The parotid glands have been the most widely studied OAR in the H&N ART literature. Brouwer's review found a mean volume reduction of 26% in this structure over a course of treatment [146], although only 5 studies report data from cohorts of more than 50 patients [191-195]. One of these papers used weekly CT to study volume change, and found that the majority of the gland volume was lost in the first half of treatment [193]. A number of statistically significant associations have been found between different parameters, and parotid gland volume loss. These include weight loss [191, 193, 196-200], dose to the parotid gland [137, 191, 193-195, 197, 201], and parotid gland density [192]. It is interesting to speculate however, that some of these associations may not be explanatory. For example, mean weight loss in HNC patients during treatment can be >10%, and >90% of patients lose weight [134]. Mean parotid gland volume loss is 26% (as stated), and gland volume loss is virtually ubiquitous. Therefore, if almost all patients are losing weight, and almost all patients have shrinking parotid glands, it is not surprising that there is a correlation between these two metrics.

The review from Brouwer et al. found 24 papers that studied differences between planned and delivered dose to the parotid glands [135, 136, 140, 196, 198, 200, 202-219], and report an average (mean) increase in mean dose of 2.2 ± 2.6Gy [146]. Work has also been done to try and identify factors that predict dose differences. The greatest reported parotid gland dose differences appear to be in patients undergoing RT for primary nasopharyngeal carcinoma [207, 208]. Univariate relationships between anatomical change metrics including reducing neck diameter [185, 196, 211], weight loss [140, 196, 200], parotid gland planned dose [218], GTV volume [216], changes in parotid gland volume and position [140, 196, 204], and parotid gland dose differences have all been reported. In a subsequent paper from Brouwer et al, the authors attempted to build a multivariate model to predict patients in whom mean parotid gland D_A would be >3Gy more than D_P [190]. Although a number of factors were statistically

significant on univariate analysis, the only parameter that remained significant on multivariate analysis was mean D_P to the parotid gland.

There are fewer published data describing anatomical changes, and dose differences to other H&N OARs. Three papers measure submandibular gland volume loss [137, 194, 220], with a mean reduction of 22% [146]. Only one paper reports dose differences to the SMGs, with a mean dose difference of 0.9Gy (D_A 52.8Gy vs D_P 51.9Gy) [204]. Two studies have analysed conformational change in the pharyngeal constrictor muscles during treatment, although neither estimated dosimetric consequences [221, 222]. One study looked at dose differences to the oral cavity [204], and there are reported data of dose differences to the whole larynx [204, 206, 209, 219], although not the supraglottic larynx as specified by current OAR segmentation guidelines [60]. There is more published data on dose differences to the spinal cord and brainstem [135, 196, 198, 200, 204, 206-211, 214, 219], although many of these studies assessed small patient cohorts, and imaging from limited timepoints during treatment.

1.7.4 ART – clinical endpoint data

There are limited data linking anatomical change and dosimetric differences. Three studies have found significant predictors of acute toxicity [195, 196, 223], although the endpoints measured are unlikely to alter clinical practice. Others have focussed on late toxicity endpoints; some by modelling NTCP changes based on dose differences [216, 224], some by measuring clinical outcomes per-se. In this work, parotid gland shrinkage during treatment was found to correlate with more pronounced reduction in saliva flow [225], worse subjective xerostomia [226, 227], and longer duration of PEG tube use [199]. A positive association between parotid gland density change - as measured by CT - and xerostomia has also been found [226]. With regards to dosimetric biomarkers, one group assessed relationships between both planned and delivered dose metrics and xerostomia, and found both to be highly predictive [218], but were unable to make a direct comparison between the two due to their relatively small sample size.

MRI has also been used to predict damage to OAR's and late toxicity endpoints, particularly xerostomia [228-233]. However, the link between imaging parameters and strategies for adapting RT plans is less explicit in this work than in research focussing on the tumour. Furthermore, most of this research compares images from before and after RT. Few studies have looked at either the parotid or submandibular glands during treatment, when validated

imaging biomarkers could potentially guide improved OAR sparing, perhaps using the ART_{OAR} approach as outlined by Heukelom and Fuller [147].

Clinical studies reporting outcomes of the implementation of ART workflows in practice are even scarcer. Two papers report the outcomes of single arm studies, in which there were no comparator groups [199, 234]. Schwartz and colleagues undertook automated adaptive replanning on 22 patients with oropharyngeal carcinoma and found promising loco-regional control results, and toxicity data comparable to that reported in IMRT-era literature [199]. Kataria et al used a similar study design with 36 patients with locally advanced HNC patients, with similar results [234].

There are 3 non-randomised cohort studies, perhaps the best known of which comes from Chen and colleagues [235]. In this study, 317 HNC patients were recruited, and 51 underwent ART on an ad-hoc basis, with the remainder constituting the comparator arm. The authors found equivalent toxicity in the 2 groups, but superior 2-year loco-regional control in the ART (88% vs 79%). Yang et al. undertook a case-control study in which 86 patients underwent ART, whilst 43 received standard IMRT [236]. Loco-regional control was again superior (97.2% vs 82.2%) and a number of QoL metrics were found to be better in the ART group. Zhao and colleagues used a similar design to treat 33 patients with ART and 66 with standard IMRT [203]. In this study, local control benefits for patients with T3+ disease, and reduced side effects for patients with N2+ disease were found for the ART group. However, the results of all 3 studies should be treated with significant caution given their small size, non-randomised designs, a significant risk of confounding variables, and a high probability of publication bias.

Perhaps the most convincing clinical data to support ART for HNC patients to date comes from Navran and colleagues [187]. In this study, which followed directly from the dosimetric planning study from the same centre cited in section 1.7.2 [186], the authors audited treatment outcomes for over 200 patients immediately before and after a change in PTV margin from 5mm to 3mm [187]. Crucially, this group had clearly demonstrated that this margin reduction did not risk under-dosing target volumes with this reduced margin because of their adaptive workflow [186]. They found that loco-regional control before and after the change was identical, but with reduced acute toxicities, and lower rates of tube-feeding dependence following radiotherapy [187].

1.7.5 ART – known unknowns

Overall therefore it can be seen that ART is an attractive strategy in HNC, and that a lot of promising work has been done to identify imaging biomarkers to guide this process. However, more data is needed to guide an intelligent approach as to who should have their radiotherapy plan adapted during treatment [146, 204], when and how, and technical developments are required to make this feasible within busy clinical departments. Clinical data, either supporting ART as a concept, or directly testing its efficacy, are limited, although there are ongoing clinical trials in this space [237, 238]. Crucially, there is as yet no published data demonstrating that clinical outcomes correlate better with delivered dose, than with planned.

It is now widely accepted that ART as a strategy may only benefit a subset of patients [146, 204], and although some work has been done to provide criteria for patient selection [190, 239, 240], there are as yet no clear guidelines by which H&N clinicians can identify patients for ART [147]. This knowledge gap is compounded by another technical development in radiotherapy – Proton Beam Therapy (PBT). The nature of proton physics means that uncertainties about the type and quantity of tissue that a proton must travel through as it traverses a patient can have a much more significant impact on the treatment plan dosimetry than is typically seen with X-Ray based treatment [241-243]. Therefore, the significant anatomical changes experienced by HNC patients may be even more problematic for proton plans, than they are for IMRT [244-246]. As the UK ramps up its own PBT centres, and seeks to provide an evidence base beyond the current list of approved indications [247], such considerations may be important.

Therefore, clear priorities for radiotherapists treating patients with HNC are to better understand the nature of differences between planned and delivered dose, to try and demonstrate that outcomes are better predicted by delivered dose, and to find and validate robust biomarkers by which patients may be intelligently selected for ART.

1.7.6 Research objectives of this dissertation

Based on the knowledge of the field, recent advances, and current problems outlined in the previous sections, the research objectives for this work were set out as follows:

- To quantify both planned and delivered radiotherapy dose to critical dose-limiting, and other important OARs in head and neck cancer patients, as accurately as possible, using daily image guidance scans, in a systematic way, in a large cohort.
- To assess statistical relationships between both planned and delivered dose, and the toxicity experienced by patients to directly test the hypothesis that delivered dose will be a stronger predictor of outcomes.
- To test, and externally validate previously reported patient, and treatment-related factors of dose differences to OARs, and to search for new ones.
- To replicate previously reported post-treatment MRI biomarkers of toxicity, and to look for such IMBs early in the treatment course as a tool for guiding ART.

1.8 Brief introduction to the VoxTox study

The VoxTox programme is an inter-disciplinary collaboration at the University of Cambridge, with a research group centred on the Department of Oncology. Its fundamental objective was to characterise relationships between planned and delivered dose at the voxel level (Vox), and resultant toxicity (Tox) in patients treated for prostate, head and neck, and CNS tumours, through careful prospective collection and curation of toxicity data and image guidance scans. It was funded via a Programme Grant from Cancer Research UK, study recruitment ran from April 2013 to April 2019, and all the work in the dissertation was undertaken within the broader framework of this research infrastructure. Details of the VoxTox study, including how it was used to provide the dataset from which the questions and objectives set out in section 1.7.6 could be addressed, are discussed in detail in Chapter 2.

1.9 References:

[1] Gregoire V, Lefebvre JL, Licitra L, Felip E, Group E-E-EGW. Squamous cell carcinoma of the head and neck: EHNS-ESMO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2010;21 Suppl 5:v184-6.
[2] OFNS. Cancer Registration Statistics, England: 2014. Cancer diagnoses and age-standardised incidence rates for all cancer sites by age, sex and region. . 2014.
[3] Hunter KD, Parkinson EK, Harrison PR. Profiling early head and neck cancer. Nat Rev Cancer. 2005;5:127-35.
[4] Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer. 2013;49:1374-403.
[5] Cleary C, Leeman JE, Higginson DS, Katabi N, Sherman E, Morris L, et al. Biological Features of Human Papillomavirus-related Head and Neck Cancers Contributing to Improved Response. Clin Oncol (R Coll Radiol). 2016;28:467-74.
[6] Cancer Research UK. Larynx cancer statistics, 2013. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/larynx-cancer. Accessed June 21, 2019.
[7] Cancer Research UK. Oral cancer statistics, 2013. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oral-cancer. Accessed June 21, 2019.
[8] Wynder EL, Stellman SD. Comparative epidemiology of tobacco-related cancers. Cancer Res. 1977;37:4608-22.
[9] Decker J, Goldstein JC. Risk factors in head and neck cancer. N Engl J Med. 1982;306:1151-5.
[10] Hashibe M, Brennan P, Chuang SC, Boccia S, Castellsague X, Chen C, et al. Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Cancer Epidemiol Biomarkers Prev. 2009;18:541-50.
[11] Sturgis EM, Cinciripini PM. Trends in head and neck cancer incidence in relation to smoking prevalence: an emerging epidemic of human papillomavirus-associated cancers? Cancer. 2007;110:1429-35.
[12] Dillner J, Knekt P, Schiller JT, Hakulinen T. Prospective seroepidemiological evidence that human papillomavirus type 16 infection is a risk factor for oesophageal squamous cell carcinoma. BMJ. 1995;311:1346.
[13] Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst. 2000;92:709-20.
[14] Mork J, Lie AK, Glattre E, Hallmans G, Jellum E, Koskela P, et al. Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. N Engl J Med. 2001;344:1125-31.
[15] Dahlstrand H, Nasman A, Romanitan M, Lindquist D, Ramqvist T, Dalianis T. Human papillomavirus accounts both for increased incidence and better prognosis in tonsillar cancer. Anticancer Res. 2008;28:1133-8.
[16] Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. J Clin Oncol. 2015;33:3235-42.
[17] Vidal L, Gillison ML. Human papillomavirus in HNSCC: recognition of a distinct disease type. Hematol Oncol Clin North Am. 2008;22:1125-42, vii.
[18] Psyrri A, Gouveris P, Vermorken JB. Human papillomavirus-related head and neck tumors: clinical and research implication. Curr Opin Oncol. 2009;21:201-5.
[19] Dahlstrom KR, Bell D, Hanby D, Li G, Wang LE, Wei Q, et al. Socioeconomic characteristics of patients with oropharyngeal carcinoma according to tumor HPV status, patient smoking status, and sexual behavior. Oral Oncol. 2015;51:832-8.
[20] D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, et al. Case-control study of human papillomavirus and oropharyngeal cancer. N Engl J Med. 2007;356:1944-56.
[21] Fakhry C, Westra WH, Li S, Cmelak A, Ridge JA, Pinto H, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst. 2008;100:261-9.
[22] Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med. 2010;363:24-35.
[23] Huang SH, O'Sullivan B. Overview of the 8th Edition TNM Classification for Head and Neck Cancer. Curr Treat Options Oncol. 2017;18:40.

[24] Mehanna H, Robinson M, Hartley A, Kong A, Foran B, Fulton-Lieuw T, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. Lancet. 2019;393:51-60.
[25] Gillison ML, Trotti AM, Harris J, Eisbruch A, Harari PM, Adelstein DJ, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. Lancet. 2019;393:40-50.
[26] Owadally W, Hurt C, Timmins H, Parsons E, Townsend S, Patterson J, 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;15:602.
[27] National Cancer Institute. SEER Cancer Statistics Review (CSR) 1975-2013. https://seer.cancer.gov/archive/csr/1975_2013/browse_csr.php?sectionSEL=20. Accessed June 21, 2019.
[28] Oxford Cancer Intelligence Unit. Profile of Head and Neck Cancers in England. 2010.
[29] Induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. The Department of Veterans Affairs Laryngeal Cancer Study Group. N Engl J Med. 1991;324:1685-90.
[30] Fertil B, Malaise EP. Intrinsic radiosensitivity of human cell lines is correlated with radioresponsiveness of human tumors: analysis of 101 published survival curves. Int J Radiat Oncol Biol Phys. 1985;11:1699-707.
[31] Ferris RL, Blumenschein G, Jr., Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2016;375:1856-67.
[32] Bhide SA, Miah AB, Harrington KJ, Newbold KL, Nutting CM. Radiation-induced xerostomia: pathophysiology, prevention and treatment. Clin Oncol (R Coll Radiol). 2009;21:737-44.
[33] Roe JW, Carding PN, Dwivedi RC, Kazi RA, Rhys-Evans PH, Harrington KJ, et al. Swallowing outcomes following Intensity Modulated Radiation Therapy (IMRT) for head & neck cancer - a systematic review. Oral Oncol. 2010;46:727-33.
[34] Strojan P, Hutcheson KA, Eisbruch A, Beitler JJ, Langendijk JA, Lee AWM, et al. Treatment of late sequelae after radiotherapy for head and neck cancer. Cancer Treat Rev. 2017;59:79-92.
[35] Gregoire V, Langendijk JA, Nuyts S. Advances in Radiotherapy for Head and Neck Cancer. J Clin Oncol. 2015;33:3277-84.
[36] Horiot JC, Le Fur R, N'Guyen T, Chenal C, Schraub S, Alfonsi S, et al. Hyperfractionation versus conventional fractionation in oropharyngeal carcinoma: final analysis of a randomized trial of the EORTC cooperative group of radiotherapy. Radiother Oncol. 1992;25:231-41.
[37] Fu KK, Pajak TF, Trotti A, Jones CU, Spencer SA, Phillips TL, et al. A Radiation Therapy Oncology Group (RTOG) phase III randomized study to compare hyperfractionation and two variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: first report of RTOG 9003. Int J Radiat Oncol Biol Phys. 2000;48:7-16.
[38] Horiot JC, Bontemps P, van den Bogaert W, Le Fur R, van den Weijngaert D, Bolla M, et al. Accelerated fractionation (AF) compared to conventional fractionation (CF) improves loco-regional control in the radiotherapy of advanced head and neck cancers: results of the EORTC 22851 randomized trial. Radiother Oncol. 1997;44:111-21.
[39] Overgaard J, Hansen HS, Specht L, Overgaard M, Grau C, Andersen E, et al. Five compared with six fractions per week of conventional radiotherapy of squamous-cell carcinoma of head and neck: DAHANCA 6 and 7 randomised controlled trial. Lancet. 2003;362:933-40.
[40] Dische S, Saunders M, Barrett A, Harvey A, Gibson D, Parmar M. A randomised multicentre trial of CHART versus conventional radiotherapy in head and neck cancer. Radiother Oncol. 1997;44:123-36.
[41] Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. Lancet. 2000;355:949-55.
[42] Pignon JP, le Maitre A, Maillard E, Bourhis J, Group M-NC. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. Radiother Oncol. 2009;92:4-14.
[43] Wendt TG, Grabenbauer GG, Rodel CM, Thiel HJ, Aydin H, Rohloff R, et al. Simultaneous radiochemotherapy versus radiotherapy alone in advanced head and neck cancer: a randomized multicenter study. J Clin Oncol. 1998;16:1318-24.
[44] Brizel DM, Albers ME, Fisher SR, Scher RL, Richtsmeier WJ, Hars V, et al. Hyperfractionated irradiation with or without concurrent chemotherapy for locally advanced head and neck cancer. N Engl J Med. 1998;338:1798-804.

[45] Calais G, Alfonsi M, Bardet E, Sire C, Germain T, Bergerot P, et al. Randomized trial of radiation therapy versus concomitant chemotherapy and radiation therapy for advanced-stage oropharynx carcinoma. J Natl Cancer Inst. 1999;91:2081-6.
[46] Olmi P, Crispino S, Fallai C, Torri V, Rossi F, Bolner A, et al. Locoregionally advanced carcinoma of the oropharynx: conventional radiotherapy vs. accelerated hyperfractionated radiotherapy vs. concomitant radiotherapy and chemotherapy--a multicenter randomized trial. Int J Radiat Oncol Biol Phys. 2003;55:78-92.
[47] Denis F, Garaud P, Bardet E, Alfonsi M, Sire C, Germain T, et al. Final results of the 94-01 French Head and Neck Oncology and Radiotherapy Group randomized trial comparing radiotherapy alone with concomitant radiochemotherapy in advanced-stage oropharynx carcinoma. J Clin Oncol. 2004;22:69-76.
[48] Adelstein DJ, Li Y, Adams GL, Wagner H, Jr., Kish JA, Ensley JF, et al. An intergroup phase III comparison of standard radiation therapy and two schedules of concurrent chemoradiotherapy in patients with unresectable squamous cell head and neck cancer. J Clin Oncol. 2003;21:92-8.
[49] Kies MS, Haraf DJ, Rosen F, Stenson K, List M, Brockstein B, et al. Concomitant infusional paclitaxel and fluorouracil, oral hydroxyurea, and hyperfractionated radiation for locally advanced squamous head and neck cancer. J Clin Oncol. 2001;19:1961-9.
[50] Garden AS, Harris J, Vokes EE, Forastiere AA, Ridge JA, Jones C, et al. Preliminary results of Radiation Therapy Oncology Group 97-03: a randomized phase ii trial of concurrent radiation and chemotherapy for advanced squamous cell carcinomas of the head and neck. J Clin Oncol. 2004;22:2856-64.
[51] Lovey J, Koronczay K, Remenar E, Csuka O, Nemeth G. Radiotherapy and concurrent low-dose paclitaxel in locally advanced head and neck cancer. Radiother Oncol. 2003;68:171-4.
[52] Calais G, Bardet E, Sire C, Alfonsi M, Bourhis J, Rhein B, et al. Radiotherapy with concomitant weekly docetaxel for Stages III/IV oropharynx carcinoma. Results of the 98-02 GORTEC Phase II trial. Int J Radiat Oncol Biol Phys. 2004;58:161-6.
[53] Huguenin P, Beer KT, Allal A, Rufibach K, Friedli C, Davis JB, et al. Concomitant cisplatin significantly improves locoregional control in advanced head and neck cancers treated with hyperfractionated radiotherapy. J Clin Oncol. 2004;22:4665-73.
[54] Cooper JS, Pajak TF, Forastiere AA, Jacobs J, Campbell BH, Saxman SB, et al. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. N Engl J Med. 2004;350:1937-44.
[55] Bernier J, Domenge C, Ozsahin M, Matuszewska K, Lefebvre JL, Greiner RH, et al. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. N Engl J Med. 2004;350:1945-52.
[56] Forastiere AA, Zhang Q, Weber RS, Maor MH, Goepfert H, Pajak TF, et al. Long-term results of RTOG 91-11: a comparison of three nonsurgical treatment strategies to preserve the larynx in patients with locally advanced larynx cancer. J Clin Oncol. 2013;31:845-52.
[57] Bonner JA, Harari PM, Giralt J, Azarnia N, Shin DM, Cohen RB, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. N Engl J Med. 2006;354:567-78.
[58] Burnet NG, Wurm R, Nyman J, Peacock JH. Normal tissue radiosensitivity--how important is it? Clin Oncol (R Coll Radiol). 1996;8:25-34.
[59] Gregoire V, Ang K, Budach W, Grau C, Hamoir M, Langendijk JA, 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;110:172-81.
[60] Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, 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;117:83-90.
[61] Gregoire V, Evans M, Le QT, Bourhis J, Budach V, Chen A, 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;126:3-24.
[62] Lee AW, Ng WT, Pan JJ, Poh SS, Ahn YC, AlHussain H, et al. International guideline for the delineation of the clinical target volumes (CTV) for nasopharyngeal carcinoma. Radiother Oncol. 2018;126:25-36.
[63] Mehanna H, Paleri V, West CM, Nutting C. Head and neck cancer--Part 1: Epidemiology, presentation, and prevention. BMJ. 2010;341:c4684.

[64] Lindberg R. Distribution of cervical lymph node metastases from squamous cell carcinoma of the upper respiratory and digestive tracts. Cancer. 1972;29:1446-9.
[65] Candela FC, Kothari K, Shah JP. Patterns of cervical node metastases from squamous carcinoma of the oropharynx and hypopharynx. Head Neck. 1990;12:197-203.
[66] Shah JP, Candela FC, Poddar AK. The patterns of cervical lymph node metastases from squamous carcinoma of the oral cavity. Cancer. 1990;66:109-13.
[67] Beetz I, Schilstra C, van der Schaaf A, van den Heuvel ER, Doornaert P, van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. Radiother Oncol. 2012;105:101-6.
[68] Cooper JS, Fu K, Marks J, Silverman S. Late effects of radiation therapy in the head and neck region. Int J Radiat Oncol Biol Phys. 1995;31:1141-64.
[69] Ellis H. Anatomy of the salivary glands. Surgery. 2012;30:569-72.
[70] Saarilahti K, Kouri M, Collan J, Kangasmaki A, Atula T, Joensuu H, et al. Sparing of the submandibular glands by intensity modulated radiotherapy in the treatment of head and neck cancer. Radiother Oncol. 2006;78:270-5.
[71] Deasy JO, Moiseenko V, Marks L, Chao KS, Nam J, Eisbruch A. Radiotherapy dose-volume effects on salivary gland function. Int J Radiat Oncol Biol Phys. 2010;76:S58-63.
[72] Murdoch-Kinch CA, Kim HM, Vineberg KA, Ship JA, Eisbruch A. Dose-effect relationships for the submandibular salivary glands and implications for their sparing by intensity modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2008;72:373-82.
[73] Goldsmith T. Videofluoroscopic evaluation of oropharyngeal swallowing. In: Som PM, Curtin HD, editors. Head and neck imaging. 4th ed. St Louis: Mosby; 2003. p. 1727 - 53.
[74] Eisbruch A, Lyden T, Bradford CR, Dawson LA, Haxer MJ, Miller AE, et al. Objective assessment of swallowing dysfunction and aspiration after radiation concurrent with chemotherapy for head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2002;53:23-8.
[75] Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. J Clin Oncol. 2008;26:3770-6.
[76] Nguyen NP, Frank C, Moltz CC, Vos P, Smith HJ, Karlsson U, et al. Impact of dysphagia on quality of life after treatment of head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2005;61:772-8.
[77] Carrara-de Angelis E, Feher O, Barros AP, Nishimoto IN, Kowalski LP. Voice and swallowing in patients enrolled in a larynx preservation trial. Arch Otolaryngol Head Neck Surg. 2003;129:733-8.
[78] Smith RV, Kotz T, Beitler JJ, Wadler S. Long-term swallowing problems after organ preservation therapy with concomitant radiation therapy and intravenous hydroxyurea: initial results. Arch Otolaryngol Head Neck Surg. 2000;126:384-9.
[79] Nguyen NP, Moltz CC, Frank C, Karlsson U, Nguyen PD, Vos P, et al. Dysphagia severity following chemoradiation and postoperative radiation for head and neck cancer. Eur J Radiol. 2006;59:453-9.
[80] Eisbruch A, Schwartz M, Rasch C, Vineberg K, Damen E, Van As CJ, 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;60:1425-39.
[81] Levendag PC, Teguh DN, Voet P, van der Est H, Noever I, de Kruijf WJ, et al. Dysphagia disorders in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscle: a dose-effect relationship. Radiother Oncol. 2007;85:64-73.
[82] Li B, Li D, Lau DH, Farwell DG, Luu Q, Rocke DM, 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;4:52.
[83] Caudell JJ, Schaner PE, Desmond RA, Meredith RF, Spencer SA, Bonner JA. 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;76:403-9.
[84] Wopken K, Bijl HP, van der Schaaf A, van der Laan HP, Chouvalova O, Steenbakkers RJ, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol. 2014;113:95-101.
[85] Fua TF, Corry J, Milner AD, Cramb J, Walsham SF, Peters LJ. Intensity-modulated radiotherapy for nasopharyngeal carcinoma: clinical correlation of dose to the pharyngo-esophageal axis and dysphagia. Int J Radiat Oncol Biol Phys. 2007;67:976-81.

[86] Dornfeld K, Simmons JR, Karnell L, Karnell M, Funk G, Yao M, 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;68:750-7.
[87] Feng FY, Kim HM, Lyden TH, Haxer MJ, Feng M, Worden FP, 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;68:1289-98.
[88] Dirix P, Abbeel S, Vanstraelen B, Hermans R, Nuyts S. Dysphagia after chemoradiotherapy for head-and-neck squamous cell carcinoma: dose-effect relationships for the swallowing structures. Int J Radiat Oncol Biol Phys. 2009;75:385-92.
[89] Alterio D, Ciardo D, Preda L, Argenone A, Caspiani O, Micera R, 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. Radiother Oncol. 2014;112:337-42.
[90] Christianen ME, Langendijk JA, Westerlaan HE, van de Water TA, Bijl HP. Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. Radiother Oncol. 2011;101:394-402.
[91] Petkar I, Rooney K, Roe JW, Patterson JM, Bernstein D, Tyler JM, 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;16:770.
[92] Schultheiss TE, Kun LE, Ang KK, Stephens LC. Radiation response of the central nervous system. Int J Radiat Oncol Biol Phys. 1995;31:1093-112.
[93] Marcus RB, Jr., Million RR. The incidence of myelitis after irradiation of the cervical spinal cord. Int J Radiat Oncol Biol Phys. 1990;19:3-8.
[94] Jeremic B, Djuric L, Mijatovic L. Incidence of radiation myelitis of the cervical spinal cord at doses of 5500 cGy or greater. Cancer. 1991;68:2138-41.
[95] Kirkpatrick JP, Van-Der-Kogel AJ, Schultheiss TE. Radiation Dose-Volume Effects in the Spinal Cord. Int J Radiation Oncology Biol Phys. 2010;76:S42-9.
[96] Pak D, Vineberg K, Feng F, Ten Haken RK, Eisbruch A. Lhermitte sign after chemo-IMRT of head-and-neck cancer: incidence, doses, and potential mechanisms. Int J Radiat Oncol Biol Phys. 2012;83:1528-33.
[97] Esik O, Csere T, Stefanits K, Lengyel Z, Safrany G, Vonoczky K, et al. A review on radiogenic Lhermitte's sign. Pathol Oncol Res. 2003;9:115-20.
[98] Webb S. The physical basis of IMRT and inverse planning. Br J Radiol. 2003;76:678-89.
[99] Cahlon O, Hunt M, Zelefsky MJ. Intensity-modulated radiation therapy: supportive data for prostate cancer. Semin Radiat Oncol. 2008;18:48-57.
[100] Burnet NG, Adams EJ, Fairfoul J, Tudor GS, Hoole AC, Routsis DS, et al. Practical aspects of implementation of helical tomotherapy for intensity-modulated and image-guided radiotherapy. Clin Oncol (R Coll Radiol). 2010;22:294-312.
[101] Barnett GC, Wilkinson J, Moody AM, Wilson CB, Sharma R, Klager S, et al. A randomised controlled trial of forward-planned radiotherapy (IMRT) for early breast cancer: baseline characteristics and dosimetry results. Radiother Oncol. 2009;92:34-41.
[102] Corletto D, Iori M, Paiusco M, Brait L, Broggi S, Ceresoli G, et al. Inverse and forward optimization of one- and two-dimensional intensity-modulated radiation therapy-based treatment of concave-shaped planning target volumes: the case of prostate cancer. Radiother Oncol. 2003;66:185-95.
[103] Longobardi B, De Martin E, Fiorino C, Dell'oca I, Broggi S, Cattaneo GM, et al. Comparing 3DCRT and inversely optimized IMRT planning for head and neck cancer: equivalence between step-and-shoot and sliding window techniques. Radiother Oncol. 2005;77:148-56.
[104] Wiezorek T, Brachwitz T, Georg D, Blank E, Fotina I, Habl G, et al. Rotational IMRT techniques compared to fixed gantry IMRT and tomotherapy: multi-institutional planning study for head-and-neck cases. Radiat Oncol. 2011;6:20.
[105] Wiehle R, Knippen S, Grosu AL, Bruggmoser G, Hodapp N. VMAT and step-and-shoot IMRT in head and neck cancer: a comparative plan analysis. Strahlenther Onkol. 2011;187:820-5.
[106] Mackie TR, Balog J, Ruchala K, Shepard D, Aldridge S, Fitchard E, et al. Tomotherapy. Semin Radiat Oncol. 1999;9:108-17.
[107] Mayles WP, Cooper T, Mackay R, Staffurth J, Williams M. Progress with Intensity-modulated radiotherapy implementation in the UK. Clin Oncol (R Coll Radiol). 2012;24:543-4.
[108] Staffurth J, Radiotherapy Development B. A review of the clinical evidence for intensity-modulated radiotherapy. Clin Oncol (R Coll Radiol). 2010;22:643-57.

[109] De Neve W, De Gersem W, Madani I. Rational use of intensity-modulated radiation therapy: the importance of clinical outcome. Semin Radiat Oncol. 2012;22:40-9.
[110] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol. 2011;12:127-36.
[111] Clark CH, Bidmead AM, Mubata CD, Harrington KJ, Nutting CM. Intensity-modulated radiotherapy improves target coverage, spinal cord sparing and allows dose escalation in patients with locally advanced cancer of the larynx. Radiother Oncol. 2004;70:189-98.
[112] Astreinidou E, Bel A, Raaijmakers CP, Terhaard CH, Lagendijk JJ. Adequate margins for random setup uncertainties in head-and-neck IMRT. Int J Radiat Oncol Biol Phys. 2005;61:938-44.
[113] van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. Int J Radiat Oncol Biol Phys. 2000;47:1121-35.
[114] Verellen D, De Ridder M, Linthout N, Tournel K, Soete G, Storme G. Innovations in image-guided radiotherapy. Nat Rev Cancer. 2007;7:949-60.
[115] Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. Int J Radiat Oncol Biol Phys. 2002;53:1337-49.
[116] National Cancer Action Team. National Radiotherapy Implementation Group Report Image Guided Radiotherapy (IGRT) Guidance for implementation and use. 2012.
[117] McBain CA, Henry AM, Sykes J, Amer A, Marchant T, Moore CM, et al. X-ray volumetric imaging in image-guided radiotherapy: the new standard in on-treatment imaging. Int J Radiat Oncol Biol Phys. 2006;64:625-34.
[118] Chen AM, Yu Y, Daly ME, Farwell DG, Benedict SH, Purdy JA. Long-term experience with reduced planning target volume margins and intensity-modulated radiotherapy with daily image-guidance for head and neck cancer. Head Neck. 2014;36:1766-72.
[119] Lagendijk JJ, Raaymakers BW, Raaijmakers AJ, Overweg J, Brown KJ, Kerkhof EM, et al. MRI/linac integration. Radiother Oncol. 2008;86:25-9.
[120] Raaijmakers AJ, Raaymakers BW, Lagendijk JJ. Integrating a MRI scanner with a 6 MV radiotherapy accelerator: dose increase at tissue-air interfaces in a lateral magnetic field due to returning electrons. Phys Med Biol. 2005;50:1363-76.
[121] Raaijmakers AJ, Raaymakers BW, van der Meer S, Lagendijk JJ. Integrating a MRI scanner with a 6 MV radiotherapy accelerator: impact of the surface orientation on the entrance and exit dose due to the transverse magnetic field. Phys Med Biol. 2007;52:929-39.
[122] Combs SE, Nusslin F, Wilkens JJ. Individualized radiotherapy by combining high-end irradiation and magnetic resonance imaging. Strahlenther Onkol. 2016;192:209-15.
[123] Schmidt MA, Payne GS. Radiotherapy planning using MRI. Phys Med Biol. 2015;60:R323-61.
[124] Choudhury A, Budgell G, MacKay R, Falk S, Faivre-Finn C, Dubec M, et al. The Future of Image-guided Radiotherapy. Clin Oncol (R Coll Radiol). 2017;29:662-6.
[125] Pollard JM, Wen Z, Sadagopan R, Wang J, Ibbott GS. The future of image-guided radiotherapy will be MR guided. Br J Radiol. 2017;90:20160667.
[126] Kupelian P, Sonke JJ. Magnetic resonance-guided adaptive radiotherapy: a solution to the future. Semin Radiat Oncol. 2014;24:227-32.
[127] Wong KH, Panek R, Bhide SA, Nutting CM, Harrington KJ, Newbold KL. The emerging potential of magnetic resonance imaging in personalizing radiotherapy for head and neck cancer: an oncologist's perspective. Br J Radiol. 2017;90:20160768.
[128] Bjerre T, Crijns S, af Rosenschold PM, Aznar M, Specht L, Larsen R, et al. Three-dimensional MRI-linac intra-fraction guidance using multiple orthogonal cine-MRI planes. Phys Med Biol. 2013;58:4943-50.
[129] Yun J, Wachowicz K, Mackenzie M, Rathee S, Robinson D, Fallone BG. First demonstration of intrafractional tumor-tracked irradiation using 2D phantom MR images on a prototype linac-MR. Med Phys. 2013;40:051718.
[130] Kontaxis C, Bol GH, Lagendijk JJ, Raaymakers BW. Towards adaptive IMRT sequencing for the MR-linac. Phys Med Biol. 2015;60:2493-509.
[131] McPartlin AJ, Li XA, Kershaw LE, Heide U, Kerkmeijer L, Lawton C, et al. MRI-guided prostate adaptive radiotherapy - A systematic review. Radiother Oncol. 2016;119:371-80.
[132] Barker JL, Jr., Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys. 2004;59:960-70.

[133] Geets X, Tomsej M, Lee JA, Duprez T, Coche E, Cosnard G, et al. Adaptive biological image-guided IMRT with anatomic and functional imaging in pharyngo-laryngeal tumors: impact on target volume delineation and dose distribution using helical tomotherapy. Radiother Oncol. 2007;85:105-15.
[134] Ottosson S, Zackrisson B, Kjellen E, Nilsson P, Laurell G. Weight loss in patients with head and neck cancer during and after conventional and accelerated radiotherapy. Acta Oncol. 2013;52:711-8.
[135] Hansen EK, Bucci MK, Quivey JM, Weinberg V, Xia P. Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2006;64:355-62.
[136] Robar JL, Day A, Clancey J, Kelly R, Yewondwossen M, Hollenhorst H, et al. Spatial and dosimetric variability of organs at risk in head-and-neck intensity-modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2007;68:1121-30.
[137] Vasquez Osorio EM, Hoogeman MS, Al-Mamgani A, Teguh DN, Levendag PC, Heijmen BJ. Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration. Int J Radiat Oncol Biol Phys. 2008;70:875-82.
[138] Han C, Chen YJ, Liu A, Schultheiss TE, Wong JY. Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy. Int J Radiat Oncol Biol Phys. 2008;70:1256-62.
[139] Loo H, Fairfoul J, Chakrabarti A, Dean JC, Benson RJ, Jefferies SJ, et al. Tumour shrinkage and contour change during radiotherapy increase the dose to organs at risk but not the target volumes for head and neck cancer patients treated on the TomoTherapy HiArt system. Clin Oncol (R Coll Radiol). 2011;23:40-7.
[140] Lee C, Langen KM, Lu W, Haimerl J, Schnarr E, Ruchala KJ, et al. Assessment of parotid gland dose changes during head and neck cancer radiotherapy using daily megavoltage computed tomography and deformable image registration. Int J Radiat Oncol Biol Phys. 2008;71:1563-71.
[141] Thomson DJ, Beasley WJ, Garcez K, Lee LW, Sykes AJ, Rowbottom CG, et al. Relative plan robustness of step-and-shoot vs rotational intensity-modulated radiotherapy on repeat computed tomographic simulation for weight loss in head and neck cancer. Med Dosim. 2016;41:154-8.
[142] Gronlund E, Johansson S, Montelius A, Ahnesjo A. Dose painting by numbers based on retrospectively determined recurrence probabilities. Radiother Oncol. 2017;122:236-41.
[143] Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys. 2010;76:S10-9.
[144] Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tome WA. Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. Int J Radiat Oncol Biol Phys. 2010;76:S135-9.
[145] Gregoire V, Jeraj R, Lee JA, O'Sullivan B. Radiotherapy for head and neck tumours in 2012 and beyond: conformal, tailored, and adaptive? Lancet Oncol. 2012;13:e292-300.
[146] Brouwer CL, Steenbakkers RJ, Langendijk JA, Sijtsema NM. Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help? Radiother Oncol. 2015;115:285-94.
[147] Heukelom J, Fuller CD. Head and Neck Cancer Adaptive Radiation Therapy (ART): Conceptual Considerations for the Informed Clinician. Semin Radiat Oncol. 2019;29:258-73.
[148] Garavello W, Ciardo A, Spreafico R, Gaini RM. Risk factors for distant metastases in head and neck squamous cell carcinoma. Arch Otolaryngol Head Neck Surg. 2006;132:762-6.
[149] Ferlito A, Shaha AR, Silver CE, Rinaldo A, Mondin V. Incidence and sites of distant metastases from head and neck cancer. ORL J Otorhinolaryngol Relat Spec. 2001;63:202-7.
[150] O'Sullivan B, Huang SH, Siu LL, Waldron J, Zhao H, Perez-Ordonez B, et al. Deintensification candidate subgroups in human papillomavirus-related oropharyngeal cancer according to minimal risk of distant metastasis. J Clin Oncol. 2013;31:543-50.
[151] Riaz N, Hong JC, Sherman EJ, Morris L, Fury M, Ganly I, et al. A nomogram to predict loco-regional control after re-irradiation for head and neck cancer. Radiother Oncol. 2014;111:382-7.
[152] Setton J, Han J, Kannarunimit D, Wuu YR, Rosenberg SA, DeSelm C, et al. Long-term patterns of relapse and survival following definitive intensity-modulated radiotherapy for non-endemic nasopharyngeal carcinoma. Oral Oncol. 2016;53:67-73.
[153] Meade S, Gaunt P, Hartley A, Robinson M, Harrop V, Cashmore J, et al. Feasibility of Dose-escalated Hypofractionated Chemoradiation in Human Papilloma Virus-negative or Smoking-associated Oropharyngeal Cancer. Clin Oncol (R Coll Radiol). 2018;30:366-74.
[154] Nutting C, Morden J, Bernstein D, Beasley M, Cosgrove V, Fisher S, et al. ART DECO (CRUK/10/018): dose escalated vs standard dose IMRT in locally advanced head and neck cancer. Radiotherapy and Oncology. 2017;122:24-.

[155] Burnet NG, Thomas SJ, Burton KE, Jefferies SJ. Defining the tumour and target volumes for radiotherapy. Cancer Imaging. 2004;4:153-61.
[156] Loimu V, Collan J, Vaalavirta L, Back L, Kapanen M, Makitie A, et al. Patterns of relapse following definitive treatment of head and neck squamous cell cancer by intensity modulated radiotherapy and weekly cisplatin. Radiother Oncol. 2011;98:34-7.
[157] Li JX, Huang SM, Jiang XH, Ouyang B, Han F, Liu S, et al. Local failure patterns for patients with nasopharyngeal carcinoma after intensity-modulated radiotherapy. Radiat Oncol. 2014;9:87.
[158] Dandekar V, Morgan T, Turian J, Fidler MJ, Showel J, Nielsen T, et al. Patterns-of-failure after helical tomotherapy-based chemoradiotherapy for head and neck cancer: implications for CTV margin, elective nodal dose and bilateral parotid sparing. Oral Oncol. 2014;50:520-6.
[159] van den Bosch S, Dijkema T, Verhoef LC, Zwijnenburg EM, Janssens GO, Kaanders JH. Patterns of Recurrence in Electively Irradiated Lymph Node Regions After Definitive Accelerated Intensity Modulated Radiation Therapy for Head and Neck Squamous Cell Carcinoma. Int J Radiat Oncol Biol Phys. 2016;94:766-74.
[160] Kjems J, Gothelf AB, Hakansson K, Specht L, Kristensen CA, Friborg J. Elective Nodal Irradiation and Patterns of Failure in Head and Neck Cancer After Primary Radiation Therapy. Int J Radiat Oncol Biol Phys. 2016;94:775-82.
[161] Weber RS, Berkey BA, Forastiere A, Cooper J, Maor M, Goepfert H, et al. Outcome of salvage total laryngectomy following organ preservation therapy: the Radiation Therapy Oncology Group trial 91-11. Arch Otolaryngol Head Neck Surg. 2003;129:44-9.
[162] Setton J, Caria N, Romanyshyn J, Koutcher L, Wolden SL, Zelefsky MJ, et al. Intensity-modulated radiotherapy in the treatment of oropharyngeal cancer: an update of the Memorial Sloan-Kettering Cancer Center experience. Int J Radiat Oncol Biol Phys. 2012;82:291-8.
[163] Mohamed AS, Rosenthal DI, Awan MJ, Garden AS, Kocak-Uzel E, Belal AM, et al. Methodology for analysis and reporting patterns of failure in the Era of IMRT: head and neck cancer applications. Radiat Oncol. 2016;11:95.
[164] Baumann M, Krause M, Hill R. Exploring the role of cancer stem cells in radioresistance. Nat Rev Cancer. 2008;8:545-54.
[165] Brizel DM, Sibley GS, Prosnitz LR, Scher RL, Dewhirst MW. Tumor hypoxia adversely affects the prognosis of carcinoma of the head and neck. Int J Radiat Oncol Biol Phys. 1997;38:285-9.
[166] Nordsmark M, Bentzen SM, Rudat V, Brizel D, Lartigau E, Stadler P, et al. Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy. An international multi-center study. Radiother Oncol. 2005;77:18-24.
[167] Isa AY, Ward TH, West CM, Slevin NJ, Homer JJ. Hypoxia in head and neck cancer. Br J Radiol. 2006;79:791-8.
[168] Thomson D, Yang H, Baines H, Miles E, Bolton S, West C, 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;26:344-7.
[169] Dirix P, Vandecaveye V, De Keyzer F, Stroobants S, Hermans R, Nuyts S. Dose painting in radiotherapy for head and neck squamous cell carcinoma: value of repeated functional imaging with (18)F-FDG PET, (18)F-fluoromisonidazole PET, diffusion-weighted MRI, and dynamic contrast-enhanced MRI. J Nucl Med. 2009;50:1020-7.
[170] Srinivasan A, Mohan S, Mukherji SK. Biologic imaging of head and neck cancer: the present and the future. AJNR Am J Neuroradiol. 2012;33:586-94.
[171] Quon H, Brizel DM. Predictive and prognostic role of functional imaging of head and neck squamous cell carcinomas. Semin Radiat Oncol. 2012;22:220-32.
[172] Jentsch C, Beuthien-Baumann B, Troost EG, Shakirin G. Validation of functional imaging as a biomarker for radiation treatment response. Br J Radiol. 2015;88:20150014.
[173] Differding S, Hanin FX, Gregoire V. PET imaging biomarkers in head and neck cancer. Eur J Nucl Med Mol Imaging. 2015;42:613-22.
[174] King AD, Chow KK, Yu KH, Mo FK, Yeung DK, Yuan J, et al. Head and neck squamous cell carcinoma: diagnostic performance of diffusion-weighted MR imaging for the prediction of treatment response. Radiology. 2013;266:531-8.
[175] Newbold K, Castellano I, Charles-Edwards E, Mears D, Sohaib A, Leach M, et al. An exploratory study into the role of dynamic contrast-enhanced magnetic resonance imaging or perfusion computed tomography for detection of intratumoral hypoxia in head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2009;74:29-37.
[176] Bernstein JM, Homer JJ, West CM. Dynamic contrast-enhanced magnetic resonance imaging biomarkers in head and neck cancer: potential to guide treatment? A systematic review. Oral Oncol. 2014;50:963-70.

[177] Hamming-Vrieze O, Navran A, Al-Mamgani A, Vogel WV. Biological PET-guided adaptive radiotherapy for dose escalation in head and neck cancer: a systematic review. Q J Nucl Med Mol Imaging. 2018;62:349-68.
[178] Gensheimer MF, Le QT. Adaptive radiotherapy for head and neck cancer: Are we ready to put it into routine clinical practice? Oral Oncol. 2018;86:19-24.
[179] Welsh L, Panek R, McQuaid D, Dunlop A, Schmidt M, Riddell A, 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;10:112.
[180] Lee N, Schoder H, Beattie B, Lanning R, Riaz N, McBride S, et al. Strategy of Using Intratreatment Hypoxia Imaging to Selectively and Safely Guide Radiation Dose De-escalation Concurrent With Chemotherapy for Locoregionally Advanced Human Papillomavirus-Related Oropharyngeal Carcinoma. Int J Radiat Oncol Biol Phys. 2016;96:9-17.
[181] Paterson C, Allwood-Spiers S, McCrea I, Foster J, McJury M, Thomson M, et al. Study of diffusion weighted MRI as a predictive biomarker of response during radiotherapy for high and intermediate risk squamous cell cancer of the oropharynx: The MeRInO study. Clin Transl Radiat Oncol. 2017;2:13-8.
[182] Teng F, Aryal M, Lee J, Lee C, Shen X, Hawkins PG, et al. Adaptive Boost Target Definition in High-Risk Head and Neck Cancer Based on Multi-imaging Risk Biomarkers. Int J Radiat Oncol Biol Phys. 2018;102:969-77.
[183] Castadot P, Lee JA, Geets X, Gregoire V. Adaptive radiotherapy of head and neck cancer. Semin Radiat Oncol. 2010;20:84-93.
[184] Schwartz DL, Garden AS, Shah SJ, Chronowski G, Sejpal S, Rosenthal DI, et al. Adaptive radiotherapy for head and neck cancer--dosimetric results from a prospective clinical trial. Radiother Oncol. 2013;106:80-4.
[185] Castelli J, Simon A, Rigaud B, Chajon E, Thariat J, Benezery K, et al. Adaptive radiotherapy in head and neck cancer is required to avoid tumor underdose. Acta Oncol. 2018;57:1267-70.
[186] van Kranen S, Hamming-Vrieze O, Wolf A, Damen E, van Herk M, Sonke JJ. Head and Neck Margin Reduction With Adaptive Radiation Therapy: Robustness of Treatment Plans Against Anatomy Changes. Int J Radiat Oncol Biol Phys. 2016;96:653-60.
[187] Navran A, Heemsbergen W, Janssen T, Hamming-Vrieze O, Jonker M, Zuur C, et al. The impact of margin reduction on outcome and toxicity in head and neck cancer patients treated with image-guided volumetric modulated arc therapy (VMAT). Radiother Oncol. 2019;130:25-31.
[188] Hamming-Vrieze O, van Kranen SR, Heemsbergen WD, Lange CA, van den Brekel MW, Verheij M, et al. Analysis of GTV reduction during radiotherapy for oropharyngeal cancer: Implications for adaptive radiotherapy. Radiother Oncol. 2017;122:224-8.
[189] Castelli J, Simon A, Lafond C, Perichon N, Rigaud B, Chajon E, et al. Adaptive radiotherapy for head and neck cancer. Acta Oncol. 2018;57:1284-92.
[190] Brouwer CL, Steenbakkers RJ, van der Schaaf A, Sopacua CT, van Dijk LV, Kierkels RG, et al. Selection of head and neck cancer patients for adaptive radiotherapy to decrease xerostomia. Radiother Oncol. 2016;120:36-40.
[191] Broggi S, Fiorino C, Dell'Oca I, Dinapoli N, Paiusco M, Muraglia A, et al. A two-variable linear model of parotid shrinkage during IMRT for head and neck cancer. Radiother Oncol. 2010;94:206-12.
[192] Fiorino C, Rizzo G, Scalco E, Broggi S, Belli ML, Dell'Oca I, et al. Density variation of parotid glands during IMRT for head-neck cancer: correlation with treatment and anatomical parameters. Radiother Oncol. 2012;104:224-9.
[193] Sanguineti G, Ricchetti F, Thomas O, Wu B, McNutt T. Pattern and predictors of volumetric change of parotid glands during intensity modulated radiotherapy. Br J Radiol. 2013;86:20130363.
[194] Wang ZH, Yan C, Zhang ZY, Zhang CP, Hu HS, Kirwan J, et al. Radiation-induced volume changes in parotid and submandibular glands in patients with head and neck cancer receiving postoperative radiotherapy: a longitudinal study. Laryngoscope. 2009;119:1966-74.
[195] Sanguineti G, Ricchetti F, Wu B, McNutt T, Fiorino C. Parotid gland shrinkage during IMRT predicts the time to Xerostomia resolution. Radiat Oncol. 2015;10:19.
[196] Ahn PH, Chen CC, Ahn AI, Hong L, Scripes PG, Shen J, et al. Adaptive planning in intensity-modulated radiation therapy for head and neck cancers: single-institution experience and clinical implications. Int J Radiat Oncol Biol Phys. 2011;80:677-85.
[197] Ajani AA, Qureshi MM, Kovalchuk N, Orlina L, Sakai O, Truong MT. A quantitative assessment of volumetric and anatomic changes of the parotid gland during intensity-modulated radiotherapy for head and neck cancer using serial computed tomography. Med Dosim. 2013;38:238-42.

[198] Ho KF, Marchant T, Moore C, Webster G, Rowbottom C, Penington H, et al. Monitoring dosimetric impact of weight loss with kilovoltage (kV) cone beam CT (CBCT) during parotid-sparing IMRT and concurrent chemotherapy. Int J Radiat Oncol Biol Phys. 2012;82:e375-82.
[199] Schwartz DL, Garden AS, Thomas J, Chen Y, Zhang Y, Lewin J, et al. Adaptive radiotherapy for head-and-neck cancer: initial clinical outcomes from a prospective trial. Int J Radiat Oncol Biol Phys. 2012;83:986-93.
[200] Wang X, Lu J, Xiong X, Zhu G, Ying H, He S, et al. Anatomic and dosimetric changes during the treatment course of intensity-modulated radiotherapy for locally advanced nasopharyngeal carcinoma. Med Dosim. 2010;35:151-7.
[201] Reali A, Anglesio SM, Mortellaro G, Allis S, Bartoncini S, Ruo Redda MG, et al. Volumetric and positional changes of planning target volumes and organs at risk using computed tomography imaging during intensity-modulated radiation therapy for head-neck cancer: an "old" adaptive radiation therapy approach. Radiol Med. 2014;119:714-20.
[202] Nishi T, Nishimura Y, Shibata T, Tamura M, Nishigaito N, Okumura M. Volume and dosimetric changes and initial clinical experience of a two-step adaptive intensity modulated radiation therapy (IMRT) scheme for head and neck cancer. Radiother Oncol. 2013;106:85-9.
[203] Zhao L, Wan Q, Zhou Y, Deng X, Xie C, Wu S. The role of replanning in fractionated intensity modulated radiotherapy for nasopharyngeal carcinoma. Radiother Oncol. 2011;98:23-7.
[204] Castadot P, Geets X, Lee JA, Gregoire V. Adaptive functional image-guided IMRT in pharyngo-laryngeal squamous cell carcinoma: is the gain in dose distribution worth the effort? Radiother Oncol. 2011;101:343-50.
[205] Wu Q, Chi Y, Chen PY, Krauss DJ, Yan D, Martinez A. Adaptive replanning strategies accounting for shrinkage in head and neck IMRT. Int J Radiat Oncol Biol Phys. 2009;75:924-32.
[206] Duma MN, Kampfer S, Schuster T, Winkler C, Geinitz H. Adaptive radiotherapy for soft tissue changes during helical tomotherapy for head and neck cancer. Strahlenther Onkol. 2012;188:243-7.
[207] Chen C, Lin X, Pan J, Fei Z, Chen L, Bai P. Is it necessary to repeat CT imaging and replanning during the course of intensity-modulated radiation therapy for locoregionally advanced nasopharyngeal carcinoma? Jpn J Radiol. 2013;31:593-9.
[208] Cheng HC, Wu VW, Ngan RK, Tang KW, Chan CC, Wong KH, et al. A prospective study on volumetric and dosimetric changes during intensity-modulated radiotherapy for nasopharyngeal carcinoma patients. Radiother Oncol. 2012;104:317-23.
[209] Graff P, Hu W, Yom SS, Pouliot J. Does IGRT ensure target dose coverage of head and neck IMRT patients? Radiother Oncol. 2012;104:83-90.
[210] Bhide SA, Davies M, Burke K, McNair HA, Hansen V, Barbachano Y, et al. Weekly volume and dosimetric changes during chemoradiotherapy with intensity-modulated radiation therapy for head and neck cancer: a prospective observational study. Int J Radiat Oncol Biol Phys. 2010;76:1360-8.
[211] Capelle L, Mackenzie M, Field C, Parliament M, Ghosh S, Scrimger R. Adaptive radiotherapy using helical tomotherapy for head and neck cancer in definitive and postoperative settings: initial results. Clin Oncol (R Coll Radiol). 2012;24:208-15.
[212] Duma MN, Kampfer S, Wilkens JJ, Schuster T, Molls M, Geinitz H. Comparative analysis of an image-guided versus a non-image-guided setup approach in terms of delivered dose to the parotid glands in head-and-neck cancer IMRT. Int J Radiat Oncol Biol Phys. 2010;77:1266-73.
[213] Fung WW, Wu VW, Teo PM. Dosimetric evaluation of a three-phase adaptive radiotherapy for nasopharyngeal carcinoma using helical tomotherapy. Med Dosim. 2012;37:92-7.
[214] Height R, Khoo V, Lawford C, Cox J, Joon DL, Rolfo A, et al. The dosimetric consequences of anatomic changes in head and neck radiotherapy patients. J Med Imaging Radiat Oncol. 2010;54:497-504.
[215] Jensen AD, Nill S, Huber PE, Bendl R, Debus J, Munter MW. A clinical concept for interfractional adaptive radiation therapy in the treatment of head and neck cancer. Int J Radiat Oncol Biol Phys. 2012;82:590-6.
[216] Marzi S, Pinnaro P, D'Alessio D, Strigari L, Bruzzaniti V, Giordano C, et al. Anatomical and dose changes of gross tumour volume and parotid glands for head and neck cancer patients during intensity-modulated radiotherapy: effect on the probability of xerostomia incidence. Clin Oncol (R Coll Radiol). 2012;24:e54-62.
[217] O'Daniel JC, Garden AS, Schwartz DL, Wang H, Ang KK, Ahamad A, et al. Parotid gland dose in intensity-modulated radiotherapy for head and neck cancer: is what you plan what you get? Int J Radiat Oncol Biol Phys. 2007;69:1290-6.
[218] Hunter KU, Fernandes LL, Vineberg KA, McShan D, Antonuk AE, Cornwall C, et al. Parotid glands dose-effect relationships based on their actually delivered doses: implications for adaptive

replanning in radiation therapy of head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2013;87:676-82.
[219] Cozzolino M, Fiorentino A, Oliviero C, Pedicini P, Clemente S, Califano G, et al. Volumetric and dosimetric assessment by cone-beam computed tomography scans in head and neck radiation therapy: a monitoring in four phases of treatment. Technol Cancer Res Treat. 2014;13:325-35.
[220] Castadot P, Geets X, Lee JA, Christian N, Gregoire V. Assessment by a deformable registration method of the volumetric and positional changes of target volumes and organs at risk in pharyngo-laryngeal tumors treated with concomitant chemo-radiation. Radiother Oncol. 2010;95:209-17.
[221] Popovtzer A, Cao Y, Feng FY, Eisbruch A. Anatomical changes in the pharyngeal constrictors after chemo-irradiation of head and neck cancer and their dose-effect relationships: MRI-based study. Radiotherapy and Oncology. 2009;93:510-5.
[222] Ricchetti F, Wu BB, McNutt T, Wong J, Forastiere A, Marur S, et al. Volumetric Change of Selected Organs at Risk during Imrt for Oropharyngeal Cancer. Int J Radiat Oncol. 2011;80:161-8.
[223] Senkus-Konefka E, Naczk E, Borowska I, Badzio A, Jassem J. Changes in lateral dimensions of irradiated volume and their impact on the accuracy of dose delivery during radiotherapy for head and neck cancer. Radiother Oncol. 2006;79:304-9.
[224] Castelli J, Simon A, Louvel G, Henry O, Chajon E, Nassef M, et al. Impact of head and neck cancer adaptive radiotherapy to spare the parotid glands and decrease the risk of xerostomia. Radiat Oncol. 2015;10:6.
[225] Teshima K, Murakami R, Tomitaka E, Nomura T, Toya R, Hiraki A, et al. Radiation-induced parotid gland changes in oral cancer patients: correlation between parotid volume and saliva production. Jpn J Clin Oncol. 2010;40:42-6.
[226] Belli ML, Scalco E, Sanguineti G, Fiorino C, Broggi S, Dinapoli N, et al. Early changes of parotid density and volume predict modifications at the end of therapy and intensity of acute xerostomia. Strahlenther Onkol. 2014;190:1001-7.
[227] Nishimura Y, Nakamatsu K, Shibata T, Kanamori S, Koike R, Okumura M, et al. Importance of the initial volume of parotid glands in xerostomia for patients with head and neck cancers treated with IMRT. Jpn J Clin Oncol. 2005;35:375-9.
[228] Dirix P, De Keyzer F, Vandecaveye V, Stroobants S, Hermans R, Nuyts S. Diffusion-weighted magnetic resonance imaging to evaluate major salivary gland function before and after radiotherapy. Int J Radiat Oncol Biol Phys. 2008;71:1365-71.
[229] Astreinidou E, Roesink JM, Raaijmakers CP, Bartels LW, Witkamp TD, Lagendijk JJ, et al. 3D MR sialography as a tool to investigate radiation-induced xerostomia: feasibility study. Int J Radiat Oncol Biol Phys. 2007;68:1310-9.
[230] Kan T, Kodani K, Michimoto K, Fujii S, Ogawa T. Radiation-induced damage to microstructure of parotid gland: evaluation using high-resolution magnetic resonance imaging. Int J Radiat Oncol Biol Phys. 2010;77:1030-8.
[231] Houweling AC, Schakel T, van den Berg CA, Philippens ME, Roesink JM, Terhaard CH, et al. MRI to quantify early radiation-induced changes in the salivary glands. Radiother Oncol. 2011;100:386-9.
[232] Cheng CC, Chiu SC, Jen YM, Chang HC, Chung HW, Liu YJ, et al. Parotid perfusion in nasopharyngeal carcinoma patients in early-to-intermediate stage after low-dose intensity-modulated radiotherapy: evaluated by fat-saturated dynamic contrast-enhanced magnetic resonance imaging. Magn Reson Imaging. 2013;31:1278-84.
[233] Juan CJ, Cheng CC, Chiu SC, Jen YM, Liu YJ, Chiu HC, et al. Temporal Evolution of Parotid Volume and Parotid Apparent Diffusion Coefficient in Nasopharyngeal Carcinoma Patients Treated by Intensity-Modulated Radiotherapy Investigated by Magnetic Resonance Imaging: A Pilot Study. PLoS One. 2015;10:e0137073.
[234] Kataria T, Gupta D, Goyal S, Bisht SS, Basu T, Abhishek A, et al. Clinical outcomes of adaptive radiotherapy in head and neck cancers. Br J Radiol. 2016;89:20160085.
[235] Chen AM, Daly ME, Cui J, Mathai M, Benedict S, Purdy JA. Clinical outcomes among patients with head and neck cancer treated by intensity-modulated radiotherapy with and without adaptive replanning. Head Neck. 2014;36:1541-6.
[236] Yang H, Hu W, Wang W, Chen P, Ding W, Luo W. Replanning during intensity modulated radiation therapy improved quality of life in patients with nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys. 2013;85:e47-54.
[237] Heukelom J, Hamming O, Bartelink H, Hoebers F, Giralt J, Herlestam T, et al. Adaptive and innovative Radiation Treatment FOR improving Cancer treatment outcomE (ARTFORCE); a randomized controlled phase II trial for individualized treatment of head and neck cancer. BMC Cancer. 2013;13:84.

[238] Esteyrie V, Gleyzolle B, Lusque A, Graff P, Modesto A, Rives M, et al. The GIRAFE phase II trial on MVCT-based "volumes of the day" and "dose of the day" addresses when and how to implement adaptive radiotherapy for locally advanced head and neck cancer. Clin Transl Radiat Oncol. 2019;16:34-9.
[239] Brown E, Owen R, Harden F, Mengersen K, Oestreich K, Houghton W, et al. Predicting the need for adaptive radiotherapy in head and neck cancer. Radiother Oncol. 2015;116:57-63.
[240] Hu YC, Tsai KW, Lee CC, Peng NJ, Chien JC, Tseng HH, et al. Which nasopharyngeal cancer patients need adaptive radiotherapy? BMC Cancer. 2018;18:1234.
[241] Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. Phys Med Biol. 2008;53:1027-42.
[242] Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions. Phys Med Biol. 2008;53:1043-56.
[243] McGowan SE, Burnet NG, Lomax AJ. Treatment planning optimisation in proton therapy. Br J Radiol. 2013;86:20120288.
[244] Liebl J, Paganetti H, Zhu M, Winey BA. The influence of patient positioning uncertainties in proton radiotherapy on proton range and dose distributions. Med Phys. 2014;41:091711.
[245] Lowe M, Albertini F, Aitkenhead A, Lomax AJ, MacKay RI. Incorporating the effect of fractionation in the evaluation of proton plan robustness to setup errors. Phys Med Biol. 2016;61:413-29.
[246] Holloway SM, Holloway MD, Thomas SJ. A method for acquiring random range uncertainty probability distributions in proton therapy. Phys Med Biol. 2017;63:01NT2.
[247] Crellin A. The Road Map for National Health Service Proton Beam Therapy. Clin Oncol (R Coll Radiol). 2018;30:277-9.

Chapter 2 – VoxTox Head and Neck

2.1 Overview

The VoxTox research collaboration is an interdisciplinary programme based at the University of Cambridge. Its fundamental objective is to better understand the relationship between radiation dose at the voxel level, and the toxicity experienced by patients following radiotherapy treatment. Funding was achieved via a programme grant from Cancer Research UK, and ethical approval was granted in February 2013 (13/EE/0008). The study was listed within the UK Clinical Research Network Study Portfolio (UK CRN ID 13716), opened in April 2013, and closed to recruitment in April 2019.

The work I have done for my PhD thesis, as described in this dissertation, has all been within the framework of the broader VoxTox programme, analysing data from patients recruited to the study, and utilising the research infrastructure and collaborations therein. This chapter describes the specifics of the broader VoxTox programme, as well as treatment and image guidance protocols for HNC patients recruited to the study. Technical details of the TomoTherapy system, core to the broader VoxTox methodology, are discussed, as are specifics of methodologies used for the extraction, and segmentation, of imaging data.

2.1.1 My Role

As described in section 2.1.2, the VoxTox study was designed, funded, and well underway before I joined the research group. Nonetheless, it is necessary to provide an overview of the study in order to give important context for subsequent chapters. On joining the research group I became an active member of the study management group, and contributed significantly to the overall running and conduct of the study. In April 2018, I submitted a protocol substantial amendment to IRAS, in order to open a sub-study of VoxTox for selected HNC patients to have multiple timepoint MRI scans during their radiotherapy treatment – the Minot-OAR sub-study. This is described in detail in Chapter 8.

Construction of the database of full patient details for VoxTox H&N patients was my work. Although colleagues provided some of the raw data, I undertook all the data analysis presented in the results section of this chapter. I designed, executed, and analysed all the data for the intra- and inter-observer contouring variability study on kVCT planning scans.

2.1.2 Acknowledgments

Neil Burnet (Professor of Radiation Oncology) and Raj Jena (Consultant Oncologist and Principle Research Associate) conceived the original idea for the VoxTox study. With assistance from Michael Sutcliffe (Collaborative Research Facilitator), Neil Burnet wrote the research grant funding application to Cancer Research UK (CRUK), and defended it at interview with Andy Parker (Professor of High Energy Physics), and Michael Sutcliffe (Professor of Biomedical Engineering). The study was funded by CRUK.

Jessica Scaife (Consultant Oncologist, Cheltenham Hospital) wrote the study protocol, as well the ethics submission, patient information sheets (PIS) and clinical reporting forms (CRFs). Amy Bates (Research Radiographer) helped Jessica Scaife to write the CRFs, and to incorporate them into the radiotherapy care management software system Mosaiq (Elekta, Stockholm, Sweden). The two patient representatives of the clinical study management group, Hilary Benn and Elizabeth Stobart, helped to write the PIS and CRFs. With the assistance of other members of her team, Amy Bates also led on approaching, recruiting and consenting patients into the VoxTox study, and subsequent late toxicity assessments. With the help of June Dean (Superintendent Radiographer), Amy Bates also wrote work instructions for image-guidance scans for patients recruited to the study.

Patients recruited to the study were treated under the care of consultant oncologists Sarah Jefferies, Richard Benson, Gill Barnett, and Rashmi Jadon. Consultant oncologists Richard Benson, Gill Barnett, Rashmi Jadon and Kerwyn Foo provided contours for the inter-observer variability data in this chapter. Gold standard contours from the Estro Falcon contouring workshop referred to in section 2.3.3 were provided by Dr Jon Cacicedo, Radiation Oncologist, Osakidetza, Barakaldo, Spain. Professor Hans Langendijk, University Medical Centre Groningen, provided gold standard contours for collaborative work on the CITOR project, which were also used for quality control of my own segmentations within VoxTox, as described in section 2.3.3.

Marina Romanchikova (Principle Research Scientist, NPL) wrote code to extract and anonymise relevant imaging, structure set, and dose data in DICOM format from vendor archive (Accuray, Sunnyvale, California, USA), and export them to the Cavendish Laboratory. Karl Harrison (Research Associate, Cavendish Laboratory) was responsible for storing, curating and processing data on Cavendish Laboratory servers. Andrew Hoole (Medical Physicist) facilitated interface between VoxTox study infrastructure, and Hospital Trust IT systems.

An overview of the design and infrastructure of the VoxTox study, focusing especially on the links between the radiotherapy department at Addenbrooke's Hospital and the Cavendish Laboratory, was published as follows:

Applying Physical Science techniques and CERN technology to an unsolved problem in radiation treatment for cancer: the multidisciplinary VoxTox research programme. Burnet NG, Scaife JE, Romanchikova M, Thomas SJ, Bates AM, Wong E, Noble DJ, et al. *IdeaSquare Journal of Expermiental Innovation, 2017; 1(1): 3*

2.2 Introduction

2.2.1 Integrated IG-IMRT with TomoTherapy

The TomoTherapy Hi-Art system is an elegant, single-platform solution for delivering integrated image guidance with IMRT (IG-IMRT) [1]. The design concept is similar to a helical CT scanner, with both the patient couch, and treatment gantry in continuous motion [2]. The patient is positioned on the carbon-fibre couch, which has 4 degrees of freedom. Specifically, the couch can accommodate left-right, superior-inferior, and anterior-posterior movements, as well as roll. The couch cannot accommodate movements to correct for pitch, or yaw [3]. These movements are represented in cartoon form in Figure 2.1. The 6MV flattening filter-free linear-accelerator is mounted on a ring gantry, which can deliver treatment from a full 360°, although in practice, treatment is delivered from 51 beam directions, each separated by approximately 7° [2-4]. Directly opposite is a xenon filled detector array [5].

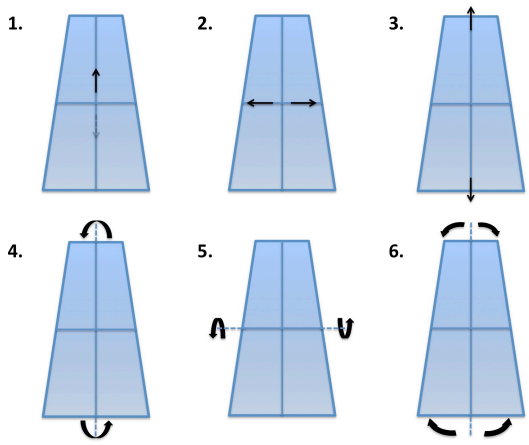


Figure 2.1: 6 theoretical degrees of freedom for a radiotherapy couch: 1. Anterior/posterior, 2. Left/right, 3. Superior/inferior, 4. Roll, 5. Pitch, 6. Yaw. The TomoTherapy Hi-Art system is capable of movements 1-4.

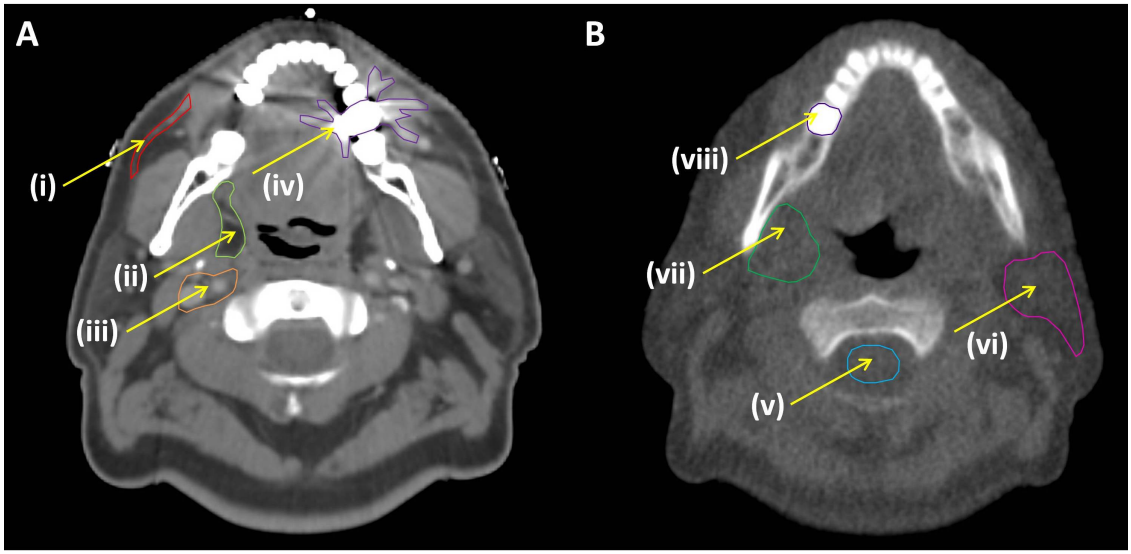


Figure 2.2 (A&B): Examples of kVCT planning CT scan with IV contrast (A) and image guidance MVCT (B) images (images from different patients). The detail in A allows resolution of detailed soft tissue structures such as parotid duct (i), parapharyngeal fat space (ii) and neurovascular bundle (iii). There is also significant streaking artefact from dental amalgam (iv). The MVCT gives sufficient detail to resolve the spinal cord (v) and parotid gland (vi), but it is very difficult to define the boundary between parapharyngeal space and medial pterygoid muscle (vii). Of note, there is less dental artefact (viii), due to higher energy MVCT X-Rays.

Standard CT scanners generate X-Rays from a potential difference often in the range of 80 - 140kV [6]. At these energies, photons predominantly interact with matter via the photoelectric effect, whereas photons in the MV energy interact mainly through Compton scatter. In theory and in practice, images generated from lower energy photon are better than higher energy equivalents. Therefore, to generate IG scans, the TomoTherapy system tunes down the nominal accelerator energy of 6MV to 3.5MV, giving a mean photon energy of 0.75MV [7]. The images produced by the system are noisier than diagnostic kVCT, with inferior soft tissue definition characteristics. However, they are adequate to identify many anatomical structures, and permit on-line registration with the planning CT [3]. An example of matched planning kVCT and IG-MVCT images is shown in Figure 2.2. The TomoTherapy system permits MVCT imaging with pre-set slice thickness spacings of 2, 4 and 6mm [7]. Thinner slice thickness confers finer spatial resolution, at the expense of longer ‘on-couch’ time for the patient, and higher radiation dose from imaging. Radiation doses associated with each slice thickness are typically 2.5, 1.3 and 0.9cGy for fine, normal and coarse settings respectively [3].

Addenbrooke’s Hospital installed its first TomoTherapy Hi-Art system in 2007, and head and neck was one of the tumour sites chosen for early work to establish effective workflows with the technology [3]. A second system was installed in 2009, and once ramp-up was complete, all patients with HNC who underwent radical radiotherapy at Addenbrooke’s were treated on

this platform, until the first unit was decommissioned in 2017. In order to maximise benefit for each patient, departmental protocol over this period was daily image guidance for all patients. Therefore, to minimise per-patient IG time, and ensure adequate throughput in a busy department, policy further stipulated that outwith exceptional circumstances, all IG should be done with coarse (6mm) slice thickness. Audit data suggest that 97% of IG done on the Hi-Art systems was with 6mm slice thickness [8].

Early work on the TomoTherapy system reported axial pixel size of MVCT images at 0.78x0.78mm [7]; images generated by the Addenbrooke's units had in-plane resolution of 0.754x0.754mm. Thus, the voxel size of MVCT images for the vast majority of patients treated on TomoTherapy Hi-Art system at Addenbrooke's, for all HNC patients in the VoxTox study, and for all subsequent work in this dissertation, is 0.754x0.754x6mm, as shown in Figure 2.3.

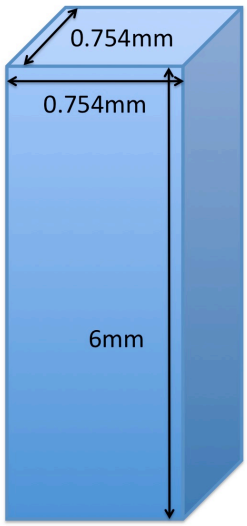


Figure 2.3: Cartoon of image guidance MVCT image voxel dimensions (distances not to scale)

2.2.2 Making the most of daily image guidance scans.

At the time the VoxTox study was conceived, IG scans were performed for one purpose only – to maximise setup accuracy. Thus, prior to the study opening, the daily MVCT scans done on the TomoTherapy units at Addenbrooke's would have been used for daily matching, but thereafter sent to a vendor archive, with no further information extracted from them. The hypotheses of the VoxTox study were based on the premise that these images may contain useful data that could be extracted and utilised for patient benefit within the workflows and resource constraints of clinical practice.

The study was therefore designed to minimise disruption at an operational level within the clinical department, whilst maximising existing academic links within the University of Cambridge, and forming news ones. In the process, a research framework involving the hospital oncology and medical physics departments, the clinical trials unit, and the University departments of High Energy Physics, Engineering and Applied Mathematics and Theoretical Physics was constructed. This process, and the resulting infra-structure is outlined in detail in the inaugural issue of the CERN IdeaSquare Journal of Experimental Innovation [9].

2.2.3 Aims of chapter 2

1. To describe the design and infrastructure of the VoxTox study, as fundamental background to all subsequent data presented in the dissertation.
2. To construct a finalised study cohort from which to address the primary hypothesis of the PhD thesis.
3. To quantify the nature and volume of the imaging, dose and toxicity data associated with this cohort
4. To quality control my own manual segmentation of H&N OARs on kVCT planning CT scans, as an important methodological step for dose accumulation.

In addressing these aims, the chapter has 3 broad themes. Firstly, a description of the clinical protocols that were used to treat HNC patients in VoxTox (section 2.3.1). Secondly, information on the structure of the VoxTox study itself, including summary data on the final study cohort, and the imaging data associated with it (sections 2.3.2, 2.4.1 - 2.4.3, 2.5.1 and 2.5.2). Finally, a description of my work to segment relevant OARs on planning CT scans, and to validate the quality of these contours (sections 2.3.3, 2.4.4, 2.5.3).

2.3 Methods

2.3.1 Treatment protocols for H&N patients at CUH

2.3.1.1 Patient pathway and radiotherapy treatment details

H&N patients in the VoxTox study were investigated and treated as per standard national guidelines, and regional and departmental protocols. Workup included full inspection of the oro/pharyngo/laryngeal mucosal surfaces, where necessary under full anaesthetic, and

pertinent imaging for TNM staging by US, CT, MRI and PET/CT. Tissue biopsies were taken from relevant sites of disease, and assessment of patient's comorbidities and fitness was performed. The use of p16 immunohistochemistry as a surrogate marker for HPV driven disease became standard through the course of the study, but did not affect management decisions for patients in the VoxTox study. All cases were discussed in the H&N MDT, and patients were referred for radiotherapy under the care of the clinicians outlined in section 2.1.2.

All head and neck patients in the VoxTox study were scanned and treated supine, and immobilised with a 5-point fixation thermoplastic shell. Radiotherapy planning kVCT images were acquired with a GE Medical Systems LightSpeed Multislice CT scanner or a Toshiba Aquilon LB CT scanner, with a voxel size of 1.074 x 1.074 x 3mm. In the early years of the VoxTox study, it was not standard practice to administer IV contrast for planning CT scans. However, from 2014 this became possible, and was increasingly used thereafter.

Planning scans were geometrically registered, saved in DICOM format, and sent to the database of the hospital clinical segmentation software – Prosoma (MedCOM, Darmstadt, Germany). Clinical segmentation of both target volumes and OARs was undertaken using the most recent consensus atlases, guidelines and clinical trial protocols [10-15], which naturally evolved over the course of the study.

Throughout the course of the study, a 3-dose level technique was used for primary treatment of pharyngo/laryngeal SCC's. Prior to November 2011, standard fractionation was 68Gy to gross disease, 60Gy to high-risk elective regions, and 54Gy in 34 fractions to intermediate risk elective regions. From November 2011, protocol changed to 65Gy, 60Gy and 54Gy in 30 fractions respectively. Standard protocol for primary disease of the oral cavity, sinuses, skin and salivary glands was maximum safe surgery with neck dissection where appropriate, with adjuvant radiotherapy at a dose of 60Gy in 30 fractions. A summary of radiotherapy treatment details, including dose, fractionation, and target volume construction is shown in table 2.1.

Once clinical contours were approved, planning scans and structure sets were transferred to the TomoTherapy treatment planning system (TPS). The first step in the workflow in this system was to downsample the planning CT scan by a factor of 2 in the axial plane (see section 2.3.2.5.2 for details). Inverse planning using an objective function was used to generate treatment plans, with homogeneous dose coverage of targets, and adequate sparing of neurological OARs given the highest priority in the TPS optimiser. A copy of a standard H&N prescription, including dose constraints and objectives for all relevant OARs, is shown in Appendix A2.1. Although this document is a recent version, it is representative of the workflow

used over the course of the VoxTox study. All finalised treatment plans were reviewed and signed by responsible clinicians in a weekly H&N planning meeting, and underwent standard patient-specific QA procedures.

Throughout the course of the study, a protocol for ART was in place for HNC patients within VoxTox. This process was known as 'Delivered Dose Assessment' – DDA. In this process, the MVCT of the day was re-imported in the TomoTherapy TPS from the treatment machine by the clinical physics team, and the daily dose calculated on this image. A rigid registration was then done between the planning CT and the daily MVCT so that contours from the planning scan could be aligned with the geometry of the MVCT. Dose metrics for these contours were then computed, and potential differences with planned dose inferred. All HNC patients underwent routine DDA on treatment days 1 and 10, regardless of progress. However, ad-hoc DDA's could be requested by any member of the treatment team for a number of reasons, including checking positional reproducibility, internal or external shape change, weight loss, or because of concern over previous trends. Three pre-specified metrics were defined, and used by radiotherapy physics to evaluate the results. These were:

1. Are there any new hot spots in any PTVs above tolerance (107%)?
2. Are there any new cold spots in any PTVs below 95%?
3. Are there any new hot spots in any OARs above tolerance?

If the answer to all 3 questions was no, then the DDA would be signed as satisfactory by physics, with no further action required. If any answer was yes, then the case was discussed with the full team including radiographers and clinicians at the weekly head and neck planning meeting. Options for outcome included no further action, repeat DDA at a specified interval, or to instigate repeat planning CT and treatment re-plan. An example of the DDA document is shown in Appendix A2.2.

Subsite	Dose/ Fractionation	GTV	High dose CTV†	High risk CTV†	Ipsilateral neck†,§	Contralateral neck†,§	PTV
Nasopharynx	70/60/57Gy 35 fractions 7 weeks	Planning CT + contrast. Fused diagnostic or planning MRI mandatory	GTV + 5mm, modified for bone and air spaces‡	- High dose CTV + 5mm - Ensure High dose CTV covered - Entire level of involved node(s)	Ib - V	II - IVa	5mm
Oropharynx	65/60/54Gy* 30 fractions 6 weeks	Planning CT + contrast. Fuse diagnostic MRI/CT or planning MRI where available/needed	GTV + 5-10mm modified for bone and air spaces‡	- Whole oropharynx; soft palate to cranial aspect of the body of the hyoid. - Medial border is midline for well lateralised tumours - Ensure High dose CTV covered - Entire level of involved node(s)	- Ib – Va¶ - RP/VIIa#,** - if bulky and/or high level II node, then add RS/VIIb.	II - IVa	5mm
Hypopharynx	65/60/54Gy* 30 fractions 6 weeks	Planning CT + contrast. Fuse diagnostic MRI/CT or planning MRI where available/needed	GTV + 10mm modified for bone and air spaces‡	- Whole hyopharynx/larynx; cranial aspect of the body of the hyoid to caudal cricoid. - Ensure High dose CTV covered - Entire level of involved node(s)	- Ib-Vb¶ - RP/VIIa#,** - add VI if apex of pyriform sinus involved.	- II-IVa - VI for apex of pyriform sinus - VIIa.	5mm
Larynx	65/60/54Gy* 30 fractions 6 weeks	Planning CT + contrast. Fuse diagnostic MRI/CT or planning MRI where available/needed	GTV + 5-10mm modified for bone and air spaces‡	- Whole larynx and pre-epiglottic space from above hyoid to inf. border of cricoid. - Ensure High dose CTV covered - Entire level of involved node(s)	- Ib-Vb¶ - RP/VIIa#,** - add VI if apex of pyriform sinus involved.	- II-IVa - VI for apex of pyriform sinus	5mm

Subsite	Dose/ Fractionation	GTV	High dose CTV	High risk CTV	Ipsilateral neck	Contralateral neck	PTV
Oral cavity††	60/54Gy 30 fractions 6 weeks	Planning CT + contrast. Fused pre-op MRI desirable	N/A (assuming primary surgery, R0 or R1 resection)	- Pre-operative GTV + 10-15mm - Ensure resection margins covered. - Pterygoid fossa for RMT tumours - Entire level of involved node(s)	- I-Iva - IVb if IVa involved - level V if N2+	- For staging up to well-lateralised T2N1: none. - more advanced stage: I-IVa	5mm
Sinus††	60/54Gy 30 fractions 6 weeks	Planning CT + contrast. Fused pre-op MRI, +/- post-op planning MRI desirable.	N/A (assuming primary surgery, R0 or R1 resection)	- Pre-operative GTV + 10mm - Ensure resection margins covered. - Cover maxillary & ethmoid sinus, nasal & pterygopalatine fossae.	- Levels I-V, guided clinically, if high risk features following neck dissection.	Rarely	5mm
Skin††	60Gy 30 fractions 6 weeks	Planning CT + contrast. Fuse pre-op volumetric imaging where available and relevant.	N/A (assuming primary surgery, R0 or R1 resection)	- Cover site of primary excision + 10-15mm margin (depending on surgical margins). - Ensure resection margins (plus flaps/reconstruction) covered.	- Levels I-V, guided clinically, if high risk features following neck dissection.	Rarely	5mm
Salivary gland††	60Gy 30 fractions 6 weeks	Planning CT + contrast. Fuse pre-op volumetric imaging where available and relevant.	N/A (assuming primary surgery, R0 or R1 resection)	- Gland bed and ductal tissue, scar, parapharyngeal space (parotid). - Consider peri-neural invasion (esp. Adenoid cystic), e.g. facial nerve to stylo-mastoid foramen.	- levels I-III guided clinically following surgery.	Rarely	5mm

Table 2.1: Summary of radiotherapy details and techniques, for head and neck patients treated at Addenbrooke's Hospital as part of the VoxTox study. Notes: *Prior to November 2011 – 68/60/57Gy, 34 fractions, 6.8 weeks. †High dose CTV = 65Gy, High risk CTV = 60Gy, Neck CTV = 54Gy unless otherwise stated. ‡GTV to CTV margins stated for primary lesions. Nodal GTV to CTV margins 5mm in most cases. §Lymph node levels as defined in RTOG guidelines [10]. ¶Level IVb added if IVa involved. #RP = retropharyngeal lymph nodes – level VIIa, RS = retrostyloid lymph nodes, level VIIb. **RS/VIIb nodes included if high/bulky level II node. ††Protocol assuming primary surgery undertaken before radiotherapy.

2.3.1.2 Concomitant therapy

The benefit of concomitant cisplatin delivered alongside radical radiation (CRT) for curative treatment of squamous cell carcinoma of the H&N is well established [16, 17]. These data show that the benefit is greater in patients with more advanced disease, and that the magnitude of benefit declines with age, with little or no benefit seen in patients over the age of 70 [17]. The benefit also comes with significant cost, as combined cisplatin-RT is significantly more toxic than radiation alone [18-21]. There is also data to support the use of the anti-EGFR monoclonal antibody Cetuximab as a radiosensitiser in this context (CetuxRT) [22]. However, more recent trials, designed to test the hypothesis that replacing concomitant cisplatin with cetuximab in patients with HPV positive disease would maintain local control and survival rates whilst reducing side effects, showed no differences in toxicity outcomes, but inferior efficacy [23, 24]. It is important to note that these data were published concurrent with the closure of VoxTox, and therefore had no impact on the treatment protocols used during the conduct of the study.

Therefore, concomitant cisplatin and cetuximab were used as standard-of-care concomitant therapies throughout the course of the study. Cetuximab was administered with a standard schedule of a 400mg/m^2 loading dose in the week prior to radiotherapy commencing, and a weekly dose of 250mg/m^2 thereafter. Cisplatin was given as weekly dose of 40mg/m^2, with the exception of nasopharyngeal primaries, where the 3-weekly 100mg/m^2 regimen was used [25]. This is slightly more controversial, as most data supporting CRT are based on the 3-weekly schedule, and this regimen is considered standard of care [26]. The rationale for this approach is based mainly on the premise that lower weekly dose may be less toxic, and provide clinicians with greater 'control' to stop chemotherapy if patients experience unacceptable side effects.

There are no randomised data directly comparing weekly 40mg/m^2 with 3-weekly 100mg/m^2. An Indian study compared weekly 30mg/m2 with the standard regimen, and found worse local control, although no differences in either PFS or OS [27]. In contrast, large retrospective analyses have suggested no clear difference in disease outcomes for weekly 40mg/m^2 versus 3-weekly 100mg/m2 [26, 28], although toxicity did indeed appear to be lower. Interestingly, there seems to be an important threshold associated with a total cumulative cisplatin dose of 200mg/m^2 [26, 29], a target that is clearly reachable with 40mg/m^2, but not 30mg/m^2, if a weekly schedule over 6 weeks is used. Unpublished audit data from our own department suggest that

in patients who consent for, and start CRT, the median number of weekly cycles delivered is 6, and the mean cumulative cisplatin dose is above the nominal threshold of 200mg/m^2.[1]

Criteria by which clinicians decided whether or not to offer concomitant systemic therapy to patients within the VoxTox study are outlined below.

Cisplatin:

- All (sufficiently fit) patients with nasopharyngeal carcinoma.
- Locally advanced (Stage III-IV, TNM 7) SCC of the oropharynx, hypopharynx or larynx, undergoing primary curative radiotherapy.
- High risk post-operative SCC (e.g. oral cavity, sinus primary), including factors such as bulky/locally-advanced disease, close or positive surgical margins, and lymph node involvement with extra-capsular extension.
- Age ≤70.
- Performance status 0-1.
- Adequate organ function.
- Absence of significant comorbidities, including renal disease, neuropathy, hearing disorders and cardiovascular disease.

Cetuximab:

- Locally advanced (Stage III-IV, TNM 7) SCC of the oropharynx, hypopharynx or larynx, undergoing primary, curative, radiotherapy, who were:
    - Fit enough for concomitant therapy (PS 0-1), but,
    - Ineligible for cisplatin for a specific reason, e.g. inadequate kidney function, pre-existing hearing problem.

2.3.2 VoxTox study design

2.3.2.1 Timeline, study cohorts and recruitment targets.

The application for a programme grant to support the VoxTox study was submitted on 9th June 2011. The application was defended at interview on 13th October 2011 and confirmation of funding was received on 8 December 2011. Ethical approval for the study was confirmed on 15 February 2013 (13/EE/0008), and the study was listed within the UK Clinical Research

[1] Dose density of weekly cisplatin delivered concurrent with radiotherapy in head and neck cancer patients. Laidley HM, Noble DJ et al. *RCR Audit Competition 2017*.

Network Study Portfolio (UK CRN ID 13716). Recruitment opened in April 2013, and closed in April 2019.

In order to maximise recruitment to the study, and to permit immediate access to imaging data to allow refinement of workflows, it was decided that recruitment should be both prospective and retrospective. Thus, patients who met both inclusion and exclusion criteria, but had already finished their treatment, could be approached as potential participants in the 'Discovery' cohort. New patients, identified through standard screening procedures as they progressed through the radiotherapy pathway, would be approached to participate in the 'Consolidation' cohort. Characteristics of both cohorts are listed below:

Discovery Cohort:

- Treated April 2008 – April 2013
- No comorbidity, baseline or acute toxicity data.
- Late toxicity data from the point of recruitment only. Therefore point prevalence analysis possible, but cumulative incidence not.
- IG scans done with standard departmental, rather than VoxTox study, protocol, resulting in a shorter FoV (see section 2.3.2.5.1 for details).
- Once extracted from archive (see section 2.3.2.5.2), imaging data available for immediate analysis.

Consolidation Cohort:

- Treated April 2013 – April 2019
- Full comorbidity, baseline, acute and late (follow up dependent) toxicity available for all patients.
- All IG scans with extended FoV according to study protocol.
- Imaging data not immediately available – to accrue with study progression.

Funding for the project via the programme grant completed in April 2017, although CRUK continued to support the study with a no-cost extension through to its completion. In the same month, HRA approval for a substantial amendment to the VoxTox protocol was approved. This included the MultIple timepoiNt MRI tO Track OAR changes in patients undergoing radical radiotherapy for head and neck cancer – Minot-OAR – substudy. This substudy was conceived and designed by me, and the objectives and design of the study are discussed in greater detail in Chapter 8. Very briefly, patients in Minot-OAR underwent 4 MRI scans, timed to coincide with radiotherapy fractions 1, 6, 16 and 26. All HNC patients recruited to VoxTox head and neck from May 2017 onwards did so within the Minot-OAR sub-study.

2.3.2.2 Inclusion and exclusion criteria

The VoxTox study was open to patients with prostate cancer, head and neck cancer, and low-grade primary malignancies of the CNS [9]. These sites were chosen based on a combination of existing data suggesting potentially meaningful differences between planned and delivered dose [30-33], data showing clear evidence of clinical benefit with dose escalation to target [34-37], and reduction to OARs [38], and departmental workload on the TomoTherapy systems. Inclusion and exclusion criteria were designed to be as broad as possible, to maximise recruitment potential, but also to ensure data heterogeneity in terms of both toxicity and OAR dose, to try and improve the discriminatory potential of dose-toxicity models.

Inclusion and exclusion criteria listed in Table 2.2 below are those that are pertinent for HNC patients, not those for the full study. All subsequent discussion pertaining to study conduct and recruitment within this dissertation will focus on the H&N cohort.

Inclusion Criteria	Exclusion Criteria
• Age > 18 years • Malignant or benign tumours of the head & neck. • Patients already treated, or suitable for treatment with radical intent, and daily IG, on the TomoTherapy platform, or equivalent technology. • Suitable for the 5 year follow up schedule. • Adequate cognitive ability to participate in interviews, and complete necessary forms and questionnaires. • Written informed consent.	• Previous radiotherapy to the treated area.

Table 2.2: Eligibility criteria for VoxTox head & neck (Adapted from tables in the study protocol, and the PhD dissertation of Dr Jessica Scaife)[2].

2.3.2.3 Toxicity assessment – clinical reporting forms

The core objective of the VoxTox study was to investigate the link between delivered dose to OARs, and toxicity. It was therefore critical to design a strategy to comprehensively capture the side effects experienced by patients, and that such data be comparable to results from clinical trials, and elsewhere in the literature. A number of different systems for assessing and

[2] PhD Thesis: Dr. Jessica Scaife, Peterhouse College, University of Cambridge, 2015.

quantifying late toxicity in HNC patients have been reported in the literature, validated, and are in widespread use. The most commonly used tools are listed below:

- Physician rated scores:
    - Radiation Therapy Oncology Group - RTOG [39].
    - Late Effects Normal Tissue, Subjective, Objective, Management - LENT/SOM [40].
    - Common Terminology Criteria for Adverse Events - CTCAE [41]

- Patients reported outcomes & quality of life tools:
    - Quality of life questionnaire - QLQ-C30 [42], and QLQ-H&N35 [43].
    - M.D. Anderson Dysphagia Inventory - MDADI [44].
    - Modified Xerostomia Questionnaire - XQ [38].

The task of capturing information in a frame of reference that would permit comparison with as many comparator studies as possible was compounded by the fact that many trials, which were either in progress, or had recently reported at the time the VoxTox protocol was being written, were not completely consistent in the tools they used to assess late toxicity, as shown in Table 2.3.

Trial	Year of publication (ref)	Scoring systems used for late toxicity:
Parsport	2011 [38]	RTOG, LENT/SOM, QLQ-C30, QLQ-H&N35, Modified-XQ.
ART-DECO	2017 [45, 46]	CTCAEv4, LENT/SOM, QLQ-C30, QLQ-H&N35, MDADI
COSTAR	2018 [14]	CTCAEv3, LENT/SOM, QLQ-C30, QLQ-H&N35.
De-ESCALaTE HPV	2019 [24]	CTCAEv4, QLQ-C30, QLQ-H&N35, MDADI
DARS	Yet to report [47]	CTCAEv4, LENT/SOM, MDADI
NIMRAD	Yet to report [13]	CTCAEv4, LENT/SOM, QLQ-C30, QLQ-H&N35, MDADI

Table 2.3: Validated toxicity assessment tools used in recent UK-based clinical trials of radiotherapy for HNC patients. Adapted from similar table in the PhD dissertation of Dr Jessica Scaife.[3]

[3] PhD Thesis: Dr Jessica Scaife, Peterhouse College, University of Cambridge, 2015.

Colleagues within the VoxTox group (Jessica Scaife and Amy Bates), therefore designed clinical reporting forms (CRFs), based on a combination of all of these toxicity assessment schemas, in such a way that completed CRFs contained sufficient information to populate as many relevant scoring systems as feasible. These CRFs were deployed electronically in the radiotherapy data management platform Mosaiq, in accordance with GCP standards. All raw toxicity data from the CRFs for the duration of the VoxTox study were also stored in the Mosaiq system. Baseline, acute, and late toxicity CRFs used for HNC patients recruited to VoxTox are shown in Appendix A3.1 – A3.3.

Scoring conformity between observers using the CRFs was also assessed and found to be satisfactory. Details of this work were undertaken before I joined the group, and are explained in detail in the PhD thesis of Dr Jessica Scaife.[4] I undertook all work to decode data collected on CRFs back into validated scores; this work is described in detail in Chapter 3.

2.3.2.4 Follow-up protocol

Patients recruited to the consolidation cohort underwent a full assessment of comorbidities, toxicities and relevant clinical data at baseline. The research radiographer team did this for the vast majority of patients. Acute toxicity assessments were mainly done by clinicians, and were done weekly during treatment, and at weeks 4, 8 and 12 post-treatment. Late toxicity and quality of life data were recorded at 6, 12, 18, 24, 36, 48 and 60 months post treatment. The research radiographer team undertook the majority of these assessments. Wherever clinic room and IT availability allowed, responses at all assessments were recorded directly into electronic versions of the CRF's within Mosaiq. When this was not possible, paper copies of the CRFs were used, and responses were subsequently uploaded into Mosaiq. Patients in the discovery cohort were subject to the same follow up schema, from the point at which they consented to participate in the study. A summary of the schedule of assessments is shown below in Figure 2. 4.

2.3.2.5 Imaging specifics

2.3.2.5.1 VoxTox MVCT protocol

Technical specifics regarding the MVCT scans done for image guidance purposes for H&N patients at Addenbrooke's Hospital are described in section 2.2.1. In order to minimise scan

[4] PhD Thesis: Dr Jessica Scaife, Peterhouse College, University of Cambridge, 2015.

time and facilitate the policy of daily IG for all patients, the standard FoV for non-study patients was restricted to the region covered by the PTV [8].

For patients recruited to the VoxTox study, the IG-MVCT FoV was extended superiorly and inferiorly, to include all parotid and submandibular gland tissue.

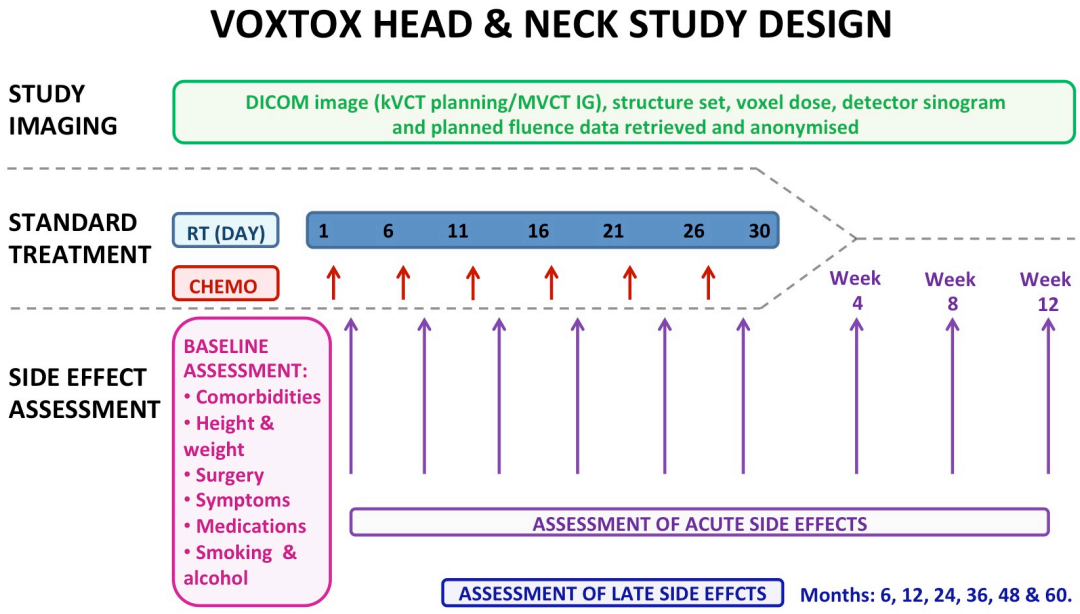


Figure 2.4: Summary timeline schema for VoxTox head and neck.

This FoV was also sufficient to include the full superior and middle pharyngeal constrictor muscle, supraglottic larynx and oral cavity in the vast majority of scans. Unfortunately the FoV was not sufficient to include all of inferior pharyngeal constrictor muscle in some patients.

Extending the FoV of MVCT scans did result in a very slight increase in both scan time, and radiation dose for patients recruited to the study. Previous work from group members showed a linear relationship between scan length and scan time ($R^2 = 0.54$), such that every additional FoV cm increased scan time by approximately 8 seconds [8]. Further work used relative weighting factors from the International Commission on Radiation Protection (ICRP) to estimate the effective dose of both standard and VoxTox study MVCT scans [48]. Using these data, it was estimated that the effective dose per-fraction with a standard IG scan was 0.32mSv, resulting in a total dose of 11.2mSv for a patient undergoing 35 fractions. Equivalent results for patients with extended FoV scans were 0.35 and 12.2mSv respectively. These correspond to respective lifetime risks of second malignancy of 1 in 1786 (0.056%) and 1 in

1639 (0.061%) respectively.[5] Therefore, the increased risk of second malignancy directly attributable to the extra image dose in the study was 0.005% (0.061 – 0.056), and the PIS for patients recruited to VoxTox H&N quoted an increase risk of 1 in 20,000 ((1/0.005) x 100). Examples of both standard and extended FoV MVCT images are shown in Figure 2.5.

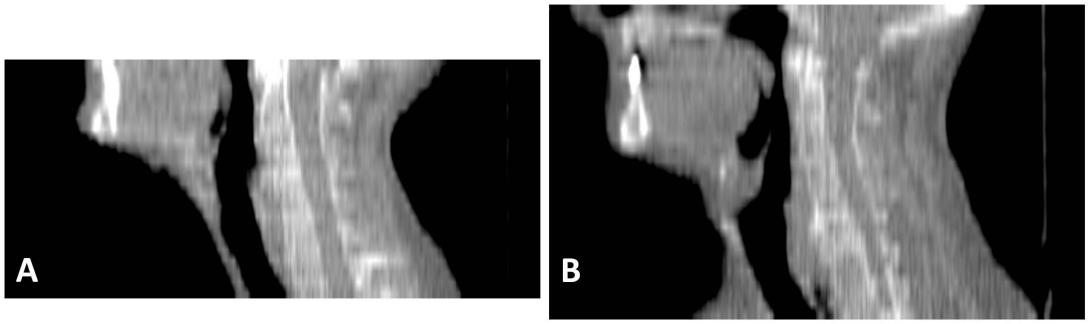


Figure 2.5 (A&B): Field of view of TomoTherapy MVCT IG scans. A – Standard protocol, B – VoxTox study protocol, note the extended FoV, especially in the cranial direction.

2.3.2.5.2 Imaging, dose, and treatment delivery data retrieval

In order to calculate delivered dose, it was necessary to have a minimum dataset for each patient. This comprised the original planning CT scan, the structure set used for treatment planning, the treatment plan and planned dose cube, all available IG-MVCT images with patient position corrections, and daily delivered fluence data in the form of delivery sinograms. At the onset of the VoxTox study, this data was stored in a variety of formats in Accuray (TomoTherapy vendor) electronic archives. Much of this data was not in DICOM format, and not easily compatible with other software platforms for working with imaging and radiotherapy dose. Marina Romanchikova, a post-doctoral researcher recruited to the group at the beginning of the VoxTox study developed a fully automated pipeline for extracting these data from archive, converting them into DICOM-RT format, and ensuring that all file headers were fully anonymised. This methodology was subsequently published [49].

An important implication of this workflow is that all the anonymised planning CT scans that were available for further work within the VoxTox framework, came from within the Accuray archive. During the period of time covered by the conduct of the study, the routine clinical workflow was for the TomoTherapy TPS to generate a planned dose cube with a 256 x 256 pixel matrix [50], in order to reduce computation time during treatment plan optimisation. Therefore, all clinical planning CT imaging data was also downsampled to the same voxel size

[5] PhD Thesis: Dr Jessica Scaife, Peterhouse College, University of Cambridge, 2015.

(2.148 x 2.148 x 3mm) on import into the TomoTherapy TPS, before treatment planning began. Therefore, because the workflow for thorough automated retrieval and anonymisation of kVCT planning scans within VoxTox depended on interrogation of the Accuray archive, the planning CTs used in this work have a voxel size of 2.148 x 2.148 x 3mm.

Once retrieved from archive, anonymised and converted into DICOM format, all imaging, dose, structure set and treatment delivery data was stored on the hospital radiotherapy research database. From there, data was sent electronically to the Cavendish Laboratory, where it was stored on the High Energy Physics cluster. Data flow, and control of computational processes within the Cavendish servers, was overseen by the Ganga task management system [51]. At my request, pertinent imaging, structure set, and planned dose data were also loaded into Prosoma research for manual segmentation. An overview of the flow of radiotherapy data in VoxTox head and neck is shown in schematic form in Figure 2.6.

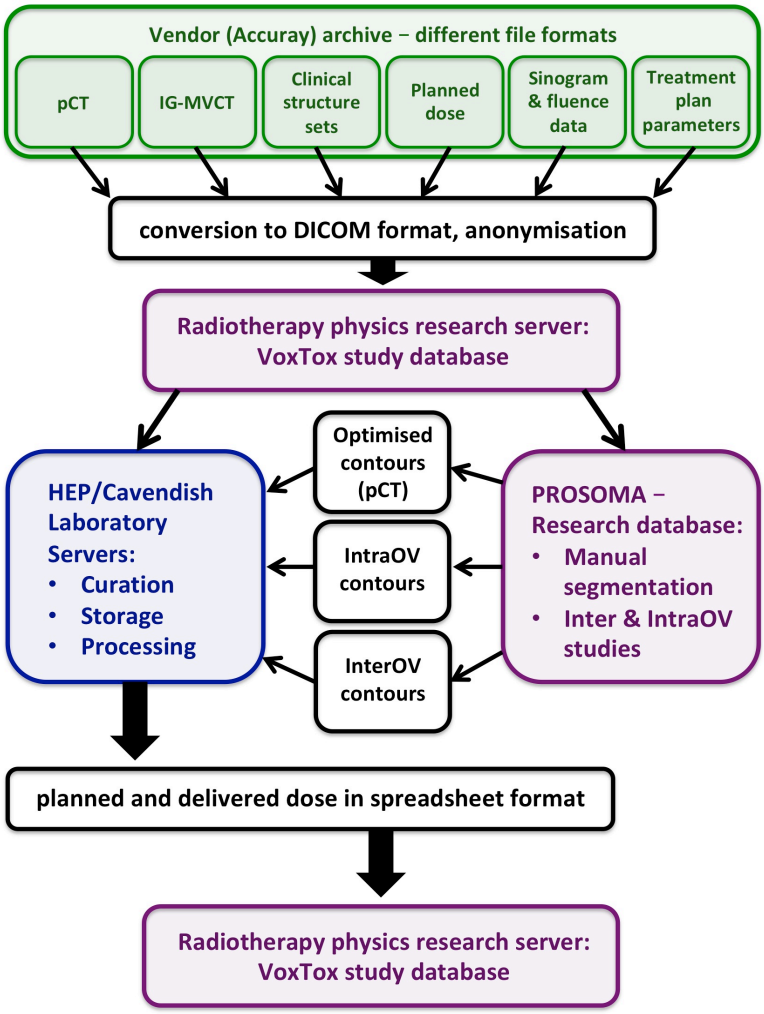


Figure 2.6: Radiotherapy data flow in VoxTox head and neck. Colours indicate location of data storage: Green – vendor archive, Purple – hospital medical physics curated databases, Blue – Cavendish laboratory, high energy physics cluster. Black boxes/text indicate data objects or processes. Arrows indicate flow of data.

2.3.3 OAR segmentation for VoxTox H&N

2.3.3.1 Defining a structure set of OARs

At the onset of the VoxTox study, the intention was to assess delivered dose to the spinal cord, parotid and submandibular glands, and the pharyngeal constrictor muscles as a single structure, as dose to these structures was known to predict the risk of important endpoints such as xerostomia, swallowing dysfunction and myelitis [38, 52-54]. On studying the literature as the project evolved, it became clear that dose to all the 'swallowing organs at risk – SWOARS' [12, 55], including the supraglottic larynx and individual constrictor muscles [54, 56, 57], as well as the minor salivary glands in the oral cavity [58, 59], is also important. Therefore, to produce planned dose metrics for these OARs, it was necessary to have high quality segmentations of these structures on the planning CT scans of VoxTox patients. The finalised 'list' of necessary structures, hereafter referred to as 'Full OARs' was:

- Right and left parotid glands
- Right and left submandibular glands
- Spinal cord
- Superior pharyngeal constrictor muscle
- Middle pharyngeal constrictor muscle
- Oral cavity
- Supraglottic larynx.

2.3.3.2 Contour quality control

Clinical structure sets were retrieved as part of the VoxTox pipeline as described in section 2.3.5.3.2, and loaded along with (downsampled) planning CT scans into the Research database of the departmental segmentation software – Prosoma. However, these segmentations could not be used for the project. Firstly, there was significant heterogeneity in which OARs had been segmented for each case, and in the nomenclature used. Second, many cases had automated OAR contours from the SPICE tool [60], which could have led to errors in DVH data. Thirdly, knowledge of the importance of the constrictor muscles, oral cavity and supraglottic larynx as normal tissue structures evolved over the course of the study, and they were not segmented as part of standard clinical protocols for much of its duration. Finally, clinical contours came from a number of different clinicians, and it is generally accepted that inter-observer variability in contouring is consistently greater than intra-observer variability [61-64]. Therefore, to ensure the most accurate, complete, and homogeneous segmentations, I

undertook to produce the full structure set detailed above for all patients in VoxTox head and neck.

To ensure that this methodology was robust, the following steps were taken to quality control my contours. Firstly, care was taken to segment in accordance with standard consensus guidelines and atlases [11, 12]. Second, I attended an ESTRO contouring workshop for H&N OARs. As preparation for this workshop, I was required to provide a full set of manual delineations of H&N OARs for a case of T3N1 (TNM v8) SCC of the right tonsil, using the web-based FALCON contouring tool [65]. In the conduct of the course, important anatomical principles were explored, and dice similarity metrics (DSC, [66]) of my contours relative to the workshop faculty were produced. Finally, during the course of my research, I established a collaboration with Professor Hans Langendijk's group from University Medical Centre Groningen. To validate the quality of my segmentations within this project, my inter-observer contouring variability relative to the Groningen group was measured. Three anonymised high-resolution planning CTs, with gold standard segmentations of the structure set listed above, were transferred electronically from Groningen to Cambridge, and uploaded into Prosoma Research for contouring. Concordance between my contours and Groningen gold-standard segmentations was assessed with DSC and mean distance to agreement (MDA, in mm) [67].

It was also important to quantify intra-observer contouring variability on downsampled planning scans. To do this, I undertook repeat segmentations of the original structure set, for a sample of 5 patients, at an interval of 8-12 weeks. To measure inter-observer variability, I recruited 4 H&N consultants from the department to contribute segmentations to an inter-observer contouring study. Participants were asked to follow a specific protocol (Appendix A4.1) and to contour 2 cases. Agreement between observers was assessed with DSC and MDA.

2.4 Results

2.4.1 Recruitment to VoxTox H&N

As the VoxTox study was not designed to test a specific intervention, no power calculation was needed, and no specific sample size objective was defined. On the contrary, as a longitudinal observational study, the designers aimed to recruit as many patients as possible within each anatomical subsite, to maximise the validity of whatever observations were made. Nonetheless, in order to consider resource implications both for the study infrastructure and

the clinical department, the protocol authors did calculate predicted recruitment rates. In addition, a pre-specified 'action level' for slow recruitment was set, to allow the study CI to investigate potential barriers to recruitment. Therefore, the protocol authors estimated a successful recruitment rate of 80%, and an overall 5YS to 60%, for VoxTox H&N, with an action level of 2/3rds of projected recruitment. Quarterly recruitment to both Discovery and Consolidation cohorts, as well as projected and actual cumulative recruitment, and pre-specified action levels, are shown in Figure 2.7.

2.4.2 VoxTox H&N - full cohort details

The baseline demographic characteristics, disease subsites, staging, dose prescription, treatment approach, and use of systemic therapy of all patients who consented to participate in VoxTox head and neck is shown in Table 2.4.

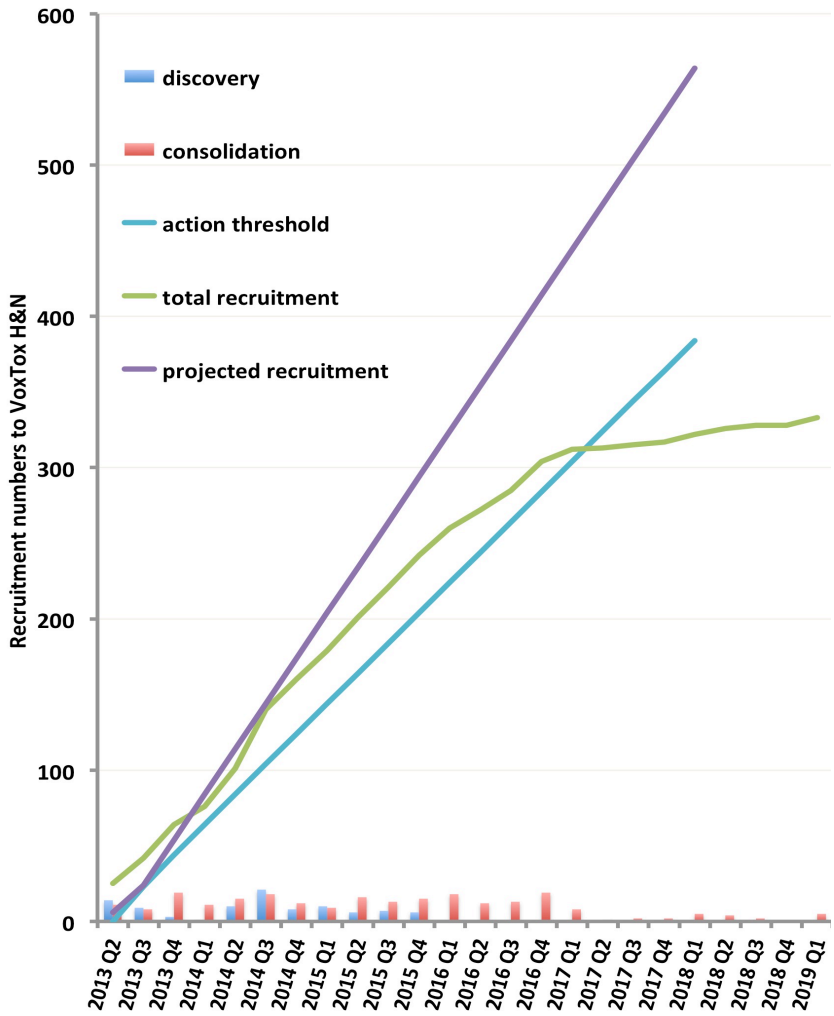


Figure 2.7: Patient recruitment numbers to VoxTox head and neck. Quarterly recruitment to Discovery (blue) and Consolidation (red) cohorts are shown as a clustered bar chart. Cumulative recruitment (green), as well as projected and threshold recruitment numbers are shown as lines.

Characteristic	Number
Age:	58.6 (10.5)*
Cohort:	
Discovery	95
Consolidation	224
Minot-OAR	18
Gender:	
Male	273 (81%)
Female	64 (19%)
Primary Disease:	
SCC	305 (90.5%)
Oropharynx	190 (56.4%)
Oral cavity	35 (10.4%)
Larynx	28 (8.3%)
Sinus	11 (3.3%)
Unknown primary	11 (3.3%)
Nasopharynx	10 (3.0%)
Hypopharynx	10 (3.0%)
Skin	8 (2.4%)
Nasal cavity	1 (0.3%)
Lacrimal gland	1 (0.3%)
Salivary Gland	26 (7.7%)
Others†	6 (1.8%)
Staging:	
T0-2	216 (64.1%)
T3-4	121 (35.9%)
N0-1	157 (46.6%)
N2+	180 (53.4%)
Dose/Fractionation:	
70/35	7 (2.1%)
68/34	35 (10.4%)
65/30	192 (57%)
60/30	90 (26.7%)
55/20	6 (1.8%)
50/20	7 (2.1%)
Neck Treatment:	
Bilateral	226 (67.1%)
Unilateral	78 (23.1%)
Primary only	33 (9.8%)
Systemic Therapy:	
Cisplatin	180 (53.4%)
Cetuximab	18 (5.3%)
None	139 (41.3%)

* mean age, standard deviation in parentheses

† Others include: 2 sarcoma, 2 paraganglioma, 2 basal cell carcinoma

Table 2.4: Baseline demographic, disease and treatment characteristics of all 337 patients recruited to VoxTox head and neck

2.4.3 Radiotherapy data for VoxTox H&N

Table 2.5 summarises the nature and volume of radiotherapy data – in DICOM format – that was transferred to the Cavendish Laboratory, and is currently stored on their servers. This includes radiotherapy imaging, treatment delivery and dose, and clinical segmentation data that was restored from archive as described in section 2.3.2.5.2. It also includes my repeated segmentations of the 'Full OARs' structure set, undertaken on downsampled planning CTs that had been reloaded into Prosoma research.

Data type	Number
No. of patient datasets:	319
Discovery	95
Consolidation	224
No. of pCTs:	328
Discovery	100
Consolidation	228
No. of MVCTs:	9112
Discovery	2888
Consolidation	6224
Mean no. of MVCTs/patient:	28.6
Discovery	30.3
Consolidation	27.8
Mean no. of slices/MVCT:	25.0
Discovery	18.5
Consolidation	28.0
No. of 'Full OAR' structure sets:	253
Discovery	38
Consolidation	215

* mean age, standard deviation in parentheses
† Others include: 2 sarcoma, 2 paraganglioma, 2 basal cell carcinoma

Table 2.5: Summary of radiotherapy datasets transferred from Addenbrooke's radiotherapy physics to the Cavendish Laboratory.

2.4.4 OAR segmentation on planning scans

2.4.4.1 Estro FALCON results.

I attended the Estro FALCON contouring workshop for H&N OARs on 27th April 2019. Data for agreement between my segmentations, and those of the course faculty are shown below in Figure 2.8.

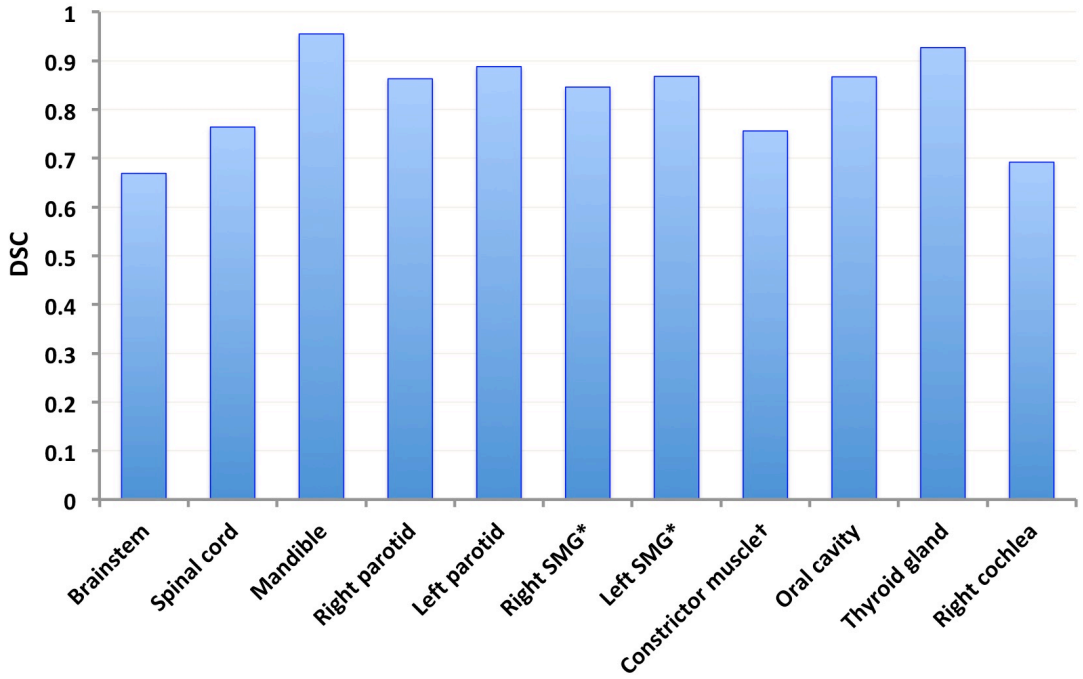


Figure 2.8: Concordance (dice similarity coefficient – DSC) between my segmentations and Estro FALCON H&N OAR contouring workshop gold-standard contours. (SMG* = submandibular gland, Constrictor muscles† - contoured as single structure, not individual muscles)

2.4.4.2 Swallowing OAR contour concordance – CITOR project

Concordance metrics between gold standard contours for swallowing OARs and spinal cord generated by clinicians at UMCG for the CITOR project, and segmentations done by me as QA for collaborative work on this study are shown in Figure 2.9 (A & B).

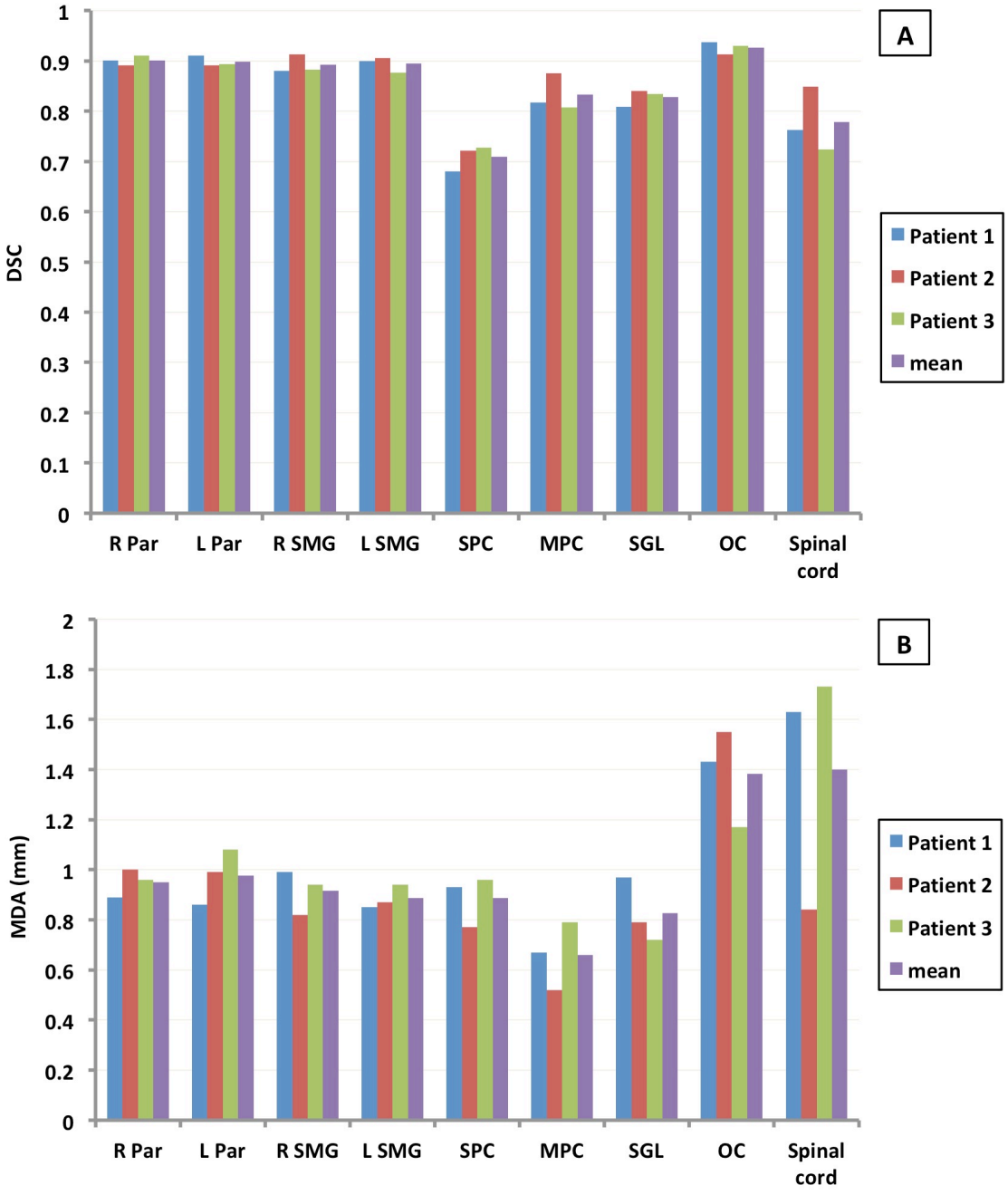


Figure 2.9 (A&B): Conformity metrics between my contours, and gold-standard segmentations from UMCG for the CITOR project. R Par & L Par = right and left parotid, R SMG & L SMG = right and left submandibular glands, SPC & MPC = superior and middle pharyngeal constrictor muscles, SGL = supraglottic larynx, OC = Oral cavity. A – Dice similarity coefficient (DSC), B – Mean distance to agreement (MDA), in mm.

2.4.4.3 Downsampled kVCT contouring – sample contours

A sample atlas, showing my segmentations of the 'Full OAR' structure set on a downsampled kVCT planning scan, is shown in Figure 2.10.

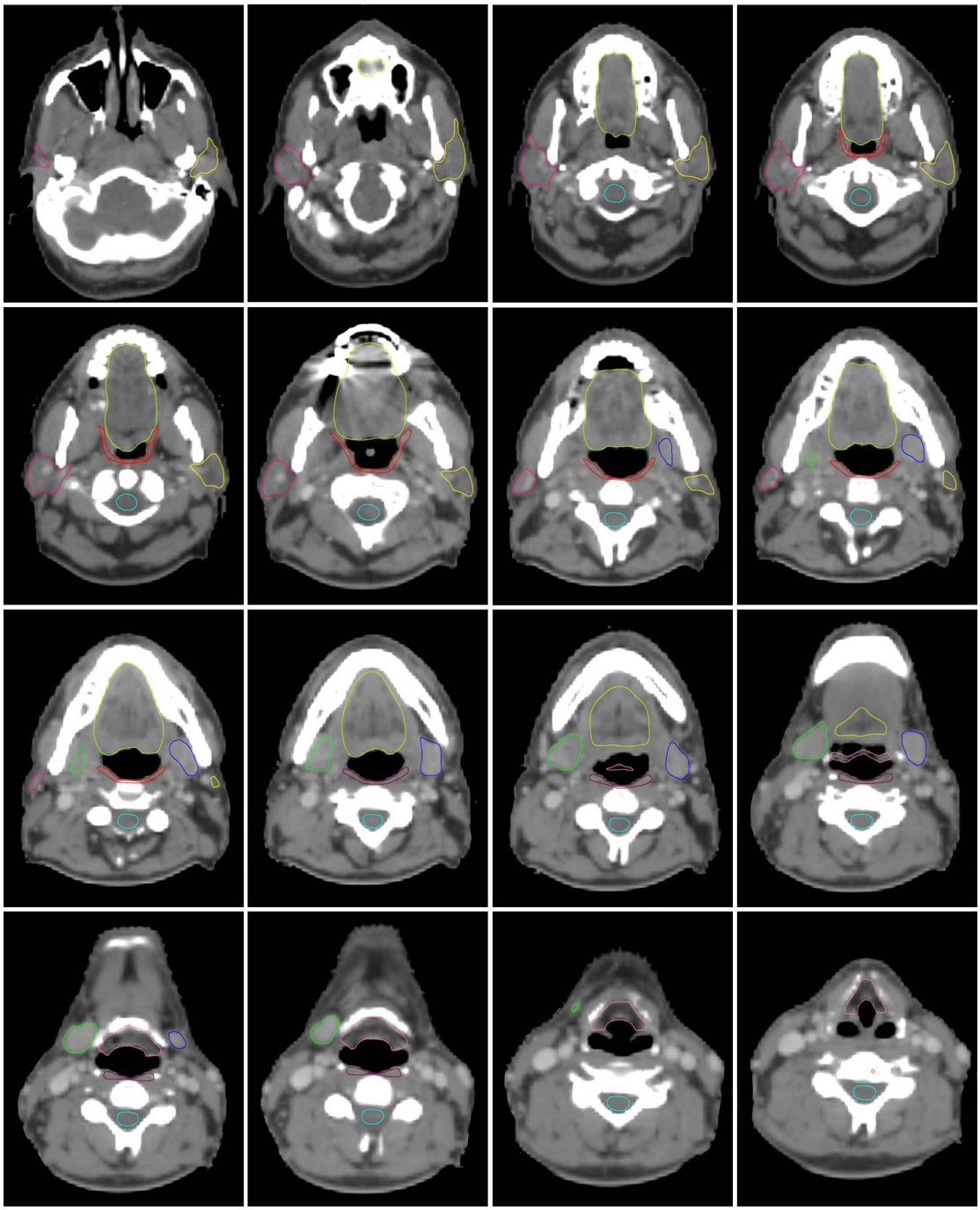


Figure 2.10: Sample atlas of Full OAR contours for a single patient. Colours: Right parotid – Ruby, Left parotid – Yellow, Oral cavity – lime, Spinal cord – cyan, Superior pharyngeal constrictor – red, Left SMG – Royal blue, Right SMG – green, Middle pharyngeal constrictor – purple, Supraglottic larynx – pink.

2.4.4.4 Downsampled kVCT contouring – Intra-observer variability

Results for intra-observer agreement between my contours on downsampled kVCT planning scans, and a second, repeat set of segmentations, are shown in Figure 2.11.

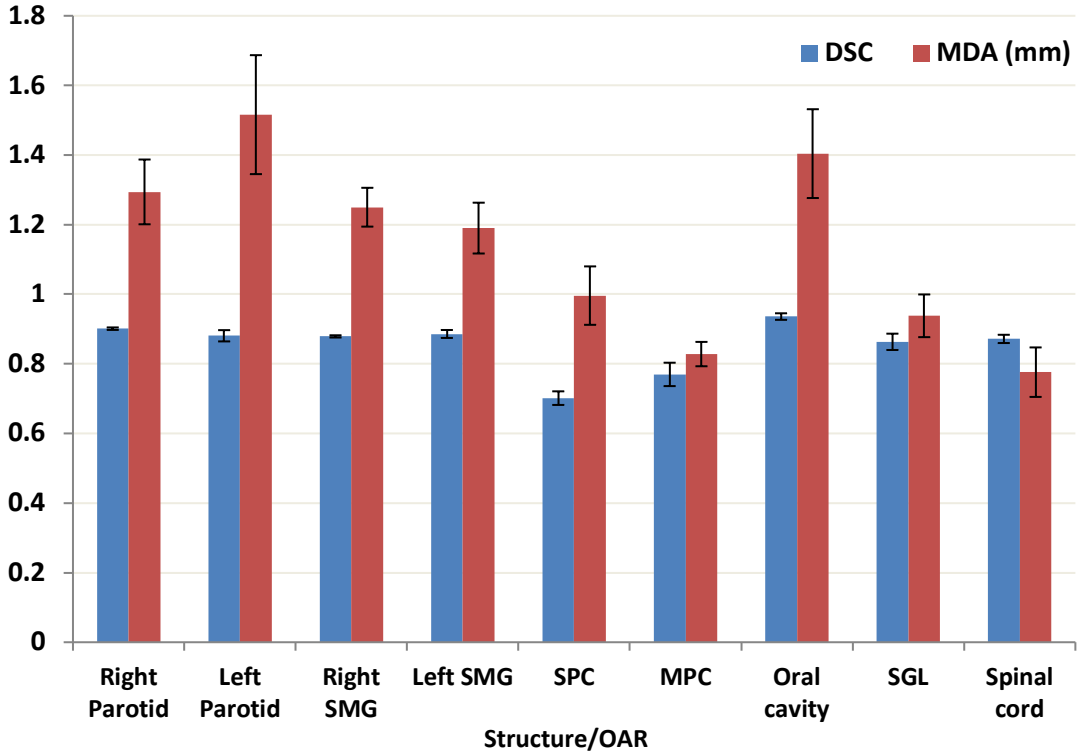


Figure 2.11: Intra-observer conformity metrics (Dice similarity coefficient – DSC, mean distance to agreement – MDA, in mm), of my repeat segmentations for 5 patients. Error bars are standard error. SMG = submandibular gland, SPC & MPC = superior and middle pharyngeal constrictor muscle, SGL = supraglottic larynx.

2.4.4.5 Downsampled kVCT contouring – Inter-observer variability

Summary results of inter-observer contouring variability between H&N site-specialists at Addenbrooke's, on downsampled kVCTs are shown in Figure 2.12. The data shown are mean DSC (A) and MDA (B) (with standard error) for all 8 structures, comparing contours between all 5 observers, for 2 patients. Therefore, results are mean values of 20 datapoints (10 combinations per patient, 2 patients). Results for each observer are shown in Tables 2.6 and 2.7. These data are mean values of 8 combinations for each individual observer.

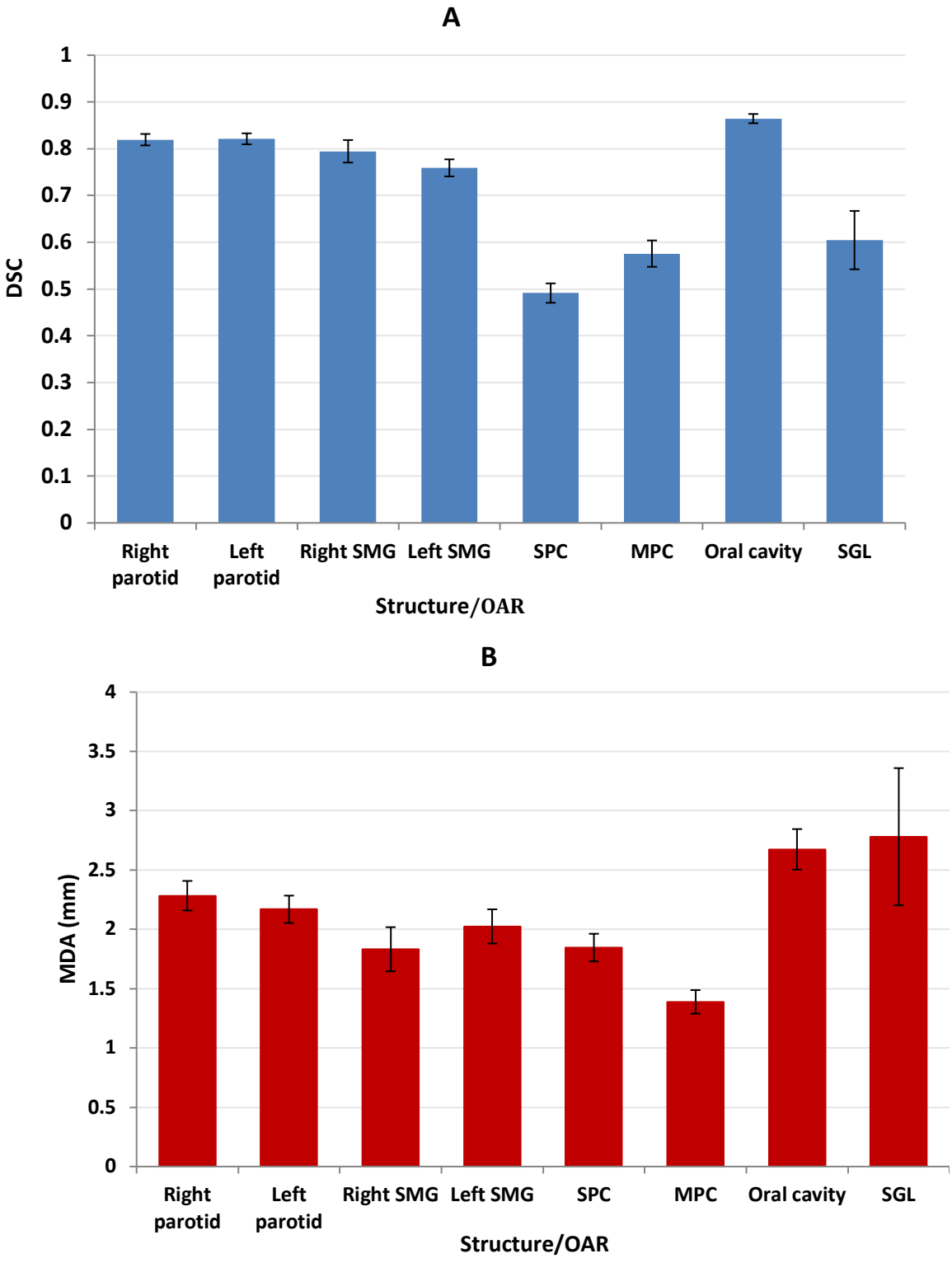


Figure 2.12 (A&B): Inter-observer conformity metrics for contours from 5 observers on downsampled kVCT scans. A - Dice similarity coefficient (DSC), B - mean distance to agreement (MDA, in mm). Error bars are standard error. SMG = submandibular gland, SPC & MPC = superior and middle pharyngeal constrictor muscle, SGL = supraglottic larynx.

	Observer 1	Observer 2	Observer 3	Observer 4	Observer 5
Right parotid	0.84 (0.01)	0.82 (0.01)	0.83 (0.01)	0.78 (0.01)	0.83 (0.01)
Left parotid	0.82 (0.01)	0.83 (0.01)	0.83 (0.01)	0.79 (0.01)	0.84 (0.01)
Right SMG	0.80 (0.03)	0.80 (0.02)	0.81 (0.02)	0.78 (0.03)	0.79 (0.03)
Left SMG	0.74 (0.02)	0.77 (0.02)	0.76 (0.02)	0.76 (0.02)	0.76 (0.02)
SPC	0.52 (0.01)	0.47 (0.02)	0.49 (0.02)	0.46 (0.02)	0.51 (0.02)
MPC	0.56 (0.03)	0.55 (0.04)	0.55 (0.03)	0.60 (0.02)	0.62 (0.02)
Oral cavity	0.86 (0.01)	0.86 (0.01)	0.87 (0.01)	0.86 (0.01)	0.87 (0.01)
SGL	0.66 (0.06)	0.65 (0.06)	0.62 (0.05)	0.45 (0.07)	0.64 (0.05)

Table 2.6: Inter-observer contouring variability (Dice similarity coefficient - DSC) for each individual observer (standard errors in parentheses). My results are observer 1. SMG = submandibular gland, SPC & MPC = superior and middle pharyngeal constrictor muscle, SGL = supraglottic larynx.

	Observer 1	Observer 2	Observer 3	Observer 4	Observer 5
Right parotid	2.12 (0.11)	2.27 (0.14)	2.12 (0.07)	2.60 (0.12)	2.31 (0.12)
Left parotid	2.19 (0.13)	2.11 (0.14)	2.05 (0.08)	2.38 (0.10)	2.12 (0.11)
Right SMG	1.74 (0.20)	1.72 (0.10)	1.66 (0.16)	2.09 (0.24)	1.96 (0.20)
Left SMG	1.96 (0.14)	1.83 (0.14)	1.93 (0.12)	2.17 (0.18)	2.23 (0.13)
SPC	1.69 (0.06)	1.89 (0.13)	1.77 (0.11)	2.04 (0.11)	1.84 (0.14)
MPC	1.45 (0.08)	1.39 (0.13)	1.50 (0.11)	1.35 (0.10)	1.26 (0.06)
Oral cavity	2.81 (0.19)	2.79 (0.22)	2.59 (0.15)	2.70 (0.17)	2.48 (0.13)
SGL	2.44 (0.59)	2.54 (0.51)	2.67 (0.57)	4.11 (0.74)	2.14 (0.24)

Table 2.7: Inter-observer contouring variability (Mean distance to agreement - MDA in mm) for each individual observer (standard errors in parentheses). My results are observer 1. SMG = submandibular gland, SPC & MPC = superior and middle pharyngeal constrictor muscle, SGL = supraglottic larynx.

2.5 Discussion

2.5.1 Study recruitment, and cohort characteristics

Figure 2.7 shows that for the first 3 years of the study, recruitment rates to VoxTox head and neck were good. Although they did not quite keep pace with the predicted recruitment line, they were well above the pre-specified action level. In the first quarter of 2016, the overall recruitment started to fall. This was because all eligible discovery cohort patients had been approached, and either agreed or declined to participate in the study. This effect can be appreciated from the individual recruitment bars for both discovery and consolidation patients. Recruitment rates for the consolidation cohort started to tail off in the first quarter of 2017 – a trend that continued through to formal recruitment cessation in April 2019. The main reason for this effect relates to staffing issues within the study group, and the Cambridge Trials Unit.

Nonetheless, the total study sample of 337 patients is a sizeable cohort with which to address the core hypotheses of VoxTox, and this thesis. By way of comparisons with recent UK clinical trials in HNC, it is considerably larger than the sample sizes of 94 in PARSPORT [38], 60 in ART-DECO [45, 46], or 110 in COSTAR [14]. It is similar to the 334 patients recruited to De-ESCALaTE HPV [24], although less than the target recruitment of 470 to NIMRAD [13]. It should be noted however that these were both multicentre studies.

However, it is important to emphasise that VoxTox is a single-centre, non-randomised longitudinal cohort study. Therefore, it is perhaps more relevant to compare the sample size with studies that were designed with a similar objective – specifically those looking in detail at radiation dose-toxicity effects, and NTCP modelling. The widely cited UMCG NTCP models of xerostomia [68], and moderate to severe dysphagia [55], were produced form sample sizes of 178 and 355 patients respectively. Another high-quality xerostomia model from a south-east Asian population used 206 patients [69], whilst training [70] and validation [71], of models predicting hypothyroidism after radiotherapy for HNC were derived from samples of 203 and 198 patients respectively. Thus the sample size of VoxTox head and neck compares favourably with high-quality reported studies that were carried out with thematically similar objectives and hypotheses.

It is well known that the demographics of patients who develop HNC are changing [72, 73]. Nonetheless, the demographics, baseline and treatment characteristics of the final cohort, as shown in Table 2.3, are representative of the broader population who undergo treatment at our

centre, and with what are presented in the literature. Ninety per cent of patients in VoxTox had SCC, matching what is classically reported [74]. The mean age of 59 is similar to average ages of 61 in the Beetz NTCP model (median), and 62 in the Wopken (mean) NTCP models [55, 68], although a little older than Lee's cohort [69]. The male female split in VoxTox head and neck – 81/19 - is also broadly similar to these two studies -71/29 and 76/24 respectively.

The disease characteristics are somewhat different however. In the Beetz paper, a much higher proportion of patients had larynx (37%) or oral cavity (35%) primaries, and 57% had T-stage 3-4 primary disease, compared with 36% in this cohort [68]. Only 39% of patients in the Beetz study received concomitant systemic therapy, compared with 58.7% in this cohort. Bilateral neck irradiation was also more common in the Beetz paper – 86% compared with 67% [68]. The Wopken cohort also had a much higher proportion of patients with larynx primaries (53%), although the T-stage proportions were more similar to VoxTox; 61% vs. 64% T0-2 [55]. However, a much greater proportion of patients in this work (66%) had N0-1 neck staging, compared with 47% in our cohort. The Wopken paper also reported more than three quarters of patients undergoing bilateral neck irradiation [55].

There are two key reasons for observed differences in baseline tumour and treatment characteristics. The first is the inclusion of non-SCC diagnoses in the VoxTox cohort, which increases the proportion of patients undergoing primary site only, or unilateral neck irradiation. The higher proportion of VoxTox patients with oropharynx primaries and more advanced neck nodal staging probably reflects the underlying evolution in the demographics of the disease. Most of the patients in these cohorts underwent treatment in the early-mid 2000's; the bulk of the VoxTox patients were treated in the period 2013-2016. It is therefore logical to conclude that the rising incidence of HPV-driven disease [72, 75], and the ongoing fall in smoking rates [76], are the main reasons for the observed differences in disease characteristics.

2.5.2 Radiotherapy data stored at the Cavendish Laboratory

It is apparent that the final cohort for VoxTox head and neck has 337 patients, whilst radiotherapy data for only 319 have undergone transfer to the Cavendish Laboratory. This is because the most recent bulk transfer of DICOM data, using the infrastructure described in section 2.3.2.5.2 was undertaken in May 2017. Patients recruited to participate in VoxTox after this point did so within the Minot-OAR substudy, as mentioned in section 2.3.2.1. Therefore, the radiotherapy, treatment toxicity, and MRI data for these patients are described and analysed separately, in the Chapter on this substudy (8), and these data are not included in the main body of the work as described in Chapters 3-7.

It is worth noting that there are 328 planning CTs for 319 patients. This because 9 patients (3%) in the cohort underwent adaptive replanning during treatment, with a new planning CT and thermoplastic shell requested, and a new treatment plan produced. Documentation of the reasons for these replans was available for 5 patients. Reasons included: anatomical and/or shape change – 4, inadequate PTV coverage – 2, weight loss – 1, higher dose to neurological OARs (spinal cord and brainstem) – 1.

The number of MVCTs per patient is slightly higher in the Discovery cohort, mainly due to the change in standard radical fractionation from 68/34 to 65/30 that occurred in November 2011. The difference is small and these data show that most MVCTs were available for most patients from both cohorts. What is more apparent is that the mean number of slices per patient is higher for consolidation patients. This is because Discovery patients were treated before the introduction of the VoxTox IG protocol, and therefore had shorter FoVs, as described in section 2.3.2.5.1.

There is also a discrepancy between the number of patients with radiotherapy data at the Cavendish Laboratory, and the number for whom I segmented the 'Full OAR' structure set. One important point here is the time consideration. Although I did not formally measure this, the typical time taken to open a relevant patient's dataset, segment the images, save and transfer the structure sets, and keep a contemporary database of progress was around 50-60 minutes. There was therefore an important trade-off between the statistical power associated with increasing the sample size, and the time taken to generate the contours. For this reason, the main focus was the Consolidation cohort, as these patients have complete baseline and late toxicity assessments. Contours were generated for 215 out of a possible 224 Consolidation patients. The remaining 9 were not segmented due to a combination of reasons including withdrawal from the study, and early post-treatment recurrence.

Contours were generated for 38 Discovery patients. Planning CTs for these patients were already uploaded and available in Prosoma Research when I started the project. Therefore, to start work on contouring protocols and data transfer workflows as described in this Chapter, to optimise methods for autosegmentation of MVCT images and computation of delivered dose using deformable image registration (DIR) as described in Chapter 4, and to increase the sample size for calculations looking exclusively at differences between planned and delivered dose as described in Chapter 6, these images were all segmented. However, because of the time constraints of this process, and the fact that these segmentations carried less scientific value than those for Consolidation patients, the decision was taken not to increase the sample size of Discovery patients beyond this number.

2.5.3 Planning scan contour quality control

Results from the Estro FALCON contouring workshop should be treated with caution, as they come from a single patient. Nonetheless, it is reassuring that for this case, Dice conformity metrics between my segmentations of key OARs including the parotid and submandibular glands, and the oral cavity, and gold standard provided by the Estro faculty, are all between 0.85 and 0.9, well above the threshold value of 0.8 for 'good agreement' [60].

Scores for the spinal cord and pharyngeal constrictor muscles were a little lower, 0.76 for both. For the spinal cord, this is because of differing interpretations of how far inferior to bring the contour. The gold standard contours came down 5cm below the inferior limit of the neck CTV, whilst my contour continued to the inferior limit of the image. The lower score for the pharyngeal constrictor muscle (contoured as single structure) is due in large part to the shape of the structure, and the nature of the DSC metric (this is discussed in more detail in Chapter 4). As the pharyngeal constrictor is a long, thin, U-shaped structure for much of its course, a small translation in the axial plane, can result in a substantial drop in DSC, for what visually appears to be a small discrepancy in the contours. A cartoon to illustrate this point is shown in Figure 2.13.

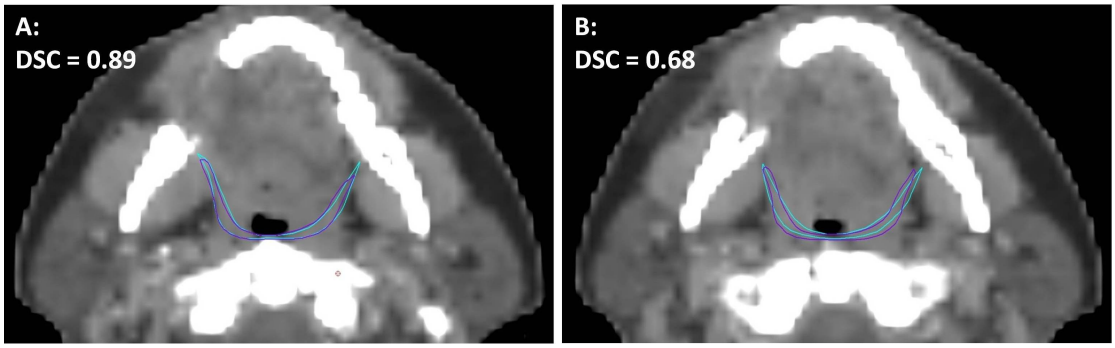


Figure 2.13 (A&B): DSC as a conformity metric for pharyngeal constrictor muscle contours. Caption A – the contours are almost super-imposed giving a DSC of 0.89. Caption B – the contours are very similar in terms of location, with similar surface areas (3cm^2 and 2.6cm^2), and distances between neighbouring points are small, but small translations lead to a much lower DSC score – 0.68.

Similar results were observed for inter-observer agreement with contours provided by UMCG. Mean DSC scores were around 0.9 for the parotid and submandibular glands, and oral cavity. They were also >0.8 for the middle pharyngeal constrictors and supraglottic larynx. Spinal cord results were a little less good, with a mean DSC of 0.78, although the result is very similar to previously published results for this structure [77]. Mean DSC for the superior pharyngeal constrictor muscle was 0.71, lower than any other structure. However mean MDA was joint

third lowest (0.89mm) emphasising the previous explanation, and the point illustrated in Figure 2.13. Overall these data compare favourably with previously published results [77-79].

Results for my own intra-observer contouring variability on downsampled kVCTs followed a very similar pattern. Mean DSC and MDA results for parotids and submandibular glands were 0.87-0.9 and 1.2-1.5mm respectively. DSC scores for oral cavity, spinal cord and supraglottic larynx were all above 0.8. As with both the Estro FALCON and UMCG data, DSC scores for the pharyngeal constrictors were slightly less good (SPC - 0.70, MPC - 0.77) but with acceptable MDA results of 1.0 and 0.83mm respectively.

Results of the intra-observer contouring variability study of H&N site-specialists at Addenbrooke's on downsampled kVCT scans, are a little less good than those seen with either the Estro FALCON, or UMCG concordance data. One reason for this is the lower image quality associated with the downsampling. Nonetheless, the data are generally reassuring. Mean DSC values across all 5 observers are near to, or above 0.8 for all of the larger OARs (parotid and submandibular glands, and oral cavity). Although DSC results for the smaller OARs (pharyngeal constrictor muscles and supraglottic larynx) are lower, this metric is less reliable for these structures for the reason outlined above. MDA data are more reassuring for both pharyngeal constrictor muscles (<2mm), but less so for the supraglottic larynx (mean MDA 2.78mm).

The explanation for this observation comes from the data in Tables 2.6 and 2.7, where it is apparent that conformity data for the SGL for Observer 4 are noticeably worse than those for other observers. On reviewing the contours themselves, this was because this observer misinterpreted the position of the arytenoid cartilages for 1 case, and started the SGL contour approximately 2cm cranial to the consensus slice of the other 4 observers. If results for observer 4 are removed, mean DSC and MDA for the SGL are 0.71 and 1.90mm respectively. More generally, 2 reassuring conclusions may be drawn from the data in Tables 2.6 and 2.7. Firstly, results for most OARs are very consistent between observers, and secondly that my data are comparable to site-specialist peers.

2.6 Conclusions

Between April 2013 and April 2019, 337 patients with head and neck cancer were recruited to VoxTox head and neck. The Consolidation cohort comprised 224, and these patients were recruited prospectively, with a full set of baseline demographic, disease, treatment and co-morbidity data collected at baseline. Due to the broad inclusion criteria, the cohort is representative of the wider population of patients with HNC treated at Addenbrooke's hospital, and is comparable to those that formed the basis of similar studies. Within the infrastructure of the VoxTox study, full radiotherapy datasets including planning and IG scans, structure sets and treatment plans were anonymised and transferred to the Cavendish Laboratory for curation. In order to provide an accurate estimate of planned dose, and as a first step in the process of calculating delivered dose, repeat manual contouring of a pre-defined set of structures of interest for 253 patients was done. Intra- and inter-observer variability experiments were undertaken to ensure the quality of these segmentations. Further details on the process by which delivered dose was calculated are described in Chapters 4 and 5, and Chapter 3 will examine the toxicity experienced by the patients in this cohort.

2.7 References:

[1] Mackie TR, Balog J, Ruchala K, Shepard D, Aldridge S, Fitchard E, et al. Tomotherapy. Semin Radiat Oncol. 1999;9:108-17.
[2] Mackie TR. History of tomotherapy. Phys Med Biol. 2006;51:R427-53.
[3] Burnet NG, Adams EJ, Fairfoul J, Tudor GS, Hoole AC, Routsis DS, et al. Practical aspects of implementation of helical tomotherapy for intensity-modulated and image-guided radiotherapy. Clin Oncol (R Coll Radiol). 2010;22:294-312.
[4] Welsh JS, Patel RR, Ritter MA, Harari PM, Mackie TR, Mehta MP. Helical tomotherapy: an innovative technology and approach to radiation therapy. Technol Cancer Res Treat. 2002;1:311-6.
[5] Schombourg K, Bochud F, Moeckli R. Stability of the Helical TomoTherapy Hi.Art II detector for treatment beam irradiations. J Appl Clin Med Phys. 2014;15:4897.
[6] Johnson TR, Krauss B, Sedlmair M, Grasruck M, Bruder H, Morhard D, et al. Material differentiation by dual energy CT: initial experience. Eur Radiol. 2007;17:1510-7.
[7] Yartsev S, Kron T, Van Dyk J. Tomotherapy as a tool in image-guided radiation therapy (IGRT): theoretical and technological aspects. Biomed Imaging Interv J. 2007;3:e16.
[8] Bates AM, Scaife JE, Tudor GS, Jena R, Romanchikova M, Dean JC, et al. Image guidance protocols: balancing imaging parameters against scan time. Br J Radiol. 2013;86:20130385.
[9] Burnet NG SJ, Romanchikova M, Thomas SJ, Bates AM, Wong E, Noble DJ, Shelley LEA, Bond SJ, Forman J, Hoole ACF, Barnett GC, Brochu FM, Simmons MPD, Jena R, Harrison K, Yeap PL, Drew A, Silvester E, Elwood P, Pullen H, Sultana A, Seah SYK, Wilson MZ, Russel SG, Benson RJ, Rimmer YL, Jefferies SJ, Taku N, Gurnell M, Powlson A, Schönlieb C-B, Cai X, Sutcliffe MPF, Parker MA. Applying physical science techniques and CERN technology to an unsolved problem in radiation treatment for cancer: the multidisciplinary 'VoxTox' research programme. CERN IdeaSquare Journal of Experimental Innovation. 2017;1:3-12.
[10] Gregoire V, Ang K, Budach W, Grau C, Hamoir M, Langendijk JA, 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;110:172-81.
[11] Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, 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;117:83-90.
[12] Christianen ME, Langendijk JA, Westerlaan HE, van de Water TA, Bijl HP. Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. Radiother Oncol. 2011;101:394-402.
[13] Thomson D, Yang H, Baines H, Miles E, Bolton S, West C, 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;26:344-7.
[14] Nutting CM, Morden JP, Beasley M, Bhide S, Cook A, De Winton E, et al. Results of a multicentre randomised controlled trial of cochlear-sparing intensity-modulated radiotherapy versus conventional radiotherapy in patients with parotid cancer (COSTAR; CRUK/08/004). Eur J Cancer. 2018;103:249-58.
[15] Owadally W, Hurt C, Timmins H, Parsons E, Townsend S, Patterson J, 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;15:602.
[16] Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer. Lancet. 2000;355:949-55.
[17] Pignon JP, le Maitre A, Maillard E, Bourhis J, Group M-NC. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. Radiother Oncol. 2009;92:4-14.
[18] Denis F, Garaud P, Bardet E, Alfonsi M, Sire C, Germain T, et al. Final results of the 94-01 French Head and Neck Oncology and Radiotherapy Group randomized trial comparing radiotherapy alone with concomitant radiochemotherapy in advanced-stage oropharynx carcinoma. J Clin Oncol. 2004;22:69-76.
[19] Cooper JS, Pajak TF, Forastiere AA, Jacobs J, Campbell BH, Saxman SB, et al. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. N Engl J Med. 2004;350:1937-44.

[20] Machtay M, Moughan J, Trotti A, Garden AS, Weber RS, Cooper JS, et al. Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: an RTOG analysis. J Clin Oncol. 2008;26:3582-9.
[21] Forastiere AA, Zhang Q, Weber RS, Maor MH, Goepfert H, Pajak TF, et al. Long-term results of RTOG 91-11: a comparison of three nonsurgical treatment strategies to preserve the larynx in patients with locally advanced larynx cancer. J Clin Oncol. 2013;31:845-52.
[22] Bonner JA, Harari PM, Giralt J, Azarnia N, Shin DM, Cohen RB, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. N Engl J Med. 2006;354:567-78.
[23] Gillison ML, Trotti AM, Harris J, Eisbruch A, Harari PM, Adelstein DJ, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. Lancet. 2019;393:40-50.
[24] Mehanna H, Robinson M, Hartley A, Kong A, Foran B, Fulton-Lieuw T, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. Lancet. 2019;393:51-60.
[25] Al-Sarraf M, LeBlanc M, Giri PG, Fu KK, Cooper J, Vuong T, et al. Chemoradiotherapy versus radiotherapy in patients with advanced nasopharyngeal cancer: phase III randomized Intergroup study 0099. J Clin Oncol. 1998;16:1310-7.
[26] Szturz P, Wouters K, Kiyota N, Tahara M, Prabhash K, Noronha V, et al. Weekly Low-Dose Versus Three-Weekly High-Dose Cisplatin for Concurrent Chemoradiation in Locoregionally Advanced Non-Nasopharyngeal Head and Neck Cancer: A Systematic Review and Meta-Analysis of Aggregate Data. Oncologist. 2017;22:1056-66.
[27] Noronha V, Joshi A, Patil VM, Agarwal J, Ghosh-Laskar S, Budrukkar A, et al. Once-a-Week Versus Once-Every-3-Weeks Cisplatin Chemoradiation for Locally Advanced Head and Neck Cancer: A Phase III Randomized Noninferiority Trial. J Clin Oncol. 2018;36:1064-72.
[28] Bauml JM, Vinnakota R, Anna Park YH, Bates SE, Fojo T, Aggarwal C, et al. Cisplatin Every 3 Weeks Versus Weekly With Definitive Concurrent Radiotherapy for Squamous Cell Carcinoma of the Head and Neck. J Natl Cancer Inst. 2018.
[29] Spreafico A, Huang SH, Xu W, Granata R, Liu CS, Waldron JN, et al. Impact of cisplatin dose intensity on human papillomavirus-related and -unrelated locally advanced head and neck squamous cell carcinoma. Eur J Cancer. 2016;67:174-82.
[30] Hatton JA, Greer PB, Tang C, Wright P, Capp A, Gupta S, et al. Does the planning dose-volume histogram represent treatment doses in image-guided prostate radiation therapy? Assessment with cone-beam computerised tomography scans. Radiother Oncol. 2011;98:162-8.
[31] Hansen EK, Bucci MK, Quivey JM, Weinberg V, Xia P. Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2006;64:355-62.
[32] Robar JL, Day A, Clancey J, Kelly R, Yewondwossen M, Hollenhorst H, et al. Spatial and dosimetric variability of organs at risk in head-and-neck intensity-modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2007;68:1121-30.
[33] Loo H, Fairfoul J, Chakrabarti A, Dean JC, Benson RJ, Jefferies SJ, et al. Tumour shrinkage and contour change during radiotherapy increase the dose to organs at risk but not the target volumes for head and neck cancer patients treated on the TomoTherapy HiArt system. Clin Oncol (R Coll Radiol). 2011;23:40-7.
[34] Dearnaley DP, Jovic G, Syndikus I, Khoo V, Cowan RA, Graham JD, et al. Escalated-dose versus control-dose conformal radiotherapy for prostate cancer: long-term results from the MRC RT01 randomised controlled trial. Lancet Oncol. 2014;15:464-73.
[35] Heemsbergen WD, Al-Mamgani A, Slot A, Dielwart MF, Lebesque JV. Long-term results of the Dutch randomized prostate cancer trial: impact of dose-escalation on local, biochemical, clinical failure, and survival. Radiother Oncol. 2014;110:104-9.
[36] Fu KK, Pajak TF, Trotti A, Jones CU, Spencer SA, Phillips TL, et al. A Radiation Therapy Oncology Group (RTOG) phase III randomized study to compare hyperfractionation and two variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: first report of RTOG 9003. Int J Radiat Oncol Biol Phys. 2000;48:7-16.
[37] Overgaard J, Hansen HS, Specht L, Overgaard M, Grau C, Andersen E, et al. Five compared with six fractions per week of conventional radiotherapy of squamous-cell carcinoma of head and neck: DAHANCA 6 and 7 randomised controlled trial. Lancet. 2003;362:933-40.
[38] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol. 2011;12:127-36.

[39] Cox JD, Stetz J, Pajak TF. Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC). Int J Radiat Oncol Biol Phys. 1995;31:1341-6.
[40] LENT SOMA tables. Radiother Oncol. 1995;35:17-60.
[41] Trotti A, Colevas AD, Setser A, Rusch V, Jaques D, Budach V, et al. CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol. 2003;13:176-81.
[42] Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85:365-76.
[43] Bjordal K, de Graeff A, Fayers PM, Hammerlid E, van Pottelsberghe C, Curran D, et al. A 12 country field study of the EORTC QLQ-C30 (version 3.0) and the head and neck cancer specific module (EORTC QLQ-H&N35) in head and neck patients. EORTC Quality of Life Group. Eur J Cancer. 2000;36:1796-807.
[44] Chen AY, Frankowski R, Bishop-Leone J, Hebert T, Leyk S, Lewin J, et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: the M. D. Anderson dysphagia inventory. Arch Otolaryngol Head Neck Surg. 2001;127:870-6.
[45] Nutting C, Morden J, Bernstein D, Beasley M, Cosgrove V, Fisher S, et al. ART DECO (CRUK/10/018): dose escalated vs standard dose IMRT in locally advanced head and neck cancer. Radiotherapy and Oncology. 2017;122:24-.
[46] Gujral DM, Miah AB, Bodla S, Richards TM, Welsh L, Schick U, et al. Final long-term results of a phase I/II study of dose-escalated intensity-modulated radiotherapy for locally advanced laryngo-hypopharyngeal cancers. Oral Oncol. 2014;50:1089-97.
[47] Petkar I, Rooney K, Roe JW, Patterson JM, Bernstein D, Tyler JM, 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;16:770.
[48] The 2007 Recommendations of the International Commission on Radiological Protection. ICRP publication 103. Ann ICRP. 2007;37:1-332.
[49] Romanchikova M, Harrison K, Burnet NG, Hoole AC, Sutcliffe MP, Parker MA, et al. Automated customized retrieval of radiotherapy data for clinical trials, audit and research. Br J Radiol. 2018;91:20170651.
[50] Thomas SJ, Eyre KR, Tudor GS, Fairfoul J. Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube. Med Phys. 2012;39:160-7.
[51] Moscicki JT, Brochu F, Ebke J, Egede U, Elmsheuser J, Harrison K, et al. GANGA: A tool for computational-task management and easy access to Grid resources. Comput Phys Commun. 2009;180:2303-16.
[52] Deasy JO, Moiseenko V, Marks L, Chao KS, Nam J, Eisbruch A. Radiotherapy dose-volume effects on salivary gland function. Int J Radiat Oncol Biol Phys. 2010;76:S58-63.
[53] Kirkpatrick JP, Van-Der-Kogel AJ, Schultheiss TE. Radiation Dose-Volume Effects in the Spinal Cord. Int J Radiation Oncology Biol Phys. 2010;76:S42-9.
[54] Eisbruch A, Schwartz M, Rasch C, Vineberg K, Damen E, Van As CJ, 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;60:1425-39.
[55] Wopken K, Bijl HP, van der Schaaf A, van der Laan HP, Chouvalova O, Steenbakkers RJ, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol. 2014;113:95-101.
[56] Jensen K, Lambertsen K, Grau C. Late swallowing dysfunction and dysphagia after radiotherapy for pharynx cancer: frequency, intensity and correlation with dose and volume parameters. Radiother Oncol. 2007;85:74-82.
[57] Dirix P, Abbeel S, Vanstraelen B, Hermans R, Nuyts S. Dysphagia after chemoradiotherapy for head-and-neck squamous cell carcinoma: dose-effect relationships for the swallowing structures. Int J Radiat Oncol Biol Phys. 2009;75:385-92.
[58] Hoebers F, Yu E, Eisbruch A, Thorstad W, O'Sullivan B, Dawson LA, et al. A pragmatic contouring guideline for salivary gland structures in head and neck radiation oncology: the MOIST target. Am J Clin Oncol. 2013;36:70-6.

[59] Eisbruch A, Kim HM, Terrell JE, Marsh LH, Dawson LA, Ship JA. Xerostomia and its predictors following parotid-sparing irradiation of head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2001;50:695-704.
[60] Thomson D, Boylan C, Liptrot T, Aitkenhead A, Lee L, Yap B, et al. Evaluation of an automatic segmentation algorithm for definition of head and neck organs at risk. Radiat Oncol. 2014;9:173.
[61] Breen SL, Publicover J, De Silva S, Pond G, Brock K, O'Sullivan B, et al. Intraobserver and interobserver variability in GTV delineation on FDG-PET-CT images of head and neck cancers. Int J Radiat Oncol Biol Phys. 2007;68:763-70.
[62] Alterio D, Ciardo D, Preda L, Argenone A, Caspiani O, Micera R, 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. Radiother Oncol. 2014;112:337-42.
[63] Piotrowski T, Gintowt K, Jodda A, Ryczkowski A, Bandyk W, Ba KB, et al. Impact of the Intra- and Inter-observer Variability in the Delineation of Parotid Glands on the Dose Calculation During Head and Neck Helical Tomotherapy. Technol Cancer Res Treat. 2015;14:467-74.
[64] Fiorino C, Reni M, Bolognesi A, Cattaneo GM, Calandrino R. Intra- and inter-observer variability in contouring prostate and seminal vesicles: implications for conformal treatment planning. Radiother Oncol. 1998;47:285-92.
[65] Eriksen JG, Salembier C, Rivera S, De Bari B, Berger D, Mantello G, et al. Four years with FALCON - an ESTRO educational project: achievements and perspectives. Radiother Oncol. 2014;112:145-9.
[66] Dice L. Measures of the Amount of Ecologic Association Between Species. Ecology. 1945;26:297-302.
[67] Chalana V, Kim Y. A methodology for evaluation of boundary detection algorithms on medical images. IEEE Trans Med Imaging. 1997;16:642-52.
[68] Beetz I, Schilstra C, van der Schaaf A, van den Heuvel ER, Doornaert P, van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. Radiother Oncol. 2012;105:101-6.
[69] Lee TF, Chao PJ, Ting HM, Chang L, Huang YJ, Wu JM, et al. Using multivariate regression model with least absolute shrinkage and selection operator (LASSO) to predict the incidence of Xerostomia after intensity-modulated radiotherapy for head and neck cancer. PLoS One. 2014;9:e89700.
[70] Ronjom MF, Brink C, Bentzen SM, Hegedus L, Overgaard J, Johansen J. Hypothyroidism after primary radiotherapy for head and neck squamous cell carcinoma: normal tissue complication probability modeling with latent time correction. Radiother Oncol. 2013;109:317-22.
[71] Ronjom MF, Brink C, Bentzen SM, Hegedus L, Overgaard J, Petersen JB, et al. External validation of a normal tissue complication probability model for radiation-induced hypothyroidism in an independent cohort. Acta Oncol. 2015;54:1301-9.
[72] Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. J Clin Oncol. 2015;33:3235-42.
[73] Dahlstrom KR, Bell D, Hanby D, Li G, Wang LE, Wei Q, et al. Socioeconomic characteristics of patients with oropharyngeal carcinoma according to tumor HPV status, patient smoking status, and sexual behavior. Oral Oncol. 2015;51:832-8.
[74] Gregoire V, Lefebvre JL, Licitra L, Felip E, Group E-E-EGW. Squamous cell carcinoma of the head and neck: EHNS-ESMO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2010;21 Suppl 5:v184-6.
[75] Dahlstrand H, Nasman A, Romanitan M, Lindquist D, Ramqvist T, Dalianis T. Human papillomavirus accounts both for increased incidence and better prognosis in tonsillar cancer. Anticancer Res. 2008;28:1133-8.
[76] Sturgis EM, Cinciripini PM. Trends in head and neck cancer incidence in relation to smoking prevalence: an emerging epidemic of human papillomavirus-associated cancers? Cancer. 2007;110:1429-35.
[77] Nelms BE, Tome WA, Robinson G, Wheeler J. Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer. Int J Radiat Oncol Biol Phys. 2012;82:368-78.
[78] Brouwer CL, Steenbakkers RJ, van den Heuvel E, Duppen JC, Navran A, Bijl HP, et al. 3D Variation in delineation of head and neck organs at risk. Radiat Oncol. 2012;7:32.
[79] Lim TY, Gillespie E, Murphy J, Moore KL. Clinically Oriented Contour Evaluation Using Dosimetric Indices Generated From Automated Knowledge-Based Planning. Int J Radiat Oncol Biol Phys. 2019;103:1251-60.

Chapter 3 – Toxicity Results

3.1 Overview

Despite technical advances in treatment delivery, patients undergoing radiotherapy for HNC are still at significant risk of moderate to severe late toxicity. In order to adequately test the core hypothesis of this thesis – that delivered dose predicts toxicity better than planned dose – it is important to show that the toxicity data on which this analysis has been based were collected in a reliable and robust manner, and that the results are broadly in line with previously published data. The methodology by which raw toxicity data gathered within the VoxTox study was processed and analysed, and 5-year toxicity results from consolidation cohort patients are presented in this chapter.

3.1.1 My Role

I wrote the draft mapping rules by which raw toxicity data collected with VoxTox CRFs were encoded into grades of recognised validated scoring systems. I also produced the finalised versions for coding, once they had been reviewed and commented on by a second expert observer. I requested the extraction of raw data – in VoxTox CRF form - from the Mosaiq system, in order to undertake this analysis. Once data in scoring system format had been returned to me, I undertook all subsequent data carpentry, checking and analysis presented in this chapter.

3.1.2 Acknowledgments

The clinical reporting forms (CRFs) used to collect the raw toxicity data presented in this Chapter were written by Jessica Scaife (Consultant Oncologist, Cheltenham Hospital), and Amy Bates (Research Radiographer), as described in Chapter 2. Amy Bates and members of her team undertook the vast majority of patient interviews in which the raw toxicity data presented in this chapter were collected. At my request, extraction of raw data from the Mosaiq database was undertaken by Kevin Skilton (Radiotherapy data manager). Rashmi Jadon (consultant H&N oncologist), reviewed the first draft of the rules mapping scores from VoxTox CRFs to recognised validated scoring systems. Based on her opinion and comments, I reviewed, amended and finalised the mapping rules so that they could be encoded in the python programming language. Karl Harrison (Research Associate, Cavendish Laboratory) wrote the parsing code to import the finalised mapping rules into python, such that all

subsequent translations of raw toxicity data from VoxTox CRF format into recognised scoring system grades could be automated. Amy Bates undertook anonymisation of raw data, after it had been extracted from Mosaiq, and transfer to the Cavendish Laboratory. Karl Harrison processed this data using the method described and sent me results in validated score format.

3.2 Introduction

Chapter 2 describes the design and infrastructure of the VoxTox study. A key component of the programme and this thesis is a detailed analysis of the toxicity that is experienced by patients undergoing radical radiotherapy for head and neck cancer.

As alluded to in Chapter 1, radiotherapy for head and neck cancer is an effective but morbid treatment. The majority of SCCs, which comprise 90% of HNC [1], arise from the tonsils, base of tongue, hypopharynx and larynx. Important normal tissue structures such as salivary glands, pharyngeal constrictor muscles, and laryngeal structures may be infiltrated by disease in these locations. If not, they are frequently adjacent to tumour, and at risk of high doses of radiation, despite the steep dose gradients that are possible with contemporary radiotherapy techniques.

It is already known that there is a clear relationship between dose to these structures, and long-term side effects such as xerostomia and swallowing dysfunction [2, 3], which can have a devastating effect on patient's quality of life [4-7]. In the era of 2-D and 3-D conformal radiotherapy, the incidence of such side effects after successful treatment was high. Based on results with these techniques [8], the landmark PARSPORT trial assumed a 1-year xerostomia rate of 90% in the control arm [9]. This trial went on to show that the concave dose distributions possible with IMRT, that could generate steep dose gradients between a target volume and an OAR, could substantially reduce the incidence of severe side effects.

Nonetheless, toxicity following radiotherapy remains a significant issue, and the problem is magnified by the changing biology and demographics of HNC [10-12], driven by the HPV epidemic [13]. As discussed in Chapter 1, this has led to contemporary clinical trials in low-risk HPV positive oropharyngeal cancer defining toxicity reduction as the primary endpoint [14-16].

Robust and consistent reporting of the toxicity associated with any cancer treatment is complex and challenging [17, 18], and this problem certainly applies to radiotherapy. Studies show that the same patient cohort, examined at the same time but using different methodologies, can

yield different results [19]. Therefore, consistent and accurate recording of treatment toxicities has long been seen as a priority for the oncology community [20].

To address this problem, researchers designed, tested and validated scoring tools in an attempt to quantify morbidity in a meaningful and consistent way. Some scores were designed specifically for using in the radiotherapy clinic [21, 22], whilst others were written for the full range of malignancies and treatment modalities [23]. In addition, there is an evolving corpus of data showing the importance of patient reported, rather than clinician recorded, treatment outcomes, and proformas to capture these data have also been developed [24-26]. A significant challenge for the initial VoxTox investigators was that recent and ongoing clinical studies and trials in the UK use a range of different scoring systems, as described in Chapter 2. Therefore, the CRFs used to collect raw toxicity within the study were designed to ensure that sufficient information would be present to populate the most commonly used scoring systems, to try and maximise comparability with results reported in the literature.

The core hypothesis of this thesis is that delivered dose metrics predict toxicity events more accurately than planned dose equivalents. However, toxicity data are noisy, and the literature shows that dose differences to many of the pertinent OARs may be relatively small [27]. Therefore, it is very important that the methodology for acquisition and processing of the toxicity data presented is robust, that the data presented are plausible, and that they are comparable with similar cohorts described in the literature.

This chapter describes in detail how toxicity data presented in this dissertation were collected, mapped onto recognised validated scoring systems, stored, collated and processed. The toxicity data are presented in detail, and decisions regarding which endpoints to carry forward for further analysis in later chapters are elucidated and explained.

3.3 Methods

3.3.1 Data collection

Details pertaining to the structure and conduct of the VoxTox study are given in Chapter 2. However, some important details of study design bear re-emphasis. The first is the 2 separate study cohorts; discovery and consolidation. As patients in the discovery cohort had neither baseline toxicity assessment, nor recording of comorbidities, and had only follow up toxicity

recorded from the point of recruitment, these patients are not included in the analysis described in this chapter. Another is the timepoints at which patients in the consolidation cohort were assessed for late toxicity; 6 months, 12 months and then annually to 5 years.

As discussed in Chapter 2, the literature reports the late effects suffered by many patients following radiotherapy for HNC using a number of different toxicity scoring systems, and VoxTox study designers aimed to collect sufficient data from every patient participant to populate all of these different scores. To do this, the information contained within them was deconstructed to 60 unique question stems, which were subsequently compiled into a study-specific CRFs. Details of this process are given in the PhD thesis of Dr Jessica Scaife.[6]

Patient interviews were undertaken by the study Research Radiographer, Amy Bates, or members of her team. Where study toxicity assessments coincided with clinical follow-up visits, these took place in person. Where this was not possible, interviews were done by telephone. Interviews were structured based on VoxTox CRFs, as shown in Appendix A3.1-3.3. Answers to the CRF questions were either transcribed directly into the Mosaiq database, or recorded on paper and digitised shortly thereafter.

An important consideration is the consistency with which different observers encode patient responses into the scores recorded on the VoxTox CRFs. This work was undertaken early in the VoxTox study, and is also presented in the PhD dissertation of Jessica Scaife.[7] The results showed good concordance between the scores recorded by Amy Bates, and clinician scores, and therefore this work was not repeated.

3.3.2 Translating CRF data to standardised endpoints

3.3.2.1 Writing mapping rules

The clinician-rated scoring systems chosen to report toxicity in this work were CTCAE v4.03 [23], Late Effects Normal Tissue, Subjective, Objective, Management - LENT/SOM [22], and Radiation Therapy Oncology Group – RTOG [21]. VoxTox CRF scores were also mapped to the quality of life scoring questionnaire, QLQ-H&N35 [25]. These were the scores that the VoxTox CRFs had been written to populate, and were also chosen to maximise comparability with available literature.

---

[6] PhD Thesis: Dr. Jessica Scaife, Peterhouse College, University of Cambridge, 2015.

[7] PhD Thesis: Dr. Jessica Scaife, Peterhouse College, University of Cambridge, 2015.

It should be noted that CTCAEv4.03, LENT/SOM, and RTOG scoring systems were designed as physician-rated toxicity assessment tools, whereas QLQ-H&N35 is a patient reported outcome metric. In the drive towards evermore patient-centric practice, such tools are gaining prominence in basic research, clinical trials and commissioning. As indicated above, all raw data collection took place at interviews with research radiographers, thus responses were not truly 'patient reported'. However, the questions in the VoxTox CRFs match those in QLQ-H&N35 very accurately (see Appendix A3). Furthermore, within interviews, patients were encouraged to answer themselves by choosing the most appropriate response, with minimal clinician influence. It was therefore decided by VoxTox investigators that mapping raw data to QLQ-H&N35 scores was meaningful and valid.

In order to facilitate efficient translation of multiple batches of raw data in VoxTox CRF form, at multiple timepoints, into recognised scores, it was decided to automate the process. By using this methodology, it ensured that the algorithms (or rules) encoding the translation would be transparent to the whole research group, and open to external scrutiny and validation. Furthermore, it meant that version control of mapping rules could be more tightly controlled, and updates and iterations of the rules could be easily managed by the underlying code.

To achieve this, it was therefore necessary to write the mapping rules in such a way that they were easily read and interpreted both by human observers and by code in the python language in-silico. Raw data, extracted from Mosaiq in VoxTox CRF form were stored in Microsoft Excel format. Spreadsheets used for toxicity mapping had patients listed by row, and CRF responses listed by column. A hypothetical example is shown below in Table 3.1.

	CH	CI	CJ	CK
Patient ID	Dry mouth	Saliva	Taste	Swallowing difficulties
VT1_H_6BA…..	2	1	1	0
VT1_H_35D…..	1	0	1	1

Table 3.1: Simplistic representation of the layout of Microsoft excels spreadsheet for mapping toxicity data from VoxTox CRF format, onto recognised scoring systems.

Mapping rules were written in Microsoft word, with the following logical operators and syntax:

IF (A>x) AND (B=y, OR C=z) THEN ScoringSystem_Endpoint = t

The capital letters (A,B,C) indicate columns, pertaining to VoxTox CRF raw data (lower case letters; x,y,z). In this example, t denotes the result, or score, in the endpoint of interest. An example rule, for Grade 2 CTCAEv4.03 dysphagia is given below:

IF (CK=1) AND (W=1) AND (((Z=1) OR (AB=1) OR (AC=1) OR (AD=1) AND (AL=1)) THEN CTCAEv4.03_dysphagia=2

The definition for CTCAE Grade 2 dysphagia is given below.

- Dysphagia – a disorder characterised by difficulty in swallowing
- Grade 2 – Symptomatic and altered eating/swallowing

In this example the excel columns referred to are as follows (the questions and their layout are shown in the CRF itself in Appendix A3.3):

- CK – Swallowing difficulty
    - 0 – None
    - 1 – Present
    - 2 – Life threatening consequences & completely unable to swallow
    - 3 – perforation or fistula of oesophagus
- W - Dietary changes
    - 0 – No
    - 1 – Yes

Columns X through to AH are contingent on column W, they are specific fields showing *how* diet has been altered.

- Z – Needs saliva substitute/water in order to eat
- AB – small bites
- AC – semi-soft food
- AD – soft food

Columns AI through to AP are also contingent on W, and responses to X->AH. These are specific fields showing *why* the diet has been altered.

- AL – swallowing difficulty.

In language therefore, this rule is therefore making the following judgement:

1. The patient reports swallowing difficulty.
2. The patient has dietary changes.

3. These dietary changes consist of at least one of; needing saliva or water substitute to eat, eating small bites, semi-soft, or soft food.
4. The reason for these dietary changes is swallowing difficulty.

In writing the mapping rules, the first step was to ensure that all combinations of clinical features that constitute a given clinical endpoint were covered. It was also important to consider which symptom, or combination of symptoms was most likely responsible for a given loss of function. Therefore, there was inevitable logical duplication within rules, and as a second step, rules were simplified, and duplicate rules were deleted.

Another important consideration is that this workflow depends entirely on the fact that the excel column denoted in the code accurately reflects the correct question in the VoxTox CRF. Therefore strict version control of both data spread sheets and mapping rule documents was followed both by myself, and Karl Harrison at the Cavendish Laboratory. Furthermore, I manually inspected the raw data excel spreadsheet, the mapping rule version being used, and the resulting scoring system data after each translation and prior to any further analysis.

3.3.2.2 Validating mapping rules

Although toxicity-scoring systems are designed to be as clear and precise as possible, there is inevitably subjective judgement inherent in the process of interpreting them. Therefore, it was important to validate the mapping rules for translating VoxTox CRF data in scoring endpoints.

To do this, a consultant oncologist colleague sub-specialising in HNC (Dr Rashmi Jadon) reviewed mapping rules for all 4 scoring systems. In the first iteration of this process, the methodology was simply to review my mapping rules, comment on their accuracy, and make suggested amendments where necessary. However, this process was found to be laborious and inefficient. Therefore, an alternative strategy was used in which the second observer simply used the same syntax described in section 3.3.2.1 to write a new set of rules, entirely independent of the first set. I then reviewed both sets of rules, and noted and quantified whether rules agreed, or had minor or major differences. Where there were subtle differences in the interpretation of which symptom, or combination of symptoms, constituted toxicity of a given grade, this was considered a minor disagreement. Where differences in the grade of toxicity associated with a given response on the CRF existed, this was considered a major disagreement.

Where disagreements existed, I considered both sets of logical rules, and either constructed a compromise between the 2, or decided to use one of the two original rules in the finalised version. This finalised version of the mapping rules was then 'locked' and sent electronically to the Cavendish laboratory. The finalised mapping rules used for translating data from VoxTox CRF to scoring system format are shown in Appendix A3.4 – A3.11.

3.3.2.3 Automating toxicity mapping

On receipt of the finalised mapping rules, colleagues at the Cavendish Laboratory deployed parsing code to import the logical operators from Microsoft Word into the python coding language. A second batch of code was written, which then imported raw CRF data in Microsoft excel format, executed the mapping rules, and exported data in .csv format as scoring system endpoint grades.

3.3.2.4 Checking automated scores

Once the automated process was complete, it was important to check that the results produced were plausible and accurate. A dataset of 5 random patients and timepoints was selected, and the 12 toxicity endpoints reported in this chapter were tested. This sample size was chosen on the a-priori expectation that it would provide a meaningful corpus data to validate the process, within a reasonable timeframe.

Using the VoxTox CRF raw data, and endpoint definitions, I produced endpoint scores using two methods, to check both the validity of the rules themselves, and the process of automation. The first was to translate CRF scores into endpoint scores simply by interpretation of the definitions, without any formalised rules. The process was then repeated, this time manually working through the finalised mapping rules to dictate how each translation should be done. The difference between the results of the automated translation, and both methods of manual translation, for each score was recorded, this giving 60 toxicity scores for each comparison.

3.3.3 Selecting endpoints

As can be seen from the CRF shown in Appendix A3.3, data on a wide range of potential toxicities were collected. This includes endpoints such as trismus, vision, and hearing. In order to maximise the usefulness of the VoxTox dataset for future work, finalised mapping rules included translations of all CRF data to relevant scoring system endpoints, wherever there was

sufficient raw data to inform the translation. However, the data carpentry, preparation and analysis necessary to present toxicity data was laborious and time-consuming. Therefore, to maximise efficiency, only endpoints pertinent to the core hypothesis of the thesis – the relationship between planned and delivered dose and toxicity – were included in this analysis. The endpoints included are as follows:

- CTCAEv4.03
    - Xerostomia
    - Salivary duct inflammation (SDI). This endpoint describes the clinical phenotype of thick, sticky, ropey saliva and oropharyngeal secretions, with or without alterations in taste.
    - Dysphagia
- LENT/SOM
    - Subjective - xerostomia
    - Management – xerostomia
    - Subjective - dysphagia
    - Subjective – taste alteration
- RTOG
    - Salivary gland
    - Oesophagus
- EORTC H&N35
    - Question 41. Have you had a dry mouth?
    - Question 42. Have you had sticky saliva?
    - Question 44. Have you had taste disturbance?

3.3.4 Statistical analysis

Multiple timepoint toxicity data were analysed in two ways. Firstly, point prevalence of the toxicity grade for data at each timepoint are presented as relative stacked bar charts. Secondly, cumulative incidence for each toxicity grade, for each endpoint was calculated, using the Kaplan-Meier method. Analyses were all conducted in Microsoft Excel.

3.4 Results

3.4.1 Final dataset for analysis

Patients recruited to the consolidation cohort (224), and the Minot-OAR sub-study (18) were available for analysis of toxicity data. 6 withdrew from the study before completing a baseline toxicity questionnaire. Median follow up for patients who did complete the baseline toxicity assessment was 23.8 months (IQR 11.7 – 39.5 months). The number of patients with completed toxicity questionnaires at baseline, 6, 12, 24, 36, 48 and 60 months respectively were 236, 209, 200, 127, 100, 55 and 15.

3.4.2 Validation of mapping rules

Agreement between version 1 of the mapping rules, and those written independently by a second expert observer are shown in Table 3.2.

Scoring system	Agree	Minor difference	Major difference
CTCAE	58 (77.3%)	14 (18.7%)	3 (4%)
LENT/SOM	33 (91.7%)	2 (5.5%)	1 (2.8%)
RTOG	123 (98.4%)	2 (1.6%)	0
EORTC H&N 35	28 (100%)	0	0

Table 3.2: Toxicity mapping rule validation. Agreement between version 1 of the mapping rules written by me (Observer 1), and those independently written by a second HNC expert (Observer 2).

3.4.3 Checking automated scores

Across all 5 patients, and 12 toxicity endpoints in the test dataset, there were no discrepancies between the automated interpretation of the mapping rules, and my manual interpretation of the mapping rules – scores were identical for all 60 endpoints. There were some differences between my interpretation of the raw CRF data, using the scoring systems themselves, and the automated results. Across all 60 endpoints, 54 (90%) were identical, 6 (10%) differed by 1 grade (e.g. 1 instead of 2, 2 instead of 3), and none by more than 1 grade. By scoring system, the results were as follows: CTCAE – 14/15 identical (93.3%), LENT/SOM – 18/20 identical (90%), RTOG – 8/10 identical (80%), EORTC H&N35 14/15 identical (93.3%).

3.4.4 Toxicity results – CTCAE

Point prevalence and cumulative incidence results, for CTCAE endpoints Xerostomia, Salivary duct inflammation and Dysphagia are shown in Figure 3.1, captions A-F.

3.4.5 Toxicity results – LENT/SOM

Point prevalence and cumulative incidence results, for LENT/SOM endpoints Subjective Xerostomia, Salivary duct inflammation and Dysphagia are shown in Figure 3.2, captions A-H.

3.4.6 Toxicity results – RTOG

Point prevalence and cumulative incidence results, for RTOG endpoints Salivary gland and Esophagus are shown in Figure 3.3, captions A-D.

3.4.7 Toxicity results – EORTC H&N-35

Point prevalence and cumulative incidence results, for H&N-35 questions 41. Have you had a dry mouth? 42. Have you had Sticky Saliva? and 44. Have you had taste disturbance? are shown in Figure 3.4, captions A-F.

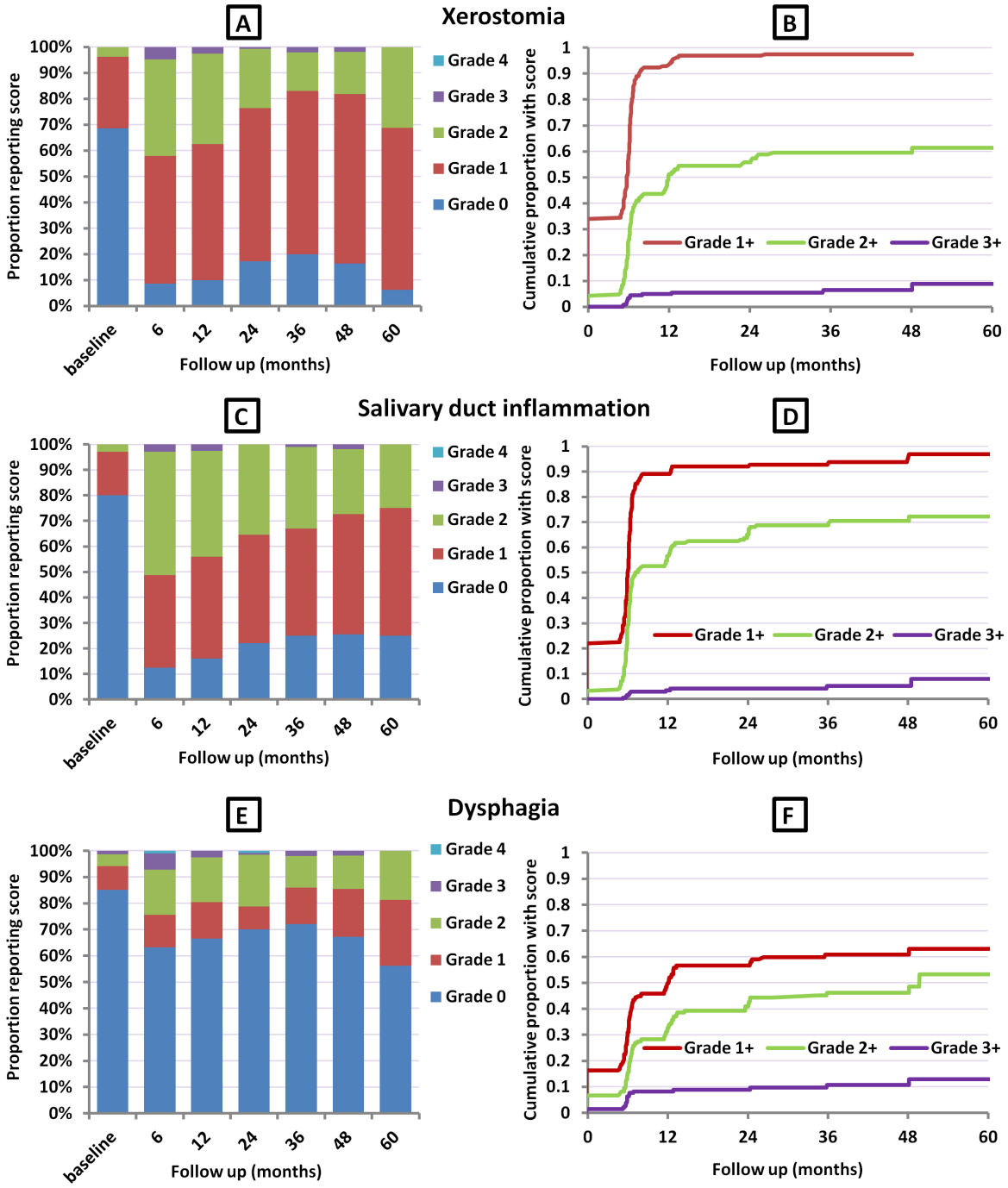


Figure 3.1: CTCAE - 5 year toxicity results for Xerostomia, Salivary duct inflammation and Dysphagia. Point prevalence results at 6, 12, 24, 36, 48 and 60 months are shown in captions A, C and E. The number of patients for analysis at each timepoints were 236, 209, 200, 127, 100, 55 and 16 respectively. Cumulative incidence results are shown in captions B, D and F; numbers at risk vary per Grade.

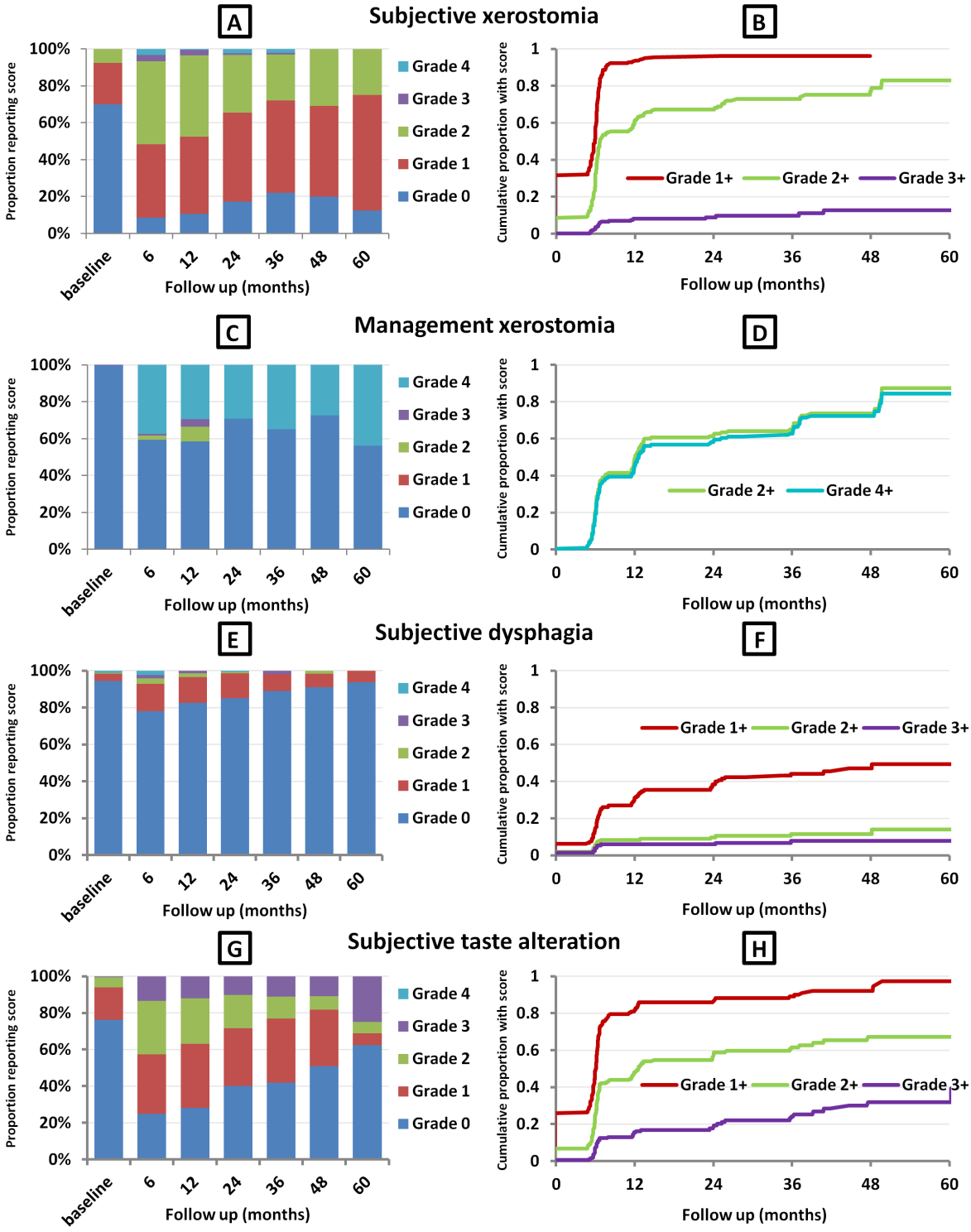


Figure 3.2: LENT/SOM - 5 year toxicity results for Subjective xerostomia, Management xerostomia, Subjective dysphagia, and Subjective taste alteration. Point prevalence results at 6, 12, 24, 36, 48 and 60 months are shown in captions A, C, E and G. The number of patients for analysis at each timepoints were 236, 209, 200, 127, 100, 55 and 16 respectively. Cumulative incidence results are shown in captions B, D, F and H; numbers at risk vary per Grade.

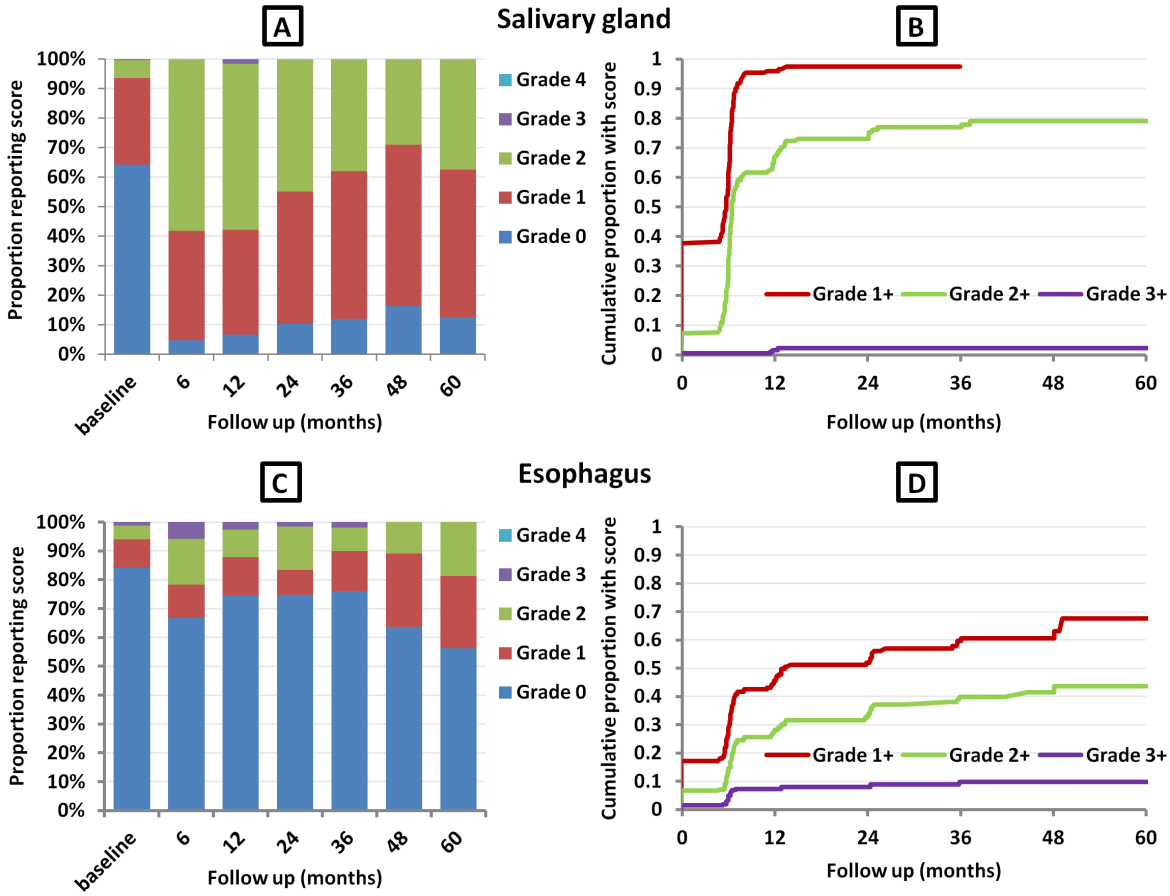


Figure 3.3: RTOG - 5 year toxicity results for Salivary gland, and Esophagus. Point prevalence results at 6, 12, 24, 36, 48 and 60 months are shown in captions A and C. The number of patients for analysis at each timepoints were 236, 209, 200, 127, 100, 55 and 16 respectively. Cumulative incidence results are shown in captions B and D; numbers at risk vary per Grade.

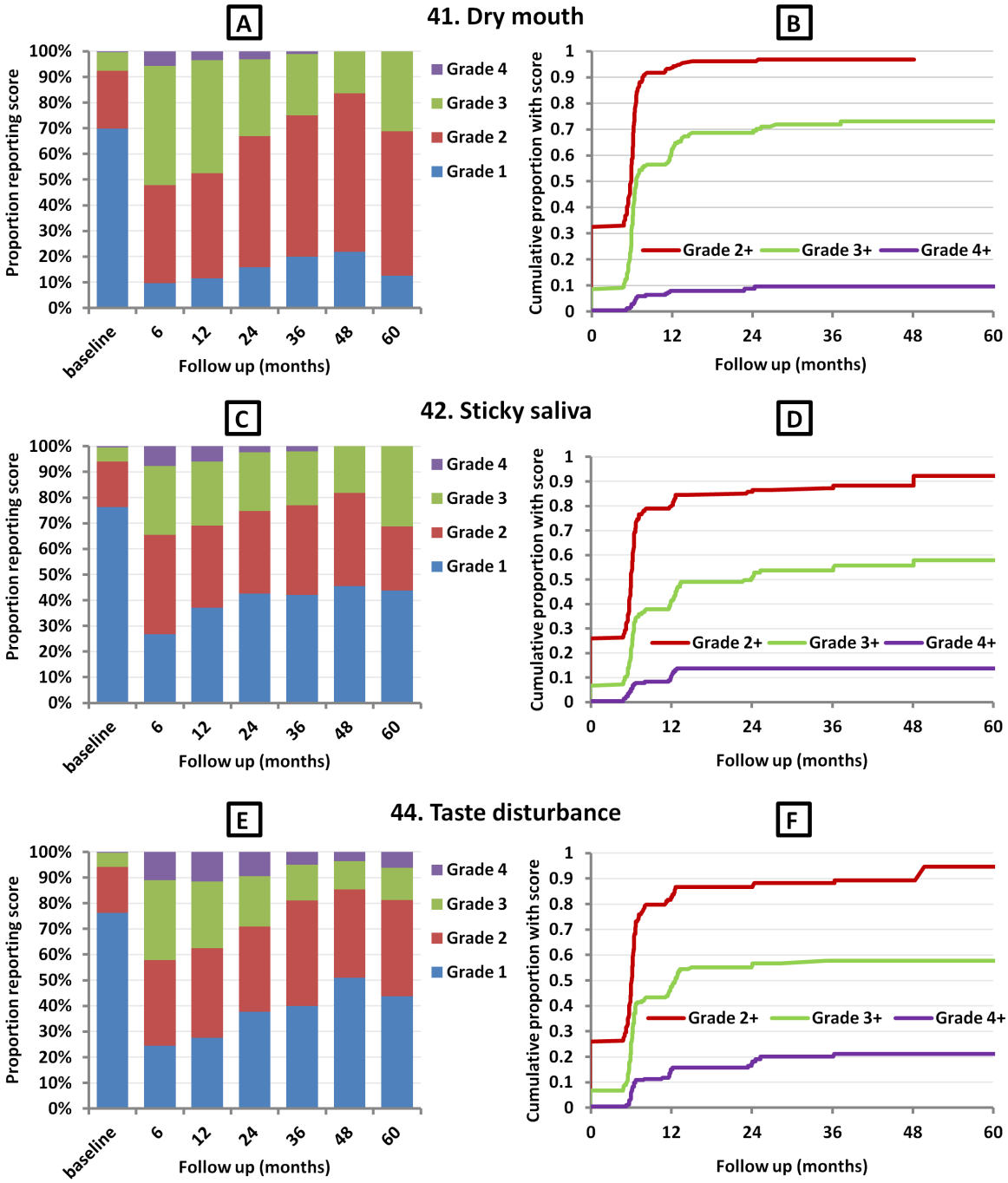


Figure 3.4: EORTC H&N-35 - 5 year toxicity results for questions 41. Have you had a dry mouth? 42. Have you had Sticky Saliva? and 44. Have you had taste disturbance? Point prevalence results at 6, 12, 24, 36, 48 and 60 months are shown in captions A, C and E. The number of patients for analysis at each timepoints were 236, 209, 200, 127, 100, 55 and 16 respectively. Cumulative incidence results are shown in captions B, D and F; numbers at risk vary per Grade.

3.5 Discussion

3.5.1 Toxicity results in context

These toxicity results broadly align with published data. The 12 and 24-month point prevalence of LENT-SOM xerostomia in the PARSPORT trial are slightly lower than the equivalent results of 47.5% and 34.6% in this work [9]. This trial also report data on swallowing dysfunction, specifically maximum reported subjective dysphagia using the LENT-SOM scale over the study period. The published result of 20% in the IMRT arm is slightly higher than the 13.9% shown here [9].

Patient and treatment factors that can affect toxicity rates must be considered when comparing results. Concomitant cisplatin chemotherapy is known to increase toxicity rates compared with radiation alone [28-30], although many of the trials that undertook these comparisons were conducted in the pre-IMRT era, and are therefore of little use for direct comparison of toxicity rates. Concomitant biological therapy with Cetuximab has also been shown to improve disease outcomes [31], although this trial did not collect late toxicity data [32]. Two recent randomised trials aimed to test the hypothesis that concomitant Cetuximab would have similar efficacy to cisplatin in patients with low risk oropharyngeal SCC, whilst reducing toxicity [14, 16]. Both trials were negative. The De-ESCALaTE HPV trial reported adverse event data in a way that precludes meaningful comparison with the results reported here [14], but the RTOG 1016 study did publish maximum recorded late toxicity over the study period using CTCAE, and rates of Grade 2+ xerostomia of 49.7% for cisplatin, and 53.6% for cetuximab are reported [16]. These figures are slightly lower than the cumulative incidence for Grade 2+ xerostomia of 61.4% shown in this work, a difference that may in part be explained by different reporting methodologies.

Radiotherapy technique is also important. It would be logical to suggest that irradiating one side of the head and neck would be less toxic than treating bilaterally, an important point as 23% of patients in this analysis underwent unilateral treatment (data shown in Chapter 2). There are comparable data in the literature for these patients. The COSTAR study randomised patients who had undergone surgery for primary tumours of the parotid glands to 3DCRT or IMRT, hypothesising that IMRT would reduce hearing loss [33]. The study also collected broader toxicity data, and reported 12-month point prevalence of Grade 1+ CTCAE xerostomia of 79%, and a cumulative incidence of Grade 2+ toxicity at 60 months of 20% in the IMRT arm [33]. Equivalent figures of 90% and 61.4% are, as expected, higher in the results presented

here, but the COSTAR data show that patients undergoing unilateral radiotherapy can still experience significant side effects.

However, direct comparisons of crude toxicity rates between interventional, and observational studies such as this, should be treated with caution, as there is greater scope for unconscious bias in reporting in the former. Furthermore, a comparison of the results presented in this chapter with toxicity rates in widely cited NTCP models is important for the work presented in Chapter 7. Beetz and colleagues used responses to the EORTC H&N35 scale to construct and validate models of both xerostomia (question 41) and sticky saliva (question 42) [2]. These models look at point prevalence 6 months after completion of radiotherapy. 83 of 161 (51%) patients reported grade 3 (quite a bit) or 4 (very much) dry mouth, almost identical to our result of 52.1%. The pattern for sticky saliva was similar, with results of 35.6% (Beetz model) and 34.5% (this work) respectively. Previous modelling of swallowing dysfunction is also based on similar event rates. Wopken et al built an NTCP model of tube-feeding dependence, which is equivalent to grade 3+ CTCAE dysphagia. At 6 months follow up, they found that this affected 10.7% of the patients in their cohort, similar to the 7.2% seen in this work.

3.5.2 Trends over time

A clear trend across all 4 scoring systems is that the highest point prevalence of most endpoints is at 6 months, with gradual improvements thereafter up to 5 years. This is most apparent with xerostomia, salivary duct inflammation and taste disturbance as scored by CTCAE, LENT-SOM and H&N35. For example, point prevalence of CTCAE xerostomia and salivary duct inflammation at 12 and 48 months are 37.5% and 44%, and 18.1% and 27.3% respectively. Results at 60 months for many of the endpoints do not seem entirely consistent with this trend, but this observation should be treated with caution due to the low number of patients (16) with 5 years follow up. That said, it is clear from the cumulative incidence plots that new cases of toxicity of all grades can and do occur later in follow up, even up to 3 or 4 years post treatment. Thus, whilst point prevalence clearly falls over time for most toxicity endpoints, longer follow up is required to capture all the patients who go on develop significant side effects.

This trend to improvement over time is more apparent with salivary gland toxicities, than with swallowing dysfunction. Although LENT-SOM subjective dysphagia appears to improve over time, this does not appear to be the case for either CTCAE dysphagia, or RTOG esophagus. Gradual improvement of xerostomia over time is a recognised phenomenon. Point prevalence

rates of 51% (grade 3+ H&N 35 dry mouth), and 35.6% (grade 3+ H&N 35 sticky saliva) at 6 months in the original Beetz model [2], improved to 40% and 25% respectively in a similar analysis with 12 month data [34]. Similarly, the PARSPORT trial data show that grade 2+ LENT-SOM subjective xerostomia improved from 38% at 12 months, to 29% at 2 years [9]. The same symptom measured by the RTOG scoring system also improved from 38% to 18% over the same period. This trial also provides potential insight into the mechanism of this effect, because similar recovery was not observed in the conventional radiotherapy arm. One hypothesis is that stem cells within the parotid glands may have been spared by IMRT in a way that did not occur with 3DCRT, thus permitting tissue regeneration and functional recovery, and this effect has been seen in mouse models [35].

3.5.3 Robustness of toxicity data

The process for automating translation of data from CRF to scoring system grades appears to be robust. Of the 264 mapping rules written in version 1, only 4 (1.5%) showed a major disagreement between observers, in which mapping rules resulted in toxicity of a different grade being attributed to a set of reported symptoms. 242 rules (91.7%) were in agreement. On testing, no differences between human and machine interpretation of the mapping rules were observed, and concordance between automated execution of the mapping rules, and subjective interpretation of the scoring system tables was seen in 54/60 (90%) of the endpoints measured.

One potential weakness in the design of the VoxTox study was the absence of the MDADI toxicity assessment tool [26]. This scoring system is widely accepted as a useful tool for assessing post-treatment swallowing outcomes, and has been used in a number of trials [14, 15, 36]. However, this scoring system comprises 20 Likert scale questions all related to swallowing, and colleagues judged that incorporating these questions into the single VoxTox CRF covering all other scoring system systems and toxicity endpoints would have made it too long, and unwieldy to use in the confines of VoxTox follow-up interviews.

3.6 Conclusions

Toxicity data within the VoxTox study were collected using bespoke CRFs, specifically designed to convey sufficient information to populate the most widely used toxicity scoring systems. An automated process, whereby human and machine interpretable logical rules translated data from CRF format to scoring system grades was developed. This process was

validated and tested, and found to be reliable and robust. The results produced by this process permitted analysis of point prevalence and cumulative incidence for CTCAE, LENT-SOM, RTOG and EORTC H&N35 scores. These results were broadly comparable to previously published results, in which similar cohorts of patients were treated with similar techniques. Some of the toxicity results shown in this chapter will be examined again in Chapter 7, which addresses whether planned or delivered dose metrics better predict these outcomes. Chapter 4 describes the methodology by which automated segmentation on IG scans was achieved, in order to calculate delivered dose.

3.7 References:

[1] Gregoire V, Lefebvre JL, Licitra L, Felip E, Group E-E-EGW. Squamous cell carcinoma of the head and neck: EHNS-ESMO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2010;21 Suppl 5:v184-6.
[2] Beetz I, Schilstra C, van der Schaaf A, van den Heuvel ER, Doornaert P, van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. Radiother Oncol. 2012;105:101-6.
[3] Wopken K, Bijl HP, van der Schaaf A, van der Laan HP, Chouvalova O, Steenbakkers RJ, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol. 2014;113:95-101.
[4] Bjordal K, Kaasa S, Mastekaasa A. Quality of life in patients treated for head and neck cancer: a follow-up study 7 to 11 years after radiotherapy. Int J Radiat Oncol Biol Phys. 1994;28:847-56.
[5] Epstein JB, Emerton S, Kolbinson DA, Le ND, Phillips N, Stevenson-Moore P, et al. Quality of life and oral function following radiotherapy for head and neck cancer. Head Neck. 1999;21:1-11.
[6] Lin A, Kim HM, Terrell JE, Dawson LA, Ship JA, Eisbruch A. Quality of life after parotid-sparing IMRT for head-and-neck cancer: a prospective longitudinal study. Int J Radiat Oncol Biol Phys. 2003;57:61-70.
[7] Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. J Clin Oncol. 2008;26:3770-6.
[8] Franzen L, Funegard U, Ericson T, Henriksson R. Parotid gland function during and following radiotherapy of malignancies in the head and neck. A consecutive study of salivary flow and patient discomfort. Eur J Cancer. 1992;28:457-62.
[9] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol. 2011;12:127-36.
[10] Cleary C, Leeman JE, Higginson DS, Katabi N, Sherman E, Morris L, et al. Biological Features of Human Papillomavirus-related Head and Neck Cancers Contributing to Improved Response. Clin Oncol (R Coll Radiol). 2016;28:467-74.
[11] Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst. 2000;92:709-20.
[12] Dahlstrom KR, Bell D, Hanby D, Li G, Wang LE, Wei Q, et al. Socioeconomic characteristics of patients with oropharyngeal carcinoma according to tumor HPV status, patient smoking status, and sexual behavior. Oral Oncol. 2015;51:832-8.
[13] Sturgis EM, Cinciripini PM. Trends in head and neck cancer incidence in relation to smoking prevalence: an emerging epidemic of human papillomavirus-associated cancers? Cancer. 2007;110:1429-35.
[14] Mehanna H, Robinson M, Hartley A, Kong A, Foran B, Fulton-Lieuw T, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. Lancet. 2019;393:51-60.
[15] Owadally W, Hurt C, Timmins H, Parsons E, Townsend S, Patterson J, 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;15:602.
[16] Gillison ML, Trotti AM, Harris J, Eisbruch A, Harari PM, Adelstein DJ, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. Lancet. 2019;393:40-50.
[17] Bentzen SM, Dorr W, Anscher MS, Denham JW, Hauer-Jensen M, Marks LB, et al. Normal tissue effects: reporting and analysis. Semin Radiat Oncol. 2003;13:189-202.
[18] Trotti A, Bentzen SM. The need for adverse effects reporting standards in oncology clinical trials. J Clin Oncol. 2004;22:19-22.
[19] Denis F, Garaud P, Bardet E, Alfonsi M, Sire C, Germain T, et al. Late toxicity results of the GORTEC 94-01 randomized trial comparing radiotherapy with concomitant radiochemotherapy for advanced-stage oropharynx carcinoma: comparison of LENT/SOMA, RTOG/EORTC, and NCI-CTC scoring systems. Int J Radiat Oncol Biol Phys. 2003;55:93-8.
[20] Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer. 1981;47:207-14.

[21] Cox JD, Stetz J, Pajak TF. Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC). Int J Radiat Oncol Biol Phys. 1995;31:1341-6.
[22] LENT SOMA tables. Radiother Oncol. 1995;35:17-60.
[23] Trotti A, Colevas AD, Setser A, Rusch V, Jaques D, Budach V, et al. CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol. 2003;13:176-81.
[24] Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85:365-76.
[25] Bjordal K, de Graeff A, Fayers PM, Hammerlid E, van Pottelsberghe C, Curran D, et al. A 12 country field study of the EORTC QLQ-C30 (version 3.0) and the head and neck cancer specific module (EORTC QLQ-H&N35) in head and neck patients. EORTC Quality of Life Group. Eur J Cancer. 2000;36:1796-807.
[26] Chen AY, Frankowski R, Bishop-Leone J, Hebert T, Leyk S, Lewin J, et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: the M. D. Anderson dysphagia inventory. Arch Otolaryngol Head Neck Surg. 2001;127:870-6.
[27] van Kranen S, Hamming-Vrieze O, Wolf A, Damen E, van Herk M, Sonke JJ. Head and Neck Margin Reduction With Adaptive Radiation Therapy: Robustness of Treatment Plans Against Anatomy Changes. Int J Radiat Oncol Biol Phys. 2016;96:653-60.
[28] Denis F, Garaud P, Bardet E, Alfonsi M, Sire C, Germain T, et al. Final results of the 94-01 French Head and Neck Oncology and Radiotherapy Group randomized trial comparing radiotherapy alone with concomitant radiochemotherapy in advanced-stage oropharynx carcinoma. J Clin Oncol. 2004;22:69-76.
[29] Trotti A, Pajak TF, Gwede CK, Paulus R, Cooper J, Forastiere A, et al. TAME: development of a new method for summarising adverse events of cancer treatment by the Radiation Therapy Oncology Group. Lancet Oncol. 2007;8:613-24.
[30] Machtay M, Moughan J, Trotti A, Garden AS, Weber RS, Cooper JS, et al. Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: an RTOG analysis. J Clin Oncol. 2008;26:3582-9.
[31] Bonner JA, Harari PM, Giralt J, Azarnia N, Shin DM, Cohen RB, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. N Engl J Med. 2006;354:567-78.
[32] Bonner JA, Harari PM, Giralt J, Cohen RB, Jones CU, Sur RK, et al. Radiotherapy plus cetuximab for locoregionally advanced head and neck cancer: 5-year survival data from a phase 3 randomised trial, and relation between cetuximab-induced rash and survival. Lancet Oncol. 2010;11:21-8.
[33] Nutting CM, Morden JP, Beasley M, Bhide S, Cook A, De Winton E, et al. Results of a multicentre randomised controlled trial of cochlear-sparing intensity-modulated radiotherapy versus conventional radiotherapy in patients with parotid cancer (COSTAR; CRUK/08/004). Eur J Cancer. 2018;103:249-58.
[34] van Dijk LV, Brouwer CL, van der Schaaf A, Burgerhof JG, Beukinga RJ, Langendijk JA, et al. CT image biomarkers to improve patient-specific prediction of radiation-induced xerostomia and sticky saliva. Radiother Oncol. 2017;122:185-91.
[35] van Luijk P, Pringle S, Deasy JO, Moiseenko VV, Faber H, Hovan A, et al. Sparing the region of the salivary gland containing stem cells preserves saliva production after radiotherapy for head and neck cancer. Sci Transl Med. 2015;7:305ra147.
[36] Petkar I, Rooney K, Roe JW, Patterson JM, Bernstein D, Tyler JM, 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;16:770.

Chapter 4 – Using deformable image registration for automated contouring of swallowing OARs

4.1 Overview

Calculating delivered dose from daily IG scans presents a number of logistical and computational challenges. Solutions to these problems require significant automation in a way that is physically plausible and computationally robust. Deformable image registration (DIR) is a well-studied and established methodology for this problem; it is widely used in the research arena, and increasingly in the clinic. This chapter describes how an open-source deformable image registration toolbox was trained and validated to calculate delivered dose to H&N OARs.

4.1.1 My Role

All planning scan (kVCT) segmentations used for contour propagation, as described in this chapter, were done by me. I designed and executed all the intra and inter-observer contouring studies described in this chapter. All experiments pertaining to the training and validation of DIR were done either following consultation with me, or at my instigation. However, the work described in this chapter was the result of very close collaboration between myself, and colleagues in department of high-energy physics (Cavendish laboratory), and I will therefore describe in some detail the precise role of each member of the research team.

Early in 2016, I made the decision to start work on automated segmentation and dose accumulation looking specifically at the spinal cord. For the rest of this year, I worked very closely with Lin Yeap, MPhil student from the Cavendish laboratory, and with Karl Harrison, Research Associate from the Cavendish Laboratory, on spinal cord auto-contouring and dose accumulation. I attended regular (minimum bi-weekly) meetings to discuss progress with this work, and help design further experiments. I provided all the reference segmentations for training and validation of the DIR algorithm for this task. I designed the inter-observer contouring study, and recruited consultants to participate in this.

Work on training and validating DIR for swallowing OARs was undertaken over 2017 and 2018 with MPhil student Shannon Seah, and Karl Harrison. I led on this work; my role was to work collaboratively to design experiments, provide reference contours, and generate both intra and inter-observer segmentation variability data. I undertook all analysis of data pertaining to training and validation metrics for deformable image registration for the swallowing OARs.

4.1.2 Acknowledgements

Lin Yeap, Shannon Seah and Karl Harrison deployed the DIR algorithm used for this work in the python coding language. Following consultation with myself and Karl Harrison, Lin Yeap performed the experiments to optimise the DIR algorithm for automated segmentation and dose accumulation to the spinal cord. He also analysed the data generated by these experiments. This work formed the basis of his MPhil thesis, which was submitted in July 2016[8]. As this work was led by Lin Yeap, and published in his MPhil thesis, I have not included any of the methodology or data for this work in this Chapter. However, this work was also published in Physics in Medicine and Biology [1], as:

Automatic contour propagation using deformable image registration to determine delivered dose to spinal cord in head-and-neck caner radiotherapy. PL Yeap, DJ Noble, K Harrison, AM Bates, NG Burnet, R Jena et al. *Phys Med Biol. 2017 Jul 12;62(15):6062-6073.*

In addition to the work described above, my role in preparing this manuscript was to write sub-sections, and heavily edit the first draft.

The approach taken for training and validating the DIR algorithm for swallowing OARs was similar to that taken in the spinal cord work. Karl Harrison helped me to design these experiments, and ran the software to process and analyse images and contours. Megan Wilson, PhD student (University College London) advised on the general approach to training and validating DIR. Marina Romanchikova, computer scientist, wrote the software for extracting images, structure sets and dose cubes from vendor archive, as described in Chapter 2. Structure sets were transferred between the hospital and Cavendish laboratory by Marina Romanchikova, and Andrew Hoole, computer scientists. Contours for this chapter were provided by the following consultant oncologists: Richard Benson, Gill Barnett, Rashmi Jadon and Kerwyn Foo.

I am grateful to Dr Eliana Vasquez-Osorio (Senior Research Fellow, University of Manchester) and Professor Jan-Jacob Sonke (Adaptive Radiotherapy group leader, NKI, Amsterdam) for helpful discussions about deformable image registration for dose accumulation.

[8] MPhil Thesis: Mr. Ping Lin Yeap, Jesus College, University of Cambridge, 2016.

4.2 Introduction

4.2.1 Background to the problem

Calculating and accumulating delivered dose to H&N OARs is a core underlying premise of the VoxTox project [2], and my PhD methodology. An accurate representation of the dose these structures actually receive, distinct from that anticipated at treatment planning permits interrogation of one of my core hypotheses – that D_A will predict toxicity events more accurately than D_P. However, this objective presents a number of theoretical and practical problems. Understanding and minimising uncertainties associated with this process is a key goal for the research community [3].

The first problem for my work was that of scale. As shown in Table 2.4 (Chapter 2), the VoxTox study has enrolled 337 patients with HNC. All patients were treated with radical intent, and therefore received between 20 and 35 treatment fractions. In order to calculate daily dose, and therefore accumulate to a final representation of delivered dose, it is necessary to know which voxels on each daily IG scan represent a given structure. Previous work in which delivered dose to H&N structures was calculated has been much smaller scale [4-6], and one established approach is manual segmentation of each pertinent image [7]. This was clearly an intractable task for this work; it would have required a structure set for all 8 OARs on approximately 10000 images. A mechanism for automated segmentation on MVCT images was therefore essential.

The next problem is how to accumulate dose. In practice, this means how best to combine the daily dose data calculated for each structure into a single per-structure representation of D_A. This remains a controversial topic within the field [8-10], and is discussed in greater detail in Chapters 5 and 6.

One approach that provides an elegant solution to the first problem, whilst also providing an established solution to the second – dose accumulation – is deformable image registration (DIR).

4.2.2 Image registration – an overview

Image registration is a process by which biological and anatomical information in one image is geometrically related to that in another [11]. Such techniques permit combining information for one subject from multiple imaging techniques, serial information on the same subject over time

from the same imaging modality, or a combination of the two [12, 13]. Image registration is widely used in the radiotherapy workflow across a range of anatomical sub-sites including brain, lung, pelvis and H&N [14-17]. In the standard workflow, it is mainly used for target volume definition, such that contemporary consensus guidelines provide clear advice as to which images are recommended to aid this process [18, 19]. Such registration is often referred to as image fusion, and in the management of HNC, common workflows involve fusing either PET, or MRI to planning CT scans.

Most image registration tools in contemporary use seek to minimise differences between geometrical or intensity-based features in each image [20]. Geometrical registration algorithms analyse the precise morphology of shapes in each image, whilst intensity-based approaches look at the absolute grey-scale value in each individual voxel in the image. For CT images, this will be Hounsfield unit information, whilst for MRI it is the relative signal intensity of a given sequence. Most image registration algorithms used in the radiotherapy clinic or research arena are intensity-based. All such software packages perform the same fundamental process, and required the same type of input data:

- **Two images.** A *fixed* image (I_F), which will provide the spatial frame of reference for the *moving* image (I_M), which undergoes some sort of transformation to more accurately match I_F.
- **A transform.** This describes and defines the ways in which I_M can be changed, to maximise similarity with I_F.
- **A similarity metric.** A variable that quantifies the similarity between two images, once the transform has taken place.
- **An optimiser.** An equation coded within the image registration algorithm that iteratively runs through a set of possible transforms to I_M, and seeks to find that which yields the optimal similarity metric result.

A schematic of this process is shown in Figure 4.1. The algorithm's finalised output may be regarded as a set of transformation parameters (T). These define how every voxel in I_M must move, relative to its starting position, in order to maximise similarity with I_F. Mathematically speaking this can be generalised as $\mathbf{X}_{I_F} = T(\mathbf{X}_{I_M})$, and can be visualised as a deformation vector field, as shown in Figure 4.2. For this work, the pCT is always regarded as I_F and the daily IG-MVCT as I_M. If a structure is segmented on I_F, the transformation parameters define which voxels on image I_M represent this structure, thereby creating an 'automated' contour on this image. This is clearly very useful as it provides a mechanism by which OAR segmentation on all MVCTs can be achieved. It also provides a useful tool to train and test a registration algorithm, as a deformed contour can be compared with ground truth segmentation.

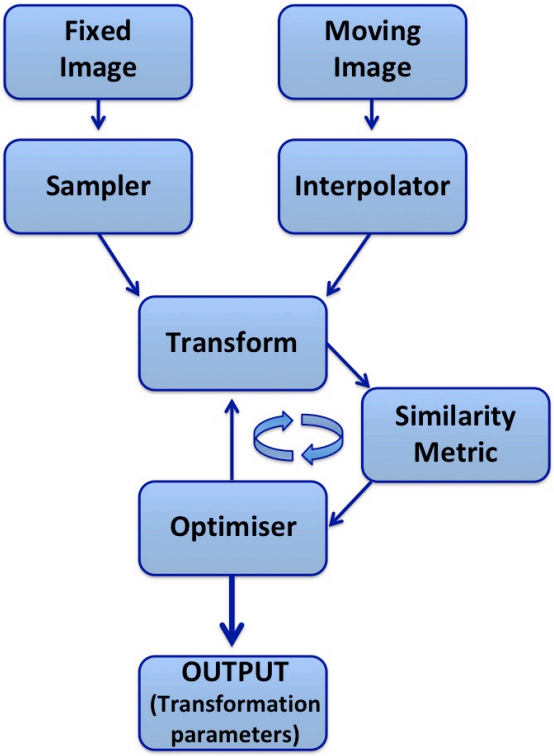


Figure 4.1: Schematic of the image registration process (adapted from similar schematic presented in Lin Yeap – MPhil thesis – 2016)[9]

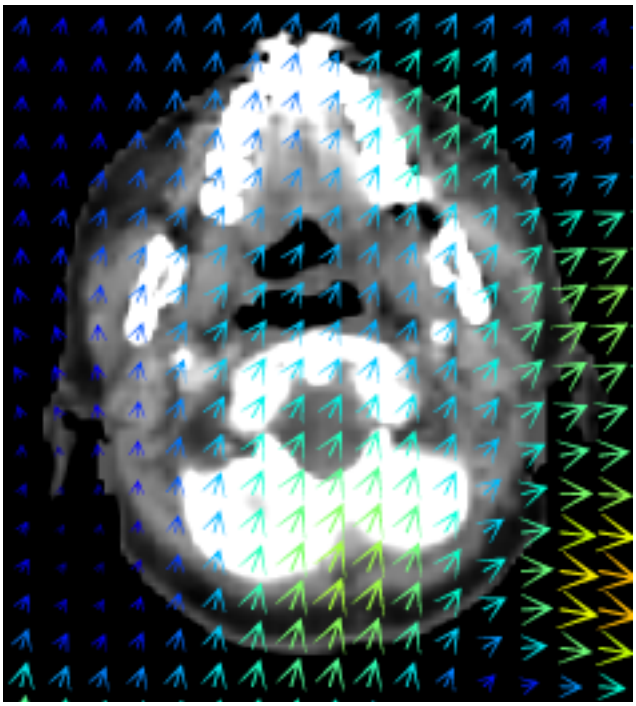


Figure 4.2: An example deformation vector field, with the magnitude and direction of deformations shown as arrows (image courtesy of Megan Wilson, PhD student, UCL).

[9] MPhil Thesis: Mr. Ping Lin Yeap, Jesus College, University of Cambridge, 2016.

4.2.3 Image registration for contour propagation in HNC radiotherapy.

Manual segmentation of planning CT scans in the radiotherapy workflow is laborious, and time-consuming. Furthermore, there is a body of opinion and an evolving literature to suggest that inter-observer variability is the 'weakest link' in this chain [21]. For these reasons, much effort has been devoted to accurate automation of the segmentation process, and specifically OARs. More recently, interest in the use of less supervised artificial intelligence (AI) based approaches has dramatically increased [22-24]. These approaches are promising, and potentially very powerful. However, for the purposes of relevant comparison with work in this chapter, subsequent discussion is limited to DIR-based techniques.

DIR for contour propagation is not new to HNC radiotherapy. A 2008 study compared how well 12 different voxel-based DIR strategies were able to transfer OAR contours from a pre- to mid-treatment kVCT scan [25]. Demons-based, level-set and fast free-form registrations were all tested, as well as combined approaches, and the addition of image pre-processing techniques. Several anatomical structures were segmented for the study, which included parotid and submandibular glands and the spinal canal, but not pharyngeal constrictor muscles. Results were pooled across all structures, and showed that all deformable approaches outperformed rigid registration, with median DSC scores between 0.78 and 0.86.

There are a number of commercially available solutions for DIR-based automated H&N OAR segmentation. These platforms mainly adopt one of two commonly used approaches to the problem. The first is referred to as Atlas-Based Automated Segmentation – ABAS. In this approach, a library of representative cases is uploaded to a database, usually containing around 20 scans. The patient's planning scan is then compared with the images in the library, and the closest match found. Thereafter, a DIR algorithm performs a registration to warp the geometry of the library image to the patient's scan, and uses the resulting transformation parameters to transfer contours across. The second strategy, known as model-based segmentation, aims to compensate for limitations in imaging quality by imposing statistically determined shape constraints, derived from ground truth segmentations [26]. Pure model-based approaches may not include DIR, but many applications of this technique adopt a hybrid strategy, in which anatomical models and image registration are combined to provide a finalised segmentation [27].

There are two relatively recent reviews, specifically addressing the auto-segmentation of H&N OARs using these techniques [28, 29]. In order to provide important context, the data published in the papers cited by these reviews [26, 30-35], where relevant to the OARs studied in this chapter, are summarised in Table 4.1. A more recent study compared the

Publication			Technical details			Results		
Ref	Year	1st author	Sample size	DIR approach	Commercial/ in-house	PGs	SMGs	PCs
29	2010	Tsuji	16	Free-form	MIMvista	DSC: 0.72	N/A	N/A
25	2011	Qazi	10	Demons	SPICE 9.8/ In-house	DSC: 0.83 HD: 5.7-5.8	DSC: 0.81-0.84 HD: 3.5-4.2	N/A
30	2011	Teguh	12	Not specified	Elekta-CMS ABAS	DSC: 0.79 MSD: 2.5	DSC: 0.70 MSD: 1.9	DSC: 0.50 MSD: 2mm
31	2013	Daisne	10	Elastic fusion	Brainlab	DSC: 0.72 MSD: 2.3	N/A	N/A
32	2014	Fritscher	18	Demons (+ geodesics)	In-house	DSC: 0.81-0.84 HD: 10-12	N/A	N/A
33	2014	Walker	40	ABAS-model hybrid	Pinnacle 9.4	DSC: 0.89	DSC: 0.73	DSC: 0.93
34	2014	Thomson	10	ABAS-model hybrid	SPICE (Pinnacle 9.4)	DSC: 0.78 MSD: 1.6 HD: 14.7	DSC: 0.70 MSD: 1.5 HD: 7.2	DSC: 0.57 MSD: 1.5 HD: 13

Table 4.1: Summary of results from publications using DIR-based automated segmentation for H&N OARs. Results are only considered for OARs considered in this Chapter; parotid glands (PGs), submandibular glands (SMGs) and pharyngeal constrictor muscles (PCs). Reported metrics are; dice similarity coefficient (DSC), mean surface distance (MSD, in mm), and Hausdorff distance HD, in mm).

performance of 6 commercially available DIR platforms across 13 Italian centres [36]. The results show DSC scores of 0.79 – 0.87 for the parotid glands.

The data across all these studies show a number of clear trends. First, that DIR is a well-established technique for OAR contour propagation in H&N radiotherapy (published data on DIR from kVCT to image guidance scans has not been addressed, and is discussed in section 4.6.1). Second, that deformable registration consistently outperforms rigid alternatives across a range of patient datasets, software platforms, and observers. Third, that results are often good but not perfect, and these tools need to be implemented with due care and attention. Nonetheless, a synthesis of the literature supports DIR as an established and evidence-based methodology for automated segmentation of H&N OARs.

4.2.4 Image registration – Elastix

Karl Harrison, Lin Yeap and I made the decision to use an open source image registration toolbox – Elastix [37, 38]. There were a number of reasons for this decision. First, the software is readily implemented within the python programming language, which made interaction and communication with other software and data management frameworks within the project straightforward [2]. Second, Elastix has a range of transformation toolboxes available, including rigid and non-rigid options. In order to generate robust optimisation data, we tested a wide range of image registration approaches. Finally, deformable registration toolbox within Elastix uses local spline-based (B-spline) transformation models [39, 40], there is extensive precedent for using such approaches in HNC radiotherapy [5, 25, 41, 42], and there is literature precedent for using Elastix with H&N imaging [43].

Elastix has the following transformation models as options [37]:

- **Translation** – I_M can be moved in 3 dimensions (x,y,z), but no rotation is permitted.
- **Rigid** – In addition to translation, I_M can be rotated relative to I_F.
- **Similarity** – I_M may be translated, rotated and scaled.

- **Affine** – permits translation, rotation, scaling and sheer of I_M [44, 45]. However, the relationship between neighbouring voxels remains the same, and parallel lines in I_M pre- and post-transform remain parallel [20].
- **B-spline (deformable)** – allows individual voxels in I_M to be moved relative to local neighbours – deformation [46]. Theoretically, this gives the deformation 3N degrees of freedom, where N = the number of voxels in the image [20]. However, the algorithm is constrained by a grid of control points, and the distance (spacing) between points on this grid restricts the distance that individual I_M voxels can move [38]. In practice, the spline function models spatial transformations between points on the grid – a parameter that the operator can optimise. If the control point spacing is too large, the algorithm may be unable to resolve details of small structures. If spacing is too small, the transform has many more degrees of freedom, which risks erratic transformations [47].

Elastix permits use of a number of different mathematical functions to produce a similarity metric between I_F and I_M once it has undergone the transform [48]. Four widely investigated and commonly used approaches were tested in this work, and are described here. These functions are presented in order of increasing complexity, with fewer assumptions about the underlying relationship between absolute voxel intensity values in each image.

- **Mean squared difference (MSD)** – assumes an equivalent relationship between voxel intensities in I_F and I_M [20, 49]. The similarity metric is the sum of the squared differences of all voxels, and the optimiser seeks to minimise this number. It is the most straightforward approach mathematically and conceptually, and the easiest to compute. It performs well when registering images of the same modality, for example four-dimensional CT [50, 51], but the underlying assumptions in this method mean it is not well suited for registering different modality images [38].

- **Normalised cross correlation (NCC)** – similar to MSD, but mathematically less strict as it assumes a linear relationship, rather than equivalence, between voxel intensities in the two images [20]. Despite normalisation of the coefficient,

a limitation of this technique is that the contribution from an individual pixel is strongly dependant on its absolute intensity value [52]. The approach has been widely used in for matching portal images and digitally reconstructed radiographs (DRRs) [53-55], but, like MSD, its underlying assumptions mean it is best suited to registering images of the same modality [20]. Despite this, it has been used to register MR and CT images of the head [56].

- **Mutual information (MI)** – this similarity metric aims to "align voxels with common probabilities of being present in their respective image sets." [20]. In other words, MI assumes a relationship between the statistical properties of a given voxel of a given intensity in I_M, being in a given location, relative to I_F. No assumptions are made about the nature of this relationship, and can be implemented without prior modifications to either image [57, 58]. As a result, it is better suited to registrations between different image modalities [38]. This technique has been successfully deployed for registering PET and SPECT with CT [59, 60], CT and MRI [61, 62], and perhaps most pertinently to this work, between CT and TomoTherapy MVCT [63].

- **Normalised mutual information (NMI)** – there is normalisation of the MI metric, based on the quantity of shared information between the 2 images. This approach is similar to MI, and has been hypothesised to have potential advantages in the presence of soft tissue compression, at the expense of greater algorithmic complexity [64].

The Elastix toolbox also permits users to make choices with regards to optimiser (the cost function equation), image sampler (whether or not all voxels, or a sample from each image, are used for registration), and interpolation (calculated intensities at non-voxel coordinates, relevant in cases where voxel sizes in I_F and I_M are not identical). These choices are largely a trade-off between computational cost and time, and accuracy. Work to inform this decision was done by Lin Yeap with little input from myself. It is described in his MPhil thesis, and for that reason is not described in any

further detail here[10]. The final decision as implemented in Elastix for all subsequent work described in this thesis is as follows:

- Optimiser: gradient descent
- Image sampler: random coordinate
- Interpolator: linear

4.2.5 Aims of chapter 4

1. To optimise and validate Elastix for the purposes of contour propagation from pCT to daily MVCT.
2. To define a workflow for subsequent dose accumulation.
3. To quantify uncertainties associated with these processes.
4. To finalise Elastix parameters for computation of D_A in the full patient cohort.

4.3 Methods

4.3.1 Background

Work on optimising Elastix for automated segmentation of the spinal cord was completed in 2016. In 2017, a consensus guideline "Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132" was published in *Med. Phys.* by Kristy Brock et al. [20]. This document (referred to hereafter as 'TG 132') comprehensively describes and defines how image registration and fusion algorithms such as Elastix should be trained, tested and validated, and states that such work should be specific to the task that the tool is being used for.

Using deformable registration for contour propagation to the swallowing OARs is a different task to the spinal cord. In contrast to the spinal cord, these structures have more complex shapes [65], and are more likely to undergo significant morphological change during treatment [66-69]. Furthermore, the anatomical context of these structures is different to the spinal cord; they are not surrounded by a hollow cylinder of bone. It was therefore decided to fully re-train, test and validate Elastix for the

[10] MPhil Thesis: Mr. Ping Lin Yeap, Jesus College, University of Cambridge, 2016.

purpose of contour propagation (and subsequent dose accumulation) to the swallowing OARs.

The overall approach to this task was similar to that used for the spinal cord. However, due to the greater complexity of this task, the methodology was also more detailed. The quantitative metrics that should be considered in assessing the performance of deformable image registration are summarised in Table III in Brock et al [20], and are reproduced below in Table 4.2.

The workflow as implemented only requires registration in one direction, with the pCT as I_F and MVCT as I_M. For this reason, consistency was not formally tested, but all other metrics listed in Table 4.2 were assessed, along with other measures of contour conformity.

Metric	Description	Target performance
Target registration error (TRE) [70]	Algorithms ability to accurately match an implanted fiducial, or naturally occurring landmark, between images	Maximum dimension of imaging voxels (2-3mm)
Mean distance to agreement (MDA) [71]	The mean distance between the surface of contours from two images, following registration	Consistent with inter-observer contouring variability (2-3mm)
Dice similarity coefficient (DSC) [72]	The proportion of 2 structures that overlap, relative to their total combined volume.	Consistent with inter-observer contouring variability (0.8 – 0.9)
Jacobian determinant [64]	Expansion or contraction of individual voxels in I_M, following registration. (1 = identical, >1 = expansion, <1 = shrinking, <0 image folding – non-physical)	No values <0. Values either side of 1 as expected clinically.
Consistency [73]	Is an algorithm independent to the direction of registration - does defining image A as I_F and B as I_M produce a different result to the inverse procedure?	Maximum dimension of imaging voxels (2-3mm)

Table 4.2: Metrics to quantify the performance of a registration algorithm – adapted from Brock et al, TG 132 [20].

4.3.2 Patient image data

Imaging data from 14 patients was used for this work. The sample size was defined on the basis of literature precedent [41], expert opinion, and a pragmatic balance between maximising available data, and time taken for image segmentation, and data from patient's in both discover and consolidation cohorts were used. All patients were treated with 65Gy in 30 fractions for SCC of the oropharynx, hypopharynx or larynx. Radiotherapy planning kVCT images were acquired with a GE Medical Systems LightSpeed Multislice CT scanner or a Toshiba Aquilon LB CT scanner, with no intravenous contrast, and a voxel size of 1.074 x 1.074 x 3mm. In the RT planning process, these images are down-sampled axially by a factor of 2 by the TomoTherapy TPS. Thus the voxel size for all kVCT images used for work in this thesis is 2.148 x 2.148 x 3mm, unless otherwise stated. The voxel size for TomoTherapy MVCT IG images is 0.754 x 0.754 x 6mm. All images were retrieved from TomoTherapy archive by Marina Romanchikova, and loaded in the research database of the clinical radiotherapy segmentation tool (Prosoma v3.3, MEDCOM, Darmstadt, Germany), for segmentation, and transfer to the Cavendish laboratory.

4.3.3 Image segmentation, intra- and inter-observer variability

The structures segmented in this work were:

- Bilateral parotid glands
- Bilateral submandibular glands
- Pharyngeal constrictor muscles (SPC and MPC combined) as a single structure)

These structures were segmented on down-sampled kVCT and IG-MVCT images according to published atlases [65]. For 4 patients, all 5 structures were contoured on all 30 MVCT scans. They were also segmented on the final (day 30) MVCT for a further 10 patients. The final fraction MVCT was chosen on the basis that it would likely show the greatest anatomical change relative to the pCT, and therefore provide the most robust test for Elastix.

To generate intra-observer variability data, I repeated segmentation of the full structure set, at an interval of 4-12 weeks. An example is shown in Figure 4.3 (A). The following MVCTs were contoured a second time:

- 2 patients – all 30 MVCTs
- 2 patients – 7 MVCTs (treatment days 1, 6, 11, 16, 21, 26, 30)
- 10 patients – final fraction (day 30) MVCT.

An inter-observer contouring study was again undertaken. Four consultant oncologists, all specialists in HNC radiotherapy, were recruited to provide reference contours. Zoom and image window settings were fixed. Participants were asked to follow a bespoke protocol for this task (Appendix A4.2), and to contour 2 MVCT images each for 2 patients (4 scans in total). Images showing segmentations of all structures for all 4 observers are shown in Figure 4.3 (B). Sample sizes for both intra- and inter-observer contouring variability data were based partly on expert opinion, partly on a pragmatic balance between sample size and time taken to generate contours.

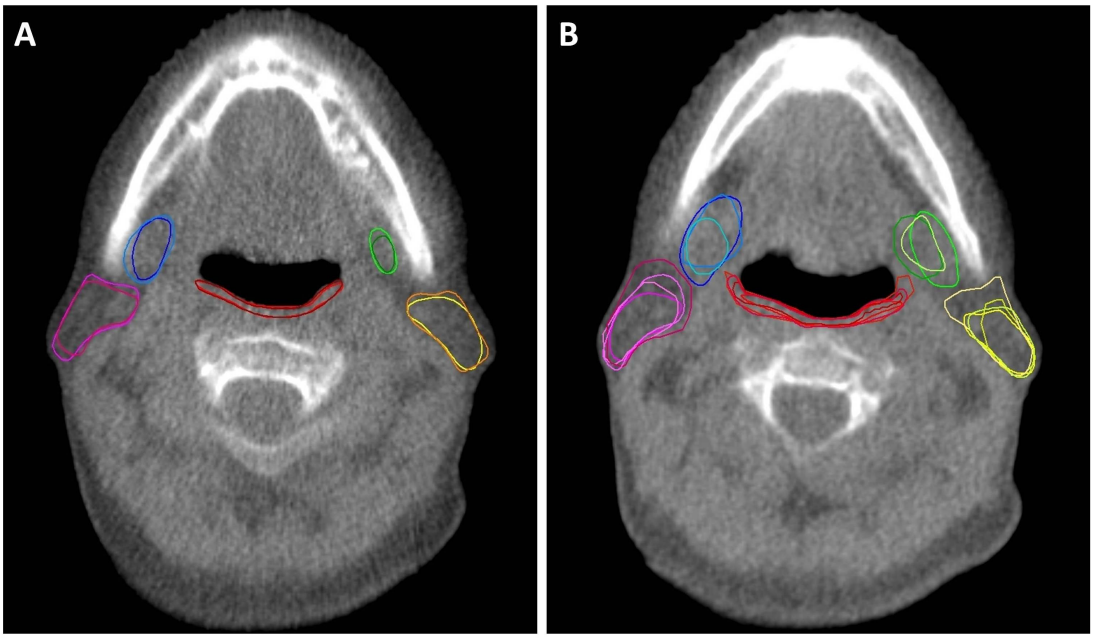


Figure 4.3 (A&B): Variability in swallowing OAR structure segmentation on H&N MVCT images. A – Intra-observer variability. B – Inter-observer variability (contours from 4 expert observers).

A combination of the following metrics was used to measure agreement between contours:

- Dice similarity coefficient (DSC) [72], Equation 4.1. This is conceptually similar but mathematically different to JCI, and gives higher scores.

$$Dice\ Similarity\ Coefficient = \frac{2|A \cap B|}{|A|+|B|} \tag{4.1}$$

- Jaccard conformity index (JCI) – the ratio of the intersection of 2 contours (A & B) to their combined area [74]. This is defined in Equation 4.2 as:

$$Jaccard\ Conformity\ Index = \frac{|A \cap B|}{|A \cup B|} \tag{4.2}$$

- Mean distance to agreement (MDA, mm) – the mean value of every possible nearest distance between surface voxels in both contours as shown in Figure 4.4 [20, 35, 71]. The metric is also known as mean surface distance, is deployed as such within Elastix and python [37], and the mathematical formulation is shown in Equation 4.3:

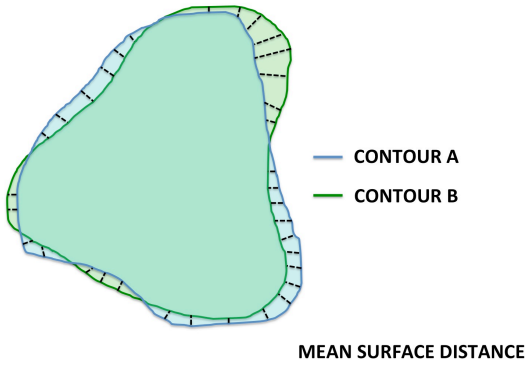


Figure 4.4: Cartoon showing concept of all possible minimum distances between voxels in contours A and B, from which mean distance to agreement (MDA) is calculated

$$MSD\ (MDA) = \frac{1}{n_S+n_{S'}}\left(\sum_{p=1}^{n_S} d(p,S')^2 + \sum_{p'=1}^{n_{S'}} d(p',S)^2\right) \tag{4.3}$$

- Hausdorff distance (HD) – the biggest possible distance in the set of all distances, where all possible points on contour A find the closest possible point on contour B. Simply put it is the biggest point difference between 2 contours, as shown in Figure 4.5.

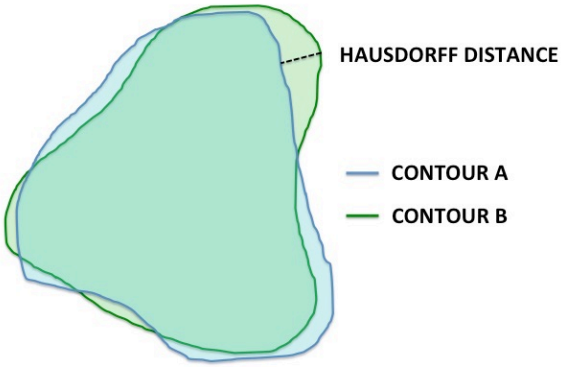


Figure 4.5: Cartoon showing the greatest possible minimum distance between contours A and B – Hausdorff distance.

- 95th centile surface distance – conceptually similar to HD, but takes the 95th centile value of the histogram of distances, rather than the maximum.
- Distance to conformity (DTC) – the maximum distance (u) from a point on the edge of (A ∩ B) to the edge of (A ∪ B) [75]. Standard convention suggests that if B is larger than A (or over-contoured), then u is positive. If B is smaller, then u is negative. This metric therefore gives an overall impression of whether or not contours are systematically larger or smaller than a reference, as shown in Figure 4.6.

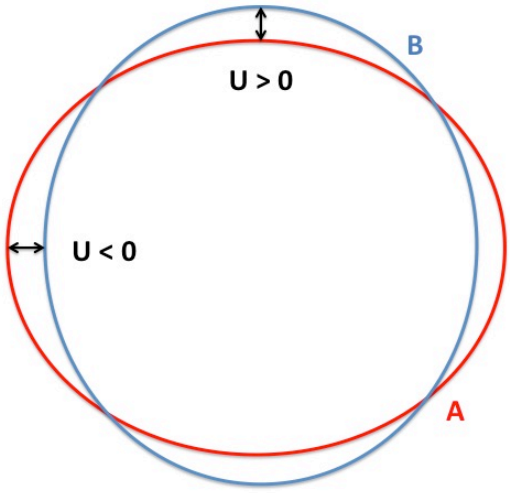


Figure 4.6: Schematic of distance to conformity (DTC) measurement

- Distance between centres (DBC) – over a sample, gives an indication of whether or not contours are systemically shifted in any of four axial directions relative to a reference [75].

- Proportional volume difference (PVD – {%}, Equation 4.4) – The volume difference between the contours. This metric does not infer one contour as being a 'gold-standard'. Rather, it measures the difference in volume between contour A (V_{CA}) and the mean of the 2 contours (V_{CA} and V_{CB}), as a function of the mean volume of these 2 contours, as shown below (NB – larger contour always regarded as contour A to ensure sign uniformity).

$$Proportional\ volume\ difference\ (\%) = \left\{ \frac{V_{CA} - \left(\frac{V_{CA}+V_{CB}}{2}\right)}{\left(\frac{V_{CA}+V_{CB}}{2}\right)} \right\} x\ 100 \qquad (4.4)$$

4.3.4 Elastix parameter tuning

Imaging and structure set data from 3 of the 4 patients for whom the parotids, submandibular glands and pharyngeal constrictors were segmented on every MVCT (3 x 30 scans – 90 in total) were used for algorithm training. The method was similar to that used for the spinal cord work. Experiments in which the following parameters were changed were undertaken:

- Registration strategy:
    - Radiographer couch shifts – the MVCT is aligned with the kVCT only by using the couch shifts implemented by the treating RTTs, for that treatment fraction.
    - Translation
    - Rigid (translation plus rotation)
    - Similarity
    - Affine
    - B-spline based deformable registration – with variation of control point spacing.
- B-spline control point spacing: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 500mm

- Similarity metrics:
    - Mean squared distance (MSD)
    - Normalised cross correlation (NCC)
    - Mutual information (MI)
    - Normalised mutual information (NMI)

To train these parameters within the Elastix toolbox, the effect of these changes on two sets of outcomes were assessed. Firstly, conformity metrics between my reference contours, and those derived from the image registration. The conformity metrics listed in the previous section were used.

Following TG 132 [20], the impact of altering these parameters on target registration error (TRE) was also assessed. A literature search did not reveal a clear consensus on which anatomical points to define, although a combination of bony and soft tissue landmarks was used in many studies [76-78]. Nonetheless, TG 132 does specify that these points need to be robust, reproducible and anatomically relevant to the underlying purpose of the registration, in this case contour propagation to swallowing OARs. In addition to these criteria, it was apparent that chosen points would need to have a good spread in the superior-inferior dimension, and be clearly identifiable on both a down-sampled pCT, and noisy MVCT images. On this basis, the following 7 anatomical points were chosen for TRE analysis:

- The inferior aspect of both Pterygoid plates - bilaterally
- Parotid gland parenchyma immediately adjacent to the inferior aspect of the mastoid process - bilaterally
- The most anterior point of the parotid gland - bilaterally
- Soft tissue immediately adjacent to the most anterior extent of the body of C2 vertebra, immediately superior to the C2/3 junction.

These points were segmented on kVCT planning scans, and 7 MVCT images (days 1, 6, 11, 16, 21, 26, 30), for 5 patients, and examples on both down-sampled pCT and MVCT images are shown in Figure 4.7 (A-C).

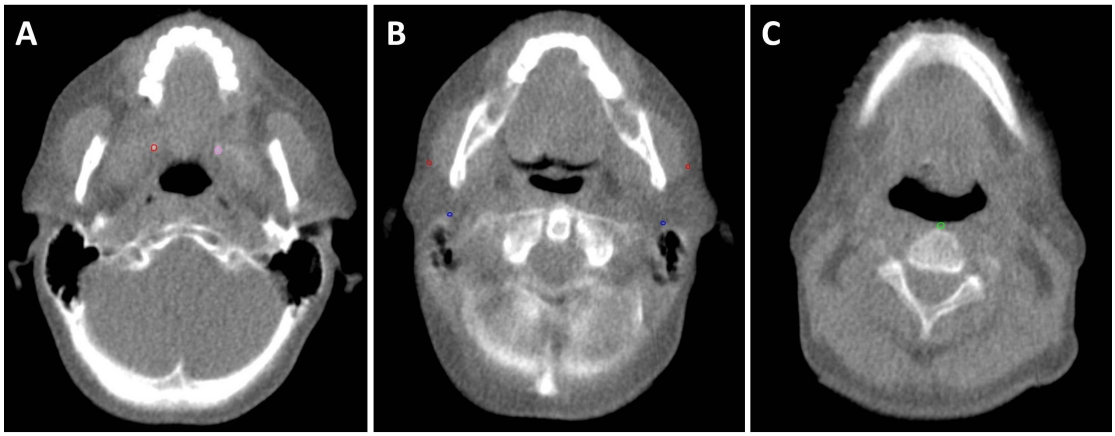


Figure 4.7 (A-C): Target registration error anatomical fiducial points. A – Inferior aspect of Pterygoid plates. B – Parotid gland; most anterior point, and immediately adjacent to Mastoid process. C – C2/3 junction.

4.3.5 Elastix parameter validation

Once the choice of Elastix parameters had been finalised, this registration strategy was tested on a separate dataset. One of the patients with all MVCTs segmented, and the 10 patient dataset with the final fraction MVCT were held back for validation (30 + 10 scans – 40 in total). This was done by producing contour conformity metrics, as described above. Analysis of TRE was also repeated on a smaller validation dataset. All 7 anatomical fiducial points were segmented on the planning kVCT and a single randomly selected MVCT for a further 5 patients.

Finally, mean Jacobian determinants for registrations between the kVCT and all 30 MVCTs, for 2 patients, were also calculated. It is difficult to interpret the results of Jacobian determinant analysis in terms of precise spatial accuracy [20], and therefore the metric has little utility in parameter tuning. However, it does show whether or not a registration is broadly anatomically and biologically plausible, and for that reason it provides a useful ‘sense check’ of algorithm output in the validation process.

4.4 Results

4.4.1 Intra- and inter-observer contouring variability

Concordance data between my first and second set of contours of the parotids, submandibular glands, and pharyngeal constrictor muscles are shown in table 4.3. These data are derived from the following patient sample:

- 2 patients where each MVCT was segmented twice (2 x 30 scans),
- 2 patients where 7 scans were re-contoured (2 x 7 scans),
- 10 patients where 1 scan was re-contoured (10 x 1 scan).

Overall these data show good concordance and consistency between the 2 contours. Taking all 84 MVCT scans that were segmented twice, mean DSC values for the parotid glands, submandibular glands and pharyngeal constrictor muscles respectively were 0.89, 0.83 and 0.76. 95% of all possible points on contour one were within 2.4mm (PGs), 3.1mm (SMGs) and 1.7mm (PCs) of the nearest neighbour on contour two. Mean volume differences between repeated contours were all less than 7.5%, and mean Hausdorff distances were 6.7, 5.6 and 6.1mms for parotids, submandibular glands and pharyngeal constrictor muscles respectively.

Results of the inter-observer variability contouring study are shown in table 4.4, and summary results for all intra-observer contouring variability data are shown in table 4.5. As expected, concordance between observers is less good than that observed with 2 contours from the same observer. Mean, combined, inter-observer DSC values were between 0.82 for parotid glands, and 0.66 for the submandibular glands. Mean DSC and JCI values were lower for the pharyngeal constrictors - 0.55 and 0.38 respectively. However, this is to be expected, as the pharyngeal constrictor muscles are long, thin, ‘U-shaped’ structures, meaning that a very small translation of an identical contour relative to its twin will result in a rapid fall in conformity metrics. The DTC data for the pharyngeal constrictors suggest that whilst observers tend to systematically segment this structure larger or smaller than colleagues, typical distances between the edges of contours are small.

	30 timepoints			7 timepoints			1 timepoint			Total		
Scans for analysis	60			14			10			84		
Structure	PGs	SMGs	PCs	PGs	SMGs	PCs	PGs	SMGs	PCs	PGs	SMGs	PCs
Median DSC	0.90 (±0.003)	0.83 (±0.006)	0.76 (±0.004)	0.92 (±0.005)	0.83 (±0.01)	0.75 (±0.02)	0.88 (±0.008)	0.84 (±0.01)	0.79 (±0.01)	0.90 (±0.002)	0.83 (±0.005)	0.76 (±0.005)
Mean DSC	0.89 (±0.002)	0.82 (±0.005)	0.76 (±0.005)	0.91 (±0.004)	0.82 (±0.008)	0.75 (±0.02)	0.88 (±0.006)	0.84 (±0.01)	0.79 (±0.01)	0.89 (±0.002)	0.83 (±0.004)	0.76 (±0.004)
Median JCI	0.81 (±0.004)	0.71 (±0.009)	0.61 (±0.006)	0.85 (±0.009)	0.71 (±0.01)	0.59 (±0.03)	0.79 (±0.01)	0.73 (±0.02)	0.65 (±0.02)	0.82 (±0.004)	0.71 (±0.007)	0.62 (±0.007)
Mean JCI	0.81 (±0.003)	0.70 (±0.007)	0.62 (±0.005)	0.84 (±0.007)	0.70 (±0.01)	0.60 (±0.02)	0.79 (±0.01)	0.73 (±0.01)	0.66 (±0.02)	0.81 (±0.003)	0.71 (±0.006)	0.62 (±0.005)
Mean MSD (mm)	0.79 (±0.01)	1.1 (±0.03)	0.58 (±0.02)	0.80 (±0.03)	1.1 (±0.06)	0.62 (±0.04)	0.87 (±0.04)	1.0 (±0.06)	0.44 (±0.04)	0.80 (±0.01)	1.1 (±0.02)	0.57 (±0.02)
Mean 95%SD (mm)	2.4 (±0.05)	3.1 (±0.08)	1.7 (±0.08)	2.4 (±0.1)	3.1 (±0.2)	2.0 (±0.2)	2.7 (±0.2)	2.7 (±0.2)	1.3 (±0.1)	2.4 (±0.4)	3.1 (±0.02)	1.7 (±0.06)
Mean DBC (mm)	1.3 (±0.06)	2.0 (±0.1)	1.7 (±0.2)	1.2 (±0.1)	1.9 (±0.2)	2.2 (±0.4)	1.5 (±0.2)	1.6 (±0.2)	1.8 (±0.4)	1.3 (±0.05)	1.9 (±0.1)	1.8 (±0.1)
Mean DTC (mm)	-0.025 (±0.007)	0.013 (±0.007)	-0.088 (±0.02)	-0.006 (±0.01)	-0.009 (±0.02)	-0.017 (±0.03)	-0.025 (±0.02)	-0.016 (±0.02)	-0.003 (±0.03)	-0.022 (±0.005)	0.005 (±0.007)	-0.066 (±0.01)
Mean HD (mm)	6.6 (±0.2)	5.7 (±0.1)	6.2 (±0.2)	6.5 (±0.5)	6.0 (±0.4)	6.0 (±0.4)	7.6 (±0.5)	4.9 (±0.3)	5.8 (±0.7)	6.7 (±0.2)	5.6 (±0.1)	6.1 (±0.2)
Mean PVD (%)	3.6 (±0.3)	5.0 (±0.4)	8.5 (±0.7)	2.3 (±0.4)	4.4 (±0.7)	3.8 (±0.8)	4.6 (±0.8)	4.8 (±0.8)	5.5 (±1.4)	3.5 (±0.2)	4.9 (±0.4)	7.4 (±0.6)

Table 4.3: Intra-observer contouring variability for parotid glands (PGs), submandibular glands (SMGs) and pharyngeal constrictors (PCs). Dice similarity coefficient (DSC), Jaccard conformity index (JCI), mean surface distance (MSD), 95% centile surface distance (95%SD), distance between centres (DBC), distance to conformity (DTC), Hausdorff distance (HD) and proportional volume difference (PVD) are reported. Standard errors reported in parentheses.

	Observer 1			Observer 2			Observer 3			Observer 4		
Structure	PGs	SMGs	PCs	PGs	SMGs	PCs	PGs	SMGs	PCs	PGs	SMGs	PCs
Median DSC	0.83 (±0.008)	0.67 (±0.03)	0.55 (±0.01)	0.83 (±0.03)	0.69 (±0.03)	0.56 (±0.03)	0.82 (±0.01)	0.76 (±0.04)	0.55 (±0.03)	0.80 (±0.01)	0.67 (±0.03)	0.51 (±0.01)
Mean DSC	0.82 (±0.007)	0.64 (±0.02)	0.54 (±0.01)	0.76 (±0.03)	0.67 (±0.03)	0.57 (±0.02)	0.81 (±0.01)	0.69 (±0.04)	0.57 (±0.03)	0.80 (±0.009)	0.64 (±0.02)	0.52 (±0.01)
Median JCI	0.70 (±0.01)	0.50 (±0.03)	0.37 (±0.01)	0.70 (±0.04)	0.53 (±0.04)	0.39 (±0.03)	0.70 (±0.02)	0.67 (±0.01)	0.38 (±0.03)	0.82 (±0.004)	0.50 (±0.03)	0.34 (±0.01)
Mean JCI	0.70 (±0.01)	0.48 (±0.02)	0.37 (±0.01)	0.63 (±0.03)	0.52 (±0.04)	0.41 (±0.02)	0.68 (±0.02)	0.66 (±0.01)	0.40 (±0.03)	0.81 (±0.003)	0.47 (±0.02)	0.35 (±0.009)
Mean DBC (mm)	1.4 (±0.2)	2.4 (±0.3)	1.4 (±0.3)	2.3 (±0.3)	3.0 (±0.3)	1.2 (±0.3)	1.6 (±0.3)	2.4 (±0.4)	1.2 (±0.2)	1.9 (±0.3)	2.9 (±0.4)	2.0 (±0.3)
Mean DTC (mm)	0.04 (±0.06)	-0.46 (±0.2)	0.10 (±0.1)	-0.08 (±0.07)	-0.86 (±0.2)	-0.50 (±0.1)	0.34 (±0.1)	1.2 (±0.2)	-0.36 (±0.1)	-0.30 (±0.09)	0.1 (±0.2)	0.76 (±0.1)
Mean HD (mm)	12.5 (±0.5)	12.6 (±0.9)	14.9 (±0.7)	12.0 (±0.7)	12.5 (±0.7)	11.7 (±0.6)	12.6 (±0.9)	12.6 (±0.9)	13.7 (±0.8)	12.7 (±0.7)	12.3 (±0.9)	13.9 (±1.0)
Mean PVD (%)	7.1 (±1.0)	14.4 (±2.6)	18.3 (±2.0)	7.1 (±1.3)	16.1 (±2.8)	24.0 (±3.9)	9.5 (±1.7)	23.1 (±3.0)	18.2 (±3.1)	9.4 (±1.5)	15.8 (±2.0)	27.4 (±3.3)

Table 4.4: Inter-observer contouring variability for parotid glands (PGs), submandibular glands (SMGs) and pharyngeal constrictors (PCs) – comparing data for each observer independently. Dice similarity coefficient (DSC), Jaccard conformity index (JCI), distance between centres (DBC), distance to conformity (DTC), Hausdorff distance (HD) and proportional volume difference (PVD) are reported. Standard errors reported in parentheses. My contours are labeled as Observer 1.

	PGs	SMGs	PCs
Median DSC	0.82 (± 0.008)	0.66 (± 0.02)	0.54 (± 0.02)
Mean DSC	0.81 (± 0.006)	0.64 (± 0.02)	0.55 (± 0.01)
Median JCI	0.69 (± 0.01)	0.49 (± 0.005)	0.37 (± 0.02)
Mean JCI	0.68 (± 0.009)	0.48 (± 0.02)	0.38 (± 0.01)
Mean DBC (mm)	1.7 (± 0.2)	2.8 (± 0.2)	1.4 (± 0.2)
Mean DTC (mm)	0.08 (± 0.06)	-0.49 (± 0.2)	-0.27 (± 0.12)
Mean HD (mm)	12.6 (± 0.5)	13.3 (± 0.7)	13.5 (± 0.6)
Mean PVD (%)	8.3 (± 1.0)	17.4 (± 1.9)	22.0 (± 2.3)

Table 4.5: Inter-observer contouring variability for parotid glands (PGs), submandibular glands (SMGs) and pharyngeal constrictors (PCs) – summary results (all observers compared against each other).

4.4.2 Elastix parameter tuning

4.4.2.1 Target registration error - TRE

TRE values were calculated for all 7 anatomical surrogate fiducials, at 7 timepoints, for 5 patients. The effect of changing registration strategy, similarity metric, and B-spline control point spacing (for deformable registrations) were investigated. Results are presented as clustered bar charts, in anatomical order, superior to inferior in Figures 4.8– 4.11.

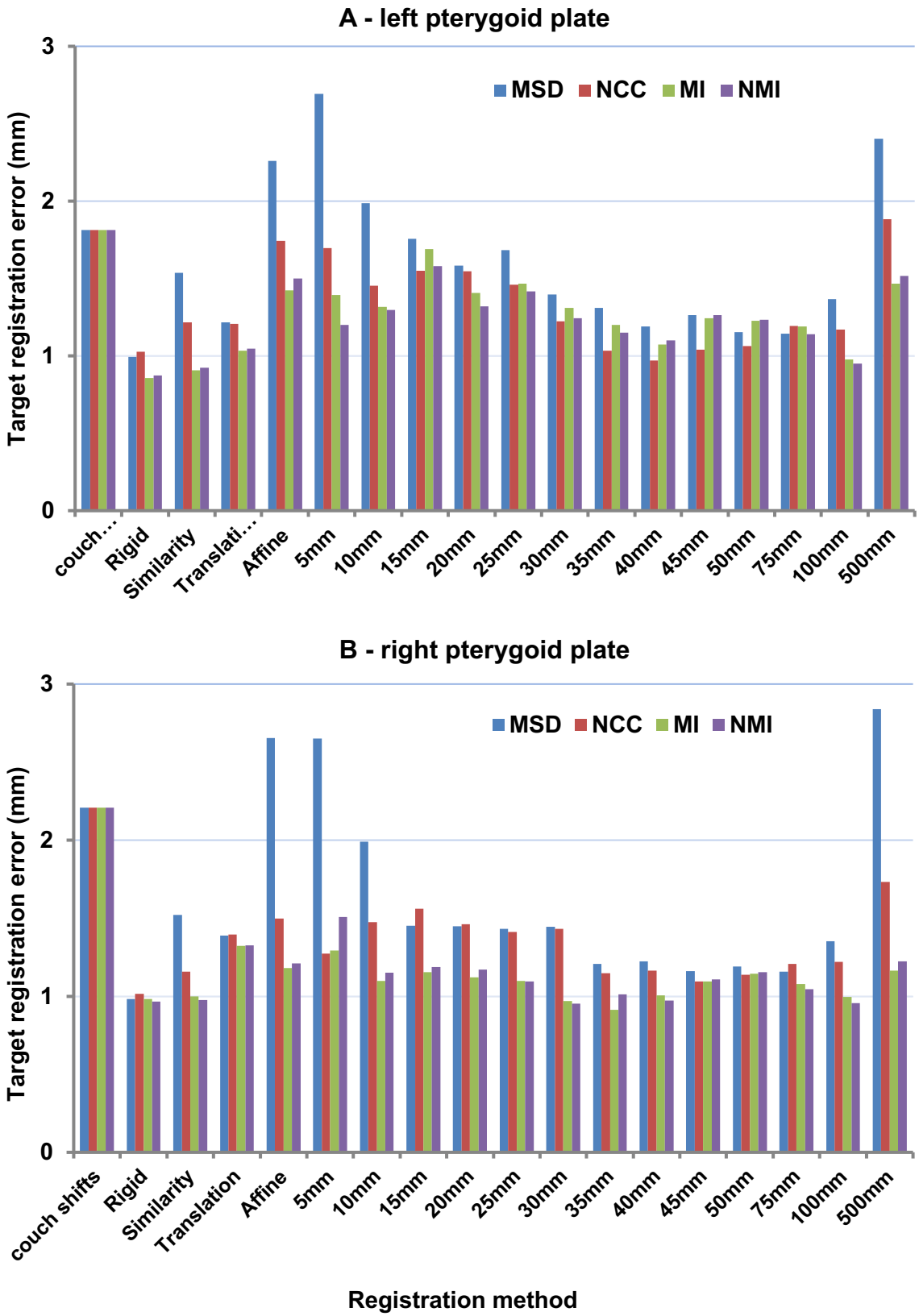


Figure 4.8 (A&B): Target registration error (TRE) results for different Elastix registration strategies, similarity metrics and control point spacing's for the Pterygoid plates, A – left, B – right.

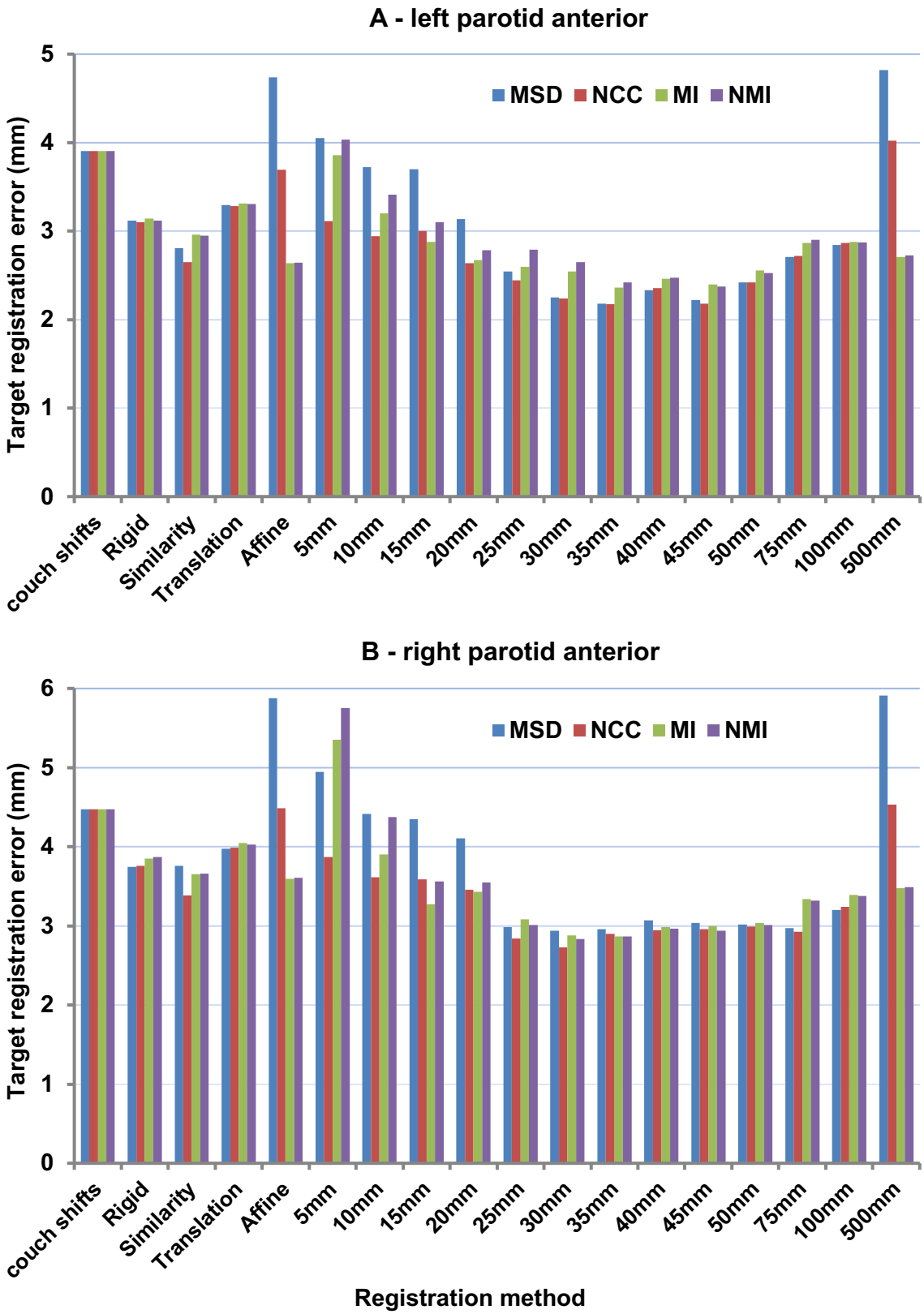


Figure 4.9 (A&B): Target registration error (TRE) results for different Elastix registration strategies, similarity metrics and control point spacing's for the 'anterior parotid' point, A – left, B – right.

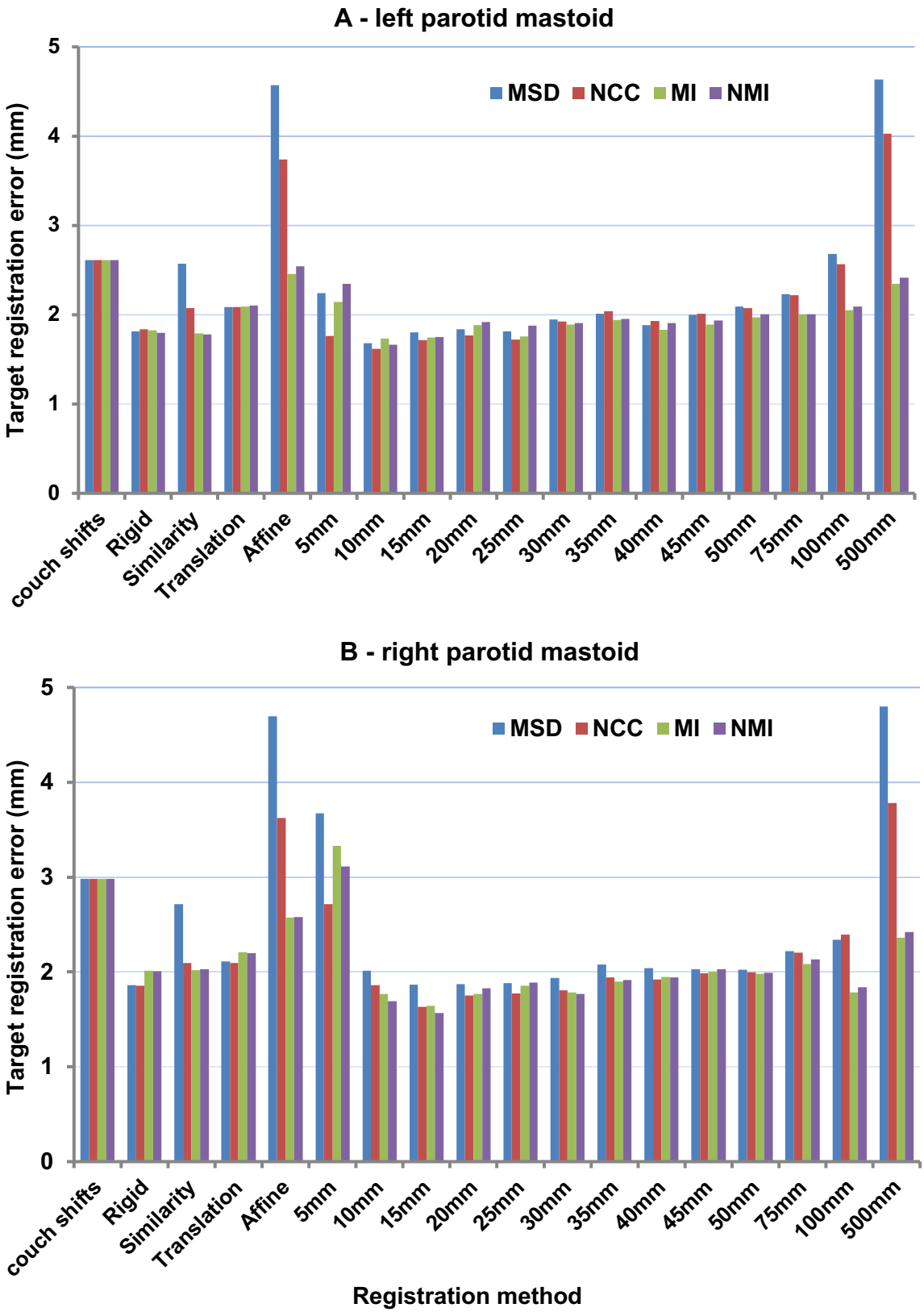


Figure 4.10 (A&B): Target registration error (TRE) results for different Elastix registration strategies, similarity metrics and control point spacing's for the 'parotid mastoid' point, A – left, B – right.

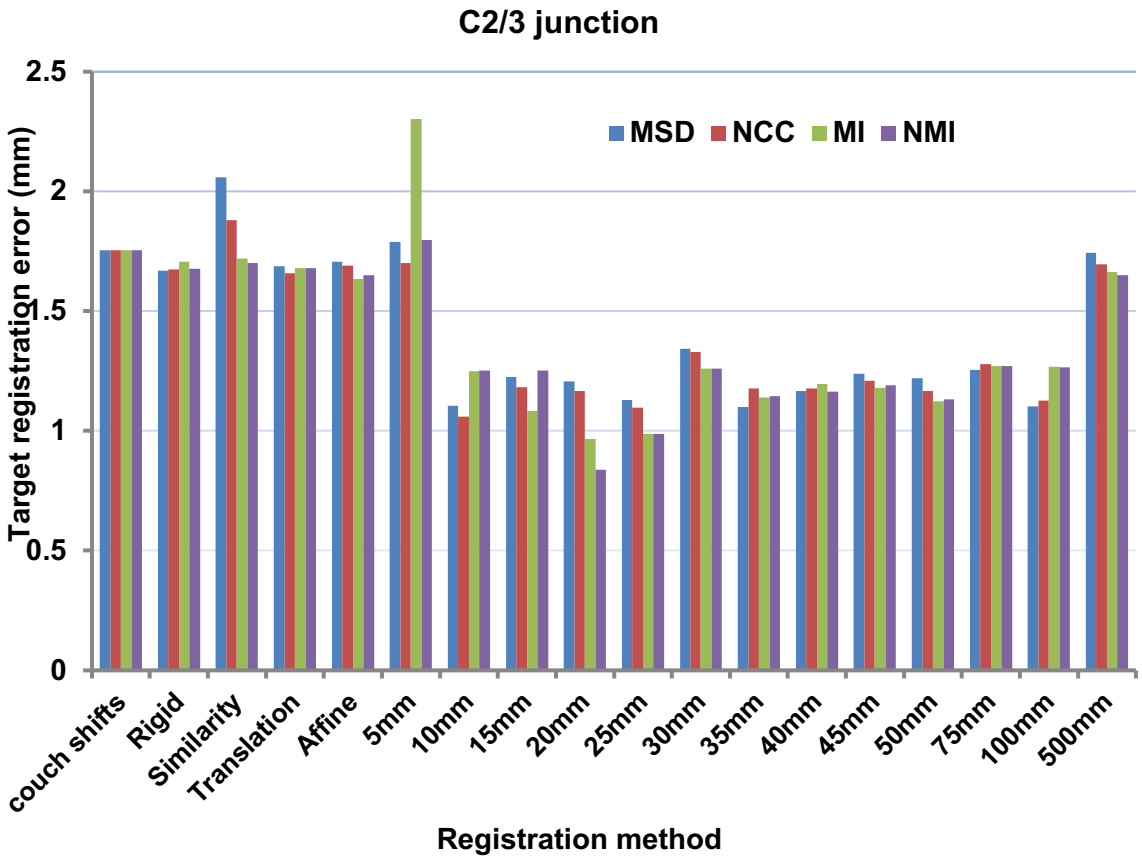


Figure 4.11: Target registration error (TRE) results for different Elastix registration strategies, similarity metrics and control point spacing's for the C2/3 junction point.

These results clearly show that overall, deformable registration approaches deliver less target registration error than non-deformable alternatives. However, these plots do not clearly indicate which similarity metric performs best. For example, the more mathematically complex similarity metrics – MI and NMI – perform better with Pterygoid plate TRE agreement, but less well with the anterior parotid point. Moreover, there is no immediately obvious trend to suggest which control point spacing offers the best agreement. For the Pterygoid plates and anterior parotid points, there is the impression of a minimum around 30-50mm, whilst for the Mastoid points and C2/3 junction the minimum appears to be around 10-25mm. In order to summarise these results, and decide on the optimum parameters for validation, a mean TRE for all registration strategies and control point spacing's was calculated. These results are shown in Table 4.6 and as a summary clustered bar chart in Figure 4.12. The algorithm with the lowest mean TRE across all image registrations and parameter options was a

deformable registration, with a 35mm B-spline control point spacing, and Mutual information (MI) as similarity metric.

Registration approach	MSD	NCC	MI	NMI	Mean
couch shifts	2.820	2.82	2.820	2.820	2.820
Rigid	2.026	2.039	2.053	2.045	2.041
Similarity	2.424	2.065	2.006	2.003	2.125
Translation	2.251	2.246	2.242	2.242	2.245
Affine	3.785	2.925	2.214	2.247	2.793
B-spline 5mm	3.149	2.304	2.810	2.822	2.771
B-spline 10mm	2.417	2.003	2.039	2.120	2.145
B-spline 15mm	2.309	2.033	1.923	1.999	2.066
B-spline 20mm	2.169	1.970	1.893	1.916	1.987
B-spline 25mm	1.924	1.820	1.835	1.867	1.861
B-spline 30mm	1.894	1.812	1.805	1.802	1.828
B-spline 35mm	1.835	1.773	**1.761**	1.781	1.788
B-spline 40mm	1.844	1.781	1.788	1.789	1.800
B-spline 45mm	1.849	1.782	1.829	1.834	1.824
B-spline 50mm	1.875	1.834	1.863	1.865	1.859
B-spline 75mm	1.955	1.964	1.976	1.974	1.967
B-spline 100mm	2.125	2.083	1.907	1.908	2.006
B-spline 500mm	3.879	3.098	2.169	2.206	2.838
Mean	1.913	1.824	1.825	1.836	

Table 4.6: Target registration error (TRE) – summary results. Non-deformable (couch shifts, rigid, similarity, translation, affine) and deformable (B-spline) with different control point spacing's (5 – 500mm) are tested. Four similarity metrics are tested; mean squared distance (MSD), normalised cross correlation (NCC), mutual information (MI) and normalised mutual information (NMI). Reported data are mean values for all 7 TRE points, across all 7 scans/registration, for all 5 patients. The lowest mean TRE (MI/B-spline 35mm) is shown in bold.

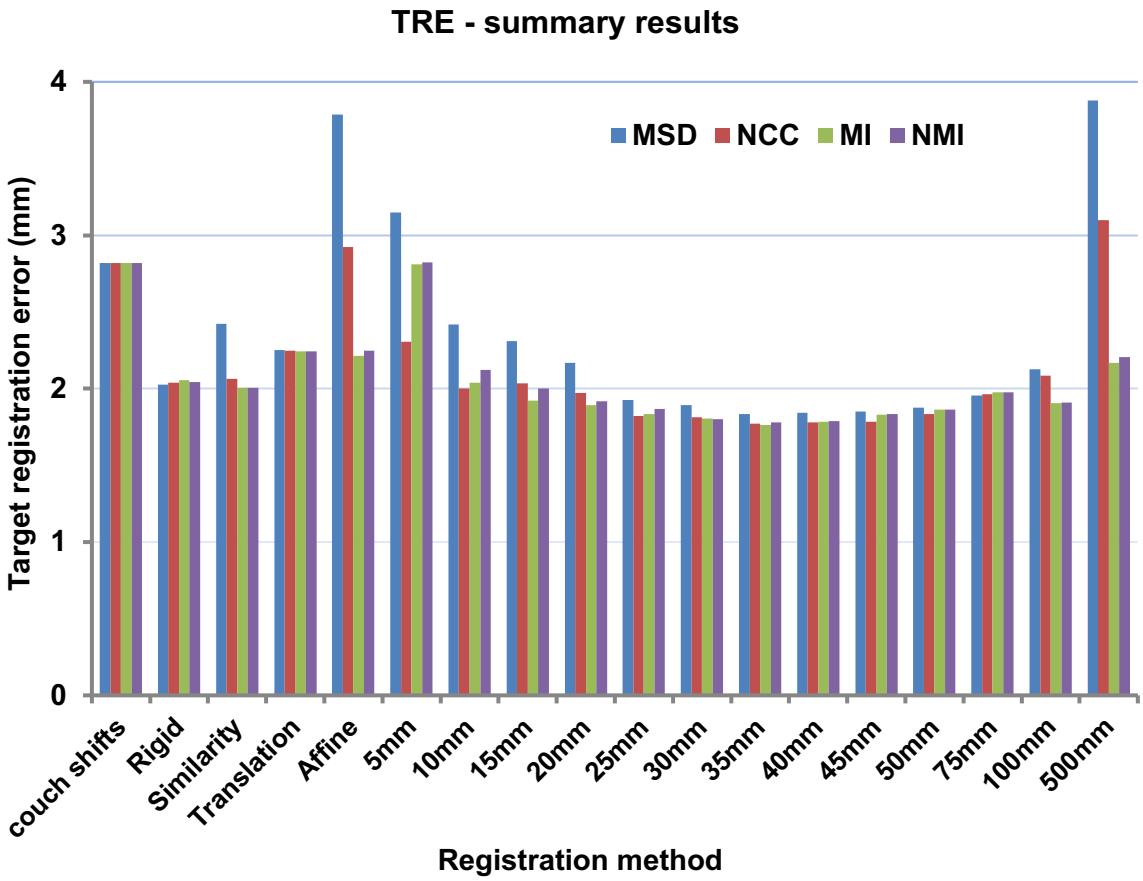


Figure 4.12: Target registration error (TRE) – summary results of data presented in Table 4.6 presented as a clustered bar chart.

4.4.2.2 Conformity metrics

As indicated in table 4.2, TG 132 recommends using DSC and MSD as metrics by which to train and assess the performance of registration algorithms [20]. However, this guideline also emphasises that in the process of commissioning, it is important to consider what the image registration application will be used for. In the context of large scale automated segmentation of OARs for dose accumulation, it was therefore important to consider other aspects of segmentation accuracy.

Specifically, it was important to try and measure the 'worst case scenario' contour with each approach – and the 95% surface distance metric was thought to be a good surrogate for this. Another relevant consideration, especially the steep dose gradients generated with IMRT plans, was the possibility of a vector translation of the contour, relative to the gold standard. This notion is also pertinent given the systematic

anatomical change often seen with HNC patients. Therefore DBC with each approach was measured. Finally, it is known that salivary glands shrink over treatment [69]. As the final objective of Elastix for this work was to generate a delivered dose volume histogram, it was therefore important to consider how well the algorithm tracked volume change over time. Therefore percentage volume differences between the deformed contour and the manual reference contours were computed. For this calculation, in contrast to the PVD metric used in the Intra- and InterOV anlaysis, my manual segmentation is taken to be the 'gold standard' contour. Therefore the reported results are simply a mean of differences between the deformed contour and the reference contour, divided by the volume of the reference contour, as a percentage.

Results for each structure are combined for each of the 5 metrics DSC, Mean SD, 95% SD, DBC, volume difference, and shown as clustered bar charts with registration strategy and similarity metric as variables in Figures 4.13 – 4.17.

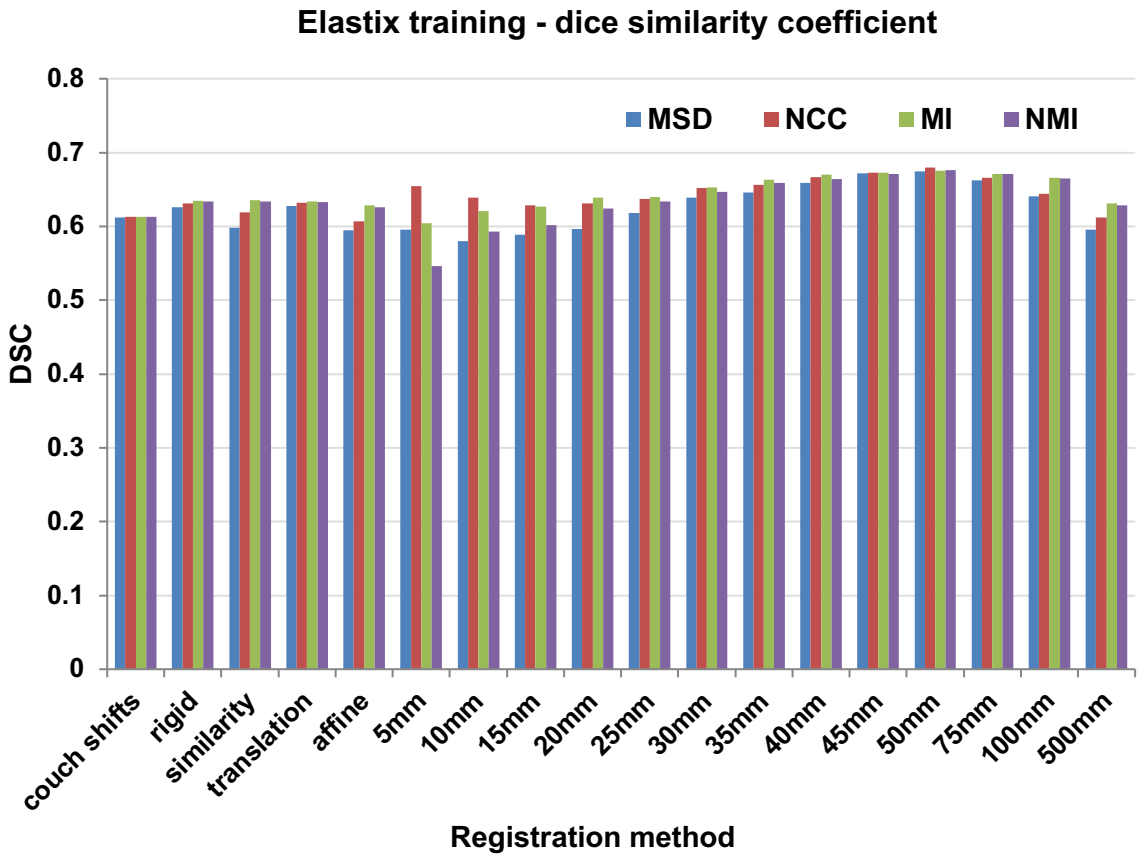


Figure 4.13: Dice similarity coefficient results - automated vs. gold standard contours for all registrations, similarity metrics, and control point spacing's.

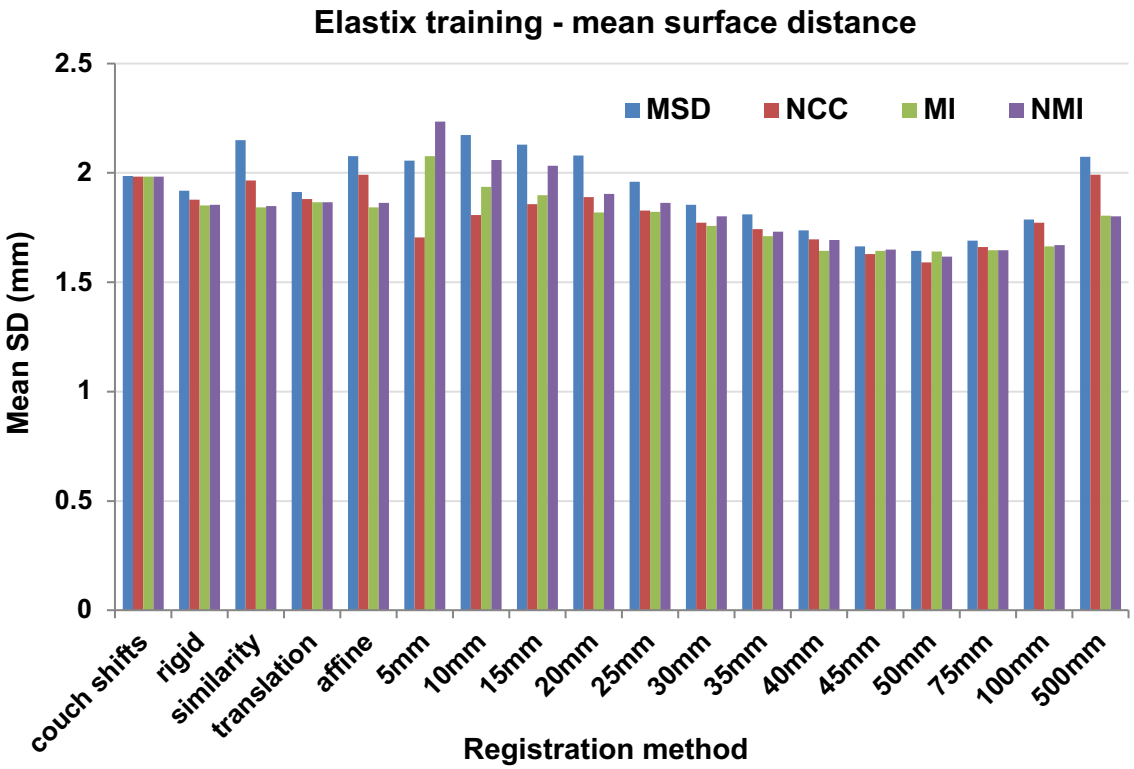


Figure 4.14: Mean surface distance results - automated vs gold standard contours for all registration approaches, similarity metrics, and control point spacing's.

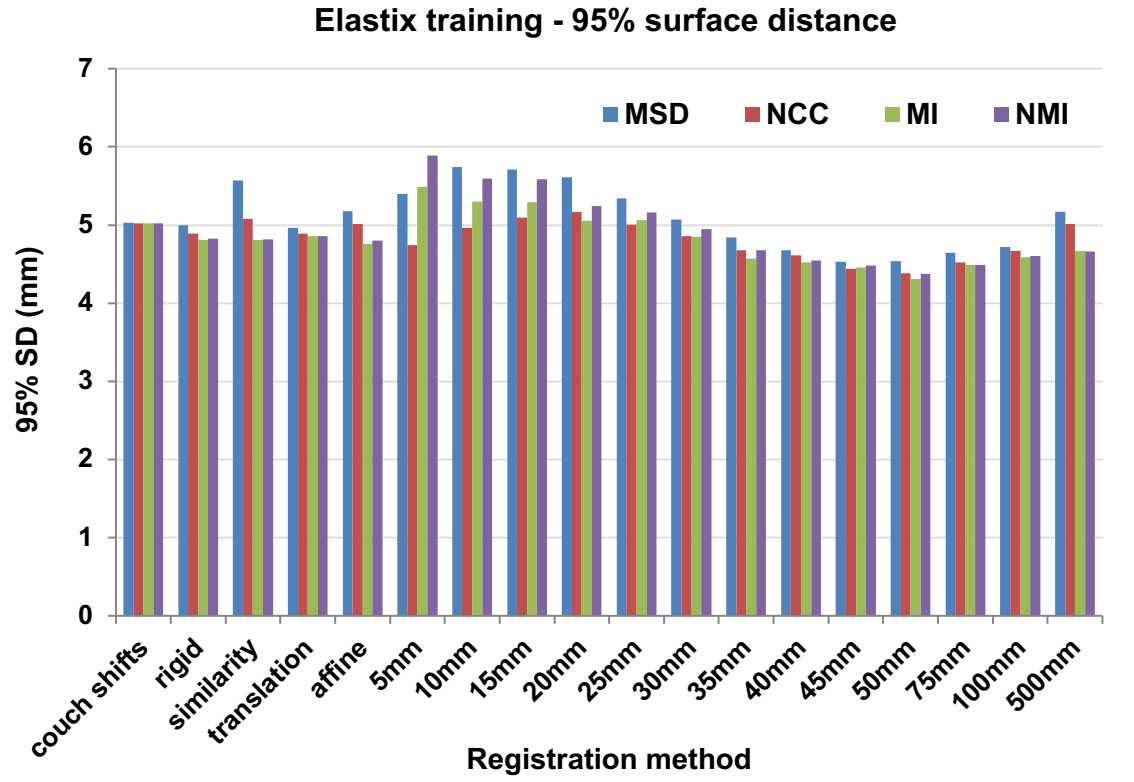


Figure 4.15: 95 centile surface distance results - automated vs. gold standard contours for all registrations, similarity metrics, and control point spacing's.

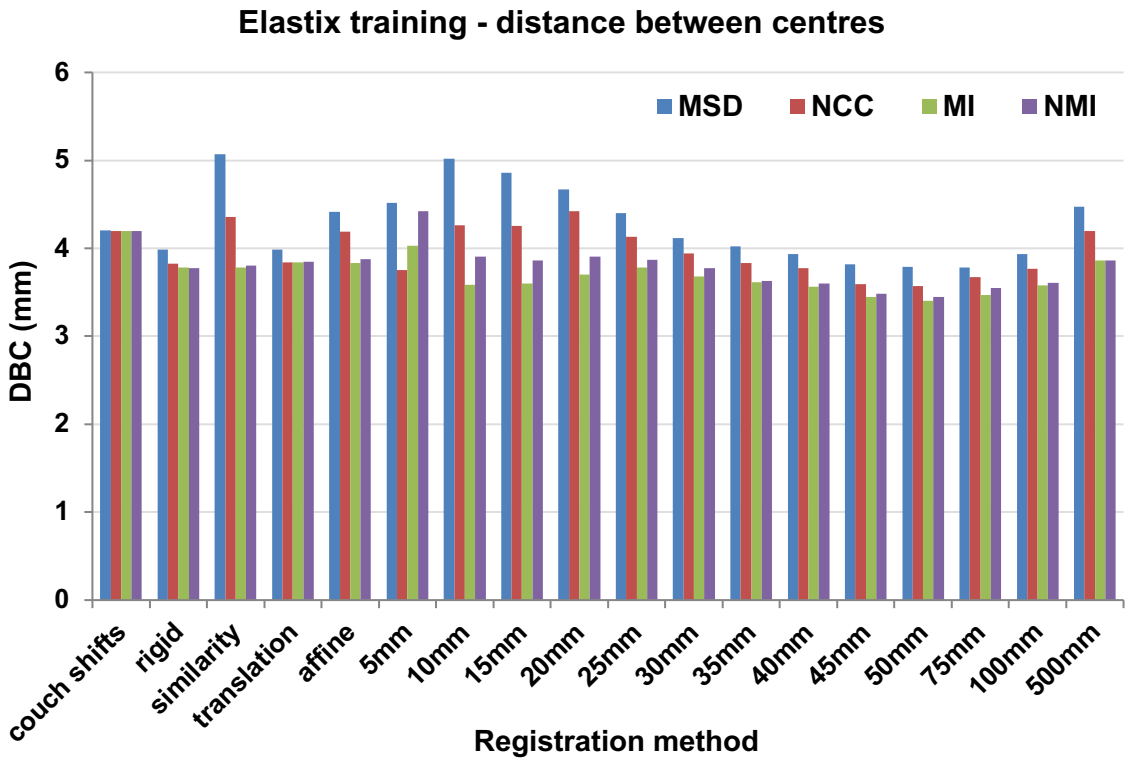


Figure 4.16: Distance between centre results - automated vs. gold standard contours for all registrations, similarity metrics, and control point spacing's.

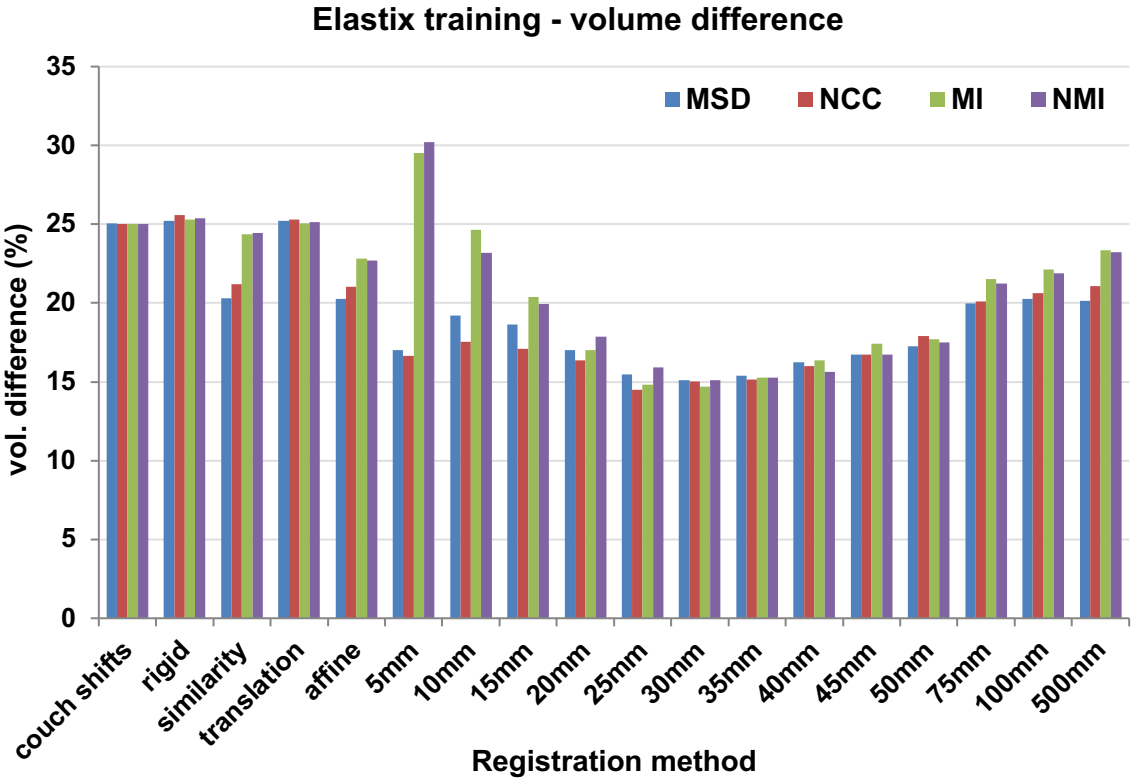


Figure 4.17: Volume difference results - automated vs. gold standard contours for all registration approaches, similarity metrics, and control point spacing's.

Some general trends are apparent from these data. Firstly, deformable strategies outperform non-deformable approaches for all metrics. Secondly, using MI and NMI as similarity metrics seems to produce marginally superior results than MSD or NCC. Next, conformity metrics (DSC & surface distance) and DBC results appear to reach an optimum around a control point spacing of 45-50mm. However, it is also clear, that the best control point spacing for accurate volume rendering is around 25-30mm.

In order to produce a single, generalisable result to make a final selection of algorithm parameters for validation, results were combined to a single score. Three scores were used for this work: DSC, mean SD and volume difference. The first reason for this was to try and adhere as closely as possible to the recommendations of TG 132. Secondly, and as discussed, it was important to consider the effect of accurate tracking of volume changes over time, especially as this results were somewhat in conflict with the conformity data. Finally, it is clear that the results for 95% SD follow a similar trend to those for mean SD. They are taken from the same distribution of underlying results, thus using both for a summary result would be to place too great an emphasis on this metric. Similarly, there is a clear overlap between the results for DBC, and DSC.

A single 'normalised score', to summarise these results was produced in the following way. For each metric (DSC, mean SD and volume difference), the results for each algorithm parameter were collated into a single table. A mean metric score, across all Elastix options was then produced. Next, each individual algorithm result was divided by the metric mean score, to give a per-metric normalised score. As DSC values closer to 1 imply better performance, whilst lower mean SD and volume difference scores are preferable, the reciprocal of the DSC division was taken, such that higher (> 1), and lower (< 1) normalised scores, imply inferior and superior performance respectively. The average of the 3 metric normalised scores was then taken, to give an overall result, as shown in Figure 4.18 and Table 4.7.

Overall, these results are broadly in agreement with the findings of the TRE analysis. Deformable registration approaches confer better results than non-deformable alternatives, and optimum results appear to be achieved with B-spline control point spacing's around 35-45mm.

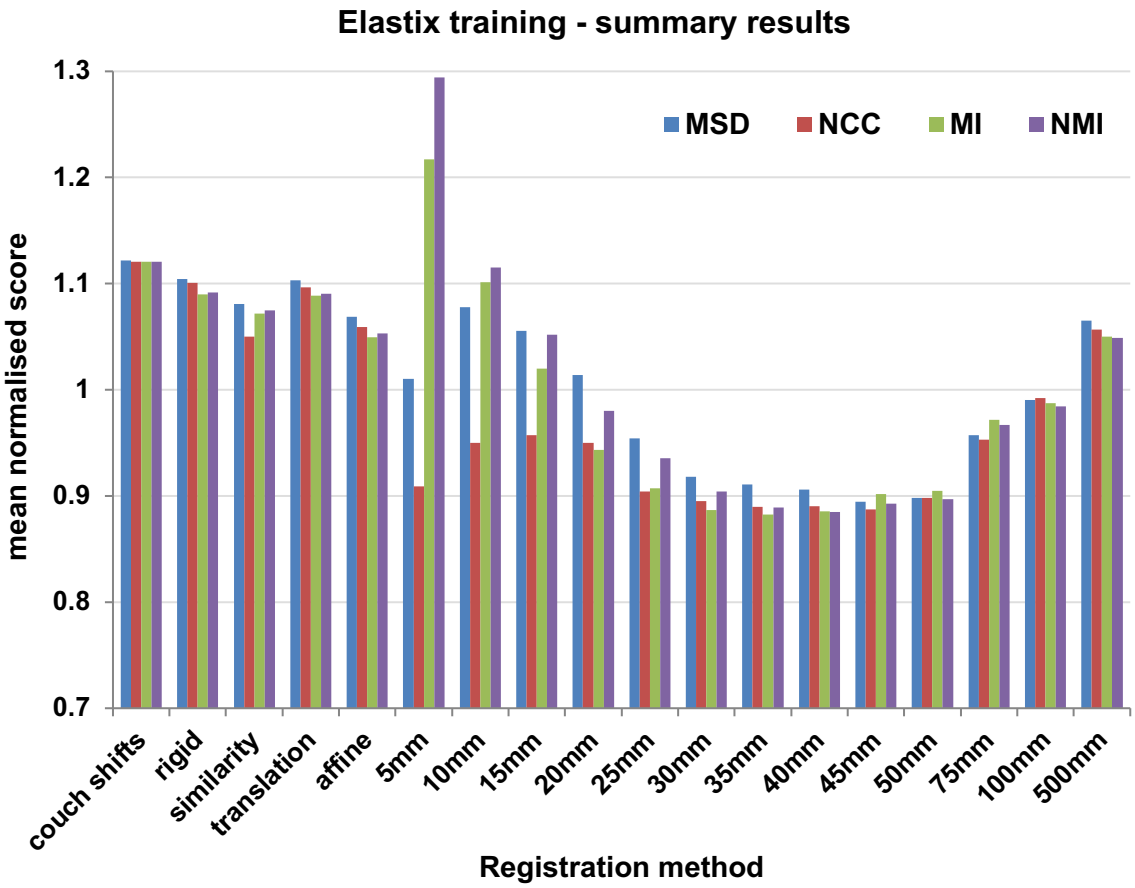


Figure 4.18: Elastix training summary results – mean of normalised scores for DSC, mean SD and volume difference, for automated vs. gold standard contours for all registration approaches, similarity metrics, and control point spacing's.

As seen with the TRE data, MI as a similarity metric performs consistently well. NCC also performs consistently well across different registration strategies, especially the smaller control-point spacing's.

Results for 5mm should be interpreted with caution, particularly those for NCC. Such small control point spacing clearly produced very poor results with both MI and NMI as similarity metrics, with a trend consistent with the data seen for 10, 15 and 20mm control point spacing's. However, the results for NCC appear to be spuriously good, an observation made more interesting by the fact that this combination of Elastix parameters produced fewer completed registrations than any of the other approaches. One possible explanation is that for NCC-based registrations in which the algorithm was struggling to find an optimum, Elastix simply didn't produce a completed result.

Registration approach	MSD	NCC	MI	NMI	Mean
couch shifts	1.121	1.121	1.121	1.121	1.121
Rigid	1.104	1.101	1.090	1.092	1.097
Similarity	1.081	1.050	1.072	1.075	1.070
Translation	1.103	1.096	1.089	1.090	1.095
Affine	1.068	1.059	1.050	1.053	1.058
B-spline 5mm	1.010	0.909	1.217	1.294	1.108
B-spline 10mm	1.078	0.950	1.101	1.115	1.061
B-spline 15mm	1.055	0.957	1.020	1.052	1.021
B-spline 20mm	1.014	0.950	0.943	0.980	0.972
B-spline 25mm	0.954	0.904	0.907	0.935	0.925
B-spline 30mm	0.918	0.895	0.887	0.904	0.901
B-spline 35mm	0.911	0.890	**0.883**	0.889	0.893
B-spline 40mm	0.906	0.890	0.886	0.885	0.892
B-spline 45mm	0.895	0.887	0.902	0.892	0.894
B-spline 50mm	0.898	0.898	0.905	0.897	0.899
B-spline 75mm	0.957	0.953	0.972	0.967	0.962
B-spline 100mm	0.990	0.992	0.987	0.985	0.989
B-spline 500mm	1.065	1.057	1.050	1.049	1.055
Mean	1.007	0.976	1.004	1.015	

Table 4.7: Elastix training summary results – mean of normalised scores for DSC, mean SD and volume difference. Non-deformable (couch shifts, rigid, similarity, translation, affine) and deformable (B-spline) with different control point spacing's (5 – 500mm) are tested. Four similarity metrics are tested; mean squared distance (MSD), normalised cross correlation (NCC), mutual information (MI) and normalised mutual information (NMI). The lowest mean normalised score (MI/B-spline 35mm) is shown in bold.

In contrast, for both MI and NMI, more data were available for analysis, suggesting that this combination of parameters allowed more poor quality registrations to complete, thereby skewing the results.

These observations notwithstanding, the general concordance between the data from both the TRE and conformity metric analysis is reassuring. On the basis of these

results, the decision was taken to proceed to validation using MI based deformable registration, with a control point spacing of 35mm.

4.4.3 Elastix validation

It was important to validate the performance of the deformable registration solution within Elastix, using these parameters. Results of the TRE analysis for a single randomly selected MVCT of 5 patients are shown below in Figure 4.19. Results for both deformable registration and translation from daily radiographer couch shifts are shown by way of comparison.

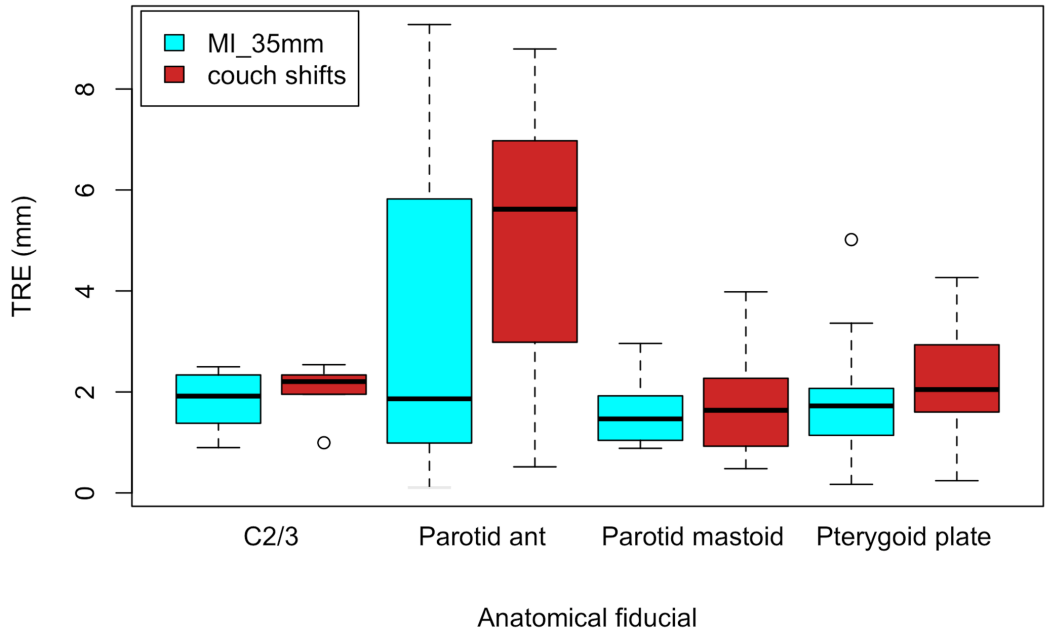


Figure 4.19: Elastix validation - TRE data for all anatomical points – sample of 5 patients with a single, randomly selected MVCT. Results of the deformable registration (Mutual information, B-spline 35mm) are shown in cyan, and translation from daily radiographer couch shifts in red for comparison.

Mean TRE values, in millimetres, for deformable registration and radiographer couch shifts respectively were; 1.81(±0.30) & 2.01(±0.27) for C2/3, 3.31(±0.94) & 4.96(±0.83) for anterior parotid, 1.61(±0.20) & 1.86(±0.36) for parotid mastoid, and 1.94(±0.43) & 2.26(±0.38) for pterygoid plates. The mean TRE value, for all anatomical points, in the

validation cohort was 2.22mm (±0.20) for deformable registration, entirely in line with the 2-3mm target specified in TG 132 [20].

Conformity analysis was also repeated for validation. Previously unseen segmentations of all 30 MVCTs for one patient, and a randomly selected MVCT for a further 10 patients were used. Results for the 1x30 and 10x1 image samples are shown separately as boxplots, in figures 4.20 and 4.21 below, and combined in summary in Table 4.8.

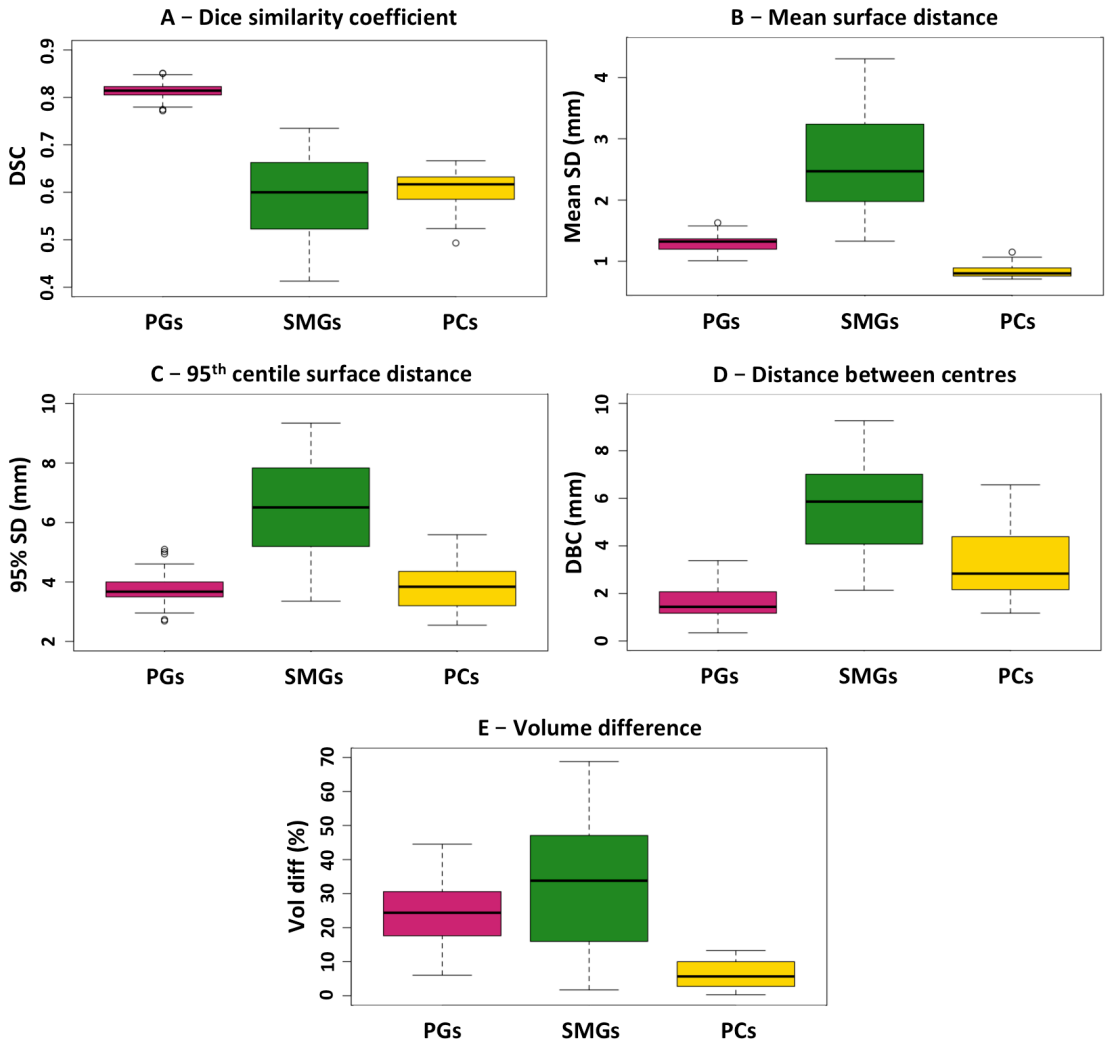


Figure 4.20 (A-E): Elastix validation – contour conformity data – sample of 1 patient, all 30 MVCTs. A – DSC, B – Mean SD, C – 95% SD, D – DBC, E – volume difference.

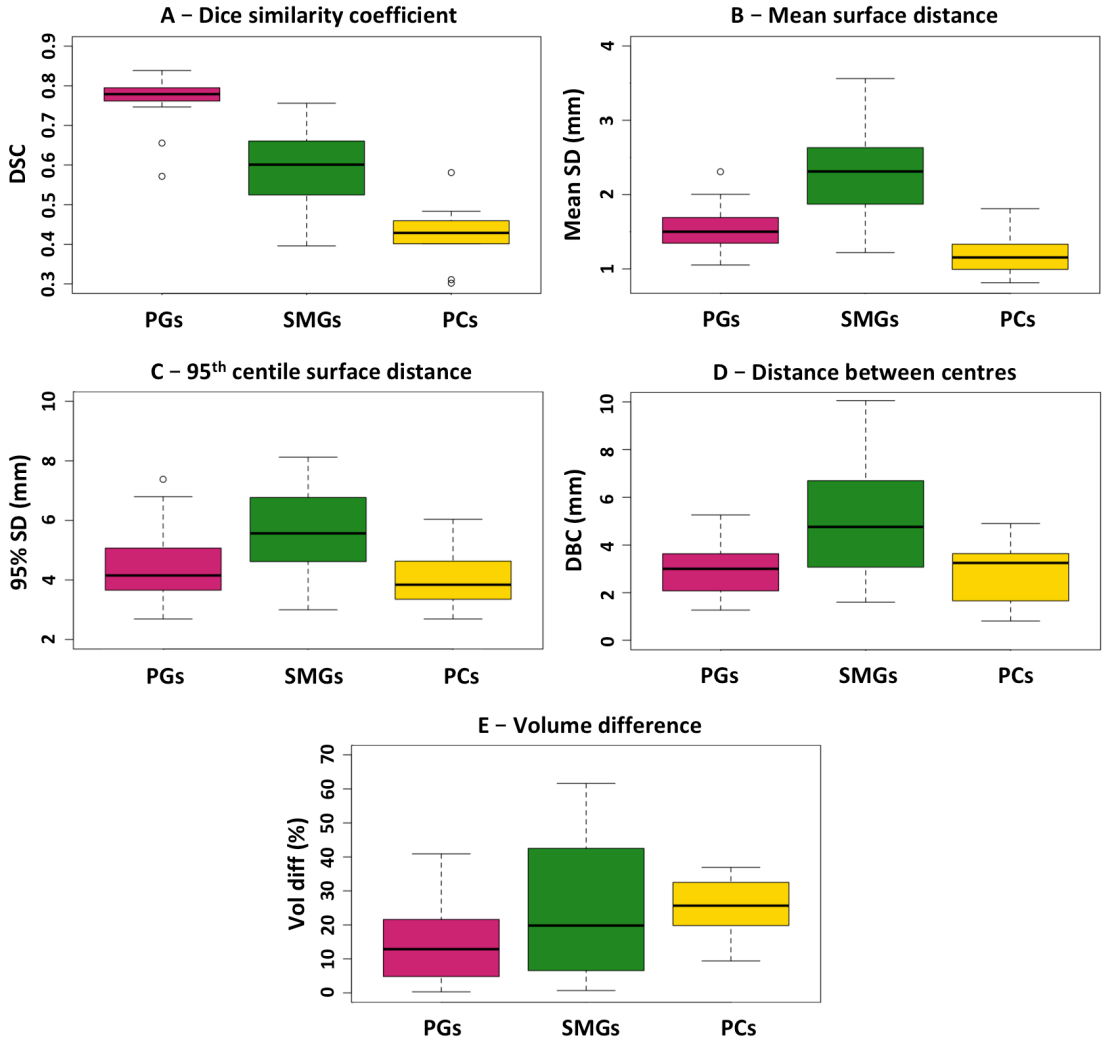


Figure 4.21 (A-E): Elastix validation – contour conformity data – sample of 10 patients, final fraction MVCT. A – DSC, B – Mean SD, C – 95% SD, D – DBC, E – volume difference.

Jacobian determinants for all registrations between the kVCT and all 30 MVCTs, for 2 patients, were also calculated. Results are shown in Figure 4.22, where the mean value of Jacobian determinants of all voxels in the structure are plotted as a function of treatment fraction. Values <1 infer shrinking voxels in I_M, and >1 voxels growing in size. Absolute values in this context are not that meaningful, but trends are broadly as expected; Jacobian determinants for all parotid glands, and 2 of 3 submandibular glands trended down over treatment, whilst both pharyngeal constrictor muscles showed slight upward trends.

	1 patient, 30 MVCTs			10 patients, 1 MVCT			Combined		
Structure	PGs	SMGs	PCs	PGs	SMGs	PCs	PGs	SMGs	PCs
Mean DSC	0.81 (±0.002)	0.59 (±0.01)	0.61 (±0.007)	0.77 (±0.01)	0.60 (±0.02)	0.43 (±0.02)	0.80 (±0.004)	0.59 (±0.01)	0.56 (±0.02)
Mean SD (mm)	1.3 (±0.02)	2.6 (±0.1)	0.84 (±0.02)	1.5 (±0.07)	2.4 (±0.2)	1.2 (±0.06)	1.4 (±0.02)	2.5 (±0.09)	0.92 (±0.04)
95% SD (mm)	3.7 (±0.07)	6.5 (±0.2)	2.4 (±0.07)	4.5 (±0.3)	5.9 (±0.5)	4.1 (±0.3)	3.9 (±0.09)	6.3 (±0.2)	2.8 (±0.2)
Mean DBC (mm)	1.6 (±0.1)	5.6 (±0.2)	3.3 (±0.3)	2.9 (±0.2)	5.0 (±0.5)	3.0 (±0.3)	1.9 (±0.1)	5.4 (±0.2)	3.2 (±0.2)
Mean DTC (mm)	0.27 (±0.01)	0.46 (±0.04)	-0.08 (±0.02)	0.09 (±0.05)	0.29 (±0.1)	-0.39 (±0.05)	0.22 (±0.02)	0.42 (±0.04)	-0.16 (±0.03)
Mean HD (mm)	11.9 (±0.1)	10.6 (±0.3)	8.4 (±0.2)	12.5 (±1.0)	9.8 (±0.7)	10.3 (±0.4)	12.0 (±0.3)	10.4 (±0.3)	9.0 (±0.3)
Mean Volume difference (%)	24.1 (±1.1)	32.2 (±2.5)	6.2 (±0.7)	13.9 (±2.4)	29.1 (±7.1)	24.9 (±2.2)	21.5 (±1.1)	31.4 (±2.5)	10.9 (±1.6)

Table 4.8: Elastix validation – contour conformity data for 1 patient x 30 scan sample, 10 patients x 1 scan sample, and results combined. Data presented per structure; parotid glands (PGs), submandibular glands (SMGs) and pharyngeal constrictors (PCs). Metrics are: Dice similarity coefficient (DSC), mean surface distance (mean SD (mm)), 95th centile surface distance (95% SD)), distance between centres (DBC), distance to conformity (DTC), Hausdorff distance (HD) and volume difference (%). Standard errors reported in parentheses.

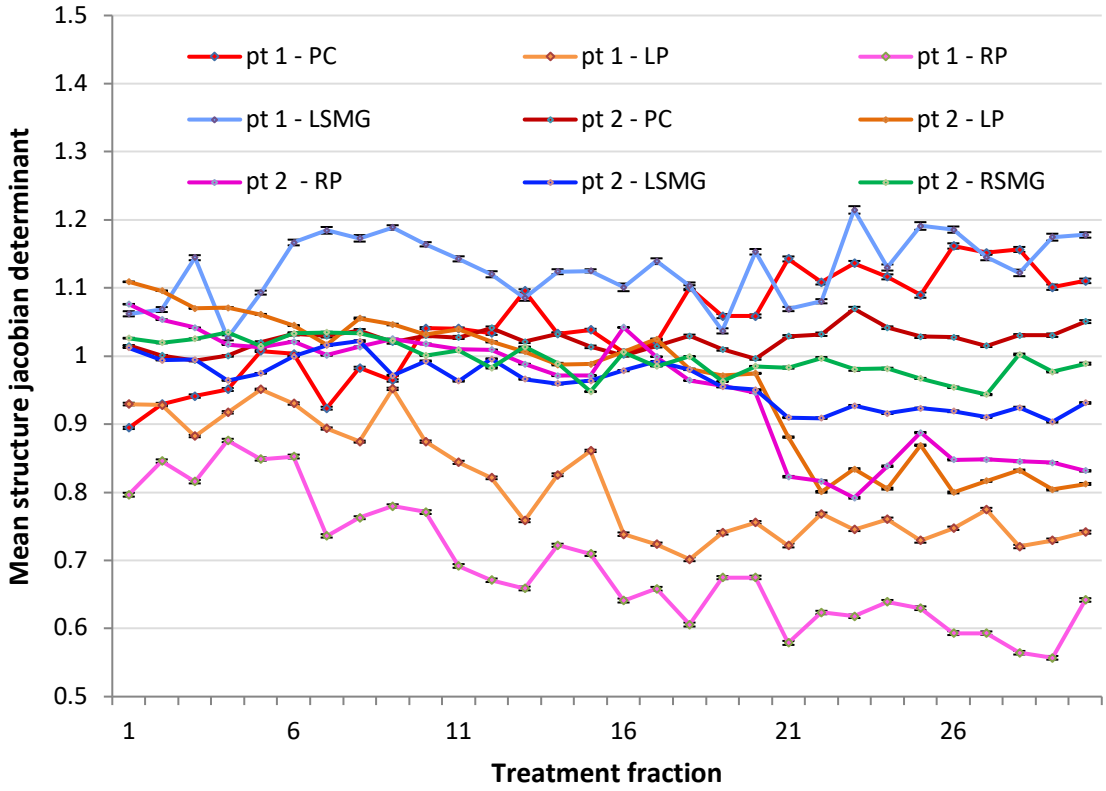
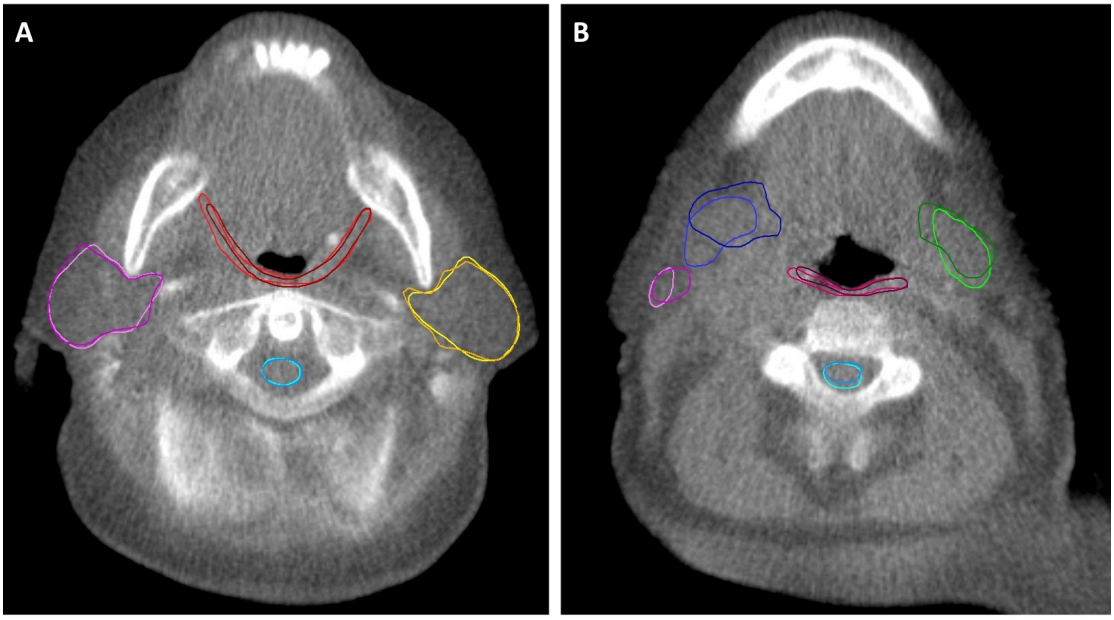


Figure 4.22: Elastix validation – Jacobian determinant data. Mean values of Jacobian determinants (with standard errors), per structure, for registrations between kVCT and every daily MVCT, for 2 patients. Line colours denote structures: red – pharyngeal constrictor muscles, purple – right parotid glands, orange – left parotid glands, blue – left submandibular gland, green – right submandibular gland (only 1 – structure absent in patient 1). Lighter shades denote data from patient 1, darker shades data from patient 2.

Overall therefore, the validation data were reassuring, with results similar to those observed with training data. Examples of both 'good' and 'bad' auto-contours generated from Elastix registrations with finalised parameters are shown in Figure 4.23.



4.23 (A&B): Examples of 'good' and 'bad' Elastix derived automated contours (darker shades) on MVCT images, with manual gold standard (lighter shades). Caption A – right parotid (pink v purple) DSC = 0.94, left parotid (yellow v orange) DSC 0.94, spinal cord (cyan v cerulean) DSC = 0.95, superior pharyngeal constrictor (red v claret) DSC = 0.75. Caption B – right parotid (pink v purple) DSC = 0.63, right submandibular gland (blue v navy) DSC = 0.59, left submandibular gland (lime v forest green) DSC = 0.76, middle pharyngeal constrictor (fuscia v burgundy) DSC = 0.75, spinal cord (cyan v cerulean) DSC = 0.80.

4.5 Discussion

4.5.1 Automated segmentation – swallowing OARs.

A number of trends are apparent from Elastix validation data for swallowing OARs. First, it is clear from both the 1 x 30 and 10 x 1 samples that MVCT parotid gland contours generated by Elastix are better than either submandibular gland, or pharyngeal constrictor contours. The combined data for the parotids show results very similar to those seen in the inter-observer variability analysis For example, mean DSC scores were 0.82 and 0.80 for inter-observer variability and Elastix respectively.

Results for the pharyngeal constrictor muscles also compare favourably. The mean DSC score is again similar – 0.54 for inter-observer and 0.56 for Elastix, and whilst the DBC is worse with image registration than a second human observer (3.2mm vs

1.4mm), this metric is of limited use for the pharyngeal constrictors given its horseshoe shape, and this deficiency is more than offset by better performance with volume difference (10.9% vs 22%), and mean Hausdorff distance (9mm vs 13.5mm). It should be noted that the method for calculating volume difference is subtly different for inter-observer and Elastix validation analyses; in the latter the volume of my first contour is taken as 'gold-standard', and the reference (denominator) by which to generate a percentage result, in the former the denominator is the mean volume of the 2 observer's contours. However, this means that reported mean volume differences are likely to be lower for inter-observer variability, meaning the result for the pharyngeal constrictors is potentially even more favourable for Elastix.

Results for the submandibular glands compare least well with the inter-observer variability results. Mean DSC is clearly lower (0.59 compared with 0.66), and whilst mean HD is better (10.4mm vs. 13.3mm), DBC is also worse (5.4mm vs. 2.8mm). These results are heavily influenced by the 1 x 30 MVCT sample, from a single patient, and it is clear that the data for the 1 x 10 MVCT sample is better for the SMG's than the 1 patient sample, in contrast to the other 2 structures.

Nonetheless, close review of the training data do confirm that deformable registration within Elastix with these parameters (B-spline, 35mm, MI) consistently performs less well with the SMGs than for other structures, both absolutely, and compared to alternative approaches, and it must be concluded that automated SMG contours generated by Elastix with these settings are inherently less robust than either parotid gland or pharyngeal constrictor contours. This conclusion is supported by the Jacobian determinant validation data. It is well documented that in most patients the SMGs shrink during treatment [69, 79, 80], with a mean value of 22% across studies suggested [81]. However, in this analysis, computed Jacobian determinant values appear to show that 1 of 3 structures assessed increased in volume during treatment, whilst the remaining 2 shrank slightly. In contrast, Jacobian determinant results for the parotids and pharyngeal constrictors conform better to literature-derived a-priori expectations. A mean value for parotid gland shrinkage of 26% has been reported [81], and all four glands saw clear trends to decreasing Jacobian determinant values during treatment. Interestingly, previous authors have reported an increase in the volume of

pharyngeal constrictor muscles during treatment [82], an effect inferred by the Jacobian data in this work.

Elastix has previously been validated in the head and neck, in the context of disease recurrence [43]. In this work, DSC scores were 0.58 and 0.62 for the SMGs, 0.72 and 0.74 for the parotids, and 0.79 for the spinal cord were reported. However, registrations in this study were between 2 kVCT images – the planning scan, and that done at recurrence. There is also data reporting the accuracy of DIR between IG CT and kVCT. Li and colleagues recently assessed ten different algorithms in this context, and specifically compared the performance of optical-flow, demons, level-set and spline based approaches in a sample of 21 HNC patients with matched planning and CBCT scans [83]. Conformity data for PGs and SMGs are not overtly stated in the paper, but reading from boxplots, typical DSC scores seem to be around 0.75-0.8 for the parotids and 0.6-0.7 for the SMGs.

Further data is available from Hvid et al, who used an intensity based free-form deformation as part of the commercially available MIM software package to look at a range of structures including the parotids, SMGs and pharyngeal constrictor muscles [84]. Their results were excellent, with DSC scores of 0.95 for the parotids, 0.85 for the SMGs, and 0.7 – 0.8 for pharyngeal constrictor muscles. The result may in part be explained by the fact that the reference contours in this work were amendments made to the transposed contours, rather than de-novo segmentations. Further comparators for deformation of parotid contours from kVCT to CBCT include DSC scores of 0.78/0.79 for Peroni et al. [85], and 0.81 for Eiland and colleagues [86].

Although there are no studies that report TRE data using the image anatomical fiducial markers described here, there is published data that provides a useful comparator. Mencarelli and colleagues implanted 4-10 small gold seeds around the primary tumour in a sample of 13 patients with oropharyngeal cancer, and tested the accuracy of registrations from planning CTs to CBCT [41]. A B-spline based deformable registration was also used for this work. Interestingly, their data show that the mean fiducial-to-fiducial 3-D vector post registration was 2.2mm – identical to the result in the validation cohort.

There is also previous work looking at this technique in the context of kVCT planning scans and TomoTherapy MVCT images [63]. Faggiano and colleagues used a B-spline free-form deformation based algorithm to auto-propagate contours from kVCT to daily MVCT images, and found a mean DSC score for the parotids of 0.77, with a mean HD value of 9.5mm – results similar to those presented here.

Overall, these results concord well with the data presented in this work. Perhaps even more importantly, most of these results satisfy suggested tolerance values stipulated in TG 132. DSC scores for the parotids align with the numerical target of 0.8-0.9, and aside from the previously outlined problems with the SMGs, all 3 structures have Dice scores comparable to the 'contouring uncertainty of the structure'. Furthermore, mean MDA values for all structures, are all within the target of 2-3mm suggested by TG 132. For example, for the pharyngeal constrictors, this result means that the mean distance from the Elastix contour to my reference is less than 1mm.

Overall therefore, these results were sufficiently robust to proceed to the next step in the process – calculation of delivered dose.

4.5.2 Automated segmentation – spinal cord.

For the reasons outlined in section 4.1, this Chapter does not contain any data pertaining to the training of Elastix for automated contouring of the spinal cord, and key data from this work have been published [1]. However, Chapter 5 will describe the process of dose accumulation to the spinal cord using automated contours from Elastix, and analyse the delivered dose data derived from this process. Therefore, the key results from this work bear brief discussion, to set the next Chapter in context.

The data suggest that conformity indices for both training and validation datasets are very similar to those observed in the inter-observer variability contouring work. Mean JCI results for Elastix training, testing and validation respectively were 0.79, 0.73 and 0.73, compared with a range of mean JCI values of 0.64 – 0.75 for 6 human observers [1]. There is limited literature on this specific subject. Two relatively recent studies have used a similar approach to segment the spinal cord [84], and spinal canal [87] on CBCT images. Results in these studies report DSC values between 0.85 and 0.9, whereas

conformity data in our paper were presented as JCI values [1], making direct comparison difficult. However, it is true to say that if computed from the same pair of matched structures, JCI values are always lower than DSC, and the difference converges as they trend towards 1. Thus, whilst it is not possible to compare directly, it is reasonable to assert that DSC values for Elastix would be >0.8, and broadly comparable to the results published by Hvid and Landry [84, 87].

4.6 Conclusions

This chapter describes work in which an open source deformable image registration software toolbox - Elastix - was trained and validated for the purposes of contour propagation from planning kVCT to image-guidance MVCT, to permit subsequent computation of delivered dose. The data presented show that with optimised parameter settings, the algorithm performed comparably to published results in the literature, in line with the uncertainty of expert observers, and within the thresholds defined by a consensus guideline. Therefore, the methodology developed in Chapter 4 was used for all subsequent delivered dose calculations.

4.7 References:

[1] Yeap PL, Noble DJ, Harrison K, Bates AM, Burnet NG, Jena R, et al. Automatic contour propagation using deformable image registration to determine delivered dose to spinal cord in head-and-neck cancer radiotherapy. Phys Med Biol. 2017;62:6062-73.
[2] Burnet NG SJ, Romanchikova M, Thomas SJ, Bates AM, Wong E, Noble DJ, Shelley LEA, Bond SJ, Forman J, Hoole ACF, Barnett GC, Brochu FM, Simmons MPD, Jena R, Harrison K, Yeap PL, Drew A, Silvester E, Elwood P, Pullen H, Sultana A, Seah SYK, Wilson MZ, Russel SG, Benson RJ, Rimmer YL, Jefferies SJ, Taku N, Gurnell M, Powlson A, Schönlieb C-B, Cai X, Sutcliffe MPF, Parker MA. Applying physical science techniques and CERN technology to an unsolved problem in radiation treatment for cancer: the multidisciplinary 'VoxTox' research programme. CERN IdeaSquare Journal of Experimental Innovation. 2017;1:3-12.
[3] Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tome WA. Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. Int J Radiat Oncol Biol Phys. 2010;76:S135-9.
[4] Thomson DJ, Beasley WJ, Garcez K, Lee LW, Sykes AJ, Rowbottom CG, et al. Relative plan robustness of step-and-shoot vs rotational intensity-modulated radiotherapy on repeat computed tomographic simulation for weight loss in head and neck cancer. Med Dosim. 2016;41:154-8.
[5] van Kranen S, Hamming-Vrieze O, Wolf A, Damen E, van Herk M, Sonke JJ. Head and Neck Margin Reduction With Adaptive Radiation Therapy: Robustness of Treatment Plans Against Anatomy Changes. Int J Radiat Oncol Biol Phys. 2016;96:653-60.
[6] Castelli J, Simon A, Rigaud B, Chajon E, Thariat J, Benezery K, et al. Adaptive radiotherapy in head and neck cancer is required to avoid tumor underdose. Acta Oncol. 2018;57:1267-70.
[7] Loo H, Fairfoul J, Chakrabarti A, Dean JC, Benson RJ, Jefferies SJ, et al. Tumour shrinkage and contour change during radiotherapy increase the dose to organs at risk but not the target volumes for head and neck cancer patients treated on the TomoTherapy HiArt system. Clin Oncol (R Coll Radiol). 2011;23:40-7.
[8] Schultheiss TE, Tome WA, Orton CG. Point/counterpoint: it is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy. Med Phys. 2012;39:6531-3.
[9] Zhong H, Chetty IJ. Caution Must Be Exercised When Performing Deformable Dose Accumulation for Tumors Undergoing Mass Changes During Fractionated Radiation Therapy. Int J Radiat Oncol Biol Phys. 2017;97:182-3.
[10] Hugo GD, Dial C, Siebers JV. In Regard to Zhong and Chetty. Int J Radiat Oncol Biol Phys. 2017;99:1308-10.
[11] Kessler ML. Image registration and data fusion in radiation therapy. Br J Radiol. 2006;79 Spec No 1:S99-108.
[12] Hill DL, Batchelor PG, Holden M, Hawkes DJ. Medical image registration. Phys Med Biol. 2001;46:R1-45.
[13] Vasquez Osorio EM, Hoogeman MS, Teguh DN, Al-Mamgani A, Kolkman-Deurloo IK, Bondar L, et al. Three-dimensional dose addition of external beam radiotherapy and brachytherapy for oropharyngeal patients using nonrigid registration. Int J Radiat Oncol Biol Phys. 2011;80:1268-77.
[14] Cattaneo GM, Reni M, Rizzo G, Castellone P, Ceresoli GL, Cozzarini C, et al. Target delineation in post-operative radiotherapy of brain gliomas: interobserver variability and impact of image registration of MR(pre-operative) images on treatment planning CT scans. Radiother Oncol. 2005;75:217-23.
[15] Hanna GG, Van Sornsen De Koste JR, Carson KJ, O'Sullivan JM, Hounsell AR, Senan S. Conventional 3D staging PET/CT in CT simulation for lung cancer: impact of rigid and deformable target volume alignments for radiotherapy treatment planning. Br J Radiol. 2011;84:919-29.
[16] Hamdan I, Bert J, Rest CCL, Tasu JP, Boussion N, Valeri A, et al. Fully automatic deformable registration of pretreatment MRI/CT for image-guided prostate radiotherapy planning. Med Phys. 2017;44:6447-55.
[17] Chuter R, Prestwich R, Bird D, Scarsbrook A, Sykes J, Wilson D, et al. The use of deformable image registration to integrate diagnostic MRI into the radiotherapy planning pathway for head and neck cancer. Radiother Oncol. 2017;122:229-35.
[18] Gregoire V, Evans M, Le QT, Bourhis J, Budach V, Chen A, 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;126:3-24.

[19] Lee AW, Ng WT, Pan JJ, Poh SS, Ahn YC, AlHussain H, et al. International guideline for the delineation of the clinical target volumes (CTV) for nasopharyngeal carcinoma. Radiother Oncol. 2018;126:25-36.
[20] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. Med Phys. 2017;44:e43-e76.
[21] Roques TW. Patient selection and radiotherapy volume definition - can we improve the weakest links in the treatment chain? Clin Oncol (R Coll Radiol). 2014;26:353-5.
[22] Tappeiner E, Proll S, Honig M, Raudaschl PF, Zaffino P, Spadea MF, et al. Multi-organ segmentation of the head and neck area: an efficient hierarchical neural networks approach. Int J Comput Assist Radiol Surg. 2019.
[23] Zhu W, Huang Y, Zeng L, Chen X, Liu Y, Qian Z, et al. AnatomyNet: Deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy. Med Phys. 2019;46:576-89.
[24] van Rooij W, Dahele M, Ribeiro Brandao H, Delaney AR, Slotman BJ, Verbakel WF. Deep learning-based delineation of head and neck organs-at-risk: geometric and dosimetric evaluation. Int J Radiat Oncol Biol Phys. 2019.
[25] Castadot P, Lee JA, Parraga A, Geets X, Macq B, Gregoire V. Comparison of 12 deformable registration strategies in adaptive radiation therapy for the treatment of head and neck tumors. Radiother Oncol. 2008;89:1-12.
[26] Qazi AA, Pekar V, Kim J, Xie J, Breen SL, Jaffray DA. Auto-segmentation of normal and target structures in head and neck CT images: a feature-driven model-based approach. Med Phys. 2011;38:6160-70.
[27] Chen A, Deeley MA, Niermann KJ, Moretti L, Dawant BM. Combining registration and active shape models for the automatic segmentation of the lymph node regions in head and neck CT images. Med Phys. 2010;37:6338-46.
[28] Lim JY, Leech M. Use of auto-segmentation in the delineation of target volumes and organs at risk in head and neck. Acta Oncol. 2016;55:799-806.
[29] Raudaschl PF, Zaffino P, Sharp GC, Spadea MF, Chen A, Dawant BM, et al. Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015. Med Phys. 2017;44:2020-36.
[30] Tsuji SY, Hwang A, Weinberg V, Yom SS, Quivey JM, Xia P. Dosimetric evaluation of automatic segmentation for adaptive IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2010;77:707-14.
[31] Teguh DN, Levendag PC, Voet PW, Al-Mamgani A, Han X, Wolf TK, et al. Clinical validation of atlas-based auto-segmentation of multiple target volumes and normal tissue (swallowing/mastication) structures in the head and neck. Int J Radiat Oncol Biol Phys. 2011;81:950-7.
[32] Daisne JF, Blumhofer A. Atlas-based automatic segmentation of head and neck organs at risk and nodal target volumes: a clinical validation. Radiat Oncol. 2013;8:154.
[33] Fritscher KD, Peroni M, Zaffino P, Spadea MF, Schubert R, Sharp G. Automatic segmentation of head and neck CT images for radiotherapy treatment planning using multiple atlases, statistical appearance models, and geodesic active contours. Med Phys. 2014;41:051910.
[34] Walker GV, Awan M, Tao R, Koay EJ, Boehling NS, Grant JD, et al. Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer. Radiother Oncol. 2014;112:321-5.
[35] Thomson D, Boylan C, Liptrot T, Aitkenhead A, Lee L, Yap B, et al. Evaluation of an automatic segmentation algorithm for definition of head and neck organs at risk. Radiat Oncol. 2014;9:173.
[36] Loi G, Fusella M, Lanzi E, Cagni E, Garibaldi C, Iacoviello G, et al. Performance of commercially available deformable image registration platforms for contour propagation using patient-based computational phantoms: A multi-institutional study. Med Phys. 2018;45:748-57.
[37] Klein S, Staring M. Elastix: the manual, September 2015. [Online; accessed Dec 2015].
[38] Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging. 2010;29:196-205.
[39] Loeckx D MF, Vandermeulen D, Suetens P. Nonrigid Image Registration Using Free-Form Deformations with a Local Rigidity Constraint. Medical Image Computing and Computer-assisted Intervention: MICCAI. 2004;1:639 - 46.
[40] Staring M, Klein S, Pluim JP. A rigidity penalty term for nonrigid registration. Med Phys. 2007;34:4098-108.
[41] Mencarelli A, van Kranen SR, Hamming-Vrieze O, van Beek S, Nico Rasch CR, van Herk M, et al. Deformable image registration for adaptive radiation therapy of head and neck cancer: accuracy and precision in the presence of tumor changes. Int J Radiat Oncol Biol Phys. 2014;90:680-7.

[42] Veiga C, Lourenco AM, Mouinuddin S, van Herk M, Modat M, Ourselin S, et al. Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm. Med Phys. 2015;42:760-9.
[43] Zukauskaite R, Brink C, Hansen CR, Bertelsen A, Johansen J, Grau C, et al. Open source deformable image registration system for treatment planning and recurrence CT scans : Validation in the head and neck region. Strahlenther Onkol. 2016;192:545-51.
[44] Kim JS, Lee JM, Lee YH, Kim JS, Kim IY, Kim SI. Intensity based affine registration including feature similarity for spatial normalization. Comput Biol Med. 2002;32:389-402.
[45] Xiong L, Viswanathan A, Stewart AJ, Haker S, Tempany CM, Chin LM, et al. Deformable structure registration of bladder through surface mapping. Medical Physics. 2006;33:1848-56.
[46] Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging. 1999;18:712-21.
[47] Brouwer CL, Kierkels RG, van 't Veld AA, Sijtsema NM, Meertens H. The effects of computed tomography image characteristics and knot spacing on the spatial accuracy of B-spline deformable image registration in the head and neck geometry. Radiat Oncol. 2014;9:169.
[48] Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. elastix: A Toolbox for Intensity-Based Medical Image Registration. Ieee T Med Imaging. 2010;29:196-205.
[49] Thevenaz P, Ruttimann UE, Unser M. A pyramid approach to subpixel registration based on intensity. IEEE Trans Image Process. 1998;7:27-41.
[50] Sarrut D, Boldea V, Miguet S, Ginestet C. Simulation of four-dimensional CT images from deformable registration between inhale and exhale breath-hold CT scans. Med Phys. 2006;33:605-17.
[51] Shekhar R, Lei P, Castro-Pareja CR, Plishker WL, D'Souza WD. Automatic segmentation of phase-correlated CT scans through nonrigid image registration using geometrically regularized free-form deformation. Med Phys. 2007;34:3054-66.
[52] Penney GP, Weese J, Little JA, Desmedt P, Hill DL, Hawkes DJ. A comparison of similarity measures for use in 2-D-3-D medical image registration. IEEE Trans Med Imaging. 1998;17:586-95.
[53] Moseley J, Munro P. A semiautomatic method for registration of portal images. Med Phys. 1994;21:551-8.
[54] Dong L, Boyer AL. An image correlation procedure for digitally reconstructed radiographs and electronic portal images. Int J Radiat Oncol Biol Phys. 1995;33:1053-60.
[55] McParland BJ, Kumaradas JC. Digital portal image registration by sequential anatomical matchpoint and image correlations for real-time continuous field alignment verification. Med Phys. 1995;22:1063-75.
[56] Studholme C, Hill DL, Hawkes DJ. Automated 3-D registration of MR and CT images of the head. Med Image Anal. 1996;1:163-75.
[57] Maes F, Vandermeulen D, Suetens P. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. Med Image Anal. 1999;3:373-86.
[58] Thevenaz P, Unser M. Optimization of mutual information for multiresolution image registration. IEEE Trans Image Process. 2000;9:2083-99.
[59] Marinelli M, Martinez-Moller A, Jensen B, Positano V, Weismuller S, Navab N, et al. Registration of myocardial PET and SPECT for viability assessment using mutual information. Med Phys. 2010;37:2414-24.
[60] Klabbers BM, de Munck JC, Slotman BJ, Langendijk HA, de Bree R, Hoekstra OS, et al. Matching PET and CT scans of the head and neck area: development of method and validation. Med Phys. 2002;29:2230-8.
[61] Roberson PL, McLaughlin PW, Narayana V, Troyer S, Hixson GV, Kessler ML. Use and uncertainties of mutual information for computed tomography/ magnetic resonance (CT/MR) registration post permanent implant of the prostate. Med Phys. 2005;32:473-82.
[62] Wu X, Dibiase SJ, Gullapalli R, Yu CX. Deformable image registration for the use of magnetic resonance spectroscopy in prostate treatment planning. Int J Radiat Oncol Biol Phys. 2004;58:1577-83.
[63] Faggiano E, Fiorino C, Scalco E, Broggi S, Cattaneo M, Maggiulli E, et al. An automatic contour propagation method to follow parotid gland deformation during head-and-neck cancer tomotherapy. Phys Med Biol. 2011;56:775-91.
[64] Rohlfing T, Maurer CR, Jr., Bluemke DA, Jacobs MA. Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. IEEE Trans Med Imaging. 2003;22:730-41.
[65] Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, 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;117:83-90.
[66] Barker JL, Jr., Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys. 2004;59:960-70.
[67] Hansen EK, Bucci MK, Quivey JM, Weinberg V, Xia P. Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2006;64:355-62.
[68] Robar JL, Day A, Clancey J, Kelly R, Yewondwossen M, Hollenhorst H, et al. Spatial and dosimetric variability of organs at risk in head-and-neck intensity-modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2007;68:1121-30.
[69] Vasquez Osorio EM, Hoogeman MS, Al-Mamgani A, Teguh DN, Levendag PC, Heijmen BJ. Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration. Int J Radiat Oncol Biol Phys. 2008;70:875-82.
[70] Fitzpatrick JM, West JB, Maurer CR, Jr. Predicting error in rigid-body point-based registration. IEEE Trans Med Imaging. 1998;17:694-702.
[71] Chalana V, Kim Y. A methodology for evaluation of boundary detection algorithms on medical images. IEEE Trans Med Imaging. 1997;16:642-52.
[72] Dice L. Measures of the Amount of Ecologic Association Between Species. Ecology. 1945;26:297-302.
[73] Yang X, Pei J, Shi J. Inverse consistent non-rigid image registration based on robust point set matching. Biomed Eng Online. 2014;13 Suppl 2:S2.
[74] Hanna GG, Hounsell AR, O'Sullivan JM. Geometrical analysis of radiotherapy target volume delineation: a systematic review of reported comparison methods. Clin Oncol (R Coll Radiol). 2010;22:515-25.
[75] Jena R, Kirkby NF, Burton KE, Hoole AC, Tan LT, Burnet NG. A novel algorithm for the morphometric assessment of radiotherapy treatment planning volumes. Br J Radiol. 2010;83:44-51.
[76] Hoffmann C, Krause S, Stoiber EM, Mohr A, Rieken S, Schramm O, et al. Accuracy quantification of a deformable image registration tool applied in a clinical setting. J Appl Clin Med Phys. 2014;15:4564.
[77] Hou J, Guerrero M, Chen W, D'Souza WD. Deformable planning CT to cone-beam CT image registration in head-and-neck cancer. Med Phys. 2011;38:2088-94.
[78] Garcia-Molla R, de Marco-Blancas N, Bonaque J, Vidueira L, Lopez-Tarjuelo J, Perez-Calatayud J. Validation of a deformable image registration produced by a commercial treatment planning system in head and neck. Phys Med. 2015;31:219-23.
[79] Wang ZH, Yan C, Zhang ZY, Zhang CP, Hu HS, Kirwan J, et al. Radiation-induced volume changes in parotid and submandibular glands in patients with head and neck cancer receiving postoperative radiotherapy: a longitudinal study. Laryngoscope. 2009;119:1966-74.
[80] Castadot P, Geets X, Lee JA, Christian N, Gregoire V. Assessment by a deformable registration method of the volumetric and positional changes of target volumes and organs at risk in pharyngo-laryngeal tumors treated with concomitant chemo-radiation. Radiother Oncol. 2010;95:209-17.
[81] Brouwer CL, Steenbakkers RJ, Langendijk JA, Sijtsema NM. Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help? Radiother Oncol. 2015;115:285-94.
[82] Ricchetti F, Wu BB, McNutt T, Wong J, Forastiere A, Marur S, et al. Volumetric Change of Selected Organs at Risk during Imrt for Oropharyngeal Cancer. Int J Radiat Oncol. 2011;80:161-8.
[83] Li X, Zhang Y, Shi Y, Wu S, Xiao Y, Gu X, et al. Comprehensive evaluation of ten deformable image registration algorithms for contour propagation between CT and cone-beam CT images in adaptive head & neck radiotherapy. PLoS One. 2017;12:e0175906.
[84] Hvid CA, Elstrom UV, Jensen K, Alber M, Grau C. Accuracy of software-assisted contour propagation from planning CT to cone beam CT in head and neck radiotherapy. Acta Oncol. 2016:1-7.
[85] Peroni M, Ciardo D, Spadea MF, Riboldi M, Comi S, Alterio D, et al. Automatic segmentation and online virtualCT in head-and-neck adaptive radiation therapy. Int J Radiat Oncol Biol Phys. 2012;84:e427-33.
[86] Eiland RB, Maare C, Sjostrom D, Samsoe E, Behrens CF. Dosimetric and geometric evaluation of the use of deformable image registration in adaptive intensity-modulated radiotherapy for head-and-neck cancer. J Radiat Res. 2014;55:1002-8.

[87] Landry G, Nijhuis R, Dedes G, Handrack J, Thieke C, Janssens G, et al. Investigating CT to CBCT image registration for head and neck proton therapy as a tool for daily dose recalculation. Med Phys. 2015;42:1354-66.

Chapter 5 – Delivered dose to the spinal cord

5.1 Overview

The spinal cord is a critical, dose-limiting OAR in radiotherapy for HNC patients. Radiation induced damage to the spinal cord can lead to transverse myelitis – a potentially catastrophic condition manifest by paresis, and sensory and autonomic dysfunction distal to the lesion. With modern radiotherapy techniques, this phenomenon is extremely rare, but dose-response curves are steep. Furthermore, there is widespread adherence to the idea that weight loss and anatomical change during treatment may lead to significant differences between planned and delivered dose to OARs, including the spinal cord. Despite this, there is little published data that quantifies delivered spinal cord dose, and still less to identify which patients are at most risk of delivered spinal cord dose exceeding that which was intended at planning. This chapter presents the methodology and results of a large-scale analysis of spinal cord delivered dose, and the factors that might predict important differences.

5.1.1 My Role

I developed all the hypotheses examined in this work, and the overall design of the study to address them was mine. I did all manual segmentations of the spinal cord on kVCT planning scans described in this chapter. These contours were used to define planned dose, and as prior structures for DIR to generate automated contours, and delivered dose calculations, on MVCTs.

I retrieved, checked and collated all data pertaining to patient staging, dose gradient, weight loss and anatomical change. I interpreted and processed raw DICOM data, generated by Karl Harrison at the Cavendish lab, in order to generate simulations of zero-IG. I undertook all analyses of planned and delivered dose, and designed and executed all experiments in which relationships between hypothesised predictors and dose differences were assessed. The results of this work were published as [1]:

Anatomical change during radiotherapy for head and neck cancer, and its effect on delivered dose to the spinal cord. DJ Noble, PL Yeap, SY Seah, K Harrison, LEA Shelley, M Romanchikova et al. *Radiother Oncol. 2019 Jan;130: 32-38.*

And the Chapter contains a number of figures, tables and sections of text from this paper.

5.1.2 Acknowledgements

As described in detail in Chapter 4, work to develop and optimise the DIR algorithm used for automated contouring and dose accumulation was led by Lin Yeap (MPhil student, Cavendish laboratory), with significant input from myself, Karl Harrison (Research Associate, Cavendish laboratory) and Shannon Seah (MPhil student, Cavendish laboratory). Karl Harrison designed the architecture to curate all the DICOM imaging, dose and structure set data on the Cavendish laboratory High Energy Physics cluster. Shannon Seah undertook final implementation of the Elastix code within python, and executed the final run to generate the delivered dose data presented in this chapter.

Marina Romanchikova, computer scientist, wrote the software for extracting images, structure sets and dose cubes from vendor archive, as described in Chapter 2. Marina Romanchikova also transferred DICOM structure set files from a research database to the Cavendish Laboratory. The algorithm for calculating daily dose on MVCT images (CheckTomo) was written by Simon Thomas (head of medical physics, Addenbrooke's Hospital). Amy Bates (clinical trials radiographer) recruited all patients to the VoxTox study, and gathered baseline data on weight and staging. Final treatment week weight data was collected by dieticians from the Addenbrooke's H&N unit.

I am grateful to Dr Eliana Vasquez-Osorio (Senior Research Fellow, University of Manchester) and Professor Jan-Jacob Sonke (Adaptive Radiotherapy group leader, NKI, Amsterdam) for helpful discussions about deformable image registration for dose accumulation.

5.2 Introduction

5.2.1 The spinal cord as an OAR in H&N radiotherapy

The spinal cord is a critical, dose-limiting OAR in radiotherapy for HNC patients [2]. In the pre-IMRT era, radiotherapy techniques for HNC were designed to avoid irradiating the cervical spinal cord in excess of its tolerance dose [3]. This caution remains paramount with modern techniques. Whilst optimising IMRT plans, most centres will prioritise adequate dose coverage of target volumes over sparing of glandular, muscular, or bony OARs such as parotid glands, pharyngeal constrictors, or mandible. In contrast however, adequate sparing of neurological OARs, including the spinal cord and brainstem, is given the highest priority in the TPS

optimiser. Thus, in the vast majority of clinical scenarios, radiotherapists will prioritise adequate protection of these neurological structures over full dose coverage of the target volumes. This principle is manifest in the protocols of contemporary H&N radiotherapy trials, where planning objectives and constraints are frequently lower than the stated organ tolerance [4-6].

This approach is ubiquitous in modern radiotherapy, because the consequences of exceeding these dose constraints are potentially catastrophic. Transverse myelitis is a condition characterised by paresis, and both sensory and autonomic dysfunction below the level of the lesion [7]. Irradiation of the spinal cord is a well-recognised cause [8], and the underlying radiobiology has been extensively studied in animal models [9-12].

L'Hermitte's syndrome (LS) is a more benign self-limiting condition, which is also an established toxicity following irradiation of the spinal cord [13]. It is characterised by electric shock-like sensations, down the spine and into the limbs on neck flexion, and is also a symptom of demyelinating conditions such as Multiple Sclerosis [14]. It is thought to be due to transient inhibition of oligo-dendrocyte activity, with subsequent reversible demyelination [15, 16]. LS risk is higher in patients whose spinal cords receive >40Gy [17, 18], although there is evidence to suggest that the incidence of LS has risen in the IMRT era. Reported risk in patients treated with field-based techniques was 3-13% [19-21], compared with 15-36% in patients treated with IMRT [13, 17, 18, 22]. It is hypothesised that this observation may be related to inhomogeneous dose distributions and dose gradient across the spinal cord [18, 22].

There are historical, single-institution series from the pre-IMRT era that report a low, but measurable risk of transverse myelitis following radiotherapy for HNC [8, 23-26]. These studies were assessed and synthesised in a review of 2008, and in the Quantec report of 2010 [27, 28]. A conclusion of this work was the clear relationship between radiotherapy dose, and the risk of transverse myelitis. These papers therefore produced a logistic dose-response model relating the risk of transverse myelitis to maximum spinal cord dose, assuming 2Gy/fraction, and $\alpha/\beta = 0.87$ [27, 28]. The model shows a typical, steep, dose-response curve. The quoted risk of transverse myelitis at 45Gy (2Gy/fraction) is 0.03%, at 50Gy it is 0.2%, whilst at 69.4Gy the risk is 50%. It is therefore clear that modest increases in spinal cord dose can lead to a dramatic increase in the risk of transverse myelitis.

5.2.2 Potential differences in spinal cord dose

As discussed in Chapter 1, significant anatomical change and weight loss are common during radical radiotherapy for HNC [29-33]. It is thought that such changes may lead to differences

between planned radiation dose, and that which is actually delivered [34]. It is further hypothesised that all internal structures, including the spinal cord, may therefore be subject to significant differences between D_P and D_A. Available literature on this topic is surprisingly scarce, and studies to date have been relatively small [33, 35-38]. Furthermore, some have focussed on the spinal canal as a surrogate for the spinal cord itself. The results of these studies suggest that dose differences are generally small, and depend on the frequency and quality of IG [35, 36].

However, to the best of my knowledge, no previous studies have examined dose differences to the spinal cord in a large prospectively collected cohort. Additionally, none have done so by accumulating dose from daily IG imaging, and no systematic attempts have been made to identify potential associations between weight loss and anatomical change and differences in spinal cord dose, or to search for alternative predictors of such differences. Furthermore, in my own clinical experience, concerns about the impact of weight loss, anatomical change and setup uncertainty on dose to the spinal cord are a common reason for instigating ART, often in the absence of any data to inform the decision.

Therefore, the primary hypothesis of this chapter was that patients experiencing significant weight loss and anatomical change during their radiotherapy treatment would be more likely to have a delivered maximum spinal cord dose higher than that anticipated at planning.

5.2.3 Aims of chapter 5:

1. To calculate delivered dose to the spinal cord in a large cohort, and quantify differences between planned and delivered dose.
2. To examine hypothesised links between weight loss, anatomical change, and spinal cord dose differences.
3. To search for links between other factors, and spinal cord dose differences. Specifically these will include; T and N stage, treatment laterality, dose gradient, and the role of image guidance.
4. To use this data to make recommendations about how to manage uncertainties around spinal cord dose during radical radiotherapy for HNC patients.

5.3 Methods

5.3.1 Cohort selection and treatment details

All patients included in this analysis were recruited to the VoxTox study, as outlined in Chapter 2. However, data in Table 2.3 show that the overall VoxTox H&N cohort is somewhat heterogeneous in terms of primary disease site, and treatment technique. Therefore, for this sub-study, it was necessary to define specific inclusion criteria to produce a cohort of patients in which these hypotheses could be adequately addressed. These inclusion criteria were:

- Treatment between 2010 and 2016
- Primary sites:
    - SCCs of the oropharynx, oral cavity, nasopharynx, hypopharynx, or larynx.
    - SCCs histologically compatible with H&N mucosal origin, but no clear primary site.
    - Primary salivary gland tumours requiring post-operative radiotherapy.
- Minimum prescription dose of 60Gy in 30 fractions.
- Neck irradiation to include at least levels II and III ipsilaterally.
- Daily TomoTherapy treatment, with all IG-MVCT images available for analysis.

Exclusion criteria for this cohort were as follows:

- Other HNC primary disease sites including skin, sinus, paraganglioma, lacrimal gland.
- No, or minimal, neck irradiation.
- Prescription dose of <60Gy in <30 fractions.

These criteria yielded a final study cohort of 133 patients, whose characteristics are summarised in table 5.1. The use of concomitant systemic therapy did not affect eligibility for this sub-study, and patients were treated with weekly cisplatin, cetuximab, or no concomitant therapy as clinically indicated.

Radiotherapy treatment for all patients in the study was in line with standard Addenbrooke's protocols, as described in Chapter 2. In brief, patients were immobilised with a 5-point fixation thermoplastic shell for their planning CT and treatment. Target volume and OAR definition were in accordance with published atlases, and a current UK trial protocol [5, 39]. For treatment planning, the spinal cord was expanded by 3mm axially to a prv. A dose objective of 46Gy and a constraint of 50Gy were defined, and given top priority, in the TPS optimiser. All patients were treated on TomoTherapy Hi-Art units, with daily IG and positional correction, with a zero-

action level approach (DIPC) [40]. The specifics of the IG workflow are shown in Appendix A2.3.

Although the VoxTox protocol stipulates that IG-MVCTs done for study patients are longer than standard, the FoV is substantially shorter than planning kVCTs, as shown in Figure 5.1. However, the superior limit of the spinal cord, at the level of the tip of the dens of C2 [39], was included in all scans.

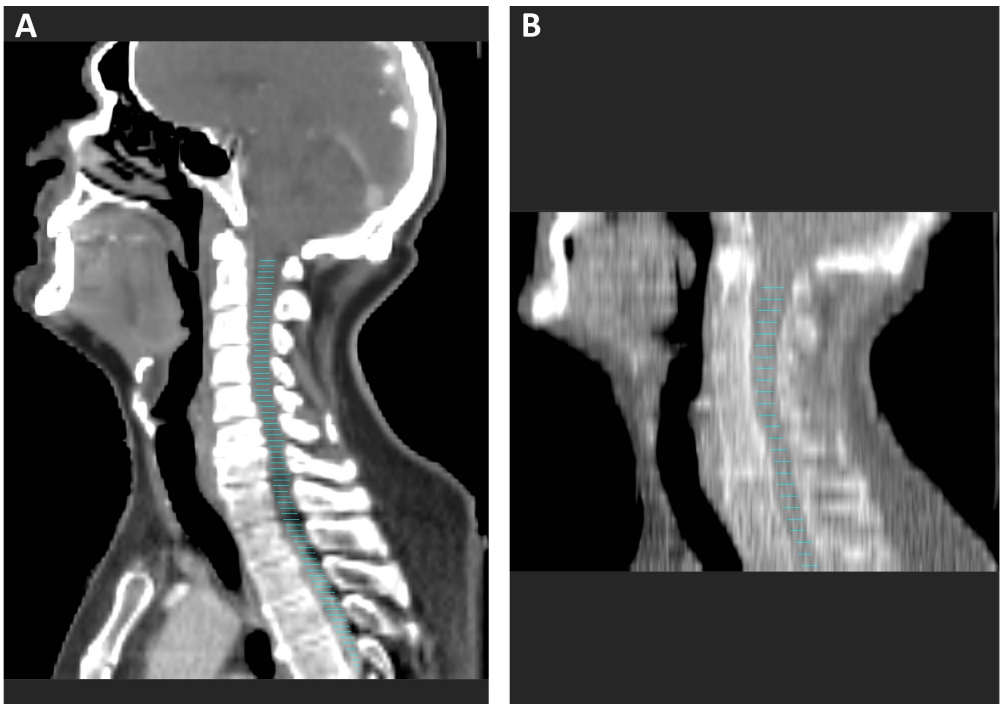


Figure 5.1 (A&B): Imaging field of view (FoV) for kVCT planning scan (A) and MVCT IG scan (B). Spinal cord contours shown in cyan.

5.3.2 Planned and delivered dose calculation

5.3.2.1 Data transfer

Patient imaging data, including all kVCT planning scans and MVCT IG images, structure set data, and treatment delivery data were extracted from vendor archive by Marina Romanchikova, as described in Chapter 2. For the purposes of this sub-study, retrieved and anonymised kVCT planning scans for this 133 patient cohort were uploaded to the research database of the department's segmentation software (Prosoma 3.3, MEDCOM, Darmstadt, Germany).

Characteristic	Number
Age	58.5 (10.1)
Gender	
Male	112 (84.2%)
Female	21 (15.8%)
Baseline weight (kg)	87.7 (19.3)
Disease	
SCC	123 (92.5%)
T0-2	77 (62.6%)
T3-4	46 (37.4%)
N0-1	45 (36.6%)
N2a-3	78 (63.4%)
Oropharynx	84 (68.3%)
Oral cavity	12 (9.8%)
Larynx	11 (8.9%)
Hypopharynx	5 (4.1%)
Nasopharynx	3 (2.4%)
Unknown primary	8 (6.5%)
Salivary Gland	10 (7.5%)
Dose/Fractionation	
70/35	3 (2.3%)
68/34*	13 (9.8%)
65/30*	88 (66.1%)
60/30	29 (21.8%)
Neck Irradiation	
Unilateral	38 (28.6%)
Bilateral	95 (71.4%)
Systemic therapy	
Cisplatin†	75 (56.4%)
Cetuximab†	11 (8.3%)
None	47 (35.3%)

* primary SCC's of the oro/hypopharynx, and larynx treated with 68Gy/34# prior to November 2011, 65Gy/30# thereafter.
† dose: cisplatin – 40mg/m^2 weekly, cetuximab 400mg/m2 loading dose, 240mg/m^2 weekly thereafter.

Table 5.1: patient characteristics – spinal cord delivered dose cohort (n = 133). For continuous variables, means and standard deviations are reported, absolute numbers and percentages for proportions.

5.3.2.2 Spinal cord segmentation

Although clinical contours were available for many patients, I re-segmented the spinal cord for all 133 patients, according to contouring guidelines [39]. For many patients in the cohort, the spinal cord contour that was used for treatment planning was an automated segmentation generated by the commercial ABAS system, SPICE. Subjectively, many of these contours appeared to delineate spinal canal, rather than spinal cord, as shown in Figure 5.2. Therefore, I decided to repeat all segmentations myself, to maximise consistency in contouring approach, and reduce potential sources of error and bias. Intra-observer contouring consistency data, and inter-observer data relative to consultant oncologist's expert in treating HNC are presented in chapter 2. Completed segmentations were then sent to the Cavendish laboratory for curation and processing.

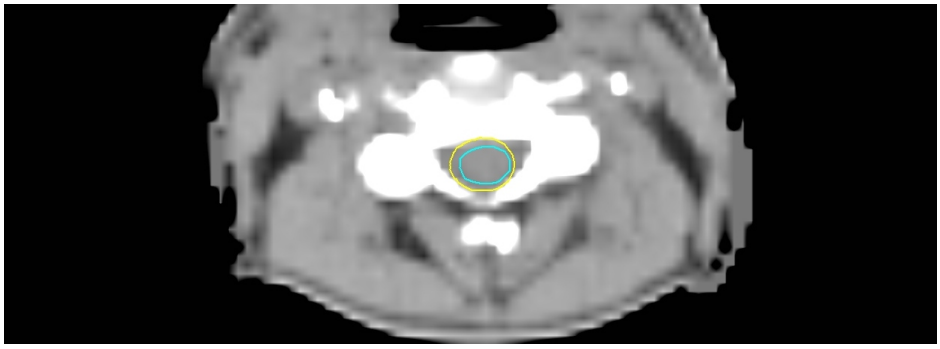


Figure 5.2: Comparison of my spinal cord contour (cyan) with ABAS based auto-contour (yellow), which appears to be defining spinal canal rather than spinal cord on this slice.

5.3.2.3 MVCT auto-contouring

An open-source deformable image registration toolbox, Elastix [41], was used to propagate spinal cord contours from kVCT to all daily MVCT images, for every patient. This process was implemented in the python coding language. The approach to training and validating Elastix for this task was very similar to that described in Chapter 4. Data on the accuracy and uncertainties associated with Elastix generated spinal cord auto-contours on Daily MVCT images are presented in the MPhil thesis of Lin Yeap[11] and the subsequent publication [42].

5.3.2.4 Daily dose calculation

Daily dose cubes, in the geometrical frame of references of the IG-MVCT, were calculated using an in-house algorithm known as CheckTomo [43]. This software was written in MATLAB

[11] MPhil Thesis: Mr. Ping Lin Yeap, Jesus College, University of Cambridge, 2016.

(MathWorks, Natick, Massachusetts, USA) by Simon Thomas (head of Medical Physics, Addenbrooke's Hospital), although the code is deployed within Python for all work in the VoxTox programme. CheckTomo was designed to be an independent check of dose calculations produced by vendor software (Accuray), and has been in use clinically in the department since 2012. It takes the DICOM image data, the treatment sinogram, tables of measured beam data, patient setup correction data, and a fast ray-trace algorithm to compute dose in each voxel of the image. Although it was written for clinical use with kVCT planning scans, TomoTherapy MVCT scans have stable HU to electron-density curves, which permit direct dose calculation against these images, thus permitting computation of daily dose cubes at scale [44]. Once daily dose cubes were computed, automated segmentations derived from Elastix were used to calculate daily delivered dose to the spinal cord.

5.3.2.5 Dose accumulation

It was also necessary to accumulate dose from every treatment fraction to a final summary representation of delivered dose (D_A). DIR, and the Elastix toolbox, lends itself well to solving this problem. As described in section 4.2.1, the image registration process defines the MVCT as I_M, so once the registration is completed, the anatomy of the MVCT will have been warped in such a way as to minimise differences with the planning kVCT (I_F). Furthermore, the transformation parameters that determine the deformation vector field of I_M, relate the geometry of each voxel of this image to the geometry of I_F – the kVCT. The daily delivered dose cube, also in DICOM format, has the same geometrical frame of reference. Thus, Elastix defines two important characteristics about the daily dose cube; firstly, which dose voxels represent which structure, and secondly the geometry of that daily dose cube in the frame of reference of the kVCT, including any changes relative to the original, undeformed MVCT frame of reference. Warping of I_M (and therefore the dose cube) is controlled by a series of connected points. As the distances between these points move in the deformation process, the effect on the dose grid is solved by linear interpolation. Thus, as the deformation produces a daily DVH in the frame of reference of the planning scan, it is anatomically plausible, and mathematically straightforward to simply sum individual daily voxel dose values over each fraction, to a final summary D_ADVH. That was the approach taken for this work. An overview of the workflow for propagating contours to daily MVCT scans, calculating daily dose, and accumulating to a final summary D_ADVH is shown in Figure 5.3.

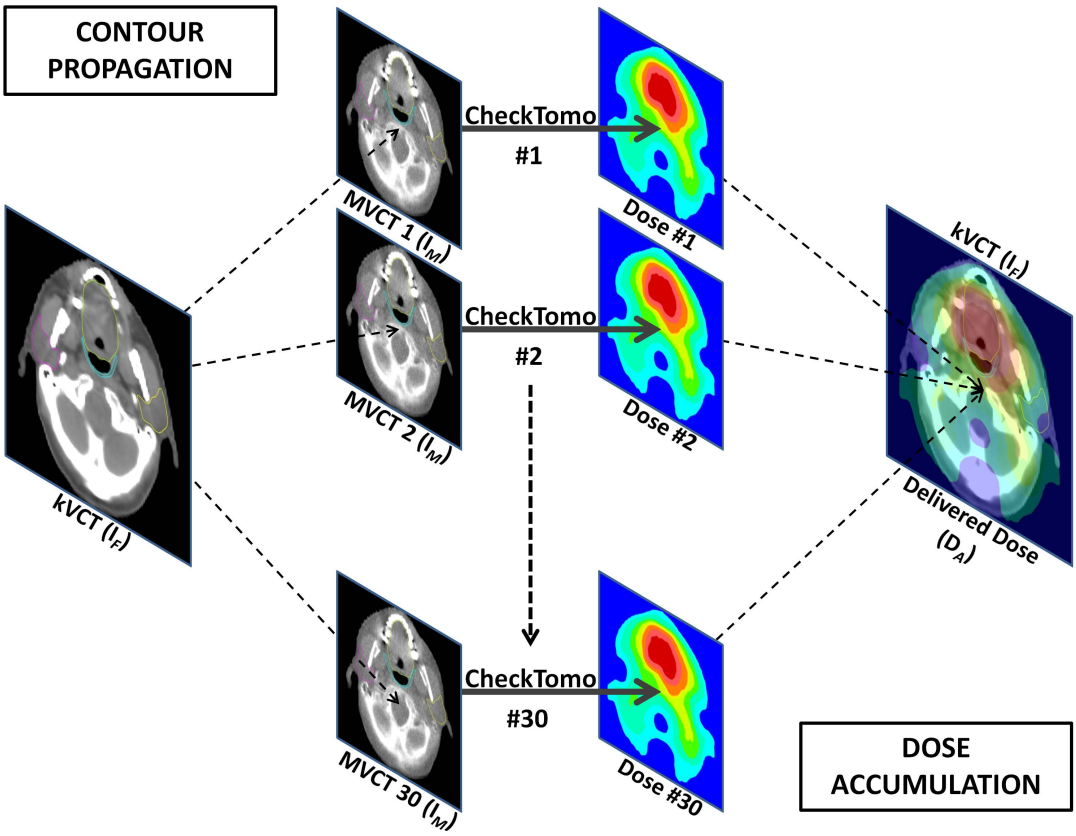


Figure 5.3: Overview of contour propagation and dose accumulation using DIR (Elastix). Contours are propagated from kVCT planning scan to daily MVCTs following image registration. Daily dose is calculated for each MVCT using CheckTomo. Finally, transformation parameters from Elastix accumulate daily dose in the geometric frame of reference of the planning kVCT, to give summary delivered dose (D_A) DVH.

5.3.2.6 Dose reporting

In order to minimise potential sources of error in comparison of planned and delivered dose, D_P DVHs were also recalculated using my repeated manual segmentations, and CheckTomo software. Delivered and planned dose data for the spinal cord, for all 133 patients in this study were sent from the Cavendish laboratory to me for further analysis. As the spinal cord is a serial organ [28], maximum dose, represented by $D_{2\%}$, was assessed. This dose metric is in-line with ICRU 83 [45]. The difference between $D_A D_{2\%}$ and $D_P D_{2\%}$ ($\Delta SC_{D2\%}$) was defined as ($D_A - D_P$), as the clinically relevant difference in this context is higher delivered dose. These results were also used for all subsequent comparison with predictive variables.

5.3.3 Anatomical change, and predictors of dose differences

Disease staging data was recorded in VoxTox baseline CRFs, as shown in Appendix A3.1. Where this was missing, or incomplete, I used the MOSAIQ data management software

(Elekta, Stockholm, Sweden), to retrieve radiotherapy booking and prescription paperwork that included this information. In line with previous work [38], binary classification was used for both T and N-stage (T0-2 vs T3-4, and N0-1 vs N2-3). Using MOSAIQ, I recorded whether or not irradiation of neck node regions had been unilateral or bilateral.

Patients in VoxTox are weighed at baseline, and weekly weights are recorded in MOSAIQ. Weight loss (ΔWL) was simply the weight recorded in the final treatment week subtracted from baseline weight. There were missing weight data for 28 patients, thus 105 were included in the analysis of ΔWL against $\Delta SC_{D2\%}$. Based on a previously published methodology [38], changes in patient separation were measured as a surrogate of anatomical change. First and final fraction MVCTs were loaded into Prosoma, and caliper measurements of the lateral neck diameter (LND) were made at the level of the C1 vertebra, and the superior thyroid notch (TN), as shown in Figure 5.4 (A-D). Using an automated function within Prosoma, external skin contours were produced on the same CT slices, and the surface area of this segmentation was recorded as slice surface area (SSA) (Figure 5.4 E-H). For each metric, changes between fractions 1 and 30 were recorded as ΔC1LND, ΔC1SSA, ΔTNLND and ΔTNSSA. One patient with very atypical setup (extreme cervical kyphosis; axial plane at CI included maxillary sinus anteriorly, and spinous process of C3 posteriorly) was excluded, leaving 132 for this analysis.

In order to assess the impact of DIPC on dose differences, the impact of no IG was simulated. The MVCT scan DICOM headers include the daily couch shifts made by radiographers and the TomoTherapy system, after IG has been done. These values were combined to compute a mean per-patient couch shift. Using Prosoma, the spinal cord was translated by the inverse of this mean shift, relative to the geometry of the planning kVCT and the planned dose cube, and $D_{2\%}$ recorded in this position. Planned $D_{2\%}$ was then subtracted from this value to give a simulated 'No IG' $D_A D_{2\%}$ value.

Finally, Prosoma was used to calculate dose gradient in the vicinity of the spinal cord. First, the spinal cord contour was grown axially by 6mm – twice the prv margin. Next, the dose wash was interrogated to find the scan slice with the maximum spinal cord dose. On this slice, the point dose on the spinal cord + 6mm ring structure (Ring structure) was recorded, at 0, 90, 180 and 270 degrees relative to the spinal cord centroid. Point dose on the corresponding vector on the spinal cord contour was also recorded, and subtracted from the 4 Ring structure point dose values. These 4 results were summed, and divided by 24 to give an overall dose gradient, in Gy/mm. This method is shown in cartoon form in Figure 5.5.

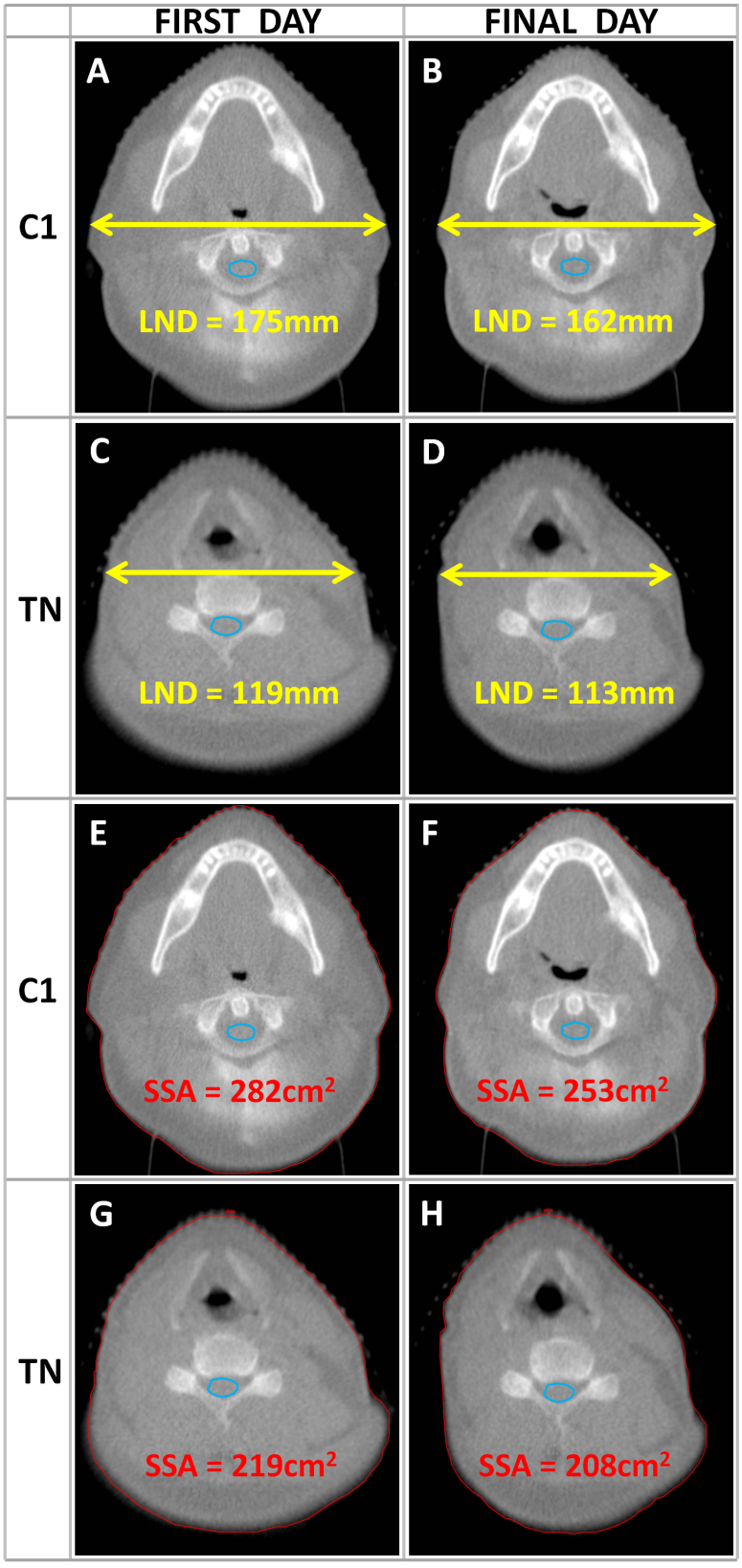


Figure 5.4 (A-H): Anatomical change parameters; lateral neck diameter (LND, captions A-D) and slice surface area (SSA, captions E-H), – measured on the IG-MVCT at the C1 vertebra (C1) and thyroid notch (TN) on the first and final treatment day.

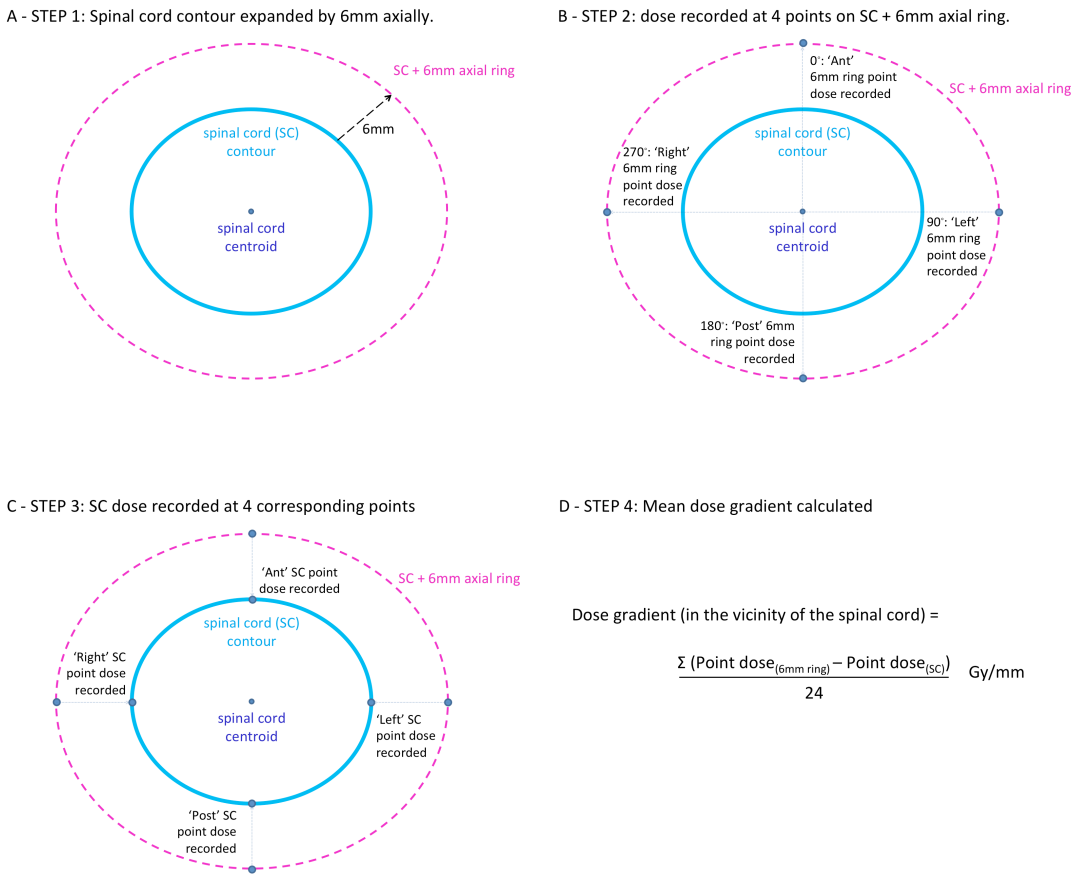


Figure 5.5 (A-D): Methodology for generalising dose gradient in the vicinity of the spinal cord. A – expansion of spinal cord contour to 'Ring structure', B – Point dose on Ring structure at 0, 90, 180 and 270 degrees recorded. C – corresponding point dose on spinal cord recorded. D – dose gradient calculation.

5.4 Results

5.4.1 Dose results

Mean $D_P D_{2\%}$ in this cohort was 36.1Gy (95% CI 35.4 to 36.9, range 22.4 to 46.4Gy). Mean $D_A D_{2\%}$ was also 36.1Gy (95% CI 35.3 to 36.8, range 22.4 to 46.3Gy). For the whole cohort of 133 patients, $\Delta SC_{D2\%}$ was normally distributed, with a mean value of -0.07Gy (95% CI -0.28 to 0.14, range -5.7 to 3.8Gy). The mean difference between D_A and D_P, independent of difference direction, was 0.9Gy (95% CI 0.76 to 1.04Gy). These data are shown as a Bland-Altmann plot and Histogram in Figures 5.6 and 5.7 respectively.

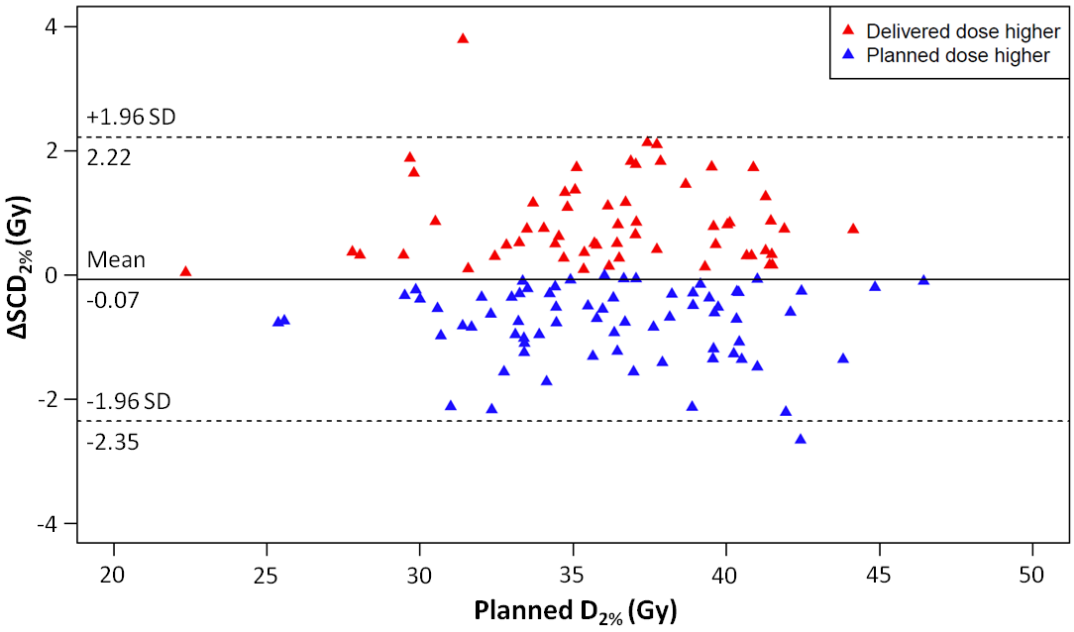


Figure 5.6: Bland-Altmann plot showing $\Delta SC_{D2\%}$ as a function of $D_PD_{2\%}$. Red triangles denote cases where delivered dose was higher, blue triangles those in which planned dose was higher.

Planned $D_{2\%}$ was higher in 72 (54.1%) of patients; delivered $D_{2\%}$ was higher in 61 (45.9%). There were only 4 patients in the cohort where D_A was 2Gy or more higher than D_P, and there were no instances where planned cord dose was within OAR tolerance, but delivered cord dose was not. As shown in Figure 5.5, there was no relationship between the absolute value of $D_PD_{2\%}$ and the direction of dose difference. Most importantly, there was no trend to suggest that patients with higher $D_PD_{2\%}$ were more likely to have a positive $\Delta SC_{D2\%}$.

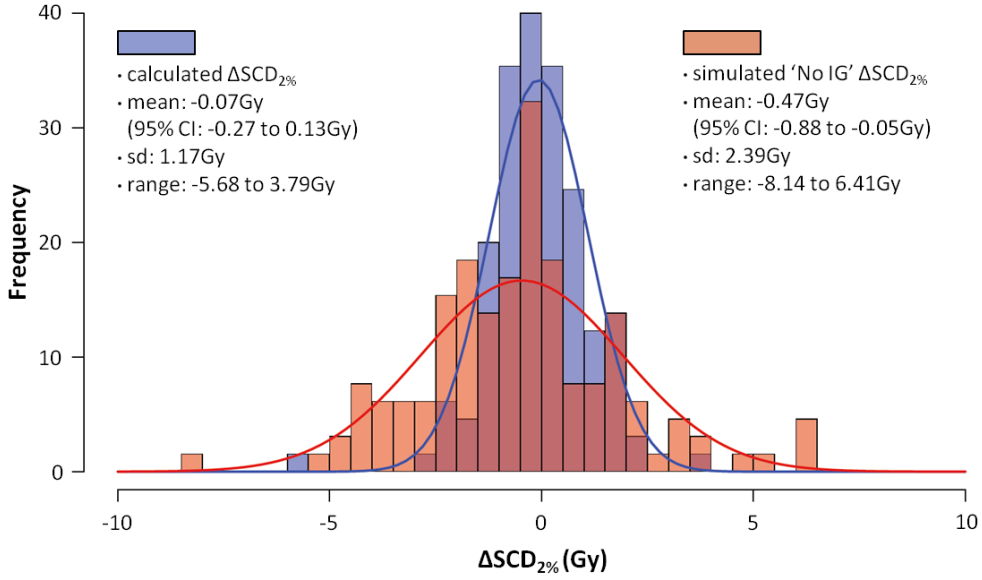


Figure 5.7: Histograms of calculated $\Delta SC_{D2\%}$ (blue), and simulated 'No IG' $\Delta SC_{D2\%}$ (red), with Gaussian distribution functions fitted to the data.

Figure 5.7 also shows data for the simulation of no image-guidance. The sample mean is similar to that seen for calculated $\Delta SC_{D2\%}$, but the distribution is much broader (st. dev. 2.39Gy compared with 1.17Gy, range -8.1 to 6.4Gy compared to -5.7 to 3.8Gy). Thus, the mean, direction agnostic, difference between planned and delivered dose was double that for calculated $\Delta SC_{D2\%}$ (1.8Gy vs 0.9Gy).

5.4.2 Anatomical change and weight loss

Results for weight loss and anatomical change during treatment are shown in Table 5.2. Relationships between weight loss and anatomical change metrics were also investigated with univariate linear regression (Table 5.3). This was firstly to understand how closely linked these changes are, and secondly to exclude potentially confounding co-linearity when conducting univariate analysis of change metrics against dose differences. Even allowing for multiple testing, all relationships were statistically significant ($p<0.005$), and most correlations were moderate (range for Pearson's product moment correlations coefficient, R, 0.28 – 0.61). Those between weight loss and shape metrics (0.28 – 0.40) were generally weaker than those between shape change metric themselves (0.37 – 0.61). There was no evidence of sufficient co-linearity to preclude separate analysis of all metrics against dose differences.

Metric	Start of treatment (95% CI)	End of treatment (95% CI)	Change (End-Start) (Abs, Rel. values) (Range, [%])	Paired t-test (P value) (95% CI)
weight (kg) (n = 105)	86.0 82.5 – 89.6	79.2 75.9 – 82.4	-6.8, -7.9 -22.1 – 6.8	< 0.001 5.8 – 8.0
C1 LND (mm)	154.3 152.1 – 156.5	141.4 139.1 – 143.7	-12.9, -8.4 -22.1 – 0.8	< 0.001 11.8 – 14.1
C1 SSA (cm^2)	225.0 219.7 – 230.3	212.9 208.2 – 217.6	-12.1, -5.4 -15.2 – 8	< 0.001 10.5 – 13.7
TN LND (mm)	123.4 118.7 – 128.1	118.1 113.7 – 122.6	-5.3, -4.3 -17.1 – 7.5	< 0.001 4.1 – 6.5
TN SSA (cm^2)	166.9 156.3 – 177.5	155.7 145.7 – 165.6	-11.2, -6.7 -20.7 – 13.6	< 0.001 9.2 – 13.3

Table 5.2: Anatomical change during treatment: weight loss (WL), lateral neck diameter (LND) and slice surface area (SSA). Measurements made at the level of C1 vertebra and superior thyroid notch (TN). n = 132 unless otherwise stated.

Test	P value	R	95%CI for R
weight loss vs. C1 LND	< 0.001	0.40	0.22 to 0.55
weight loss vs. C1 SSA	< 0.001	0.36	0.19 to 0.52
weight loss vs. TN LND	0.003	0.28	0.10 to 0.45
weight loss vs. TN SSA	< 0.001	0.40	0.23 to 0.55
C1 LND vs. C1 SSA	< 0.001	0.50	0.36 to 0.62
C1 LND vs. TN LND	< 0.001	0.37	0.22 to 0.51
C1 LND vs. TN SSA	< 0.001	0.50	0.36 to 0.62
C1 SSA vs. TN LND	< 0.001	0.47	0.33 to 0.60
C1 SSA vs. TN SSA	< 0.001	0.48	0.34 to 0.60
C4 LND vs. C4 SSA	< 0.001	0.61	0.49 to 0.71

Table 5.3: Univariate relationships between weight loss during treatment and individual anatomical change metrics (Pearson's product moment correlation coefficient, R)

5.4.3 Predictors of spinal cord dose differences

Relationships between neck irradiation strategy, T and N stage are shown in Figure 5.8. Patients who received radiotherapy to neck nodal regions on both sides of the neck were no more likely to see delivered spinal cord $D_{2\%}$ being higher than planned (Relative $\Delta SC_{D2\%}$ was 0.23% for UNI, -0.34% for BNI, two-sample t-test p = 0.39, 95% CI -0.73 to 1.88%). Neither did higher T-stage predict for higher $\Delta SC_{D2\%}$; mean $\Delta SC_{D2\%}$ was -0.31% for T0-2, and 0.04% for T3-4 (p = 0.56, 95% CI -1.56 to 0.86%). The data suggest a weak relationship between higher N-stage and higher $\Delta SC_{D2\%}$ values ($\Delta SC_{D2\%}$ was -0.86% for N0-1, 0.20% for N2-3), although this did not reach statistical significance (p = 0.088, 95% CI -2.29 to 0.16%).

Steeper dose gradient in the vicinity of the spinal cord did predict for higher $\Delta SC_{D2\%}$ (Figure 5.9). On univariate linear regression $r^2 = 0.27$ ($p < 0.001$). Mean dose gradient in patients for whom $\Delta SC_{D2\%}$ was positive was 0.74Gy/mm, compared with 0.28Gy/mm in patients where $\Delta SC_{D2\%}$ was negative (95% CI for difference in means 0.34 – 0.57Gy/mm, $p < 0.001$, two-sample t-test).

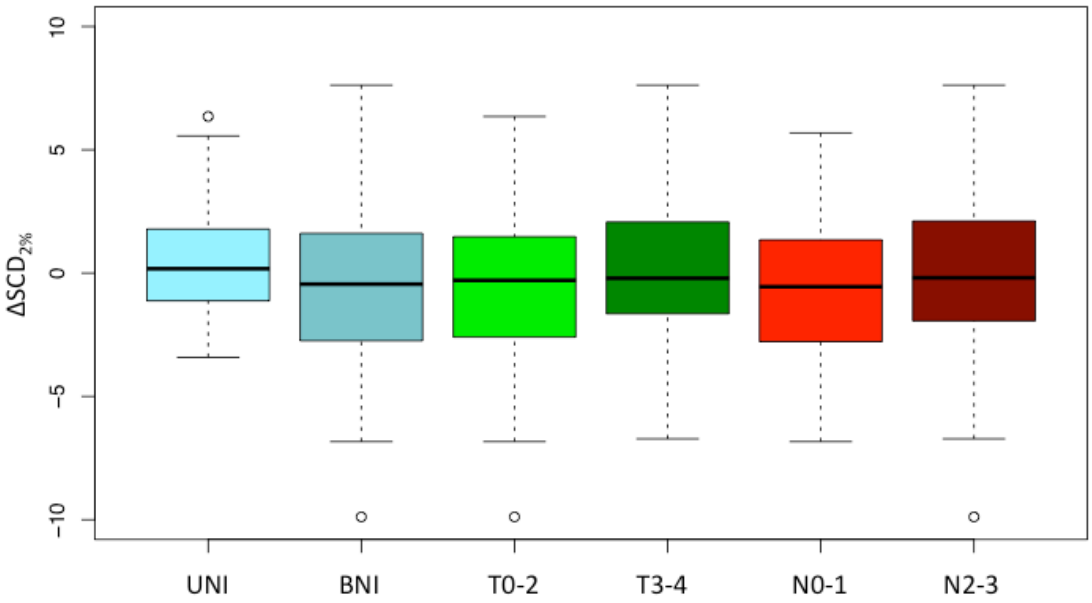


Figure 5.8: Univariate relationships between neck irradiation strategy (unilateral – UNI vs. bilateral – BNI, staging (T and N) and relative dose differences ($D_A - D_P$, $\Delta SC_{D2\%}$) to the spinal cord.

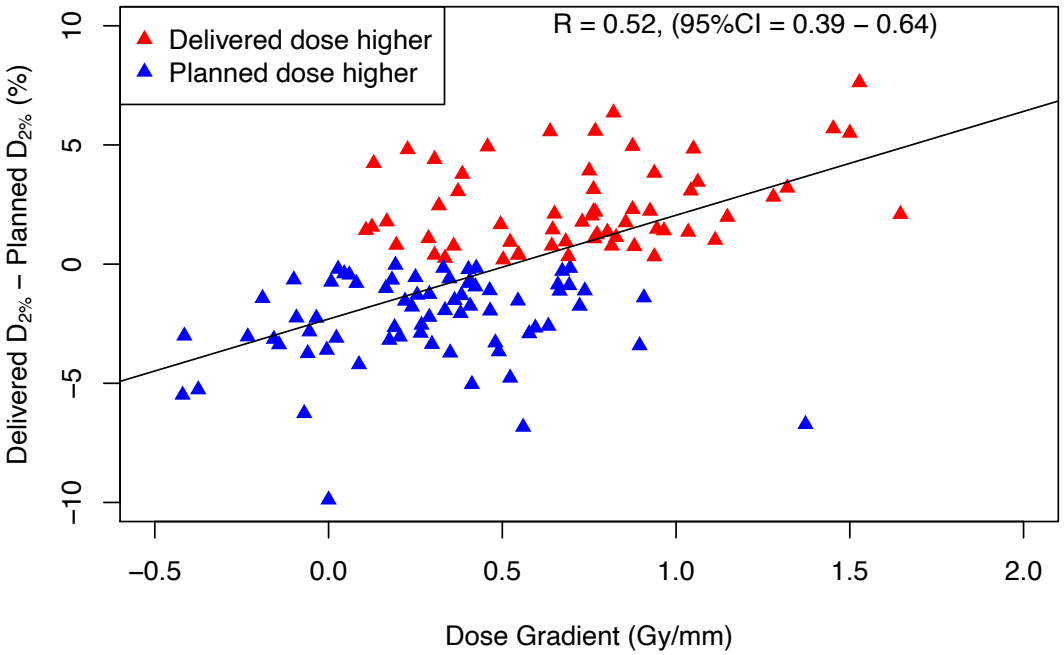


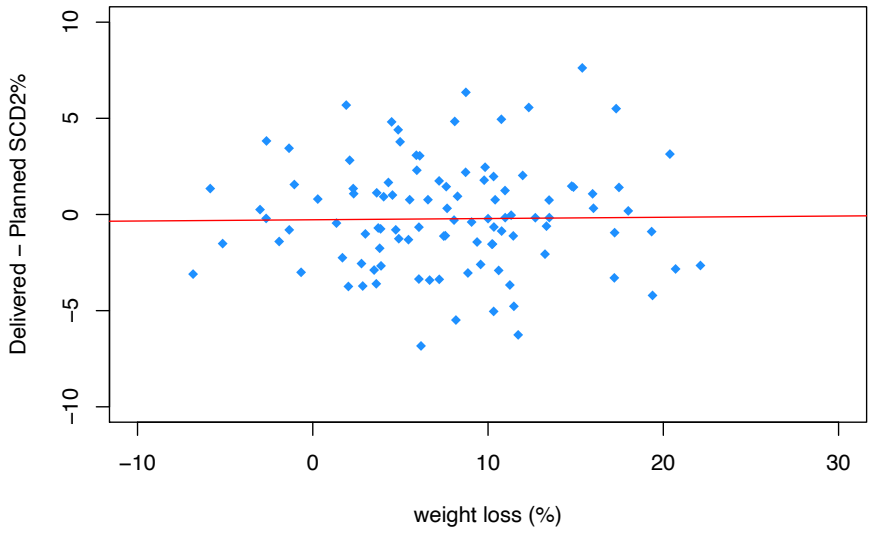
Figure 5.9: Scatter plot of relative spinal cord dose differences ($\Delta SC_{D2\%}$) as a function of dose gradient in the vicinity of the spinal cord.

Univariate linear regression was also used to look for potential relationships between weight loss and anatomical change and $\Delta SC_{D2\%}$. These data are shown as scatter plots in Figure 5.10 (A-E), and summarised in Table 5.4. As can be seen from these results, no meaningful relationships were observed ($r^2 < 0.05$ for all models).

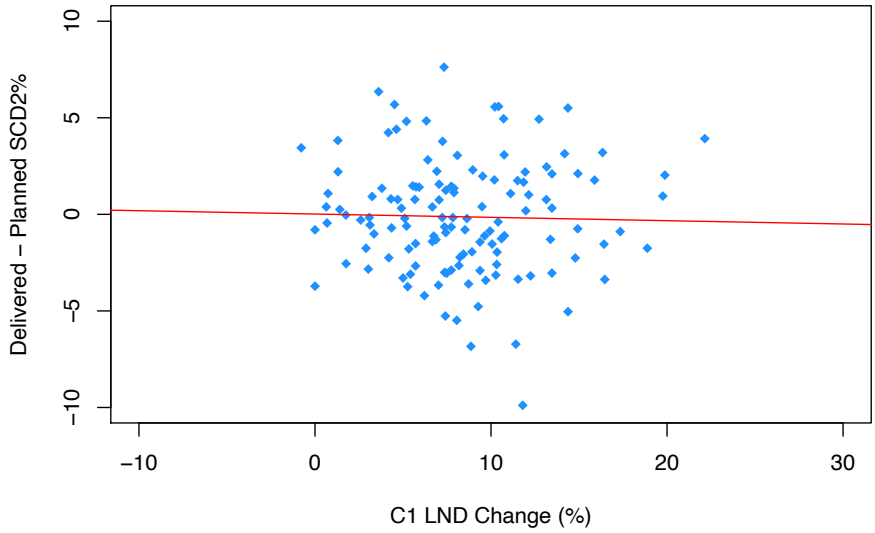
Model	Intercept	B	r^2	P
ΔWL vs. $\Delta SCD_{2\%}$	-0.27	0.006	< 0.001	0.90
ΔC1LND vs. $\Delta SCD_{2\%}$	0.02	-0.017	< 0.001	0.80
ΔC1SSA vs. $\Delta SCD_{2\%}$	0.59	-0.14	0.02	0.08
ΔTNLND vs. $\Delta SCD_{2\%}$	0.13	-0.06	0.002	0.26
ΔTNSSA vs. $\Delta SCD_{2\%}$	-0.12	< 0.001	< 0.001	0.99

Table 5.4: Summary results of univariate linear regression models of weight loss (WL) and shape change (lateral neck diameter – LND, slice surface area – SSA) during treatment against changes in spinal cord dose ($SCD_{2\%}$).

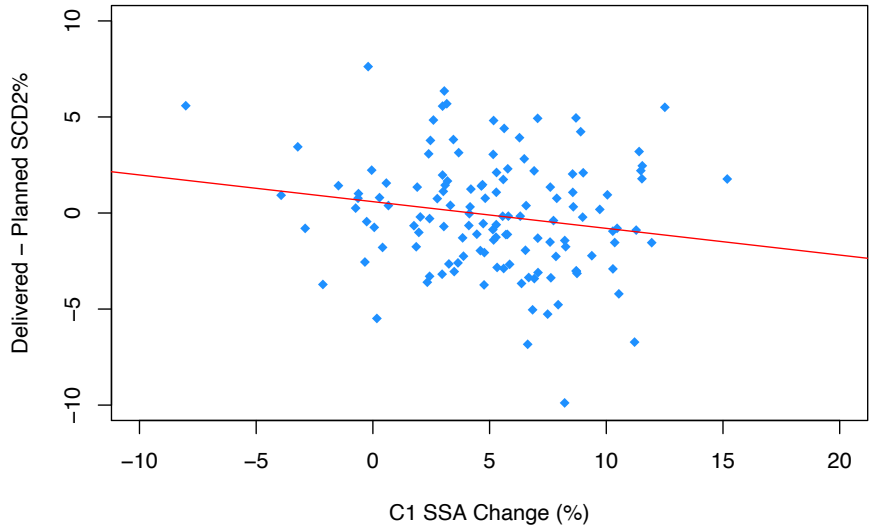
A – weight loss versus $\Delta SC_{D2\%}$



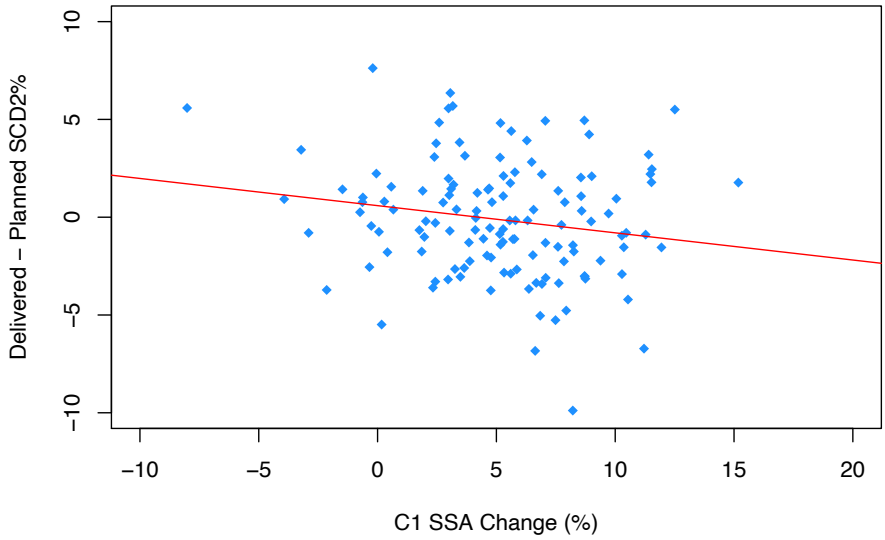
B – Separation at C1 versus $\Delta SC_{D2\%}$



C – Slice surface area change at C1 versus $\Delta SC_{D2\%}$



D – Separation at C4 (Thyroid notch) versus $\Delta SC_{D2\%}$



E – Slice surface area change at C4 (Thyroid notch) versus $\Delta SC_{D2\%}$

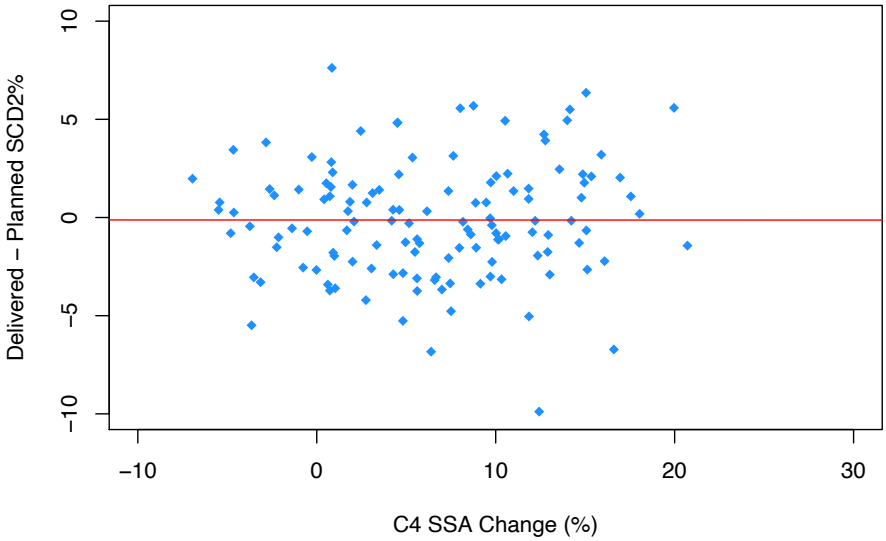


Figure 5.10 (A-E): Scatter plots of univariate relationships between weight loss & anatomical change and spinal cord dose.

5.5 Discussion

5.5.1 Spinal cord dose

This is not the first attempt to quantify differences between planned and delivered dose to the spinal cord in HNC patients [36, 37, 46-50]. However, the cohorts analysed in these studies were factors smaller (10-20 patients), than the sample of 133 used for this work. Furthermore, this is the first study to analyse spinal cord delivered dose at scale, whilst systematically measuring anatomical change and weight loss during treatment, and attempting to understand relationships between the two.

The magnitude of differences in SC dose seen in this work (0.9Gy, 2.5% of planned dose) is broadly similar to previously reported data (2.1% - 4.9%) [37, 46-50]. However, these studies found SC delivered dose to be systematically higher than planned, in contrast to data presented here. Other authors have not observed such clear systematic differences; Robar and colleagues report a sample mean of 0.3% (sd. 4.7%) for $\Delta D_{max\ 1\ cc}$, similar to our mean $\Delta SCD_{2\%}$ of -0.07Gy (-0.2% of mean $D_P D_{2\%}$, sd. 3.4%) [32], and a more recent study found a mean $SC\Delta D_{max}$ of 0.4Gy (in plans with a 5mm CTV to PTV margin) [50]. Differences between planned and delivered dose on the TomoTherapy system have also been reported. Using daily MVCT-IG on a cohort of 20 HNC patients undergoing BNI, Duma et al found that 51% of treatment fractions had a D_{max} higher than planned, and an overall difference of 1.2% from the plan [35]. The same authors found a 'systematic deviation' between planned and accumulated D_{max} in 75% of patients [36], similar to the 74.4% (99/133) of patients in this study who had a delivered $SCD_{2\%}$ > 1% different to planned $D_{2\%}$.

Nonetheless, the discrepancy between data presented here, and studies in which delivered SC dose is systematically higher, merits further discussion. One possible explanation is the frequency of imaging for dose accumulation. Some researchers have accumulated dose from scheduled (kVCT) rescans, and interpolated dose between timepoints (Castadot - 4 scans, Ahn – 3 scans, Bhide – 4 scans, Cheng – 2 scans) [37, 46, 47, 49], whilst others use weekly CBCT [48]. Interestingly, all these studies reported systematically higher delivered dose. In contrast, authors using daily IG images to accumulate dose saw smaller systematic differences (Duma et al (MVCT), D_A 0.16Gy higher; van Kranen et al (CBCT), D_A 0.4Gy) [36, 50].

PRV margins may also be relevant, and reporting on their use is inconsistent. Graff and colleagues did not find greater dose differences for a 4mm PRV than for the SC itself [51]. However, Castadot et al found that the difference in SC-PRV (4mm margin) Dmax (1.9Gy) was more than twice that seen for the cord [46], lending credence to the notion that PRV driven

optimization results in steep dose gradients away from the cord itself, and a more homogeneous 'dose-island' within. Thus anatomical change and setup error may result in significant differences to PRV dose, without substantial changes to cord dose itself. The data presented here support this logic; delivered SC dose was systematically higher than planned in patients with a steep dose gradient in the vicinity of the cord itself.

Image guidance policy may also be important. In the 'No IG' simulation, mean (direction agnostic) $\Delta SCD_{2\%}$ was found to be double the calculated values where daily IG was used (1.8Gy vs 0.9Gy). This supports the findings of previous studies, where daily IG use is associated with smaller dose differences [36, 50], and where a direct relationship between frequency of IG, and the magnitude of dose difference is shown [35]. In line with these data, a plausible explanation for the small dose differences seen in this cohort is the policy of DIPC that is used by the H&N unit at Addenbrooke's hospital.

5.5.2 Anatomical change and predictors of dose differences

The data provide no evidence to suggest that patients undergoing bilateral neck treatment would be more likely to see higher delivered SC doses. Furthermore, the results show no effect of disease T-stage on $\Delta SCD_{2\%}$, in line with previous work [38]. A possible relationship between more advanced nodal disease and higher SC dose is suggested, although statistical significance was not reached. Interestingly, N-stage is an important parameter in models that predict for the need for ART [52].

The observed mean weight loss of 7.9% is similar to previously published figures (5 to 11.3%) [29, 30, 47, 48, 53, 54]. Crucially, no relationship was seen between weight loss, and higher than planned SC doses, a point on which the literature lacks consensus. The general notion that weight loss leads to significant dosimetric changes is commonly held [34, 55], and one study has shown a link between weight loss and changes in SC dose [56]. Others have not [38, 48], a finding replicated here. This study is substantially larger than any which has previously addressed this question, and helps to clarify this point.

Patients undergoing radical RT for HNC may undergo shape change that is independent of weight loss, and studies have shown that reducing neck diameter is common during treatment [38, 54, 57]. In-silico models have suggested that reducing neck diameter may lead to higher than planned dose to the SC and brainstem [58]. Capelle et al found a significant correlation between reducing LND at the TN and $\Delta SCD_{2\%}$, although no relationship for reduction at C1 [38]. Ahn and colleagues found a significant correlation (R = 0.3) between reduction at the level of the 'mandibular joint' and increased SC dose, a surprising result given that this

structure is superior to the foramen magnum in most patients (in the axial plane) [37, 38]. In this work, significant reductions in both lateral separation and axial surface area at the level of both the C1 vertebra, and the Thyroid Notch were observed. The shape change data presented here are similar in magnitude to those previously reported [37, 38, 57], but no relationships between these changes and a systematic increase in cord dose were seen.

I suggest three possible explanations for this observation. Firstly the concept suggested by Graff and colleagues [51], that the spinal cord may be preserved from significant dosimetric change due to its central location, and the use of a PRV margin. This leads to the second point, that dosimetric differences are likely to be random, with minimal impact from systematic error [32]. Finally, perhaps most importantly, and based on our simulation of D_A in the absence of IG and the logic of Duma et al [35], the fact that these patients underwent DIPC as part of their treatment protocol may be crucial to the small dose differences observed.

5.5.3 Problems and limitations of dose accumulation

The data presented in our published work on automated contouring of the spinal cord with DIR [42], and in Chapter 4, illustrate that whilst there is uncertainty associated with the geometrical transposition of segmentations from one image to another, these uncertainties can at least be understood, and quantified [59]. However, there are also uncertainties associated with the subsequent processes of dose deformation and accumulation. The uncertainties are sufficient that there is debate as to whether or not it is reasonable to do it at all [60].

Opponents make the valid point that, as yet, there is no way to directly calibrate dose accumulation calculations against measurement [60]. Others argue that dose calculations are based on equations that divide energy by mass, and in many clinical scenarios, mass may be gained or lost. Whilst this is unlikely to be the case for the spinal cord it is highly likely in the case of shrinking glands. Thus, where dose is warped based on intensity informed deformations of underlying anatomy, the principle of conservation of energy may be violated [61]. The counter-argument to this assertion is that where mass is lost between time points, the biological material that constituted this mass is also lost. Importantly, the energy that led to this process – presumably via a cell death mechanism – is also ‘lost’, meaning that where mass is lost in such a system, there is no need to conserve energy [62].

These arguments are theoretical and abstract, but research has also been done to develop methods for quantifying computational uncertainties [63, 64]. Using these approaches, Veiga and colleagues tested 4 different DIR algorithms and used each to warp the geometry of ‘dose

of the day' calculations back to the planning CT space [65]. They found that whilst geometrical matching was similar, the mean squared dose difference of deformed dose cubes was approximately 2%, and that observed discrepancies were greater in regions of poorer image quality, and higher dose gradients.

In summary, it is clear that this remains a controversial topic, and further research is needed in the field. Better, more anatomically realistic DIR algorithms are needed, as well as improved understanding of the underlying physical, biological and anatomical processes that drive change during radiotherapy [65]. However, the methodology described in this chapter represents the most sophisticated, robust, and evidence based approach currently available for computing delivered dose at scale, in order to test the fundamental underlying hypotheses concerning delivered dose.

5.6 Conclusions

This chapter describes what is, to the best of my knowledge, the biggest study of delivered dose to the spinal cord in patients undergoing radical radiotherapy for HNC. Observed differences between planned, and delivered maximum dose to the spinal cord were small.

In designing this study, I hypothesised that patients who experienced greater weight loss during treatment, and those who underwent greater anatomical change, would be more likely to receive a delivered dose to the spinal cord that was higher than anticipated at planning. The data presented in this chapter clearly show that this is not the case. Furthermore, there was no evidence to suggest that patients receiving bilateral neck irradiation are more likely to have higher delivered cord dose than those receiving unilateral treatment only. Higher T-stage did not predict higher delivered cord dose, but whilst it did not reach statistical significance the data suggest that patients with more advanced nodal disease maybe at slightly greater risk of delivered dose to the cord being higher than planned. The only metric that clearly predicted for higher delivered dose was a steeper dose gradient in the region of the spinal cord itself.

All the patients whose data was analysed in this work were treated with daily IG with positional correction, and a zero-action level policy. By simulating the effect of no IG, it has been shown that this policy is a key reason for the small dose differences observed. Thus, in patients who undergo radical radiotherapy for HNC using this technique, weight loss and shape change during treatment do not mandate replanning for spinal cord safety.

5.7 References:

[1] Noble DJ, Yeap PL, Seah SYK, Harrison K, Shelley LEA, Romanchikova M, et al. Anatomical change during radiotherapy for head and neck cancer, and its effect on delivered dose to the spinal cord. Radiother Oncol. 2019;130:32-8.
[2] Schultheiss TE, Kun LE, Ang KK, Stephens LC. Radiation response of the central nervous system. Int J Radiat Oncol Biol Phys. 1995;31:1093-112.
[3] Sailer SL, Sherouse GW, Chaney EL, Rosenman JG, Tepper JE. A comparison of postoperative techniques for carcinomas of the larynx and hypopharynx using 3-D dose distributions. Int J Radiat Oncol Biol Phys. 1991;21:767-77.
[4] Miah AB, Bhide SA, Guerrero-Urbano MT, Clark C, Bidmead AM, St Rose S, et al. Dose-escalated intensity-modulated radiotherapy is feasible and may improve locoregional control and laryngeal preservation in laryngo-hypopharyngeal cancers. Int J Radiat Oncol Biol Phys. 2012;82:539-47.
[5] Thomson D, Yang H, Baines H, Miles E, Bolton S, West C, 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;26:344-7.
[6] Meade S, Gaunt P, Hartley A, Robinson M, Harrop V, Cashmore J, et al. Feasibility of Dose-escalated Hypofractionated Chemoradiation in Human Papilloma Virus-negative or Smoking-associated Oropharyngeal Cancer. Clin Oncol (R Coll Radiol). 2018;30:366-74.
[7] Beh SC, Greenberg BM, Frohman T, Frohman EM. Transverse myelitis. Neurol Clin. 2013;31:79-138.
[8] Abbatucci JS, Delozier T, Quint R, Roussel A, Brune D. Radiation myelopathy of the cervical spinal cord: time, dose and volume factors. Int J Radiat Oncol Biol Phys. 1978;4:239-48.
[9] Ang KK, van der Kogel AJ, van der Schueren E. The effect of small radiation doses on the rat spinal cord: the concept of partial tolerance. Int J Radiat Oncol Biol Phys. 1983;9:1487-91.
[10] Ang KK, Price RE, Stephens LC, Jiang GL, Feng Y, Schultheiss TE, et al. The tolerance of primate spinal cord to re-irradiation. Int J Radiat Oncol Biol Phys. 1993;25:459-64.
[11] Bijl HP, van Luijk P, Coppes RP, Schippers JM, Konings AW, van der Kogel AJ. Unexpected changes of rat cervical spinal cord tolerance caused by inhomogeneous dose distributions. Int J Radiat Oncol Biol Phys. 2003;57:274-81.
[12] Bijl HP, van Luijk P, Coppes RP, Schippers JM, Konings AW, van Der Kogel AJ. Regional differences in radiosensitivity across the rat cervical spinal cord. Int J Radiat Oncol Biol Phys. 2005;61:543-51.
[13] Youssef B, Shank J, Reddy JP, Pinnix CC, Farha G, Akhtari M, et al. Incidence and predictors of Lhermitte's sign among patients receiving mediastinal radiation for lymphoma. Radiat Oncol. 2015;10:206.
[14] Brownlee WJ, Hardy TA, Fazekas F, Miller DH. Diagnosis of multiple sclerosis: progress and challenges. Lancet. 2017;389:1336-46.
[15] St Clair WH, Arnold SM, Sloan AE, Regine WF. Spinal cord and peripheral nerve injury: current management and investigations. Semin Radiat Oncol. 2003;13:322-32.
[16] Atkinson SL, Li YQ, Wong CS. Apoptosis and proliferation of oligodendrocyte progenitor cells in the irradiated rodent spinal cord. Int J Radiat Oncol Biol Phys. 2005;62:535-44.
[17] Pak D, Vineberg K, Feng F, Ten Haken RK, Eisbruch A. Lhermitte sign after chemo-IMRT of head-and-neck cancer: incidence, doses, and potential mechanisms. Int J Radiat Oncol Biol Phys. 2012;83:1528-33.
[18] Laidley HM, Noble DJ, Barnett GC, Forman JR, Bates AM, Benson RJ, et al. Identifying risk factors for L'Hermitte's sign after IMRT for head and neck cancer. Radiat Oncol. 2018;13:84.
[19] Fein DA, Marcus RB, Jr., Parsons JT, Mendenhall WM, Million RR. Lhermitte's sign: incidence and treatment variables influencing risk after irradiation of the cervical spinal cord. Int J Radiat Oncol Biol Phys. 1993;27:1029-33.
[20] Esik O, Csere T, Stefanits K, Lengyel Z, Safrany G, Vonoczky K, et al. A review on radiogenic Lhermitte's sign. Pathol Oncol Res. 2003;9:115-20.
[21] Leung WM, Tsang NM, Chang FT, Lo CJ. Lhermitte's sign among nasopharyngeal cancer patients after radiotherapy. Head Neck. 2005;27:187-94.
[22] Ko HC, Powers AR, Sheu RD, Kerns SL, Rosenstein BS, Krieger SC, et al. Lhermitte's Sign following VMAT-Based Head and Neck Radiation-Insights into Mechanism. PLoS One. 2015;10:e0139448.
[23] Atkins HL, Tretter P. Time-dose considerations in radiation myelopathy. Acta Radiol Ther Phys Biol. 1966;5:79-94.

[24] McCunniff AJ, Liang MJ. Radiation tolerance of the cervical spinal cord. Int J Radiat Oncol Biol Phys. 1989;16:675-8.
[25] Marcus RB, Jr., Million RR. The incidence of myelitis after irradiation of the cervical spinal cord. Int J Radiat Oncol Biol Phys. 1990;19:3-8.
[26] Jeremic B, Djuric L, Mijatovic L. Incidence of radiation myelitis of the cervical spinal cord at doses of 5500 cGy or greater. Cancer. 1991;68:2138-41.
[27] Schultheiss TE. The radiation dose-response of the human spinal cord. Int J Radiat Oncol Biol Phys. 2008;71:1455-9.
[28] Kirkpatrick JP, Van-Der-Kogel AJ, Schultheiss TE. Radiation Dose-Volume Effects in the Spinal Cord. Int J Radiation Oncology Biol Phys. 2010;76:S42-9.
[29] Barker JL, Jr., Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys. 2004;59:960-70.
[30] Ottosson S, Zackrisson B, Kjellen E, Nilsson P, Laurell G. Weight loss in patients with head and neck cancer during and after conventional and accelerated radiotherapy. Acta Oncol. 2013;52:711-8.
[31] Hansen EK, Bucci MK, Quivey JM, Weinberg V, Xia P. Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2006;64:355-62.
[32] Robar JL, Day A, Clancey J, Kelly R, Yewondwossen M, Hollenhorst H, et al. Spatial and dosimetric variability of organs at risk in head-and-neck intensity-modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2007;68:1121-30.
[33] Han C, Chen YJ, Liu A, Schultheiss TE, Wong JY. Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy. Int J Radiat Oncol Biol Phys. 2008;70:1256-62.
[34] Brouwer CL, Steenbakkers RJ, Langendijk JA, Sijtsema NM. Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help? Radiother Oncol. 2015;115:285-94.
[35] Duma MN, Kampfer S, Schuster T, Aswathanarayana N, Fromm LS, Molls M, et al. Do we need daily image-guided radiotherapy by megavoltage computed tomography in head and neck helical tomotherapy? The actual delivered dose to the spinal cord. Int J Radiat Oncol Biol Phys. 2012;84:283-8.
[36] Duma MN, Schuster T, Aswathanarayana N, Fromm LS, Molls M, Geinitz H, et al. Localization and quantification of the delivered dose to the spinal cord. Predicting actual delivered dose during daily MVCT image-guided tomotherapy. Strahlenther Onkol. 2013;189:1026-31.
[37] Ahn PH, Chen CC, Ahn AI, Hong L, Scripes PG, Shen J, et al. Adaptive planning in intensity-modulated radiation therapy for head and neck cancers: single-institution experience and clinical implications. Int J Radiat Oncol Biol Phys. 2011;80:677-85.
[38] Capelle L, Mackenzie M, Field C, Parliament M, Ghosh S, Scrimger R. Adaptive radiotherapy using helical tomotherapy for head and neck cancer in definitive and postoperative settings: initial results. Clin Oncol (R Coll Radiol). 2012;24:208-15.
[39] Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, 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;117:83-90.
[40] Burnet NG, Adams EJ, Fairfoul J, Tudor GS, Hoole AC, Routsis DS, et al. Practical aspects of implementation of helical tomotherapy for intensity-modulated and image-guided radiotherapy. Clin Oncol (R Coll Radiol). 2010;22:294-312.
[41] Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging. 2010;29:196-205.
[42] Yeap PL, Noble DJ, Harrison K, Bates AM, Burnet NG, Jena R, et al. Automatic contour propagation using deformable image registration to determine delivered dose to spinal cord in head-and-neck cancer radiotherapy. Phys Med Biol. 2017;62:6062-73.
[43] Thomas SJ, Eyre KR, Tudor GS, Fairfoul J. Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube. Med Phys. 2012;39:160-7.
[44] Thomas SJ, Romanchikova M, Harrison K, Parker MA, Bates AM, Scaife JE, et al. Recalculation of dose for each fraction of treatment on TomoTherapy. Br J Radiol. 2016;89:20150770.
[45] ICRU., Measurements ICoRUa. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). ICRU Report 83. Journal of the ICRU. 2010;Volume 10:pages 1-106.

[46] Castadot P, Geets X, Lee JA, Gregoire V. Adaptive functional image-guided IMRT in pharyngo-laryngeal squamous cell carcinoma: is the gain in dose distribution worth the effort? Radiother Oncol. 2011;101:343-50.
[47] Bhide SA, Davies M, Burke K, McNair HA, Hansen V, Barbachano Y, et al. Weekly volume and dosimetric changes during chemoradiotherapy with intensity-modulated radiation therapy for head and neck cancer: a prospective observational study. Int J Radiat Oncol Biol Phys. 2010;76:1360-8.
[48] Ho KF, Marchant T, Moore C, Webster G, Rowbottom C, Penington H, et al. Monitoring dosimetric impact of weight loss with kilovoltage (kV) cone beam CT (CBCT) during parotid-sparing IMRT and concurrent chemotherapy. Int J Radiat Oncol Biol Phys. 2012;82:e375-82.
[49] Cheng HC, Wu VW, Ngan RK, Tang KW, Chan CC, Wong KH, et al. A prospective study on volumetric and dosimetric changes during intensity-modulated radiotherapy for nasopharyngeal carcinoma patients. Radiother Oncol. 2012;104:317-23.
[50] van Kranen S, Hamming-Vrieze O, Wolf A, Damen E, van Herk M, Sonke JJ. Head and Neck Margin Reduction With Adaptive Radiation Therapy: Robustness of Treatment Plans Against Anatomy Changes. Int J Radiat Oncol Biol Phys. 2016;96:653-60.
[51] Graff P, Hu W, Yom SS, Pouliot J. Does IGRT ensure target dose coverage of head and neck IMRT patients? Radiother Oncol. 2012;104:83-90.
[52] Brown E, Owen R, Harden F, Mengersen K, Oestreich K, Houghton W, et al. Predicting the need for adaptive radiotherapy in head and neck cancer. Radiother Oncol. 2015;116:57-63.
[53] Langius JA, Doornaert P, Spreeuwenberg MD, Langendijk JA, Leemans CR, van Bokhorst-de van der Schueren MA. Radiotherapy on the neck nodes predicts severe weight loss in patients with early stage laryngeal cancer. Radiother Oncol. 2010;97:80-5.
[54] Mazzola R, Ricchetti F, Fiorentino A, Di Paola G, Fersino S, Giaj Levra N, et al. Cachexia induces head and neck changes in locally advanced oropharyngeal carcinoma during definitive cisplatin and image-guided volumetric-modulated arc radiation therapy. Eur J Clin Nutr. 2016;70:738-42.
[55] Brouwer CL, Steenbakkers RJ, van der Schaaf A, Sopacua CT, van Dijk LV, Kierkels RG, et al. Selection of head and neck cancer patients for adaptive radiotherapy to decrease xerostomia. Radiother Oncol. 2016;120:36-40.
[56] Wang X, Lu J, Xiong X, Zhu G, Ying H, He S, et al. Anatomic and dosimetric changes during the treatment course of intensity-modulated radiotherapy for locally advanced nasopharyngeal carcinoma. Med Dosim. 2010;35:151-7.
[57] Senkus-Konefka E, Naczk E, Borowska I, Badzio A, Jassem J. Changes in lateral dimensions of irradiated volume and their impact on the accuracy of dose delivery during radiotherapy for head and neck cancer. Radiother Oncol. 2006;79:304-9.
[58] Chen C, Fei Z, Chen L, Bai P, Lin X, Pan J. Will weight loss cause significant dosimetric changes of target volumes and organs at risk in nasopharyngeal carcinoma treated with intensity-modulated radiation therapy? Med Dosim. 2014;39:34-7.
[59] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. Med Phys. 2017;44:e43-e76.
[60] Schultheiss TE, Tome WA, Orton CG. Point/counterpoint: it is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy. Med Phys. 2012;39:6531-3.
[61] Zhong H, Chetty IJ. Caution Must Be Exercised When Performing Deformable Dose Accumulation for Tumors Undergoing Mass Changes During Fractionated Radiation Therapy. Int J Radiat Oncol Biol Phys. 2017;97:182-3.
[62] Hugo GD, Dial C, Siebers JV. In Regard to Zhong and Chetty. Int J Radiat Oncol Biol Phys. 2017;99:1308-10.
[63] Salguero FJ, Saleh-Sayah NK, Yan C, Siebers JV. Estimation of three-dimensional intrinsic dosimetric uncertainties resulting from using deformable image registration for dose mapping. Med Phys. 2011;38:343-53.
[64] Murphy MJ, Salguero FJ, Siebers JV, Staub D, Vaman C. A method to estimate the effect of deformable image registration uncertainties on daily dose mapping. Med Phys. 2012;39:573-80.
[65] Veiga C, Lourenco AM, Mouinuddin S, van Herk M, Modat M, Ourselin S, et al. Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm. Med Phys. 2015;42:760-9.

Chapter 6 – Delivered dose to the swallowing OARs

6.1 Overview

Toxicity following radical radiotherapy for HNC remains a significant problem. Radiation dose to crucial OARs, including salivary glands, oral cavity, and swallowing structures predicts the risk of long-term side effects. The dose response curves that model these risks are steep, so small differences in dose may be clinically important. It is a widely thought that the weight loss and anatomical change seen in HNC patient during treatment may lead to differences between planned and delivered dose. Studies have attempted to quantify these differences, and to identify factors that predict them. However, many of these studies have been small, and have used sparse timepoint imaging data during treatment to estimate delivered dose. This chapter presents the methodology and results of a large-scale analysis to estimate delivered dose to swallowing OARs, to investigate previously suggested predictors of dose differences, and to try and find new ones.

6.1.1 My Role

All the hypotheses investigated in this chapter were mine. I did all the manual segmentations of OARs for the calculation of both planned and delivered dose as described in 6.3 Methods. I collated and analysed all the baseline, demographic, disease and treatment information presented in this chapter, and defined the final patient cohort for analysis. All the analysis of dose data, and relationships between hypothesised predictors and dose differences were done by me. The work presented in this chapter formed the basis of a number of conference abstracts, including the following published abstract:

Predictors of dose differences to swallowing OARs in patients undergoing radiotherapy for HNC. Noble DJ, Harrison K, Wilson M, Hoole A, Thomas SJ, Burnet NG, Jena R.

- *Poster presentation:* International Conference of Head & Neck Oncology, Barcelona, March 2019.
- *Published in:* Radiotherapy and Oncology 132:65-66.

6.1.2 Acknowledgements

Amy Bates (clinical trials radiographer) and her team recruited all the patients whose data is used in this chapter to the VoxTox study, and collected all the baseline, demographic, disease

and treatment information presented in this chapter. Dieticians and Speech & Language Therapists from the clinical H&N team recorded the patient weight data used in this work. As described in Chapter 2, computer codes for extraction of radiotherapy data from archive, anonymisation and transfer from the hospital to the Cavendish Laboratory were written by Marina Romanchikova (computer scientist) and executed by her, and Andrew Hoole (computer scientist). Work on the training and validation of the DIR software for automated segmentation of MVCT images, and dose accumulation, was done by very close collaboration between myself and Karl Harrison (Research Associate, Cavendish laboratory), with valuable input from Megan Wilson (Physics PhD student, UCL)

6.2 Introduction

Radical radiotherapy for HNC remains an effective but morbid treatment [1]. Despite advances in radiotherapy techniques, persistent side effects and reduced quality of life remain a significant problem [2]. Dry mouth (xerostomia) and swallowing dysfunction are amongst the most common, and most disruptive toxicities experienced by patients [3-7]. Radiation dose to parotid, submandibular, and minor salivary glands in the oral cavity predicts long-term side effects such as xerostomia, thick sticky saliva and taste disturbance [8-12], whilst dose to the pharyngeal constrictor muscles and supraglottic larynx determines the risk of long-term swallowing dysfunction [13-17]. Furthermore, the curves of NTCP models that attempt to quantify relationships between radiation dose to these OARs and toxicity endpoints are steep [8, 14], meaning that small increases in dose may have clinically important consequences.

As described in Chapters 1 and 5, HNC patients can undergo significant anatomical change during treatment. Changes include weight loss [18], shrinkage and shape change [19-21], involution and shrinkage of the tumour, and also changes in the size, shape and position of OARs themselves [22-25]. This can lead to differences between the dose of radiation that both the tumour and OARs receive over a course of treatment, and that which was anticipated at planning [26, 27]. Adaptive radiotherapy – ART – has been suggested as a way of mitigating these changes [2], by maintaining dose coverage in the presence of anatomical change, by escalating dose to the tumour, reducing dose to OARs, or by a combination of these objectives [28].

These ideas are based on research in which authors have attempted to quantify dose differences, and to find factors that can reliably predict these differences [20, 21, 29-37]. However, many of these studies looked at small cohorts of patients (20 or less), and have used

a limited number of images of patient anatomy during treatment in their attempts to estimate delivered dose [20, 21, 29-34]. Furthermore, whilst there are a number of studies that have assessed the parotid glands, there are far fewer data looking at dose differences to other OARs, including the submandibular glands, oral cavity, pharyngeal constrictor muscles and laryngeal substructures [36]. Partly because of this, ART workflows in current practice are heterogeneous, and clinicians lack definitive data on which to base ART protocols.

Therefore, the objectives of this chapter were as follows:

- To quantify dose differences to all the OARs involved in long term xerostomia and swallowing dysfunction – specifically both parotid and submandibular glands, the oral cavity, superior and middle pharyngeal constrictor muscles, and the supraglottic larynx.
- To do so in a systematic way, using daily imaging to track anatomical change, in as large a cohort as possible.
- To investigate previously published relationships between these dose differences, and factors that may influence them.
- To investigate other potential predictors of dose differences.
- To use these results to make suggestions as to which patients may benefit from ART for toxicity reduction.

6.3 Methods

6.3.1 Cohort selection and treatment details

All data used in this chapter were collected from patients recruited to the VoxTox study, as described in Chapter 2. Patients were treated according to standard departmental techniques and protocols as outlined in Chapters 2 and 5. Key objectives of this chapter were to calculate D_A for as a large cohort as possible, and to try identify patterns of dose differences based on parameters such as weight loss, primary site, use of concomitant systemic therapy, pre-RT surgery. Therefore, in contrast to the approach taken in Chapter 5, the inclusion criteria for this sub-study were broadened to include the full range of primary sites and treatment regimens shown in Table 2.4.

However, the methodology for calculating both planned and delivered dose involved manual segmentation of all OARs on the pCT, and was time-consuming. Therefore, the focus was on the consolidation cohort, as these patients also had toxicity data available for analysis, as described in Chapter 2. There were therefore 253 patients with a full structure set of OARs.

For 10 of these patients, it was not possible to compute D_A, for a combination of reasons. Seven of these patients started, but did not complete treatment due to inter-current illness or disease progression. For the remainder, there were issues relating to the IG scans, and deformable image registration workflow. These included; large numbers of missing IG scans, IG scans with truncated FoVs, and errors in the way information about post-IG couch shifts were encoded in DICOM header information. Therefore, 243 patients had both D_P and D_A available for analysis, and the demographic, disease, and treatment details of cohort are summarised in Table 6.1.

Characteristic	Number
Age	58.8 (10.6)
Gender	
Male	201 (82.7%)
Female	42 (17.3%)
Baseline weight (kg)	84.2 (17.6)
Histology & Primary site	
SCC	223 (91.8%)
Oropharynx	147 (60.5%)
Larynx	21 (8.6%)
Oral cavity	19 (7.8%)
Sinus	9 (3.7%)
Nasopharynx	8 (3.3%)
Hypopharynx	7 (2.9%)
Unknown primary	7 (2.9%)
Skin	5 (2.1%)
Salivary gland	16 (6.6%)
Adenocarcinoma	6 (2.5%)
Acinic cell	6 (2.5%)
Adenoid cystic	3 (1.2%)
Mucoepidermoid	1 (1.2%)
Others	4 (1.6%)
T-stage	
Tis/T1	71 (29.2%)
T2	83 (34.2%)
T3	35 (14.4%)
T4	54 (22.2%)
N-stage	
N0	72 (29.6%)
N1	31 (12.8%)
N2a	35 (14.4%)
N2b	80 (32.9%)

N2c	20 (8.2%)
N3	5 (2.1%)
Surgery	
Neck dissection	73 (30.0%)
Parotidectomy	14 (5.8%)
Laryngectomy	5 (2.1%)
Dose/Fractionation	
70/35	6 (2.5%)
68/34*	16 (6.6%)
65/30*	159 (65.4%)
60/30	53 (21.8%)
55/20	4 (1.6%)
50/20	5 (2.1%)
Neck Irradiation	
Unilateral	68 (28.0%)
Bilateral	169 (69.5%)
None	6 (2.5%)
Systemic therapy	
Cisplatin†	133 (54.7%)
Cetuximab†	17 (7.0%)
None	93 (38.3%)
* primary SCC's of the oro/hypopharynx, and larynx treated with 68Gy/34# prior to November 2011, 65Gy/30# thereafter. † dose: cisplatin – 40mg/m^2 weekly, cetuximab 400mg/m2 loading dose, 240mg/m^2 weekly thereafter.	

Table 6.1: patient characteristics – swallowing OAR delivered dose cohort (n = 243). For continuous variables, means and standard deviations are reported, absolute numbers and percentages for proportions.

6.3.2 Dose calculation

Patient imaging data, including all kVCT planning scans and MVCT IG images, structure set data, and treatment delivery data were extracted from vendor archive by Marina Romanchikova, as described in Chapter 2. Anonymised, downsampled kVCTs for all 253 patients were uploaded the research database of segmentation software (Prosoma 3.3, MEDCOM, Darmstadt, Germany). Structures were contoured by a single observer according to recognised guidelines [38, 39], with intra- and inter-observer variability data as shown in Chapter 2. Where structures had been surgically resected prior to RT planning, the same ROI descriptor with the suffix '_resected' (e.g. "Right SMG_resected") was used in the structure set to ensure confidence in data completeness rather than error by omission. Radiotherapy

planning imaging and planned dose data were used to define primary tumour laterality for each patient, so that 'right' and 'left' salivary gland structures could be encoded as 'ipsilateral' and 'contralateral' for analysis.

Completed segmentations were exported to the Cavendish laboratory using the data framework described in Chapter 2. B-spline based deformable image registration (Elastix) [40], and a ray-tracing dose calculation algorithm [41], were used to compute daily dose on IG-MVCTs, as described in Chapter 5. In a small number of cases, daily MVCT imaging datasets were missing. In such circumstances, interpolation was used. Specifically, per-voxel dose values from the previous fraction were doubled in the dose accumulation process. Where more than 50% of IG MVCTs were missing, patients were excluded from analysis. For some patients, the superior extent of the MVCT was too low, missing off the top few slices of the gland. This was a much greater problem for the discovery cohort than the consolidation. In these cases, missing MVCT data were substituted in from the pCT. The numbers of missing slices for both cohorts are shown in Table 6.2 (data for the full cohort of VoxTox patients with imaging data at the Cavendish lab – 319 patients). Where more than half the 'expected' numbers of slices were missing, the IG scan was discarded, and dose interpolation from the previous fraction was used. This happened with 8% of MVCT scans for patients in the discovery cohort, and did not occur for patients in the consolidation cohort.

No. of missing slices per scan	Discovery (2888 total)	Consolidation (6224 total)
0	544 (18.9%)	4313 (69.3%)
1	357 (12.4%)	1265 (20.3%)
2	483 (16.7%)	517 (8.3%)
3	420 (14.5%)	103 (1.7%)
4	504 (17.5%)	26 (0.4%)
5	362 (12.5%)	0
6	168 (5.8%)	0
7	50 (1.7%)	0

Table 6.2: Number of MVCT scans missing superior (cranial) slices with parotid gland tissue (data for all 319 patients with imaging data on Cavendish laboratory servers shown).

Daily delivered dose values were accumulated to a final summary D_ADVH. Previous work has examined relationships between a number of dose volume parameters to OARs, and their resultant toxicities [10, 15, 42]. However, the most widely used dose parameter in the literature is mean dose to a structure [8, 10, 13, 14]. Furthermore, co-linearity between dose metrics is

a well-known problem when examining dose-response relationships [43, 44]. Therefore, the decision was made to summarise both D_P and D_A as mean dose to each structure. Similar to the spinal cord work, dose differences were defined as ($D_{meanA} - D_{meanP}$), as higher delivered dose is more clinically important.

6.3.3 Dose difference predictors

Weight loss and anatomical change data were collected and assessed using the same methodology as described in Chapters 2 and 5. As the cohort used for this work was substantially larger than that in Chapter 5, weight loss data on more patients (206) were available for analysis. The same anatomical change data (LND and SSA at the level of the C1 vertebra and thyroid notch) collected for the work on spinal cord dose difference predictors were used (data on 132 patients were available). Information on disease primary site, staging, concomitant therapy and any pre-RT surgery were all collected during the baseline VoxTox assessment, as described in Chapters 2 and 5. Neck irradiation strategy was recorded by checking RT booking paperwork and plan information in Mosaiq.

Previous work has suggested a link between changes in parotids during treatment and dose differences to the gland [20, 29]. In these studies, patient sample sizes were 23 and 10 respectively, and volume was calculated by manual ROI segmentation. That approach was not practical for the sample size in this work, and an automated approach was needed. As described in Chapter 4, the matrix of Jacobian determinant values generated following a deformable registration provides information about per-voxel volume differences between the two images in the registration. The absolute value of a Jacobian determinant for any given voxel is degenerate, and a number of different mappings can exist. However, taken as a mean value over thousands of voxels representing the parotid glands, we hypothesised that this would be a reasonable surrogate of volume change in the gland. Therefore, for each patient, the mean Jacobian determinant value of all voxels regarded as belonging to the parotid glands on the final treatment fraction MVCT image was recorded, and correlation statistics between mean Jacobian determinant to each gland, and dose differences, were calculated.

6.3.4 Statistical analysis

Descriptive statistics (mean, standard deviation, 95% confidence intervals, median and IQR) were used to assess D_P, D_A, and dose difference data. An unsupervised dendrogram and heatmap approach were used to visualise patterns in dose differences between different OARs. Relationships between binary and ordinal variables (primary site, staging, concomitant

therapy and surgery) and dose differences were assessed by student t-test and ANOVA, and visualised as box-plots. Parametric data distributions were checked by histogram and manual inspection. Hypothesised relationships between continuous variables (weight loss, anatomical change, final fraction parotid ROI mean Jacobian) and OAR dose differences were investigated with univariate linear regression models, and Pearson's product-moment correlation coefficient (r), and visualised with scatter plots. Data distributions were plotted as histograms to confirm parametric distributions [36]. All statistical analysis in this work was done in R statistical software (R Notebook, R version 3.4.0). P-values <0.05 were considered statistically significant, and Bonferroni corrections to account for multiple testing were used where appropriate.

6.4 Results

6.4.1 Dose differences to swallowing OARs

Table 6.3 shows mean (with standard deviation) and median (with IQR) cohort values of both planned and delivered mean dose to each of the 8 swallowing OARs.

	Planned dose (Gy)		**Delivered dose (Gy)**	
	Mean (s.d.)	**Median (IQR)**	**Mean (s.d.)**	**Median (IQR)**
IPG (n = 231)	40.0 (11.6)	42.3 (35.4 – 47.2)	41.6 (11.9)	43.6 (37.1 – 49.5)
CPG (n = 243)	25.4 (12.6)	28.6 (13.4 – 35.0)	26.3 (13.2)	29.0 (14.0 – 36.7)
ISMG (n = 181)	58.5 (13.2)	62.8 (59.3 – 63.9)	59.7 (13.4)	64.2 (60.1 – 65.7)
CSMG (n = 237)	43.8 (19.1)	53.4 (30.2 – 56.6)	45.0 (19.5)	54.6 (30.2 – 57.7)
SPC (n = 243)	53.5 (12.8)	59.1 (50.4 – 61.7)	54.3 (13.0)	59.8 (50.9 – 62.6)
MPC (n = 243)	53.5 (13.7)	58.4 (51.6 – 60.7)	54.1 (14.0)	58.9 (52.1 – 62.0)
SGL (n = 233)	52.7 (14.7)	56.7 (51.6 – 61.3)	53.6 (15.0)	58.0 (51.8 – 62.0)
OC (n = 243)	48.9 (12.8)	53.8 (44.6 – 57.3)	49.3 (13.0)	54.4 (45.6 – 57.8)

Table 6.3: Average (mean and median) mean doses in Gy to each swallowing OAR. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. The number of patients with valid dose data for analysis is shown in parentheses.

These data are shown as a single summary boxplot in Figure 6.1, and histograms showing the distribution of dose differences for each structure are shown in Figure 6.2.

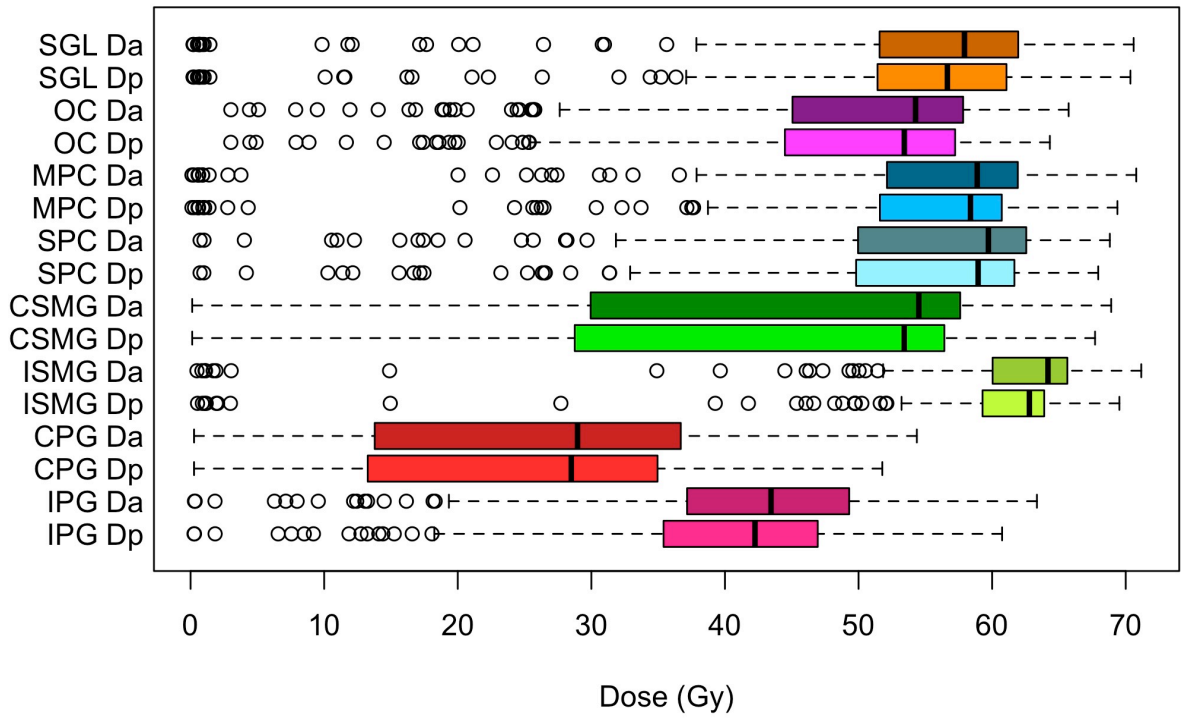


Figure 6.1: Boxplot of (mean) planned (D_P) and delivered (D_A) dose values for each swallowing OAR. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity.

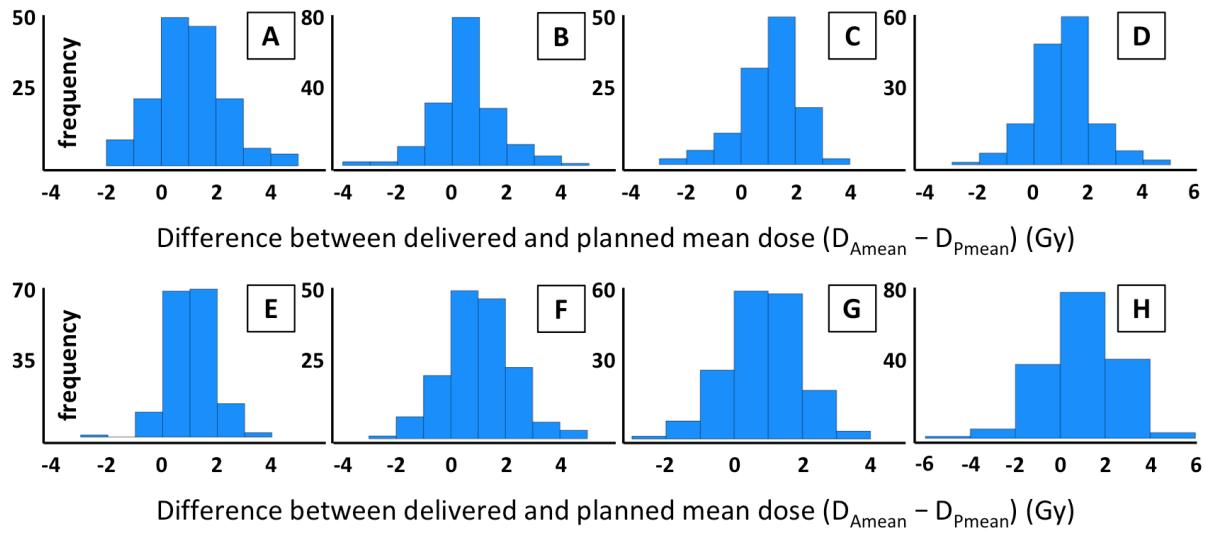


Figure 6.2 (A-H): Histograms of differences between planned (D_P) and delivered (D_A) dose for each swallowing OAR. A = Ipsilateral parotid gland, B = Contralateral parotid gland, C = Ipsilateral submandibular gland, D = Contralateral submandibular gland, E = Superior pharyngeal constrictor muscle, F = Middle pharyngeal constrictor muscle, G = Supraglottic larynx, H = Oral cavity.

Mean differences between planned and delivered dose, with 95% confidence intervals, were as follows: IPG = 1.56Gy (1.37 – 1.74), CPG = 0.94Gy (0.77 – 1.11), ISMG = 1.24Gy (1.11 – 1.36), CSMG = 1.17Gy (1.05 – 1.29), SPC = 0.81Gy (0.71 – 0.91), MPC = 0.68Gy (0.55 – 0.82), SGL = 0.98Gy, OC = 0.44Gy (0.30 – 0.57).

Data distributions of per-patient delivered minus planned dose, for each OAR, are shown in Figure 6.2. It can be seen that for some OARs (e.g. CPG and OC), distributions are broader than was seen for others (e.g. SPC). Thus, whilst mean dose differences across the full sample are relatively small for all OARs, there is much greater range in the number of individual patients with potentially clinically significant dose deltas to different OARs. Numbers of patients with a dose difference of 2Gy or more by OAR were as follows: IPG = 78 (32.1%), CPG = 46 (18.9%), ISMG = 31 (12.8%), CSMG = 39 (16%), SPC = 8 (3.3%), MPC = 25 (10.3%), SGL = 14 (5.8%), OC = 54 (22.2%). A threshold difference of 2Gy was chosen for 2 reasons, firstly because it approximates to the delivery of an extra treatment fraction to a high-risk (60Gy) CTV, and secondly on the basis of similar work in the literature [45].

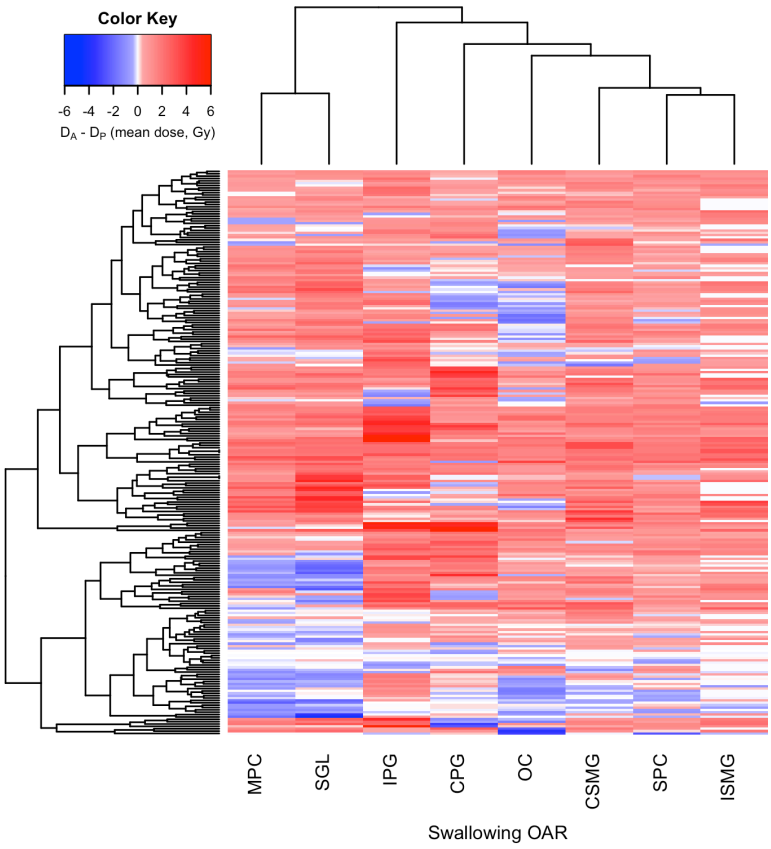


Figure 6.3: Unsupervised dendrogram and heatmap of dose differences to all swallowing OARs. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Red = D_A higher, Blue = D_P higher.

A relevant clinical question is whether or not potentially clinically significant dose differences occur in the same, or in different patients. Therefore, an unsupervised dendrogram and heatmap analysis, showing OARs clustered by dose differences was done, and is shown in Figure 6.3. This shows that some patients had delivered doses that were higher than planned to all OARs, whilst other patients saw predominantly lower doses. No patients had a dose difference of 2Gy or more to all 8 OARs. One patient saw a dose differences of 2Gy or more to 6 OARs, only the IPG (1.40Gy) and SGL (1.22Gy) being under this threshold. Nine patients (3.7%) had a dose difference of 2Gy or more to 5 OARs. Interestingly, these 9 cases shared a number of features. All 9 had SCCs of the oropharynx, all of them had an N-stage of 2b or greater, and received radiotherapy to both sides of the neck. Eight of 9 received concomitant cisplatin, whilst only 1 underwent neck dissection prior to radiation.

6.4.2 Anatomical change and weight loss vs. dose differences

Data on anatomical change at the level of the C1 vertebra and thyroid notch are the same as those presented in Chapter 5. Weight loss data was available for 206 patients in this cohort. Mean weight loss (with standard errors) was 5.6(±0.34)kg, 6.8(±0.36)%. Results of univariate linear regression, and correlation statistics, between absolute dose differences (in Gy) and relative weight loss (in %) are given in Table 6.4. With Bonferroni correction, a p-value <0.0063 is regarded as statistically significant. Therefore, weak but significant correlations were found between relative weight loss, and dose differences to the pharyngeal constrictor muscles, contralateral SMG, and supraglottic larynx.

OAR	r	95%CI (r)	p-value
IPG	0.16	0.01 – 0.29	0.033
CPG	0.068	-0.07 – 0.21	0.33
ISMG	0.21	0.05 – 0.35	0.011
CSMG	0.25	0.11 – 0.37	<0.001
SPC	0.23	0.10 – 0.36	<0.001
MPC	0.31	0.17 – 0.43	<0.001
SGL	0.23	0.09 – 0.36	0.001
OC	0.02	-0.12 – 0.16	0.77

Table 6.4: Univariate relationships between relative weight loss (%) and absolute dose differences (Gy) to all swallowing OARs. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Bonferroni correction: p-value <0.0063 considered statistically significant.

Results of univariate analysis between anatomical change metrics at the level of the CI vertebra and thyroid notch are shown in Table 6.5 and 6.6 respectively.

	C1 vertebra ΔLND (%)			C1 vertebra ΔSSA (%)		
	r	95%CI (r)	p-value	r	95%CI (r)	p-value
IPG	0.19	0.01 – 0.36	0.035	0.16	-0.02 – 0.33	0.080
CPG	0.16	-0.02 – 0.32	0.077	0.15	-0.02 – 0.32	0.086
ISMG	0.28	0.08 – 0.46	0.0078	0.23	0.02 – 0.41	0.033
CSMG	0.37	0.21 – 0.51	<0.001	0.28	0.11 – 0.43	0.002
SPC	0.29	0.12 – 0.44	<0.001	0.15	-0.02 – 0.32	0.087
MPC	0.43	0.28 – 0.56	<0.001	0.27	0.11 – 0.43	0.002
SGL	0.33	0.17 – 0.48	<0.001	0.21	0.04 – 0.37	0.017
OC	0.06	-0.11 – 0.24	0.47	0.01	-0.16 – 0.19	0.88

Table 6.5: Univariate relationships between relative anatomical change at the level of the C1 vertebra (lateral neck dimension – LND, slice surface area – SSA) (%) and absolute dose differences (Gy) to all swallowing OARs.

	Thyroid notch ΔLND (%)			Thyroid notch ΔSSA (%)		
	r	95%CI (r)	p-value	r	95%CI (r)	p-value
IPG	0.06	-0.12 – 0.23	0.51	0.11	-0.06 – 0.29	0.21
CPG	0.30	0.13 – 0.45	<0.001	0.14	-0.03 – 0.31	0.11
ISMG	0.18	-0.03 – 0.37	0.10	0.22	0.02 – 0.41	0.033
CSMG	0.28	0.11 – 0.44	0.001	0.27	0.10 – 0.43	0.002
SPC	0.17	0 – 0.33	0.05	0.11	-0.07 – 0.28	0.23
MPC	0.37	0.20 – 0.51	<0.001	0.40	0.24 – 0.54	<0.001
SGL	0.27	0.10 – 0.43	0.002	0.37	0.21 – 0.52	<0.001
OC	-0.05	-0.22 – 0.12	0.56	0	-0.17 – 0.17	0.99

Table 6.6: Univariate relationships between relative anatomical change at the level of the Thyroid notch (lateral neck dimension – LND, slice surface area – SSA) (%) and absolute dose differences (Gy) to all swallowing OARs. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Bonferroni correction: p-value <0.0063 considered statistically significant.

6.4.3 Primary disease site vs. dose differences

Separate ANOVA tests were undertaken, to investigate hypothesised relationships between primary disease site, and dose differences to each organ. Due to low numbers in some disease primary sites, some were grouped together to maintain statistical power. Groupings were made on the basis of anatomical proximity, and similar RT planning and delivery techniques. The disease primary site groups (with patient numbers) were as follows: carcinoma unknown primary (CUP) 7, hypopharynx & larynx 27, nasopharynx 8, oral cavity 19, oropharynx 147, salivary gland & sinus & skin 35. Data are shown in Figure 6.4, and results of ANOVA analyses for each OAR are shown in Table 6.7.

OAR	DF	Sum Sq	F value	p-value
IPG	5	45.2	5.0	<0.001
CPG	5	58.9	7.4	<0.001
ISMG	5	18.7	4.4	<0.001
CSMG	5	35.6	8.6	<0.001
SPC	5	17.5	7.0	<0.001
MPC	5	36.4	4.5	<0.001
SGL	5	33.8	2.4	0.036
OC	5	5.11	0.9	0.47

Table 6.7: Results of ANOVA tests of differences between planned and delivered dose by primary disease site. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Bonferroni correction: p-value <0.0063 considered statistically significant.

Although the p-values for all ANOVA tests apart from that for the oral cavity suggest significant differences between groups, visual inspection of the data in Figure 6.4 suggest no clear pattern for most OARs. The exceptions are both parotid glands where dose differences are conspicuously higher in patients with nasopharyngeal primary disease, and this observation was born out on Tukey sub-group analysis.

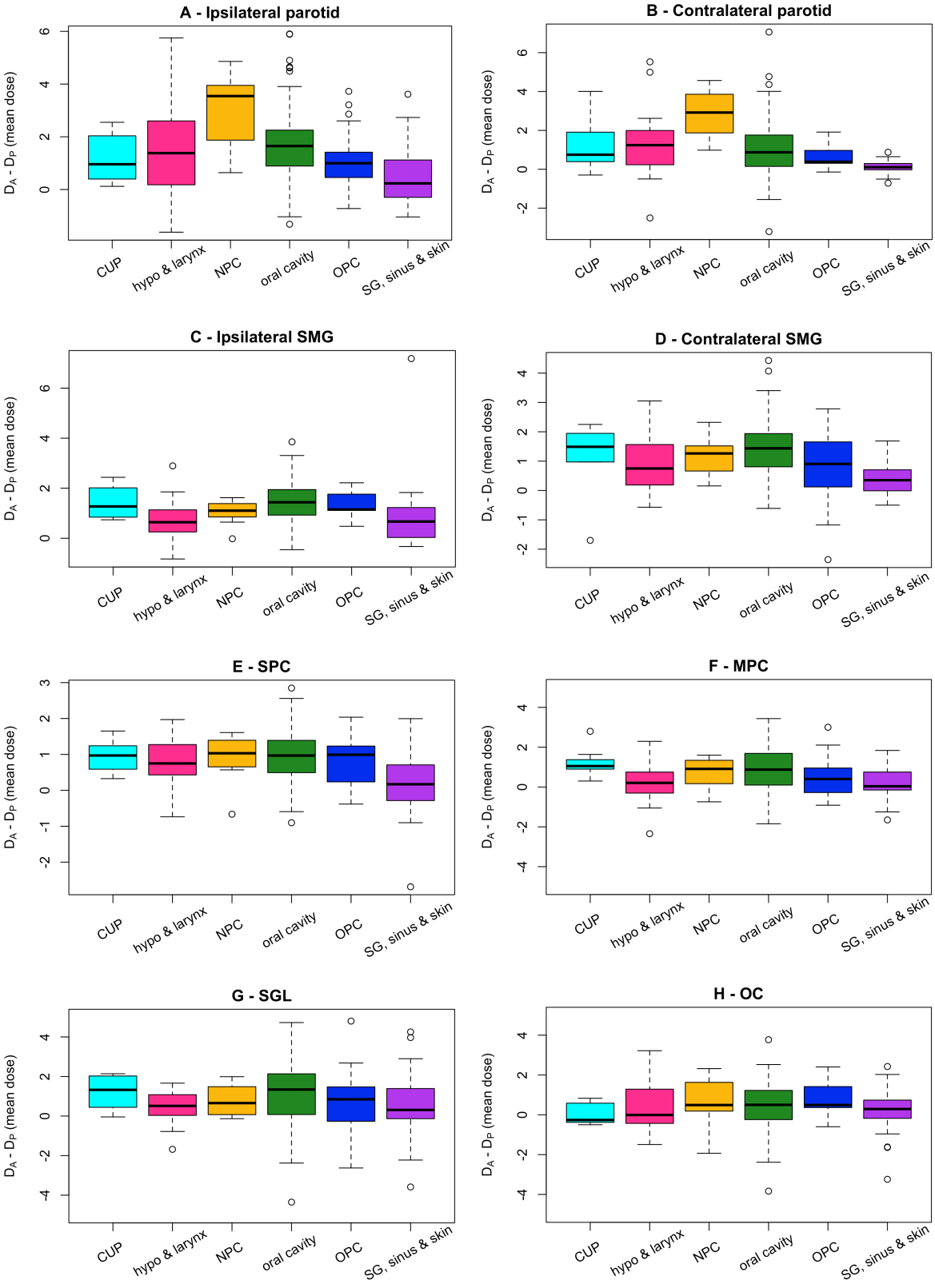


Figure 6.4 (A-H): Boxplots of swallowing OAR dose differences by primary disease. CUP = carcinoma unknown primary, NPC = nasopharynx, OPC = oropharynx. A - Ipsilateral parotid gland, B - Contralateral parotid gland, C - Ipsilateral submandibular gland, D - Contralateral submandibular gland, E - Superior pharyngeal constrictor muscle, F - Middle pharyngeal constrictor muscle, G - Supraglottic larynx, H - Oral cavity.

6.4.4 T&N stage vs. dose differences

Potential relationships between T and N staging and dose differences to OARs were also investigated. With Bonferroni corrections for 8 student t-tests, statistical significance was defined by a p-value <0.0063. Data for both variables were binarised (T0-2 vs. T3-4 & N0-1 vs. N2-3) in line with previous methodologies [37, 46]. Table 6.8 shows mean values for D_{meanA} – D_{meanP} (Gy), and the p-value and 95% CIs of the t-test assessing their significance, for T-stage. Table 6.9 shows data in the same format for N-stage. The same data are shown as boxplots in Figures 6.5 and 6.6. For T-stage, the only OAR for which higher T-stage predicted higher D_A with statistical significance (given multiple testing) was the supraglottic larynx. Differences for the oral cavity narrowly missed the pre-defined alpha value of 0.0063, but the relationship is certainly suggestive. For N-stage, highly significant relationships between more advanced nodal disease in the neck, and higher delivered dose, were seen for both submandibular glands, both superior and middle pharyngeal constrictor muscles, and the supraglottic larynx. The relationship was borderline for the contralateral parotid (p=0.009), and not significant for either the ipsilateral parotid, or oral cavity.

	T-stage 0-2	**T-stage 3-4**	**p-value**	**95% CI**
	Mean D_A-D_P (Gy)	**Mean D_A-D_P (Gy)**		
IPG	1.48	1.69	0.28	-0.59 to 0.17
CPG	0.93	0.96	0.84	-0.39 to 0.32
ISMG	1.19	1.30	0.42	-0.36 to 0.15
CSMG	1.14	1.23	0.50	-0.34 to 0.17
SPC	0.81	0.81	0.96	-0.20 to 0.19
MPC	0.56	0.87	0.028	-0.58 to -0.03
SGL	0.75	1.38	0.003	-1.02 to -0.21
OC	0.30	0.68	0.007	-0.66 to -0.11

Table 6.8: Differences between planned and delivered mean dose to all swallowing OARs, with data split by binarised T-stage (T0-2 vs T3-4). IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Bonferroni correction: p-value <0.0063 considered statistically significant.

	N-stage 0-1	N-stage 2+	p-value	95% CI
	Mean D_A-D_P (Gy)	Mean D_A-D_P (Gy)		
IPG	1.39	1.68	0.12	-0.66 to 0.08
CPG	0.68	1.13	0.009	-0.80 to -0.11
ISMG	0.84	1.61	<0.001	-0.99 to -0.53
CSMG	0.75	1.49	<0.001	-0.97 to -0.51
SPC	0.53	1.02	<0.001	-0.67 to -0.30
MPC	0.25	0.99	<0.001	-0.99 to -0.49
SGL	0.48	1.31	<0.001	-1.21 to -0.44
OC	0.30	0.54	0.074	-0.52 to 0.02

Table 6.9: Differences between planned and delivered mean dose to all swallowing OARs, with data split by binarised N-stage (N0-1 vs N2+). IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Bonferroni correction: p-value <0.0063 considered statistically significant

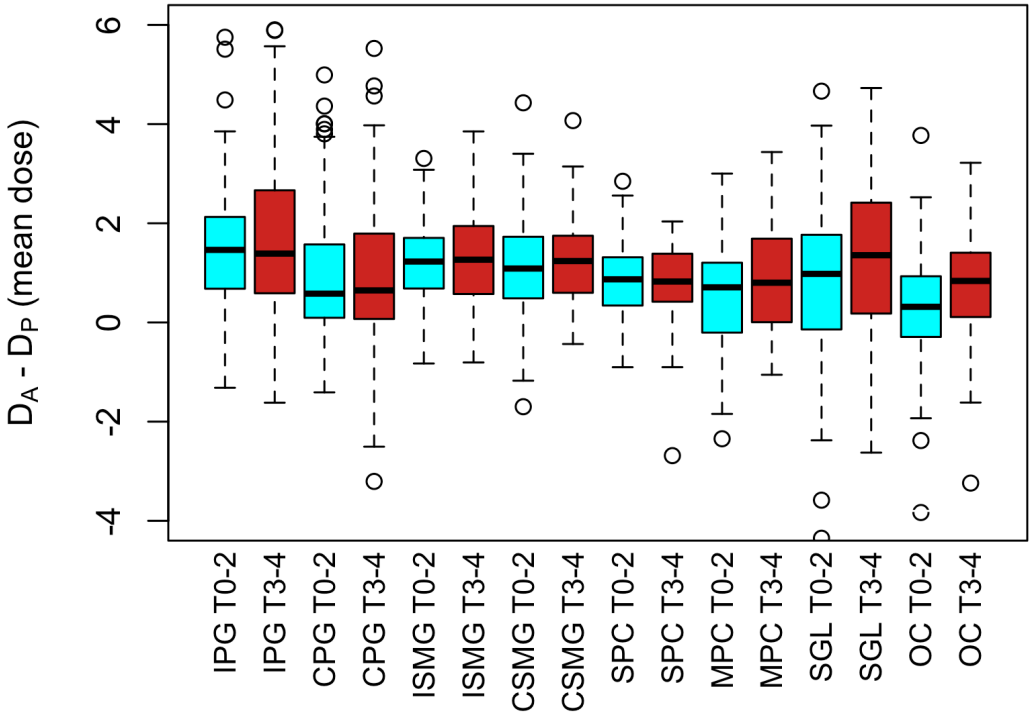


Figure 6.5: Boxplot of dose differences to swallowing OARs by T-stage, binarised as T0-2 (cyan boxes) vs. T3-4 (red boxes). Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity.

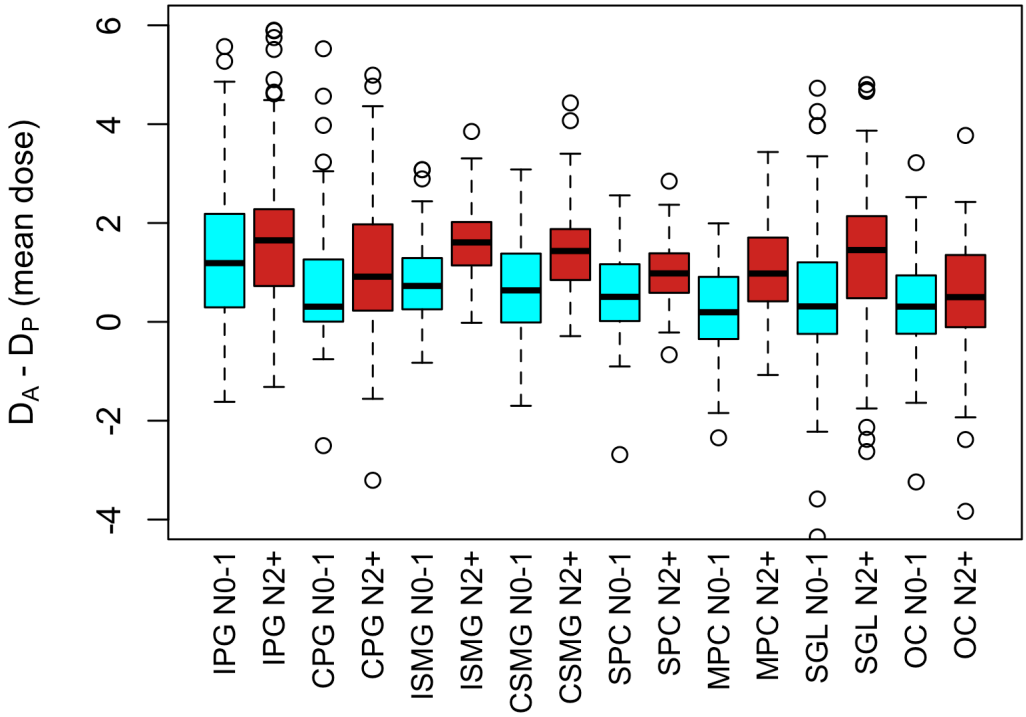


Figure 6.6: Boxplot of dose differences to swallowing OARs by N-stage, binarised as N0-1 (cyan boxes) vs. N2+ (red boxes). Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity.

6.4.5 Neck dissection vs. dose differences

Figure 6.7 shows a boxplot of dose differences to each OAR, with data split by whether or not patients underwent a neck dissection prior to definitive radiotherapy. On direct inspection of the data, no clear pattern was apparent for most OARs. Nonetheless, separate student t-tests were performed for each OAR, and statistical significance determined by a p-value <0.0063. Mean $D_A - D_P$ to the Ipsilateral parotid was 1.68Gy in patients who did not have a neck dissection, and 1.26Gy in patients who did (p=0.025). For the contralateral parotid, respective values were 1.04Gy and 0.71Gy (p=0.052). P-values for all other OARs were >0.2. Thus, with Bonferroni correction, neck dissection prior to RT was not a statistically significant predictor of dose differences to any OARs, although suggestive relationships for both parotid glands were seen.

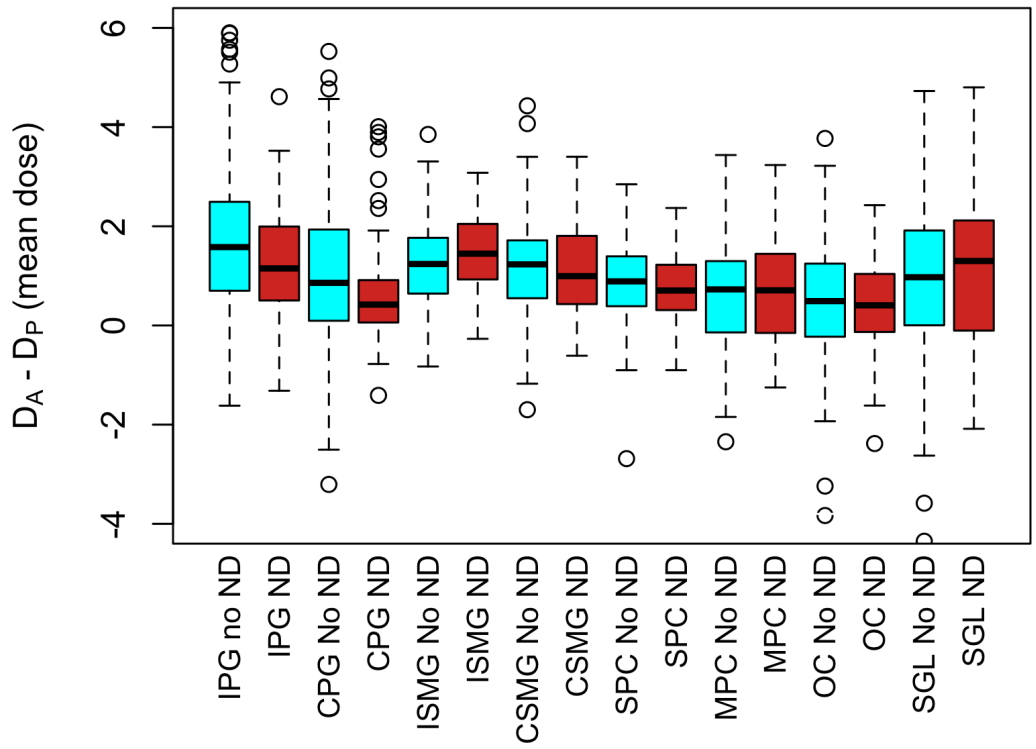


Figure 6.7: Boxplot of dose differences to all swallowing OARs, split by whether or not patients underwent a primary neck dissection (ND) – red boxes, or not (no ND) – cyan boxes. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity.

6.4.6 Planned OAR dose vs. dose differences

Previous studies have shown relationships between planned dose to parotid glands and dose differences [19, 35, 47]. Putative relationships between planned (mean) dose and dose differences to all swallowing OARs were investigated with univariate linear regression models, and Pearson correlation coefficients. Clearly significant relationships were seen for both submandibular glands, superior and middle pharyngeal constrictors, and the contralateral parotid gland. Weak/borderline relationships were seen for ipsilateral parotid gland, and supraglottic larynx. The results are summarised in Table 6.10:

	Planned dose vs dose difs to OARs		
	r	95%CI (r)	p-value
IPG	0.19	0.06 – 0.31	0.004
CPG	0.38	0.26 – 0.48	<0.001
ISMG	0.36	0.22 – 0.48	<0.001
CSMG	0.34	0.23 – 0.45	<0.001
SPC	0.29	0.17 – 0.40	<0.001
MPC	0.29	0.16 – 0.40	<0.001
SGL	0.19	0.06 – 0.31	0.005
OC	0.14	0.02 – 0.26	0.027

Table 6.10: Pearson product-moment correlation coefficients (with 95% confidence intervals and p-values) for univariate relationships between mean planned dose (Gy), and absolute dose differences (Gy) to all swallowing OARs. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Bonferroni correction: p-value <0.0063 considered statistically significant.

6.4.7 Neck RT strategy vs. dose differences

For this analysis, radiotherapy plans were considered as either treating both sides of the neck or not. Therefore patients were considered as having undergone unilateral or bilateral treatment. It seems intuitive that this factor might affect differences between planned and delivered dose to OARs, and it too has been investigated in previous studies. Therefore, each OAR was tested separately to ascertain whether neck RT strategy might predict higher than planned delivered dose. Results of each test are presented in Table 6.11, and the data are shown in Figure 6.8.

	Unilateral	Bilateral	p-value	95% CI
	Mean D_A-D_P (Gy)	Mean D_A-D_P (Gy)		
IPG	0.95	1.80	<0.001	0.51 to 1.20
CPG	0.23	1.25	<0.001	0.79 to 1.26
ISMG	0.70	1.41	<0.001	0.47 to 0.95
CSMG	0.74	1.33	<0.001	0.31 to 0.88
SPC	0.38	1.00	<0.001	0.40 to 0.84
MPC	0.24	0.87	<0.001	0.33 to 0.91
SGL	0.55	1.14	0.010	0.14 to 1.04
OC	0.28	0.49	0.12	-0.06 to 0.49

Table 6.11: Differences between planned and delivered mean dose to swallowing OARs by neck irradiation strategy (unilateral vs. bilateral). P-values and 95% confidence intervals from t-tests between distributions. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Bonferroni correction: p-value <0.0063 considered statistically significant.

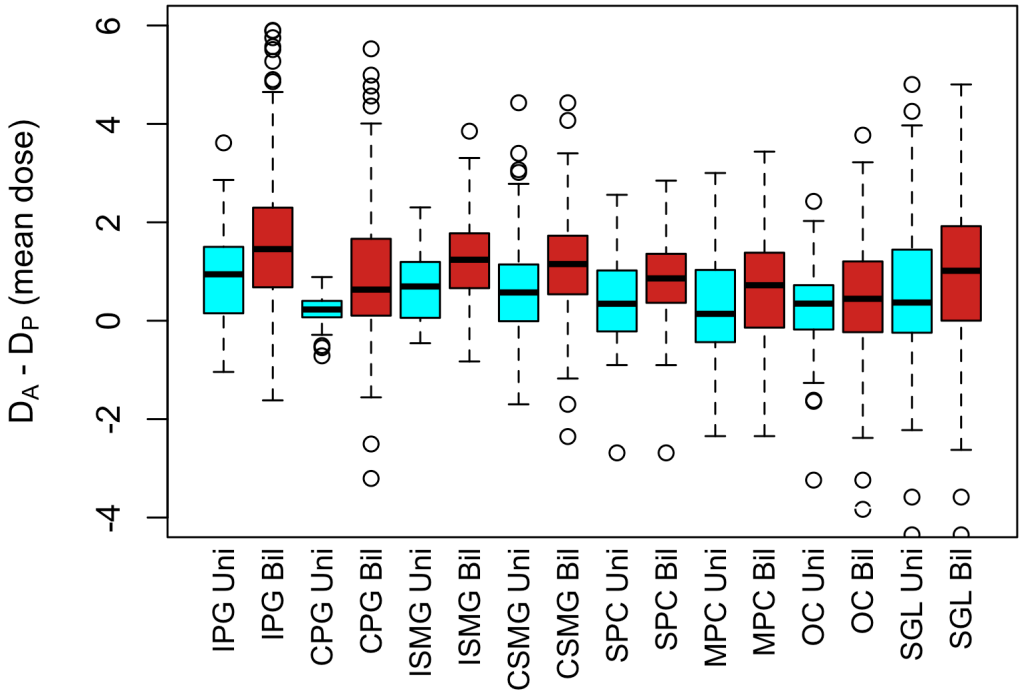


Figure 6.8: Boxplot of swallowing OARs dose differences by neck RT strategy - Unilateral neck RT (Uni, cyan boxes) or bilateral neck RT (Bil, red boxes). IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity.

6.4.8 Systemic therapy vs. dose differences

Receipt of concomitant systemic therapy was also investigated as a predictor of dose differences to OARs. As shown in Table 6.1, the numbers of patients receiving concomitant cisplatin, cetuximab, or no concomitant therapy were 133, 17 and 93 respectively. Due to the (relatively) low number of patients receiving cetuximab, the data were first plotted as 8 separate boxplots to visualise the relationship between each group (plots not shown). For most OARs, the dose difference data distributions for patients receiving cisplatin and cetuximab were very similar. Therefore, to simplify the analysis and improve statistical power these 2 groups of patients were combined to a single group who received systemic concomitant therapy (SCT). As previous, 8 separate t-tests (with Bonferroni corrections) were performed, with the results presented in Table 6.11. The data are shown in Figure 6.9.

	No SCT	**SCT**	**p-value**	**95% CI**
	Mean D_A-D_P (Gy)	**Mean D_A-D_P (Gy)**		
IPG	1.22	1.77	0.004	0.18 to 0.92
CPG	0.52	1.22	<0.001	0.51 to 1.22
ISMG	0.78	1.48	<0.001	0.45 to 0.96
CSMG	0.73	1.41	<0.001	0.44 to 0.92
SPC	0.48	1.02	<0.001	0.35 to 0.73
MPC	0.27	0.94	<0.001	0.40 to 0.93
SGL	0.54	1.21	0.002	0.25 to 1.08
OC	0.32	0.50	0.18	-0.09 to 0.45

Table 6.12: Differences between planned and delivered mean dose to swallowing OARs by use of concomitant systemic therapy - cisplatin or cetuximab (SCT) vs. none (no SCT). P-values and 95% confidence intervals from t-tests between distributions. IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. Bonferroni correction: p-value <0.0063 considered statistically significant.

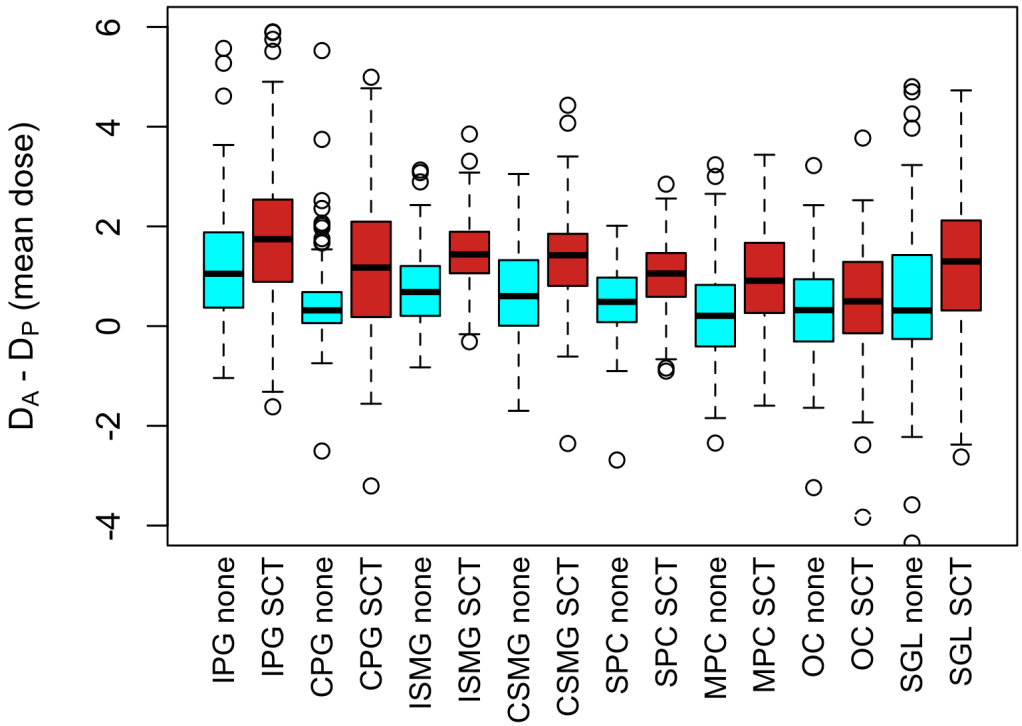


Figure 6.9: Boxplot of swallowing OARs dose differences by concomitant systemic therapy – no systemic therapy (none, cyan boxes) or cisplatin/cetuximab (SCT, red boxes). IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity.

6.4.9 Parotid gland volume change vs. dose differences

Mean per-gland Jacobian determinant values were calculated as described in section 6.3.3. Matched Jacobian determinant and dose difference data were available for 218 patients. As a smaller Jacobian value indicates greater shrinkage, and previous work has found a positive correlation between shrinkage and dose differences, a value of 1 – Jacobian determinant was taken for analysis for each patient. Both relationships were statistically significant, although the relationship was stronger for the ipsilateral parotid gland. The Pearson product-moment correlation coefficient was 0.34 (95% CI 0.22-0.46, $p<0.0001$), for the ipsilateral parotid, and 0.20 (95% 0.07-0.32, $p=0.004$) for the ipsilateral parotid gland. These data are shown in Figure 6.10.

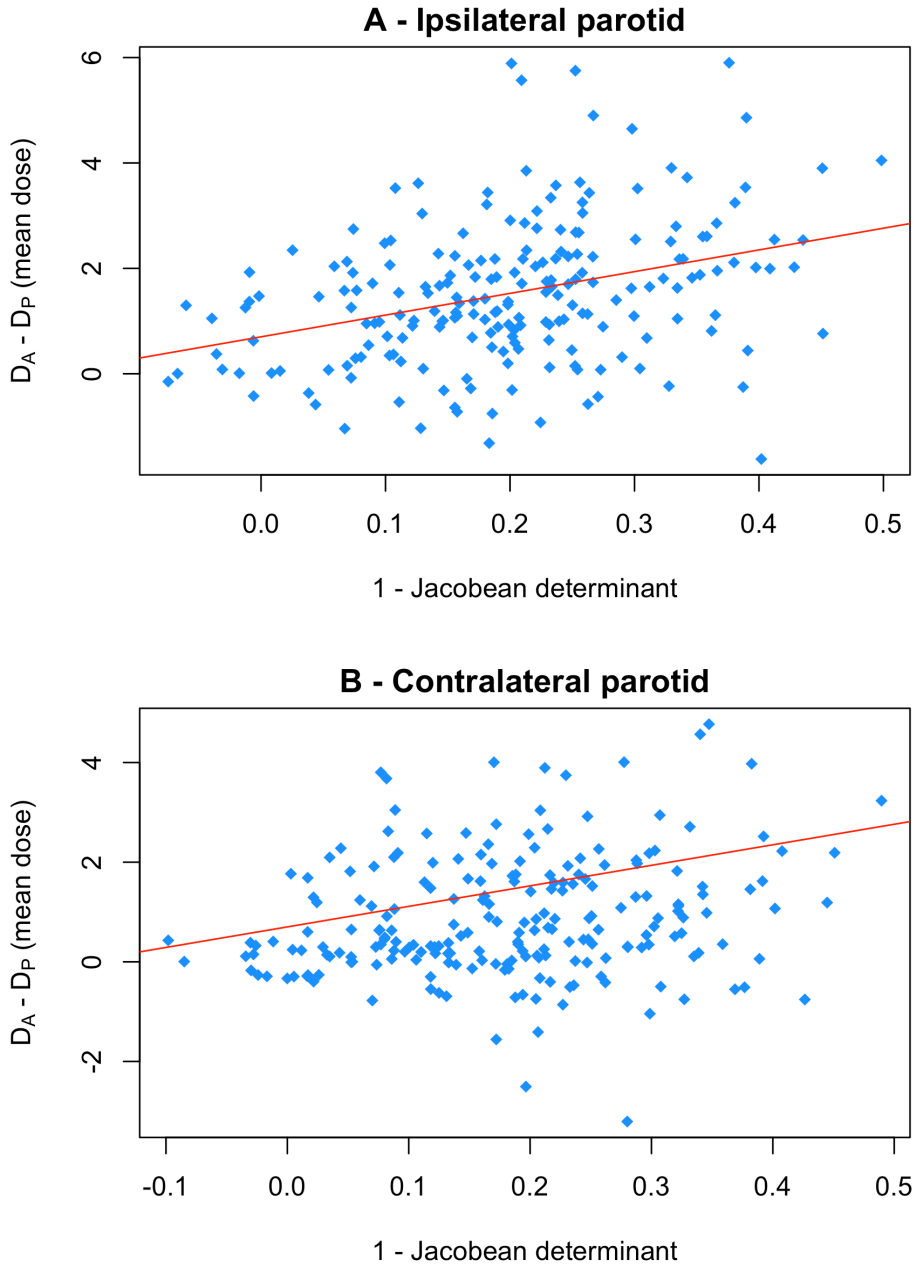


Figure 6.10 (A&B): Relationship between volume loss, and dose differences to the parotid gland. Scatter plots of 1 – Jacobian determinant against dose differences for the ipsilateral (A) and contralateral (B) parotid glands.

6.5 Discussion

6.5.1 Dose Differences

These data show that delivered dose is greater than planned for the swallowing OARs studied in this chapter. The greatest difference was to the ipsilateral parotid gland (1.56Gy, 95%CI 1.37 – 1.74), and the least was seen by the oral cavity (0.44Gy, 95%CI 0.30 – 0.57). These data broadly concur with previous studies, although the magnitudes of differences for some OARs are smaller than has been previously observed. For example, a review by Brouwer et al found that mean dose difference to the parotid glands was 2.2Gy [36]. There are fewer published data on other OARs, although one study reported submandibular gland dose difference results similar to those shown here (D_P 51.9Gy vs. D_A 52.8Gy) [29].

There are a number of possible explanations for this discrepancy. Two factors, discussed in Chapter 5 as possible reasons for observed differences in spinal cord statistics, may also be relevant here. The first is that all patients in this study were treated with DIPC and a zero-action level policy [48], in contrast to some comparable studies [23, 49, 50], leading to smaller dose differences due to setup error. The second is that daily imaging was used for dose calculation, meaning that substantial possible dose differences calculated on a 'one-off' imaging timepoint are not magnified by interpolation. Interestingly, the dose difference results to most OARs are more similar to those reported by a more recent publication in which daily IG was used both for setup correction and dose calculation [45].

The intrinsic uncertainties associated with DIR, as outlined in Chapter 4, may also be important, although the 'summary' performance statistics for Elastix in this work are comparable to other studies where similar data are available. Another explanation is the attempt to make the cohort as broad and representative by including patients with a range of primary disease sites. In contrast to other groups in which narrower cohorts of patients with pharyngeal SCC's have been studied [34, 45], there were a significant number of patients in this work who had undergone treatment for primary disease of the sinuses, salivary glands, or skin. Treatment plans for primary disease at these sites tend to have smaller target volumes, often treating only unilaterally, and lower prescription doses. They are also less likely to receive concomitant systemic therapy, or undergo significant weight loss. The data in this chapter show relationships between these factors and higher delivered dose to most of the OARs being studied. Therefore it is logical to suggest that the inclusion of these patients into the final cohort may have reduced mean dose differences to most OAR

6.5.2 Predictors of dose differences

The relationship between primary disease site and dose differences to all swallowing OARs was assessed by ANOVA. Primary site was found to be a statistically significant predictor of dose differences to all OARs except the supraglottic larynx and oral cavity. On inspection of the boxplots in Figure 6.4, there is certainly the suggestion that mean dose differences to all OARs are consistently lower for patients with primary disease of the salivary glands, skin or sinuses, as suggested in the previous section. This result is as expected, as treatment plans for these disease sites rarely included elective target volumes in the neck. This means that absolute doses to some OARs are likely to be lower, and that plan dosimetry is less likely to be affected by changing neck anatomy. Despite the statistical significance, the only other relationship that clearly appears from inspection of the data is that patients with primary nasopharynx cancer (NPC) have bigger dose differences to both parotid glands. It should be noted that there were only 8 NPC patients in the cohort. Therefore, despite the large overall sample size, the numbers for this specific question are small.

Nonetheless, this result is both entirely plausible, and supported by the literature. Two comparable but smaller studies were based on cohorts in which half the patients had NPC [32, 51], and found mean dose differences the parotids of 3Gy and 5Gy respectively. Another study in which all 19 patients had NPC saw dose differences of up to 10Gy to the parotids [52]. NPC treatment plans tend to have higher parotid gland doses, with mean doses often >30Gy [32]. Higher planned dose may well be an independent risk factor for greater dose differences [31], and it may also lead to very steep dose gradient in the region of the parotid glands. Thus, if there is anatomical change in the region of the parotids due to shrinkage of primary of nodal tumour, or involution of the glands themselves [19, 20, 47, 53], this will lead to greater dose deltas. Furthermore, patients with nasopharyngeal primary disease are more likely to have bulky nodal disease, more likely to undergo concomitant systemic therapy, and less likely to undergo primary neck dissection, all factors associated with greater dose differences according to previous work, and the data in this chapter.

Relationships between planned dose to OARs and dose differences were assessed. Significant associations between planned dose and dose differences were found for all OARs except the oral cavity, and strongly significant relationships were found for the submandibular glands, pharyngeal constrictor muscles, and contralateral parotid glands. Absolute planned dose to the parotid glands has been shown to correlate with eventual dose differences [31], but there are no comparable data for other structures. Planned SMG dose has been shown to

correlate with volume changes in one study [35], although this finding was not replicated elsewhere [25].

The notion that volume change per-se may be an independent risk factor for greater dose differences is supported both by previous research [29], and by data presented in this chapter. However, it is worth noting that there is an important methodological difference. Previous studies have measured volume change in small patient samples with manual segmentation of parotid glands [49, 50]. In contrast, the parameter used in this work was the Jacobian determinant [54], which was calculated in an automated fashion for a large sample. Whilst the Jacobian determinant is an inherently noisy metric, this is an interesting finding, as it relates a direct numerical output from the DIR process directly to a clinically relevant endpoint – dose differences to OARs. This approach was only undertaken for the parotid glands because of concerns about noise in the Jacobian data, and because it was novel. However, the findings suggest it would be an interesting approach to follow for the other swallowing OARs.

The effect of both weight loss and anatomical change on dose differences was also investigated. Both have been widely investigated as predictors of dose differences to a number of H&N OARs, and found to be significant predictors in most cases [20, 21, 30, 34]. Overall, these findings are supported by this work, although there are a number of discrepancies. Statistically significant associations between weight loss and dose differences were found for both superior and middle pharyngeal constrictors, contralateral submandibular gland, and supraglottic larynx, but not other OARs. In contrast, Lee et al [32], and Wang and colleagues [34], both found strong correlations between weight loss and dose increase to the parotid glands. Relationships between anatomical change metrics and dose differences were, in general, stronger than those observed for weight loss, and were comparable to those reported in the literature [20, 21, 30]. Again, the strongest relationships were seen for the pharyngeal constrictor muscles, contralateral submandibular gland, and supraglottic larynx.

Interestingly, the strongest relationships between N-stage and dose differences were found for the same structures. T-tests on dose differences for patients with N2+ compared with N0-1 disease were highly significant for the pharyngeal constrictors, submandibular glands and supraglottic larynx. P-values for the other 3 OARs did reach statistical significance with an α of 0.0063, but visual inspection of the data for all OARs in Figure 6.6 does suggest that delivered dose is consistently higher to all OARs for patients with more advanced neck disease. Although this finding per-se is new, it is worth noting that previous authors have found higher N-stage to be an independent risk factor for clinician-directed ART [45].

T-stage was less predictive than N-stage. The only relationship that reached clear statistical significance was the supraglottic larynx (p=0.003), and the oral cavity was borderline significant (p=0.007). This is not entirely surprising. Firstly, pathological lymph nodes tend to have fairly isotropic shapes, being either spherical or ellipsoid, meaning that maximum single dimension, as used in staging, is a good surrogate for volume. In contrast, primary disease may have a more complex shape, whereby the link between maximum single dimension and volume is much less robust. Furthermore, T-stage is also affected by the relationship of the tumour with nearby normal anatomy, meaning that a relatively small tumour may be staged T4 due to invasion into neighbouring structures. It therefore follows that the link between N-stage and internal anatomical change during treatment – and therefore dose differences – may be less noisy than that for T-stage.

The data relating pre-RT neck dissection to dose differences hint at support for this hypothesis. For both parotid glands, patients who underwent neck dissection for bulky neck disease before radiotherapy appeared to experience lower dose differences than those who did not (both relationships borderline significant, IPG p=0.025, CPG p=0.052). However, pre-RT neck dissection did not appear to affect the risk of higher delivered doses to other OARs, and to the best of my knowledge, there are no published data addressing this specific question.

Other aspects of treatment strategy clearly do influence the risk of greater dose differences to swallowing OARs. Patients receiving concomitant systemic therapy were significantly more likely to have higher D_A to all OARs except the oral cavity. Brouwer et al found that concomitant chemotherapy was a significant univariate predictor of greater dose differences to the parotids [37], but the effect on other OARs has not been studied. A similar pattern was seen for neck irradiation strategy. As might be expected, bilateral neck treatment was significantly associated with higher delivered dose to all OARs except the supraglottic larynx and oral cavity. However, what is interesting and perhaps counter-intuitive is to note that differences were highly significant for both *ipsilateral* paired structures. Dose differences to the ipsilateral parotid and submandibular gland in patients undergoing bilateral neck treatment were double those found in patients whose treatment was unilateral.

Perhaps the best explanation is a theme that runs through all these univariate analyses, that of co-linearity and confounding. In this instance, it is likely that patients undergoing bilateral treatment are also more likely to have more advanced disease, to receive concomitant systemic therapy, and to have nasopharyngeal primaries, all of which have been shown to be predict higher delivered doses. This problem is pertinent elsewhere in the analysis. For example, it is well known that concomitant systemic therapy worsens side effects [55-57],

which may lead to greater weight loss and anatomical change, which are independent predictors of dose differences [20, 32, 34]. Therefore, whilst the data presented in this chapter are interesting and of value, the next step is to undertake multivariate analysis of which factors predict significant dose differences to all swallowing OARs.

6.6 Conclusions

This is the largest study to assess differences between planned and delivered dose to the key OARs associated with dry mouth, thick sticky saliva, and swallowing problems, in patients who have undergone radical RT for HNC. It is one of few to have done so using daily imaging for dose accumulation, and to assess OARs other than the parotid glands. Calculated dose differences to the parotids are slightly lower than those previously reported, which may reflect greater heterogeneity in this patient cohort. Dose differences to other OARs appear comparable to those reported in the literature, although there are fewer data available for comparison. Although mean dose differences across the cohort were relatively small (range 0.44Gy to 1.56Gy), a substantial proportion of patients in the cohort had potentially clinically significant dose differences to some OARs, of 2Gy or more. These patient's would therefore have benefited more from treatment adaptation to reduce the risk of toxicity.

A number of factors were found to predict dose differences to OARs, although different factors were relevant for each OAR. Patients with NPC appear to be at greater risk of larger dose differences to the parotid glands, a finding that validates similar published results, and supports planned ART for patients with NPC in an attempt to reduce the risk of moderate to severe late xerostomia. Weight loss, anatomical change and more advanced neck disease all predicted greater dose differences to the SMGs, and the midline structures – pharyngeal constrictor muscles and supraglottic larynx, suggesting that these factors may help to select patients for ART, if reduction of swallowing dysfunction is an important clinical objective. Concomitant systemic therapy seems to increase risk of higher delivered dose to all OARs. Interestingly, patients who had clinically significant dose differences to more than half of the OARs studied shared similar case features. All of them had primary SCC of the oropharynx, with advanced neck nodal disease, and were treated with bilateral neck RT, whilst the majority (8/9) also received concomitant cisplatin. ART for these patients is likely to have reduced the risk of late toxicity.

6.7 References

[1] Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. J Clin Oncol. 2008;26:3770-6.
[2] Gregoire V, Jeraj R, Lee JA, O'Sullivan B. Radiotherapy for head and neck tumours in 2012 and beyond: conformal, tailored, and adaptive? Lancet Oncol. 2012;13:e292-300.
[3] Bjordal K, Kaasa S, Mastekaasa A. Quality of life in patients treated for head and neck cancer: a follow-up study 7 to 11 years after radiotherapy. Int J Radiat Oncol Biol Phys. 1994;28:847-56.
[4] Epstein JB, Emerton S, Kolbinson DA, Le ND, Phillips N, Stevenson-Moore P, et al. Quality of life and oral function following radiotherapy for head and neck cancer. Head Neck. 1999;21:1-11.
[5] Lin A, Kim HM, Terrell JE, Dawson LA, Ship JA, Eisbruch A. Quality of life after parotid-sparing IMRT for head-and-neck cancer: a prospective longitudinal study. Int J Radiat Oncol Biol Phys. 2003;57:61-70.
[6] Nguyen NP, Frank C, Moltz CC, Vos P, Smith HJ, Karlsson U, et al. Impact of dysphagia on quality of life after treatment of head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2005;61:772-8.
[7] Nguyen NP, Moltz CC, Frank C, Karlsson U, Nguyen PD, Vos P, et al. Dysphagia severity following chemoradiation and postoperative radiation for head and neck cancer. Eur J Radiol. 2006;59:453-9.
[8] Beetz I, Schilstra C, van der Schaaf A, van den Heuvel ER, Doornaert P, van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. Radiother Oncol. 2012;105:101-6.
[9] Saarilahti K, Kouri M, Collan J, Kangasmaki A, Atula T, Joensuu H, et al. Sparing of the submandibular glands by intensity modulated radiotherapy in the treatment of head and neck cancer. Radiother Oncol. 2006;78:270-5.
[10] Deasy JO, Moiseenko V, Marks L, Chao KS, Nam J, Eisbruch A. Radiotherapy dose-volume effects on salivary gland function. Int J Radiat Oncol Biol Phys. 2010;76:S58-63.
[11] Murdoch-Kinch CA, Kim HM, Vineberg KA, Ship JA, Eisbruch A. Dose-effect relationships for the submandibular salivary glands and implications for their sparing by intensity modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2008;72:373-82.
[12] Sapir E, Tao Y, Feng F, Samuels S, El Naqa I, Murdoch-Kinch CA, et al. Predictors of Dysgeusia in Patients With Oropharyngeal Cancer Treated With Chemotherapy and Intensity Modulated Radiation Therapy. Int J Radiat Oncol Biol Phys. 2016;96:354-61.
[13] Levendag PC, Teguh DN, Voet P, van der Est H, Noever I, de Kruijf WJ, et al. Dysphagia disorders in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscle: a dose-effect relationship. Radiother Oncol. 2007;85:64-73.
[14] Wopken K, Bijl HP, van der Schaaf A, van der Laan HP, Chouvalova O, Steenbakkers RJ, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol. 2014;113:95-101.
[15] Li B, Li D, Lau DH, Farwell DG, Luu Q, Rocke DM, 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;4:52.
[16] Dornfeld K, Simmons JR, Karnell L, Karnell M, Funk G, Yao M, 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;68:750-7.
[17] Dirix P, Abbeel S, Vanstraelen B, Hermans R, Nuyts S. Dysphagia after chemoradiotherapy for head-and-neck squamous cell carcinoma: dose-effect relationships for the swallowing structures. Int J Radiat Oncol Biol Phys. 2009;75:385-92.
[18] Ottosson S, Zackrisson B, Kjellen E, Nilsson P, Laurell G. Weight loss in patients with head and neck cancer during and after conventional and accelerated radiotherapy. Acta Oncol. 2013;52:711-8.
[19] Broggi S, Fiorino C, Dell'Oca I, Dinapoli N, Paiusco M, Muraglia A, et al. A two-variable linear model of parotid shrinkage during IMRT for head and neck cancer. Radiother Oncol. 2010;94:206-12.
[20] Ahn PH, Chen CC, Ahn AI, Hong L, Scripes PG, Shen J, et al. Adaptive planning in intensity-modulated radiation therapy for head and neck cancers: single-institution experience and clinical implications. Int J Radiat Oncol Biol Phys. 2011;80:677-85.

[21] Capelle L, Mackenzie M, Field C, Parliament M, Ghosh S, Scrimger R. Adaptive radiotherapy using helical tomotherapy for head and neck cancer in definitive and postoperative settings: initial results. Clin Oncol (R Coll Radiol). 2012;24:208-15.
[22] Barker JL, Jr., Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys. 2004;59:960-70.
[23] Hansen EK, Bucci MK, Quivey JM, Weinberg V, Xia P. Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2006;64:355-62.
[24] Robar JL, Day A, Clancey J, Kelly R, Yewondwossen M, Hollenhorst H, et al. Spatial and dosimetric variability of organs at risk in head-and-neck intensity-modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2007;68:1121-30.
[25] Vasquez Osorio EM, Hoogeman MS, Al-Mamgani A, Teguh DN, Levendag PC, Heijmen BJ. Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration. Int J Radiat Oncol Biol Phys. 2008;70:875-82.
[26] Han C, Chen YJ, Liu A, Schultheiss TE, Wong JY. Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy. Int J Radiat Oncol Biol Phys. 2008;70:1256-62.
[27] Loo H, Fairfoul J, Chakrabarti A, Dean JC, Benson RJ, Jefferies SJ, et al. Tumour shrinkage and contour change during radiotherapy increase the dose to organs at risk but not the target volumes for head and neck cancer patients treated on the TomoTherapy HiArt system. Clin Oncol (R Coll Radiol). 2011;23:40-7.
[28] Heukelom J, Fuller CD. Head and Neck Cancer Adaptive Radiation Therapy (ART): Conceptual Considerations for the Informed Clinician. Semin Radiat Oncol. 2019;29:258-73.
[29] Castadot P, Geets X, Lee JA, Gregoire V. Adaptive functional image-guided IMRT in pharyngo-laryngeal squamous cell carcinoma: is the gain in dose distribution worth the effort? Radiother Oncol. 2011;101:343-50.
[30] Castelli J, Simon A, Louvel G, Henry O, Chajon E, Nassef M, et al. Impact of head and neck cancer adaptive radiotherapy to spare the parotid glands and decrease the risk of xerostomia. Radiat Oncol. 2015;10:6.
[31] Hunter KU, Fernandes LL, Vineberg KA, McShan D, Antonuk AE, Cornwall C, et al. Parotid glands dose-effect relationships based on their actually delivered doses: implications for adaptive replanning in radiation therapy of head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2013;87:676-82.
[32] Lee C, Langen KM, Lu W, Haimerl J, Schnarr E, Ruchala KJ, et al. Assessment of parotid gland dose changes during head and neck cancer radiotherapy using daily megavoltage computed tomography and deformable image registration. Int J Radiat Oncol Biol Phys. 2008;71:1563-71.
[33] Marzi S, Pinnaro P, D'Alessio D, Strigari L, Bruzzaniti V, Giordano C, et al. Anatomical and dose changes of gross tumour volume and parotid glands for head and neck cancer patients during intensity-modulated radiotherapy: effect on the probability of xerostomia incidence. Clin Oncol (R Coll Radiol). 2012;24:e54-62.
[34] Wang X, Lu J, Xiong X, Zhu G, Ying H, He S, et al. Anatomic and dosimetric changes during the treatment course of intensity-modulated radiotherapy for locally advanced nasopharyngeal carcinoma. Med Dosim. 2010;35:151-7.
[35] Wang ZH, Yan C, Zhang ZY, Zhang CP, Hu HS, Kirwan J, et al. Radiation-induced volume changes in parotid and submandibular glands in patients with head and neck cancer receiving postoperative radiotherapy: a longitudinal study. Laryngoscope. 2009;119:1966-74.
[36] Brouwer CL, Steenbakkers RJ, Langendijk JA, Sijtsema NM. Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help? Radiother Oncol. 2015;115:285-94.
[37] Brouwer CL, Steenbakkers RJ, van der Schaaf A, Sopacua CT, van Dijk LV, Kierkels RG, et al. Selection of head and neck cancer patients for adaptive radiotherapy to decrease xerostomia. Radiother Oncol. 2016;120:36-40.
[38] Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, 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;117:83-90.

[39] Christianen ME, Langendijk JA, Westerlaan HE, van de Water TA, Bijl HP. Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. Radiother Oncol. 2011;101:394-402.
[40] Klein S, Staring M. Elastix: the manual, September 2015. [Online; accessed Dec 2015].
[41] Thomas SJ, Eyre KR, Tudor GS, Fairfoul J. Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube. Med Phys. 2012;39:160-7.
[42] Eisbruch A, Kim HM, Terrell JE, Marsh LH, Dawson LA, Ship JA. Xerostomia and its predictors following parotid-sparing irradiation of head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2001;50:695-704.
[43] El Naqa I, Bradley J, Blanco AI, Lindsay PE, Vicic M, Hope A, et al. Multivariable modeling of radiotherapy outcomes, including dose-volume and clinical factors. Int J Radiat Oncol Biol Phys. 2006;64:1275-86.
[44] Dean JA, Wong KH, Welsh LC, Jones AB, Schick U, Newbold KL, et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. Radiother Oncol. 2016;120:21-7.
[45] van Kranen S, Hamming-Vrieze O, Wolf A, Damen E, van Herk M, Sonke JJ. Head and Neck Margin Reduction With Adaptive Radiation Therapy: Robustness of Treatment Plans Against Anatomy Changes. Int J Radiat Oncol Biol Phys. 2016;96:653-60.
[46] Brown E, Owen R, Harden F, Mengersen K, Oestreich K, Houghton W, et al. Predicting the need for adaptive radiotherapy in head and neck cancer. Radiother Oncol. 2015;116:57-63.
[47] Sanguineti G, Ricchetti F, Thomas O, Wu B, McNutt T. Pattern and predictors of volumetric change of parotid glands during intensity modulated radiotherapy. Br J Radiol. 2013;86:20130363.
[48] Burnet NG, Adams EJ, Fairfoul J, Tudor GS, Hoole AC, Routsis DS, et al. Practical aspects of implementation of helical tomotherapy for intensity-modulated and image-guided radiotherapy. Clin Oncol (R Coll Radiol). 2010;22:294-312.
[49] Ho KF, Marchant T, Moore C, Webster G, Rowbottom C, Penington H, et al. Monitoring dosimetric impact of weight loss with kilovoltage (kV) cone beam CT (CBCT) during parotid-sparing IMRT and concurrent chemotherapy. Int J Radiat Oncol Biol Phys. 2012;82:e375-82.
[50] Bhide SA, Davies M, Burke K, McNair HA, Hansen V, Barbachano Y, et al. Weekly volume and dosimetric changes during chemoradiotherapy with intensity-modulated radiation therapy for head and neck cancer: a prospective observational study. Int J Radiat Oncol Biol Phys. 2010;76:1360-8.
[51] Nishi T, Nishimura Y, Shibata T, Tamura M, Nishigaito N, Okumura M. Volume and dosimetric changes and initial clinical experience of a two-step adaptive intensity modulated radiation therapy (IMRT) scheme for head and neck cancer. Radiother Oncol. 2013;106:85-9.
[52] Cheng HC, Wu VW, Ngan RK, Tang KW, Chan CC, Wong KH, et al. A prospective study on volumetric and dosimetric changes during intensity-modulated radiotherapy for nasopharyngeal carcinoma patients. Radiother Oncol. 2012;104:317-23.
[53] Fiorino C, Rizzo G, Scalco E, Broggi S, Belli ML, Dell'Oca I, et al. Density variation of parotid glands during IMRT for head-neck cancer: correlation with treatment and anatomical parameters. Radiother Oncol. 2012;104:224-9.
[54] Rohlfing T, Maurer CR, Jr., Bluemke DA, Jacobs MA. Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. IEEE Trans Med Imaging. 2003;22:730-41.
[55] Denis F, Garaud P, Bardet E, Alfonsi M, Sire C, Germain T, et al. Final results of the 94-01 French Head and Neck Oncology and Radiotherapy Group randomized trial comparing radiotherapy alone with concomitant radiochemotherapy in advanced-stage oropharynx carcinoma. J Clin Oncol. 2004;22:69-76.
[56] Trotti A, Pajak TF, Gwede CK, Paulus R, Cooper J, Forastiere A, et al. TAME: development of a new method for summarising adverse events of cancer treatment by the Radiation Therapy Oncology Group. Lancet Oncol. 2007;8:613-24.
[57] Machtay M, Moughan J, Trotti A, Garden AS, Weber RS, Cooper JS, et al. Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: an RTOG analysis. J Clin Oncol. 2008;26:3582-9.

Chapter 7 – Is Delivered Dose a Better Predictor of Toxicity?

7.1 Overview

Adaptive radiotherapy is widely regarded as a promising approach for reducing toxicity following radiotherapy for HNC. Although a number of studies have shown that delivered dose to H&N OARs may be different to planned, and that treatment adaption can reduce these dose differences, there are very few data showing any clinical benefit from ART. Crucially, there is no published work asking whether delivered dose is a better predictor of toxicity than planned. This chapter presents work in which an NTCP modelling approach, using both planned and delivered dose metrics, was used to address this hypothesis, in an attempt to provide much needed clinical endpoint data to support ART as a worthwhile approach in HNC patients.

7.1.1 My Role

Although the underlying hypothesis regarding relationships between planned and delivered dose and toxicity were core to the broader VoxTox programme, ideas regarding its specific implications for HNC patients, and how they might fit into the broader landscape of ART for HNC were exclusively mine. Furthermore, I developed and executed all of the methodologies described in this chapter. Collation, curation, and carpentry of all baseline demographic, disease, and treatment data used in this chapter were done by me. Planned and delivered dose data were produced by the processes described in Chapters 4, 5 and 6, and toxicity results were generated by the workflow described in Chapter 3. I then undertook all the data analysis and toxicity modelling work described in the chapter.

7.1.2 Acknowledgements

As described in Chapters 2 and 3, Amy Bates (Research Radiographer) and members of her team recruited all the patients whose data is used in this Chapter to the VoxTox study, and undertook follow up interviews to collect raw toxicity data. Karl Harrison (Research Associate, Cavendish Laboratory), with assistance from Lin Yeap (MPhil Student, Cavendish Laboratory), Shannon Seah (MPhil Student, Cavendish Laboratory), and Megan Wilson (PhD student, University College London) had a significant role in the generation of planned and delivered dose data used in this Chapter. Leila Shelley (Medical Physicist, Edinburgh Cancer Centre) provided guidance in the process of developing the methodology for using NTCP modelling as a tool for comparing planned and delivered dose metrics as toxicity predictors, and I am

grateful to Professor Hans Langendijk (Professor of Radiation Oncology, University Medical Centre, Groningen) for helpful conversations on this topic.

7.2 Introduction

The data presented in Chapter 6, and similar studies in the literature, have all looked at differences between planned and delivered dose, and tried to finds ways of predicting these differences [1-3]. This work is largely based on the underlying premise that dose differences are important, and where dose to an OAR of interest is higher than anticipated at planning, this scenario should be avoided if possible. There is plentiful indirect evidence to support this notion; not least clinical data where dose reductions to specific OARs have resulted in substantially lower toxicity rates [4-7]. The idea that dose is an important and independent predictor of toxicity is also extensively supported by in-silico modelling work [8-11], and on-going randomised trials hypothesise that dose reduction will also reduce toxicity [12]. Therefore, it is uncontroversial to state that dose matters, and the strength of this assertion has underpinned the design of studies which have tried to test the efficacy of ART in the clinic [13-15].

This research has taken place at the level of individual patients; the common concept being that once dose differences exceed a pre-defined level in a given patient, replanning should occur, because an inferred threshold in the difference of toxicity probability has been breached. This is the notion of either ART_{ex_aequo}, or ART_{OAR} as defined by Heukelom and Fuller [16], and seems a logical approach to improve outcomes in HNC, where cure rates are improving, but late toxicity remains a significant problem [6, 17]. However, to the best of my knowledge, the question of whether D_A predicts toxicity more accurately than D_P on a *population* basis remains unanswered, and has been highlighted as a research priority [18]. This is potentially important, as it could improve the quality and accuracy of toxicity prediction models, an important objective per-se. Furthermore, such a result would emphasise that whilst differences between D_P and D_A to most OARs, in most patients, are small, larger differences *are* clinically relevant, because they can affect toxicity outcome prediction at a population level.

In prostate cancer, there is published work from the VoxTox group showing that delivered rectal dose surface map-based metrics predict rectal bleeding and proctitis more accurately than equivalent planned dose metrics [19]. However, these findings have not been replicated in other OARs and toxicities in other tumour sites. Furthermore, the methodology used in this study only looked at direct univariate relationships between radiotherapy dose and toxicity

endpoints. Other important patient, treatment and disease factors that may have affected outcome were not considered. For the work in this chapter, it was therefore decided to investigate univariate relationships between planned and delivered radiotherapy dose and toxicity, but also to assess dose-toxicity relationships in the context of multivariate analyses.

The classic approach to multivariate modelling of dose-toxicity response relationships is generically known as normal tissue complication probability (NTCP) modelling [20, 21]. Different NTCP modelling approaches assume different underlying mathematical relationships between radiation dose and the probability of a given event, but most generate sigmoidal curves to describe the relationship between dose to OARs and toxicity risk [21-24]. The best known technique for predicting toxicity following RT for HNC is the Lyman-Kutcher-Burman model [22], but this has been criticised for over-simplifying dose-volume effects, and failure to capture the complex relationships between dose, treatment time, clinical factors and toxicity outcomes [21]. The use of logistic regression-based models is theoretically sound [21, 24], is now the standard methodology for HNC toxicity modelling [8-10, 25, 26], and was the approach taken in this work.

Therefore, the research questions of this chapter were, firstly, to test the hypothesis that univariate relationships between delivered radiotherapy dose to relevant OARs and specific toxicity endpoints would be stronger than equivalent planned dose relationships. Secondly to investigate whether or not both planned and delivered OAR dose metrics could be used to build, test and validate multivariate toxicity prediction (NTCP) models. And finally, to see if D_A-based NTCP models would predict toxicity significantly more accurately than their D_P-based equivalents.

To address these questions, the objectives of this Chapter were as follows:

- To examine univariate relationships between demographic, disease and treatment variables, as well as planned and delivered dose metrics, and toxicity.
- To incorporate candidate variables from univariate analysis into multivariate NTCP models that relate planned radiation dose to important OARs to toxicity events in HNC patients recruited to the VoxTox study.
- To use the same methodology to produce equivalent models for delivered dose metrics.
- To compare the performance of planned and delivered dose-based models to address the hypothesis that delivered dose predicts toxicity events more accurately than planned dose.

7.3 Methods

7.3.1 Cohort selection and treatment details

The patient population for this analysis was based on the same cohort used for analysis of toxicity results in Chapter 3, and dose differences in Chapter 6. However, in order to build robust predictive models of endpoints of interest, it was necessary to optimise homogeneity in the final dataset for analysis. Previous multivariate toxicity modelling studies have included a range of baseline factors such as pre-RT surgery and pre-treatment symptoms, smoking and drinking habits and staging information [8-10]. Therefore, only patients recruited to the consolidation cohort of the VoxTox study were included in this sub-study. The timepoint chosen for toxicity assessment was 1-year point-prevalence – this is discussed in more detail in section 7.3.3. As shown in Chapter 3, 200 patients had 1 year follow up data. Of these, it was not possible to compute D_A for 2 patients, for reasons described in Chapter 6 (section 3.1). Therefore, a final cohort of 198 patients was available for analysis, and details are shown in Table 7.1.

Characteristic	Number
Age	58.9 (10.6)
Gender	
Male	164 (82.8%)
Female	34 (17.2%)
Weight	
Baseline (kg)	85.2 (18.1)
Weight loss (%)	7.4 (5.2)
Histology & Primary site	
SCC	180 (90.9%)
Oropharynx	120 (60.6%)
Larynx	16 (8.8%)
Oral cavity	13 (6.6%)
Sinus	9 (4.5%)
Nasopharynx	6 (3.0%)
Hypopharynx	4 (2.0%)
Unknown primary	7 (3.5%)
Skin	5 (2.5%)
Salivary gland	15 (7.6%)
Acinic cell	6 (3.0%)
Adenocarcinoma	5 (2.5%)
Adenoid cystic	3 (1.5%)
Mucoepidermoid	1 (0.5%)
Others	3 (1.5%)

T-stage	
Tis/T1	56 (28.3%)
T2	67 (33.8%)
T3	30 (15.2%)
T4	45 (22.7%)
N-stage	
N0	59 (29.8%)
N1	22 (11.1%)
N2a	29 (14.7%)
N2b	67 (33.8%)
N2c	18 (9.1%)
N3	3 (1.5%)
Surgery	
Neck dissection	57 (28.8%)
Parotidectomy	13 (6.6%)
Laryngectomy	4 (2.0%)
Dose/Fractionation	
70/35	6 (3.0%)
65/30	137 (69.2%)
60/30	47 (23.7%)
55/20	3 (1.5%)
50/20	5 (2.5%)
Neck Irradiation	
Unilateral	58 (29.3%)
Bilateral	134 (67.7%)
None	6 (3.0%)
Systemic therapy	
Cisplatin*	108 (54.5%)
Cetuximab*	14 (7.1%)
None	76 (38.4%)
Smoking	
Current smoker	13 (6.6%)
Ex-smoker	100 (50.5%)
Never smoker	57 (28.8%)
Missing data	28 (14.1%)
Mean pack years	17.5 (21.1)
Alcohol	
Current drinker	117 (59.1%)
Ex-drinker	41 (20.7%)
Never drinker	12 (6.1%)
Missing data	28 (14.1%)
Mean units/week	14.2 (23.6)

* dose: cisplatin – 40mg/m^2 weekly, cetuximab 400mg/m2 loading dose, 240mg/m^2 weekly thereafter.

Table 7.1: Patient characteristics – planned and delivered dose versus toxicity cohort (n = 198). For continuous variables, means and standard deviations are reported, absolute numbers and percentages for proportions.

Radiotherapy treatment planning and delivery, including immobilisation, simulation, target volume definition, dose/fractionation schedules, and concomitant therapy were all done according to standard departmental protocols as described in Chapter 2.

7.3.2 OAR segmentation and dose calculation

Workflows for the computation of planned and delivered dose were as described in Chapters 2, 4, 5 and 6. For each patient in this analysis, RT treatment planning CTs were restored to the research database of Prosoma (v3.3 MEDCOM, Darmstadt, Germany). I undertook segmentation of a defined set of relevant OARs for each patient, according to recognised contouring atlases [27, 28]. Quality control data for these segmentations is shown in Chapter 2. Once completed, finalised structure sets of segmented OARs and planning CT datasets were used to calculate planned dose DVHs, as described in Chapter 2. Daily MVCT-IG images, and deformable image registration, were then used to compute delivered dose to each OAR, as described in Chapters 4 and 5. Both planned and delivered dose calculation were done using the same dose calculation algorithm – CheckTomo [29] - in order to ensure consistency.

7.3.3 Toxicity endpoints

Studies investigating and reporting late effects of radiotherapy for HNC consistently show that point prevalence for most toxicities peak around 6-12 months, and gradually improve thereafter [4]. This finding has been replicated in VoxTox, as shown by the data in Chapter 3. Therefore, assessment of point-prevalence at 12 months meant that the majority of moderate to severe toxicity events in a study population will have been captured, and the literature supports this approach [25].

Although clinical trials often report toxicity data across a range of side effects [4, 30-32], most toxicity modeling papers focus on either a single endpoint [10, 26], or a small number [8, 25]. For this work, the starting objective was to investigate relationships between radiotherapy dose to a number of OARs and a range of toxicities. Furthermore, the data presented in Chapter 7 show that mean differences between planned and delivered dose are small, in the order of 1Gy for many OARs. With dose differences of these magnitudes, we hypothesised that it would be unlikely to find big differences in the predictive capacities of delivered and planned dose metrics, and there were a number of reasons for this assertion. First, whilst many published predictive models appear to perform well, they are not perfect, and model performance metrics such as area under receiver-operator characteristic (ROC) curves have substantial confidence

intervals [8-10], partly because toxicity endpoint data are inevitably noisy. It could also be argued that as model performance improves, the incremental benefit from superior predictor metric data declines. Another reason for a theoretical limit on the performance of dose-toxicity response models is that they take no account of individual radiosensitivity [33].

Therefore, a more realistic objective for this work was to examine a range of toxicity endpoints, and look for a trend in the pattern of model performance. On this basis, multivariate logistic regression toxicity prediction models, using both planned and delivered dose metrics, were constructed and validated for the following endpoints:

- 12 month_CTCAEv4.03_xerostomia_Gr2+
- 12 month_CTCAEv4.03_salivary duct inflammation_Gr2+
- 12 month_CTCAEv4.03_dysphagia_Gr2+
- 12 month_EORTC H&N35_Q41 – Have you had a dry mouth?_Gr3+
- 12 month_EORTC H&N35_Q42 – Have you had sticky saliva?_Gr3+
- 12 month_EORTC H&N35_Q44 – Have you had taste disturbance_Gr3+
- 12 month_LENT/SOM_Salivary gland subjective_Gr2+
- 12 month_LENT/SOM_Salivary gland management_Gr2+

In order to undertake logistic regression, toxicity results for each endpoint were binarised as outlined above. Thresholds for binarisation were chosen on the basis of literature precedent [8, 25], and to define a set of moderate-to-severe toxicity events.

7.3.4 Statistical analysis

7.3.4.1 General approach

For the reasons described in section 7.3.3, it was thought unlikely that the predictive power of delivered dose-based dose-toxicity models would substantially outperform planned dose equivalents. Therefore, the approach taken was to repeat the same methodology for 8 toxicity endpoints, and observe for a trend in differences of model performance. A brief overview of the methodology used is described below, with more detailed explanations in the following sections.

First, univariate regression was undertaken between all hypothesised predictors and binarised toxicity endpoints. This included all baseline factors such as age, baseline weight and weight loss as continuous variables, gender as a binary variable, and baseline symptoms as an ordinal variable. Staging data were included, and binarised such that T-stage was recorded as T0-2 (0) vs. T3-4 (1), and N-stage as N0-1 (0) vs. N2+ (1). Analysis also included smoking

(pack years), and alcohol consumption (estimated units/week) as continuous variables. Treatment factors were assessed, including the use of concomitant systemic therapy (by cisplatin vs. cetuximab vs. none as a nominal variable, and binarised as cisplatin Y/N, and systemic therapy Y/N), and pre-RT surgery (neck dissection and parotidectomy as binary variables, removal of SMGs as an ordinal variable). Finally, univariate relationships between mean D_P, and mean D_A to each OAR were recorded. Next, correlation statistics between candidate variables showing statistically significant or borderline univariate relationships with toxicity were calculated, and where co-linearity occurred, the variable with weaker association was excluded. All remaining variables were then included in a multivariate logistic regression, the output of which was a regression equation [21, 24, 26], expressed as shown in Equation 7.1:

$$S = \beta_0 + (\beta_1 V_1)..+(\beta_n V_n) \tag{7.1}$$

Normal tissue complication probability values were then calculated from this parameter according to the formula shown in Equation 7.2 [8, 10].

$$NTCP = \frac{1}{(1+ e^{-S})} \tag{7.2}$$

In the first instance this process was undertaken with planned dose metrics, and delivered dose metrics were not used. Planned dose models were optimised, calibrated and internally validated, and model performance was recorded. At this point, construction of the multivariate logistic regression model for each endpoint was repeated, directly substituting delivered dose metrics for planned dose. The process of variable selection was not repeated in this step, so non dose-based variables were simply retained in model construction. Delivered dose-based models were then optimised, calibrated and internally validated in an identical fashion to planned dose-based equivalents. Performance metrics were recorded and compared with planned dose model data.

All data storage and carpentry for this work was done in Microsoft excel. Forrest plots for model ROC curve AUCs and histograms of discrimination slope performance (section 7.4.5) were also produced in Microsoft excel. All univariate and multivariate regression analysis, and receiver operator characteristic (ROC) curve analysis was done in SPSS version 25 (IBM, Armonk, New York, USA). Most model calibration, discrimination and performance data were derived directly from SPSS output. Discrimination slope analyses were done in R version 3.4.0.

7.3.4.2 Univariate analyses and variable selection

In order to characterise crude relationships between each variable and the outcomes of interest, univariate logistic regression analyses were done for each binarised endpoint [34]. Analysis output included a Wald chi-squared value, odds ratio with 95% confidence intervals, and p-value. Results were recorded and collated, and all metrics with a p-value <0.1 were considered candidate variables for multivariate analysis. However, there was a significant risk of co-linearity between candidate variables [8], and this was quantified as follows. First, a matrix of Pearson correlation coefficients between each candidate variable was constructed. Next, tolerance and variance inflation factor (VIF) [25, 34], were calculated as shown in Equations 7.3 and 7.4:

$$Tolerance\ (T) = 1 - R^2 \tag{7.3}$$

$$VIF = \frac{1}{T} \tag{7.4}$$

A threshold VIF of 5 [34], equivalent to correlation coefficient of 0.8 [25], was used. Where such co-linearity existed, the variable with the stronger univariate relationship with outcome was retained, and the co-linear metric was discarded.

7.3.4.3 Model building, and internal validation

Selected candidate variables were then included in a multivariate logistic regression model. Again, the risk of over-fitting to training data is high, and the choice of methodology for variable selection with the multivariate model is important [21, 35]. Based on similar work in this field, forward variable selection by likelihood ratio was used [8, 26]. In this process, variables are added to the model in a stepwise fashion, and the statistical significance of differences in model performance is tested at each step. Once the p-value of adding parameters exceeds 0.05, variable selection is completed, and no further parameters are added. Rather than randomly splitting the dataset into training and validation cohorts, internal validation was done with 1000 bootstraps for each analysis, and 10-fold cross validation [8, 10, 25]. Analysis output was a logistic regression equation as shown in Equation 7.1, and NTCP values for each value were calculated with equation 7.2.

7.3.4.4 Assessing model performance

Model performance was assessed using standard tools [36, 37]. Overall model performance was summarised by Nagelkerke's R^2 statistic [8, 38], and the -2 log-likelihood ratio [25]. The

capacity of equivalent planned and delivered dose-based models to discriminate between cases with and without the endpoint of interest was tested using the respective area under the curve (AUC) of the ROC curves [8, 10, 25, 39, 40]. The discrimination slope – the absolute difference between NTCP values in patients with and without the endpoint of interest – was also calculated for both planned and delivered dose models [8, 10, 25]. An example discrimination slope plot is shown in Figure 7.1.

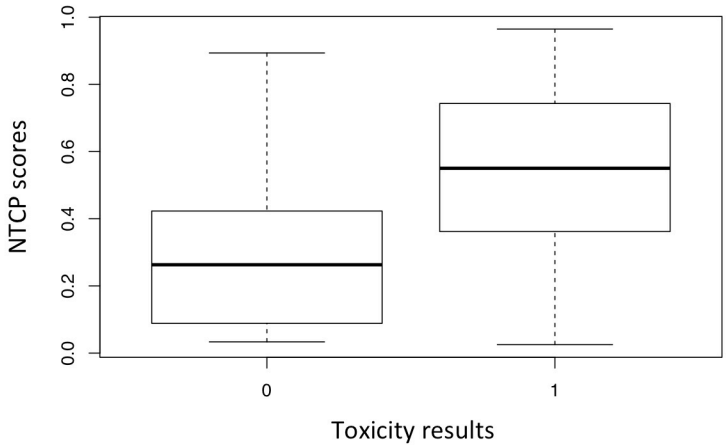


Figure 7.1: Example boxplot for computation of discrimination slope score. Mean NTCP values are calculated for those patients with (1) and without (0) the toxicity endpoint of interest (all endpoints for this analysis are binary).

Overall model calibration was analysed with a Hosmer-Lemeshow test with 10 groups [8, 10, 25], in which the expected and observed proportions of patients with toxicity are recorded. The 10 groups are partitioned to ensure that the observed event rates span the spectrum of predicted risk. From this 10x4 contingency table, a scatter plot of observed vs. expected risk was produced, with a line of best-fit. An example is shown in Figure 7.2.

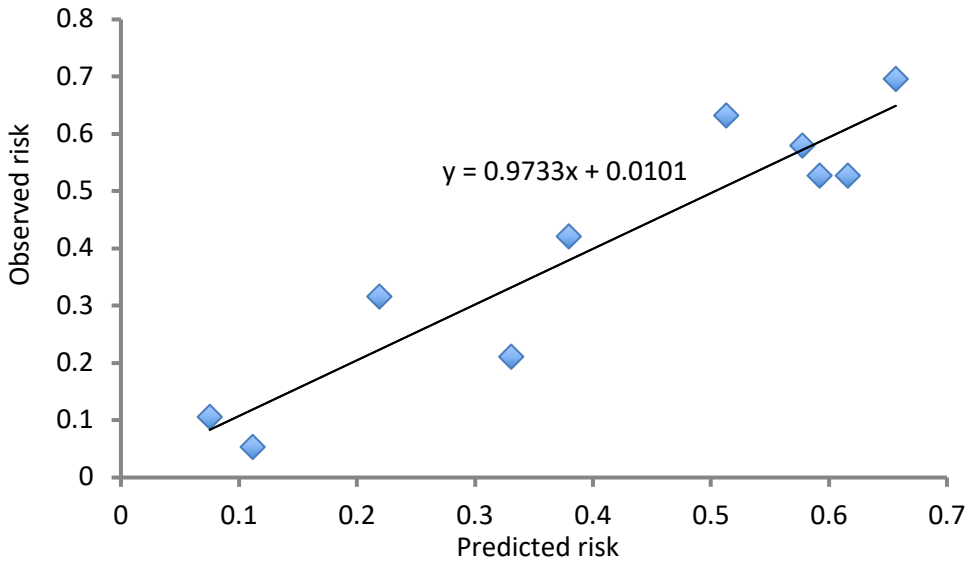


Figure 7.2: Example NTCP model calibration plot, showing slope and intercept of calibration slope.

The slope and intercept of this line were recorded and the Brier score was calculated, as shown in Equation 7.5:

$$Brier\ score = \frac{\sum(Obs-Exp)^2}{N} \quad (7.5)$$

7.4 Results

7.4.1 Toxicity data

Baseline and 1-year point prevalence toxicity data for all 8 endpoints were available for 198 patients. The mean age was 59 (range 26 to 88), and full details of the cohort are shown in Table 7.1. Full toxicity data, including baseline symptoms, are shown in Chapter 3. In order to undertake logistic regression analysis, toxicity results were split into binary groups, as described in section 7.3.3, although baseline toxicity scores were treated as ordinal variables in the regression analysis. Proportions of patients with the binary toxicity endpoint of interest, at baseline and at 12 months, for all endpoints, are shown in Table 7.2.

		CTCAEv4.03	
	Xerostomia Gr2+	SDI Gr2+	Dysphagia Gr2+
baseline	7 (3.5%)	5 (2.5%)	11 (5.6%)
12 month	75 (37.9%)	88 (44.4%)	39 (19.7%)
		EORTC H&N35	
	Q41 Gr3+	Q42 Gr3+	Q44 Gr3+
baseline	16 (8.1%)	12 (6.1%)	16 (8.1%)
12 month	95 (48%)	62 (31.3%)	74 (37.4%)
		LENT/SOM	
	Salivary gland subj. Gr2+		Salivary gland man. Gr1+
baseline	16 (8.1%)		1 (0.5%)
12 month	94 (47.5%)		82 (41.4%)

Table 7.2: Proportions of patients with binary outcomes of interest, at baseline and 12 months, for all 8 toxicity endpoints.

7.4.2 Univariate analyses

Full results of all univariate regression analyses are shown in Appendix A5. Truncated tables of univariate analysis results, showing only relationships with p-values <0.1, are shown in Tables 7.3 – 7.10 (for all tables, standard abbreviations are used: IPG = Ipsilateral parotid gland, CPG = Contralateral parotid gland, ISMG = Ipsilateral submandibular gland, CSMG = Contralateral submandibular gland, SPC = Superior pharyngeal constrictor muscle, MPC = Middle pharyngeal constrictor muscle, SGL = Supraglottic larynx, OC = Oral cavity. D_P indicates mean planned dose to an OAR, D_A indicates mean delivered dose).

Variable	Nature	Wald	p-value	OR	95% CI
weight loss (%)	cont	5.49	0.019	1.073	1.012 – 1.137
N stage	binary	9.84	0.002	2.71	1.453 – 5.040
cisplatin	nominal	10.9	0.001	2.99	1.562 – 5.730
SACT Y/N	binary	10.2	0.001	2.83	1.494 – 5.344
cisplatin Y/N	binary	10.4	0.001	2.71	1.476 – 4.957
parotidectomy	binary	3.91	0.048	0.13	0.016 – 0.982
IPG_D_P	cont	9.60	0.002	1.053	1.019 – 1.088
CPG_D_P	cont	28.8	<0.001	1.083	1.052 – 1.116
ISMG_D_P	cont	4.12	0.042	1.050	1.002 – 1.101
CSMG_D_P	cont	21.1	<0.001	1.056	1.032 – 1.081
SPC_D_P	cont	17.3	<0.001	1.100	1.052 – 1.151
MPC_D_P	cont	11.9	0.001	1.083	1.035 – 1.132
OC_D_P	cont	15.8	<0.001	1.073	1.036 – 1.111
SGL_D_P	cont	12.5	<0.001	1.078	1.034 – 1.124
IPG_D_A	cont	9.89	0.002	1.053	1.020 – 1.087
CPG_D_A	cont	30.8	<0.001	1.084	1.053 – 1.115
ISMG_D_A	cont	4.18	0.041	1.050	1.002 – 1.100
CSMG_D_A	cont	21.1	<0.001	1.055	1.031 – 1.080
SPC_D_A	cont	17.4	<0.001	1.099	1.051 – 1.149
MPC_D_A	cont	12.2	<0.001	1.078	1.034 – 1.125
OC_D_A	cont	16.3	<0.001	1.073	1.037 – 1.111
SGL_D_A	cont	12.7	<0.001	1.079	1.035 – 1.125
baseline Xer	ordinal	6.47	0.011	1.976	1.169 – 3.342

Table 7.3: Results of univariate logistic regression analyses with grade 2+ CTCAEv4.03 xerostomia as the primary endpoint. Candidate variables only shown.

Variable	Nature	Wald	p-value	OR	95% CI
gender (male)	binary	3.36	0.067	0.49	0.234 – 1.050
cetuximab	nominal	6.90	0.009	5.04	1.508 – 16.84
cisplatin	nominal	14.0	<0.001	3.37	1.785 – 6.367
SACT Y/N	binary	15.7	<0.001	3.53	1.891 – 6.575
cisplatin Y/N	binary	9.79	0.002	2.53	1.415 – 4.533
CPG_D_P	cont	11.8	0.001	1.043	1.018 – 1.068
CSMG_D_P	cont	13.0	<0.001	1.032	1.015 – 1.050
SPC_D_P	cont	6.56	0.010	1.036	1.008 – 1.064
MPC_D_P	cont	7.26	0.007	1.037	1.010 – 1.065
SGL_D_P	cont	8.01	0.005	1.036	1.011 – 1.063
CPG_D_A	cont	13.9	<0.001	1.045	1.021 – 1.069
CSMG_D_A	cont	13.1	<0.001	1.032	1.014 – 1.049
SPC_D_A	cont	6.75	0.009	1.036	1.009 – 1.063
MPC_D_A	cont	7.66	0.006	1.038	1.011 – 1.065
SGL_D_A	cont	8.25	0.004	1.037	1.012 – 1.062
baseline SDI	ordinal	3.70	0.054	1.81	0.989 – 3.296

Table 7.4: Results of univariate logistic regression analyses with grade 2+ CTCAEv4.03 salivary duct inflammation as the primary endpoint. Candidate variables only shown.

Variable	Nature	Wald	p-value	OR	95% CI
N stage	binary	3.17	0.075	2.00	0.932 – 4.300
cisplatin	nominal	5.26	0.022	2.61	1.150 – 5.905
SACT Y/N	binary	4.62	0.032	2.43	1.081 – 5.450
cisplatin Y/N	binary	5.60	0.018	2.51	1.171 – 5.394
alcohol (units/wk)	cont	3.60	0.058	1.006	1.000 – 1.012
IPG_D_P	cont	7.25	0.007	1.060	1.016 – 1.105
CPG_D_P	cont	8.01	0.005	1.046	1.014 – 1.079
CSMG_D_P	cont	7.65	0.006	1.039	1.011 – 1.068
SPC_D_P	cont	10.5	0.001	1.129	1.049 – 1.215
MPC_D_P	cont	6.82	0.009	1.094	1.023 – 1.170
OC_D_P	cont	8.01	0.005	1.069	1.021 – 1.120
SGL_D_P	cont	7.12	0.008	1.082	1.021 – 1.147
IPG_D_A	cont	7.52	0.006	1.060	1.017 – 1.106
CPG_D_A	cont	8.35	0.004	1.045	1.014 – 1.077
CSMG_D_A	cont	7.87	0.005	1.040	1.012 – 1.069
SPC_D_A	cont	10.6	0.001	1.130	1.050 – 1.217
MPC_D_A	cont	7.20	0.007	1.090	1.024 – 1.162
OC_D_A	cont	8.67	0.003	1.073	1.024 – 1.124
SGL_D_A	cont	7.30	0.007	1.084	1.022 – 1.148

Table 7.5: Results of univariate logistic regression analyses with grade 2+ CTCAEv4.03 dysphagia as the primary endpoint. Candidate variables only shown.

Variable	Nature	Wald	P value	OR	95% CI
gender (male)	binary	4.45	0.035	0.44	0.203 – 0.943
weight loss	cont	4.21	0.040	1.067	1.003 – 1.134
weight loss (%)	cont	5.65	0.017	1.072	1.012 – 1.136
N stage	binary	8.01	0.005	2.32	1.295 – 4.157
cisplatin	nominal	4.07	0.044	1.85	1.017 – 3.350
SACT Y/N	binary	3.55	0.060	1.75	0.978 – 3.127
cisplatin Y/N	binary	4.18	0.041	1.81	1.025 – 3.184
IPG_D_P	cont	6.69	0.01	1.039	1.009 – 1.069
CPG_D_P	cont	14.5	<0.001	1.048	1.023 – 1.073
ISMG_D_P	cont	4.80	0.028	1.043	1.004 – 1.082
CSMG_D_P	cont	16.4	<0.001	1.036	1.018 – 1.054
SPC_D_P	cont	17.8	<0.001	1.079	1.041 – 1.117
MPC_D_P	cont	13.4	<0.001	1.068	1.031 – 1.107
OC_D_P	cont	15.6	<0.001	1.060	1.030 – 1.092
SGL_D_P	cont	12.5	<0.001	1.055	1.024 – 1.087
IPG_D_A	cont	7.75	0.005	1.041	1.012 – 1.071
CPG_D_A	cont	15.7	<0.001	1.048	1.024 – 1.072
ISMG_D_A	cont	4.94	0.026	1.044	1.005 – 1.084
CSMG_D_A	cont	16.6	<0.001	1.036	1.018 – 1.053
SPC_D_A	cont	18.4	<0.001	1.079	1.042 – 1.118
MPC_D_A	cont	13.7	<0.001	1.066	1.030 – 1.103
OC_D_A	cont	16.0	<0.001	1.060	1.030 – 1.091
SGL_D_A	cont	12.5	<0.001	1.055	1.024 – 1.086

Table 7.6: Results of univariate logistic regression analyses with grade 3+ EORTC H&N35 Q41 – have you had a dry mouth? as the primary endpoint. Candidate variables only shown.

Variable	Nature	Wald	p-value	OR	95% CI
gender (male)	binary	3.06	0.080	0.50	0.239 – 1.085
weight loss (%)	cont	3.07	0.080	1.055	0.994 – 1.119
cetuximab	nominal	2.88	0.089	2.81	0.853 – 9.278
cisplatin	nominal	5.27	0.022	2.21	1.122 – 4.336
SACT Y/N	binary	5.89	0.015	2.27	1.171 – 4.400
cisplatin Y/N	binary	3.58	0.059	1.82	0.979 – 3.378
CPG_D_P	cont	11.2	0.001	1.046	1.019 – 1.074
ISMG_D_P	cont	2.83	0.093	1.035	0.994 – 1.077
CSMG_D_P	cont	12.3	<0.001	1.039	1.017 – 1.061
SPC_D_P	cont	5.20	0.023	1.037	1.005 – 1.071
MPC_D_P	cont	7.27	0.007	1.054	1.014 – 1.095
OC_D_P	cont	3.35	0.067	1.027	0.998 – 1.056
SGL_D_P	cont	8.92	0.003	1.063	1.021 – 1.107
CPG_D_A	cont	12.5	0.000	1.047	1.021 – 1.074
ISMG_D_A	cont	2.75	0.097	1.033	0.994 – 1.073
CSMG_D_A	cont	12.4	0.000	1.039	1.017 – 1.061
SPC_D_A	cont	5.40	0.020	1.038	1.006 – 1.070
MPC_D_A	cont	7.41	0.007	1.052	1.014 – 1.091
OC_D_A	cont	3.61	0.058	1.027	0.999 – 1.056
SGL_D_A	cont	9.03	0.003	1.064	1.022 – 1.108

Table 7.7: Results of univariate logistic regression analyses with grade 3+ EORTC H&N35 Q42 – have you had sticky saliva? as the primary endpoint. Candidate variables only shown.

Variable	Nature	Wald	p-value	OR	95% CI
cisplatin	nominal	7.75	0.005	2.49	1.310 – 4.738
SACT Y/N	binary	7.86	0.005	2.46	1.312 – 4.623
cisplatin Y/N	binary	6.39	0.011	2.16	1.189 – 3.922
CPG_D_P	cont	8.95	0.003	1.038	1.013 – 1.064
CSMG_D_P	cont	11.4	0.001	1.032	1.013 – 1.051
SPC_D_P	cont	10.7	0.001	1.059	1.023 – 1.097
MPC_D_P	cont	7.18	0.007	1.043	1.011 – 1.076
OC_D_P	cont	5.32	0.021	1.033	1.005 – 1.061
SGL_D_P	cont	7.25	0.007	1.039	1.011 – 1.069
CPG_D_A	cont	8.71	0.003	1.036	1.012 – 1.060
CSMG_D_A	cont	11.5	0.001	1.032	1.013 – 1.051
SPC_D_A	cont	10.7	0.001	1.058	1.028 – 1.094
MPC_D_A	cont	7.32	0.007	1.042	1.012 – 1.074
OC_D_A	cont	4.65	0.031	1.029	1.003 – 1.056
SGL_D_A	cont	6.86	0.009	1.037	1.009 – 1.065
baseline_Q44	ordinal	8.31	0.004	2.08	1.264 – 3.428

Table 7.8: Results of univariate logistic regression analyses with grade 3+ EORTC H&N35 Q44 – taste alteration? as the primary endpoint. Candidate variables only shown.

Variable	Nature	Wald	p-value	OR	95% CI
gender (male)	binary	4.72	0.030	0.43	0.198 – 0.920
weight loss	cont	4.00	0.046	1.065	1.001 – 1.132
weight loss (%)	cont	5.73	0.017	1.073	1.013 – 1.137
N stage	binary	10.7	0.001	2.68	1.486 – 4.833
cisplatin	nominal	4.80	0.028	1.95	1.073 – 3.549
SACT Y/N	binary	4.25	0.039	1.85	1.031 – 3.313
cisplatin Y/N	binary	4.84	0.028	1.89	1.072 – 3.341
IPG_D_P	cont	8.50	0.004	1.046	1.015 – 1.077
CPG_D_P	cont	15.6	<0.001	1.050	1.025 – 1.076
ISMG_D_P	cont	4.91	0.027	1.044	1.005 – 1.085
CSMG_D_P	cont	15.6	<0.001	1.035	1.017 – 1.053
SPC_D_P	cont	17.3	<0.001	1.077	1.040 – 1.115
MPC_D_P	cont	12.6	<0.001	1.063	1.028 – 1.100
OC_D_P	cont	16.1	<0.001	1.063	1.032 – 1.094
SGL_D_P	cont	11.8	0.001	1.052	1.022 – 1.083
IPG_D_A	cont	8.92	0.003	1.045	1.015 – 1.076
CPG_D_A	cont	17.1	<0.001	1.050	1.026 – 1.075
ISMG_D_A	cont	5.04	0.025	1.046	1.006 – 1.088
CSMG_D_A	cont	16.0	<0.001	1.035	1.018 – 1.052
SPC_D_A	cont	17.7	<0.001	1.077	1.040 – 1.115
MPC_D_A	cont	12.7	<0.001	1.061	1.027 – 1.096
OC_D_A	cont	16.4	<0.001	1.062	1.031 – 1.093
SGL_D_A	cont	12.1	0.001	1.053	1.023 – 1.083
baseline_LSXer	ordinal	2.76	0.097	1.46	0.934 – 2.295

Table 7.9: Results of univariate logistic regression analyses with grade 2+ LENT/SOM Salivary gland subjective as the primary endpoint. Candidate variables only shown.

Variable	Nature	Wald	p-value	OR	95% CI
N stage	binary	7.70	0.006	2.34	1.283 – 4.250
cisplatin	nominal	20.0	<0.001	4.50	2.328 – 8.716
SACT Y/N	binary	17.4	<0.001	3.96	2.074 – 7.552
cisplatin Y/N	binary	21.1	<0.001	4.26	2.296 – 7.920
neck diss	binary	2.90	0.089	0.59	0.322 – 1.083
parotidectomy	binary	3.35	0.067	0.24	0.051 – 1.107
IPG_D_P	cont	3.31	0.069	1.026	0.998 – 1.056
CPG_D_P	cont	19.3	<0.001	1.059	1.032 – 1.087
ISMG_D_P	cont	5.60	0.018	1.078	1.013 – 1.147
CSMG_D_P	cont	22.4	<0.001	1.053	1.031 – 1.076
SPC_D_P	cont	17.7	<0.001	1.092	1.048 – 1.137
MPC_D_P	cont	13.4	<0.001	1.086	1.039 – 1.135
OC_D_P	cont	15.9	<0.001	1.068	1.034 – 1.104
SGL_D_P	cont	11.3	0.001	1.059	1.024 – 1.096
IPG_D_A	cont	3.24	0.072	1.025	0.998 – 1.053
CPG_D_A	cont	19.3	<0.001	1.056	1.031 – 1.082
ISMG_D_A	cont	6.19	0.013	1.087	1.018 – 1.162
CSMG_D_A	cont	22.5	<0.001	1.053	1.031 – 1.075
SPC_D_A	cont	18.2	<0.001	1.093	1.049 – 1.139
MPC_D_A	cont	14.3	<0.001	1.087	1.041 – 1.135
OC_D_A	cont	16.0	<0.001	1.067	1.033 – 1.101
SGL_D_A	cont	11.8	0.001	1.063	1.027 – 1.101

Table 7.10: Results of univariate logistic regression analyses with grade 2+ LENT/SOM Salivary gland management as the primary endpoint. Candidate variables only shown.

A number of variables consistently appeared as univariate predictors of many of these toxicity endpoints, with a p-value threshold for significance <0.05. Correction for multiple testing was not done for these analyses, as the primary purpose of this analysis was variable pre-selection for multivariate testing.

Male gender was associated with lower risk of EORTC H&N35 Q41 (OR 0.44, p=0.035) and LENT/SOM subjective xerostomia (0.43, p=0.03), whilst more advanced N-stage was associated with CTCAE xerostomia (OR 2.71, p=0.002), EORTC H&N35 Q41 (OR 2.32, p=0.005), and LENT/SOM subjective xerostomia (OR 2.68, p=0.001) and xerostomia management (OR 2.34, p=0.006). Proportional weight loss during treatment was found to increase the risk of CTCAE xerostomia (OR 1.073, p=0.019), EORTC H&N35 Q41 (OR 1.072,

p=0.017), and LENT/SOM subjective xerostomia (OR 1.073, p=0.017), whilst pre-RT parotidectomy seemed to be protective for CTCAE xerostomia (OR 0.13, p=0.048). Not surprisingly, the use of concomitant systemic therapy also increased the risk of various toxicities.

The use of cisplatin specifically (as a nominal variable) had the strongest associations for dry mouth across endpoints; CTCAE xerostomia (OR 2.99, p=0.001), EORTC H&N35 Q41 (OR 1.85, p=0.044), and LENT/SOM subjective xerostomia (OR 1.95, p=0.028) and xerostomia management (OR 4.50, p<0.001). This was also true for CTCAE dysphagia (OR 2.61, p=0.022). However, Cetuximab use was also independently associated with CTCAE salivary duct inflammation (OR 5.04, p=0.009). For this endpoint, the use of any systemic therapy as a binary variable had a stronger univariate association than the use of cisplatin as a binary variable (OR 3.53, p<0.001 vs OR 2.53, p=0.001), and this was also true of EORTC H&N35 Q42 – have you had sticky saliva? (OR 2.27, p=0.015 vs OR 1.82, p=0.059).

A range of dosimetric parameters were associated with higher risk of all toxicity endpoints. Broadly, patterns were as expected. For planned dose metrics, dose to the contralateral parotid gland had the strongest associations with CTCAE xerostomia (OR 1.083, p<0.001), EORTC Q41 (OR 1.048, p<0.001) and LENT/SOM subjective xerostomia (OR 1.050, p<0.001). Dose to the contralateral submandibular gland had the strongest relationships with CTCAE salivary duct inflammation (OR 1.032, p<0.001) and EORTC H&N35 questions 42 (sticky saliva) (OR 1.039, p<0.001) and 44 (taste disturbance) (OR 1.032, p=0.001). Dose to the superior pharyngeal constrictor muscle had the strongest association with CTCAE dysphagia (OR 1.129, p=0.001). It is also worth noting that across all endpoints, there was a trend for Wald chi-squared statistics, odds ratios and p-values to be marginally stronger for delivered dose metrics than for planned. Examples include dose to the contralateral parotid with CTCAE xerostomia (Wald 28.8, OR 1.083 vs Wald 30.8, OR 1.084), and dose the superior pharyngeal constrictor muscle with CTCAE dysphagia (Wald 10.5, OR 1.129 vs. Wald 10.6, OR 1.130).

Co-linearity diagnostics were run between all parameters classified as candidate variables across all endpoints. The only association that breached the pre-defined threshold was between mean dose to the middle pharyngeal constrictor muscle, and supraglottic larynx, which had a tolerance of 0.092, and a VIF of 10.82. Thus, mean dose values to these 2 structures were never included in the same multivariate analysis. One or the other was selected on a model-by-model basis depending on the stronger univariate relationship. Full details are given in Appendix A5.

7.4.3 Multivariate analyses and model specifics

Full details of all the parameters included in the first step of model production are given in Appendix A5. This includes significance p-values of adding each parameter to the model. Final logistic regression equations after internal validation with bootstrapping, with finalised parameters and coefficients, are shown below for all 8 toxicity endpoints:

Grade 2+ CTCAEv4.03 xerostomia:

- Planned dose: S = 1.263(baseline_xer) + 0.097(CPG_D_P) – 3.546
- Delivered dose: S = 1.307(baseline_xer) + 0.098(CPG_D_A) – 3.688

Grade 2+ CTCAEv4.03 salivary duct inflammation:

- Planned dose: S = 0.918(SACT Y/N) + 0.019(CSMG_D_P) – 1.660
- Delivered dose: S = 0.912(SACT Y/N) + 0.018(CSMG_D_A) – 1.662

Grade 2+ CTCAEv4.03 dysphagia:

- Planned dose: S = 0.004(alcohol units/wk) + 0.086(SPC_D_P) + 0.071(SGL_D_P) – 10.401
- Delivered dose: S = 0.004(alcohol units/wk) + 0.089(SPC_D_A) + 0.068(SGL_D_A) – 10.597

Grade 3+ EORTC H&N35 Q41 – have you had a dry mouth:

- Planned dose: S = 0.037(weight loss %) + 0.052(OC_D_P) + 0.023(CPG_D_P) – 3.535
- Delivered dose: S = 0.036(weight loss %) + 0.051(OC_D_A) + 0.025(CPG_D_A) – 3.569

Grade 3+ EORTC H&N35 Q42 – have you had sticky saliva:

- Planned dose: S = 0.038(CSMG_D_P) – 2.559
- Delivered dose: S = 0.038(CSMG_D_P) – 2.596

Grade 3+ EORTC H&N35 Q44 – have you had taste disturbance:

- Planned dose: S = 0.815(baseline_H&N35_Q44) + 0.033(CSMG_D_P) – 3.072
- Delivered dose: S = 0.815(baseline_H&N35_Q44) + 0.033(CSMG_D_A) – 3.104

Grade 2+ LENT/SOM Salivary gland xerostomia subjective:

- Planned dose: S = 0.696(baseline_LS_SubXer) + 0.041(CPG_D_P) + 0.043(OC_D_P) – 3.534
- Delivered dose: S = 0.714(baseline_LS_SubXer) + 0.043(CPG_D_A) + 0.041(OC_D_A) – 3.560

Grade 2+ LENT/SOM Salivary gland xerostomia management:

- Planned dose: S = 0.912(cisplatin Y/N) + 0.040(CSMG_D_P) – 2.731
- Delivered dose: S = 0.895(cisplatin Y/N) + 0.040(CSMG_D_A) – 2.753

It is worth noting that the inclusion of alcohol consumption as a factor in the CTCAE dysphagia model is a new finding, which has not been reported elsewhere in the literature. Given the absence of a clear causal relationship between alcohol consumption and swallowing dysfunction, and the risk of false positive findings with multiple testing, it may be that this is a spurious finding.

7.4.4 Model performance

Overall performance for both planned and delivered dose-based models was assessed with Nagelkerke's R^2 statistic and -2 log likelihood (-2LL) scores. These data are summarised in Table 7.11. Overall, the best performing models were for CTACE xerostomia and LENT/SOM xerostomia management. It can be seen that higher Nagelkerke R^2 statistics for delivered dose-based models, indicating better model performance, were found for 6/8 toxicity endpoints, with identical performance for 2. Lower -2LL scores also indicate superior model performance, and the same pattern was observed, with lower scores for the same 6/8 endpoints, and identical scores for 2.

Toxicity endpoint	Planned dose models		Delivered dose models	
	Nagelkerke's R^2	-2LL	Nagelkerke's R^2	-2LL
CTCAE xerostomia	0.316	211	0.336	207
CTCAE SDI	0.137	245	0.137	245
CTCAE dysphagia	0.195	150	0.203	149
EORTC H&N35 Q41	0.171	237	0.180	236
EORTC H&N35 Q42	0.111	224	0.112	224
EORTC H&N35 Q44	0.149	234	0.151	233
LENT/SOM Subj Xer	0.200	242	0.210	240
LENT/SOM Man Xer	0.240	225	0.240	225

Table 7.11: NTCP model performance summary, for planned and delivered dose-based models. (-2LL = -2 log likelihood statistic)

7.4.5 Model discrimination

The capacity of models to predict which patients would, and would not develop the toxicity endpoint of interest was tested with ROC curves, and discrimination slope scores. Results for AUCs of ROC curves are summarised in Table 7.12. ROC curves for both planned and delivered dose-based models for each endpoint are shown in Figure 7.3 – 7.10. These data are summarised in a Forrest plot in Figure 7.11.

Toxicity endpoint	Planned dose models		Delivered dose models	
	AUC	95% CI	AUC	95% CI
CTCAE xerostomia	0.787	0.723 – 0.851	0.801	0.739 – 0.863
CTCAE SDI	0.670	0.595 – 0.746	0.668	0.593 – 0.744
CTCAE dysphagia	0.749	0.666 – 0.831	0.752	0.668 – 0.836
EORTC H&N35 Q41	0.690	0.616 – 0.764	0.697	0.624 – 0.771
EORTC H&N35 Q42	0.636	0.557 – 0.715	0.641	0.563 – 0.720
EORTC H&N35 Q44	0.681	0.604 – 0.758	0.682	0.605 – 0.758
LENT/SOM Subj Xer	0.735	0.665 – 0.804	0.741	0.673 – 0.810
LENT/SOM Man Xer	0.735	0.665 – 0.805	0.735	0.665 – 0.804

Table 7.12: Areas under ROC curves, with 95% confidence intervals, for both planned and delivered dose-based NTCP models for all 8 toxicity endpoint

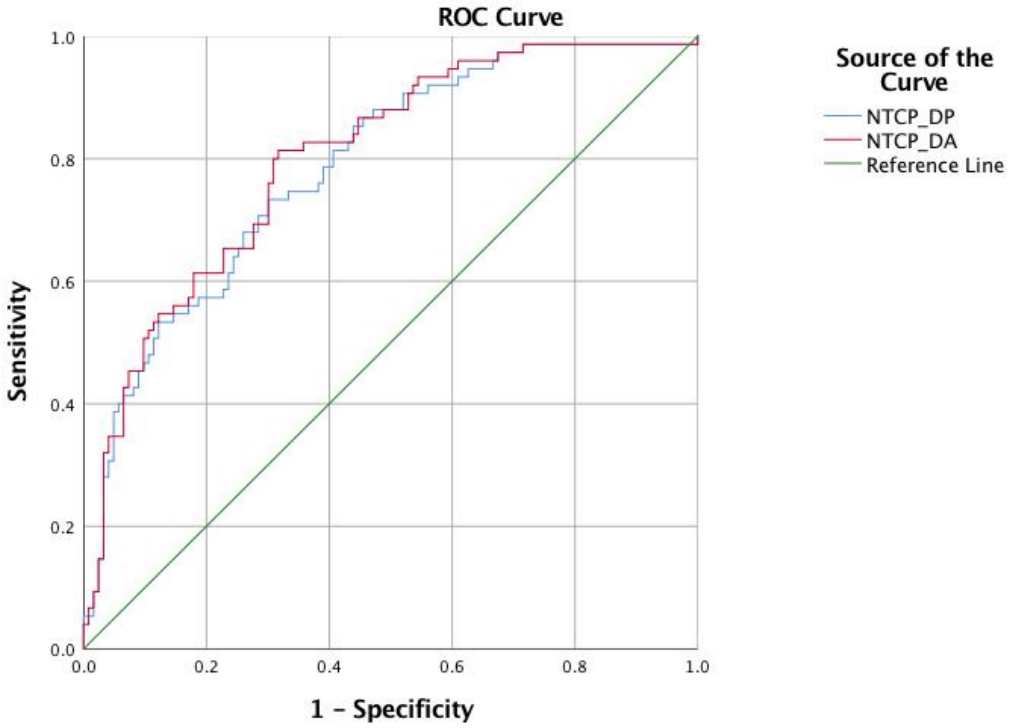


Figure 7.3: Receiver operator characteristic curves for both planned and delivered dose-based NTCP models, for grade 2+ CTCAE v4.03 xerostomia

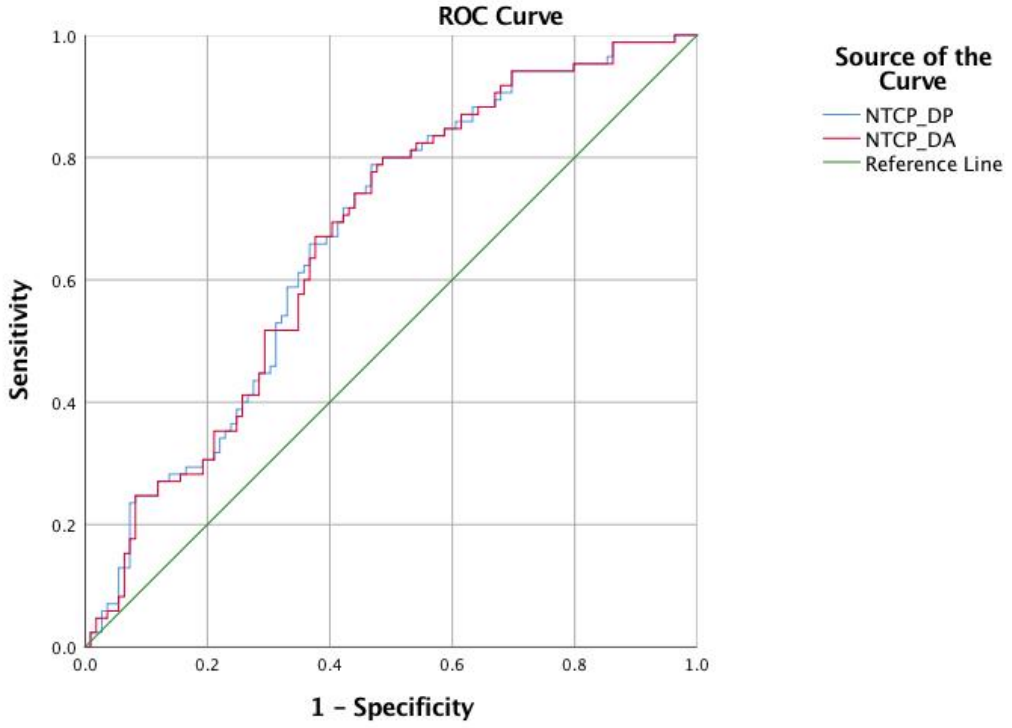


Figure 7.4: Receiver operator characteristic curves for both planned and delivered dose-based NTCP models, for grade 2+ CTCAE v4.03 salivary duct inflammation.

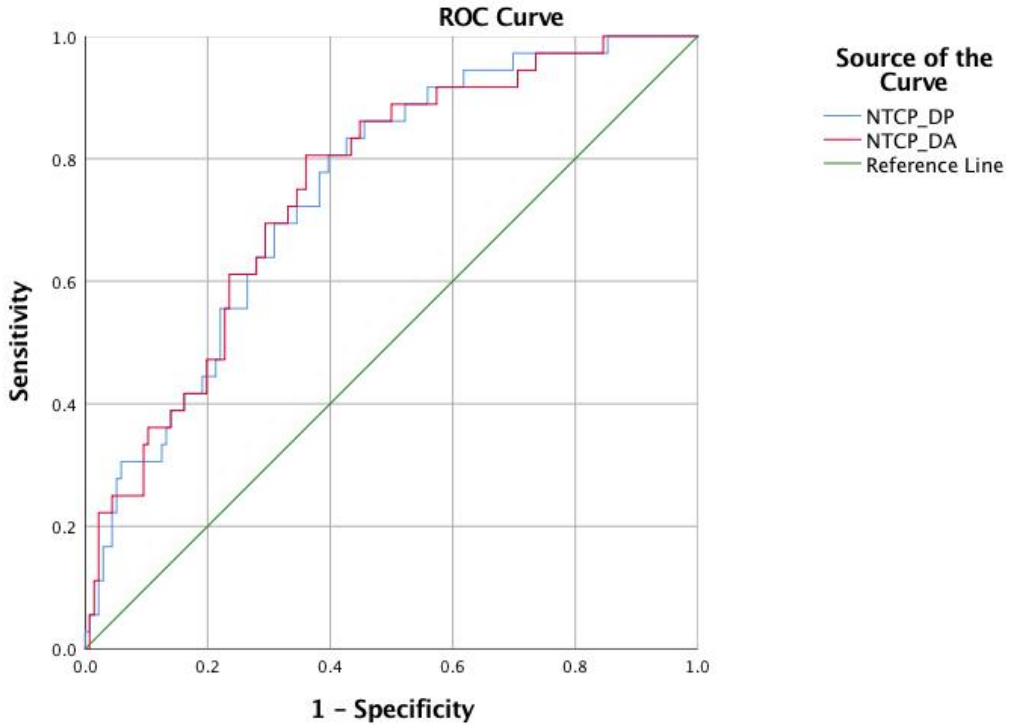


Figure 7.5: Receiver operator characteristic curves for both planned and delivered dose-based NTCP models, for grade 2+ CTCAE v4.03 dysphagia.

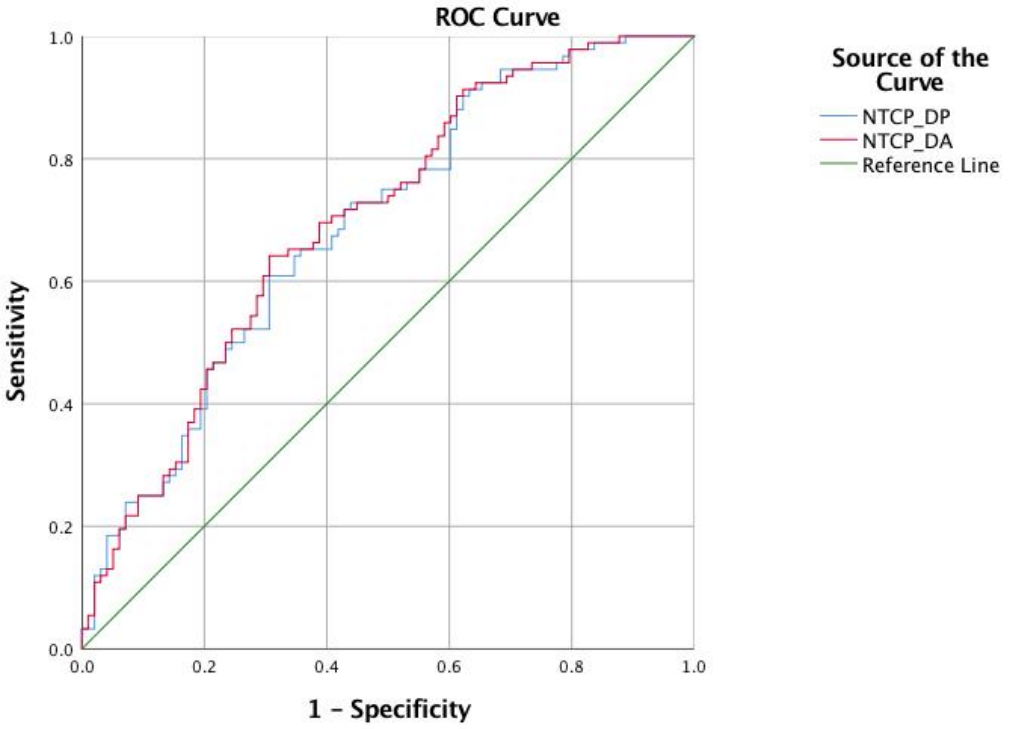


Figure 7.6: Receiver operator characteristic curves for both planned and delivered dose-based NTCP models, for grade 3+ EORTC H&N 35 Q41 – have you had a dry mouth?

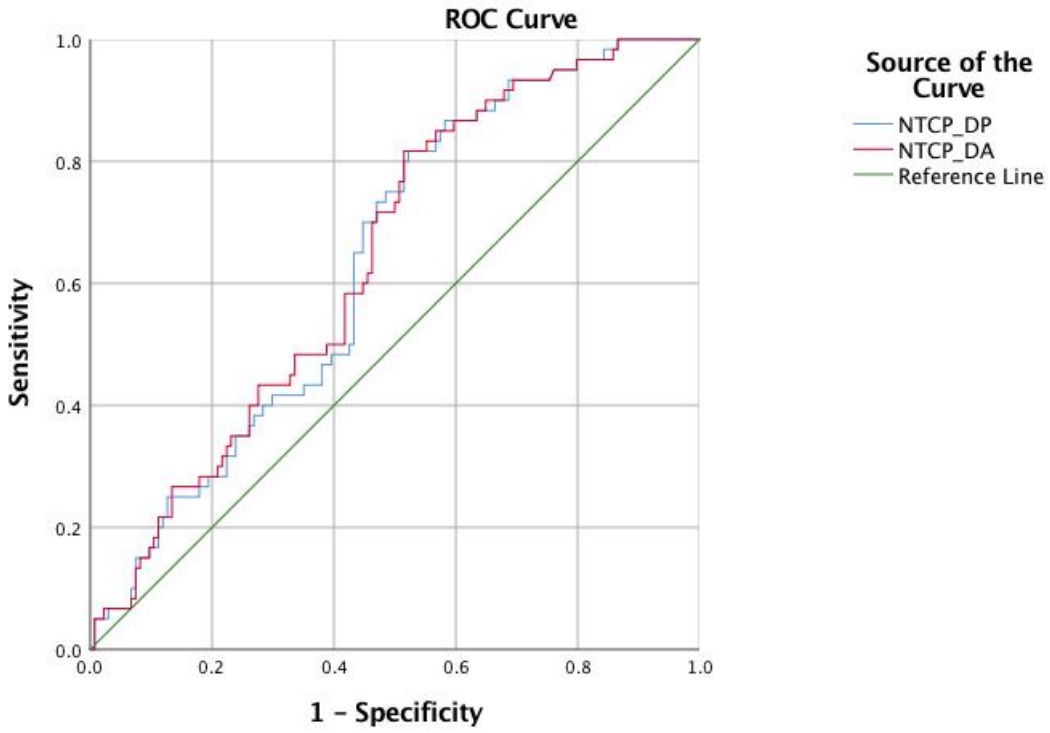


Figure 7.7: Receiver operator characteristic curves for both planned and delivered dose-based NTCP models, for grade 3+ EORTC H&N 35 Q42 – have you had sticky saliva?

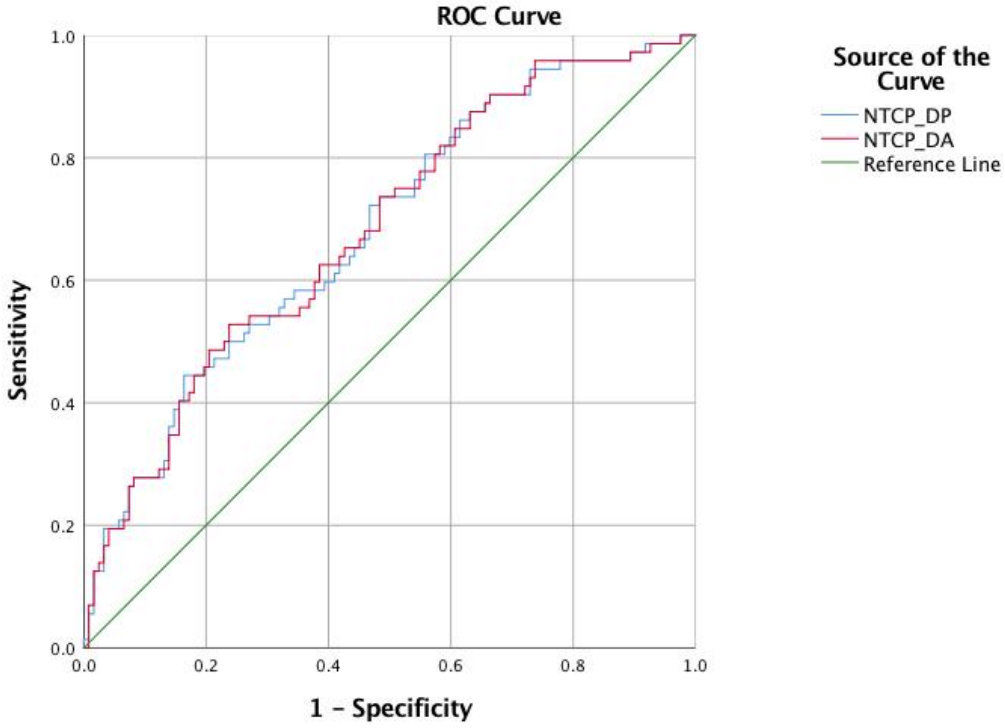


Figure 7.8: Receiver operator characteristic curves for both planned and delivered dose-based NTCP models, for grade 3+ EORTC H&N 35 Q44 – have you had taste disturbance?

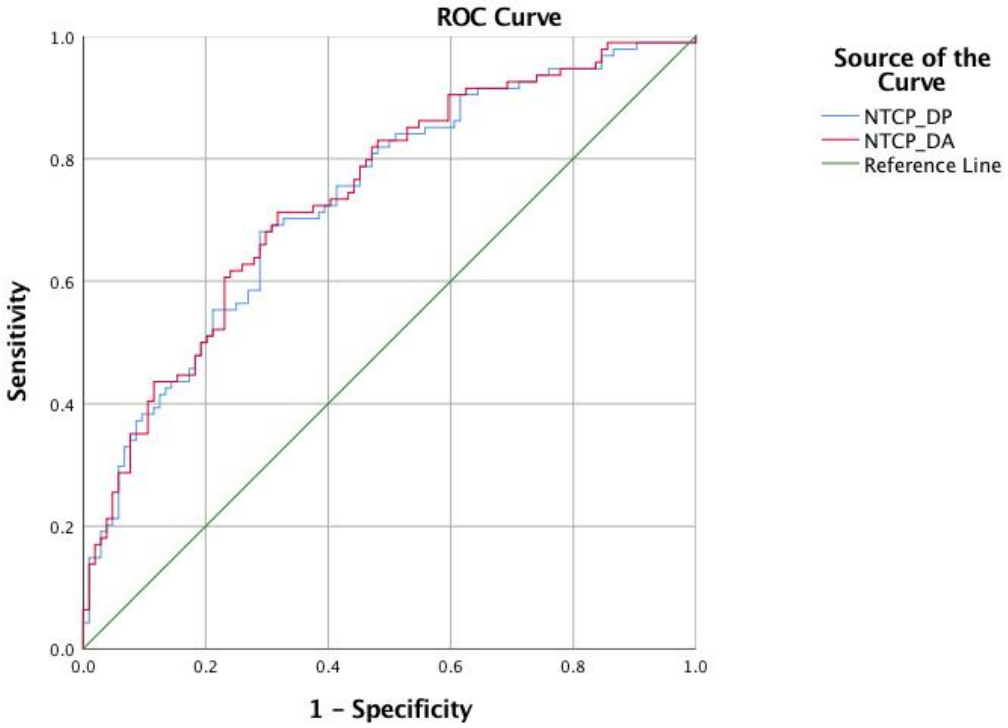


Figure 7.9: Receiver operator characteristic curves for both planned and delivered dose-based NTCP models, for grade 2+ LENT/SOM salivary gland subjective xerostomia

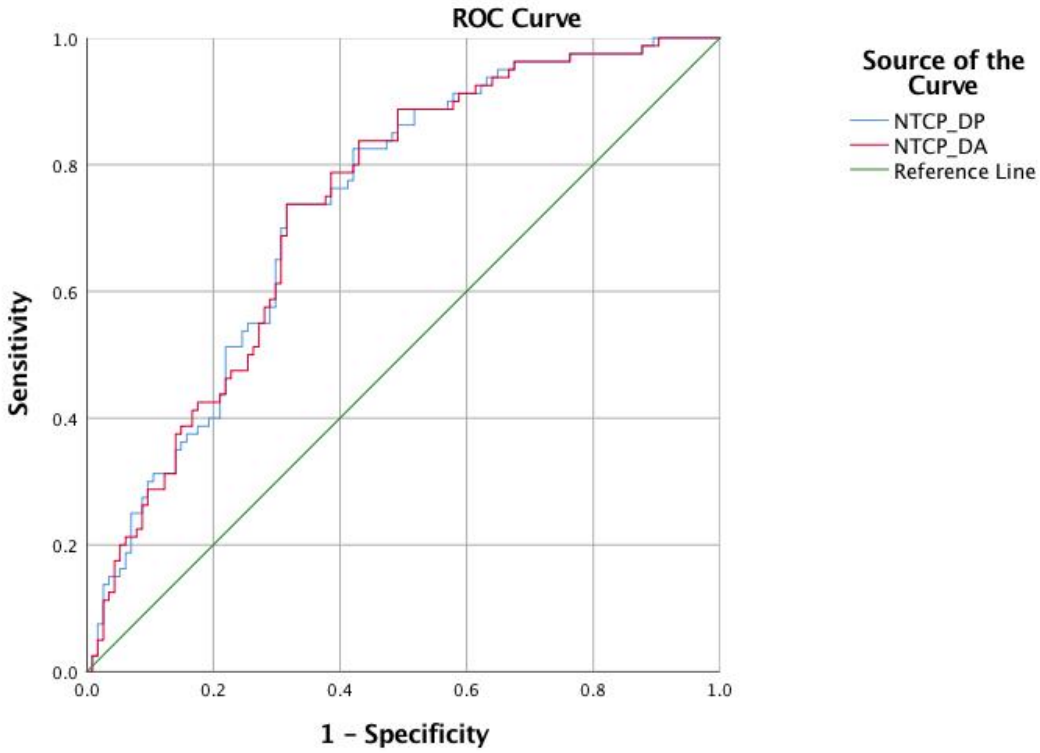


Figure 7.10: Receiver operator characteristic curves for both planned and delivered dose-based NTCP models, for grade 1+ LENT/SOM salivary gland subjective xerostomia

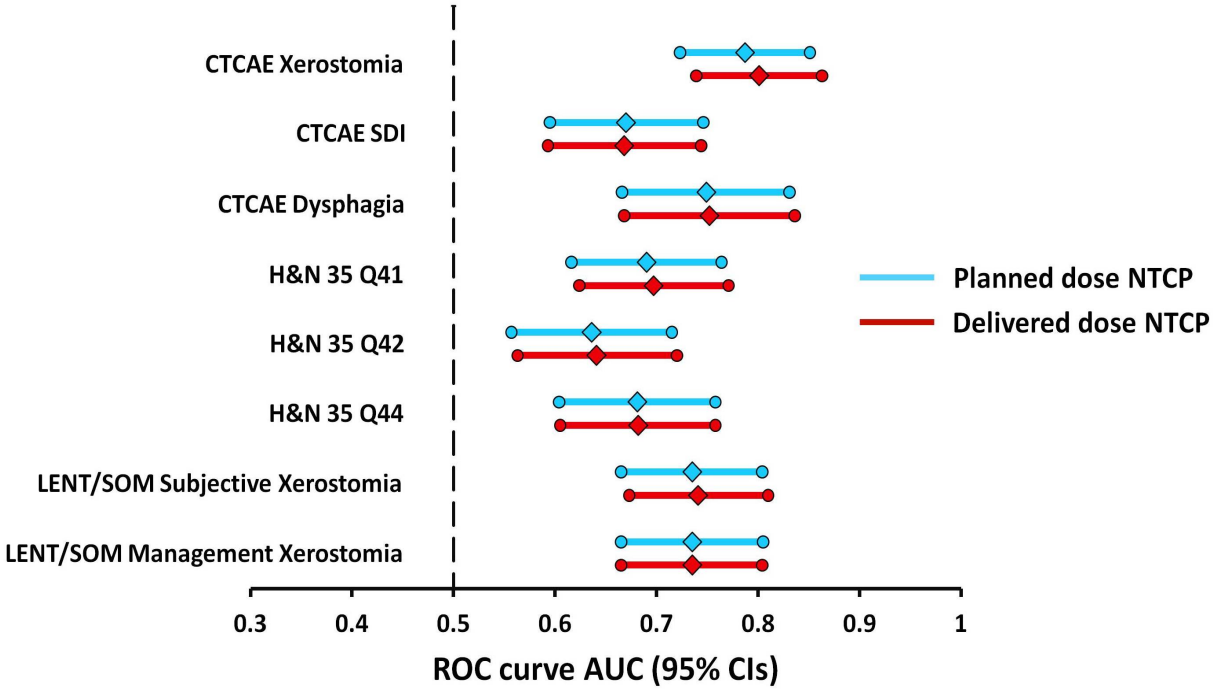


Figure 7.11: Forrest plot of ROC curve AUCs and 95% confidence intervals for both planned and delivered dose-based NTCP models.

Discrimination slope results for both planned and delivered dose-based NTCP models for all 8 toxicity endpoints are summarised in Table 7.13, and shown as a clustered bar chart in shown in Figure 7.12.

Toxicity endpoint	Planned dose models	Delivered dose models
	Disc. slope	Disc. slope
CTCAE xerostomia	0.241 (±0.031)	0.259 (±0.032)
CTCAE SDI	0.103 (±0.022)	0.104 (±0.022)
CTCAE dysphagia	0.057 (±0.020)	0.058 (±0.02)
EORTC H&N35 Q41	0.126 (±0.024)	0.128 (±0.024)
EORTC H&N35 Q42	0.071 (±0.016)	0.073 (±0.016)
EORTC H&N35 Q44	0.110 (±0.023)	0.111 (±0.023)
LENT/SOM Subj Xer	0.154 (±0.025)	0.162 (±0.026)
LENT/SOM Man Xer	0.176 (±0.026)	0.178 (±0.026)

Table 7.13: Discrimination slope results for planned and delivered dose-based NTCP models for all 8 toxicity endpoints. Standard errors shown in parentheses.

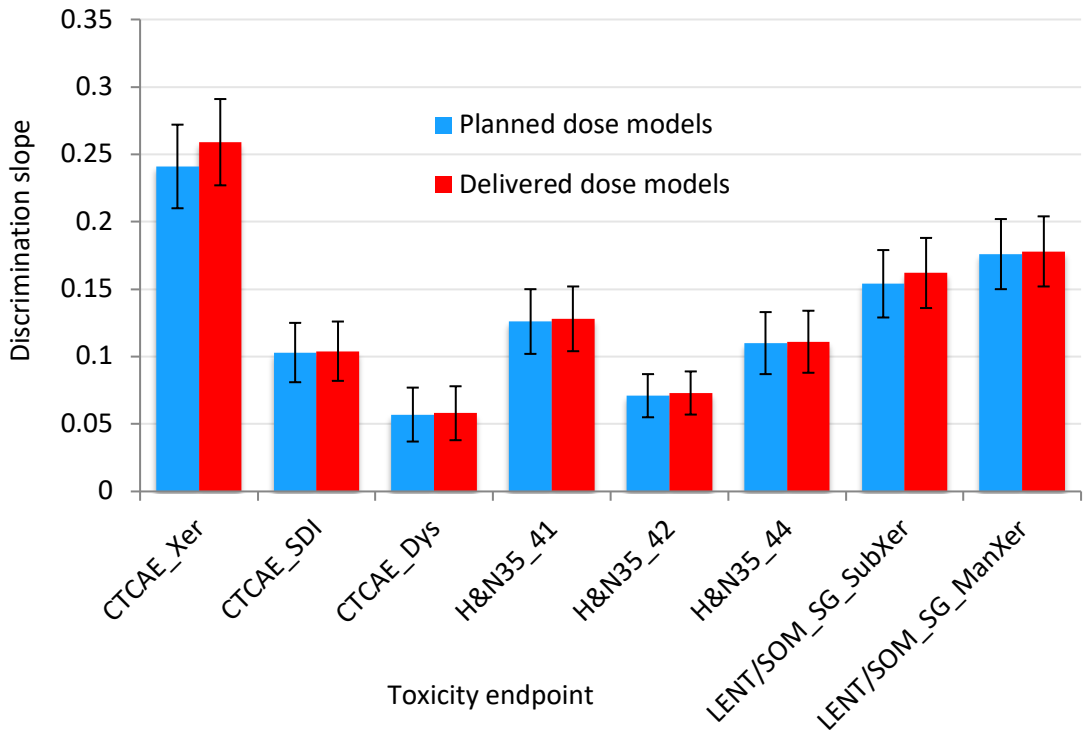


Figure 7.12: Clustered bar chart of discrimination slope results (with standard errors) for both planned and delivered dose-based NTCP models, for all 8 toxicity endpoints.

Similar trends to those in the multivariate model performance summary results are seen. The best performing models with both planned and delivered dose metrics, using AUC of ROC curves and discrimination slope data, were for CTCAE xerostomia, and LENT/SOM subjective xerostomia, and xerostomia management. Models predicting more severe symptoms of thick, sticky saliva using both CTCAE (salivary duct inflammation) and EORTC H&N 35 (Q42) were less discriminatory. A trend observed in result of the univariate analysis was also seen with multivariate models. Although differences are minimal, AUCs were superior for delivered dose-based models for 6/8 toxicity endpoints, whilst discrimination slope results suggested marginally better capacity to distinguish cases from controls for all 8 delivered dose models.

7.4.6 Model calibration

Calibration data for all models are shown in Table 7.14. Overall model calibration is summarised by the Hosmer-Lemeshow (HL) chi-squared statistic, and p-value. A lower HL score, and higher p-value indicate better calibration. Data for calibration slope, intercept and Brier score are also presented. Perfect model calibration would result in a calibration slop of 1, intercept of 0, and Brier score of 0.

Toxicity endpoint	Planned dose models					Delivered dose models				
	HL χ^2	p-val	CS	Int	BS	HL χ^2	p-val	CS	Int	BS
CTCAE xerostomia	2.83	0.95	1.02	-0.01	<0.01	7.91	0.44	1.02	-0.01	<0.01
CTCAE SDI	5.44	0.71	1.02	-0.01	<0.01	7.81	0.45	1.03	-0.01	<0.01
CTCAE dysphagia	2.68	0.95	1.08	-0.02	<0.01	4.34	0.83	1.06	-0.01	<0.01
EORTC H&N35_Q41	10.2	0.25	0.96	0.02	0.13	4.4	0.82	0.94	0.03	<0.01
EORTC H&N35_Q41	8.01	0.43	0.95	0.02	<0.01	10.4	0.24	0.96	0.01	0.01
EORTC H&N35_Q41	3.62	0.89	0.98	0.01	<0.01	5.39	0.72	0.97	0.01	<0.01
LENT/SOM Subj Xer	9.51	0.30	1.02	-0.01	0.01	6.00	0.65	1.04	-0.02	<0.01
LENT/SOM Man Xer	4.09	0.85	0.97	0.01	<0.01	5.54	0.70	0.97	0.01	<0.01

Table 7.14: Calibration results for both planned and delivered dose-based models. HL χ^2 = Hosmer Lemeshow chi-squared statistic, p-val = p-value, CS = calibration slope, Int = intercept, BS = brier score.

Overall, model calibration results were reassuring; all Hosmer-Lemeshow tests were negative (all p-values >0.2), and all Brier scores were <0.25, indicating adequate calibration of all models.

7.5 Discussion

7.5.1 Toxicity data

Toxicity endpoints were selected to reflect moderate to severe symptoms 12 months after radiotherapy completed. Although a number of NTCP modelling studies have used 6-month data [8-10, 26], others have used 12-month results [25], and this seemed a pragmatic compromise between peak event rate and a slightly larger sample with 6-month data, and the more persistent lasting symptoms seen at later timepoints. Where possible, the thresholds chosen to define presence of a specific toxicity were based on literature precedent. For example, the threshold of grade 3+ for EORTC H&N 35 questions 41 and 42 was based on the work of Beetz et al [8].

An exception was the choice of grade 2+ CTCAEv4.03 dysphagia. The NTCP model of Wopken and colleagues predicts the probability of on-going PEG-tube dependence at 6 months [10]. However, this endpoint has 2 disadvantages. Firstly, a low event rate of 10.7%, partially compensated for by a larger sample size of 355 patients. Secondly, whilst this threshold improves the objectivity of the endpoint, it does not include a number of patients with significant, life-altering symptoms. Of note, an earlier analysis of the same cohort was based on grade 2+ toxicity outcomes [9]. Grade 2 CTCAEv4.03 dysphagia is defined as "symptomatic and altered eating/swallowing". Therefore, this threshold was chosen to ensure this degree of toxicity was being measured, and to optimise the event rate for analysis. Where no precedent existed, selected thresholds were also based on a balance between sufficient event-rate for robust model training and validation, and ensuring that the predicted phenotype was clinically relevant. For example, grade 2 LENT/SOM salivary gland xerostomia management is defined as "Occasional saliva substitute, sugarless candy or gum, sialagogues". Therefore, any patient with this grade (or worse) requires some degree of active management to cope with their xerostomia, with resultant lifestyle consequences.

As discussed in Chapter 3, toxicity results presented here are similar to published data from similar cohorts, and in clinical trials [4]. Importantly, event-rates for many of the binary toxicity endpoints chosen were comparable to those in previous NTCP-modelling work. For example, the 6-month event rates used for the Beetz xerostomia model were very similar to data in this cohort [8]. This is important, as it makes metrics for model performance, discrimination and validation seen in this work more comparable with literature precedent. The discrepancy with the event-rate in the Wopken dysphagia model is an exception [10], but the results of the earlier Christianen analyses are more directly comparable, as they were based on grade 2+ events [9].

7.5.2 Univariate analysis

A number of baseline demographic, disease and treatment factors were associated with all the toxicity endpoints assessed. Percentage weight loss as a continuous variable was associated with CTCAEv4.03 xerostomia, EORTC H&N 35 dry mouth and sticky saliva and LENT/SOM subjective xerostomia. Previous NTCP models for xerostomia and sticky saliva did not find this effect [8], although this may be because weight loss data were not available for analysis. Weight loss was previously found to be a univariate predictor of dysphagia by Wopken et al [10]. There are 2 plausible mechanisms for this effect. The first is to hypothesise that patients treated with large high-dose volumes, bilateral neck treatment and concomitant systemic therapy would develop more severe acute side effects earlier in treatment, leading to poor

alimentation and greater weight loss. These patients would also be likely to receive higher doses to the parotid glands. Another is that weight loss leads to higher delivered doses to the parotid glands, although data in Chapter 6 show that the relationship between weight loss and CPG dose differences are non-significant.

The data in this Chapter suggest that men were less likely to develop xerostomia and thick sticky saliva, although these relationships were all of borderline significance. This effect was not observed in previous NTCP models [8-10], and there is no clearly plausible mechanism to explain it. If anything the contrary might be expected – with men (relatively) more likely to have SCC's, hence higher absolute doses and bilateral neck treatment, and women more likely to have salivary gland primaries, with lower doses and unilateral neck treatment. The fact that parotidectomy was *protective* for CTCAEv4.03 xerostomia and LENT/SOM xerostomia management strongly suggests that this cannot be the mechanism. The most likely explanation therefore is that this is a statistical anomaly, due to the substantial gender imbalance in the sample (83% men, 17% women), an imbalance that is broadly similar to that observed in similar NTCP modelling work [8, 10, 25].

Staging consistently appears as a univariate predictor of toxicity outcomes, with N-stage specifically associated with xerostomia, thick sticky saliva and dysphagia [9, 10, 25]. This effect was also observed in these data, where strong univariate associations between more advanced N-stage, and higher probabilities of moderate to severe xerostomia and dysphagia were observed. A similar pattern was seen for baseline symptoms. The presence or absence of symptoms before radiotherapy consistently predicts subsequent toxicity [8-10, 25, 26], and again this finding was replicated here, where strong univariate associations between baseline symptoms and xerostomia, salivary duct inflammation and taste disturbance were observed.

On univariate analysis, the use of concomitant systemic therapy increased the risk of all toxicity endpoints. This is entirely in keeping with the literature, which has consistently reported the same effect [30, 41-43]. Interestingly however, there may be some deeper subtleties apparent from the data, partly limited by the small patient numbers treated with cetuximab. The use of cisplatin specifically – as a nominal variable - has the strongest univariate association with CTAEv4.03 xerostomia and dysphagia, and LENT/SOM xerostomia scores. Odds ratios for the use of any systemic therapy (including cetuximab) are still significant but weaker. Furthermore, the use of cetuximab as a nominal variable was not significantly associated with toxicity for any of these endpoints. This suggests that using cisplatin as a concomitant therapy increases the risk of moderate to severe late toxicity for these endpoints more than cetuximab, a finding not substantiated by 2 large recent RCTs [32, 44].

In contrast, the opposite effect is seen for CTCAEv4.03 salivary duct inflammation. The use of cetuximab has an odds ratio of 5.04, higher than that for cisplatin, and the strongest univariate effect for concomitant therapy is the use of any systemic therapy as a binary variable. The same is true for EORTC H&N35 sticky saliva. This infers that whilst cisplatin confers the greatest risk for xerostomia and dysphagia, cetuximab use may be particularly associated with an increased risk of this very specific toxicity. It should be restated however that the number of patients receiving cetuximab was small (14/198), therefore these data are no more than suggestive.

Strong relationships between OAR dose and side effects were observed for most endpoints. However, 2 important trends emerge. Firstly, when looking at planned dose metrics, the strongest univariate relationships between OAR doses and most toxicity endpoints were those expected, either a-priori, or from the literature [8-10, 45-48]. Examples include CPG dose for CTCAEv4.03 xerostomia, CSMG dose for salivary duct inflammation, and SPC dose for dysphagia. Secondly, direct comparison of univariate relationships for equivalent planned and delivered OAR doses consistently show very marginally stronger associations with delivered dose metrics. For all 8 endpoints, the Wald chi-squared statistic for the strongest univariate OAR planned dose-response relationship is lower than the delivered dose equivalent.

However, it is important to emphasise that absolute differences between univariate toxicity relationships for both planned and delivered dose were very small, and none were statistically significant. In some cases odds ratios were identical, and in all cases confidence intervals overlapped. Given the meagre differences between planned and delivered dose in the cohort, this result is not surprising. Therefore, whilst the result is interesting and suggestive, it cannot be definitively concluded that delivered dose is a better univariate predictor of toxicity.

7.5.3 Multivariate models

All NTCP models were relatively simple, and the maximum number of terms in any model (excluding constants) was 3. This is reassuring as it reduced the likelihood of major over-fitting to training data, and the recommended limit of ≥10 events per variable was respected for all models [49]. Robust model validation is also crucial to reduce the risk of over-fitting [50-52]. The TRIPOD statement describes 4 levels of quality and robustness in prognostic modelling [50]. Level 1 models are produced from a single dataset with no validation attempted. In level 2 analyses, internal validation is done, whilst level 3 and 4 involve genuine external validation

on datasets that are entirely separate to the training cohort. As the NTCP models produced in this work included delivered dose metrics, and this is the first study to do so, true external validation was, by definition, not possible. In such circumstances, there has been a vogue to split a single dataset into training and validation cohorts. However, the TRIPOD consensus statement on prognostic modelling states that "randomly splitting a dataset into training and validation cohorts is thought to be external validation – but is in fact a weak form of internal validation" [50]. This is equivalent to the example of an erroneous type 2 analysis in the opinion piece of Zwanenburg and Löck [51]. The TRIPOD statement also advises that internal validation methods include calibration and discrimination, with boot-strapping and cross-validation as expected approaches for internal validation [50]. The methodology used in this work was designed to follow these consensus guidelines as closely as possible.

The parameters included in the planned dose NTCP models were similar to previous publications in many cases. For example, the CTCAEv4.03 and LENT/SOM subjective xerostomia models included only baseline symptoms and mean dose to the contralateral parotid gland, identical to the Beetz xerostomia model [8] (although it should be reiterated that the endpoint for this analysis was EORTC H&N35 Q41). A study by Van Dijk and colleagues sought to improve xerostomia prediction by incorporating CT texture features of the parotid glands into their models, which also included baseline toxicity scores [25]. NTCP values for the Beetz sticky saliva model were most affected by mean dose to the contralateral submandibular gland, as were the CTCAEv4.03 SDI and EORTC H&N35 Q42 models in this work. The dysphagia regression equation in this work included mean dose to the superior pharyngeal constrictor muscle, as did the models of both Christianen [9], and Wopken et al [10]. As noted in section 7.4.3, this model also included alcohol consumption, which may have been a spurious result.

Model performance was satisfactory. All planned and delivered dose models were significant predictors of toxicity outcomes (ROC curve AUC lower 95% CI > 0.5 in all cases). Planned dose-based model performance and discrimination scores were consistent with previously published results in some cases, and slightly inferior in others. Xerostomia prediction using EORTC H&N35 Q41 in this work is summarised by a ROC curve AUC and discrimination slope of 0.69 and 0.13, very similar to values of 0.68 and 0.10 in the Beetz model [8]. For sticky saliva the model in this work was slightly inferior to the Beetz equivalent [8]; ROC curve AUCs and discrimination slopes were 0.64 vs. 0.70 and 0.07 vs. 0.12 respectively. Interestingly, performance for both CTCAEv4.03 and LENT/SOM xerostomia models was superior to the EORTC endpoint in this work. More recent EORTC endpoint based xerostomia and sticky saliva models from Van Dijk et al reported respective AUCs of 0.75 and 0.74, more similar to

the CTCAEv4.03 and LENT/SOM based results reported here. ROC curve AUC and discrimination slope scores for the CTCAEv4.03 dysphagia model in this work (0.75 and 0.06) are lower than those seen in similar work by Christianen and Wopken et al [9, 10], but it should be noted that these studies had access to a larger sample of patients (~350).

Similar to the univariate analysis, the NTCP model results suggest that delivered dose-based multivariate toxicity prediction models consistently perform marginally better than planned dose-based equivalents. In 6/8 cases, both Nagelkerke R^2 and -2LL statistics for delivered dose models were slightly better than their planned dose equivalents, and in 2/8 cases they were identical. This is despite the fact that preliminary model building and training was done with planned dose metrics. Delivered dose-based models had higher ROC curve AUCs in 6/8 cases, with 1 identical result, and 1 case where the planned dose model was marginally superior. The discrimination slope result was better for all 8 delivered dose models. Calibration results were satisfactory in all cases, with consistently close alignment between predicted and observed event rates for both planned and delivered dose based models.

However, it is important to reiterate that all these differences were very small, with substantial overlap of all confidence intervals. Overall model performance is similar for both planned and delivered dose models, and no statistical significance of the differences that were observed was sought or tested for. However, this was entirely expected. As shown in Chapter 6, mean differences between planned and delivered dose across this patient sample were relatively small, and it was therefore highly unlikely that large differences in model performance would be seen. There is literature precedent for this approach, observation and subsequent conclusions. In the Van Dijk study of parotid gland CT IMBs, the authors concluded that including textural features improved predictive performance on the basis of ROC curve AUCs for xerostomia and sticky saliva respectively improving from 0.75 to 0.77 and 0.74 to 0.77, and discrimination slope scores of 0.19 to 0.21 and 0.15 to 0.18 [25]. In all cases 95% confidence intervals overlapped substantially.

These data therefore suggest that whilst mean dose differences to H&N OARs across a population may be small, they may be significant enough to cause a very subtle effect in the predictive capabilities of predictive toxicity models. The very similar performance of planned and delivered dose-based models shows that they predict the same outcome for most patients. It is only in the relatively rare cases where dose differences are substantial that model predictions will be noticeably different. Therefore, if the objective of ART is toxicity reduction, careful patient selection is crucial.

7.6 Conclusions

In this work, relationships between mean planned and delivered radiotherapy dose to OARs crucial for saliva production and swallowing, and patient toxicity measured with 3 scoring systems and 8 discrete endpoints, were assessed. On both univariate and multivariate analysis, delivered dose metrics were marginally superior predictors of toxicity outcomes than planned dose equivalents, but differences were very small, and not statistically significant. Given the small OAR D_P versus D_A dose deltas at a population level, such subtle differences in toxicity prediction are not surprising emphasising that many patients do not need ART, and careful patient selection is crucial to identify those that do.

7.7 References:

[1] Ho KF, Marchant T, Moore C, Webster G, Rowbottom C, Penington H, et al. Monitoring dosimetric impact of weight loss with kilovoltage (kV) cone beam CT (CBCT) during parotid-sparing IMRT and concurrent chemotherapy. Int J Radiat Oncol Biol Phys. 2012;82:e375-82.
[2] Brown E, Owen R, Harden F, Mengersen K, Oestreich K, Houghton W, et al. Predicting the need for adaptive radiotherapy in head and neck cancer. Radiother Oncol. 2015;116:57-63.
[3] Brouwer CL, Steenbakkers RJ, van der Schaaf A, Sopacua CT, van Dijk LV, Kierkels RG, et al. Selection of head and neck cancer patients for adaptive radiotherapy to decrease xerostomia. Radiother Oncol. 2016;120:36-40.
[4] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol. 2011;12:127-36.
[5] Cmelak A, Li SL, Marur S, Zhao WQ, Westra WH, Chung CH, et al. Symptom reduction from IMRT dose deintensification: Results from ECOG 1308 using the Vanderbilt Head and Neck Symptom Survey version 2 (VHNSS V2). Journal of Clinical Oncology. 2015;33.
[6] Kelly JR, Husain ZA, Burtness B. Treatment de-intensification strategies for head and neck cancer. European Journal of Cancer. 2016;68:125-33.
[7] Navran A, Heemsbergen W, Janssen T, Hamming-Vrieze O, Jonker M, Zuur C, et al. The impact of margin reduction on outcome and toxicity in head and neck cancer patients treated with image-guided volumetric modulated arc therapy (VMAT). Radiother Oncol. 2019;130:25-31.
[8] Beetz I, Schilstra C, van der Schaaf A, van den Heuvel ER, Doornaert P, van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. Radiother Oncol. 2012;105:101-6.
[9] Christianen MEMC, Schilstra C, Beetz I, Muijs CT, Chouvalova O, Burlage FR, et al. Predictive modelling for swallowing dysfunction after primary (chemo)radiation: Results of a prospective observational study. Radiotherapy and Oncology. 2012;105:107-14.
[10] Wopken K, Bijl HP, van der Schaaf A, van der Laan HP, Chouvalova O, Steenbakkers RJ, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol. 2014;113:95-101.
[11] Castelli J, Simon A, Louvel G, Henry O, Chajon E, Nassef M, et al. Impact of head and neck cancer adaptive radiotherapy to spare the parotid glands and decrease the risk of xerostomia. Radiat Oncol. 2015;10:6.
[12] Owadally W, Hurt C, Timmins H, Parsons E, Townsend S, Patterson J, 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;15:602.
[13] Schwartz DL, Garden AS, Thomas J, Chen Y, Zhang Y, Lewin J, et al. Adaptive radiotherapy for head-and-neck cancer: initial clinical outcomes from a prospective trial. Int J Radiat Oncol Biol Phys. 2012;83:986-93.
[14] Chen AM, Daly ME, Cui J, Mathai M, Benedict S, Purdy JA. Clinical outcomes among patients with head and neck cancer treated by intensity-modulated radiotherapy with and without adaptive replanning. Head Neck. 2014;36:1541-6.
[15] Kataria T, Gupta D, Goyal S, Bisht SS, Basu T, Abhishek A, et al. Clinical outcomes of adaptive radiotherapy in head and neck cancers. Br J Radiol. 2016;89:20160085.
[16] Heukelom J, Fuller CD. Head and Neck Cancer Adaptive Radiation Therapy (ART): Conceptual Considerations for the Informed Clinician. Semin Radiat Oncol. 2019;29:258-73.
[17] Masterson L, Moualed D, Liu ZW, Howard JEF, Dwivedi RC, Tysome JR, et al. De-escalation treatment protocols for human papillomavirus-associated oropharyngeal squamous cell carcinoma: A systematic review and meta-analysis of current clinical trials. European Journal of Cancer. 2014;50:2636-48.
[18] Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tome WA. Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. Int J Radiat Oncol Biol Phys. 2010;76:S135-9.
[19] Shelley LEA, Scaife JE, Romanchikova M, Harrison K, Forman JR, Bates AM, et al. Delivered dose can be a better predictor of rectal toxicity than planned dose in prostate radiotherapy. Radiother Oncol. 2017;123:466-71.
[20] Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys. 2010;76:S10-9.

[21] Xu CJ, van der Schaaf A, Schilstra C, Langendijk JA, van't Veld AA. Impact of Statistical Learning Methods on the Predictive Power of Multivariate Normal Tissue Complication Probability Models. Int J Radiat Oncol. 2012;82:E677-E84.
[22] Lyman JT. Complication Probability as Assessed from Dose Volume Histograms. Radiation Research. 1985;104:S13-S9.
[23] Jackson A, Kutcher GJ, Yorke ED. Probability of Radiation-Induced Complications for Normal-Tissues with Parallel Architecture Subject to Nonuniform Irradiation. Medical Physics. 1993;20:613-25.
[24] El Naqa I, Bradley J, Blanco AI, Lindsay PE, Vicic M, Hope A, et al. Multivariable modeling of radiotherapy outcomes, including dose-volume and clinical factors. Int J Radiat Oncol Biol Phys. 2006;64:1275-86.
[25] van Dijk LV, Brouwer CL, van der Schaaf A, Burgerhof JG, Beukinga RJ, Langendijk JA, et al. CT image biomarkers to improve patient-specific prediction of radiation-induced xerostomia and sticky saliva. Radiother Oncol. 2017;122:185-91.
[26] Beasley W, Thor M, McWilliam A, Green A, Mackay R, Slevin N, et al. Image-based Data Mining to Probe Dosimetric Correlates of Radiation-induced Trismus. Int J Radiat Oncol. 2018;102:1330-8.
[27] Christianen ME, Langendijk JA, Westerlaan HE, van de Water TA, Bijl HP. Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. Radiother Oncol. 2011;101:394-402.
[28] Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, 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;117:83-90.
[29] Thomas SJ, Eyre KR, Tudor GS, Fairfoul J. Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube. Med Phys. 2012;39:160-7.
[30] Denis F, Garaud P, Bardet E, Alfonsi M, Sire C, Germain T, et al. Final results of the 94-01 French Head and Neck Oncology and Radiotherapy Group randomized trial comparing radiotherapy alone with concomitant radiochemotherapy in advanced-stage oropharynx carcinoma. J Clin Oncol. 2004;22:69-76.
[31] Bonner JA, Harari PM, Giralt J, Azarnia N, Shin DM, Cohen RB, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. N Engl J Med. 2006;354:567-78.
[32] Mehanna H, Robinson M, Hartley A, Kong A, Foran B, Fulton-Lieuw T, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. Lancet. 2019;393:51-60.
[33] Barnett GC, West CM, Dunning AM, Elliott RM, Coles CE, Pharoah PD, et al. Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nat Rev Cancer. 2009;9:134-42.
[34] Schaake W, van der Schaaf A, van Dijk LV, Bongaerts AHH, van den Bergh ACM, Langendijk JA. Normal tissue complication probability (NTCP) models for late rectal bleeding, stool frequency and fecal incontinence after radiotherapy in prostate cancer patients. Radiotherapy and Oncology. 2016;119:381-7.
[35] Dean JA, Wong KH, Welsh LC, Jones AB, Schick U, Newbold KL, et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. Radiother Oncol. 2016;120:21-7.
[36] Vergouwe Y, Moons KGM, Steyerberg EW. External Validity of Risk Models: Use of Benchmark Values to Disentangle a Case-Mix Effect From Incorrect Coefficients. Am J Epidemiol. 2010;172:971-80.
[37] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the Performance of Prediction Models A Framework for Traditional and Novel Measures. Epidemiology. 2010;21:128-38.
[38] Gerds TA, Cai TX, Schumacher M. The performance of risk prediction models. Biometrical J. 2008;50:457-79.
[39] Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. Am J Epidemiol. 2004;159:882-90.
[40] Hosmer DW, Hosmer T, leCessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med. 1997;16:965-80.

[41] Cooper JS, Pajak TF, Forastiere AA, Jacobs J, Campbell BH, Saxman SB, et al. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. N Engl J Med. 2004;350:1937-44.
[42] Machtay M, Moughan J, Trotti A, Garden AS, Weber RS, Cooper JS, et al. Factors associated with severe late toxicity after concurrent chemoradiation for locally advanced head and neck cancer: an RTOG analysis. J Clin Oncol. 2008;26:3582-9.
[43] Forastiere AA, Zhang Q, Weber RS, Maor MH, Goepfert H, Pajak TF, et al. Long-term results of RTOG 91-11: a comparison of three nonsurgical treatment strategies to preserve the larynx in patients with locally advanced larynx cancer. J Clin Oncol. 2013;31:845-52.
[44] Gillison ML, Trotti AM, Harris J, Eisbruch A, Harari PM, Adelstein DJ, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. Lancet. 2019;393:40-50.
[45] Saarilahti K, Kouri M, Collan J, Kangasmaki A, Atula T, Joensuu H, et al. Sparing of the submandibular glands by intensity modulated radiotherapy in the treatment of head and neck cancer. Radiother Oncol. 2006;78:270-5.
[46] Deasy JO, Moiseenko V, Marks L, Chao KS, Nam J, Eisbruch A. Radiotherapy dose-volume effects on salivary gland function. Int J Radiat Oncol Biol Phys. 2010;76:S58-63.
[47] Murdoch-Kinch CA, Kim HM, Vineberg KA, Ship JA, Eisbruch A. Dose-effect relationships for the submandibular salivary glands and implications for their sparing by intensity modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2008;72:373-82.
[48] Levendag PC, Teguh DN, Voet P, van der Est H, Noever I, de Kruijf WJ, et al. Dysphagia disorders in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscle: a dose-effect relationship. Radiother Oncol. 2007;85:64-73.
[49] Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49:1373-9.
[50] Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med. 2015;162:W1-W73.
[51] Zwanenburg A, Lock S. Why validation of prognostic models matters? Radiotherapy and Oncology. 2018;127:370-3.
[52] Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of radiomic signature development: The need for safeguards. Radiotherapy and Oncology. 2019;130:2-9.

Chapter 8 – Multiple tImepoiNt mri tO Track organ at risk (OAR) changes in patients undergoing radical radiotherapy for head and neck cancer: the MINOT-OAR study

8.1 Overview

Chapters 2-7 all describe work whose underlying principle is better targeted ART for toxicity reduction, based on a greater understanding of delivered dose to important OARs. However, this approach may be driven by other sources of information, not least advanced imaging techniques such as MRI. In this Chapter, I describe the methodology and some preliminary results from the Minot-OAR study – a sub-study of VoxTox. In this study, participants were invited to undergo MRI scans immediately prior to their first fraction of radiotherapy, and before fractions 6, 16, and 26. The imaging protocol was designed to include a detailed anatomical sequence, diffusion-weighted imaging, and a sequence to permit estimation of the relative fat fraction within a given structure. Concurrent with this, the side effects experienced by patients were recorded using the standard VoxTox toxicity assessment tools and methods as described in Chapters 2 and 3. The objectives of this work were to improve our understanding of the nature, and mechanisms of change within the major salivary glands during radiotherapy treatment, to try and validate previously published imaging biomarkers of toxicity and to try and find such IMBs as early as possible in a course of treatment, in order to maximise the theoretical benefit of ART for toxicity reduction.

8.1.1 My Role

The idea for the Minot-OAR sub-study was mine. Whilst conducting the literature review to understand both current state-of-the-art, and areas where more research were required, my understanding of the role of MRI in H&N radiotherapy increased significantly, not the least the potential and problems associated with combined MRI-linear accelerator units – MR-Linac. A synthesis of this understanding was published as the following opinion piece:

The future of image-guided radiotherapy – is image everything?
Noble DJ, Burnet NG. Br J Radiol. 2018 Jul;91(1087):20170894

I undertook all necessary consultations with experts and stakeholders prior to drafting the initial study proposal, and led on the process of gaining regulatory and ethical approval for the study via a substantial amendment to the VoxTox study itself. Along with Amy Bates (Research Radiographer), I led on recruitment to the study, and undertook many of the toxicity

assessment interviews. I completed a course in MRI safety in order to personally supervise study scans. I undertook all the data collation, image and data analysis described in this chapter. Although recruitment to the Minot-OAR study has completed, image analysis is ongoing, and toxicity data continue to mature. The data presented in this chapter represent the most up to date progress with this work.

8.1.2 Acknowledgements

Amy Bates and staff from the CCTC (Cambridge Cancer Trials Centre) provided assistance in the process of drafting, editing and submitting necessary documents for the substantial amendment to the HRA. Vicky Lupson (Superintendent Radiographer, Wolfson Brain Imaging Centre - WBIC) provided support, assistance and guidance in providing relevant costings, and documentation, to undertake imaging on WBIC scanners. Karen Welsh (WBIC radiographer) helped to design the imaging protocol, and undertook many of the scans in the study. All scans were done by members of the WBIC radiographer team. Recruitment to Minot-OAR, as well as toxicity assessments, were undertaken by Amy Bates, other members of the research radiographer team, and myself. Dr Tilak Das (consultant H&N radiologist) provided input into the imaging protocol, as well as governance reports for study scans, and inter-observer variability data. All other governance reports were performed by Dr Dan Scoffings (consultant H&N radiologist). I received assistance in the process of extracting imaging data from PACS, as well as anonymisation and storing of images in the University Department of Radiology from Dr Fulvio Zaccagna (PhD student, Department of Radiology), and the Department of Radiology IT team, and Dr Zaccagna also provided input and advice on the methodology for image analysis. I am grateful for the input of Drs Ferdia Gallagher (CRUK Senior Cancer Research Fellow, Honorary Consultant Radiologist, PhD second supervisor), Andrew Priest (MRI physicist and Affiliated Lecturer, Department of Radiology), and Martin Graves (MRI physicist and Affiliated Lecturer, Department of Radiology, Head of MR Physics and Radiology IT, CUHNFT) in discussions regarding the science, and implementation of the imaging protocol.

8.2 Introduction

8.2.1 Imaging-driven ART

Chapters 5, 6 and 7 have shown that delivered dose to OARs can differ from doses anticipated during planning, that these differences may be clinically significant, and that patient, disease and treatment factors may all help to guide clinicians in the broader objective of delivering

personalised radiotherapy. However, Heukelom and Fullers recent and thorough review shows that other sources of data may also be very helpful in driving intelligent ART [1], specifically imaging modalities which are not currently standard components of radiotherapy workflows, such as PET-CT and MRI done specifically for treatment adaptation. There has already been considerable interest in deriving and validating IMBs from both of these modalities [2], and active research continues in this space [3, 4].

Much of this research has focussed on disease, both the primary tumour and loco-regional lymph nodes [5, 6]. PET-CT IMBs such as changes in the avidity of FDG and F-MISO [3, 7], and functional MRI techniques such as dynamic contrast enhancement (DCE) [8-12] and diffusion weighted MRI [13-21], have been studied extensively. The underlying aim of this research is ART_{amplio}, ART_{reduco}, or ART_{totale}, as defined by Heukelom and Fuller [1]; target volumes and doses are amended on the basis of new information about disease, no new information about the morphology or physiology of OARs is derived, and any reduction in treatment toxicity results from reduced dose to OARs from this process, rather than by selecting patients for preferential or enhanced OAR sparing based on new data about that structure.

However, both imaging modalities have also been used to study the effect of radiation on OARs themselves, especially the salivary glands [22-25], with the overall aim of finding IMBs that are sensitive and specific for toxicity events. MRI in particular has been used to study changes in the morphology, function and physiology of salivary glands [24]. Techniques that have been used, and measured parameters include quantification of morphological change [24, 26, 27], DCE [22, 23, 27], and diffusion weighted imaging (DWI) [21, 24, 27-30].

8.2.2 Diffusion-weighted imaging theory

DWI is a technique whereby images are created from differences in the signal characteristics of water diffusion in different tissues. Given a large enough sample of water molecules, followed for a sufficient duration of time, the distribution of individual vectors will be normally distributed around a mean of 0 [31]. The standard deviation of this distribution will be proportional to the diffusion coefficient, and the time over which the sample is studied. In human tissues, the movement of water molecules is not ‘free’, but restricted by microstructures such as cell membranes, and organelles [31]. In DWI, a pair of opposing magnetic fields dephase and rephase water molecules [29]. If water molecules are stationary, no dephasing will occur. However, if there is movement, then dephasing will occur, and this is depicted as signal loss on DWI [29]. The magnitude of signal loss depends on 2 factors; firstly, the amount

of water molecule movement, and secondly the strength and duration of the gradient pulse [29]. This latter factor can be summarised as the b-value of the imaging sequence [30].

The complexity of this system in tissues is increased by the presence of blood vessels, and tissue perfusion. At the level of individual voxels, the organisation of capillary networks is thought to be pseudo-random, hence the motion of fluid in these vessels should be incoherent [32, 33]. This leads to the notion of Intravoxel Incoherent Motion (IVIM), or imaging of microscopic translational motion that occurs in voxels on MR imaging [32, 33]. Therefore, this motion has 2 distinct components; random Brownian diffusion of water molecules, and tissue perfusion within capillary beds [32]. The summary quantification of these 2 processes together is known as the Apparent Diffusion Coefficient (ADC) [34]. However, the relative contribution of these 2 components of ADC varies, depending on the b-values used in the acquisition of the images. At lower b-values (50-150mm/s2) blood flow, and hence perfusion contribute significantly to the acquired signal [30]. At higher b-values (500-1000mm/s2) contribution from perfusion to overall signal is suppressed [25]. Furthermore, by using multiple b-values, it is possible to model the signal decay, and estimate the relative contributions of both perfusion fraction (f) and diffusion coefficient (D) from the summary ADC statistic [30].

Using multiple b-values not only permits quantification of mean diffusion in tissues, but also interrogation of the shape of the diffusion data distribution. In Diffusion Kurtosis Imaging (DKI), the underlying principle is that tissues with a higher degree of cellular complexity, for example with more densely packed cell membranes or organelles, will deliver a less Gaussian distribution of water diffusion data [35]. The degree of kurtosis quantifies how Gaussian (or not) the distribution of data is, and a high degree of diffusional kurtosis is thought to associate with greater complexity at a microstructural level [35, 36]. However, these effects are more prominent with stronger gradients, hence higher b-values are required for this technique, with typical protocol for DKI involving multiple b-values up to 2500mm/s2 [37].

Therefore, a number of research groups have hypothesised that radiation effects upon salivary glands may have caused microstructural changes within these organs, such as necrosis, atrophy, and inflammation, that could be quantified by these imaging techniques, and might predict subsequent toxicity events [24, 25]. For example, research groups have studied absolute values of ADC, and relative changes in this metric, before and after radiation [24, 25, 29]. ADC values before and after stimulation of the salivary glands [25, 29, 38] have been studied, and more recently DKI techniques have also been used to study radiation induced salivary gland damage [39, 40]. Much of this work has studied the changes in such metrics before and after radiotherapy [24, 25, 29]. A smaller number of papers report on changes

during a course of treatment [21, 27, 41], but to the best of my knowledge, none have studied how these metrics change at multiple timepoints during a course of treatment.

Furthermore, uncertainty remains as to the biological mechanisms that underlie the considerable morphological changes that occur during treatment, and the toxicity that many patients continue to experience [24, 25, 42]. In-vitro and animal model work suggests that damage to the plasma cell membrane of acini cells is followed by DNA damage, then death and subsequent lysis of acinar progenitor cells [24, 43-47], and suggested mechanisms of volume reduction following radiotherapy include acinar loss [24, 45, 46] and fibrosis [24, 46]. The fact that parotid glands can recover their volume up to a year or more after radiotherapy goes some way towards supporting this idea. It is also supported by animal model work showing that volume recovery post-radiation is heavily dependent on dose to the glands stem cell compartment [48]. Despite this work however, a better understanding of the aetiology of radiation induced salivary gland injury is needed [25, 42], and would help to improve the design of clinical studies whose objective is toxicity reduction.

8.2.3 Objectives of the Minot-OAR study

We hypothesised that studying changes in parotid volume and ADC at multiple timepoints during treatment might help to better elucidate some of these mechanisms. We further hypothesised that having an estimate of how the relative fat content of salivary glands changed during treatment, in parallel with volume and ADC data, might augment this analysis. A number of authors have demonstrated a relationship between weight loss and reduction in gland volume [49-52]. Therefore, it could be suggested that in most patients, relative fat fraction would *reduce* during a course of radiotherapy, as patients lose weight and become catabolic. If this were true, it would follow that there would therefore be a negative correlation between gland volume reduction, and changes in estimated fat fraction. Conversely, studies have shown that IMBs inferring greater fat content in parotid parenchyma at baseline associate with toxicity [53-55]. Thus, it could also be hypothesised that gland volume reduction is due to loss of acini, from a direct cytotoxic effect [24, 45, 46], and in these circumstances, an *increase* in relative fat fraction during treatment might follow. Therefore, we hypothesised that attempts to quantify the relative fat content of salivary glands over the course of treatment might, in addition to ADC data, help to inform the underlying mechanisms of volume loss, and potentially toxicity.

Another gap in the literature is the relative lack of studies directly comparing these MR IMBs with toxicity events. An exception is the work of Juan et al, who showed a negative association between parotid volume and toxicity grade, and a positive one with parotid gland ADC [24].

Whilst interesting, these findings have 2 problems. Firstly, the timing of MRIs showing these relationships was >1 year after radiotherapy, well beyond the point at which meaningful intervention could reduce the risk of the event. Secondly, they have not been externally validated by other datasets and research groups.

Therefore, this piece of work aimed to address the following objectives:

1. To measure imaging metrics including volume, ADC and an estimate of fat fraction in salivary glands at multiple timepoints during treatment, thereby trying to deduce a pattern of how they change over time, and gain fresh insight into mechanisms of damage of the salivary glands.
2. To measure gland volume and ADC values towards the end of treatment, alongside toxicity data, to permit external validation of previously reported post-RT IMBs of toxicity.
3. To study these IMBs early enough in treatment, that if they subsequently proved to be reliable predictors of toxicity events, they could guide early treatment adaptation, and thus maximum benefit, in selected patients.

This work was undertaken as a sub-study of the broader VoxTox study – called 'Multiple tImepoiNt mri tO Track organ at risk (OAR) changes in patients undergoing radical radiotherapy for head and neck cancer: the Minot-OAR study'.

8.3 Methods

8.3.1 Study design

8.3.1.1 Ethical approval, governance and funding

The underlying aims and objectives of Minot-OAR were very similar to those of the broader VoxTox study. Furthermore, the research infrastructure for collection and curation of toxicity data was already in place. Therefore, it was decided to open Minot-OAR as a sub-study of VoxTox, rather than opening as a stand-alone project.

To this end, a substantial amendment to VoxTox was submitted to the HRA on 17th March 2017. In addition to tracked and clean versions of the VoxTox study protocol, this also included a patient information sheet (PIS), consent form and GP letter specific to Minot-OAR. Confirmation of approval for the amendment, by both the HRA and relevant research ethics

committee (REC), was received on 28th March 2017. Copies of the amended VoxTox protocol, PIS, consent form and GP letter are included in Appendix A6. As Minot-OAR was a sub-study of VoxTox, it was co-sponsored by the University of Cambridge and Cambridge University Hospitals NHS Foundation Trust. It was conducted according to the principles of Good Clinical Practice (GCP) guidelines, the spirit and the letter of the declaration of Helsinki, ICH-GCP Guidelines, the protocol study protocol, and applicable local regulatory requirements.

All MR imaging within the Minot-OAR sub-study was performed in collaboration with, and on hardware belonging to the Wolfson Brain Imaging Centre (WBIC – a facility of the Department of Clinical Neurosciences, University of Cambridge). Written confirmation of the research agreement with the WBIC, pending receipt of confirmatory communication from the HRA/REC, and defining the cost of scans, was also received on the 28th March 2017. With written permission from responsible persons representing the Cambridge Cancer Centre, the study was funded via the consumables budgets of my Addenbrooke's Charitable Trust (ACT), and Cancer Research UK Clinical Research Fellowships.

8.3.1.2 Study schedule

As outlined in the introduction, a point of difference for Minot-OAR in comparison with published literature was the idea of MR imaging at multiple timepoints during treatment. A key objective was to search for MR IMB's early in treatment, as a tool for guiding toxicity-reducing ART, and it was therefore important to schedule a scan that could address this hypothesis. However, it was also important to have a scan at, or near to the end of treatment, so that previously described anatomical changes, and links between IMBs and toxicity could be assessed and potentially validated. Another important consideration was to try and maximise homogeneity of scan timings between study participants, thereby reducing the potential for temporal noise in the imaging metrics measured. Although there was no literature precedent in this regard, we postulated that delivery of concomitant chemotherapy shortly before MRI scans could affect results, potentially by iatrogenically affecting patient's hydration status, or by enhanced cytotoxic, or anti-inflammatory effects. Timings were therefore scheduled to ensure that scans were not being done within 48 hours of delivery of concomitant chemotherapy. Therefore, taking all these considerations into account, the schedule for the Minot-OAR study was to perform MRI at timepoints, as outlined below:

- Baseline – immediately prior to fraction 1 (baseline)
- Beginning of week 2 – fraction 6 (timepoint 1)
- Beginning of week 4 – fraction 16 (timepoint 2)
- Beginning of week 6 – fraction 26 (timepoint 3)

It was decided not to make any changes to the standard VoxTox schedule of events with regards to toxicity assessments. These were therefore collected according to the schedule outlined in Chapter 2. The only subtle difference was to stipulate collection of toxicity data within Minot-OAR for a minimum of 2 years, compared with 5 years in the broader VoxTox study. Figure 8.1 shows a cartoon schema of the Minot-OAR study (this is the same as Figure 2.4 – with the addition of the MRI study schedule).

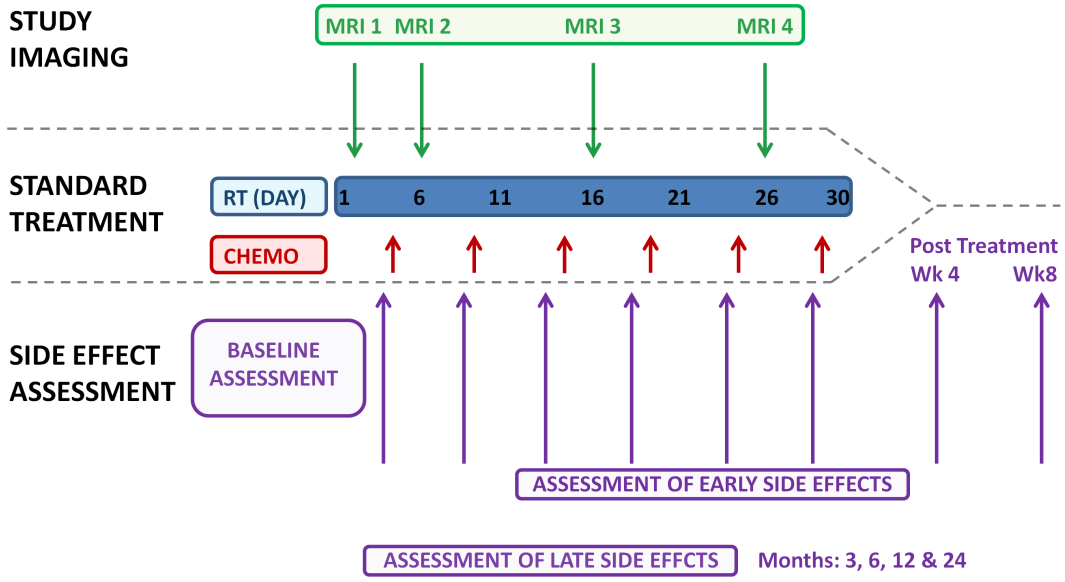


Figure 8.1: Schedule of events for patients recruited to the VoxTox Minot-OAR sub-study.

8.3.1.3 Imaging protocol

All MR imaging within Minot-OAR was done on Siemens (Siemens Healthcare, Erlangen, Germany) 3T MRI scanners at the WBIC. The vast majority of imaging was done on a Skyra*fit* unit. However, unexpected machine maintenance during the conduct of the study, meant the decision was taken to perform 2 scans on the WBIC 3T Prisma*fit* unit. For all scans, patients were positioned supine, supported by vendor's cushions within a 64 channel head and neck coil. It was not possible to image patients in their treatment position or thermoplastic shells

The specifics of the imaging protocol were devised to satisfy as many important objectives as possible, within a limited timeframe. Fields of view for all scan sequences were defined to ensure that both parotid, and submandibular glands were fully encompassed on all patients, with a small safety margin superior and inferior. An a-priori expectation of the study was that the severe acute side effects experienced by many patients undergoing radical RT for HNC in the latter weeks of treatment, could make scans 3 and 4 difficult to tolerate, and reduce concordance with the full study protocol. Therefore, it was decided to try and keep total time

on the scanner within 20 minutes to make the protocol as tolerable as possible for patients, and maximise participation with all 4 scans.

First, it was crucial to have a high quality anatomical sequence on which to measure volumetric changes, and define regions of interest (ROIs) for other metrics. Based on literature precedent [24], test images on healthy volunteers, and input from a consultant head and neck radiologist (Dr Tilak Das), it was decided to use a fat-saturated heavily T2-weighted sequence (so called BLADE – parameters: TE = 90ms, TR = 4380ms, matrix = 320 x 320, in-plane resolution = 0.6875 x 0.6875mm^2, slice thickness = 3mm, distance factor = 10%, FOV = 220mm).

To measure ADC, a readout-segmented multi-shot echo-planar imaging diffusion weighted sequence (so called RESOLVE – parameters: TE1 = 117ms, TE2 = 162ms, TR = 7240ms, matrix size = 200 x 200, in-plane resolution = 1.3 x 1.3 mm^2 , slice thickness = 4mm, distance factor = 10%, FOV = 250mm) was performed with b values of 0, and 700s/mm^2. The b-value of 700mm/s^2 was based on similar published work, and is a compromise between being low enough to retain high signal intensity, whilst having high enough signal intensity to exclude contributions from tissue perfusion [25, 56]. Previous groups working in this field have investigated salivary gland ADC values before and after stimulation of salivary flow, using ascorbic acid tablets or lemon juice [25, 29, 57], and we aimed to investigate the effects of salivary gland stimulation in this work. However, these studies undertook imaging before, and substantially after RT; thus scans were done in the absence of any acute mucositis. Using ascorbic acid in the presence of significant mucositis would not have been ethical, and the link between food cues, and increased salvation is well documented [58, 59]. Furthermore, experiments have specifically shown increased salivation following simulated consumption of sour food [60]. Therefore, we hypothesised that an audio-visual stimulus could be used instead of ascorbic acid or lemon juice as a means of stimulating saliva flow. This was done in the following way. Immediately after the pre-stimulation DWI sequence, the patient was shown a video of a volunteer biting into a lemon. Concurrent with the video, this message was read out over the intercom; *"I want you to imagine now, eating a large, juicy lemon. Imagine how this tastes in your mouth, how it feels. Really bite down into the lemon. Imagine the smell and citrus taste ".* On completion of the message, a second, post-stimulation DWI sequence was taken, with identical parameters to the first.

Finally, we wanted a sequence that would give a robust numerical estimate of changes in relative fat fraction during treatment. Recent work has shown that the mDIXON method can used to estimate the fat fraction of salivary glands [61, 62]. Therefore, the imaging protocol also included a DIXON sequence (so called VIBE-DIXON - parameters: TR = 6.68ms, TE1 =

2.46ms, TE2 = 3.69ms, flip angle = 12.0 deg, matrix size = 352 x 352, in-plane resolution = 0.7 x 0.7mm, slice thickness = 1mm, distance factor = 20%, FOV = 250mm), which generated both 'fat' and 'water' images, from which fat fraction could be estimated.

Therefore, the finalised imaging protocol, with scan times, was as follows:

- Localiser: 11 seconds
- Coronal T1: 41 seconds
- VIBE-DIXON: 3 minutes 57 seconds
- T2 BLADE: 4 minutes 16 seconds
- DWI (RESOLVE) – pre-stimulation: 2 minutes 47 seconds
- Audio-visual salivary gland stimulation: 30 seconds
- DWI (RESOLVE) – pre-stimulation: 2 minutes 47 seconds

Thus total time inside the scanner room for patient set-up and imaging was well inside the pre-specified target of 20 minutes.

8.3.1.4 Patient and treatment details

Selection criteria for the study were defined in order to maximise homogeneity of the patient sample, in terms of both disease characteristics, and treatment protocols, and to minimise the possibility of confounding variables affecting toxicity rates. They were also intended to minimise crossover with concurrent national clinical trials being run in the department at the time recruitment [63, 64]. Therefore, finalised inclusion and exclusion were as follows:

Inclusion criteria:

- Age: between the ages of 18 and 75 (inclusive). The upper age limit reflects clinical departmental protocol for the use of concomitant systemic therapy alongside RT (it is very unusual for this to be given to patients over the age of 70).
- Histologically confirmed squamous cell carcinoma of the oropharynx, larynx or hypopharynx.
- Undergoing treatment with radical radiotherapy to a dose of 65Gy in 30 fractions over 6 weeks (or equivalent radio-biological dose).
- Radiotherapy delivered as IMRT with daily image guidance using TomoTherapy or equivalent technology.
- Radiotherapy treatment volumes including primary site and lymph node regions on at least one side of the neck.

- Systemically well enough, and consented for concomitant weekly systemic therapy with cisplatin (40mg/m^2).
- Baseline performance status 0 or 1.
- Written informed consent.

Exclusion criteria:

- Previous radiotherapy to the treated area.
- Surgery prior to radiotherapy as part of definitive treatment for this malignancy:
    - Specifically, neck dissection with or without pharyngo-laryngectomy prior to radiotherapy.
    - Staging surgery was not considered to be an exclusion criterion. Thus general anaesthetic with pan-endoscopy, tonsillectomy and/or debulking biopsy of primary site, or biopsy of involved lymph node did not exclude patients from the Minot-OAR study.
- Patients who did not require concomitant systemic therapy.
- Patients receiving concomitant systemic therapy other than cisplatin (e.g. cetuximab)
- Patients not fit enough for concomitant systemic therapy. Note that occasionally patients have their weekly systemic therapy stopped before the full 6 cycles are administered, due to concerns about toxicity. Patients who were consented to Minot-OAR, and received at least 2 weekly cycles of concomitant systemic therapy, remained in the study. Those who had their systemic therapy plan discontinued before starting treatment, or after 1 concomitant cycle, were be removed from the study.
- Patients whose treatment volumes did not include any neck nodal volumes.
- Documented claustrophobia.
- Implanted metal or electronic devices, which were a contra-indication to MRI. Specifically:
    - Heart pacemakers & implantable cardiac defibrillators.
    - Insulin pumps.
    - Implanted hearing aids.
    - Implanted aneurysm clips.
    - Neuro-stimulators.
    - Metallic bodies in the eye.

The study was not designed to detect a specific effect size; therefore, a power calculation was not necessary. Similar previous studies have recruited samples of 5-20 patients [21, 24, 25, 29], and based on their findings we hypothesised that a sample of 15 patients would address

the study aims and objectives. Suitable patients were identified via the Addenbrooke's Hospital HNC MDT meeting.

Characteristic	Number
Age (years)	55.1 (5.7)
Gender:	
Male	16 (88.8%)
Female	2 (11.2%)
Primary site:	
Oropharynx	16 (88.8%)
Larynx	1 (5.6%)
Hypopharynx	1 (5.6%)
T-stage:	
Tis/T1	5 (27.8%)
T2	8 (44.4%)
T3	4 (22.2%)
T4	1 (5.6%)
N-stage:	
N1	2 (11.2%)
N2a	5 (27.8%)
N2b	8 (44.4%)
N2c	3 (16.6%)
Salivary gland mean dose (Gy):	
IPG	41.8 (8.7)
CPG	32.8 (7.8)
ISMG	63.3 (2.1)
CSMG	54.2 (5.0)
Baseline values:	
Parotid gland volume (ml)	26.1 (6.2)
Submandibular gland volume (ml)	8.9 (1.9)
Parotid gland ADC ($x10^{-3}$ mm/s^2)	1.69 (0.19)
Submandibular gland ADC ($x10^{-3}$ mm/s^2)	1.77 (0.23)
MRI scans completed:	
1	18 (100%)
2	17 (94.4%)
3	16 (88.8%)
4	11 (61.1%)

Table 8.1: patient characteristics – Minot-OAR sub-study (n = 18, 12 for ADC data). For continuous variables, means and standard deviations are reported, absolute numbers and percentages for proportions. IPG = ipsilateral parotid gland, CPG = contralateral parotid gland, ISMG = ipsilateral submandibular gland, CSMG = contralateral submandibular gland.

Using VoxTox recruitment data from the previous 12 months, it was estimated that with the selection criteria outlined above, a recruitment rate of 50%, and an in-study attrition rate of 33%, 8 patients per year would complete the full study protocol. On this basis, the objective was to recruit participants over a 2-year period. As anticipated, the study ran for 2 years, from April 2017 to April 2019. During this time 18 patients were recruited. Patient and treatment details are shown in Table 8.1.

8.3.2 Image Analysis

On completion of each scan, images were sent from the scanner software directly to the hospital PACS system. At this point scans were not anonymised, and DICOM headers included all patient specific details including name, date of birth and hospital numbers. They were reviewed by a site-specialist radiologist to ensure there were no new or unexpected clinical findings, and a governance report was issued. From PACS, I imported all anatomical (BLADE) T2 sequences into the research database of the Prosoma system. Scan windowing was optimised for each scan, and I used the segmentation tools within Prosoma to contour both parotid and submandibular glands on a slice-by-slice basis. Examples of contours within Prosoma are shown in Figure 8.2. Prosoma automatically generated volumes from these 2-D contours, which were recorded in an excel spreadsheet.

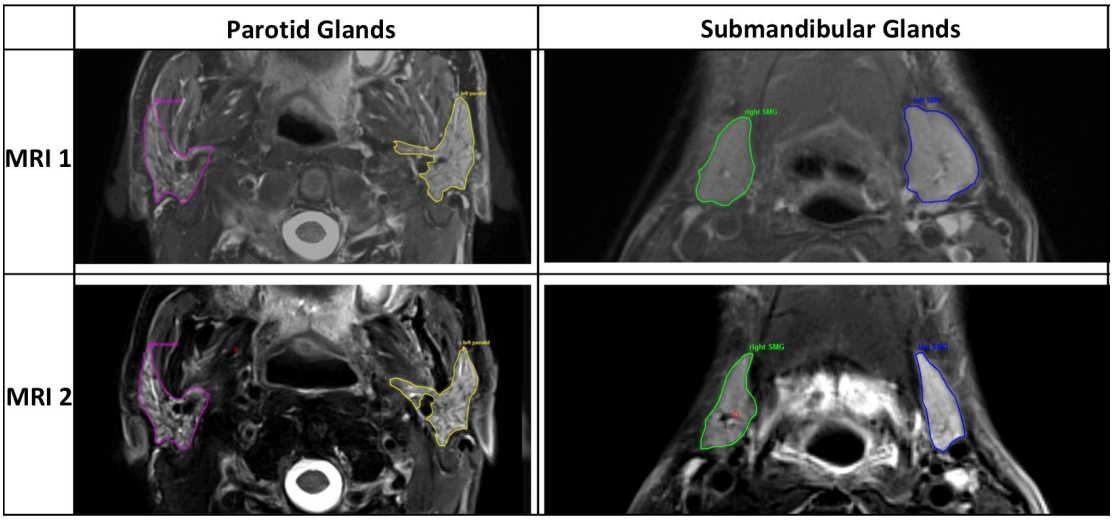


Figure 8.2: Segmentations of salivary glands on T2-BLADE sequences using Prosoma software, for volume measurement. Note the volume loss between scans 1 (baseline) and 2 (radiotherapy fraction 6).

For work on ADC on fat-fraction calculations, images were anonymised, and imported into a database of the University of Cambridge, Department of Radiology. From here, they were loaded into Osirix software (Pixmeo SARL, Bernex, Switzerland) for further analysis.

Within Osirix, b0 and b700 images were used to calculate ADC maps on a pixel-by-pixel basis using the following equation:

$$ADC = \frac{ln(SI_0/SI_{700})}{b_{700} - b_0} \quad (8.1)$$

Where SI_0 and SI_{700} respectively represent the signal intensities measured at b-values 0 and 700s/mm^2. A split-screen user interface was used to load ADC maps next to T2-BLADE sequences. Polygonal ROI's for parotid and submandibular glands were segmented on the BLADE sequence, and saved. Following previous methodologies [24, 25], and to reduce noise from these structures on ADC analysis, large blood vessels and salivary duct structures were excluded from contours, unless completely surrounded by gland parenchyma. A rigid image registration tool within Osirix was then used to fuse the BLADE and ADC map to the same geometric frame of reference. Manual corrections of registrations were possible, and performed to optimise relative positions of blood vessels and salivary duct structures, and CSF surrounding the spinal cord. Once optimised, registrations were locked, and ROIs were copied from BLADE images onto ADC maps. This workflow is shown in Figure 8.3. Per-slice ADC data were then exported in csv format, and mean values for each ROI were calculated.

8.3.3 Inter-observer variability

To ensure reproducibility of both gland volume and ADC data, measurements were made by a second expert observer (consultant head & neck/neuro-radiologist). For 2 patients, the parotid and salivary glands were segmented in Osirix on the T2-BLADE images, for baseline and final (timepoint 3) scans. Volumes of these ROIs were recorded. Image fusion was then performed between the T2-BLADE and ADC maps, ROIs were transposed over, and ADC measurements made using the same methodology described in section 8.3.2. Therefore 16 datapoints were available for direct comparison with my observations, for both gland volume and ADC.

8.3.4 Toxicity measurement

As described in section 8.3.1.2, the Minot-OAR sub-study used the standard VoxTox toxicity assessment workflow, as described in detail in Chapters 2 and 3. CTCAEv4.03 was used for

toxicity reporting, with 2 endpoints of interest – xerostomia and salivary duct inflammation (SDI). As discussed in section 3.3.3, SDI describes the clinical phenotype of thick, sticky ropey saliva and oropharyngeal secretions, with or without alterations in taste.

Due to the relative immaturity of the toxicity data, both acute and toxicity results were of interest. However, there were challenges associated with using acute toxicity data as an endpoint. First, all systems for scoring radiotherapy toxicity (and this applies to CTCAEv4.03) are relatively low resolution Likert scales. Second, in this cohort of 18 patients all of whom received concomitant cisplatin and radiation to both sides of the neck, moderate to severe acute toxicity rates by the latter weeks of treatment were predictably high (see Figure 8.11 for details). This meant that event rates would have been too high to provide a good discriminative endpoint for hypothesised IMB's. Furthermore, it was not clear how to choose a specific timepoint during treatment at which to measure acute toxicity. Therefore, for acute toxicity, it was decided to measure the total cumulative burden of acute side effects over treatment, in the hope that this would give a broader spread of results against which to compare imaging data. This was done by adding toxicity scores recorded at each weekly assessment during treatment, and the 6-week post-treatment score, for each endpoint, to give a cumulative acute toxicity score (CATS). For late toxicity, 6 and 12-month point prevalence were collated, and 12-month data are reported consistent with the methodology used in Chapter 7.

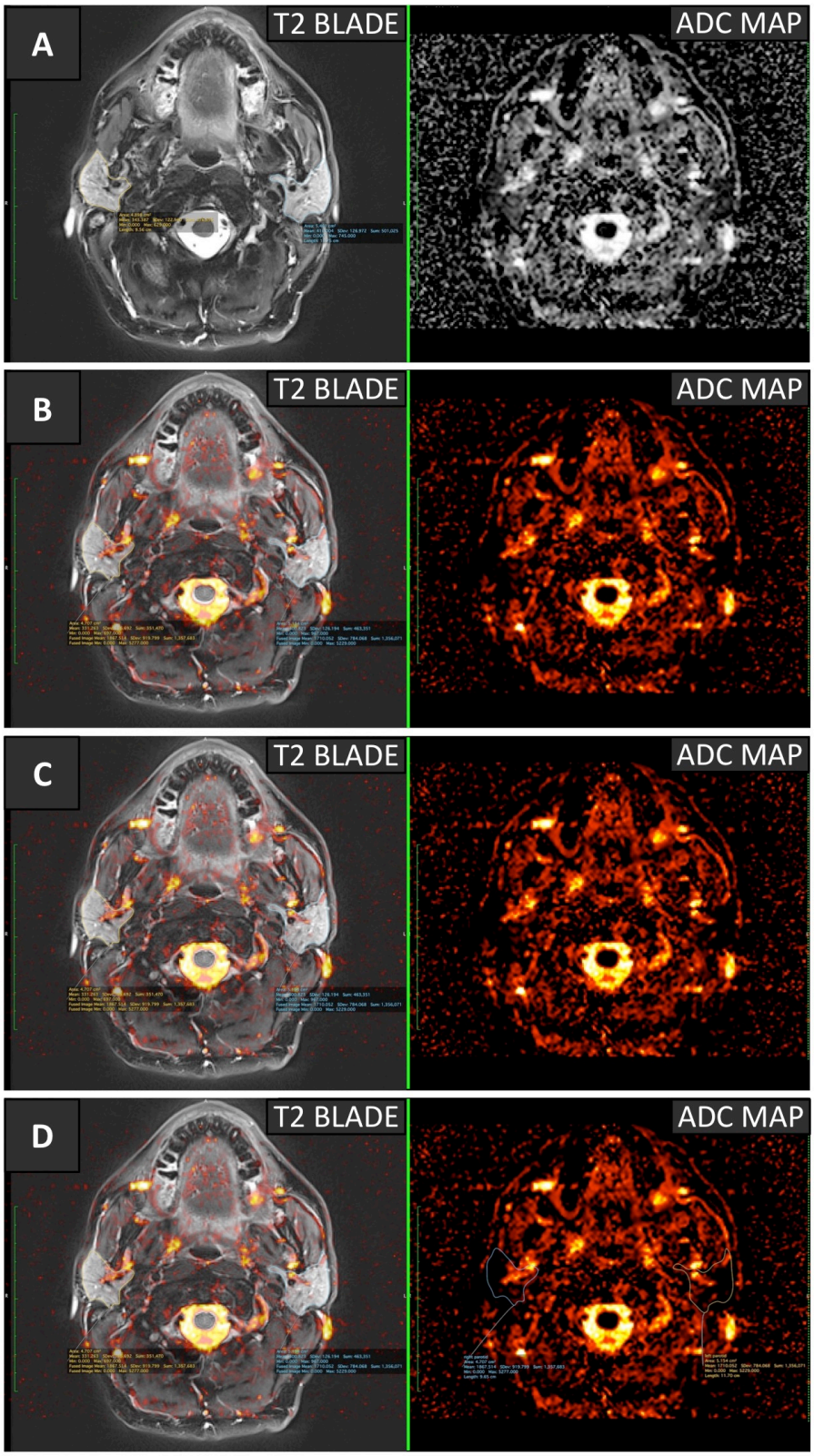


Figure 8.3 (A-D): Technique for measuring ADC. A – Anatomical sequence (T2 Blade) and ADC map loaded side by side (split screen). ROIs completed on T2 image and saved. B – ADC switched to hot/cold render, rigid registration between T2 and ADC map. C – minor manual corrections of registration, based on relative positions of salivary ducts, blood vessels and CSF surrounding the spinal cord. D – ROIs copied from T2 to ADC map, and ADC data exported in csv format.

8.4 Results

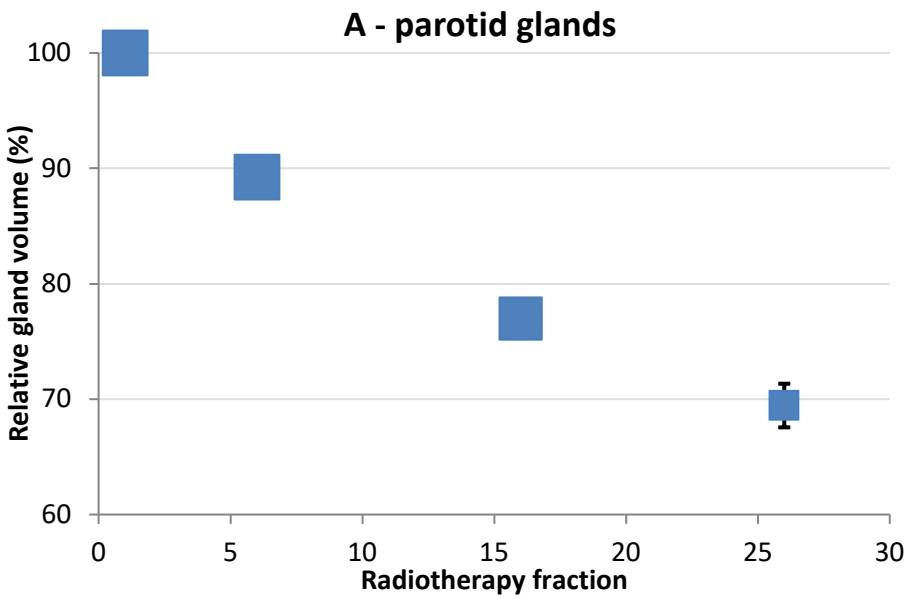
Between April 2017 and April 2019, 18 patients were recruited to participate in the Minot-OAR sub-study. Their details, including study MRI completion rates are shown in Table 8.1. 11/18 participants (61%) completed all 4 study MRI scans, confirming a-priori concerns regarding scan tolerability in the latter stages of treatment. One patient had the baseline scan, but suffered a surgical emergency (perforated colon), unrelated to HNC treatment, shortly after radiotherapy commenced, and was not fit enough for scans 2-4. Therefore, there were 36 parotid glands for analysis at baseline, 34 at timepoint 1, 32 at timepoint 2, and 22 at timepoint 3. One patient had an absent left submandibular gland. Therefore, respective numbers of this structure for analysis were 35, 33, 31 and 21.

8.4.1 Salivary gland volume changes over time

Mean baseline values for both parotid and submandibular gland volumes are shown in Table 8.1, and subsequent relative volume reduction are reported. Mean volume reductions (with standard errors), for the parotid glands at timepoints 1, 2 and 3 respectively were: 10.8%, 23.0%, 30.6%. Mean volume reductions (with standard errors), for the submandibular glands at timepoints 1, 2 and 3 respectively were: 11.4%, 25.4%, 32.2%. These data are shown in Figure 8.4.

The effect of radiotherapy mean [planned] dose on relative reduction in gland volume at timepoints 1 and 3 was investigated. These data are shown as scatter plots in Figure 8.5. Observed relationships were weak, or non-existent. For the parotid glands, R^2 for dose versus gland shrinkage at timepoints 1 and 3 respectively were 0.14 and 0.08, with p-values of 0.030 and 0.20. For the submandibular glands, R^2 values were 0.02 and 0.01, with p-values of 0.43 and 0.70 respectively.

However, on inspection of the raw data, it was clear that in most patients, the gland receiving the higher dose saw greater volume reduction at both timepoints. For the parotids, the gland with the higher mean planned dose shrunk more at timepoint 1 in 14/17 cases (p=0.013, two-tailed binomial test), and 9/11 cases at timepoint 3 (p=0.065, two-tailed binomial test). For the submandibular glands, respective proportions were 12/16 (p=0.077, two-tailed binomial test), and 6/10 (p=0.75, two tailed binomial test).



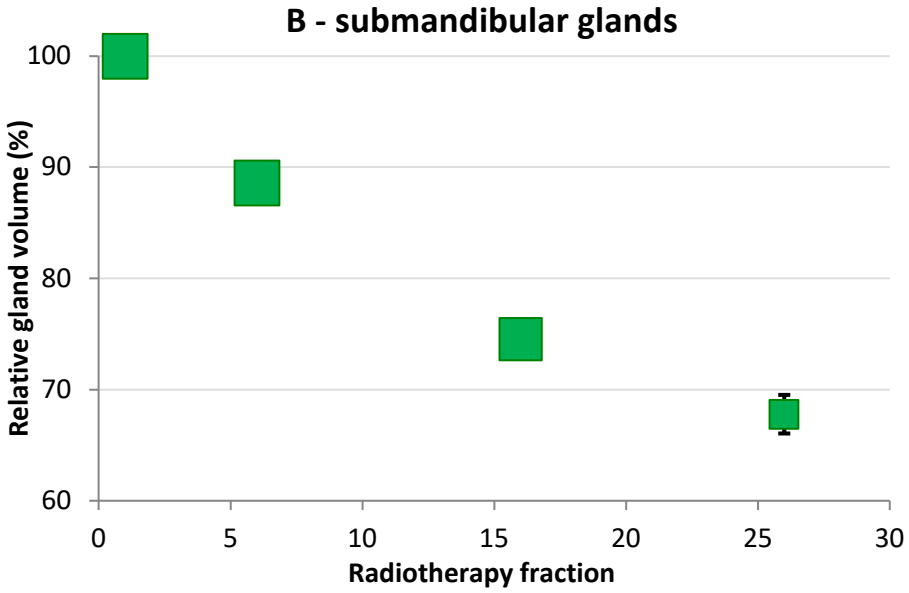


Figure 8.4 (A&B): Salivary gland volume reduction during treatment. A – parotid glands, B – submandibular glands. Marker sizes are proportional to sample size at each timepoint (18, 17, 16, 11 respectively). Error bars are standard error.

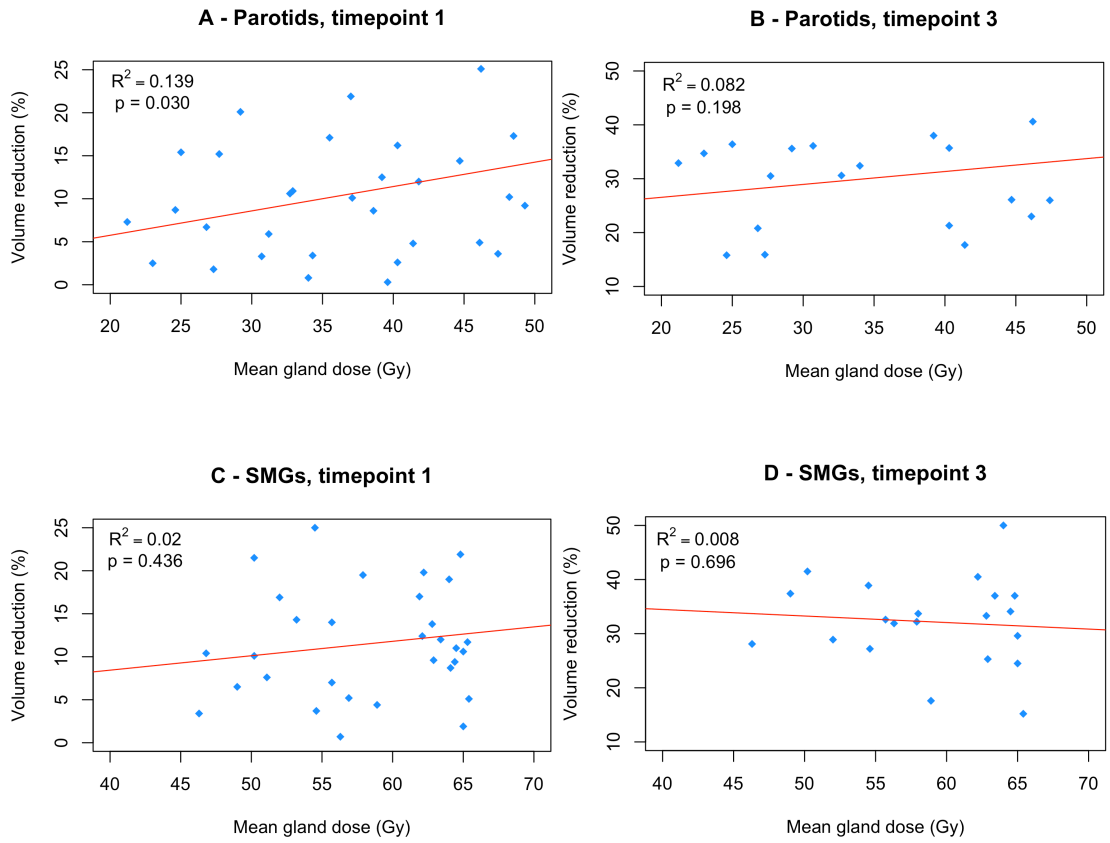


Figure 8.5 (A-D): Relationships between mean dose to salivary glands (A&B – parotid glands, C&D – submandibular glands), and relative volume reduction at timepoints 1 (A&C) and 3 (B&D).

To test this further, mean dose to, and mean volume reduction of both paired structures was calculated for each patient. For each gland, mean dose and volume reduction values were subtracted from observed values, and plotted against each other. These data are shown in Figure 8.6, and relationships are much more convincing. For the parotid glands, R^2 and p-values at timepoint 1 were 0.38 and <0.001 respectively. At timepoint 3, results were 0.44 and <0.001. For the submandibular glands, R^2 and p-values at timepoint 1 were also significant (0.29 and 0.001), but this was not the case at timepoint 3 (R^2 = 0.01, p=0.67).

These results suggest that whilst dose may be an important determinant in gland volume reduction, the strength of the effect is reduced by substantial variation between patients, independent of dose. To test this hypothesis further, volume reduction in all paired glands (right v left, parotid and submandibular glands together), at timepoints 1, 2 and 3, was compared. The data are shown in Figure 8.7, and results appear to confirm this idea – relatively strong and statistically significant relationships were observed at timepoints 1 and 2, with a weaker/borderline relationship seen for timepoint 3.

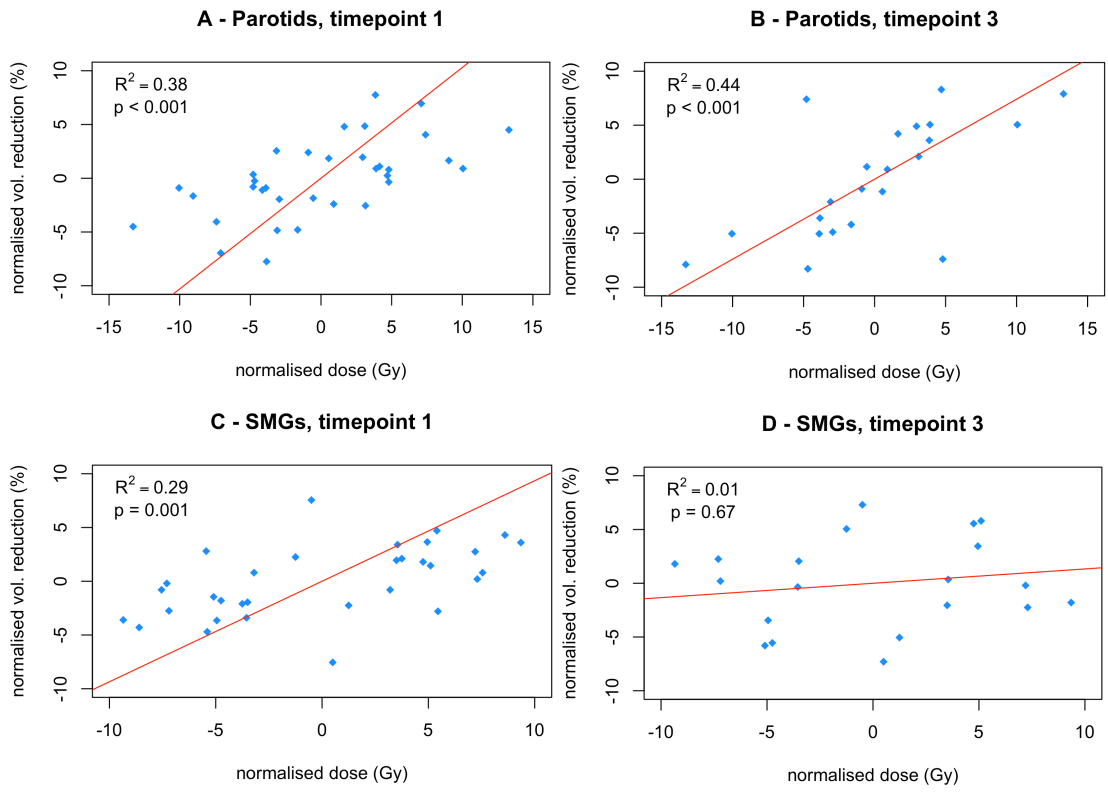


Figure 8.6 (A-D): Relationships between mean gland dose and relative volume reduction, normalised to mean values in each patient. A&B – parotid glands, C&D – submandibular glands

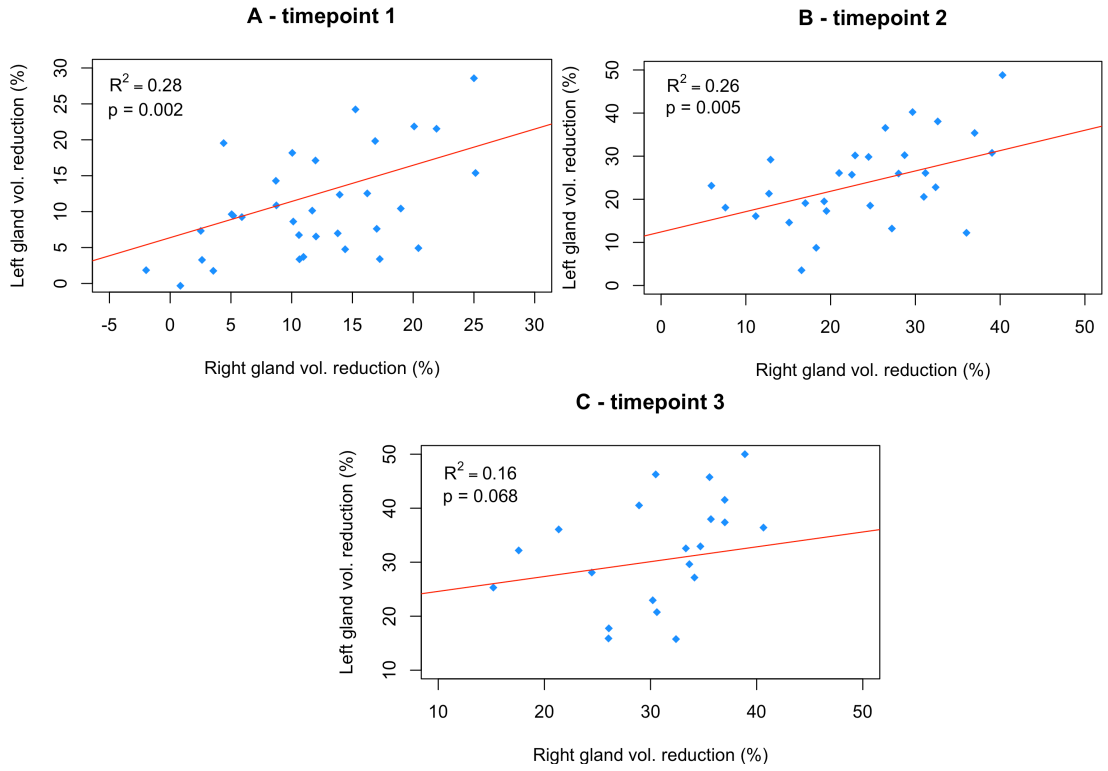


Figure 8.7 (A-C): Relationships between right and left salivary gland (both parotids and submandibular) volume reduction in each patient, at timepoints 1 (A), 2 (B), and 3 (C).

8.4.2 ADC changes over time

ADC data was available for 12 patients. All 12 had the baseline, and timepoint 1 scans. Ten had the timepoint 2 scan, and 8 managed all 4. Mean baseline ADC values for both parotid and submandibular glands are shown in Table 8.1. Mean ADC values (x10^{-3}, standard errors in parentheses) were 1.79 (0.03), 1.92 (0.04) and 2.03 (0.04) s/mm^2 at timepoints 1, 2 and 3. For the submandibular glands, respective values were 1.78 (0.03), 1.94 (0.05) and 2.14 (0.06) s/mm^2. A summary of these data is presented in boxplot format in Figure 8.8. Differences between distributions at timepoints 2 and 3 versus baseline are apparent from the plot, but differences at timepoint 1 were formally tested with paired t-tests. For the parotid glands the difference (1.69 vs. 1.79) was statistically significant ($p<0.001$). There was no difference for the SMGs (1.78 vs. 1.78, $p=0.82$). Data for individual patients, for both structures, are shown as spaghetti plots in Figure 8.9 (A&B).

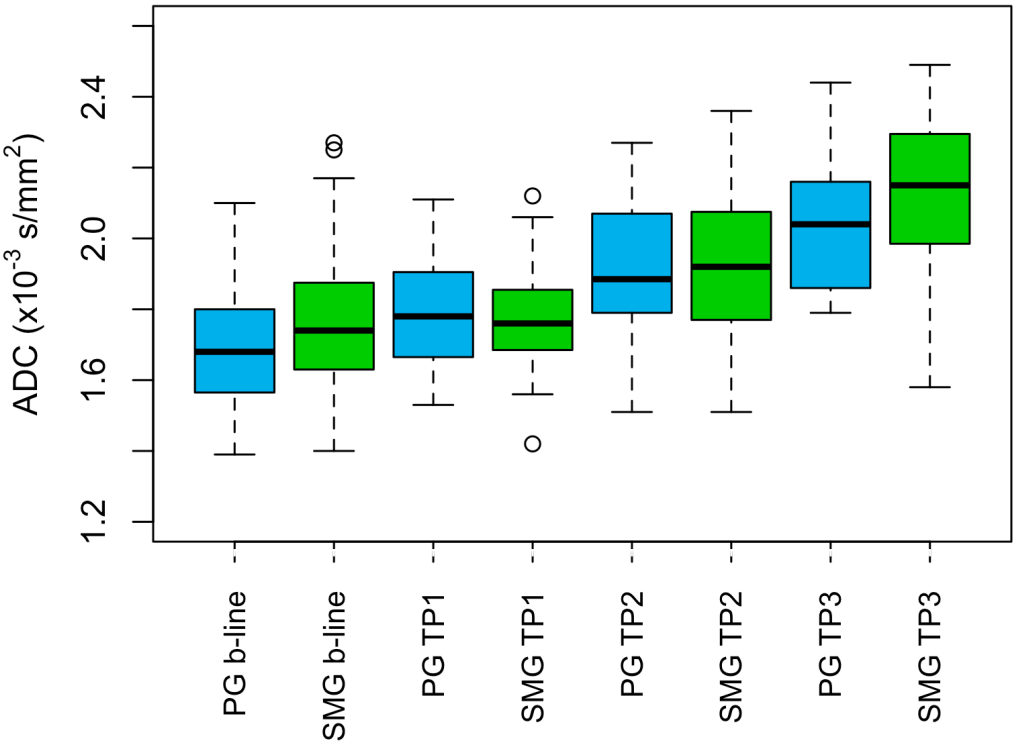


Figure 8.8: Absolute ADC values for parotid (blue) and submandibular glands (green) at baseline, and MRI timepoints 1, 2 and 3.

Proportional changes in ADC values, relative to baseline values, or a per-patient basis, are shown in Figure 8.10 (A&B). Changes in parotid gland ADC are apparent at timepoint 1 (fraction 6), and mean values appear to increase in a linear fashion throughout the rest of

treatment. For the SMGs, changes at timepoint 1 are minimal, but a noticeable increase in ADC is observed thereafter.

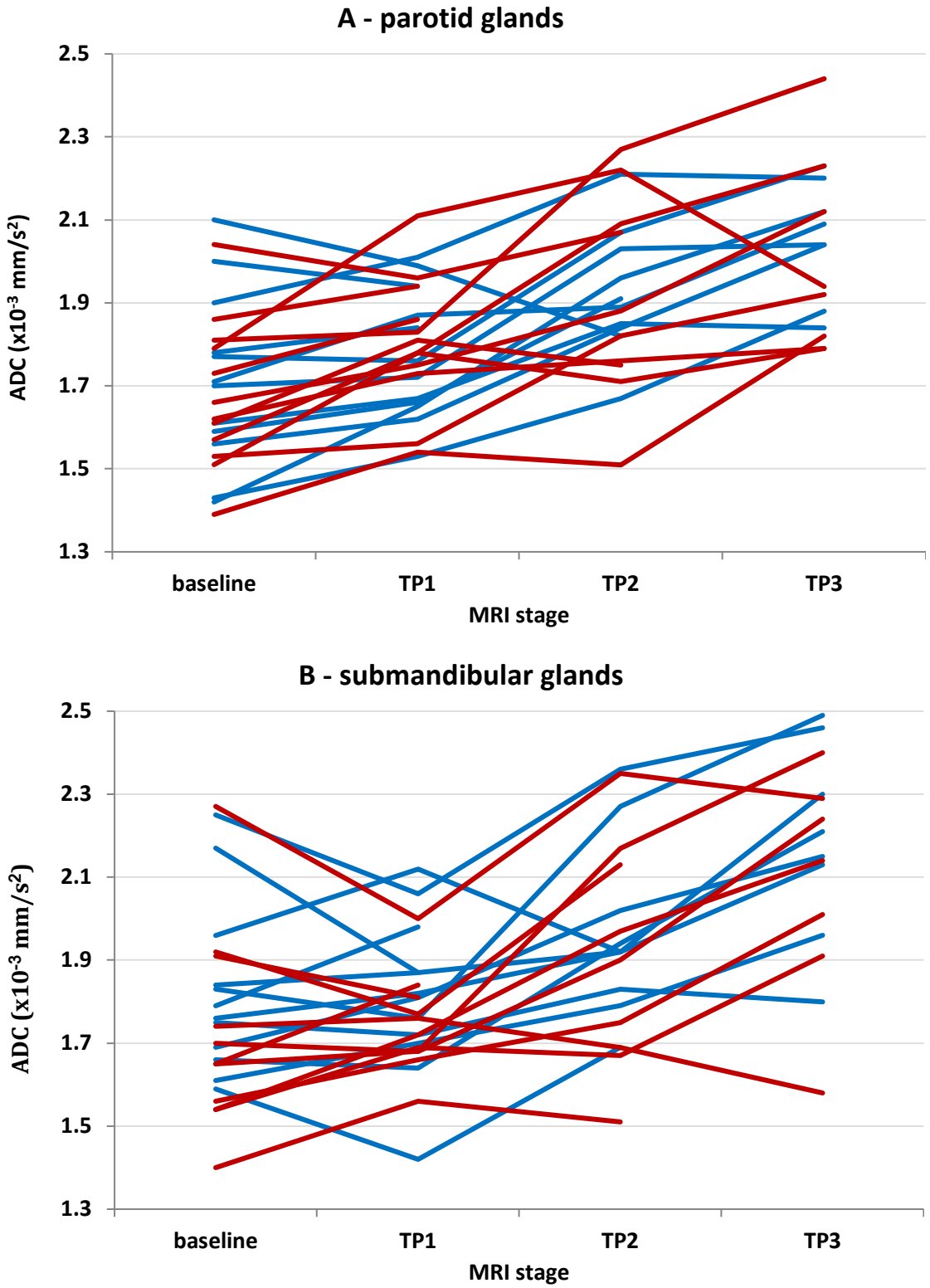


Figure 8.9 (A&B): Trend in mean parotid (A) and submandibular gland (B) ADC values over time, for each individual patient. Right glands shown in red, left in blue.

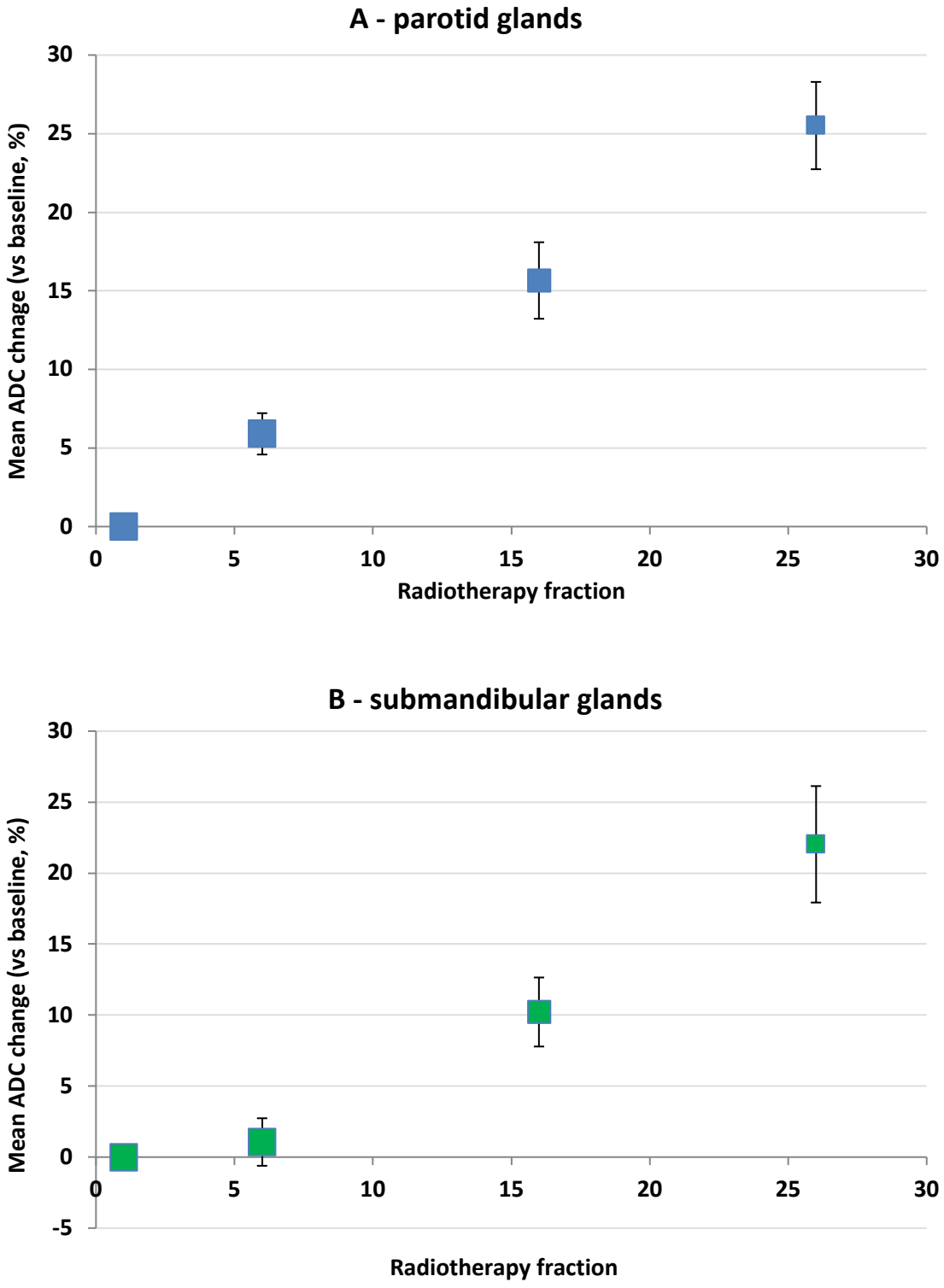


Figure 8.10 (A&B): Mean relative changes in parotid (A) and submandibular gland (B) ADC values over time. Marker sizes are proportional to sample size at each timepoint (12, 12, 10, 8 respectively). Error bars are standard error.

8.4.3 Toxicity results

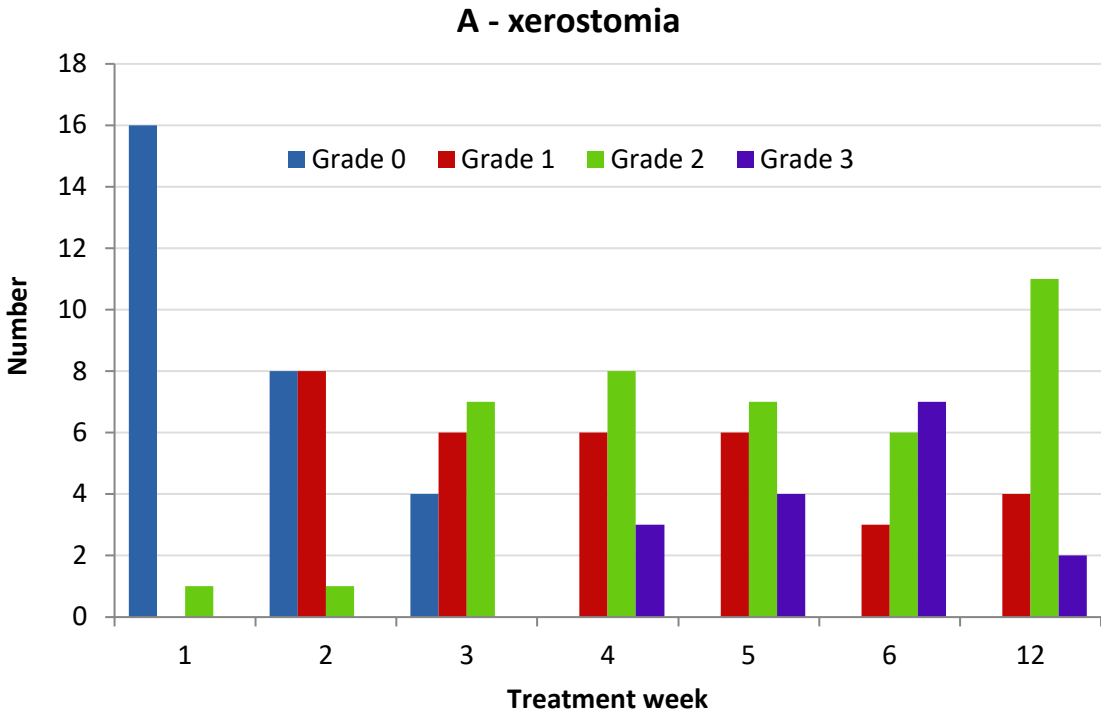
Weekly CTCAEv4.03 acute toxicity results for patients in the Minot-OAR study are shown in Figure 8.11. In both cases there is a clear trend for worsening side effects during treatment. Whilst there was partial resolution by week 12 (6 weeks post treatment), 13/17 (76.5%) patients still reported Grade 2+ xerostomia, and 11/17 (64.7%) Grade 2+ salivary duct inflammation. The acute toxicity result used for comparison with imaging metrics was the cumulative score over all 7 timepoints – cumulative acute toxicity score (CATS) – for reasons discussed in section 8.3.4. Mean CATS for xerostomia and SDI were 9.7 and 10.8 respectively. Importantly, ranges for these means were 6 - 15 (xerostomia), and 7 - 15 (SDI), emphasising that this endpoint gives a finer resolution representation of the overall burden of acute side effects.

Late toxicity results at 6 and 12 months post-treatment were collated and reviewed. 6-month data were available for 11 patients, 12-month data for 9, although one of these was the patient who was unable to participate in scans at timepoints 1-3 due to a surgical emergency, so numbers available for analysis against imaging data were 10 and 8 respectively. Even though the sample size was slightly smaller, a preliminary analysis of the 12 month toxicity data was done, firstly because of the changing nature of event rates over time as shown in Chapter 3, and the fact that 12 month data is a more reliable measure of true late toxicity, and secondly as this was the temporal endpoint used for the core analyses in Chapter 7. At 12 months, 3/8 patients had grade 1 CTCAEv4.03 xerostomia, and 5/8 grade 2. Two of 8 had grade 0 salivary duct inflammation, 2/8 grade 1, and 4/8 grade 2.

8.4.4 Volume changes and acute toxicity

An advantage of using CATS as the acute toxicity metric was that it provided a broader spread of results to be used a dependant variable in analyses against hypothesised IMBs. Figure 8.12 shows CATS results for xerostomia plotted against parotid gland volume reduction (A&B), and salivary duct inflammation plotted against submandibular gland volume reduction (C&D), at timepoints 1 (A&C) and 3 (B&D). In all analyses, the independent variable is mean per-patient gland volume reduction, so that there is one observation per patient. As the sample sizes were relatively small, and the dependent variable was ordinal rather than truly continuous, Spearman's rank correlation tests were done in each case. None were statistically significant. For parotid gland volume reduction against xerostomia, Spearman's rho (with p-values in parentheses) at timepoints 1 and 3 respectively were 0.40 (0.11) and 0.37 (0.26). For

submandibular gland volume reduction against salivary duct inflammation, Spearman's rho (with p-values in parentheses) at timepoints 1 and 3 respectively were 0.16 (0.55) and -0.02 (0.96).



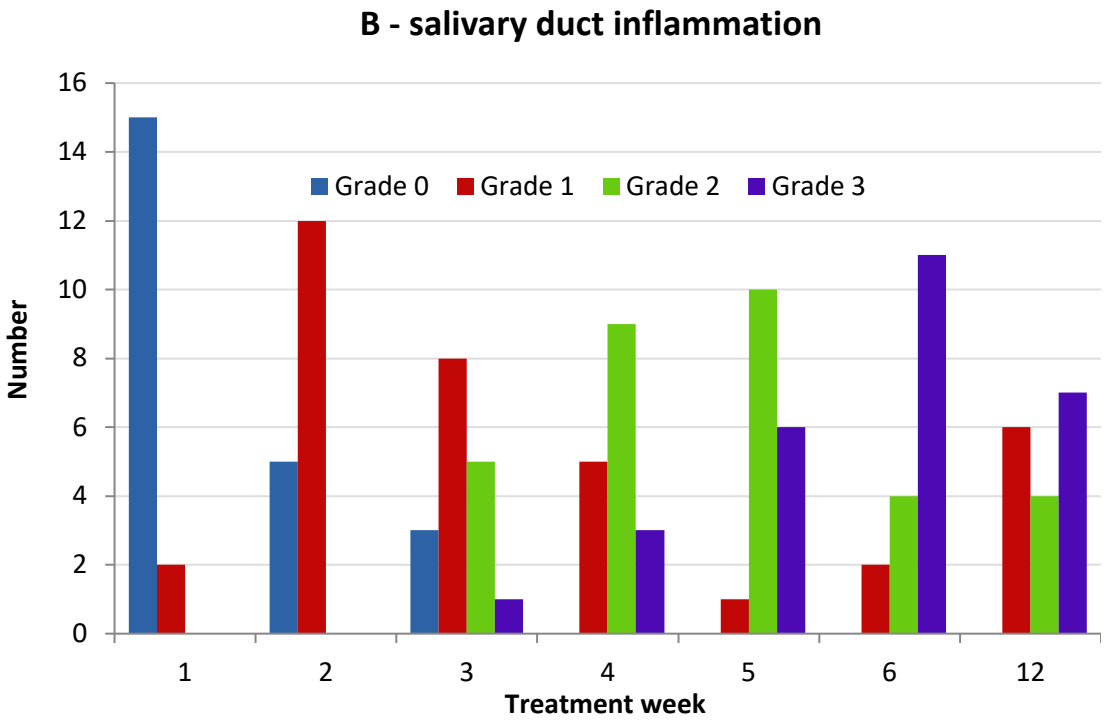


Figure 8.11 (A&B): Weekly CTCAEv4.03 acute toxicity scores for patients in the Minot-OAR sub-study. A – xerostomia, B – salivary duct inflammation.

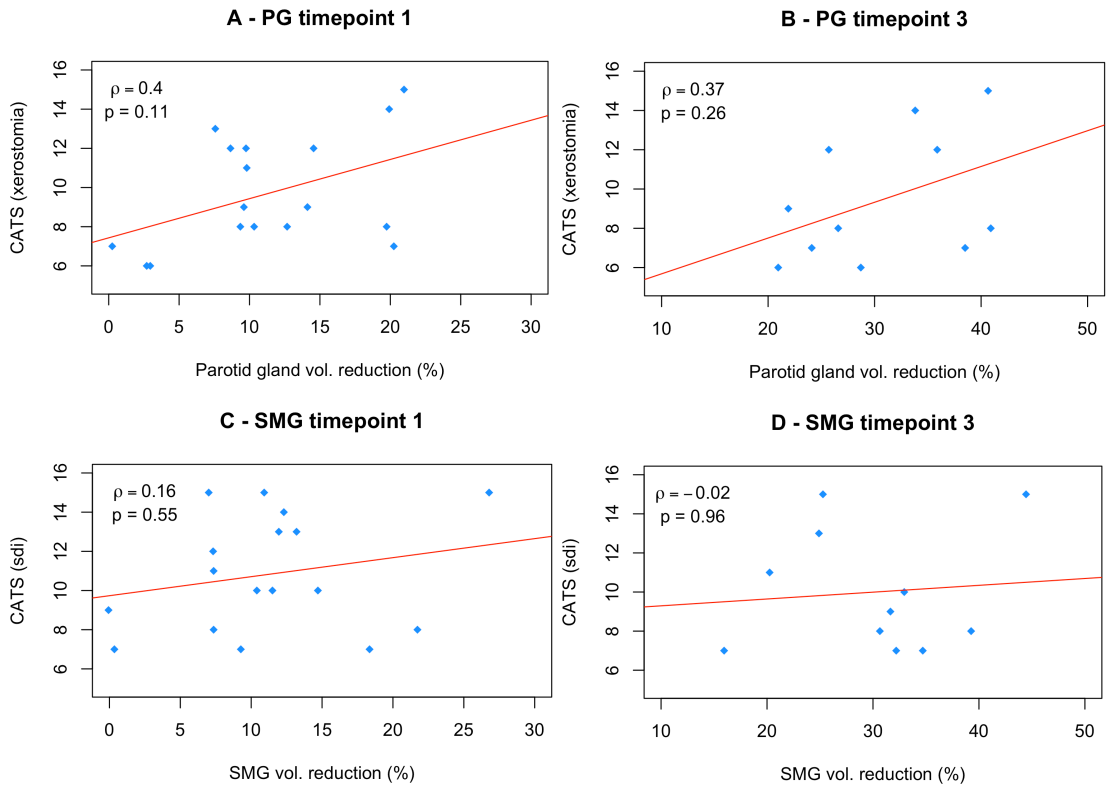


Figure 8.12 (A-D): Salivary gland volume reduction and cumulative acute toxicity (CATS). A&B – parotid gland volume reduction and xerostomia at timepoints 1 and 3. C&D – submandibular gland volume reduction and salivary duct inflammation at timepoints 1 and 3. Gland volume reduction data mean values of both paired structures, per-patient.

Data were also split into two independent groups based on a pragmatic threshold CATS of <10, and 10 and above. Gland volume reductions at each timepoint, split by this threshold, are shown in Figure 8.13. For this analysis, the volume of both glands, rather than a mean of paired structure values, were included. Wilcoxon signed rank tests were used to test the null hypothesis that the ranks of gland volume reduction were no different in both groups, and results were as follows. At timepoint 1, median parotid gland volume reduction for patients with a xerostomia CATS <10 was 9.35 (range 0.3 – 25.1), compared with 10.75 (2.5 – 21.9) for those with a score of 10 and above (p=0.44). At timepoint 3, equivalent results were 28.15 (15.8 – 46.3), and 35.15 (20.8 – 45.7) (p=0.17). For submandibular gland volume reduction and salivary duct inflammation, median values were 11 (-2 – 21.9) for <10 and 10.5 (3.4 – 28.6) for 10 and above (p=0.61) at timepoint 1, and 34.1 (27.2 – 41.5) and 30.15 (15.2 - 50) (p=0.13) at timepoint 3.

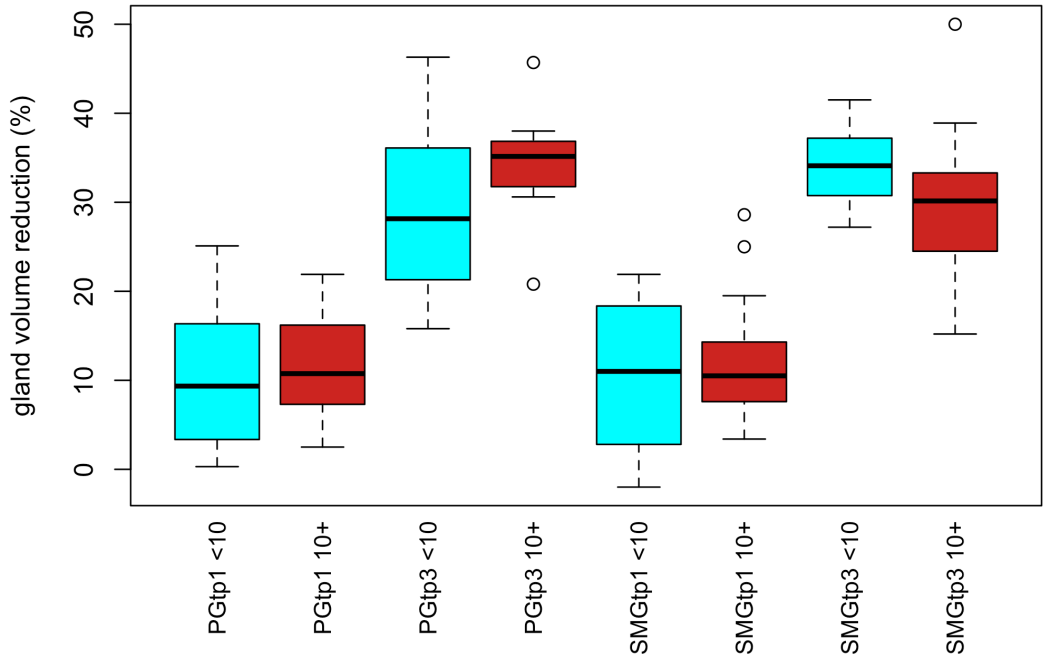


Figure 8.13: Salivary gland volume reduction split by a CATS threshold of <10 and 10 and above. Box-and-whisker plots 1-4 show parotid gland volume reduction at timepoints 1 (1 and 2) and 3 (3 and 4), with a toxicity endpoint of xerostomia. Box-and-whisker plots 5-8 show parotid gland volume reduction at timepoints 1 (5 and 6) and 3 (7 and 8), with a toxicity endpoint of salivary duct inflammation.

8.4.5 Volume changes and late toxicity

Mean reduction in salivary gland volume of patients who reported grade 0-1 toxicity at 12 months was compared with that observed in patients with grade 2+ side effects. Parotid gland volumes were assessed against xerostomia, and submandibular gland volumes with salivary duct inflammation. Results suggested that gland volume loss was greater in patients who experience Gr 2+ toxicity, but sample sizes were small, and formal statistical testing was therefore not done. Mean parotid gland volume reduction (with standard errors in parentheses) at timepoints 1 and 3, for patients with grade 0-1 and 2+ toxicity respectively were 9.7±3.9% and 13.5±2.4%, and 29.4±3.9% and 34.2±4.8%. Results for submandibular gland volume reduction were 12.8±3.1% and 14.7±2.9%, and 34.7±1.6% and 37.5±4.8% respectively. These data are shown in Figure 8.14.

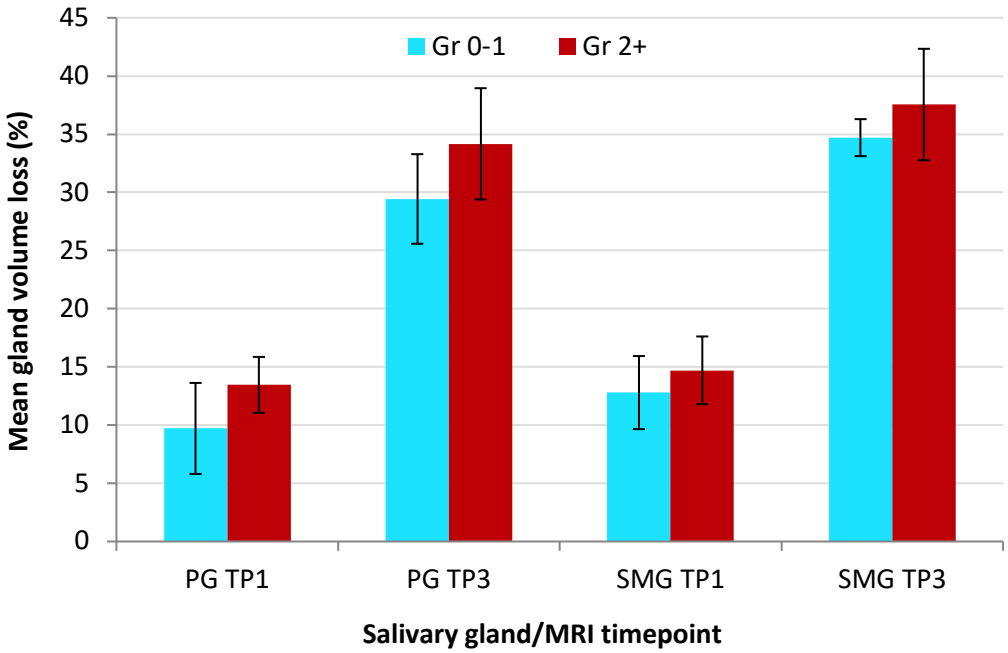


Figure 8.14: Salivary gland volume loss versus late toxicity (12 month point prevalence), at MRI timepoints 1 (TP1) and 3 (TP3). CTCAEv4.03 xerostomia for changes in parotid glands (PG), and salivary duct inflammation for submandibular glands (SMG). Cyan columns represent mean volume loss of all glands in patients with grade 0-1 toxicity, red columns those with grade 2+ toxicity. Error bars are standard error of mean value.

8.4.6 ADC changes and toxicity

A preliminary investigation of hypothesised relationships between changes in ADC values during treatment, and both acute and late toxicity, was done. Changes in mean parotid gland ADC values (mean value for paired structure) were analysed against xerostomia scores, and submandibular gland results were assessed against salivary duct inflammation data. However, these analyses were limited by the immaturity of the toxicity data, and small sample size.

For acute toxicity, matched CATS and MRI timepoint 1 ADC change data were available for 12 patients. For MRI timepoint 3, 8 datasets were available. As an initial screening analysis, data for each hypothesised relationship were plotted as scatter plots (not shown), with ADC change, and CATS as independent and dependent variables respectively. No patterns emerged for parotid gland ADC change and xerostomia CATS (R^2 <0.05 for both timepoint 1, and timepoint 3 data). There was a suggestion that patients whose SMGs had a greater rise in ADC at timepoint 1 (fraction 6), had *less* acute toxicity. Mean ADC change for patients with a CATS of <10 (n=5) was 6.2(±1.9)%, compared with -2.1(±2.4)%, for those with a CATS of 10 or greater (n=7). Due to the small sample size, no formal significance testing was

undertaken, the observation was not borne out on the MRI timepoint 3 data, and no conclusions should be drawn from this result at this stage.

Twelve-month point prevalence was used for assessment of late toxicity. Again this was a preliminary analysis given the sample size available, and statistical testing was not done. However, the data suggest that patients with lower rises in ADC values relative to baseline experience worse side effects. Mean parotid gland ADC increase (with standard errors in parentheses) at timepoints 1 and 3, for patients with grade 0-1 and 2+ toxicity respectively were 6.5±3.6% and 3.0±2.3%, and 26.4±2.5% and 21.3±5.5%. Results for submandibular gland volume reduction were 6.7±0.6% and -2.7±2.8%, and 25.0±5.8% and 11.8±10.6% respectively. These data are shown in Figure 8.15.

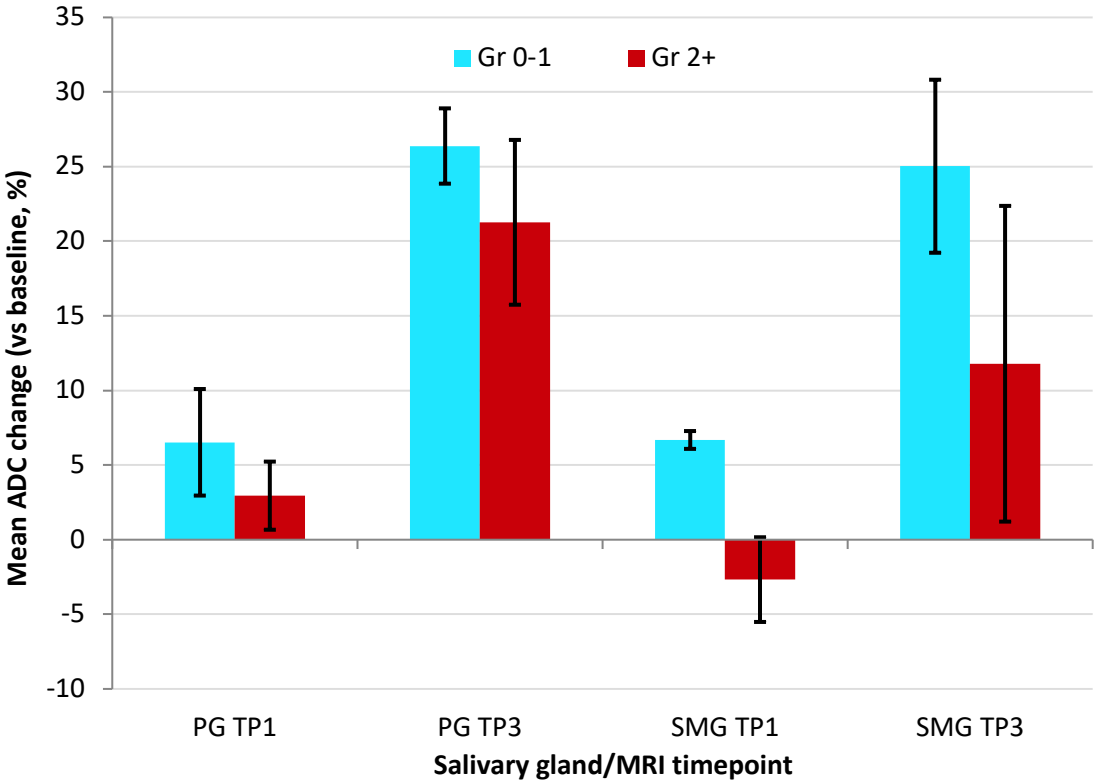


Figure 8.15: Salivary gland ADC change versus late toxicity (12 month point prevalence), at MRI timepoints 1 (TP1) and 3 (TP3). CTCAEv4.03 xerostomia for changes in parotid glands (PG), and salivary duct inflammation for submandibular glands (SMG). Cyan columns represent ADC change of all glands in patients with grade 0-1 toxicity, red columns those with grade 2+ toxicity. Error bars are standard error of mean value.

8.4.7 Inter-observer variability

Concordance between the gland volumes measured by myself (Obs 1) and a consultant radiologist (Obs 2) were generally good. For all 16 structures, the mean difference between

the measurements made by myself, and my colleague (Obs 1 – Obs 2) was 0.44mls (st dev 1.67), or 7.1% (st dev 18.0%) of the mean of the 2 observations. Mean proportional volume differences for the SMGs were larger (8.3%) than those of the parotid glands (5.9%). These data are summarised in a Bland-Altmann plot in Figure 8.16. Agreement for ADC measurements was also reassuring. Across 16 observations, the mean value of (Obs 1- Obs 2) was 0.046 x 10^{-3} s/mm^2 (st dev 0.124) as shown in a further Bland-Altmann plot in Figure 8.17. The mean proportional difference between the observations was 2.0% (st dev 6.6%).

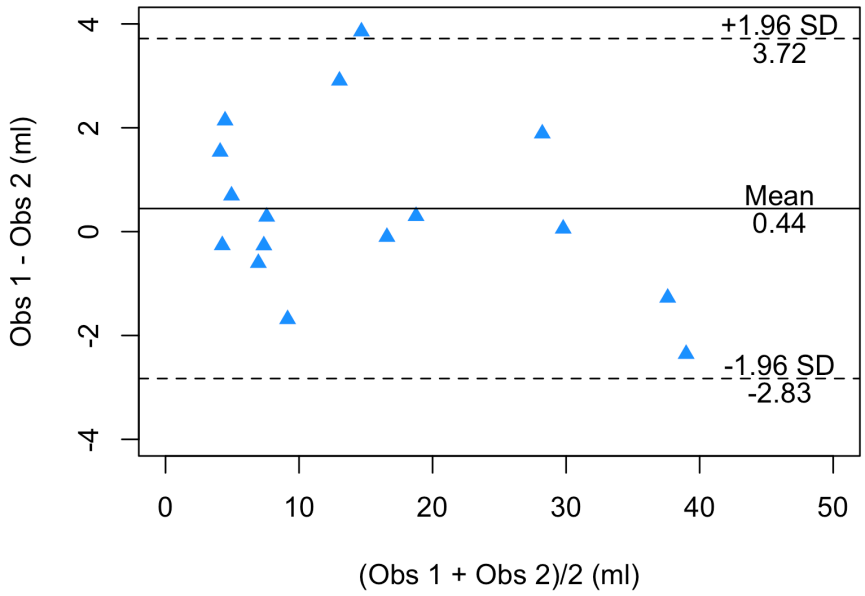


Figure 8.16: Minot-OAR sub-study, inter-observer variability – gland volume measurement. Data from 2 patients, at 2 timepoints (baseline, and timepoint 3)

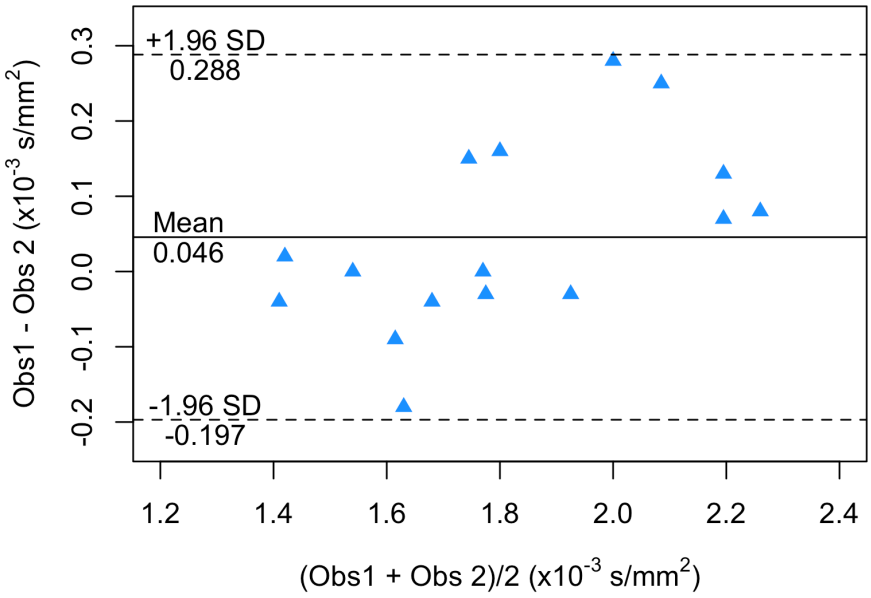


Figure 8.17: Minot-OAR sub-study, inter-observer variability – ADC measurement. Data from 2 patients, at 2 timepoints (baseline, and timepoint 3).

8.5 Discussion

8.5.1 Study recruitment

This is the first study to use MRI to investigate the major salivary glands in detail at multiple timepoints during radiotherapy for HNC. The rate of participant enrolment into the study was satisfactory, the final sample size of 18 matched our pre-defined objective, and recruitment completed within 2 years. As anticipated, there was a significant attrition rate within the study, despite having an imaging protocol carefully designed to minimise patient discomfort and time on the scanner. Only 11 of the 18 patients who underwent the baseline MRI were well enough to have the final MRI at fraction 26.

8.5.2 Gland volume change

A convincing pattern of volume decrease over time was apparent for both parotid and submandibular glands. As shown by the data in Figure 8.4, the rate at which volume is lost in both structures appears to follow an exponential decay curve. This finding fits with in-vitro and animal models of cytotoxicity [43, 48, 65, 66], but it is the first time it has been convincingly shown in human patients, because of the need for high-quality, multiple timepoint imaging during treatment. Interestingly, the mean parotid gland volume loss of 18.2% at fraction 10 measured by Marzi and colleagues fits extremely well with the data presented here [27].

It is also clear that in both structures, and most patients, there is significant volume loss by fraction 6; mean volume loss in both structures was greater than 10%. Two important and interesting conclusions can be drawn from this observation. The first is that conformational changes in these OARs happens early in treatment, so if ART is being undertaken because of concerns about this evolving geometry, and its impact on dosimetry, it would also be logical to perform this early in treatment. Second, it suggests that if volume change transpires to be a predictive IMB of toxicity, relevant differences between patient groups may also be apparent early in a course of treatment.

The data in Figure 8.5 suggest that the relationship between mean radiotherapy dose to salivary glands, and volume reduction, is unconvincing, in contrast to previous studies. However, the results in Figure 8.6 show that in most patients, the (paired) gland receiving the higher radiation dose shrinks more. This suggests that dose is an important predictor of volume reduction, but that its effect may be weaker than that of individual radio-sensitivity. This notion

is supported by the data in Figure 8.7, which shows a convincing correlation between paired gland volume reduction in each patient, independent of dose. This further supports the widely accepted concept that individual radio-sensitivity is important, and may have a substantial impact upon toxicity phenotype [67].

8.5.3 ADC changes during treatment

The data presented in section 8.4.2 demonstrate clearly that both parotid and submandibular glands undergo substantial changes in ADC during treatment. By fraction 26, the mean increase in ADC in the parotid glands was 25.5%, whilst for the SMGs it was 22.0%. These results are broadly similar to previously published data. Juan et al found a 35.7% increase in parotid gland ADC on scans between the last fraction of radiotherapy, and 100 days later [24], whilst Loimu et al found a 22.3% in pre-stimulated gland ADC 6 months after radiation [25].

The data in Figures 8.8, and 8.10A show a visible rise in ADC by fraction 6, and this was found to statistically significant. This effect was not seen for the SMG's, where mean ADC at fraction 6 was the same as that observed at baseline ($1.78 mm/s^2$). Figure 8.10 suggests that the ADC increase of the parotid glands during treatment follows a linear trend, whilst the pattern for the SMGs is less clear.

8.5.4 Predictors of toxicity

Patterns of acute toxicity suffered by participants in the study followed a typical pattern. The data in Figure 8.8 show a gradual progression of side effect severity during radiotherapy, with partial improvement by week 12 (6 weeks post treatment). In most cases, neither volume, nor ADC change, at timepoints 1 or 3, showed a convincing relationship with acute toxicity, as measured by the CATS. A possible exception is shown in the boxplot in Figure 8.10, in which the distribution of volume changes for patients with a CATS of 10+ (median 35.2%) appears to be different to those with a CATS <10 (median 28.2%). However, this was not significant on Mann-Whitney U test ($p=0.16$). Furthermore, the same plot suggests a possible relationship in the other direction between SMG volume change and SDI CATS, with less gland shrinkage in patients with worse side effects. Therefore, it must be concluded that no clear or significant relationship between gland volume change and acute side effects were observed in this study.

Relationships between IMBs and late toxicity are potentially more interesting and clinically significant. However, the immature toxicity data, and incomplete ADC dataset, mean that this was only a preliminary analysis, and no firm conclusions can yet be drawn. The data in Figure

8.11 suggest that there was more gland shrinkage in patients who experienced Gr2+ toxicity compared with those who did not. This is similar to patterns seen in a similar study [24], but it must be re-iterated that the sample size was small, as were absolute differences between the 2 groups, and no formal significance testing was done – therefore no clear conclusions can be drawn from this preliminary analysis.

As the toxicity data matures, 12 month point prevalence data will become available for the full cohort, increasing the statistical power of the analysis. Furthermore, data for later timepoints will also accumulate, potentially aiding future analyses. As shown in Chapter 3, the point prevalence of both the toxicity endpoints used in this Chapter gradually declines over time. It could therefore be argued that toxicity measured at later timepoints may be more useful for distinguishing between those patients with a truly 'severe' phenotype, and those in whom there has been some recovery of normal tissue architecture and function. If this is the case, it may relationships with the IMBs measured in this study are clearer and stronger.

Intriguingly, the results seen in Figure 8.12 hint at a relationship between ADC and late toxicity events counter to published results [24, 25, 68]. Previous authors have presented data suggesting that higher parotid gland ADC on DWI done post-radiotherapy correlates with worse toxicity [24, 25, 68]. In contrast the data shown in Figure 8.12 infer the opposite, that patients who experience a greater rise in gland ADC during treatment (at fraction 6 and 26) are *less* likely to experience significant late toxicity. Again, this is a preliminary analysis, with immature toxicity data, with an incomplete ADC dataset, and because of this, tests for statistical significance have not yet been done, and no firm conclusions can yet be drawn from these results.

However, if this pattern in the data persists with a larger sample size, and statistical significance is found, it may not be conceptually incompatible with previous results. Juan et al's study of ADC and toxicity in patients treated for nasopharyngeal carcinoma found higher ADC values in the parotid glands of patients with more xerostomia [24]. However, this measurement was on DWI done more than 1-year post treatment. The authors suggest a number of possible mechanisms for this observation, including loss of gland acini, and vasogenic oedema [24]. Although acute inflammation leads to increased vascular permeability, which in turn should lead to higher ADC values, authors have shown that acute inflammation can also cause a fall in ADC, in conditions such as acute Multiple Sclerosis [69], and acute demyelinating encephalomyelitis [70], as increasing cellularity from immune infiltrates restrict the Brownian motion of water. Therefore, one possible explanation is that in the hyper-acute phase (during treatment), radiation induced parenchymal injury leads to an acute inflammatory

reaction with an immune infiltrate and a fall in ADC. Over time, this regresses, and the loss of acini reduces the restriction of diffusion, leading to a rise in ADC. If this proves to be the case, it further suggests that DWI early in treatment may not prove to be a specific IMB of salivary gland related side effects, because of the variability of ADC changes in the presence of acute inflammation [69, 70].

8.5.5 Study strengths and weaknesses

A major strength of this study is the scheduling of 3 MRI scans during treatment, and the fact that all imaging was done on a 3T scanner. The specific scan timing protocol is also strength, reducing possible noise from day-to-day changes in imaging parameters. The cohort is homogeneous in terms of primary disease site, radiotherapy treatment details and the use of concomitant therapy. The fact that inter-observer variability for both volume and ADC measurements was tested, and found to be low and acceptable is also a strength. Weaknesses include the use of only 2 b-values for DWI sequences, although this decision was based on similar previous work [25]. The use of a new and untested technique for salivary gland stimulation is a potential weakness, although it was necessitated by the ethical concerns of using established methods in the acute setting. Immature late toxicity data mean that analyses with these endpoints were only preliminary, and will be repeated as data mature. Finally, the participant attrition rate through the study, whilst anticipated, did limit the sample size available for timepoint 3 images.

8.6 Conclusions

This is the first study to use MRI to investigate changes in the major salivary glands at multiple timepoints during radiotherapy for HNC. The data show that both parotid and submandibular glands lose volume over treatment, and that the rate of volume change is highest early in treatment. Similarly, ADC values in both structures increased steadily during therapy, with a measurable difference between mean ADC at baseline and fraction 6, observed for the parotid glands. No associations between these imaging metrics and acute toxicity were observed, and data were insufficiently mature to permit full analysis of late toxicity results.

8.7 References:

[1] Heukelom J, Fuller CD. Head and Neck Cancer Adaptive Radiation Therapy (ART): Conceptual Considerations for the Informed Clinician. Semin Radiat Oncol. 2019;29:258-73.
[2] Jentsch C, Beuthien-Baumann B, Troost EG, Shakirin G. Validation of functional imaging as a biomarker for radiation treatment response. Br J Radiol. 2015;88:20150014.
[3] Heukelom J, Hamming O, Bartelink H, Hoebers F, Giralt J, Herlestam T, et al. Adaptive and innovative Radiation Treatment FOR improving Cancer treatment outcomE (ARTFORCE); a randomized controlled phase II trial for individualized treatment of head and neck cancer. BMC Cancer. 2013;13:84.
[4] Paterson C, Allwood-Spiers S, McCrea I, Foster J, McJury M, Thomson M, et al. Study of diffusion weighted MRI as a predictive biomarker of response during radiotherapy for high and intermediate risk squamous cell cancer of the oropharynx: The MeRInO study. Clin Transl Radiat Oncol. 2017;2:13-8.
[5] Dirix P, Vandecaveye V, De Keyzer F, Stroobants S, Hermans R, Nuyts S. Dose painting in radiotherapy for head and neck squamous cell carcinoma: value of repeated functional imaging with (18)F-FDG PET, (18)F-fluoromisonidazole PET, diffusion-weighted MRI, and dynamic contrast-enhanced MRI. J Nucl Med. 2009;50:1020-7.
[6] Houweling AC, Wolf AL, Vogel WV, Hamming-Vrieze O, van Vliet-Vroegindeweij C, van de Kamer JB, et al. FDG-PET and diffusion-weighted MRI in head-and-neck cancer patients: implications for dose painting. Radiother Oncol. 2013;106:250-4.
[7] Bittner MI, Wiedenmann N, Bucher S, Hentschel M, Mix M, Weber WA, et al. Exploratory geographical analysis of hypoxic subvolumes using 18F-MISO-PET imaging in patients with head and neck cancer in the course of primary chemoradiotherapy. Radiother Oncol. 2013;108:511-6.
[8] Newbold K, Castellano I, Charles-Edwards E, Mears D, Sohaib A, Leach M, et al. An exploratory study into the role of dynamic contrast-enhanced magnetic resonance imaging or perfusion computed tomography for detection of intratumoral hypoxia in head-and-neck cancer. Int J Radiat Oncol Biol Phys. 2009;74:29-37.
[9] Chikui T, Kawano S, Kawazu T, Hatakenaka M, Koga S, Ohga M, et al. Prediction and monitoring of the response to chemoradiotherapy in oral squamous cell carcinomas using a pharmacokinetic analysis based on the dynamic contrast-enhanced MR imaging findings. Eur Radiol. 2011;21:1699-708.
[10] Shukla-Dave A, Lee NY, Jansen JF, Thaler HT, Stambuk HE, Fury MG, et al. Dynamic contrast-enhanced magnetic resonance imaging as a predictor of outcome in head-and-neck squamous cell carcinoma patients with nodal metastases. Int J Radiat Oncol Biol Phys. 2012;82:1837-44.
[11] Agrawal S, Awasthi R, Singh A, Haris M, Gupta RK, Rathore RK. An exploratory study into the role of dynamic contrast-enhanced (DCE) MRI metrics as predictors of response in head and neck cancers. Clin Radiol. 2012;67:e1-5.
[12] Bernstein JM, Homer JJ, West CM. Dynamic contrast-enhanced magnetic resonance imaging biomarkers in head and neck cancer: potential to guide treatment? A systematic review. Oral Oncol. 2014;50:963-70.
[13] Kim S, Loevner L, Quon H, Sherman E, Weinstein G, Kilger A, et al. Diffusion-weighted magnetic resonance imaging for predicting and detecting early response to chemoradiation therapy of squamous cell carcinomas of the head and neck. Clin Cancer Res. 2009;15:986-94.
[14] Vandecaveye V, Dirix P, De Keyzer F, de Beeck KO, Vander Poorten V, Roebben I, et al. Predictive value of diffusion-weighted magnetic resonance imaging during chemoradiotherapy for head and neck squamous cell carcinoma. Eur Radiol. 2010;20:1703-14.
[15] Hatakenaka M, Nakamura K, Yabuuchi H, Shioyama Y, Matsuo Y, Ohnishi K, et al. Pretreatment apparent diffusion coefficient of the primary lesion correlates with local failure in head-and-neck cancer treated with chemoradiotherapy or radiotherapy. Int J Radiat Oncol Biol Phys. 2011;81:339-45.
[16] Vandecaveye V, Dirix P, De Keyzer F, Op de Beeck K, Vander Poorten V, Hauben E, et al. Diffusion-weighted magnetic resonance imaging early after chemoradiotherapy to monitor treatment response in head-and-neck squamous cell carcinoma. Int J Radiat Oncol Biol Phys. 2012;82:1098-107.
[17] Lambrecht M, Van Herck H, De Keyzer F, Vandecaveye V, Slagmolen P, Suetens P, et al. Redefining the target early during treatment. Can we visualize regional differences within the target volume using sequential diffusion weighted MRI? Radiother Oncol. 2014;110:329-34.
[18] Hauser T, Essig M, Jensen A, Laun FB, Munter M, Maier-Hein KH, et al. Prediction of treatment response in head and neck carcinomas using IVIM-DWI: Evaluation of lymph node metastasis. Eur J Radiol. 2014;83:783-7.

[19] Chen Y, Liu X, Zheng D, Xu L, Hong L, Xu Y, et al. Diffusion-weighted magnetic resonance imaging for early response assessment of chemoradiotherapy in patients with nasopharyngeal carcinoma. Magn Reson Imaging. 2014;32:630-7.
[20] Matoba M, Tuji H, Shimode Y, Toyoda I, Kuginuki Y, Miwa K, et al. Fractional change in apparent diffusion coefficient as an imaging biomarker for predicting treatment response in head and neck cancer treated with chemoradiotherapy. AJNR Am J Neuroradiol. 2014;35:379-85.
[21] Doornaert P, Dahele M, Ljumanovic R, de Bree R, Slotman BJ, Castelijns JA. Use of diffusion-weighted magnetic resonance imaging (DW-MRI) to investigate the effect of chemoradiotherapy on the salivary glands. Acta Oncol. 2015;54:1068-71.
[22] Juan CJ, Chen CY, Jen YM, Liu HS, Liu YJ, Hsueh CJ, et al. Perfusion characteristics of late radiation injury of parotid glands: quantitative evaluation with dynamic contrast-enhanced MRI. Eur Radiol. 2009;19:94-102.
[23] Houweling AC, Schakel T, van den Berg CA, Philippens ME, Roesink JM, Terhaard CH, et al. MRI to quantify early radiation-induced changes in the salivary glands. Radiother Oncol. 2011;100:386-9.
[24] Juan CJ, Cheng CC, Chiu SC, Jen YM, Liu YJ, Chiu HC, et al. Temporal Evolution of Parotid Volume and Parotid Apparent Diffusion Coefficient in Nasopharyngeal Carcinoma Patients Treated by Intensity-Modulated Radiotherapy Investigated by Magnetic Resonance Imaging: A Pilot Study. PLoS One. 2015;10:e0137073.
[25] Loimu V, Seppala T, Kapanen M, Tuomikoski L, Nurmi H, Makitie A, et al. Diffusion-weighted magnetic resonance imaging for evaluation of salivary gland function in head and neck cancer patients treated with intensity-modulated radiotherapy. Radiother Oncol. 2017;122:178-84.
[26] Wang ZH, Yan C, Zhang ZY, Zhang CP, Hu HS, Kirwan J, et al. Radiation-induced volume changes in parotid and submandibular glands in patients with head and neck cancer receiving postoperative radiotherapy: a longitudinal study. Laryngoscope. 2009;119:1966-74.
[27] Marzi S, Farneti A, Vidiri A, Di Giuliano F, Marucci L, Spasiano F, et al. Radiation-induced parotid changes in oropharyngeal cancer patients: the role of early functional imaging and patient-/treatment-related factors. Radiat Oncol. 2018;13:189.
[28] Zhang Y, Ou D, Gu Y, He X, Peng W, Mao J, et al. Diffusion-weighted MR imaging of salivary glands with gustatory stimulation: comparison before and after radiotherapy. Acta Radiol. 2013;54:928-33.
[29] Dirix P, De Keyzer F, Vandecaveye V, Stroobants S, Hermans R, Nuyts S. Diffusion-weighted magnetic resonance imaging to evaluate major salivary gland function before and after radiotherapy. Int J Radiat Oncol Biol Phys. 2008;71:1365-71.
[30] Marzi S, Forina C, Marucci L, Giovinazzo G, Giordano C, Piludu F, et al. Early radiation-induced changes evaluated by intravoxel incoherent motion in the major salivary glands. J Magn Reson Imaging. 2015;41:974-82.
[31] Koh DM, Padhani AR. Diffusion-weighted MRI: a new functional clinical technique for tumour imaging. Br J Radiol. 2006;79:633-5.
[32] Zhang L, Murata Y, Ishida R, Ohashi I, Yoshimura R, Shibuya H. Functional evaluation with intravoxel incoherent motion echo-planar MRI in irradiated salivary glands: a correlative study with salivary gland scintigraphy. J Magn Reson Imaging. 2001;14:223-9.
[33] Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology. 1986;161:401-7.
[34] Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology. 1988;168:497-505.
[35] Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med. 2005;53:1432-40.
[36] Jensen JH, Helpern JA. MRI quantification of non-Gaussian water diffusion by kurtosis analysis. NMR Biomed. 2010;23:698-710.
[37] Lazar M, Jensen JH, Xuan L, Helpern JA. Estimation of the orientation distribution function from diffusional kurtosis imaging. Magn Reson Med. 2008;60:774-81.
[38] Astreinidou E, Roesink JM, Raaijmakers CP, Bartels LW, Witkamp TD, Lagendijk JJ, et al. 3D MR sialography as a tool to investigate radiation-induced xerostomia: feasibility study. Int J Radiat Oncol Biol Phys. 2007;68:1310-9.
[39] Zhou N, Chen W, Pan X, He J, Yan J, Zhou Z, et al. Early evaluation of radiation-induced parotid damage with diffusion kurtosis imaging: a preliminary study. Acta Radiol. 2018;59:212-20.

[40] Qian W, Xu XQ, Zhu LN, Ma G, Su GY, Bu SS, et al. Preliminary study of using diffusion kurtosis imaging for characterizing parotid gland tumors. Acta Radiol. 2019;60:887-94.
[41] Zhang Q, Wei YM, Qi YG, Li BS. Early Changes in Apparent Diffusion Coefficient for Salivary Glands during Radiotherapy for Nasopharyngeal Carcinoma Associated with Xerostomia. Korean J Radiol. 2018;19:328-33.
[42] Vissink A, van Luijk P, Langendijk JA, Coppes RP. Current ideas to reduce or salvage radiation damage to salivary glands. Oral Dis. 2015;21:e1-10.
[43] Konings AW, Coppes RP, Vissink A. On the mechanism of salivary gland radiosensitivity. Int J Radiat Oncol Biol Phys. 2005;62:1187-94.
[44] Nagler RM. The enigmatic mechanism of irradiation-induced damage to the major salivary glands. Oral Dis. 2002;8:141-6.
[45] Stephens LC, King GK, Peters LJ, Ang KK, Schultheiss TE, Jardine JH. Acute and late radiation injury in rhesus monkey parotid glands. Evidence of interphase cell death. Am J Pathol. 1986;124:469-78.
[46] Radfar L, Sirois DA. Structural and functional injury in minipig salivary glands following fractionated exposure to 70 Gy of ionizing radiation: an animal model for human radiation-induced salivary gland injury. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2003;96:267-74.
[47] Li J, Shan Z, Ou G, Liu X, Zhang C, Baum BJ, et al. Structural and functional characteristics of irradiation damage to parotid glands in the miniature pig. Int J Radiat Oncol Biol Phys. 2005;62:1510-6.
[48] van Luijk P, Pringle S, Deasy JO, Moiseenko VV, Faber H, Hovan A, et al. Sparing the region of the salivary gland containing stem cells preserves saliva production after radiotherapy for head and neck cancer. Sci Transl Med. 2015;7:305ra147.
[49] Ho KF, Marchant T, Moore C, Webster G, Rowbottom C, Penington H, et al. Monitoring dosimetric impact of weight loss with kilovoltage (kV) cone beam CT (CBCT) during parotid-sparing IMRT and concurrent chemotherapy. Int J Radiat Oncol Biol Phys. 2012;82:e375-82.
[50] Wang X, Lu J, Xiong X, Zhu G, Ying H, He S, et al. Anatomic and dosimetric changes during the treatment course of intensity-modulated radiotherapy for locally advanced nasopharyngeal carcinoma. Med Dosim. 2010;35:151-7.
[51] Broggi S, Fiorino C, Dell'Oca I, Dinapoli N, Paiusco M, Muraglia A, et al. A two-variable linear model of parotid shrinkage during IMRT for head and neck cancer. Radiother Oncol. 2010;94:206-12.
[52] Sanguineti G, Ricchetti F, Thomas O, Wu B, McNutt T. Pattern and predictors of volumetric change of parotid glands during intensity modulated radiotherapy. Br J Radiol. 2013;86:20130363.
[53] van Dijk LV, Brouwer CL, van der Schaaf A, Burgerhof JG, Beukinga RJ, Langendijk JA, et al. CT image biomarkers to improve patient-specific prediction of radiation-induced xerostomia and sticky saliva. Radiother Oncol. 2017;122:185-91.
[54] van Dijk LV, Thor M, Steenbakkers R, Apte A, Zhai TT, Borra R, et al. Parotid gland fat related Magnetic Resonance image biomarkers improve prediction of late radiation-induced xerostomia. Radiother Oncol. 2018;128:459-66.
[55] van Dijk LV, Noordzij W, Brouwer CL, Boellaard R, Burgerhof JGM, Langendijk JA, et al. (18)F-FDG PET image biomarkers improve prediction of late radiation-induced xerostomia. Radiother Oncol. 2018;126:89-95.
[56] Thoeny HC, De Keyzer F, Boesch C, Hermans R. Diffusion-weighted imaging of the parotid gland: Influence of the choice of b-values on the apparent diffusion coefficient value. J Magn Reson Imaging. 2004;20:786-90.
[57] Habermann CR, Gossrau P, Kooijman H, Graessner J, Cramer MC, Kaul MG, et al. Monitoring of gustatory stimulation of salivary glands by diffusion-weighted MR imaging: comparison of 1.5T and 3T. AJNR Am J Neuroradiol. 2007;28:1547-51.
[58] Wooley OW, Wooley SC. Relationship of Salivation in Humans to Deprivation, Inhibition and the Encephalization of Hunger. Appetite. 1981;2:331-50.
[59] Spence C. Mouth-Watering: The Influence of Environmental and Cognitive Factors on Salivation and Gustatory/Flavor Perception. J Texture Stud. 2011;42:157-71.
[60] Keesman M, Aarts H, Vermeent S, Hafner M, Papies EK. Consumption Simulations Induce Salivation to Food Cues. Plos One. 2016;11.
[61] Kise Y, Chikui T, Yamashita Y, Kobayashi K, Yoshiura K. Clinical usefulness of the mDIXON Quant the method for estimation of the salivary gland fat fraction: comparison with MR spectroscopy. Br J Radiol. 2017;90:20160704.
[62] Zhou N, Chu C, Dou X, Chen W, He J, Yan J, et al. Early evaluation of radiation-induced parotid damage in patients with nasopharyngeal carcinoma by T2 mapping and mDIXON Quant imaging: initial findings. Radiat Oncol. 2018;13:22.

[63] Thomson D, Yang H, Baines H, Miles E, Bolton S, West C, 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;26:344-7.
[64] Owadally W, Hurt C, Timmins H, Parsons E, Townsend S, Patterson J, 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;15:602.
[65] Vissink A, s-Gravenmade EJ, Ligeon EE, Konings WT. A functional and chemical study of radiation effects on rat parotid and submandibular/sublingual glands. Radiat Res. 1990;124:259-65.
[66] Paardekooper GM, Cammelli S, Zeilstra LJ, Coppes RP, Konings AW. Radiation-induced apoptosis in relation to acute impairment of rat salivary gland function. Int J Radiat Biol. 1998;73:641-8.
[67] Barnett GC, West CM, Dunning AM, Elliott RM, Coles CE, Pharoah PD, et al. Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nat Rev Cancer. 2009;9:134-42.
[68] Zhou N, Guo T, Zheng H, Pan X, Chu C, Dou X, et al. Apparent diffusion coefficient histogram analysis can evaluate radiation-induced parotid damage and predict late xerostomia degree in nasopharyngeal carcinoma. Oncotarget. 2017;8:70226-38.
[69] Tievsky AL, Ptak T, Farkas J. Investigation of apparent diffusion coefficient and diffusion tensor anisotrophy in acute and chronic multiple sclerosis lesions. AJNR Am J Neuroradiol. 1999;20:1491-9.
[70] Bernarding J, Braun J, Koennecke HC. Diffusion- and perfusion-weighted MR imaging in a patient with acute demyelinating encephalomyelitis (ADEM). J Magn Reson Imaging. 2002;15:96-100.

Chapter 9 – Summary & directions for future research

9.1 Summary of presented work

Radiotherapy treatment for HNC has evolved almost beyond recognition in recent decades. The widespread implementation of IMRT, along with ever-better IG techniques has helped to reduce treatment toxicity [1]. However, the incidence of HNC is rising, driven by the HPV epidemic [2, 3]. It is now known that cure rates for many of these patients are expected to be higher than historical series, independent of treatment techniques, because of the inherent biological characteristics of the disease [4]. Concurrent with this, radiation-based radical treatment regimens for HNC are toxic, the prevalence of moderate to severe late sequelae remains unacceptably high, and there is consensus within the community that de-escalation strategies are urgently needed [5-7]. ART allows treatment teams to modify radiotherapy plans once a course of treatment is underway, and is a promising option for achieving this objective [8]. ART may be achieved with a variety of techniques, and for different reasons, many of which are based on the fundamental underlying premise that the dose of radiation that is finally delivered to important structures such as the tumour target, and healthy organs at risk, is different to that which was anticipated at treatment planning [9]. However, with contemporary workflows it is laborious, resource intensive, and available data suggest that it will have minimal benefit for many patients [10]. Therefore, research to better understand the potential dosimetric implications of ART, and to find ways of intelligently targeting this approach to the patients who will benefit most, is much needed [11].

The CRUK-funded VoxTox programme recruited patients with prostate and CNS tumours, as well as those undergoing radical treatment for HNC. The over-arching objectives of the study were to calculate both planned and delivered radiation dose to OARs in patients recruited to the study, to prospectively collect high-resolution toxicity data, and analyse relationships between the two. This was technically possible because of the deployment of 2 TomoTherapy units at Addenbrooke's Hospital in 2007 and 2009, and the use of daily MVCT IG.

Chapter 2 provides technical background on TomoTherapy technology, details as to the design and conduct of the VoxTox study, and standard treatment pathways for patients with HNC in Addenbrooke's Hospital. The study opened to recruitment in April 2013, and closed in April 2019. During this time, a total of 337 patients were recruited to the VoxTox head & neck cohort. For 319 of these patients, all available imaging and radiotherapy treatment data were anonymised and stored at the Cavendish laboratory. For 253 of these patients, I undertook

manual contouring of the following OARs; bilateral parotid glands, bilateral submandibular glands, superior and middle pharyngeal constrictor muscles, oral cavity, and supraglottic larynx, and spinal contours for 133 patients. These segmentations permitted calculation of planned dose, and a crucial step in the process of calculating delivered dose. Conformity metrics of my contours relative to other expert observers, and intra-observer consistency with myself were consistently good, with most DSC scores >0.8.

Chapter 3 describes the methodologies by which treatment toxicity data were collected, stored and processed. Because recent trials and modelling studies in HNC have used a wide range of toxicity endpoints, VoxTox investigators aimed to ensure that results could be presented with as many of these validated scoring systems as possible. Therefore, raw data in the study were collected using study specific reporting forms, populated at interviews or telephone calls between study participants, and the research radiographer team. A parsing code, based on a series of human interpretable logical rules written by me, translated these data into scoring system results, in an automated fashion. Point prevalence, and cumulative incidence to 5 years follow up are presented for the following toxicity endpoints: CTCAEv4.03 xerostomia, salivary duct inflammation and dysphagia, LENT/SOM subjective xerostomia, xerostomia management, subjective dysphagia, and subjective taste alteration, RTOG salivary gland and oesophagus, and EORTC H&N-35 question 41 – dry mouth, question 42 – sticky saliva, and question 44 – taste disturbance. Peak point prevalence of all toxicities was found at 6 months post-treatment, and gradually improved thereafter. Results were found to be broadly consistent with those reported in the literature [1, 12-15].

The process of calculating delivered dose is non-trivial, and controversial [16-18]. Amongst a number of problems was the challenge of providing accurate segmentation of multiple OARs, on around 10000 images, a task that clearly required substantial automation. The most widely used methodology in this field is deformable image registration (DIR), a technique that attempts to provide a common geometric frame of reference between images at different timepoints in treatment in a way that reflects underlying anatomical changes. Chapter 4 describes the process by which an open-source DIR toolbox (Elastix) was trained, tested, and validated to map planning kVCT and IG MVCT images, together. By providing this geometric frame of reference, my manual contours performed on planning CT images could be transposed to image guidance scans, permitting the calculation of daily dose, which could then be accumulated to an estimation of total delivered dose. Experiments testing the performance of Elastix, whilst systematically changing both assumptions and parameters within the algorithm were performed, and a pragmatic optimum found for use within the rest of the project.

According to my clinical experience, and the literature [19-21], a common reason for replanning radiotherapy for HNC patients during a course of treatment is concern about excess dose to the spinal cord. Often, this is due to an a-priori assumption that weight loss and anatomical change during treatment will lead to delivered doses to the spinal cord that are higher than planned. In Chapter 5, delivered dose to the spinal cord was calculated, compared to planned dose, and this hypothesis was directly tested. Data from 133 patients was used of this analysis, and dose differences were found to be small. Mean absolute difference between planned and delivered dose in this cohort was 0.9Gy, and only 4 patients had delivered spinal cord dose >2Gy higher than planned. Calculated delivered spinal cord dose did not breach organ tolerance in any of the patients in this cohort. Importantly, no relationships between weight loss or any other measured metric of anatomical change, were found. The only factor that did predict for higher spinal cord delivered dose was a steeper dose gradient in the vicinity of the spinal cord. The work presented in Chapter 5 was published in Radiotherapy & Oncology [22].

The same methodology was used to compute both planned and delivered dose to other normal tissue, the so-called 'swallowing OARS'; parotid glands, submandibular glands, superior and middle pharyngeal constrictor muscles, oral cavity, and supraglottic larynx. A larger cohort of 253 patients was used for this work, which is presented in Chapter 6. The data in this Chapter show that delivered dose is greater than planned for all of the swallowing OARs, with most (mean) differences in the order of 1Gy. The greatest difference was to the ipsilateral parotid gland (1.56Gy, 95%CI 1.37 – 1.74), and the least was seen by the oral cavity (0.44Gy, 95%CI 0.30 – 0.57). These data broadly concur with previous studies, although the magnitudes of differences for some OARs were smaller than has been previously observed.

Univariate analyses were undertaken to look at individual factors that may influence whether or not delivered dose is higher than planned to specific organs. On ANOVA, patients with primary disease of the nasopharynx (NPC) were more likely to have higher delivered doses to the parotid glands. Although the sample size was small (there were only 8 patients with NPC in the study), these results fit with previously published data [23, 24]. Both weight loss and anatomical change during treatment showed statistically significant associations with higher delivered dose to the midline OARs (pharyngeal constrictor muscles and supraglottic larynx), and SMGs. However, relationships were relatively weak (Pearson's product moment correlation coefficient, r, ranged from 0.11 – 0.43). Other factors that consistently predicted for higher OAR dose were higher N-stage, and the use of concomitant systemic therapy. There was also a suggestion that pre-radiotherapy neck dissection may be protective for some OARs. However, it should be noted that there may be co-linearity between these factors, and multivariate analysis is needed to improve characterization of the factors that predict clinically

significant dose differences between planned and delivered. This is discussed in greater detail in section 9.2.

Despite this, work was done to try and find common factors in cases with clinically relevant dose differences. Because radical radiotherapy for HNC is generally delivered in fraction sizes of around 2Gy, a dose difference of 2Gy or greater was deemed to be clinically significant, because it implies that the OAR in question is receiving the equivalent of an extra fraction of treatment. There is also precedent for this approach [25]. No patients had a dose difference of 2Gy or more to all 8 swallowing OARs, but 9 of 253 patients (3.7%) had a delivered minus planned dose difference of 2Gy or more, to more than half (5/8) of the OARs considered. All of these patients had advanced nodal disease in the neck (N2b or greater), and received radiotherapy to both sides of the neck. Eight of 9 also received concomitant cisplatin chemotherapy.

All of this work depends on the premise that radiotherapy dose to OARs is a crucial biomarker of treatment toxicity. By this logic, it could be hypothesised that delivered radiation dose to OARs should predict toxicity outcomes more accurately than planned dose. This was a core hypothesis of the VoxTox study, and was directly tested in Chapter 7. Because a number of factors including patient demographics, co-morbidities, and disease and treatment characteristics may influence toxicity events in addition to radiotherapy dose, a multivariate analysis was used for this work. The standard approach to this problem in the literature is logistic regression-based Normal Tissue Complication Probability (NTCP) modeling [14, 15, 26], and this method was employed in this work. NTCP models predicting the occurrence of moderate to severe symptoms of dry mouth, changes to the quality and consistency of saliva, and swallowing difficulties, 1-year post treatment, were trained and validated for 8 toxicity endpoints. In the first step, this was done using planned dose to OARs. The process was then repeated using delivered dose metrics, and the performance of planned and delivered dose-based models were compared and contrasted. Overall model performance was satisfactory, and comparable to published results [14, 15, 27]. Although absolute results for delivered dose-base models were marginally superior to planned for most endpoints, the models themselves were very similar, with substantial overlap of performance metric confidence intervals, and no statistically significant differences were observed. These results were not surprising given the small differences between planned and delivered dose results themselves, and emphasise that ART for toxicity reduction is unlikely to be useful in most patients.

Chapter 8 presents preliminary data from Minot-OAR, a substudy of the broader VoxTox programme. This study approaches the same problem – how best to select patients for ART

to reduce toxicity – from a different angle. The concept was to use both anatomical and functional MRI sequences to better characterise and understand the changes that salivary glands undergo during a course of radiotherapy treatment, and to identify imaging features that might predict toxicity events, by scanning patients at multiple timepoints during treatment. To do this, a cohort of 18 patients were invited to be scanned on a 3T MRI scanner with an anatomical T2 sequence, as well as diffusion weighted and fat-fraction quantifying sequences, at fractions 1, 6, 16 and 26 of radiotherapy. Data from this study confirm previous findings that the salivary glands shrink, but also show that apparent diffusion coefficient (ADC) rises, as patients progress through their radiotherapy treatment. Interestingly, these changes appear to start as early as the first week. No clinically, or statistically significant associations between imaging parameters and cumulative acute radiation toxicity over the full course of treatment were observed. Immature late toxicity data meant that analyses of imaging features as predictors of these endpoints were undertaken with small numbers, and were preliminary only.

9.2 Synthesis & conclusions

Until such time as new technology and workflows permit accurate, rapid, and robust on-set imaging, segmentation, and dose re-optimisation in a near fully automated fashion, ART for patients being treated for HNC will remain time, effort, and resource intensive. In addition, there is limited clinical data showing a clear benefit for this technique in unselected populations. It is therefore important to understand better what the dosimetric impact of ART might be, which patients have the greatest differences between planned and delivered dose to important OARs, and whether other imaging techniques undertaken during treatment can predict toxicity events, and thus help to identify patients who will benefit most from this approach. The key results of this thesis provide much needed data to help address some of these questions, and are listed below:

- Differences between planned and delivered maximum dose to the spinal cord in patients treated with daily image guidance are small, and rarely clinically relevant.
- Contrary to widely held a-priori assumptions; weight loss and anatomical change during treatment do not increase the risk of higher than planned dose to the spinal cord.
- Delivered dose to all other OARs assessed in this work is consistently higher than planned, although differences are small in most patients.
- A small number of patients do have clinically significant dose differences to salivary glands, oral cavity, supraglottic larynx, and swallowing muscles.

- Univariate analyses suggest that factors including N-stage, and the use of concomitant systemic therapy, in addition to weight loss and anatomical change may pre-dispose to bigger differences between planned and delivered dose.
- Toxicity prediction models based on delivered dose data have very similar performance metrics to those generated with planned dose, emphasising that dose differences across the full cohort were small, and that ART is unlikely to confer clinically significant reductions in toxicity risk for most patients.
- Multi-timepoint MRI during radiotherapy shows architectural and functional changes in salivary glands starting soon after radiotherapy starts, and evolving as treatment progresses. However, data were insufficiently mature to adequately address whether these changes in imaging features predict for toxicity events.

9.3 Future directions

As indicated in section 9.1, the next logical step from work done in this thesis is a multivariate analysis of factors that predict clinically significant differences between planned and delivered dose to OARs. In the first instance, I would envisage defining clinical thresholds of patients who experience the biggest dose differences, and are therefore likely to benefit most from replanning. These thresholds might be based on absolute dose differences between planned and delivered dose, such as the 2Gy threshold used in Chapter 6, or combinations of dose differences to different OARs. The methodology could also be extended to use either published NTCP models, or the NTCP models generated in this work, to estimate the difference in toxicity risk that these dose differences confer. Thereafter, different approaches might be taken to multivariate analysis. A logistic regression approach similar to that used in Chapter 7 would be appropriate, but more unsupervised categorisation approaches could also be used.

The striking similarity between planned and delivered dose based NTCP models in Chapter 7 suggest that further work on delivered dose-based toxicity prediction is unlikely to be of much value, outwith clinical situations in which bigger differences between planned and delivered dose might be expected. One such scenario is Proton Beam Therapy (PBT). Because of the very steep dose gradients generated by proton Bragg peaks, and the importance of range uncertainty, PBT plans are much more sensitive to subtle differences in patient anatomy and setup than X-Ray plans [28]. Therefore, if similar contouring, immobilisation, and IG strategies are used for a PBT and X-Ray plan for the same patient, one might anticipate bigger dose deltas with the former. Therefore, as PBT is increasingly used to treat HNC [29], research into

delivered PBT dose based NTCP modelling may be an interesting and useful avenue to pursue.

Another interesting development in radiation oncology in recent years is the field of radiomics – asking whether the large quantities of information encoded within an image, and the patterns therein, can help to predict clinical outcomes. This approach has already been tried in HNC, with imaging modalities including pre-RT PET-CT, MRI, and planning CT [30-32]. There is less published data in which radiomics approaches have been applied to IG scans in an attempt to improve outcome prediction, although it is an area of interest, and has been tried with CBCT during RT for lung cancer [33]. One reason for this is that IG scans are inherently noisier and of lower quality than diagnostic, or treatment planning equivalents. Conversely however, IG scans may contain previously untapped imaging information on inherent radiosensistivity. It is known that there is a substantial spread in individual radiosensitivity within the population [34], but finding a simple, reliable and cheap test to measure this has proved difficult. It may be that radiomic analysis of IG scans, taken after a patient has started RT - and which they will be having anyway – can help to bridge this gap.

Finally, there is a clear and obvious need for greater automation in this field. If it were possible to automatically segment structures of interest on an IG scan, and re-compute daily dose, in an accurate, robust manner, in near real-time with the patient on-set, most research into ART would become obsolete. This would be 4D, rather 3D radiotherapy, and provided the workflows could be shown to be reliable, and without significant cost implications, it would be difficult to justify not using such an approach to treat patients. However, the technical challenges in reaching this destination remain daunting, and we are a long way from achieving this objective. However, work into fully automated machine-learning based auto-segmentation approaches is starting to appear [35, 36]. If these are shown to robust and reliable, and can be implemented across a range of hardware and software infrastructures, the light at the end of the tunnel will have come a little closer.

9.4 References:

[1] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol. 2011;12:127-36.
[2] Dahlstrand H, Nasman A, Romanitan M, Lindquist D, Ramqvist T, Dalianis T. Human papillomavirus accounts both for increased incidence and better prognosis in tonsillar cancer. Anticancer Res. 2008;28:1133-8.
[3] Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. J Clin Oncol. 2015;33:3235-42.
[4] Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med. 2010;363:24-35.
[5] Mehanna H, Robinson M, Hartley A, Kong A, Foran B, Fulton-Lieuw T, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. Lancet. 2019;393:51-60.
[6] Gillison ML, Trotti AM, Harris J, Eisbruch A, Harari PM, Adelstein DJ, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. Lancet. 2019;393:40-50.
[7] Owadally W, Hurt C, Timmins H, Parsons E, Townsend S, Patterson J, 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;15:602.
[8] Gregoire V, Jeraj R, Lee JA, O'Sullivan B. Radiotherapy for head and neck tumours in 2012 and beyond: conformal, tailored, and adaptive? Lancet Oncol. 2012;13:e292-300.
[9] Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tome WA. Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. Int J Radiat Oncol Biol Phys. 2010;76:S135-9.
[10] Brouwer CL, Steenbakkers RJ, Langendijk JA, Sijtsema NM. Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help? Radiother Oncol. 2015;115:285-94.
[11] Heukelom J, Fuller CD. Head and Neck Cancer Adaptive Radiation Therapy (ART): Conceptual Considerations for the Informed Clinician. Semin Radiat Oncol. 2019;29:258-73.
[12] Mehanna H, Robinson M, Hartley A, Kong A, Foran B, Fulton-Lieuw T, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. Lancet. 2019;393:51-60.
[13] Gillison ML, Trotti AM, Harris J, Eisbruch A, Harari PM, Adelstein DJ, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. Lancet. 2019;393:40-50.
[14] Beetz I, Schilstra C, van der Schaaf A, van den Heuvel ER, Doornaert P, van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. Radiother Oncol. 2012;105:101-6.
[15] Wopken K, Bijl HP, van der Schaaf A, van der Laan HP, Chouvalova O, Steenbakkers RJ, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol. 2014;113:95-101.
[16] Schultheiss TE, Tome WA, Orton CG. Point/counterpoint: it is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy. Med Phys. 2012;39:6531-3.
[17] Zhong H, Chetty IJ. Caution Must Be Exercised When Performing Deformable Dose Accumulation for Tumors Undergoing Mass Changes During Fractionated Radiation Therapy. Int J Radiat Oncol Biol Phys. 2017;97:182-3.
[18] Hugo GD, Dial C, Siebers JV. In Regard to Zhong and Chetty. Int J Radiat Oncol Biol Phys. 2017;99:1308-10.
[19] Han C, Chen YJ, Liu A, Schultheiss TE, Wong JY. Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy. Int J Radiat Oncol Biol Phys. 2008;70:1256-62.
[20] Duma MN, Kampfer S, Schuster T, Aswathanarayana N, Fromm LS, Molls M, et al. Do we need daily image-guided radiotherapy by megavoltage computed tomography in head and neck helical tomotherapy? The actual delivered dose to the spinal cord. Int J Radiat Oncol Biol Phys. 2012;84:283-8.

[21] Duma MN, Schuster T, Aswathanarayana N, Fromm LS, Molls M, Geinitz H, et al. Localization and quantification of the delivered dose to the spinal cord. Predicting actual delivered dose during daily MVCT image-guided tomotherapy. Strahlenther Onkol. 2013;189:1026-31.
[22] Noble DJ, Yeap PL, Seah SYK, Harrison K, Shelley LEA, Romanchikova M, et al. Anatomical change during radiotherapy for head and neck cancer, and its effect on delivered dose to the spinal cord. Radiother Oncol. 2019;130:32-8.
[23] Nishi T, Nishimura Y, Shibata T, Tamura M, Nishigaito N, Okumura M. Volume and dosimetric changes and initial clinical experience of a two-step adaptive intensity modulated radiation therapy (IMRT) scheme for head and neck cancer. Radiother Oncol. 2013;106:85-9.
[24] Cheng HC, Wu VW, Ngan RK, Tang KW, Chan CC, Wong KH, et al. A prospective study on volumetric and dosimetric changes during intensity-modulated radiotherapy for nasopharyngeal carcinoma patients. Radiother Oncol. 2012;104:317-23.
[25] van Kranen S, Hamming-Vrieze O, Wolf A, Damen E, van Herk M, Sonke JJ. Head and Neck Margin Reduction With Adaptive Radiation Therapy: Robustness of Treatment Plans Against Anatomy Changes. Int J Radiat Oncol Biol Phys. 2016;96:653-60.
[26] Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys. 2010;76:S10-9.
[27] Christianen MEMC, Schilstra C, Beetz I, Muijs CT, Chouvalova O, Burlage FR, et al. Predictive modelling for swallowing dysfunction after primary (chemo)radiation: Results of a prospective observational study. Radiotherapy and Oncology. 2012;105:107-14.
[28] Müller BS, Duma MN, Kampfer S, Nill S, Oelfke U, Geinitz H, et al. Impact of interfractional changes in head and neck cancer patients on the delivered dose in intensity modulated radiotherapy with protons and photons. Phys Med. 2015;31:266-72.
[29] Leeman JE, Romesser PB, Zhou Y, McBride S, Riaz N, Sherman E, et al. Proton therapy for head and neck cancer: expanding the therapeutic window. Lancet Oncol. 2017;18:e254-e265
[30] van Dijk LV, Brouwer CL, van der Schaaf A, Burgerhof JG, Beukinga RJ, Langendijk JA, et al. CT image biomarkers to improve patient-specific prediction of radiation-induced xerostomia and sticky saliva. Radiother Oncol. 2017;122:185-91.
[31] van Dijk LV, Thor M, Steenbakkers RJHM, Apte A, Zhai TT, Borra R, et al. Parotid gland fat related Magnetic Resonance image biomarkers improve prediction of late radiation-induced xerostomia. Radiother Oncol. 2018;128:459-466.
[32] van Dijk LV, Noordzij W, Brouwer CL, Boellaard R, Burgerhof JGM, Langendijk JA, et al. 18F-FDG PET image biomarkers improve prediction of late radiation-induced xerostomia. Radiother Oncol. 2018;126:89-95.
[33] van Timmeren JE, van Elmpt W, Leijenaar RTH, Reymen B, Monshouwer R, Bussink J, et al. Longitudinal radiomics of cone-beam CT images from non-small cell lung cancer patients: Evaluation of the added prognostic value for overall survival and locoregional recurrence. Radiother Oncol. 2019;136:78-85.
[34] Barnett GC, West CM, Dunning AM, Elliott RM, Coles CE, Pharoah PD, et al. Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nat Rev Cancer. 2009;9:134-42.
[35] Kosmin M, Ledsam J, Romera-Paredes B, Mendes R, Moinuddin S, de Souza D, et al. Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer. Radiother Oncol. 2019 Jun;135:130-140.
[36] Macomber MW, Phillips M, Tarapov I, Jena R, Nori A, Carter D, et al. Autosegmentation of prostate anatomy for radiation treatment planning using deep decision forests of radiomic features. Phys Med Biol. 2018;63:235002

Appendices

A1 Publications relating to thesis

A1.1 First author papers
A1.2 Other papers
A1.3 First author published abstracts

A1. Publications related to thesis

A1.1 First author papers

Squamous cell carcinoma of the oral cavity, oropharynx and upper oesophagus.
Noble DJ, Jefferies SJ. *Medicine (United Kingdom) 47(5): 269-274.*

Anatomical change during radiotherapy for head and neck cancer, and its effect on delivered dose to the spinal cord. Noble DJ, Yeap PL, Seah SYK, Harrison K, Shelley LEA, Romanchikova M, Bates AM, Zheng Y, Barnett GC, Benson RJ, Jefferies SJ, Tomas SJ, Jena R, Burnet NG. *Radiother Oncol. 2019 Jan;130:32-38.*

The future of image-guided radiotherapy – is image everything?
Noble DJ, Burnet NG. *Br J Radiol. 2018 Jul;91(1087):20170894.*

A1.2 Other papers

Target volume concepts in radiotherapy and their implications for imaging
Burnet NG, Noble DJ, Paul A, Whitfield GA, Delorme S. *Radiologe. 2018 Aug;58(8):708-721.*

Automatic contour propagation using deformable image registration to determine delivered dose to spinal cord in head-and-neck cancer radiotherapy.
Yeap PL, Noble DJ, Harrison K, Bates AM, Burnet NG, Jena R, Romanchikova M, Sutcliffe MPF, Thomas SJ, Barnett GC, Benson RJ, Jefferies SJ, Parker MA. *Phys Med Biol. 2017 Jul 12;62(15):6062-6073.*

Applying physical science techniques and CERN technology to an unsolved problem in radiation treatment for cancer: the multidisciplinary VoxTox research programme.
Burnet NG, Scaife JE, Romanchikova M, Thomas SJ, Bates, AM, Wong E, Noble DJ, Shelley LEA, Bond SJ, Forman JR, Hoole ACF, Barnett GC, Brochu FM, Simmons MPD, Jena R, Harrison K, Yeap PL, Drew A, Silvester E, Elwood P, Pullen H, Sultana A, Seah SYK, Wilson MZ, Russell SG, Benson RJ, Rimmer YL, Jefferies SJ, Taku N, Gurnell M, Powlson AS, Schönlieb C-B, Cai X, Sutcliffe MPF, Parker MA. *CERN IdeaSquare Journal of Expermiental Innovation, 2017; 1(1): 3*

A1.3 First author published abstracts

Assessing dynamic change in salivary gland function using MRI during chemoradiation for head and neck cancer (HNC). Noble DJ, Zaccagnia F,, Bates AM, Das T, Lupson V, Welsh K, Barnett GC, Rashmi Jadon R, Benson RJ, Burnet NG, Jena R, Gallagher FA.

- *Oral presentation:* UKIO, Liverpool, June 2019.
- *Published in:* BIR Publications: Short Paper Presentations | Proceedings of UK Radiological Conference 2019. https://doi.org/10.1259/conf-pukrc.2019.A-short-paper-presentations

Univariate toxicity associations are stronger with delivered than planned dose in HNC patients.
Noble DJ, Harrison K, Hoole A, Wilson M, Thomas SJ, Bates AM, Shelley LEA, Burnet NG, Jena R.

- *Poster presentation:* Estro 39, Milan, April 2019.
- *Published in:* Radiotherapy and Oncology 133:S539-S540

Predictors of dose differences to swallowing OARs in patients undergoing radiotherapy for HNC.
Noble DJ, Harrison K, Wilson M, Hoole A, Thomas SJ, Burnet NG, Jena R.

- *Poster presentation:* International Conference of Head & Neck Oncology, Barcelona, March 2019.
- *Published in:* Radiotherapy and Oncology 132:65-66.

CTCAE-based comprehensive toxicity assessment following radical radiotherapy for head & neck squamous cell carcinoma (HNSCC). Noble DJ, Bates AM, Scaife JE, Gemmill J, Benson RJ, Barnett GC, Jefferies SJ, Burnet NG, Jena R.

- *Poster presentation:* NCRI, Glasgow, November 2018. **(nominated for RCR Ross award)**
- *Published in:* British Journal of Cancer 119, 1-49(2018). 15-16.

A2 Documents relating to standard Radiotherapy treatment & workflows at Addenbrooke's Hospital

A2.1 Radiotherapy prescription – head and neck
A2.2 Dose difference assessment (DDA document)
A2.3 Radiographer work instructions – TomoTherapy

Radiotherapy Prescription – Head & Neck

CP/Head&Neck Prescription

Prescription – Head & Neck

Fields in RED are mandatory

Stage T: *
N: *
M: M0* No Distant Metastasis
Category CAT1* Treated Radically with curative intent with no breaks in treatment
Primary ☒ Follow up OPA Hospital: Addenbrookes
Concession ☐ Sub C ☐ If Other please specify:
Concession No. Sub E ☐ Follow up 6 week post treatment

Hospital no: {Ident.IDA@U}
Surname: {Patient.Last_Name@U}
First names: {Patient.First_Name@U}
Date of birth: {Admin.Birth_Date@d6b}
NHS No: {Ident.IDB@U}

Phase	1 – single phase			
Target area	Head and Neck			
Fields description	As Planned			
Field size (cm)	As planned			
Energy	6MV			
Total Dose (Gy) State maximum dose prescription				
No. of fractions				
Dose per fraction (Gy)				
Where Specified	Median Dose			
Fractions per week	5			
Overall treatment time				
Practitioner Initials				

Chemotherapy:
no chemotherapy

Minimum interval to RT: if chemo - 2 hrs

Bloods (please tick)			
Mon	☐	Tue	☐
Wed	☐	Thur	☐
Fri	☐	None	☐
Once Only	☐		

Additional Comments:

Allow non clinician authorisation of treatment plan provided clinical Protocol Followed ☐

CT Volume Definition Head and Neck Standard

		☐	☐	☐	☐	☐	
Dose (Gy)	PTV3	70	65	-	-	-	
	PTV2	60	60	60	-	-	
	PTV1	56	54	54	55	50	
Fractions		35	30	30	20	20	

Volumes defined in Prosoma
Fields marked * are non standard

SP		ProSoma Comment	

Additional Comments:

Hospital no:	{Ident.IDA@U}
Surname:	{Patient.Last_Name@U}
First names:	{Patient.First_Name@U}
Date of birth:	{Admin.Birth_Date@d6b}
NHS No:	{Ident.IDB@U}

Margins to be used (cm)	All	*AP*	*Lateral*	*Sup-Inf*
CTV1 –PTV1	☐ 0.5cm	cm	cm	cm
CTV2 –PTV2	☐ 0.5cm	cm	cm	cm
CTV3 –PTV3	☐ 0.5cm	cm	cm	cm

OARs not outlined will not be considered

Use?	**Organ**	**Objective dose (i.e. soft limit, goal)**		**Absolute dose constraint**	
☐	PTV1 / PTV2 / PTV3	97% - 105% to cover		95% - 107% to cover	
☐	Spinal Cord PRV	46Gy		50Gy	
☐	Brainstem PRV	48Gy		52Gy	
☐	Hypothalamus	44Gy		-	
☐	Pituitary	45Gy		-	
☐	Optic chiasm PRV	50Gy		55Gy	
☐	Larynx	44Gy		-	
☐	Pharyngeal Constrictors	50Gy (Mean Dose)		-	
☐	Mandible	No hot spots (102%)		-	
		Right	**Left**	**Right**	**Left**
☐	Parotid	20Gy (Mean Dc	20Gy (Mean Dc	-	-
☐	Submandibular Gland	35Gy (Mean Dc	35Gy (Mean Dc	-	-
☐	Lens	6Gy	6Gy	-	-
☐	Cornea	30Gy	30Gy	-	-
☐	Retina	50Gy	50Gy	-	-
☐	Globe	40Gy	40Gy	45Gy	45Gy
☐	Lacrimal Gland	26Gy	26Gy	-	-
☐	Optic Nerve PRV	50Gy	50Gy	55Gy	55Gy
☐	Cochlea	35Gy	35Gy	45Gy	45Gy

☐ Tick here if plan has become more complex than originally indicated on Action Sheet.
Name of planner informed:

Tomotherapy delivered dose assessment form

LA4.004

DDA REQUEST

Plan Name:

Machine:

Date of MVCT scan

Treatment site:

Fraction number

Hospital no:	{Ident.IDA@U}
Surname:	{Patient.Last_Name@U}
First names:	{Patient.First_Name@U}
DOB:	{Admin.Birth_Date@d6b}
NHS No:	{Ident.IDB@U}

Reason for assessment		Tick all applicable	Comments
Routine assessment as per protocol		☐	
Check positional reproducibility		☐	
Physics Request		☐	
External shape change seen	+	☐	
	-	☐	
Internal structure changes seen	Target	☐	
	OAR	☐	
Repeat - trends previously seen	Shape change	☐	
	Dose differences	☐	
Other		☐	

Total shims added at this date? If yes, state amount	Head	Shoulder

Quality Checklist request to Physics-Planning opened? TP:DDA ☐

Requested by:		**Date**	

ASSESSMENT PREPARATION

Data prep. & batched by: *(level 2 trained staff as per TP2.04)*		**Date**	

Tomotherapy delivered dose assessment form

LA4.004

EVALUATION ANALYSIS RESULTS

Are there any new hot spots in any PTVs above tolerance? :
Are there any new cold spots in any PTVs below 95%? :
Are there any new hot spots in any OARs above tolerance? :

Comments (if required)

Is there a screenshot added to illustrate the above : Add Screenshot

Action	Tick applicable	Comments
None (within tolerances)	☐	
Must review before next fraction (must communicate to set)	☐	
Discuss at next review meeting (state date)	☐	
Review other (state date by which you expect conclusion)	☐	

Signed: *(level 2 trained staff as per TP2.04)*		**Date**	

Oncologist Decisions

Comments if required

Signed

Closed by:		**Date**	

Work Instruction

LA3.110
April 2017

Treating with Tomotherapy

1 Scope

This document provides work instructions for Radiographer and Assistant Practitioners in the use of Tomotherapy.

2 Purpose

To ensure safe and accurate radiotherapy treatment delivery

3 Undertaken By

Radiographers who have achieved the relevant competency levels.

3.1 To download a treatment

- Open MOSAIQ by clicking on the Mosaiq icon located on the desktop

- Change the **Queue Location** when calling the patient into the treatment room by following SG3.10;
- From the **Home Screen** > select **Quick Schedule** > **Location** > **Patient** > **Open Chart**
- The **Chart** screen will open with the patient selected;
- Click on **D and I** icon > **Site Set Up** for immobilisation information.
- Select **RO Treat** icon > on the **Treatment Chart Page, c**heck the correct Gantry number is displayed for todays treatment. If this needs to be changed select **Tx calander** > **Change** > **Define** then reschedule fields as appropriate > click **ok** > **ok** then return to **Tx chart** > **Treat**.
- In Treatment Delivery Table select Send Plan.
- Reselect D and I icon > Site Set Up for immobilisation information.

Where Mosaiq does not communicate with tomotherapy and there are multiple plans or phases please ensure the correct plan is selected.
Any re-plans and subsequent phases **will all be active.** Take extra care when selecting plans by ensuring that the plan names are the same and correct for that plan/ phase.

3.2 Viewing the Trt Plan

- Click on **DOCUMENTS** icon > pull window to extra screen and maximise to fit.
- Click on **Plan Document** which will open plan in the Document Viewer window.
- Scroll through plan to view DVH legends and/or dose wash pictures to inform slice selection.

3.3 To pull treatment detils onto Tomotherapy Opertaing System.

- Select OIS icon
- Select Patient /Plan Box opens > check patient details > OK if correct.
- Select slice thickness > C=6, N=4, F=2.
- Select slices to identify PTV
- Click on Accept Slices > Prepare Scan> in prepare patient/room box click Ok if correct

3.4 On Mosaiq

- On the Chart screen select **On Trt Info** tab.
- Check **Patient Quality Check List** to see if any actions are required.
- Read annotations in **Pre-Trt/On Trt** Tabs in **Quick Notes**.

4 In the treatment room

- Set up immobilisation as per Site Setup – the sliding bar on the right of this box can be used to view additional information if required.
- On **Day 1** record location points for immobilisation devices in Site Set Up
- Identify the patient as per OHR/I001 > check this with the information in the Manual Patient Verification screen and personal equipment if being used.

5 Patient Positioning

- Instruct patient to avoid heavy breathing during the scan
- Ensure the couch lateral is at zero
- Position patient as per CT planning scan position
- Using the Positioning Control Panel (PCP) touch screen: select **Main** > **Setup**>**Yes**
- Bed moves to approximate set up position
- Align tattoos with the red lasers
  - Set height first
  - For large movements use manual arrow controls with the dead man handle button
  - For fine movements use touch screen: **Step Move**> set mm> **Move Y** or **Move Z**
- Ref **Position** > Reset **Origin** (resets laser value to zero)
- Record **Y** and **Z couch values** (planned set up position)
- Touch screen: select **Main** > **Ready** > **Yes** (couch moves into bore ready for scan)

If Treatment plan includes bolus, ensure bolus is positioned correctly prior to scan

6 Perform Scan

- Wait for prompt in What's Next text box
- On Status consol turn key to left (image) to Scan
- When green light illuminated press **Start**
- When scan complete wait for prompt in What's Next text box & turn key to program

7 TomoImage registration

- Wait for prompt in 'What's Next' text box
- Select Registration tab
- In [Automatic Registration Control] select **Bone & soft tissue** technique
- Use **Fine resolution**
- Select **Incomplete Field of View** if patient contour outside FOV
- Select **Translations + Pitch + Yaw + Roll**
- Select **Start**
- Confirm pitch, yaw and roll are within tolerances.
- Select translational and /or rotational corrections to be applied automatically (translations or translations and roll)
- Select **Start** > select **Yes** to re-set values to zero
- Using **Orientation** options check registration on slices throughout volume in transverse, sagital and coronal planes
- To change colour right click the scan box **[in Scan Image Control box]**
- Use **Balance** and / or **Checker** slider to optimise registration

8 Treatment delivery

8.1 Post registration outside room

- Select **Treat** tab
- Use mouse to highlight the next available treatment fraction (date does not have to match)
- Click on **Prepare Treatment**
- If dates conflict, warning message will appear, click **Yes** to continue
- Confirm patient position click **OK**
- Click on **Apply** (De-select any roll corrections <0.5°)
- Pop-up box appears indicating that system is connecting to the OCSS
- Couch control box opens (Fig 1)

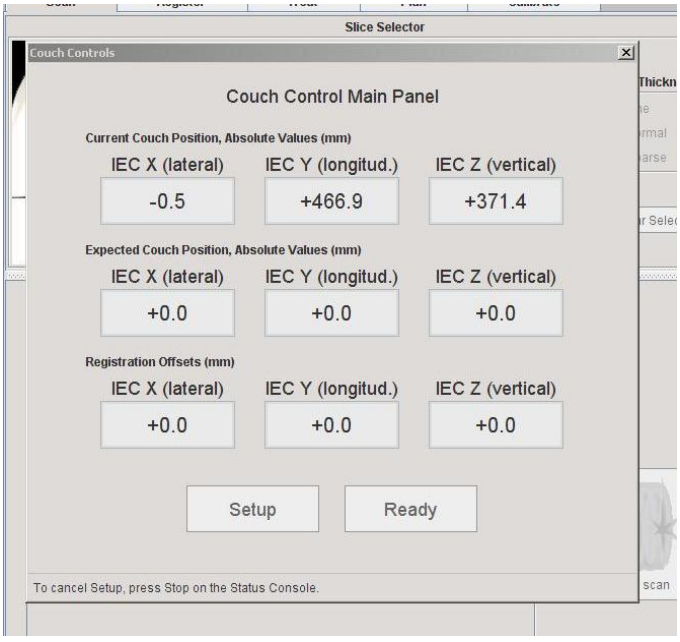


Figure 1: Couch controls box

- When set-up tab is blue, select: this will apply any moves to the couch displayed in the Registration Offsets display
- The user confirmation box will then open (Fig 2)

- Confirm expected moves with current displayed, if correct select "Yes"
- On the couch control panel select "Ready" when blue and couch will move into ready to treat position. Record couch longitudinal

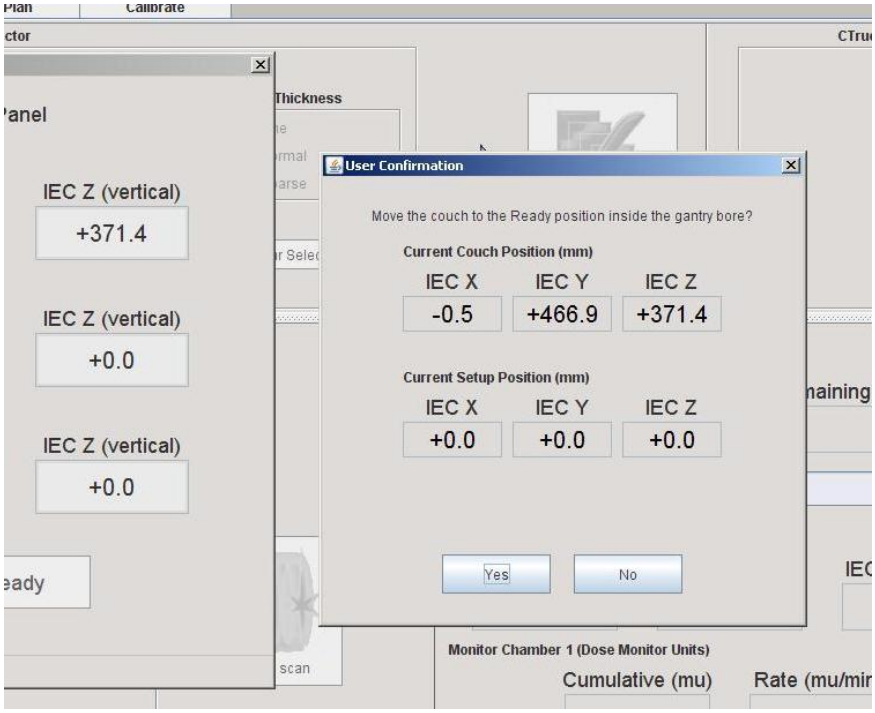


Figure 2: User Confirmation box

- Wait for prompt in **What's Next** text box
- Status console: turn key to **Treat** (right), press **start** when green light illuminated
- if the system requires more than 300secs to prepare a procedure, a dialog box is displayed on the Treat tab
- On dialog box, click, **Wait,** if box shows a second time linac will shutdown
- **Treatment Status display** displays time elapsed and remaining in seconds as well as current Y values in mm. the current Z value should not change during treatment
- Ensure couch Y moving during treatment
- Monitor dose rate during treatment
- Open the on treatment assessment as per SG3.05 and fill out as per LA3.135.
- Wait for prompt in **What's Next** text box
- Status control: turn key to **Program**

8.2 Inside the room

The in-room controls in place of the OCSS can be used at radiographer discretion or for an unrecoverable interrupt

- Touch screen: press **Main** > **Set up** > **Yes**
- Red lasers and couch will be in treatment position. Lasers should be running through tattoos
- Touch screen: press **Ref. Position** to verify correct adjustments have been performed
- Laser Y and z values should agree with recorded shifts
- Touch screen: select **Reset Origin**

- Record Y and Z couch positions (registered ready position)
- Touch screen: press **Main** > **Ready**> **Yes**

9 Interrupts during Scanning & Treatment Delivery

9.1 Interruption during a scan

- Will always require a rescan
- Wait for prompt in What's Next text box
- Status control: turn key to Program
- When system interruption message appears, click on **Details**
- Annotate fault log

9.2 Interruption before treatment initiated

- Touch screen: **Main** > **Set up** > **Yes**. Ensure couch position is the same as the Set Up Y and Z values. If not manual adjustment will be required
- Touch screen: Press **Main**> **Ready** > **Yes**

9.3 Recoverable Interruption during treatment

- Record the Couch Y position and the remaining time in seconds
- Wait for prompt in What's Next text box
- Status control: turn key to **Program**
- When system interruption message appears, click on **Details**
- Annotate in fault log
- Enter password
- Right click on the interrupted treatment fraction and **Generate a Completion Procedure**
- Type password, then click OK
- A dialog box displays the duration of the completion procedure. Confirm that the Duration Time Left is not more than 15 seconds longer than the remaining time recorded at the time of the interruption
- Select OK
- A **Scheduled Completion** plan will appear and be automatically highlighted
- Click on the **Prepare Treatment** icon
- Read and acknowledge laser position, roll correction and rescan warning messages.
- Confirm patient position, press **OK**
- In Room, **Main** > **Set up** > **Yes**. Ensure couch position is the same as the Set Up Y and Z values. If not manual adjustment will be required.
- Red lasers should coincide with tattoos unless patient has moved
- If patient has moved cancel procedure, rescan, register then deliver scheduled completion.
- Touch screen: press **Ref Position.** Laser Y and Z values should be near to zero.
- Touch screen: **Reset Origin**
- Record Y and Z positions Touch screen: Press **Main**> **Ready** > **Yes**
- Couch will move into bore to start treatment. Record Y position and compare to interrupted Y position.
- Status console: turn key to **Treat** (right), when green light illuminated, press **Start.**

9.4 Unrecoverable

- Shutdown as per TS3.71 and restart the system
- Select patient, disease and plan number
- Red lasers will NOT be through tattoos. They should be opposite of the shifts if the patient has not moved.
- Follow instructions as per 12.3 above

9.5 Removing a patient from the couch in an emergency

- If a scan or treatment has to be abandoned in an emergency patient can be removed from couch using manual controls.
- The machine must be turned off before the emergency release button on the couch will work.
- The Yellow couch horizontal release button located at foot of couch will allow the couch to be retracted fully.
- To manually lower, retract couch top fully, and lower couch by pressing the button on the left hand side under the feet end of the couch top.

10 Completion of treatment

Once a treatment has been completed the relevant status should be applied to Mosaiq following SG3.11

11 Monitoring compliance with the effectiveness of this document

a) Process for monitoring compliance and effectiveness: Review of Safety Learning Events, as recorded on the Trust QSIS reporting system, for non-compliance and the results presented to the Patient Safety & Governance group - the minutes of these meeting are retained for a minimum of 5 years.

b) Standards/Key Performance Indicators: This process forms part of a quality system accredited to International Standard BS EN ISO 9001:2015. The effectiveness of the process will be monitored in accordance with the methods given in the quality manual (QM1.00).

12 Associated documents

TS3.71 Tomotherapy (LA3 & LA4) start-up and shutdown

Equality and Diversity Statement

This document complies with the Cambridge University Hospitals NHS Foundation Trust service Equality and Diversity statement.

Disclaimer

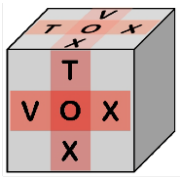
It is **your** responsibility to check against the electronic library that this printed out copy is the most recent issue of the document.

Document management. *Notify any changes required to the Cancer QA Team*

Approval:	Head of Radiotherapy, Specialist Radiographer
Date approved:	20/04/2017
Date on e-library:	20/04/2017
Owning Department:	Cancer Division QA Team - Radiotherapy
Author(s):	Specialist Radiographer
File name:	p_la3_110_treating_with_tomo
Keywords:	Tomotherapy treat
Review period:	Managed by QPulse/Cancer QA Team
Version number:	5
Local reference:	LA3.110
Media ID:	

To reference the **latest version** of this document use the **link** on *Local reference/MediaID* above

A3 Toxicity data – collection & processing



VoxTox – Linking Radiation dose at the voxel level with toxicity

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

Obs View Label:	VOX:H&N baseline Assessment V2
Obs View Desciption:	VOX:H&N baseline Assessment
Active Date:	21/11/2013
Code description	Vox:H&N base Ass (VHA2)

Please ensure that all information in red is completed as this is mandatory data

VHB2:TREATMENT INFO

VHB2: HEIGHT (CM) _____

VHB2: WEIGHT (KG) _____

VHB2:SMOKING

VHB2:SMOKER	0 - CURRENT SMOKER	
	1 - EX-SMOKER	
	2 - NEVER SMOKED	

VHB2:SMOKING AGE STARTED _____

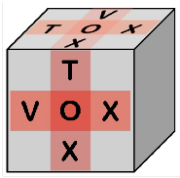
VHB2:SMOKING AGE STOPPED _____

VHB2:AVERAGE NO PER DAY _____

VHB2:TYPE OF SMOKING	0 - CIGARETTES	
	1 - ROLL UPS	
	2 - CIGARS	
	3 - PIPE	
	4 - OTHER	

VHB2:ALCOHOL

VHB2:ALCOHOL	0 - CURRENT DRINKER	
	1 -EX-DRINKER	



VoxTox – Linking Radiation dose at the voxel level with toxicity

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

	2 - NEVER DRUNK	

VHB2:ALCOHOL AGE STARTED _____

VHB2:ALCOHOL AGE STOPPED _____

VHB2:AVERAGE UNITS PER WEEK _____

VHB2:PRE-EXISTING MORBIDITIES

VHB2:DIABETES MELLITUS	0. NO	
	1. YES	
PLEASE CHECK ALL THAT APPLY	VHB2: (DM)DIET CONTROL	
	VHB2: (DM)METFORMIN	
	VHB2: (DM)INSULIN	
	VHB2: (DM) OTHER	
VHB2:CARDIOVASCULAR DISEASE	0. NO	
	1. YES	
VHB2:HYPERTENSION	0. NO	
	1. YES	
VHB2: IS PT ON STATINS	0. NO	
	1. YES	
VHB2: IS PT ON ACE INHIBITORS	0. NO	
	1. YES	
VHB2:CONNECTIVE TISSUE DISEASE	0. NO	
	1. YES	
VHB2:DERMATITIS	0. NO	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

	1. YES	
VHB2: RETINOPATHY	0. NO	
	1. YES	
VHB2: RETINAL DETACHMENT	0. NO	
	1. YES	
VHB2: CATARACT	0. NO	
	1. YES	

VHB2:OTHER DIAGNOSES _____

VHB2:DENTAL

VHB2:ANY TEETH IN SITU	0 - NO	
	1- YES	
VHB2: EXTRACTION PAST 3 MONTHS	0 NO	
	1. YES	

VHB2:PEG/RIG

VHB2:PEG/RIG IN SITU	0- NO	
	1 - YES	

VHB2:OPERATIVE INTERVENTION

INSTRUCTION: **OTHER THAN A TUMOUR BIOPSY, HAS THIS PATIENT HAD SURGERY? PLEASE CHECK BOX FOR ALL RELEVANT PROCEDURES THAT APPLY.**

VHB2:SURGERY TO PRIMARY TUMOUR	0 - NO	
	1 -YES	

VHB2:DATE OF SURGERY _____

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:SURGERY INVOVED BONE		
VHB2: LARYNGECTOMY		
VHB2:NECK DISSECTION		
VHB2:DATE OF DISSECTION		

VHB2:NEOADJUVANT CHEMO

VHB2:NEOADJUVANT CHEMOTHERAPY	0- NO	
	1- YES	

VHB2: CHEMOTHERAPY TYPE	0- CISPLATIN & 5FU	
	1 - TPF	
	2- OTHER	
VHB2: OTHER CHEMOTHERAPY		

VHB2:TUMOUR SITE

VHB2:TUMOUR SITE	0 - SUPRAGLOTTIC OR SUBGLOTTIC LARYNX	
	1 - GLOTTIC LARYNX	
	2 - NASAL CAVITY	
	3 - PARANASAL SINUSES	
	4 - OROPHARYNX	
	5 - HYPOPHARYNX	
	6 - LIP	
	7 - ORAL CAVITY	
	8 - NASOPHARYNX	
	9 - MAJOR SALIVARY GLAND TUMOUR	
	10 - NOT KNOWN AT THIS TIME	
	11 - OTHER	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:IF OTHER PLEASE STATE _____

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:STAGING

INSTRUCTION: **PLEASE REFER TO TNM 7.0 2010**

VHB2:STAGING T: _____

VHB2:STAGING N: _____

VHB2:TUMOUR HISTOLOGY

INSTRUCTION: **PLEASE SELECT HISTOLOGY. IF NOT SHOWN THEN PLEASE COMPLETE OTHER HISTOLOGY BOX.**

VHB2:TUMOUR HISTOLOGY		
	0 - SQUAMOUS CELL CARCINOMA	
	1 - UNDIFFERENTIATED CARCINOMA	
	2 - ADENOCARCINOMA	
	3 - ADENOSQUAMOUS CARCINOMA	
	4 - ADENOMA	
	5 - MUCOEPIDERMOID CARCINOMA	
	6 - ADENOID CYSTIC CARCINOMA	
	7 - NOT KNOWN AT THIS TIME	
	8 - OTHER	

VHB2:OTHER HISTOLOGY _____

_____

VHB2:TUMOUR GRADE

VHB2:TUMOUR GRADE		
	0 - WELL	
	1 - MODERATELY	
	2 - POORLY DIFFERENTIATED	
	3 - NOT KNOWN AT THIS TIME	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:DIET CHANGES	0- NO	
	1- YES	
VHB2:DIET (SELECT ALL THAT APPLY)	OCC. SALIVA SUBSTITUTE	
	FREQ SALIVA SUBSTITUTE	
	NEEDS SUBS/WATER IN ORDER TO EAT	
	NON-ACIDIC FOODS ONLY	
	SMALL BITES	
	SEMI-SOFT FOOD	
	SOFT FOOD	
	EATS SOFT FOOD BUT DIFFICULT	
	LIQUIDS ONLY	
	TEMPORARY NG TUBE	
	IV FLUIDS	
	TPN IE NUTRITION INTO VEIN	
	PERMANENT FEEDING TUBE	
VHB2:REASON FOR ALTERED DIET (SELECT ALL THAT APPLY)	DRY MOUTH	
	SALIVA	
	TASTE CHANGES	
	SWALLOWING DIFFICULTY	
	WEIGHT LOSS	
	TRISMUS	
	CHEWING DIFFICULTY	
OTHER REASON		
VHB2: DRUG FOR HEAD & NECK REGION	0-NO	
	1- YES	
VHB2: DRUGS FOR HEAD AND NECK (SELECT ALL THAT APPLY)	OCCASIONAL ANALGESIA	
	REGULAR ANALGESIA	
	ANTACIDS	
	NUTRITIONAL SUPPLEMENTS	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

IRON	
STEROIDS	
FLUORIDES FOR TEETH	
MUSCLE RELAXANTS	
DIURETIC/WATER MEDICATION	
OCC CREAMS AND MEDS FOR SKIN	
DAILY CREAMS AND MEDS FOR SKIN	
OCCASIONAL EARDROPS/OINTMENTS	
REGULAR EARDROPS/OINTMENTS	

VHB2:SKIN

INSTRUCTION: **PLEASE RECORD WORST SYMPTOMS DURING PAST MONTH**

VHB2: SKIN DERMATITIS	0 – NONE	
	1 – FAINT ERYTHEMA OR DRY DESQUAMATION	
	2 – MODERATE/TENDER/BRIGHT ERYTHEMA/ PATCHY MOIST DESQUAMATION	
	3 – WIDESPREAD MOIST DESQUAMATION/ BLEEDING INDUCED BY MINOR TRAUMA	
	4 – SKIN NECROSIS/ ULCERATION OF FULL THICKNESS DERMIS/ HAEMORRHAGE	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2: SCALINESS AND ROUGHNESS	0 - NONE	
	1 - PRESENT BUT NOT CAUSING SYMPTOMS	
	2 - CAUSING SYMPTOMS	
	3 - REQUIRING CONSTANT ATTENTION	
VHB2: SKIN SENSATION	0 - NORMAL	
	1 - SENSITIVE OR ITCHY	
	2 - PAINFUL AT TIMES	
	3 - PAINFUL MOST OF THE TIME	
VHB2:PRURITUS	0 – NONE	
	1 - MILD AND LOCALIZED	
	2 -INTENSE/ WIDESPREAD/ INTERMITTENT/SCRATCHING CHANGES/ORAL MEDS	
	3. CONSTANT/ LIMITING SLEEP/ ORAL STEROIDS/ IMMUNOSUPRESSANTS	
VHB2: SKIN PIGMENTATION	0 – NONE	
	1 -SLIGHT AREA LIGHTER/DARKER SKIN CAUSING NO PSYCHOSOCIAL IMPACT	
	2 - SLIGHT AREA LIGHTER/DARKER SKIN CAUSING PSYCHOSOCIAL IMPACT	
	3 - MARKED AREA LIGHTER/DARKER SKIN	
VHB2: TELANGIECTASIA	0 - NONE	
	1 - MILD	
	2 - MODERATE	
	3 - SEVERE	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2: ATROPHY (THINNING)	0 - NONE	
	1 - PRESENT BUT NOT CAUSING SYMPTOMS	
	2 - CAUSING SYMPTOMS	
	3 - CAUSING SOME LIMITATION OF FUNCTION	
	4 - CAUSING SEVERE LOSS OF FUNCTION	
	5 - NON-HEALING ULCER	

VHB2: FIBROSIS (THICK HARD TISSUE)	0 - NONE	
	1 - PRESENT BUT NOT CAUSING SYMPTOMS	
	2 - CAUSING SYMPTOMS	
	3 - CAUSING SOME LIMITATION OF FUNCTION	
	4 - CAUSING SEVERE LOSS OF FUNCTION	

VHB2: SKIN SWELLING	0 - NONE	
	1 - PRESENT BUT NOT CAUSING SYMPTOMS	
	2 - CAUSING SYMPTOMS	
	3 - CAUSING SOME LIMITATION OF FUNCTION	
	4 - CAUSING SEVERE LOSS OF FUNCTION	

VHB2: HAIR LOSS	0 - NONE	
	1 - THINNING BUT NO PSYCHOSOCIAL IMPACT	
	2 - THINNING AND PSYCHOSOCIAL IMPACT	
	3 - LOST IN PATCHES	
	4 - COMPLETE LOSS	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:MOUTH AND JAW

INSTRUCTION: PLEASE RECORD WORST SYMPTOMS DURING PAST MONTH

VHB2:DRY MOUTH

0 - NO SYMPTOMS	
1 - MILD SYMPTOMS WITHOUT SIGNIFICANT DIETARY ALTERATION	
2 - MODERATE SYMPTOMS/ ORAL INTAKE ALTERATIONS	
3 - UNABLE TO ADEQUATELY NOURISH ORALLY	

VHB2:SALIVA AND TASTE

0 - NO SYMPTOMS	
1 - SLIGHTLY THICKENED SALIVA OR SLIGHTLY ALTERED TASTE E.G. METALLIC	
2 - THICK, STICKY SALIVA/ CHANGED TASTE/ DIET/ SECRETION-INDUCED SYMPTOMS	
3 - SEVERE SECRETION-INDUCED SYMPTOMS/ TUBE FEEDING/ TPN	
4 - LIFE-THREATENING CONSEQUENCES OR URGENT INTERVENTION INDICATED	
5- ACUTE SALIVARY GLAND NECROSIS	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:ORAL MUCOSA	0 - NO INFLAMMATION OF ORAL MUCOSA	
	1 -ASYMPTOMATIC/ MILD SYMPTOMS AND INTERVENTION NOT INDICATED	
	2 -MODERATE PAIN / MODIFIED DIET BUT NOT INTERFERING WITH ORAL INTAKE	
	3 - SEVERE PAIN/ INTERFERING WITH ORAL INTAKE	
	4 - LIFE-THREATENING CONSEQUENCES / URGENT INTERVENTION INDICATED	
	5 - ULCERATION/ HAEMORRHAGE/ NECROSIS	

VHB2: DENTURES	0 - NOT NEEDED	
	1 - FIT WELL	
	2 - LOOSE FIT	
	3 - NOT ABLE TO USE	

VHB2: PAIN IN JAW	0 - NONE	
	1 - OCCASIONAL AND MILD	
	2 - SOMETIMES BUT BEARABLE	
	3 - OFTEN PRESENT OR SEVERE	
	4 - CONSTANT AND VERY SEVERE	

VHB2: PAIN IN TEETH	0 - NONE	
	1 - OCCASIONAL AND MILD	
	2 - SOMETIMES BUT BEARABLE	
	3 - OFTEN PRESENT OR SEVERE	
	4 - CONSTANT AND VERY SEVERE	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:TRISMUS	0 – NONE	
	1 - DECREASED RANGE OF MOVEMENT (ROM) BUT NO IMPAIRMENT IN EATING	
	2 - DECREASED ROM REQUIRING SMALL BITES, SOFT FOODS OR PUREES	
	3 - DECREASED ROM UNABLE TO ADEQUATELY NOURISH/HYDRATE ORALLY	
VHB2: OSTEORADIONECROSIS	0 – NONE	
	1 - PRESENT BUT NOT CAUSING SYMPTOMS	
	2 - PRESENT AND CAUSING SYMPTOMS	
	3 - SEVERE SYMPTOMS OR LIMITING SELF-CARE	
	4 - LIFE-THREATENING CONSEQUENCES/URGENT INTERVENTION	
VHB2: PHARYNX AND OESPHAGUS		
INSTRUCTION:	PLEASE RECORD WORST SYMPTOMS DURING PAST MONTH	
VHB2: PAIN ON SWALLOWING	0 - NO SYMPTOMS	
	1 - MILD PAIN ON SWALLOWING	
	2 - MODERATE PAIN ON SWALLOWING	
	3 - SEVERE PAIN ON SWALLOWING/TUBE FEEDING/ IV FLUIDS/ TPN	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:DYSPHAGIA	0 - NO SYMPTOMS	
	1 - SYMPTOMS BUT ABLE TO EAT REGULAR DIET	
	2 - SYMPTOMS AND ALTERED EATING/SWALLOWING	
	3 - SEVERELY ALTERED EATING/SWALLOWING/TUBE FEEDING/TPN/IV FLUIDS	
	4 - LIFE-THREATENING CONSEQUENCES/URGENT INTERVENTION	

BLEEDING IN VOMIT OR BLACK, TARRY STOOLS	0 - NEVER	
	1 - OCCASIONAL	
	2 - SOMETIMES	
	3 - EVERY DAY	

VHB2:LARYNX

INSTRUCTION: **PLEASE RECORD WORST SYMPTOMS DURING PAST MONTH**

VHB2:HOARSENESS	0 – NO SYMPTOMS	
	1 – MILD/ INTERMITTENT/VOICE CHANGE BUT FULLY UNDERSTANDABLE	
	2 – MODERATE/ PERSISTENT VOICE CHANGES/ UNDERSTANDABLE ON TELEPHONE	
	3 – SEVERE CHANGES/ WHISPERED SPEECH/ VOICE AID UP TO 50%	
	4 - NON-UNDERSTANDABLE/ APHONIC/ VOICE AID OVER 50%/ WRITTEN COMM.	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2: LARYNGEAL INFLAMMATION	0 – NONE	
	1 – MILD SORE THROAT	
	2 – MODERATE SORE THROAT/ NEEDING NON-NARCOTIC ANALGESIA	
	3 – SEVERE SORE THROAT INCLUDING NEEDING NARCOTIC ANALGESIA	
	4 - TRACHEOSTOMY/INTUBATION NEEDED FOR BREATHING/COUGHING UP BLOOD	

VHB2: BREATHING	0 - NO PROBLEMS	
	1 - OCCASIONAL DIFFICULTY	
	2 - DIFFICULTY SOMETIMES	
	3 - BREATHING DIFFICULT MOST OF THE TIME	
	4 - BREATHING ALWAYS NOISY	

VHB2:EARS

INSTRUCTION: **PLEASE RECORD WORST SYMPTOMS DURING PAST MONTH**

VHB2:EAR INFLAMMATION	0 – NONE	
	1 - ITCHING AND FLAKINESS INSIDE EAR CANAL/ NOT NEEDING ANY TREATMENT	
	2 - NEEDING EAR DROPS FOR INFLAMMATION	
	3 - SEVERE INFLAMMATION WITH DISCHARGE OR PERFORATED EARDRUM	
	4 - INFLAMMATION/INFECTION/ NECROSIS OF BONE DUE TO EAR INFLAMMATION	
	5 - URGENT OPERATION NEEDED	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2: EAR PAIN	0 – NONE	
	1 - MILD PAIN	
	2 - MODERATE PAIN OR LIMITING INSTRUMENTAL ADL	
	3 - SEVERE PAIN OR LIMITING SELF-CARE ADL	
VHB2:TINNITUS	0 – NONE	
	1 - MILD SYMPTOMS AND INTERVENTION NOT INDICATED	
	2 - MODERATE SYMPTOMS OR LIMITING INSTRUMENTAL ADL	
	3 - SEVERE SYMPTOMS OR LIMITING SELF-CARE ADL	
VHB2:HEARING	0 – NORMAL	
	1 - REDUCED HEARING DETECTED ON TESTING BUT PATIENT NOT NOTICED THIS	
	2 - PATIENT AWARE OF REDUCED/ INTERVENTION NOT INDICATED	
	3 - HEARING AID/ INTERVENTION INDICATED/ LIMITING SELF-CARE ADL	
	4 - COMPLETE DEAFNESS ONE/ BOTH EARS AND NOT CORRECTABLE	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2: EAR SKIN	0 - NO SYMPTOMS	
	1 - ITCHING AND FLAKINESS INSIDE EAR CANAL	
	2 - INFLAMMATION IN EAR CANAL	
	3 - NON-HEALING ULCER	
	4 - INFLAMMATION/INFECTION/NECROSIS OF BONE DUE TO EAR INFLAMMATION	
	5 - URGENT OPERATIVE INTERVENTION	

VHB2:EYES

INSTRUCTION: **PLEASE RECORD WORST SYMPTOMS DURING PAST MONTH**

VHB2: VISION	0 - NO PROBLEMS	
	1 - SOME CHANGES IN VISION	
	2 - SEVERE LOSS OF VISION	
	3 - BLINDNESS IN AFFECTED EYE	

VHB2:PHOTOPHOBIA	0 - NONE	
	1 - PRESENT BUT VISION NORMAL	
	2 - PRESENT AND REDUCED VISION	
	3 - PRESENT AND MAJOR REDUCTION IN VISION	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:EYE INFLAMMATION	0 – NONE	
	1 - MILD SYMPTOMS AND INTERVENTION NOT INDICATED	
	2 – INTERVENTION INDICATED/ LIMITING INSTRUMENTAL ADL	
	3 - CORNEAL ULCERATION	
	4 - DECLINE IN VISION OR LIMITING SELF-CARE ADL	
	5 - PERFORATION OR BLINDNESS IN AFFECTED EYE	
VHB2:DRY EYE	0 – NONE	
	1 - MILD SYMPTOMS RELIEVED BY LUBRICANTS	
	2 - SYMPTOMATIC OR MULTIPLE AGENTS INDICATED OR LIMITING INSTRUMENTAL ADL	
	3 - DECREASE IN VISION DUE TO DRY EYE OR LIMITING SELF CARE ADL	
VHB2:WATERING EYE	0 -NONE	
	1 - PRESENT BUT NO INTERVENTION INDICATED	
	2 - INTERVENTION INDICATED	
	3- OPERATIVE INTERVENTION INDICATED	
VHB2:GLAUCOMA	0 – NONE	
	1 - SINGLE TOPICAL AGENT	
	2 - MULTIPLE AGENTS INDICATED OR LIMITING INSTRUMENTAL ADL	
	3 - NOTICEABLE VISUAL FIELD DEFECTS/ OPERATION / LIMITING SELF-CARE ADL	
	4 - BLINDNESS DUE TO GLAUCOMA	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

SPINAL CORD

TINGLING OR ELECTRIC SENSATION DOWN BACK	0 - NEVER	
	1 - OCCASIONAL AND MILD	
	2 - SOMETIMES BUT BEARABLE	
	3 - OFTEN PRESENT OR SEVERE	
	4 - CONSTANT AND VERY SEVERE	

NUMBNESS	0 – NONE	
	1 - VERY SLIGHT	
	2 – PRESENT	
	3 - TOTAL LOSS OF SENSATION AND DANGER OF SELF-INJURY	
	4 - LIFE-THREATENING CONSEQUENCES/URGENT INTERVENTION	

MUSCLE WEAKNESS IN LIMBS	0 – NONE	
	1 - VERY SLIGHT LOSS	
	2 – PRESENT	
	3 - NO MOVEMENT IN LIMB	
	4 - LIFE-THREATENING CONSEQUENCES/URGENT INTERVENTION	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

BLADDER/BOWEL INCONTINENCE	0 - NO PROBLEMS	
	1 - OCCASIONAL INCONTINENCE	
	2 - INCONTINENT SOMETIMES	
	3 - OFTEN INCONTINENT	
	4 - COMPLETELY INCONTINENT	

VHB2:BRACHIAL PLEXOPATHY

INSTRUCTION: **PLEASE RECORD WORST SYMPTOMS DURING PAST MONTH - BRACHIAL PLEXOPATHY IS DAMAGE TO THE BRACHIAL PLEXUS**

VHB2:BRACHIAL PLEXOPATHY	0 – NONE	
	1 - ASYMPTOMATIC AND CLINICAL OBSERVATION ONLY / NO INTERVENTION	
	2 - MODERATE SYMPTOMS OR LIMITING INSTRUMENTAL ADL	
	3 - SEVERE SYMPTOMS OR LIMITING SELF CARE ADL	

VHB2:HYPOTHYROIDISM

VHB2:HYPOTHYROIDISM	0 – NONE	
	1 - ASYMPTOMATIC/ CLINICAL OBSERVATIONS/ NO INTERVENTION	
	2 - SYMPTOMATIC/ THYROID REPLACEMENT INDICATED/ LIMITING INSTRUMENTAL ADL	
	3 - SEVERE SYMPTOMS/ LIMITING SELF CARE ADL/ HOSPITALIZATION	
	4 - LIFE-THREATENING CONSEQUENCES OR URGENT INTERVENTION INDICATED	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:GENERAL

INSTRUCTION: **PLEASE RECORD WORST SYMPTOMS DURING PAST MONTH**

VHB2:ANOREXIA

0. NONE	
1. LOSS OF APPETITE BUT NO SIGNIFICANT DECREASE IN ORAL INTAKE	
2. ORAL INTAKE DECREASED WITHOUT SIGNIFICANT IMPACT	
3 - INADEQUATE ORAL CALORIE OR FLUID INTAKE/FEEDING TUBE/TPN	
4 - LIFE-THREATENING CONSEQUENCES/URGENT INTERVENTION	

VHB2:NAUSEA

0 – NONE	
1 - LOSS OF APPETITE DUE TO NAUSEA BUT NO CHANGE IN EATING HABITS	
2 - ORAL INTAKE DECREASED DUE TO NAUSEA WITHOUT SIGNIFICANT IMPACT	
3 - INADEQUATE ORAL CALORIE/ FLUID INTAKE/TUBE FEEDING/TPN	

VHB2:VOMITING

0 – NONE	
1 - 1-2 EPISODES IN 24 HOURS	
2 - 3-5 EPISODES IN 24 HOURS	
3 - >= 6 EPISODES IN 24 HOURS/ TUBE FEEDING/ TPN/ HOSPITALIZATION	
4 - LIFE-THREATENING CONSEQUENCES OR URGENT INTERVENTION	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2:CONSTIPATION	0 – NONE	
	1 - OCCASIONAL/INTERMITTENT SYMPTOMS/ OCCASIONAL LAXATIVES, DIET MODIFICATION	
	2 - PERSISTENT SYMPTOMS/REGULAR LAXATIVES/LIMITING INSTRUMENTAL ADL	
	3 - MANUAL EVACUATION NEEDED/ LIMITING SELF CARE ADL	
	4 - LIFE-THREATENING CONSEQUENCES/ URGENT INTERVENTION INDICATED	

VHB2:FATIGUE	0. NONE	
	1. FATIGUE RELIEVED BY REST	
	2. FATIGUE NOT RELIEVED BY REST OR LIMITING INSTRUMENTAL ADL	
	3. FATIGUE LIMITING SELF CARE ADL	

VHB2: WHO PERFORMANCE STATUS	0 – FULLY ACTIVE AND NO CHANGE	
	1 – UNABLE TO CARRY OUT HEAVY PHYSICAL WORK, BUT CAN DO ANYTHING ELSE	
	2 – ACTIVE MORE THAN HALF THE DAY/ SELF CARING/ UNABLE TO WORK	
	3 – IN BED OR SITTING MORE THAN HALF THE DAY/ REDUCED ABLITY TO SELF CARE	
	4 – IN BED OR A CHAIR ALL THE TIME AND UNABLE TO SELF CARE	

Baseline toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHB2: OTHER COMMENTS

VoxTox – Linking Radiation dose at the voxel level with toxicity

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

Obs View Label: VOX:H&N Acute Assessment V2
Obs View Desciption: VOX:H&N Acute Assessment
Active Date: 11/11/2013
Code description Vox:H&N Acute Ass (VHA1.2)

Please ensure that all information in red is completed as this is mandatory data

VHA1.1: WEIGHT (KG) _____

VHA1.1:PRE-EXISTING MORBIDITIES

SINCE LAST ASSESSMENT :-
PLEASE RECORD ANY NEW DIAGNOSES, PROCEDURES OR SURGERY SINCE LAST STUDY ASSESSMENT

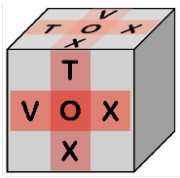
VHA1.1: NEW MEDICAL PROBLEMS?	0- NO	
	1- YES	

VHA1.1: DESCRIBE PROBLEM _____

VHA1.1:SMOKING

VHA1.1:SMOKING STATUS CHANGED?	0- NO	
	1- YES	

VHA1.1: DESCRIBE SMOKING CHANGE _____



VoxTox – Linking Radiation dose at the voxel level with toxicity

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1.1:ALCOHOL

VHA1.1:ALCOHOL STATUS CHANGED?	0- NO	
	1- YES	

VHA1.1 : DESCRIBE ALCOHOL CHANGE

_____

_____

_____

VHA1.1: CONCURRENT CHEMOTHERAPY

VHA1.1:CONCURRENT CHEMOTHERAPY?	0- NO	
	1- YES	
VHA1.1:CHEMO TYPE	0- CISPLATIN	
	1- CETUXMAB	
	2- OTHER CHEMO	

VHA1.1:OTHER CHEMO _____

VHA1.1: SKIN

PLEASE RECORD WORST SYMPTOMS SINCE LAST ASSESSMENT

VHA1.1:RADIATION DERMATITIS	0. NONE	
	1 .FAINT ERYTHEMA OR DRY DESQUAMATION	
	2 MODERATE/ TENDER/BRIGHT ERYTHEMA / PATCHY MOIST DESQUAMATION	
	3. WIDESPREAD MOIST DESQUAMATION / BLEEDING INDUCED BY MINOR TRAUMA	
	4. SKIN NECROSIS/ ULCERATION OF FULL THICKNESS DERMIS/ HAEMORRHAGE	

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1.1:PRURITUS	0. NONE	
	1. MILD AND LOCALIZED	
	2. INTENSE/ WIDESPREAD/ INTERMITTENT/ SCRATCHING CHANGES/ ORAL MEDS	
	3. CONSTANT/ LIMITING SLEEP/ ORAL STEROIDS/ IMMUNOSUPPRESSIANTS	
VHA1.1:ALOPECIA	0 . NO CHANGE FROM PRE-RT BASELINE	
	1. SOME HAIR LOSS	
	2. >=50% PRE-RT BASELINE OR ASSOCIATED WITH PSYCHOLOGICAL IMPACT	

VHA1.1:MOUTH

VHA1.1:DRY MOUTH	0. NO SYMPTOMS	
	1. MILD SYMPTOMS WITHOUT SIGNIFICANT DIETARY ALTERATION	
	2. MODERATE SYMPTOMS/ ORAL INTAKE ALTERATIONS	
	3 .UNABLE TO ADEQUATELY NOURISH ORALLY	
VHA1.1:SALIVA AND TASTE	0 . NO SYMPTOMS	
	1. SLIGHTLY THICKENED SALIVA OR SLIGHTLY ALTERED TASTE E.G. METALLIC	
	2 . THICK, STICKY SALIVA/ CHANGED TASTE/ DIET/ SECRETION-INDUCED SYMPTOMS	
	3. SEVERE SECRETION-INDUCED SYMPTOMS/ TUBE FEEDING/ TPN	
	4. LIFE-THREATENING CONSEQUENCES OR URGENT INTERVENTION INDICATED	
	5. ACUTE SALIVARY GLAND NECROSIS	

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1.1:ORAL MUCOSA	0. NO INFLAMMATION OF ORAL MUCOSA	
	1. ASYMPTOMATIC / MILD SYMPTOMS AND INTERVENTION NOT INDICATED	
	2. MODERATE PAIN / MODIFIED DIET ; NOT INTERFERING WITH ORAL INTAKE	
	3. SEVERE PAIN / INTERFERING WITH ORAL INTAKE	
	4. LIFE-THREATENING CONSEQUENCES OR URGENT INTERVENTION INDICATED	
	5. ULCERATION / HAEMORRHAGE / NECROSIS	

VHA1.1:TRISMUS	0 . NONE	
	1. DECREASED RANGE OF MOVEMENT (ROM) BUT NO IMPAIRMENT IN EATING	
	2 . DECREASED ROM REQUIRING SMALL BITES/ SOFT FOODS/ PUREES	
	3. DECREASED ROM UNABLE TO ADEQUATELY NOURISH / HYDRATE ORALLY	

VHA1.1: PHARYNX AND OESPHAGUS

VHA1.1:DYSPHAGIA	0. NO SYMPTOMS	
	1. SYMPTOMS BUT ABLE TO EAT REGULAR DIET	
	2. SYMPTOMS AND ALTERED EATING/SWALLOWING	
	3. SEVERELY ALTERED EATING/ SWALLOWING/ TUBE FEEDING/TPN/IV FLUID	
	4. LIFE-THREATENING CONSEQUENCES INCLUDING COMPLETE OBSTRUCTION	

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1.1: PAIN ON SWALLOWING	0. NO SYMPTOMS	
	1. MILD PAIN ON SWALLOWING	
	2. MODERATE PAIN ON SWALLOWING	
	3- SEVERE PAIN ON SWALLOWING/ TUBE FEEDING/ IV FLUIDS/ TPN	
	4 -ULCERATION, PERFORATION OR FISTULA	

VHA1.1:LARYNX

VHA1.1:HOARSENESS	0.NO SYMPTOMS	
	1. MILD/ INTERMITTENT/ VOICE CHANGE BUT FULLY UNDERSTANDABLE	
	2. MODERATE/ PERSISTENT VOICE CHANGES/ UNDERSTANDABLE ON TELEPHONE	
	3. SEVERE CHANGES/ WHISPERED SPEECH/ VOICE AID UP TO 50%	
	4 - NON-UNDERSTANDABLE/ APHONIC/ VOICE AID OVER 50%/ WRITTEN COMM.	

LARYNGEAL INFLAMMATION	0. NONE	
	1. MILD SORE THROAT	
	2. MODERATE SORE THROAT/ NON-NARCOTIC ANALGESIA	
	3. SEVERE THROAT PAIN INCLUDING NEEDING NARCOTIC ANALGESIA	
	4. TRACHEOSTOMY/ INTUBATION NEEDED FOR BREATHING/ COUGHING UP BLOOD	

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1.1:EARS

VHB1:EARS INFLAMMATION	0. NONE	
	1. ITCHING AND FLAKINESS INSIDE EAR CANAL NOT NEEDING ANY TREATMENT	
	2. NEEDING EAR DROPS FOR INFLAMMATION	
	3. SEVERE INFLAMMATION WITH DISCHARGE OR PERFORATED EARDRUM	
	4. INFLAMMATION/INFECTION/ NECROSIS OF BONE DUE TO EAR INFLAMMATION	
	5. URGENT OPERATION NEEDED	

VHA1.1: EAR PAIN	0 . NONE	
	1. MILD PAIN	
	2. MODERATE PAIN OR LIMITING INSTRUMENTAL ADL	
	3. SEVERE PAIN OR LIMITING SELF-CARE ADL	

VHA1.1:TINNITUS	0. NONE	
	1. MILD SYMPTOMS AND INTERVENTION NOT INDICATED	
	2. MODERATE SYMPTOMS OR LIMITING INSTRUMENTAL ADL	
	3. SEVERE SYMPTOMS OR LIMITING SELF-CARE ADL	

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1.1:HEARING	0 .NORMAL	
	1. REDUCED HEARING DETECTED ON TESTING BUT PATIENT NOT NOTICED THIS	
	2. PATIENT AWARE OF REDUCED HEARING/ INTERVENTION NOT INDICATED	
	3. HEARING AID OR INTERVENTION INDICATED OR LIMITING SELF-CARE ADL	
	4 - COMPLETE DEAFNESS ONE OR BOTH EARS AND NOT CORRECTABLE	

VHA1.1:EYES

VHA1.1:EYE INFLAMMATION	0. NONE	
	1 . MILD SYMPTOMS AND INTERVENTION NOT INDICATED	
	2. INTERVENTION INDICATED OR LIMITING INSTRUMENTAL ADL	
	3. CORNEAL ULCERATION	
	4. DECLINE IN VISION OR LIMITING SELF-CARE ADL	
	5. PERFORATION OR BLINDNESS IN AFFECTED EYE	

VHA1.1:DRY EYE	0. NONE	
	1. MILD SYMPTOMS RELIEVED BY LUBRICANTS	
	2. SYMPTOMATIC / MULTIPLE AGENTS INDICATED/ LIMITING INSTRUMENTAL ADL	
	3. DECREASE IN VISION DUE TO DRY EYE OR LIMITING SELF CARE ADL	

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1.1:WATERING EYE	0. NONE	
	1.PRESENT AND STABLE COMPARED WITH BASELINE	
	2. PRESENT AND INCREASED COMPARED WITH BASELINE BUT NO INTERVENTION	
	3. INTERVENTION INDICATED	
	4. OPERATIVE INTERVENTION INDICATED	
VHA1.1:GLAUCOMA	0. NONE	
	1. SINGLE TOPICAL AGENT	
	2. MULTIPLE AGENTS INDICATED/ LIMITING INSTRUMENTAL ADL	
	3. NOTICEABLE VISUAL FIELD DEFECTS/ OPERATION /LIMITING SELF-CARE ADL	
	4. BLINDNESS DUE TO GLAUCOMA	

VHA1.1:GENERAL

VHA1.1:ANOREXIA	0. NONE	
	1. LOSS OF APPETITE BUT NO SIGNIFICANT DECREASE IN ORAL INTAKE	
	2. ORAL INTAKE DECREASED WITHOUT SIGNFICANT IMPACT	
	3. INADEQUATE ORAL CALORIE OR FLUID INTAKE/TUBE FEEDING/TPN	
	4. LIFE THREATING CONSEQUENCES / URGENT INTERVENTION	

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

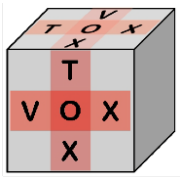
VHA1.1:NAUSEA	0. NONE	
	1. LOSS OF APPETITE DUE TO NAUSEA BUT NO CHANGE IN EATING HABITS	
	2 .ORAL INTAKE DECREASED DUE TO NAUSEA WITHOUT SIGNIFICANT IMPACT	
	3. INADEQUATE ORAL CALORIE OR FLUID INTAKE	
VHA1.1:VOMITING	0. NONE	
	1. 1-2 EPISODES IN 24 HOURS	
	2. 3-5 EPISODES IN 24 HOURS	
	3. >= 6 EPISODES IN 24 HOURS OR TUBE FEEDING, TPN OR HOSPITALIZATION	
	4. LIFE-THREATENING CONSEQUENCES OR URGENT INTERVENTION	
VHA1.1:CONSTIPATION	0. NONE	
	1. OCCASIONAL/INTERMITTENT/ OCCASIONAL LAXATIVES/ DIET MODIFICATION	
	2. PERSISTENT SYMPTOMS/ REGULAR LAXATIVES/LIMITING INSTURMENTAL ADL	
	3. MANUAL EVACUATION NEEDED OR LIMITING SELF CARE ADL	
	4. LIFE-THREATENING CONSEQUENCES/ URGENT INTERVENTION INDICATED	
VHA1.1:FATIGUE	0. NONE	
	1. FATIGUE RELIEVED BY REST	
	2. FATIGUE NOT RELIEVED BY REST OR LIMITING INSTRUMENTAL ADL	
	3. FATIGUE LIMITING SELF CARE ADL	

Acute toxicity clinical reported form for those within head and neck region consolidation & discovery cohort

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1.1: WHO PERFORMANCE STATUS	0. FULLY ACTIVE AND NO CHANGE	
	1. UNABLE TO CARRY OUT HEAVY PHYSICAL WORK, BUT CAN DO ANYTHING ELSE	
	2. ACTIVE MORE THAN HALF THE DAY/ SELF CARING/ UNABLE TO WORK	
	3. IN BED OR SITTING MORE THAN HALF THE DAY AND REDUCED ABLITY TO SELF CARE	
	4. IN BED OR A CHAIR ALL THE TIME AND UNABLE TO SELF CARE	

VPB1: OTHER COMMENTS



VoxTox – Linking Radiation dose at the voxel level with toxicity

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

.WEIGHT IN KG _____

SMOKER		
	0 - NEVER SMOKED	
	1 - EX-SMOKER	
	2 - CURRENT	

AVERAGE NO PER DAY _____

TYPE OF SMOKING		
	0 - CIGARETTES	
	1 - ROLL UPS	
	2 - CIGARS	
	3 - PIPE	
	4 - OTHER	

ALCOHOL

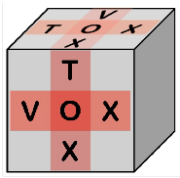
ALCOHOL		
	0 - CURRENT	
	1 -EX-DRINKER	
	2 - NEVER DRUNK	

AVERAGE NO UNITS PER WEEK _____

NEW MEDICAL DIAGONIS		
	0. NO	
	1. YES	

PLEASE STATE _____

.PERFORMANCE STATUS		
	0 – FULLY ACTIVE AND NO CHANGE	
	1 – UNABLE TO CARRY OUT HEAVY PHYSICAL WORK, BUT CAN DO ANYTHING ELSE	
	2 – ACTIVE MORE THAN HALF THE DAY/ SELF CARING/ UNABLE TO WORK	
	3 – IN BED OR SITTING MORE THAN HALF THE DAY/ REDUCED ABLITY TO SELF CARE	
	4 – IN BED OR A CHAIR ALL THE TIME AND UNABLE TO SELF CARE	



VoxTox – Linking Radiation dose at the voxel level with toxicity

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

.CHANGE IN PS/ADL C/W PRE-RT	0. NO CHANGE	
	1. SOME CHANGE	
	2. DRAMATIC CHANGE	

.REASON/S FOR REDUCED PS/ADL (SELECT ALL THAT APPLY)	DRY MOUTH	
	SALIVA AND TASTE	
	SKIN SENSATION	
	HEARING LOSS	
	TINNITUS	
	NUMBNESS	
	MUSCLE WEAKNESS IN LIMBS	
	PHOTOPHOBIA	

.OTHER REASON _____

.TIREDNESS	0. NONE	
	1. MILD AND RELEIVED BY REST	
	2. MODERATE OR NOT RELEIVED BY REST	
	3. SEVERE OR LIMITING SELF CARE	

DIETARY CHANGES	0. NO	
	1. YES	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

.DIET (SELECT ALL THAT APPLY)	OCCAS. SALIVA SUBSTITUTE	
	FREQ. SALIVA SUBSTITUTE/WATER	
	NEEDS SALIVA SUBSTITUTE/WATER IN ORDER TO EAT	
	NON-ACIDIC FOODS ONLY	
	SMALL BITES	
	SEMI-SOFT FOOD	
	SOFT FOOD	
	SOFT FOOD BUT DIFFICULT	
	LIQUIDS ONLY	
	TEMPORARY NG TUBE	
	TPN IE NUTRITION INTO VEIN	
	PERMANENT FEEDING TUBE	
	IV FLUIDS	

.REASON FOR ALTERED DIET (SELECT ALL THAT APPLY)	DRY MOUTH	
	SALIVA	
	TASTE CHANGES	
	SWALLOWING DIFFICULTY	
	WEIGHT LOSS	
	TRISMUS	
	CHEWING DIFFICULTY	
	OTHER REASON	

MEDICAL INTERVENTION	0. NO	
	1. YES	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

.MEDICAL INTERVENTIONS (SELECT ALL THAT APPLY)		
	SURGERY TO HEAD AND NECK	
	WASHOUT/CLEANSE	
	ANTIBIOTICS	
	OCCASIONAL DILATATION	
	REGULAR DILATATION	
	OCCASIONAL TRANSFUSION	
	REGULAR TRANSFUSIONS	
	HUMIDIFIER/NEBULISER	
	TOOTH FILLING	
	EXTRACTION OF SOME TEETH	
	EXTRACTION OF ALL TEETH	
	HYPERBARIC OXYGEN	
	HEARING AID	
	HOSPITAL INPATIENT STAY	
	OTHER MEDICAL INTERVENTION	
	TEMPORARY TRACHEOSTOMY	
	PERMANENT TRACHEOSTOMY	

.REASON FOR INTERVENTION (SELECT ALL THAT APPLY)		
	ULCER	
	SWALLOWING DIFFICULTIES	
	WEIGHT LOSS	
	BREATHING DIFFICULTIES	
	HOARSENESS	
	TOOTH DECAY	
	OSTEORADIONECROSIS	
	OEDEMA	
	SKIN FIBROSIS	
	EARS	
	BLEEDING	
	EYES	

OTHER REASON _____

DRUGS FOR HEAD & NECK		
	0. NO	
	1. YES	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

DRUGS FOR HEAD AND NECK (SELECT ALL THAT APPLY)	
ANALGESIA	
REGULAR ANALGESIA	
ANTACIDS	
NUTRITIONAL SUPPLEMENTS	
IRON	
STEROIDS	
FLUORIDES FOR TEETH	
MUSCLE RELAXANTS	
DIURETIC/WATER MEDICATION	
CREAMS AND MEDS FOR SKIN	
DAILY CREAMS AND MEDS FOR SKIN	
OCCASIONAL EARDROPS/OINTMENTS	
REGULAR EARDROPS/OINTMENTS	
EYE DROPS/OINMENTS	

MOUTH AND THROAT

PAIN - MOUTH AND THROAT MUCOSA AND SWALLOW	
0. NONE	
1. OCCASIONAL AND MILD	
2. SOMETIMES BUT BEARABLE	
3. OFTEN PRESENT OR SEVERE	
4. CONSTANT AND VERY SEVERE	

PAIN MANAGEMENT	
0. NONE NEEDED	
1. OCCASIONAL ANALGESIA	
2. REGULAR ANALGESIA	
3. NARCOTIC ANALGESIA	
4. SURGERY	

DRY MOUTH	
0. NO SYMPTOMS	
1. MILD SYMPTOMS	
2. MODERATE SYMPTOMS	
3. COMPLETE DRYNESS	

SALIVA	
0. NO SYMPTOMS	
1. SLIGHTLY THICKENED SALIVA	
2. THICK STICKY SALIVA	
3. NO SALIVA	
4. NON-HEALING ULCER	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

TASTE	0. NO SYMPTOMS	
	1. OCCASIONAL AND SLIGHTLY ALTERED TASTE	
	2. SOMETIMES ALTERED TASTE	
	3. TASTE ALMOST ALWAYS ALTERED	
SWALLOWING DIFFICULTY	0. NONE	
	1. PRESENT	
	2. LIFE THREATENING CONSEQUENCES & COMPLETELY UNABLE TO SWALLOW	
	3. 2. PERFORATION OR FISTULA OESOPHAGUS	
BLEEDING IN VOMIT OR BLACK/TARRY STOOLS	1. NEVER	
	2. OCCASIONAL	
	3. SOMETIMES	
	4. EVERY DAY	

JAW

PAIN IN JAW	0. NONE	
	1. OCCASIONAL AND MILD	
	2. SOMETIMES BUT BEARABLE	
	3. OFTEN PRESENT OR SEVERE	
	4. CONSTANT AND VERY SEVERE	
JAW PAIN MANAGEMENT	0. NONE NEEDED	
	1. OCCASIONAL ANALGESIA	
	2. REGULAR ANALGESIA	
	3. NARCOTIC ANALGESIA	
	4. SURGERY	
DENTURES	0. NOT NEEDED	
	1. FIT WELL	
	2. LOOSE FIT	
	3. NOT ABLE TO USE	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

TRISMUS	0. NONE	
	1. MILD	
	2. MODERATE	
	3. SEVERE	
	4. UNABLE TO OPEN MOUTH AT ALL	
OSTEORADIONECROSIS	0. NONE	
	1. PRESENT BUT NOT CAUSING SYMPTOMS	
	2. PRESENT AND CAUSING SYMPTOMS	
	3. SEVERE SYMPTOMS OR LIMITING SELF CARE	
	4. LIFE THREATENING/ URGENT INTERVENTION	
PAIN IN TEETH	0. NONE	
	1. OCCASIONAL AND MILD	
	2. SOMETIMES BUT BEARABLE	
	3. OFTEN PRESENT OR SEVERE	
	4. CONSTANT AND VERY SEVERE	
TOOTH PAIN MANAGEMENT	0. NONE NEEDED	
	1. OCCASIONAL ANALGESIA	
	2. REGULAR ANALGESIA	
	3. NARCOTIC ANALGESIA	
	4. EXTRACTION	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

LARYNX

HOARSE VOICE	0. NO SYMPTOMS	
	1. MILD/ INTERMITTENT/ VOICE CHANGE BUT FULLY UNDERSTANDABLE	
	2. MODERATE/ PERSISTENT VOICE CHANGES/ UNDERSTANDABLE ON TELEPHONE	
	3. SEVERE CHANGES/ WHISPERED SPEECH/ VOICE AID UP TO 50%	
	4. NON-UNDERSTANDABLE/ APHONIC/ VOICE AID OVER 50%/ WRITTEN COMM.	
	5. NECROSIS	

BREATHING	0. NO PROBLEMS	
	1. OCCASIONAL DIFFICULTY	
	2. DIFFICULTY SOMETIMES	
	3. BREATHING DIFFICULT MOST OF THE TIME	
	4. BREATHING ALWAYS NOISY	

SKIN IN TREATED AREA

SCALINESS OR ROUGHNESS	0. NONE	
	1. PRESENT BUT NOT CAUSING SYMPTOMS	
	2. CAUSING SYMPTOMS	
	3. REQUIRING CONSTANT ATTENTION	

SKIN SENSATION	0. NORMAL	
	1. SENSITIVE OR ITCHY	
	2. PAINFUL AT TIMES	
	3. PAINFUL MOST OF THE TIME	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

SKIN SWELLING	0. NONE	
	1. PRESENT BUT NOT CAUSING SYMPTOMS	
	2. CAUSING SYMPTOMS	
	3. CAUSING SOME LIMITATION OF FUNCTION	
	4. CAUSING SEVERE LOSS OF FUNCTION	
HAIR LOSS	0. NONE	
	1. THINNING	
	2. LOST IN PATCHES	
	3. COMPLETE LOSS	
CHANGE IN COLOUR	0. NONE	
	1. SLIGHT OR TEMPORARY & NO PSYCOLOGICAL IMPACT	
	2. SLIGHT OR TEMPORARY & PSYCOLOGICAL IMPACT	
	3. MARKED	
TELANGIECTASIA	0. NONE	
	1. MILD	
	2. MODERATE	
	3. SEVERE	
FIBROSIS (THICK HARD SKIN)	0. NONE	
	1. PRESENT BUT NOT CAUSING SYMPTOMS	
	2. CAUSING SYMPTOMS	
	3. CAUSING SOME LIMITATION OF FUNCTION	
	4. CAUSING SEVERE LOSS OF FUNCTION	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

ATROPHY (THINNING)	0. NONE	
	1. PRESENT BUT NOT CAUSING SYMPTOMS	
	2. CAUSING SYMPTOMS	
	3. INFECTION IN THE AREA	
	4. CAUSING SEVERE LOSS OF FUNCTION	
	5. NON-HEALING ULCER	

EARS

EAR SKIN	0. NO SYMPTOMS	
	1. ITCHY & FLAKENESS INSIDE EAR CANAL	
	2. INFLAMATION IN EAR CANAL	
	3. NO HEALING ULCER	
	4. INFLAMATION, INFECTION, NECROSIC BONE	
	5. URGENT OPERATION	

PAIN IN EAR	0. NONE	
	1. OCCASIONAL AND MILD	
	2. SOMETIMES BUT BEARABLE	
	3. OFTEN PRESENT OR SEVERE	
	4. CONSTANT AND VERY SEVERE	

EAR PAIN MANAGEMENT	0. NONE NEEDED	
	1. OCCASIONAL ANALGESIA	
	2. REGULAR ANALGESIA	
	3. ORAL/FEEDING TUBE NARCOTIC ANALGESIA	
	4. IV, IM OR SC NARCOTIC ANALGESIA	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

TINNITUS	0. NEVER	
	1. OCCASIONAL	
	2. SOMETIMES	
	3. MOST OF THE TIME	
	4. ALWAYS	
HEARING	0. NORMAL	
	1. SLIGHT LOSS	
	2. DIFFICULTIES WITH FAINT SPEECH	
	3. DIFFICULTIES WITH LOUD SPEECH	
	4. COMPLETE DEAFNESS	
SPINAL CORD		
TINGLING OR ELECTRIC SENSATION DOWN BACK	0. NEVER	
	1. OCCASIONAL AND MILD	
	2. SOMETIMES BUT BEARABLE	
	3. OFTEN PRESENT OR SEVERE	
	4. CONSTANT AND VERY SEVERE	
TINGLING OR ELECTRIC MANAGEMENT	0. NONE NEEDED	
	1. OCCASIONAL ANALGESIA	
	2. INTERMITTENT HIGH DOSE STEROIDS/ REGULAR ANALGESIA	
	3. PERSISTENT HIGH DOSE STEROIDS	
NUMBNESS	0. NONE	
	1. VERY SLIGHT	
	2. PRESENT	
	3. TOTAL LOSS OF SENSATION/ DANGER OF SELF INJURY	
	4. LIFE THREATENING CONSEQUENCE/ URGENT INTERVENTION	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

MUSCLE WEAKNESS IN LIMBS	0. NONE	
	1. VERY SLIGHT LOSS	
	2. PRESENT	
	3. NO MOVEMENT IN LIMB	
	4. LIFE THREATENING CONSEQUENCE/ URGENT INTERVENTION	

BLADDER AND BOWEL CONTINENCE	0. NO PROBLEMS	
	1. OCCASIONAL INCONTINENCE	
	2. INCONTINENT SOMETIMES	
	3. OFTEN INCONTINENT	
	4. COMPLETELY INCONTINENT	

MANAGEMENT OF INCONTINENCE	0. NONE NEEDED	
	1. OCCASIONAL PADS	
	2. SOMETIMES PADS	
	3. REGULAR PADS OR INTERMITTENT CATHETER	
	4. CONSTANT PADS OR INDWELLING CATHETER	

EYES

PLEASE SELECT TOSE THAT APPLY	CORNEAL ULCER	
	RETINOPATHY	
	RETINAL DETACHMENT	
	GLAUCOMA	
	CATARACT	

VISION	0. NO PROBLEM	
	1. SOME CHANGES IN VISION	
	2. SEVERE LOSS IN VISION	
	3. BLINDNESS IN AFFECTED EYE	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

PHOTOPHOBIA	0. NONE	
	1. PRESENT BUT VISION NORMAL	
	2. PRESENT & REDUCED VISION	
	3. PRESENT & MAJOR REDUCTION IN VISION	
EYE PAIN/ DRYNESS	0. NONE	
	1. OCCASSIONAL & MILD	
	2. SOMETIMES BUT BEARABLE	
	3. OFTEN PRESENT OR SERVERE	
	4. CONSTANT AND VERY SERVERE	
EYE PAIN MANAGEMENT	0. NONE NEEDED	
	1. OCCASIONAL ANALGESIA	
	2. REGULAR ANALGESIA	
	3. ORAL/FEEDING TUBE NARCOTIC ANALGESIA	
	4. IV, IM OR SC NARCOTIC ANALGESIA	
WATERING EYES	0. NONE	
	1. OCCASIONAL	
	2. SOMETIMES	
	3. MOST OF THE TIME	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

GENERAL

HYPOTHYRODISM	0. NONE	
	1. ASYMPTOMATIC/ CLNICALLY OBERSERVED/ NO INTERVENTION	
	2. SYMPTOMATIC THYROID REPLACEMENT INDICATED/ LIMITING ADL	
	3. SERVERE SYMPTOMS/HOSPITALIZATION/ LIMITING SELF CARE ADL	
	4. LIFE THREATENING CONSEQUENCE/ URGENT INTERVENTION	

ANOREXIA	0. NONE	
	1. 1. LOSS OF APPETITE BUT NO SIGNIFICANT DECREASE IN ORAL INTAKE	
	2. 2. ORAL INTAKE DECREASED BUT NO MEDICAL INTERVENTION	
	3. REQUIRING IV FLUIDS DUE TO ANOREXIA	
	4. REQUIRING FEEDING TUBE OR TPN DUE TO ANOREXIA	

Late Toxicity Clinical Reported Form For Those Within the Head and Neck Consolidation Cohorts

Visit Number		Date of birth	
Patients initials		VoxTox number	

VHA1:NAUSEA	0. NONE	
	1. LOSS OF APPETITE DUE TO NAUSEA BUT NO CHANGE IN EATING HABITS	
	2. ORAL INTAKE DECREASED DUE TO NAUSEA WITHOUT SIGNIFICANT IMPACT	
	3. INADEQUATE ORAL CALORIE OR FLUID INTAKE	

VHA1:VOMITING	0. NONE	
	1. 1-2 EPISODES IN 24 HOURS	
	2. 3-5 EPISODES IN 24 HOURS	
	3. >= 6 EPISODES IN 24 HOURS OR TUBE FEEDING, TPN OR HOSPITALIZATION	
	4. LIFE-THREATENING CONSEQUENCES OR URGENT INTERVENTION	

VHA1:CONSTIPATION	0. NONE	
	1. OCCASIONAL/INTERMITTENT/ OCCASIONAL LAXATIVES/ DIET MODIFICATION	
	2. PERSISTENT SYMPTOMS/ REGULAR LAXATIVES/LIMITING INSTURMENTAL ADL	
	3. MANUAL EVACUATION NEEDED OR LIMITING SELF CARE ADL	
	4. LIFE-THREATENING CONSEQUENCES/ URGENT INTERVENTION INDICATED	

HNC baseline_v2 toxicity data mapping rules – CTCAE v4.03

Mapped endpoints:

Ear pain

External ear inflammation

Hearing impaired

Tinnitus

Cataract

Corneal ulcer

Dry eye

Glaucoma

Retinopathy

Dry mouth

Dysphagia

Nausea

Salivary duct inflammation

Anorexia

Trismus

Hoarseness

Rules:

IF (CX=1) THEN CTCAE_4.03:ear_pain=1

IF (CX=2) THEN CTCAE_4.03:ear_pain=2

IF (CX=3) THEN CTCAE_4.03:ear_pain=3

IF (CW=1) OR (CW=2) THEN CTCAE_4.03:external_ear_inflammation=1

IF (CW=3) THEN CTCAE_4.03:external_ear_inflammation=2

IF (CW=4) THEN CTCAE_4.03:external_ear_inflammation=3

IF (CW=5) THEN CTCAE_4.03:external_ear_inflammation=4

IF (CY=1) THEN CTCAE_4.03:tinnitus=1

IF (CY=2) THEN CTCAE_4.03:tinnitus=2

IF (CY=3) THEN CTCAE_4.03:tinnitus=3

IF (CZ=1) THEN CTCAE_4.03:hearing_impaired=1

IF (CZ=2) THEN CTCAE_4.03:hearing_impaired=2

IF (CZ=3) THEN CTCAE_4.03:hearing_impaired=3

IF (CZ=4) THEN CTCAE_4.03:hearing_impaired=4

IF (AB=1) AND (DB=0) THEN CTCAE_4.03:cataract=1

IF (AB=1) AND (DB=1) THEN CTCAE_4.03:cataract=2

IF (AB=1) AND (DB=2) THEN CTCAE_4.03:cataract=3

IF (AB=1) AND (DB=3) THEN CTCAE_4.03:cataract=4

IF (DD=3) THEN CTCAE_4.03:corneal_ulcer=2

IF (DD=4) THEN CTCAE_4.03:corneal_ulcer=3

IF (DD=5) THEN CTCAE_4.03:corneal_ulcer=4

IF (DE=1) THEN CTCAE_4.03:dry_eye=1

IF (DE=2) THEN CTCAE_4.03:dry_eye=2

IF (DE=3) THEN CTCAE_4.03:dry_eye=3

EYE PAIN MAPPING NOT POSSIBLE FROM BASELINE DATA

IF (DG=1) THEN CTCAE_4.03:glaucoma=1

IF (DG=2) THEN CTCAE_4.03:glaucoma=2

IF (DG=3) THEN CTCAE_4.03:glaucoma=3

IF (DG=4) THEN CTCAE_4.03:glaucoma=4

IF (AA=1) AND (DB=0) THEN CTCAE_4.03:retinopathy=1

IF (AA=1) AND (DB=1) THEN CTCAE_4.03:retinopathy=2

IF (AA=1) AND (DB=2) THEN CTCAE_4.03:retinopathy=3

IF (AA=1) AND (DB=3) THEN CTCAE_4.03:retinopathy=4

IF (CI=1) OR (CJ=1) THEN CTCAE_4.03:dry_mouth=1

IF (CI=2) THEN CTCAE_4.03:dry_mouth=2

IF (CI>0) AND (CJ=1) AND (AW=1) AND (((AX=1) OR (AY=1) OR (AZ=1)) AND ((BG=1) OR (BH=2))) THEN CTCAE_4.03:dry_mouth=2

IF (CI=3) THEN CTCAE_4.03:dry_mouth=3

IF (CI=2) AND (CJ>2) THEN CTCAE_4.03:dry_mouth=3

IF (CI=2) AND (CJ>1) AND (AW=1) AND (BF=1) AND ((BG=1) OR (BH=1)) THEN CTCAE_4.03:dry_mouth=3

IF (CR=1) THEN CTCAE_4.03:dysphagia=1

IF (CR=2) THEN CTCAE_4.03:dysphagia=2

IF (CR=3) THEN CTCAE_4.03:dysphagia=3

IF (CR=4) THEN CTCAE_4.03:dysphagia=4

IF (DO=1) THEN CTCAE_4.03:nausea=1

IF (DO=2) THEN CTCAE_4.03:nausea=2

IF (DO=3) THEN CTCAE_4.03:nausea=3

IF (CJ=1) THEN CTCAE_4.03:salivary_duct_inflammation=1

IF (CJ=2) THEN CTCAE_4.03:salivary_duct_inflammation=2

IF (CJ=3) THEN CTCAE_4.03:salivary_duct_inflammation=3

IF (CJ>3) THEN CTCAE_4.03:salivary_duct_inflammation=4

IF (DN=1) THEN CTCAE_4.03:anorexia=1

IF (DN=2) THEN CTCAE_4.03:anorexia=2

IF (DN=3) THEN CTCAE_4.03:anorexia=3

IF (DN=4) THEN CTCAE_4.03:anorexia=4

IF (CO=1) THEN CTCAE_4.03:trismus=1

IF (CO=2) THEN CTCAE_4.03:trismus=2

IF (CO=3) THEN CTCAE_4.03:trismus=3

IF (CO=1) THEN CTCAE_4.03:trismus=1

IF (CT=1) THEN CTCAE_4.03:hoarseness=1

IF (CT=2) THEN CTCAE_4.03:hoarseness=2

IF (CT>2) THEN CTCAE_4.03:hoarseness=3

HNC baseline_V2 toxicity data mapping rules - LENT SOM(A)

Mapped endpoints:

Ear – Subjective (pain) NOT POSSIBLE TO DISTINGUISH BETWEEN GRADES 3 AND 4 – CODED AS GRADE 3.

Ear – Subjective (tinnitus) NOT POSSIBLE TO DISTINGUISH BETWEEN GRADES 3 AND 4 – CODED AS GRADE 3.

Ear - Subjective (hearing)

Ear - Objective (skin)

Ear – Management (pain) GRADE 4 CANT BE CODED.

Ear – Management (skin)

Ear – Management (hearing loss) CANNOT BE MAPPED

Mucosa – oral and pharyngeal - Subjective (pain)

Mucosa – oral and pharyngeal – Subjective (dysphagia)

Mucosa – oral and pharyngeal – Subjective (taste alteration)

Mucosa – oral and pharyngeal – Objective (weight)

Mucosa – oral and pharyngeal – Management (pain)

Mucosa – oral and pharyngeal – Management (ulcer) CANNOT BE MAPPED

Mucosa – oral and pharyngeal – Management (dysphagia) (GRADES 3 AND 4 CANNOT BE MAPPED, AS THE BASELINE CRF DOESN'T CONTAIN INFOMRATION ABOUT THE NATURE OF ANALGESIA USED [NARCOTIC OR NOT])

Mucosa – oral and pharyngeal – Management (taste alteration)

Salivary gland – Subjective (xerostomia)

Salivary gland – Objective (saliva)

Salivary gland – Management (xerostomia)

Salivary gland – Global (LENT score)

Mandible – Subjective (pain)

Mandible – Subjective (mastication)

Mandible – Subjective (denture use)

Mandible – Subjective (trismus)

Mandible – Management (pain) CANNOT BE MAPPED

Mandible – Management (trismus and mastication)

Larynx – Subjective (voice/hoarseness)

Larynx – Subjective (breathing)

Skin/subcutaneous tissue – Subjective (scaliness/roughness)

Skin/subcutaneous tissue – Subjective (sensation) (GRADE 4 CANNOT BE MAPPED)

Skin/subcutaneous tissue – Objective (edema)

Skin/subcutaneous tissue – Objective (alopecia)

Skin/subcutaneous tissue – Objective (pigmentation)

Skin/subcutaneous tissue – Objective (telangiectasia)

Skin/subcutaneous tissue – Objective (fibrosis/scar)

Skin/subcutaneous tissue – Objective (atrophy/contraction)

<u>Rules:</u>

LENT_SOM(A):mucosa_oral_&_pharyngeal(Objective_weight)=E

IF (CK=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_pain)=1

IF (CK=2) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_pain)=2

IF (CK=3) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_pain)=3

IF (CK>3) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_pain)=4

IF (CK>1) AND (BN=1) AND (BO=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_pain)=1

IF (CK>1) AND (BN=1) AND (BP=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_pain)=2

IF (CK>2) AND (BN=1) AND (BP=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_pain)=3

IF (CK>3) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_pain)=4

IF (CI=1) THEN LENT_SOM(A):salivary_gland(Subjective_xerostomia)=1

IF (CI=2) THEN LENT_SOM(A):salivary_gland(Subjective_xerostomia)=2

IF (CI=3) AND (AW=1) AND ((AX=1) OR (AY=1) OR (AZ=1) OR (BA=1) OR (BB=1) OR (BC=1)) AND (BG=1) THEN LENT_SOM(A):salivary_gland(Subjective_xerostomia)=3

IF (CI=3) AND (AW=1) AND ((BD=1) OR (BE=1) OR (BF=1)) AND (BG=1)THEN LENT_SOM(A):salivary_gland(Subjective_xerostomia)=4

IF (CI=0) AND ((CJ=1) OR (CJ=2)) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=1

IF ((CI=1) OR (CI=2)) AND (CJ=0) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=1

IF (CI=2) AND ((CJ=1) OR (CJ=2)) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=2

IF (CI=3) AND ((CJ=1) OR (CJ=2)) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=3

IF (CI=3) AND ((CJ=3) OR (CJ=4)) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=4

IF (CI>1) AND (AW=1) AND (AX=1) AND ((BG=1) OR (BH=1)) THEN LENT_SOM(A):salivary_gland(Management_xerostomia)=2

IF (CI>1) AND (AW=1) AND (AY=1) AND ((BG=1) OR (BH=1)) THEN LENT_SOM(A):salivary_gland(Management_xerostomia)=3

IF (CI>1) AND (AW=1) AND (AZ=1) AND ((BG=1) OR (BH=1)) THEN LENT_SOM(A):salivary_gland(Management_xerostomia)=4

IF (CJ=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_taste_alteration)=1

IF (CJ=2) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_taste_alteration)=2

IF (CJ=3) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_taste_alteration)=3

IF (CJ>1) AND (AW=1) AND (BA=1) AND (BI=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_taste_alteration)=1

IF (CJ>1) AND (AW=1) AND (BB=1) AND (BI=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_taste_alteration)=2

IF (CJ>1) AND (AW=1) AND ((BC=1) OR (BD=1)) AND (BI=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_taste_alteration)=3

IF (CJ>1) AND (AW=1) AND ((BE=1) OR (BF=1)) AND (BI=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_taste_alteration)=4

IF (CR>0) AND (AW=1) AND ((BA=1) OR (BB=1) OR (BC=1)) AND (BJ=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=1

IF (CR>1) AND (AW=1) AND (BD=1) AND (BJ=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=2

IF (CR>2) AND (AW=1) AND (BE=1) AND (BJ=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=3

IF (CR>2) AND (AW=1) AND (BF=1) AND (BJ=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=4

IF (CR>0) AND (AW=1) AND ((AX=1) OR (AY=1) OR (AZ=1)) AND (BJ=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_dysphagia)=1

IF (CR>1) AND (AW=1) AND ((BA=1) OR (BB=1) OR (BC=1) OR (BD=1) OR (BE=1)) AND (BJ=1) AND ((BO=1) OR (BP=1)) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_dysphagia)=2

LENT_SOM(A):salivary_gland_global=(LENT_SOM(A):salivary_gland(Subjective_xerostomia) + LENT_SOM(A):salivary_gland(Objective_saliva)+ LENT_SOM(A):salivary_gland(Management_xerostomia))/3

IF (CM=1) THEN LENT_SOM(A):mandible(Subjective_pain)=1

IF (CM=2) THEN LENT_SOM(A):mandible(Subjective_pain)=2

IF (CM=3) THEN LENT_SOM(A):mandible(Subjective_pain)=3

IF (CM=4) THEN LENT_SOM(A):mandible(Subjective_pain)=4

IF (CM>1) AND ((BA=1) OR (BB=1) OR (BC=1)) AND (BL=1) THEN LENT_SOM(A):mandible(Subjective_mastication)=2

IF (CM>1) AND ((BD=1) OR (BE=1) OR (BF=1)) AND (BL=1) THEN LENT_SOM(A):mandible(Subjective_mastication)=3

IF (CL=2) THEN LENT_SOM(A):mandible(Subjective_denture_use)=2

IF (CL=3) THEN LENT_SOM(A):mandible(Subjective_denture_use)=3

IF (CO=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=1

IF (CO=2) THEN LENT_SOM(A):mandible(Subjective_trismus)=2

IF (CO=2) AND (AW=1) AND ((BD=1) OR (BE=1)) AND (BK=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=3

IF (CO=3) THEN LENT_SOM(A):mandible(Subjective_trismus)=4

IF (CP>1) AND (AW=1) AND ((BB=1) OR (BC=1)) AND (BK=1) THEN LENT_SOM(A):mandible(Management_trismus&mastication)=2

IF (CP>1) AND (AW=1) AND (((BD=1) OR (BE=1)) AND ((BK=1) OR (BL=1)) AND (BN=1) AND (BV=1)) THEN LENT_SOM(A):mandible(Management_trismus&mastication)=3

IF (CP>1) AND (AW=1) AND (BF=1) AND (BK=1) THEN LENT_SOM(A):mandible(Management_trismus&mastication)=4

IF (CT=1) THEN LENT_SOM(A):larynx(Subjective_voice/hoarseness)=1

IF (CT=2) THEN LENT_SOM(A):larynx(Subjective_voice/hoarseness)=2

IF (CT=3) THEN LENT_SOM(A):larynx(Subjective_voice/hoarseness)=3

IF (CT=4) THEN LENT_SOM(A):larynx(Subjective_voice/hoarseness)=4

IF (CT>0) AND (CV=1) THEN LENT_SOM(A):larynx(Subjective_breathing)=1

IF (CT>0) AND (CV=2) THEN LENT_SOM(A):larynx(Subjective_breathing)=2

IF (CT>0) AND (CV=3) THEN LENT_SOM(A):larynx(Subjective_breathing)=3

IF (CT>0) AND (CV=4) THEN LENT_SOM(A):larynx(Subjective_breathing)=4

IF (BZ=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_scaliness/roughness)=1

IF (BZ=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_scaliness/roughness)=2

IF (BZ=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_scaliness/roughness)=3

IF (CA=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_sensation)=1

IF (CA=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_sensation)=2

IF (CA=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_sensation)=3

IF (CG=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_edema)=1

IF (CG=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_edema)=2

IF (CG=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_edema)=3

IF (CG=4) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_edema)=4

IF (CC=1) OR (CC=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_pigmentation_change)=1

IF (CC=3) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_pigmentation_change)=2

IF (CD=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_telangiectasia)=1

IF (CD=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_telangiectasia)=2

IF (CD=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_telangiectasia)=3

IF (CF=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_fibrosis_scar)=1

IF (CF=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_fibrosis_scar)=2

IF (CF=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_fibrosis_scar)=3

IF (CF=4) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_fibrosis_scar)=4

IF (CE=1) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_atrophy_contraction)=1

IF (CE=2) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_atrophy_contraction)=2

IF (CE=3) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_atrophy_contraction)=3

IF (CE=4) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_atrophy_contraction)=4

IF (CH=1) OR (CH=2) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_alopecia_scalp)=1

IF (CH=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_alopecia_scalp)=2

IF (CH=4) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_alopecia_scalp)=3

IF (CW=1) THEN LENT_SOM(A):ear(Objective_skin)=1

IF (CW=2) THEN LENT_SOM(A):ear(Objective_skin)=2

IF (CW=3) THEN LENT_SOM(A):ear(Objective_skin)=3

IF (CW=4) THEN LENT_SOM(A):ear(Objective_skin)=4

IF (CW=1) AND (BN=1) AND (BX=1) THEN LENT_SOM(A):ear(Management_skin)=1

IF (CW=2) AND (BN=1) AND (BX=1) THEN LENT_SOM(A):ear(Management_skin)=2

IF (CW=3) OR (CW=4) THEN LENT_SOM(A):ear(Management_skin)=3

IF (CW=5) THEN LENT_SOM(A):ear(Management_skin)=4

IF (CX=1) THEN LENT_SOM(A):ear(Subjective_pain)=1

IF (CX=2) THEN LENT_SOM(A):ear(Subjective_pain)=2

IF (CX=3) THEN LENT_SOM(A):ear(Subjective_pain)=3

IF ((CX=1) OR (CX=2)) AND (BN=1) AND (BO=1) THEN
LENT_SOM(A):ear(Management_pain)=1

IF ((CX=1) OR (CX=2)) AND (BN=1) AND (BP=1) THEN
LENT_SOM(A):ear(Management_pain)=2

IF (CX=3) AND (BN=1) AND (BP=1) THEN LENT_SOM(A):ear(Management_pain)=3

IF (CY=1) THEN LENT_SOM(A):ear(Subjective_tinnitus)=1

IF (CY=2) THEN LENT_SOM(A):ear(Subjective_tinnitus)=2

IF (CY=3) THEN LENT_SOM(A):ear(Subjective_tinnitus)=3

IF (CZ=1) THEN LENT_SOM(A):ear(Subjective_hearing)=1

IF (CZ=2) THEN LENT_SOM(A):ear(Subjective_hearing)=2

IF (CZ=3) THEN LENT_SOM(A):ear(Subjective_hearing)=3

IF (CZ=4) THEN LENT_SOM(A):ear(Subjective_hearing)=4

HNC baseline_v2 toxicity data mapping rules – EORTC QLQ – H&N 35

Mapped endpoints:

31. Have you had pain in your mouth?

32. Have you had pain in your jaw?

40. Have you had problems opening your mouth wide?

41. Have you had a dry mouth?

42. Have you had sticky saliva?

44. Have you had problems with your sense of taste?

46. Have you been hoarse?

Rules:

IF (CQ=0) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_mouth?=1

IF (CQ=1) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_mouth?=2

IF (CQ=2) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_mouth?=3

IF (CQ=3) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_mouth?=4

IF (CI=0) THEN EORTC_QLQ-H&N35:have_you_had_a_dry_mouth?=1

IF (CI=1) THEN EORTC_QLQ-H&N35:have_you_had_a_dry_mouth?=2

IF (CI=2) THEN EORTC_QLQ-H&N35:have_you_had_a_dry_mouth?=3

IF (CI=3) THEN EORTC_QLQ-H&N35:have_you_had_a_dry_mouth?=4

IF (CJ=0) THEN EORTC_QLQ-H&N35:have_you_had_sticky_saliva?=1

IF (CJ=1) THEN EORTC_QLQ-H&N35:have_you_had_sticky_saliva?=2

IF (CJ=2) THEN EORTC_QLQ-H&N35:have_you_had_sticky_saliva?=3

IF (CJ>2) THEN EORTC_QLQ-H&N35:have_you_had_sticky_saliva?=4

IF (CJ=0) THEN EORTC_QLQ-H&N35:have_you_had_problems_with_your_sense_of_taste?=1

IF (CJ=1) THEN EORTC_QLQ-H&N35:have_you_had_problems_with_your_sense_of_taste?=2

IF (CJ=2) THEN EORTC_QLQ-H&N35:have_you_had_problems_with_your_sense_of_taste?=3

IF (CJ>2) THEN EORTC_QLQ-H&N35:have_you_had_problems_with_your_sense_of_taste?=4

IF (CM=0) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_jaw?=1

IF (CM=1) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_jaw?=2

IF (CM=2) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_jaw?=3

IF (CM>2) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_jaw?=4

IF (CO=0) THEN EORTC_QLQ-
H&N35:have_you_had_problems_opening_your_mouth_wide?=1

IF (CO=1) THEN EORTC_QLQ-
H&N35:have_you_had_problems_opening_your_mouth_wide?=2

IF (CO=2) THEN EORTC_QLQ-
H&N35:have_you_had_problems_opening_your_mouth_wide?=3

IF (CO=3) THEN EORTC_QLQ-
H&N35:have_you_had_problems_opening_your_mouth_wide?=4

IF (CT=0) THEN EORTC_QLQ-H&N35:have_you_been_hoarse?=1

IF (CT=1) THEN EORTC_QLQ-H&N35:have_you_been_hoarse?=2

IF (CT=2) THEN EORTC_QLQ-H&N35:have_you_been_hoarse?=3

IF (CT>2) THEN EORTC_QLQ-H&N35:have_you_been_hoarse?=4

HNC baseline_V2 toxicity data mapping rules - RTOG

Mapped endpoints:

Skin

Salivary glands

Spinal cord

Eye

Larynx

Esophagus

Joint (TMJ)

Rules:

IF (CC>0) OR (CD=1) OR (CE=1) OR (CH=1) OR (CH=2) THEN RTOG_late:skin=1

IF (CD=2) OR (CE=2) OR (CE=3) OR (CH=3) THEN RTOG_late:skin=2

IF (CD=3) OR (CE=4) THEN RTOG_late:skin=3

IF (CE=5) THEN RTOG_late:skin=4

IF (CI=1) OR (CJ=1) THEN RTOG_late:salivary_glands=1

IF (CI=2) OR (CJ=2) THEN RTOG_late:salivary_glands=2

IF (CI=3) OR (CJ=3) THEN RTOG_late:salivary_glands=3

IF (CJ>3) THEN RTOG_late:salivary_glands=4

IF (DH=1) OR (DH=2) THEN RTOG_late:spinal_cord=1

IF (DH=3) OR (DH=4) THEN RTOG_late:spinal_cord=2

IF (DD=1) THEN RTOG_late:eyes=1

IF (DD=2) OR (DD=3) OR (DG=1) THEN RTOG_late:eyes=2

IF (DB=2) OR (DD=4) OR (DG=3) THEN RTOG_late:eyes=3

IF (DB=3) OR (DD=5) OR (DG=4) THEN RTOG_late:eyes=4

IF ((CT=1) OR (CT=2)) AND (CV<2) THEN RTOG_late:larynx=1

IF ((CT=1) OR (CT=2) OR (CT=3)) AND (CV=2) THEN RTOG_late:larynx=2

IF (CT=4) OR (CV=3) OR (CV=4) THEN RTOG_late:larynx=3

IF (CU=4) THEN RTOG_late:larynx=4

IF (CR=1) AND (AW=0) AND (CQ<2) THEN RTOG_late:esophagus=1

IF (CR=1) AND (AW=1) AND (BA=1) AND (BJ=1) THEN RTOG_late:esophagus=1

IF (CR>1) AND (AW=1) AND ((BB=1) OR (BC=1) OR (BD=1)) AND (BJ=1) THEN RTOG_late:esophagus=2

IF (CR>1) AND (CQ>1) THEN RTOG_late:esophagus=2

IF (CR>1) AND (AW=1) AND ((BE=1) OR (BF=1)) AND (BJ=1) THEN RTOG_late:esophagus=3

IF (BF=1) AND (BJ=1) AND (CQ=3) THEN RTOG_late:esophagus=3

IF (CR=3) AND (W=1) AND (AL=1) THEN RTOG_late:esophagus=3

IF (CR=3) AND (CQ=3) THEN RTOG_late:esophagus=3

IF (CR=4) THEN RTOG_late:esophagus=4

IF (CO=1) THEN RTOG_late:joint_TMJ=1

IF (CO=1) AND (CM>1) THEN RTOG_late:joint_TMJ=2

IF (CO=2) THEN RTOG_late:joint_TMJ=2

IF (CO=2) AND (CM>2) THEN RTOG_late:joint_TMJ=3

IF (CO=3) THEN RTOG_late:joint_TMJ=3

IF (CO=3) AND (CM=4) THEN RTOG_late:joint_TMJ=4

HNC late toxicity data mapping rules – CTCAE v4.03

Mapped endpoints:

Ear pain

External ear inflammation

Hearing impaired

Tinnitus

Cataract

Corneal ulcer

Dry eye

Eye pain

Glaucoma

Retinopathy

Dry mouth

Dysphagia

Nausea

Salivary duct inflammation

Anorexia

Trismus

Hoarseness

Rules:

IF (DE=1) OR (DE=2) THEN CTCAE_4.03:ear_pain=1

IF (DE=2) AND ((DF=1) OR (DF=2)) THEN CTCAE_4.03:ear_pain=2

IF (DE=3) THEN CTCAE_4.03:ear_pain=2

IF (DE=3) AND (DF>0) THEN CTCAE_4.03:ear_pain=3

IF (DE=4) THEN CTCAE_4.03:ear_pain=3

IF (DD=1) THEN CTCAE_4.03:external_ear_inflammation=1

IF (DD=2) THEN CTCAE_4.03:external_ear_inflammation=2

IF (DD=3) THEN CTCAE_4.03:external_ear_inflammation=3

IF (DD=4) THEN CTCAE_4.03:external_ear_inflammation=4

IF (DG=1) OR (DG=2) THEN CTCAE_4.03:tinnitus=1

IF (DG=3) THEN CTCAE_4.03:tinnitus=2

IF (DG=4) THEN CTCAE_4.03:tinnitus=3

IF (DH=1) THEN CTCAE_4.03:hearing_impaired=1

IF (DH=2) THEN CTCAE_4.03:hearing_impaired=2

IF (DH=3) THEN CTCAE_4.03:hearing_impaired=3

IF (DH=4) THEN CTCAE_4.03:hearing_impaired=4

IF (DS=1) AND (DT=0) THEN CTCAE_4.03:cataract=1

IF (DS=1) AND (DT=1) THEN CTCAE_4.03:cataract=2

IF (DS=1) AND (DT=2) THEN CTCAE_4.03:cataract=3

IF (DS=1) AND (DT=3) THEN CTCAE_4.03:cataract=4

IF (DO=1) AND ((DT=1) OR (DW=1) OR (DW=2)) THEN CTCAE_4.03:corneal_ulcer=2

IF (DO=1) AND ((DT=2) OR (DW=3) OR (DW=4)) THEN CTCAE_4.03:corneal_ulcer=3

IF (DO=1) AND (DT=3) THEN CTCAE_4.03:corneal_ulcer=4

IF ((DV=1) OR (DV=2)) AND (DW=0) THEN CTCAE_4.03:dry_eye=1

IF ((DV=3) OR (DV=4)) AND (DW<2) THEN CTCAE_4.03:dry_eye=2

IF (DV>2) AND (DW<3) AND ((DT=1) OR (DT=2)) THEN CTCAE_4.03:dry_eye=3

IF ((DV=1) OR (DV=2)) AND ((DW=0) OR (DW=1)) THEN CTCAE_4.03:eye_pain=1

IF ((DV=2) OR (DV=3)) AND (DW=2) THEN CTCAE_4.03:eye_pain=2

IF (DV=3) AND (DW=3) THEN CTCAE_4.03:eye_pain=3

IF (DV=4) AND (DW>2) THEN CTCAE_4.03:eye_pain=3

IF (DR=1) AND (DT=0) THEN CTCAE_4.03:glaucoma=1

IF (DR=1) AND (DT=1) THEN CTCAE_4.03:glaucoma=2

IF (DR=1) AND (DT=2) THEN CTCAE_4.03:glaucoma=3

IF (DR=1) AND (DT=3) THEN CTCAE_4.03:glaucoma=4

IF (DP=1) AND (DT=0) THEN CTCAE_4.03:retinopathy=1

IF (DP=1) AND (DT=1) THEN CTCAE_4.03:retinopathy=2

IF (DP=1) AND (DT=2) THEN CTCAE_4.03:retinopathy=3

IF (DP=1) AND (DT=3) THEN CTCAE_4.03:retinopathy=4

IF (CH=1) OR (CI=1) OR (CI=2) THEN CTCAE_4.03:dry_mouth=1

IF (CH=2) AND ((CI=0) OR (CI=1)) AND (W=0) THEN CTCAE_4.03:dry_mouth=1

IF (CH>1) AND (CI>1) THEN CTCAE_4.03:dry_mouth=2

IF (CH=3) AND ((CI=0) OR (CI=1)) AND (W=0) THEN CTCAE_4.03:dry_mouth=2

IF (CH>1) AND (W=1) AND ((AI=1) OR (AJ=1)) AND ((X=1) OR (Y=1) OR (Z=1) OR (AB=1) OR (AC=1) OR (AD=1)) THEN CTCAE_4.03:dry_mouth=2

IF (CH=3) AND (CI=3) THEN CTCAE_4.03:dry_mouth=3

IF (CH>1) AND (CI>1) AND (W=1) AND ((AI=1) OR (AJ=1)) AND ((AE=1) OR (AF=1) OR (AG=1) OR (AH=1)) THEN CTCAE_4.03:dry_mouth=3

IF (CK=1) AND (W=0) THEN CTCAE_4.03:dysphagia=1

IF (CK=1) AND (W=1) AND (((Z=1) OR (AB=1) OR (AC=1) OR (AD=1)) AND (AL=1)) THEN CTCAE_4.03:dysphagia=2

IF ((AG=1) OR (AH=1)) AND ((AL=1) OR (CF>2) OR (CK=1)) THEN CTCAE_4.03:dysphagia=3

IF (W=1) AND (AL=1) AND (CK=1) AND ((AE=1) OR (AF=1)) THEN CTCAE_4.03:dysphagia=3

IF (CK>1) THEN CTCAE_4.03:dysphagia=4

IF (DZ=1) THEN CTCAE_4.03:nausea=1

IF (DZ=2) THEN CTCAE_4.03:nausea=2

IF (DZ=3) THEN CTCAE_4.03:nausea=3

IF (CI=1) THEN CTCAE_4.03:salivary_duct_inflammation=1

IF (CJ>0) THEN CTCAE_4.03:salivary_duct_inflammation=1

IF (CI=2) THEN CTCAE_4.03:salivary_duct_inflammation=2

IF (CJ=3) THEN CTCAE_4.03:salivary_duct_inflammation=2

IF ((CI=1) OR (CJ>1)) AND (W=1) AND ((AJ=1) OR (AK=1)) AND ((X=1) OR (Y=1) OR (Z=1) OR (AA=1)) THEN CTCAE_4.03:salivary_duct_inflammation=2

IF ((CI=1) OR (CJ>1)) AND (AQ=1) AND (AW=1) THEN CTCAE_4.03:salivary_duct_inflammation=2

IF ((CI=1) OR (CJ>1)) AND (N=1) AND (P=1) THEN CTCAE_4.03:salivary_duct_inflammation=2

IF (CI>1) AND (CJ=3) AND (W=1) AND ((AE=1) OR (AF=1) OR (AG=1) OR (AH=1)) AND ((AJ=1) OR (AK=1)) THEN CTCAE_4.03:salivary_duct_inflammation=3

IF (CI>1) AND (CJ=3) AND (N=2) AND (P=1) AND (W=1) AND ((AJ=1) OR (AK=1)) THEN CTCAE_4.03:salivary_duct_inflammation=3

IF (EB=1) THEN CTCAE_4.03:anorexia=1

IF (EB=2) THEN CTCAE_4.03:anorexia=2

IF (EB=3) THEN CTCAE_4.03:anorexia=3

IF (EB=4) THEN CTCAE_4.03:anorexia=4

IF (AN=1) THEN CTCAE_4.03:trismus=1

IF (AN=2) AND (W=0) THEN CTCAE_4.03:trismus=1

IF (AN=2) AND (W=1) AND (AN=0) THEN CTCAE_4.03:trismus=1

IF (AN=2) AND (W=1) AND ((AB=1) OR (AC=1) OR (AD=1)) AND (AN=1) THEN CTCAE_4.03:trismus=2

IF (AN=3) THEN CTCAE_4.03:trismus=2

IF (AN=3) AND (W=1) AND ((AE=1) OR (AF=1) OR (AG=1) OR (AH=1)) AND (AN=1) THEN CTCAE_4.03:trismus=3

IF (CT=1) THEN CTCAE_4.03:hoarseness=1

IF (CT=2) THEN CTCAE_4.03:hoarseness=2

IF (CT>2) THEN CTCAE_4.03:hoarseness=3

HNC late toxicity data mapping rules - LENT SOM(A)

Mapped endpoints:

Ear – Subjective (pain)

Ear – Subjective (tinnitus)

Ear - Subjective (hearing)

Ear - Objective (skin)

Ear – Management (pain)

Ear – Management (skin)

Ear – Management (hearing loss)

Mucosa – oral and pharyngeal - Subjective (pain)

Mucosa – oral and pharyngeal – Subjective (dysphagia)

Mucosa – oral and pharyngeal – Subjective (taste alteration)

Mucosa – oral and pharyngeal – Objective (weight)

Mucosa – oral and pharyngeal – Management (pain)

Mucosa – oral and pharyngeal – Management (ulcer)

Mucosa – oral and pharyngeal – Management (dysphagia)

Mucosa – oral and pharyngeal – Management (taste alteration)

Salivary gland – Subjective (xerostomia)

Salivary gland – Objective (saliva)

Salivary gland – Management (xerostomia)

Salivary gland – Global (LENT score)

Mandible – Subjective (pain)

Mandible – Subjective (mastication)

Mandible – Subjective (denture use)

Mandible – Subjective (trismus)

Mandible – Management (pain)

Mandible – Management (trismus and mastication)

Larynx – Subjective (voice/hoarseness)

Larynx – Subjective (breathing)

Skin/subcutaneous tissue – Subjective (scaliness/roughness)

Skin/subcutaneous tissue – Subjective (sensation)

Skin/subcutaneous tissue – Objective (edema)

Skin/subcutaneous tissue – Objective (alopecia)

Skin/subcutaneous tissue – Objective (pigmentation)

Skin/subcutaneous tissue – Objective (ulcer/necrosis)

Skin/subcutaneous tissue – Objective (telangiectasia)

Skin/subcutaneous tissue – Objective (fibrosis/scar)

Skin/subcutaneous tissue – Objective (atrophy/contraction

<u>Rules:</u>

LENT_SOM(A):mucosa_oral_&_pharyngeal(Objective_weight)=D

IF ((BF=1) OR (CI=4)) AND (AQ=1) AND (AS=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_ulcer)=2

IF ((BF=1) OR (CI=4)) AND (AQ=1) AND (AT=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_ulcer)=3

IF ((BF=1) OR (CI=4)) AND (AQ=1) AND (AR=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_ulcer)=4

IF (CF=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_pain)=1

IF (CF=2) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_pain)=2

IF (CF=3) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_pain)=3

IF (CF=4) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_pain)=4

IF (CG=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_pain)=1

IF (CG=2) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_pain)=2

IF (CG=3) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_pain)=3

IF (CG=4) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_pain)=4

IF (CH=1) THEN LENT_SOM(A):salivary_gland(Subjective_xerostomia)=1

IF (CH=2) THEN LENT_SOM(A):salivary_gland(Subjective_xerostomia)=2

IF (CH=3) THEN LENT_SOM(A):salivary_gland(Subjective_xerostomia)=3

IF (CH=3) AND (N>0) AND (P=1) THEN
LENT_SOM(A):salivary_gland(Subjective_xerostomia)=4

IF (CH=3) AND (((AD=1) OR (AE=1) OR (AF=1) OR (AG=1) OR ((AH=1)) AND (O=1) OR (AJ=1))) THEN LENT_SOM(A):salivary_gland(Subjective_xerostomia)=4

IF (CH=0) AND ((CI=1) OR (CI=2)) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=1

IF ((CH=1) OR (CH=2)) AND (CI=0) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=1

IF (CH=2) AND ((CI=1) OR (CI=2)) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=2

IF (CH=3) AND ((CI=1) OR (CI=2)) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=3

IF (CH=3) AND ((CI=3) OR (CI=4)) THEN LENT_SOM(A):salivary_gland(Objective_saliva)=4

IF (CH>O) AND (X=1) AND (Y=0) AND (Z=0) THEN
LENT_SOM(A):salivary_gland(Management_xerostomia)=2

IF (CH>O) AND (Y=1) AND (Z=0) THEN
LENT_SOM(A):salivary_gland(Management_xerostomia)=3

IF (CH>O) AND (Z=1) THEN LENT_SOM(A):salivary_gland(Management_xerostomia)=4

IF (CJ=1) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_taste_alteration)=1

IF (CJ=2) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_taste_alteration)=2

IF (CJ=3) THEN LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_taste_alteration)=3

IF (CJ>1) AND (W=1) AND ((AA=1) OR (AB=1)) AND (AK=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_taste_alteration)=1

IF (CJ>1) AND (W=1) AND (AC=1) AND (AK=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_taste_alteration)=2

IF (CJ>1) AND (W=1) AND ((AD=1) OR (AE=1)) AND (AK=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_taste_alteration)=3

IF (CJ>1) AND (W=1) AND ((AF=1) OR (AG=1) OR (AH=1)) AND (AK=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_taste_alteration)=4

IF (CK>0) AND (W=1) AND (AL=1) AND ((AC=1) OR (AD=1)) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=1

IF (CK>0) AND (W=1) AND (AE=1) AND (AL=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=2

IF (CK>0) AND (W=1) AND (AF=1) AND (AL=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=3

IF (CK>0) AND (W=1) AND ((AG=1) OR (AH=1)) AND (AL=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=4

IF (CK>1) AND (W=1) AND (AL=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Subjective_dysphagia)=4

IF (CK>1) AND (Z=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_dysphagia)=1

IF (CK>1) AND (CF>1) AND ((CG=1) OR (CG=2)) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_dysphagia)=2

IF (CK>1) AND (CF>1) AND (CG=3) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_dysphagia)=3

IF (CK>1) AND (CF>1) AND (AH=1) THEN
LENT_SOM(A):mucosa_oral_&_pharyngeal(Management_dysphagia)=4

LENT_SOM(A):salivary_gland_global=(LENT_SOM(A):salivary_gland(Subjective_xerostomia)
+ LENT_SOM(A):salivary_gland(Objective_saliva)+
LENT_SOM(A):salivary_gland(Management_xerostomia))/3

IF (CM=1) THEN LENT_SOM(A):mandible(Subjective_pain)=1

IF (CM=2) THEN LENT_SOM(A):mandible(Subjective_pain)=2

IF (CM=3) THEN LENT_SOM(A):mandible(Subjective_pain)=3

IF (CM=4) THEN LENT_SOM(A):mandible(Subjective_pain)=4

IF (CM>1) AND (((AB=1) OR (AC=1) OR (AD=1)) AND (AO=1))THEN
LENT_SOM(A):mandible(Subjective_mastication)=2

IF (CM>1) AND (((AE=1) OR (AF=1) OR (AG=1) OR (AH=1)) AND (AO=1))THEN
LENT_SOM(A):mandible(Subjective_mastication)=3

IF (CM>0) AND (CN=1) THEN LENT_SOM(A):mandible(Management_pain)=1

IF (CM>1) AND (CN=2) THEN LENT_SOM(A):mandible(Management_pain)=2

IF (CM>2) AND (CN=3) THEN LENT_SOM(A):mandible(Management_pain)=3

IF (CM>2) AND (CN=4) THEN LENT_SOM(A):mandible(Management_pain)=4

IF (CO=2) THEN LENT_SOM(A):mandible(Subjective_denture_use)=2

IF (CO=3) THEN LENT_SOM(A):mandible(Subjective_denture_use)=3

IF (CP=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=1

IF (CP=2) THEN LENT_SOM(A):mandible(Subjective_trismus)=1

IF (CP=2) AND (W=1) AND (AN=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=2

IF (CP=2) AND (W=1) AND ((AB=1) OR (AC=1) OR (AD=1)) AND (AN=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=2

IF (CP=2) AND (W=1) AND ((AE=1) OR (AF=1) OR (AG=1) OR (AH=1)) AND (AN=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=3

IF (CP=3) AND (W=1) AND (AN=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=3

IF (CP=3) AND (W=1) AND ((AE=1) OR (AF=1) OR (AG=1) OR (AH=1)) AND (AN=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=4

IF (CP=4) AND (W=1) AND (AN=1) THEN LENT_SOM(A):mandible(Subjective_trismus)=4

IF (CP>1) AND (W=1) AND ((AC=1) OR (AD=1)) AND (AN=1) THEN LENT_SOM(A):mandible(Management_trismus&mastication)=2

IF (CP>1) AND (W=1) AND ((((AE=1) OR (AF=1)) AND (AN=1)) OR ((AQ=1) AND (BQ=1) AND ((AT=1) OR (BY=1)))) THEN LENT_SOM(A):mandible(Management_trismus&mastication)=3

IF (CP>1) AND (W=1) AND ((AG=1) OR (AH=1)) AND (AN=1) THEN LENT_SOM(A):mandible(Management_trismus&mastication)=4

IF (CT=1) THEN LENT_SOM(A):larynx(Subjective_voice/hoarseness)=1

IF (CT=2) THEN LENT_SOM(A):larynx(Subjective_voice/hoarseness)=2

IF (CT=3) THEN LENT_SOM(A):larynx(Subjective_voice/hoarseness)=3

IF (CT=4) THEN LENT_SOM(A):larynx(Subjective_voice/hoarseness)=4

IF (CT>0) AND (CU=1) THEN LENT_SOM(A):larynx(Subjective_breathing)=1

IF (CT>0) AND (CU=2) THEN LENT_SOM(A):larynx(Subjective_breathing)=2

IF (CT>0) AND (CU=3) THEN LENT_SOM(A):larynx(Subjective_breathing)=3

IF (CT>0) AND (CU=4) THEN LENT_SOM(A):larynx(Subjective_breathing)=4

IF (CV=1) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Subjective_scaliness/roughness)=1

IF (CV=2) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Subjective_scaliness/roughness)=2

IF (CV=3) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Subjective_scaliness/roughness)=3

IF (CW=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_sensation)=1

IF (CW=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_sensation)=2

IF (CW=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Subjective_sensation)=3

IF (CW=3) AND (N>0) AND (Q=1) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Subjective_sensation)=4

IF (CX=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_edema)=1

IF (CX=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_edema)=2

IF (CX=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_edema)=3

IF (CX=4) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_edema)=4

IF (CY=1) OR (CY=2) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_pigmentation_change)=1

IF (CY=3) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_pigmentation_change)=2

IF (CZ=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_telangiectasia)=1

IF (CZ=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_telangiectasia)=2

IF (CZ=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_telangiectasia)=3

IF (DA=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_fibrosis_scar)=1

IF (DA=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_fibrosis_scar)=2

IF (DA=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_fibrosis_scar)=3

IF (DA=4) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_fibrosis_scar)=4

IF (DB=1) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_atrophy_contraction)=1

IF (DB=2) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_atrophy_contraction)=2

IF (DB=3) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_atrophy_contraction)=3

IF (DB=4) THEN
LENT_SOM(A):skin/subcutaneous_tissue(Objective_atrophy_contraction)=4

IF (DC=1) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_alopecia_scalp)=1

IF (DC=2) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_alopecia_scalp)=2

IF (DC=3) THEN LENT_SOM(A):skin/subcutaneous_tissue(Objective_alopecia_scalp)=3

IF (DD=1) THEN LENT_SOM(A):ear(Objective_skin)=1

IF (DD=2) THEN LENT_SOM(A):ear(Objective_skin)=2

IF (DD=3) THEN LENT_SOM(A):ear(Objective_skin)=3

IF (DD=4) THEN LENT_SOM(A):ear(Objective_skin)=4

IF ((DD=1) OR (DD=2)) AND (BQ=1) AND (CC=1) THEN
LENT_SOM(A):ear(Management_skin)=1

IF ((DD=2) OR (DD=3)) AND (((AQ=1) AND (AT=1)) OR ((BQ=1) AND (CD=1))) THEN
LENT_SOM(A):ear(Management_skin)=2

IF (DD=4) AND (AQ=1) AND (AR=1) AND (BM=1) THEN
LENT_SOM(A):ear(Management_skin)=4

IF (DE=1) THEN LENT_SOM(A):ear(Subjective_pain)=1

IF (DE=2) THEN LENT_SOM(A):ear(Subjective_pain)=2

IF (DE=3) THEN LENT_SOM(A):ear(Subjective_pain)=3

IF (DE=4) THEN LENT_SOM(A):ear(Subjective_pain)=4

IF ((DE=1) OR (DE=2) OR (DE=3)) AND (DF=1) THEN
LENT_SOM(A):ear(Management_pain)=1

IF (DE>1) AND (DF=2) THEN LENT_SOM(A):ear(Management_pain)=2

IF (DE>2) AND (DF=3) THEN LENT_SOM(A):ear(Management_pain)=3

IF (DE>2) AND (DF=4) THEN LENT_SOM(A):ear(Management_pain)=4

IF (DG=1) THEN LENT_SOM(A):ear(Subjective_tinnitus)=1

IF (DG=2) THEN LENT_SOM(A):ear(Subjective_tinnitus)=2

IF (DG=3) THEN LENT_SOM(A):ear(Subjective_tinnitus)=3

IF (DG=4) THEN LENT_SOM(A):ear(Subjective_tinnitus)=4

IF (DH=1) THEN LENT_SOM(A):ear(Subjective_hearing)=1

IF (DH=2) THEN LENT_SOM(A):ear(Subjective_hearing)=2

IF (DH=3) THEN LENT_SOM(A):ear(Subjective_hearing)=3

IF (DH=4) THEN LENT_SOM(A):ear(Subjective_hearing)=4

IF (DH>1) and (BA=1) THEN LENT_SOM(A):ear(Management_hearing_loss)=3

HNC late toxicity data mapping rules – EORTC QLQ – H&N 35

Mapped endpoints:

31. Have you had pain in your mouth?

32. Have you had pain in your jaw?

40. Have you had problems opening your mouth wide?

41. Have you had a dry mouth?

42. Have you had sticky saliva?

44. Have you had problems with your sense of taste?

46. Have you been hoarse?

Rules:

IF (CF=0) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_mouth?=1

IF (CF=1) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_mouth?=2

IF (CF=2) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_mouth?=3

IF (CF>2) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_mouth?=4

IF (CH=0) THEN EORTC_QLQ-H&N35:have_you_had_a_dry_mouth?=1

IF (CH=1) THEN EORTC_QLQ-H&N35:have_you_had_a_dry_mouth?=2

IF (CH=2) THEN EORTC_QLQ-H&N35:have_you_had_a_dry_mouth?=3

IF (CH=3) THEN EORTC_QLQ-H&N35:have_you_had_a_dry_mouth?=4

IF (CI=0) THEN EORTC_QLQ-H&N35:have_you_had_sticky_saliva?=1

IF (CI=1) THEN EORTC_QLQ-H&N35:have_you_had_sticky_saliva?=2

IF (CI=2) THEN EORTC_QLQ-H&N35:have_you_had_sticky_saliva?=3

IF (CI=3) THEN EORTC_QLQ-H&N35:have_you_had_sticky_saliva?=4

IF (CJ=0) THEN EORTC_QLQ-H&N35:have_you_had_problems_with_your_sense_of_taste?=1

IF (CJ=1) THEN EORTC_QLQ-H&N35:have_you_had_problems_with_your_sense_of_taste?=2

IF (CJ=2) THEN EORTC_QLQ-H&N35:have_you_had_problems_with_your_sense_of_taste?=3

IF (CJ=3) THEN EORTC_QLQ-H&N35:have_you_had_problems_with_your_sense_of_taste?=4

IF (CM=0) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_jaw?=1

IF (CM=1) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_jaw?=2

IF (CM=2) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_jaw?=3

IF (CM>2) THEN EORTC_QLQ-H&N35:have_you_had_pain_in_your_jaw?=4

IF (CP=0) THEN EORTC_QLQ-
H&N35:have_you_had_problems_opening_your_mouth_wide?=1

IF (CP=1) THEN EORTC_QLQ-
H&N35:have_you_had_problems_opening_your_mouth_wide?=2

IF (CP=2) THEN EORTC_QLQ-
H&N35:have_you_had_problems_opening_your_mouth_wide?=3

IF (CP>2) THEN EORTC_QLQ-
H&N35:have_you_had_problems_opening_your_mouth_wide?=4

IF (CT=0) THEN EORTC_QLQ-H&N35:have_you_been_hoarse?=1

IF (CT=1) THEN EORTC_QLQ-H&N35:have_you_been_hoarse?=2

IF (CT=2) THEN EORTC_QLQ-H&N35:have_you_been_hoarse?=3

IF (CT>2) THEN EORTC_QLQ-H&N35:have_you_been_hoarse?=4

HNC late toxicity data mapping rules – RTOG

Mapped endpoints:

Skin

Salivary glands

Spinal cord

Eye

Larynx

Esophagus

Joint (TMJ)

Rules:

IF (CY>0) OR (CZ=1) OR (DB=1) OR (DC=1) OR (DC=2) THEN RTOG_late:skin=1

IF (CZ=2) OR (DB=2) OR (DC=3) THEN RTOG_late:skin=2

IF (CZ=3) OR (DB=4) THEN RTOG_late:skin=3

IF (DB=5) THEN RTOG_late:skin=4

IF (CH=1) OR (CI=1) OR (CJ=1) OR (CJ=2) THEN RTOG_late:salivary_glands=1

IF (CH=2) OR (CI=2) OR (CJ=3) THEN RTOG_late:salivary_glands=2

NO DESCRIPTION FOR RTOG GRADE 3

IF (CH=3) AND (CH>1) AND (CI>1) AND (N=2) AND (P=1) AND (W=1) AND ((AI=1) OR (AJ=1) OR (AK=1)) AND ((AF=1) OR (AG=1) OR (AH=1)) THEN RTOG_late:salivary_glands=4

IF (DI=1) OR (DI=2) THEN RTOG_late:spinal_cord=1

IF (DI=3) OR (DI=4) THEN RTOG_late:spinal_cord=2

IF (DS=1) AND (DT=0) THEN RTOG_late:eyes=1

IF (DO=1) AND (DT=0) THEN RTOG_late:eyes=1

IF (DS=1) AND (DT>0) THEN RTOG_late:eyes=2

IF (DO=1) AND (DT=1) THEN RTOG_late:eyes=2

IF (DP=1) AND ((DT=0) OR (DT=1)) THEN RTOG_late:eyes=2

IF (DR=1) AND ((DT=0) OR (DT=1)) THEN RTOG_late:eyes=2

IF (DO=1) AND (DT>1) THEN RTOG_late:eyes=3

IF (DP=1) AND (DT>1) THEN RTOG_late:eyes=3

IF (DR=1) AND (DT>1) THEN RTOG_late:eyes=3

IF (DT=3) THEN RTOG_late:eyes=4

IF ((CT=1) OR (CT=2)) AND (CU<2) THEN RTOG_late:larynx=1

IF ((CT=1) OR (CT=2) OR (CT=3)) AND (CU=2) THEN RTOG_late:larynx=2

IF (CT=4) OR (CU=3) OR (CU=4) THEN RTOG_late:larynx=3

IF (CT=5) THEN RTOG_late:larynx=4

IF (CK=1) AND (W=0) AND (CF<2) THEN RTOG_late:esophagus=1

IF (CK=1) AND (W=1) AND (AB=1) AND (AL=1) THEN RTOG_late:esophagus=1

IF (CK=1) AND (W=1) AND ((AC=1) OR (AD=1) OR (AE=1)) AND (AL=1) THEN RTOG_late:esophagus=2

IF (CK=1) AND (W=1) AND (((AF=1) OR (AG=1) OR (AH=1)) AND (AL=1)) THEN RTOG_late:esophagus=3

IF ((AG=1) OR (AH=1)) AND ((AL=1) OR (CF>0) OR (CK=1)) THEN RTOG_late:esophagus=3

IF (CK=2) AND (W=1) AND (AL=1) THEN RTOG_late:esophagus=3

IF (CK=3) THEN RTOG_late:esophagus=4

IF (CP=1) THEN RTOG_late:joint_TMJ=1

IF (CP=1) AND (CM>1) AND (CN>0) THEN RTOG_late:joint_TMJ=2

IF (CP=2) THEN RTOG_late:joint_TMJ=2

IF (CP=2) AND (CM>2) AND (CN>1) THEN RTOG_late:joint_TMJ=3

IF (CP=3) THEN RTOG_late:joint_TMJ=3

IF (CP=4) OR (CN=4) THEN RTOG_late:joint_TMJ=4

VoxTox Head & Neck

PROTOCOL FOR INTER-OBSEVER CONTOURING: PAROTIDS, SUBMANDIBULAR GLANDS, SUPERIOR & MIDDLE PHARYNGEAL CONSTRICTOR MUSCLES, SUPRAGLOTTIC LARYNX & ORAL CAVITY ON DOWNSAMPLED kVCT SCANS – 'swallowing OARS'

Thank you very much for participating in this part of the study. The objective is to quantify agreement between observers when contouring swallowing OARs on downsampled kVCTs used in VoxTox. We would like you to contour the structures specified for 2 patients – this should take about one to one and half hours of your time.

PROCEDURE

1. Opening Prosoma Research

- Open Prosoma, the database login window will appear.
- Under “Database Location” select “Research” from the drop-down menu next to Phonebook-Entry:
- Type in username and password.
- Click OK.
- Click “Load Data from Database”.

2. Opening the first patient

- From the list of patients on the left of the screen select “VT1_H_78386K1L”
- There are 8 simulation plans under this Dicom Series. The bottom 4 (numbered 7-10) have been prepared for each of you to contour. They have been labelled “kVCT InterOV [your initials]”.
- Each consultant is allocated to the following simulation plan; RJB (sim plan 7), GCB (8), RJ1 (9), KF (10).
- Please open your simulation plan to contour.

3. Contouring the first scan

- The scan has been preset to x2.8 zoom, and windowed on standard soft tissue settings. In contrast to studies that you have contributed to on MVCT images, please feel free to change zoom and window settings as preferred to optimise your contouring – suggestions below.
    - The window settings as saved are optimised for the parotids and SMGs
    - Wider window setting with a lower centre may help to better visualise the superior aspect of the epiglottis when drawing the SGL.
- Click on “VOI editor”.
- Under “VOI Definition” there will be 8 volumes labelled; “right parotid - INITIALS”, “left parotid - INITIALS”, “right SMG - INITIALS”, “left SMG - INITIALS”, “SPC - INITIALS” (superior pharyngeal constrictors), “MPC - INITIALS” (middle pharyngeal constrictors), “SGL – INITIALS” (supraglottic larynx), “OC – INITIALS” (oral cavity). The labels and colours are preset – please don’t change these.

- Under “VOI Creation” select your preferred contouring tool (under the “Contours” tab) for each structure. Use whichever contouring tool that you would use clinically for each structure.[1] Then click “Draw and Edit”.
- Attached with this protocol are 2 consensus contouring atlases – Brouwer 2015, and Christianen 2011. Please follow the guidelines in Brouwer for the parotids, SMGs and oral cavity, where necessary. Please use the Christianen atlas for segmenting the pharyngeal constrictors and supraglottic larynx.
- Where the borders of structures are difficult to identify (for example with the pharyngeal constrictors), use surrogate anatomy to guide contours, and use the interpolate function in Prosoma.
- For larger structures (eg parotids and oral cavity) interpolation may be used, but please try to segment at least every second slice.

4. Saving your contours

- When you have completed your contouring on this scan click on “file i/o” in the top left hand corner of the screen.
- Under “Simulation Plan” on the right of the screen click on “Save to DB”.
- Please write: ‘kVCT InterOV – [your INITIALS] – complete’ as a comment and click “OK”.
- Thank you for contouring you first scan.

5. Contouring the next scans.

- Please repeat the same procedure for patient VT1_H_81AB51K1.
- The simulation plans for each consultant are as follows; sim plan 3 RJB, 4 GCB, 6 RJ1, 7 KF.
- Once complete please save as for the first scan.

This completes your contouring. Thank you very much for taking the time to participate. Please email me to let me know when you have finished your contouring – I’m happy to disseminate the data once I have analysed the contours.

[1] Please ensure that you are consistent with your use of contouring tools between cases.

VoxTox Head & Neck

PROTOCOL FOR INTER-OBSEVER CONTOURING OF PAROTIDS, SUBMANDIBULAR GLANDS AND SUPERIOR PHARYNGEAL CONSTRICTOR MUSCLES

Thank you very much for participating in this part of the study. The objective is to compare the accuracy of the automated contouring algorithm with segmentation done by expert observers. We would like you to contour the structures specified on 2 daily IG MVCT's for 2 patients (ie 4 scans in total). Based on our experience of MVCT contouring, we would anticipate this taking one to one and half hours of your time.

PROCEDURE

1. Opening Prosoma Research

- Open Prosoma, the database login window will appear.
- Under “Database Location” select “Research” from the drop-down menu next to Phonebook-Entry:
- Type in username and password.
- Click OK.
- Click “Load Data from Database”.

2. Opening the first patient

- From the list of patients on the left of the screen select “VT1_H_2E54B1K1”
- Scroll down to “Dicom Series 31 CT HFS”. This is the MVCT for day 1 of this patient’s treatment.
- There are 7 simulation plans under this Dicom Series. The bottom 5 have been prepared for each of you to contour. They have been labelled “consultant name: Salivary Gland & PC Inter-observer contouring study”.
- Each consultant is allocated to the following simulation plan; TA (sim plan 101), GCB (102), RJB (103), NGB (104), SJJ (105).
- Please open your simulation plan to contour.

3. Contouring the first scan

- The scan has been preset to x1.8 zoom and windowing (width 637 HU & centre 32HU). The windowing has been optimised for structure identification and to ensure consistency between observers. Please keep to these presets whilst contouring.
- Click on “VOI editor”.
- Under “VOI Definition” there will be 5 volumes labelled; “right parotid” in sky blue, “left parotid” in yellow, “right SMG” in light green, “left SMG” in royal blue, “pharyngeal constrictors” in red.
- Under “VOI Creation” select your preferred contouring tool (under the “Contours” tab) for each structure. Use whichever contouring tool that you would use clinically for each structure.[1] Then click “Draw and Edit”.
- On the MVCT’s, structures are more easily identified on some slices than others. Start on an easier slice and proceed.
- Please contour each slice upon which you are confident that you can identify the structure of interest.

[1] Please ensure that you are consistent with your use of contouring tools between cases.

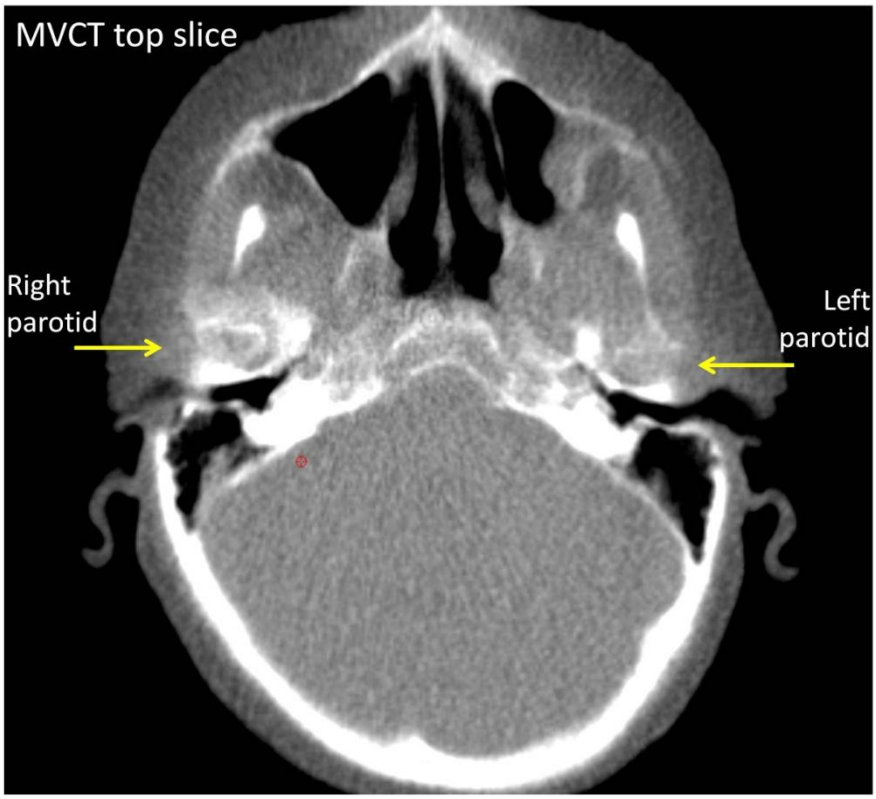
- Consider using interpolation between these slices.
- Please aim to contour at least every 3rd slice.
- Where structures cannot be identified at all (for example with the pharyngeal constrictors), use surrogate anatomy to guide contours, and use the interpolate function in Prosoma.

4. Parotid Glands

- Please contour the parotid glands based on the RTOG atlas of Gregoire and colleagues, see link below:

https://www.rtog.org/CoreLab/ContouringAtlases/HNAtlases.aspx

- The protocol for VoxTox image guidance MVCT's states that the superior limit of the scan should ensure the whole parotid is included. In practice this is often quite tight, and the superior border of the gland may be the top slice of the scan. An example is shown below;



5. Submandibular glands

- Please follow the same procedure for the submandibular glands, using the RTOG atlas as a guide.
- Both submandibular glands will be well within the field of view of the scans.

- Both patients in this study have their submandibular glands in-situ (neither had undergone neck dissection prior to radiotherapy).

6. Pharyngeal constrictors.

- In VoxTox, manual segmentation of the pharyngeal constrictor muscles are based on an atlas published by the Italian Radiation Oncology Group – AIRO (Alterio et al, Radiotherapy and Oncology, 2014, paper attached).
- This paper has two approaches for contouring the superior pharyngeal constrictors (PSCM) atlases, which they compare. The atlases for these techniques are on the third page of the paper, (page 329). For the purposes of this exercise, please use the red contours ('optimised' PSCM) as a guide.

7. Saving your contours

- When you have completed your contouring on this scan click on "file i/o" in the top left hand corner of the screen.
- Under "Simulation Plan" on the right of the screen click on "Save to DB".
- Please write: 'your name'_IOV study contours as a comment and click "OK".
- Thank you for contouring you first scan.

8. Contouring 3 more scans.

- Please repeat the same procedure on the day 16 scan of the same patient (VT1_H_2E54B1K1). This is Dicom Series 17 CT HFS.
- The simulation plans for each consultant are as follows; sim plan 106 TA, 107 GCB, 108 RJB, 109 NGB, 110 SJJ.
- Once complete please save as for the first scan. This completes the first patient.
- The zoom is the same for the other patient. Windowing is very similar and has been optimised to view all the relevant structures as well as possible.
- The second patient is VT1_H_623B02K1.
    - Day 1 is Dicom Series 28. Simulation plans for each consultant are as follows; 132 TA, 133 GCB, 134 RJB, 135 NGB, 136 SJJ.
    - Day 16 is Dicom Series 13. Simulation plans for each consultant are as follows; 137 TA, 138 GCB, 139 RJB, 140 NGB, 141 SJJ.

This completes your contouring. Thank you very much for taking the time to participate. Please email me to let me know when you have finished your contouring; we hope to have results of Inter-observer variability within a few weeks of everyone submitting their contours.

A5 Full details of univariate & multivariate NTCP model building and validation (Chapter 7)

FULL RESULTS OF UNIVARIATE ANALYSIS, VARIABLE SELECTION, MODEL BUILDING AND VALIDATION – PLANNED AND DELIVERED DOSE NTCP MODELS

Gr2+ CTCAE Xerostomia

Univariate associations (regression)

Variable	Nature	Wald	P value	OR	95% CI
Gender (male)	binary	2.52	0.112	0.58	0.260 – 1.152
Age	cont	0.47	0.492	0.99	0.964 – 1.018
Baseline wt	cont	2.69	0.101	0.986	0.969 – 1.003
Weight loss (kg)	cont	3.82	0.051	1.064	1.000 – 1.131
Weight loss (%)	**cont**	**5.488**	**0.019**	**1.073**	**1.012 – 1.137**
T stage	ordinal	0.004	0.947	1.008	0.792 – 1.283
Bin T stage	binary	1.173	0.279	1.385	0.768 – 2.495
Bin N stage	**binary**	**9.838**	**0.002**	**2.706**	**1.453 – 5.040**
cetuximab	nominal	0.883	0.347	1.790	0.531 – 6.030
cisplatin	**nominal**	**10.931**	**0.001**	**2.992**	**1.562 – 5.730**
SACT Y/N	**binary**	**10.205**	**0.001**	**2.826**	**1.494 – 5.344**
Cisplatin Y/N	**binary**	**10.370**	**0.001**	**2.705**	**1.476 – 4.957**
Neck diss	binary	0.047	0.828	0.937	0.521 – 1.686
parotidectomy	**binary**	**3.910**	**0.048**	**0.125**	**0.016 – 0.982**
SMGectomy	ordinal	0.045	0.833	0.938	0.516 – 1.704
Smoker (PY)	cont	0.033	0.855	0.999	0.984 – 1.013
Alcohol (total units)	cont	0.308	0.579	1.002	0.996 – 1.007
IPG_DP	**cont**	**9.599**	**0.002**	**1.053**	**1.019 – 1.088**
CPG_DP	**Cont**	**28.809**	**0.000**	**1.083**	**1.052 – 1.116**
ISMG_DP	**Cont**	**4.118**	**0.042**	**1.050**	**1.002 – 1.101**
CSMG_DP	**Cont**	**21.123**	**0.000**	**1.056**	**1.032 – 1.081**
SPC_DP	**Cont**	**17.326**	**0.000**	**1.100**	**1.052 – 1.151**
MPC_DP	**Cont**	**11.949**	**0.001**	**1.083**	**1.035 – 1.132**
OC_DP	**Cont**	**15.787**	**0.000**	**1.073**	**1.036 – 1.111**
SGL_DP	**Cont**	**12.464**	**0.000**	**1.078**	**1.034 – 1.124**
IPG_DA	**cont**	**9.892**	**0.002**	**1.053**	**1.020 – 1.087**
CPG_DA	**Cont**	**30.815**	**0.000**	**1.084**	**1.053 – 1.115**
ISMG_DA	**Cont**	**4.180**	**0.041**	**1.050**	**1.002 – 1.100**
CSMG_DA	**Cont**	**21.089**	**0.000**	**1.055**	**1.031 – 1.080**
SPC_DA	**Cont**	**17.449**	**0.000**	**1.099**	**1.051 – 1.149**
MPC_DA	**Cont**	**12.196**	**0.000**	**1.078**	**1.034 – 1.125**
OC_DA	**Cont**	**16.304**	**0.000**	**1.073**	**1.037 – 1.111**
SGL_DA	**Cont**	**12.673**	**0.000**	**1.079**	**1.035 – 1.125**
Baseline Xer	**ordinal**	**6.466**	**0.011**	**1.976**	**1.169 – 3.342**

Co-linearity diagnostics

Test	Tolerance	VIF
Bin N vs Cisplatin	0.857	1.167
Bin N vs parotidectomy	0.977	1.024
Bin N vs Bline_Xer	0.999	1.001
Bin N vs IPG_DP	0.863	1.158
Bin N vs CPG_DP	0.862	1.160
Bin N vs ISMG_DP	0.852	1.174
Bin N vs CSMG_DP	0.821	1.217
Bin N vs SPC_DP	0.779	1.284
Bin N vs MPC_DP	0.823	1.215
Bin N vs OC_DP	0.785	1.274
Bin N vs SGL_DP	0.818	1.223
Cisplatin vs parotidectomy	0.887	1.127
Cisplatin vs Bline_Xer	0.981	1.020
Cisplatin vs IPG_DP	0.938	1.066
Cisplatin vs CPG_DP	0.711	1.406
Cisplatin vs ISMG_DP	0.884	1.131
Cisplatin vs CSMG_DP	0.672	1.487
Cisplatin vs SPC_DP	0.739	1.353
Cisplatin vs MPC_DP	0.822	1.216
Cisplatin vs OC_DP	0.810	1.234
Cisplatin vs SGL_DP	0.829	1.206
parotidectomy vs Bline_Xer	0.975	1.026
parotidectomy vs IPG_DP	0.987	1.013
parotidectomy vs CPG_DP	0.867	1.153
parotidectomy vs ISMG_DP	0.999	1.001
parotidectomy vs CSMG_DP	0.813	1.230
Parotidectomy vs SPC_DP	0.935	1.069
parotidectomy vs MPC_DP	0.923	1.083
parotidectomy vs OC_DP	0.943	1.060
parotidectomy vs SGL_DP	0.889	1.124
Bline_Xer vs IPG_DP	0.998	1.002
Bline_Xer vs CPG_DP	0.978	1.022
Bline_Xer vs ISMG_DP	0.992	1.008
Bline_Xer vs CSMG_DP	0.994	1.006
Bline_Xer vs SPC_DP	1	1
Bline_Xer vs MPC_DP	0.998	1.002
Bline_Xer vs OC_DP	0.997	1.003
Bline_Xer vs SGL_DP	0.997	1.003
IPG_DP vs CPG_DP	0.844	1.185
IPG_DP vs ISMG_DP	0.604	1.656
IPG_DP vs CSMG_DP	0.928	1.077
IPG_DP vs SPC_DP	0.603	1.659
IPG_DP vs MPC_DP	0.865	1.156

IPG_DP vs OC_DP	0.680	1.471
IPG_DP vs SGL_DP	0.905	1.105
CPG_DP vs ISMG_DP	0.770	1.299
CPG_DP vs CSMG_DP	0.282	3.541
CPG_DP vs SPC_DP	0.535	1.868
CPG_DP vs MPC_DP	0.676	1.478
CPG_DP vs OC_DP	0.692	1.444
CPG_DP vs SGL_DP	0.699	1.430
ISMG_DP vs CSMG_DP	0.611	1.637
ISMG_DP vs SPC_DP	0.366	2.731
ISMG_DP vs MPC_DP	0.340	2.945
ISMG_DP vs OC_DP	0.483	2.071
ISMG_DP vs SGL_DP	0.465	2.149
CSMG_DP vs SPC_DP	0.460	2.173
CSMG_DP vs MPC_DP	0.356	2.808
CSMG_DP vs OC_DP	0.648	1.543
CSMG_DP vs SGL_DP	0.393	2.542
SPC_DP vs MPC_DP	0.416	2.403
SPC_DP vs OC_DP	0.214	4.673
SPC_DP vs SGL_DP	0.508	1.967
MPC_DP vs OC_DP	0.613	1.631
MPC_DP vs SGL_DP	**0.092**	**10.821**
OC_DP vs SGL_DP	0.630	1.587

Parameters put into model for LR based forward selection:

- Weight loss (%)
- Bin N
- Cisplatin Y/N
- Parotidectomy
- Bline_Xer
- IPG_DP
- CPG_DP
- ISMG_DP
- CSMG_DP
- SPC_DP
- OC_DP
- SGL_DP

NB. MPC_DP omitted due to colinearity with SGL_DP. (SGl_DP UVA associations stronger)

Model performance with this procedure:

- Nagelkerke R2 = 0.297
- -2LL = 146.460
- HL: chi-squared = 5.773, p = 0.673

Significance of adding each parameter:

- **CPG_DA. P = 0.000 (Step 1)**
- **Bline_Xer p = 0.005 (Step 2)**
- Weight loss (%) = 0.272
- Cisplatin Y/N p = 0.501
- Parotidectomy p =0.623
- IPG p =0.342
- ISMG p=0.543
- CSMG p=0.263
- SPC p=0.140
- OC p=0.279
- SGL p=0.438
- Bin N p=0.668

Therefore stopped at step 2, with CPG and Bline_Xer as model parameters

Model generation repeated for planned dose:

- Only these 2 parameters
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 1.263(Bline_Xer) + 0.097(CPG_DP) -3.546

Overall:

- Nagel R2 = 0.316
- -2LL = 210.493

Discrimination:

- ROC AUC = 0.787 (SE = 0.033, 95% CI 0.723 – 0.851)
- Discrimination slope =0.241 (SE = 0.031)

Calibration:

- HL: chi-squared = 2.831, p = 0.945
- Cal slope = 1.0214

- Intercept = -0.0092
- Brier score = 0.0028

Model generation repeated for delivered dose:

- Only these CPG_DA and Bline_Xer
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 1.307(Bline_Xer) + 0.098(CPG_DA) -3.688

Overall:

- Nagel R2 = 0.336
- -2LL = 206.514

Discrimination:

- ROC AUC = 0.801 (SE = 0.032, 95% CI 0.739 – 0.863)
- Discrimination slope =0.259 (SE = 0.032)

Calibration:

- HL: chi-squared = 7.906, p = 0.443
- Cal slope = 1.0226
- Intercept = -0.0097
- Brier score = 0.0084

Gr2+ CTCAE SDI

Univariate associations (regression)

Variable	Nature	Wald	P value	OR	95% CI
Gender (male)	binary	3.358	0.067	0.496	0.234 – 1.050
Age	cont	0.001	0.974	1	0.974 – 1.027
Baseline wt	cont	1.068	0.302	0.992	0.976 – 1.008
Weight loss (kg)	cont	0.032	0.859	1.005	0.948 – 1.067
Weight loss (%)	cont	0.369	0.543	1.017	0.963 – 1.075
T stage	ordinal	0.014	0.904	0.986	0.779 – 1.247
Bin T stage	binary	0.039	0.844	1.060	0.595 – 1.888
Bin N stage	binary	2.106	0.147	1.532	0.861 – 2.727
cetuximab	**nominal**	**6.903**	**0.009**	**5.040**	**1.508 – 16.844**
cisplatin	**nominal**	**14.038**	**0.000**	**3.371**	**1.785 – 6.367**
SACT Y/N	**binary**	**15.710**	**0.000**	**3.526**	**1.891 – 6.575**
Cisplatin Y/N	**binary**	**9.788**	**0.002**	**2.533**	**1.415 – 4.533**
Neck diss	binary	1.131	0.287	0.729	0.407 – 1.305
parotidectomy	binary	2.383	0.123	0.353	0.094 – 1.324
SMGectomy	ordinal	0.388	0.533	0.830	0.461 – 1.492
Smoker (PY)	cont	1.243	0.265	0.992	0.977 – 1.006
Alcohol (total units)	cont	0.085	0.770	1.001	0.995 – 1.007
IPG_DP	cont	0.739	0.390	1.012	0.985 – 1.039
CPG_DP	**Cont**	**11.772**	**0.001**	**1.043**	**1.018 – 1.068**
ISMG_DP	Cont	0.806	0.369	1.011	0.987 – 1.036
CSMG_DP	**Cont**	**13.025**	**0.000**	**1.032**	**1.015 – 1.050**
SPC_DP	**Cont**	**6.558**	**0.010**	**1.036**	**1.008 – 1.064**
MPC_DP	**Cont**	**7.258**	**0.007**	**1.037**	**1.010 – 1.065**
OC_DP	Cont	2.664	0.103	1.020	0.996 – 1.046
SGL_DP	**Cont**	**8.006**	**0.005**	**1.036**	**1.011 – 1.063**
IPG_DA	cont	1.090	0.296	1.014	0.988 – 1.040
CPG_DA	Cont	13.871	0.000	1.045	1.021 – 1.069
ISMG_DA	Cont	0.863	0.353	1.011	0.988 – 1.036
CSMG_DA	Cont	13.067	0.000	1.032	1.014 – 1.049
SPC_DA	Cont	6.750	0.009	1.036	1.009 – 1.063
MPC_DA	Cont	7.661	0.006	1.038	1.011 – 1.065
OC_DA	Cont	2.398	0.122	1.019	0.995 – 1.043
SGL_DA	Cont	8.252	0.004	1.037	1.012 – 1.062
Baseline SDI	ordinal	3.701	0.054	1.805	0.989 – 3.296

Additonal colinearity diagnostics

Test	Tolerance	VIF
SACT Y/N vs CPG_DP	0.711	1.406
SACT Y/N vs CSMG_DP	0.672	1.487
SACT Y/N vs SPC_DP	0.739	1.353
SACT Y/N vs MPC_DP	0.822	1.216
SACT Y/N vs SGL_DP	0.829	1.206

Parameters put into model for LR based forward selection:

- SACT Y/N
- CPG_DA
- CSMG_DA
- SPC_DA
- SGL_DA

NB. MPC_DA omitted due to colinearity with SGL_DA (SGl_DA UVA associations stronger)

Model performance with this procedure:

- Nagelkerke R2 = 0.114
- -2LL = 243.942
- HL: chi-squared = , p =

Significance of adding each parameter:

- **SACT Y/N p=0.000 (Step 1)**
- **CSMG_DA p=0.058 (Step2)**
- CPG_DA. P = 0.229
- SPC p=0.657
- SGL p=0.125

Therefore stopped at step 2, with SACT Y/N and CSMG_DA as parameters

Model generation repeated for planned dose:

- Only these 2 parameters
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 0.918(chemo Y/N) + 0.019(CSMG_DP) -1.660

Overall:

- Nagel R2 = 0.137
- -2LL = 245.036

Discrimination:

- ROC AUC = 0.670 (SE=0.038, 95%CI 0.595-0.746)
- Discrimination slope = 0.103 (SE = 0.022)

Calibration:

- HL: chi-squared = 5.437, p = 0.710
- Cal slope = 1.0192
- Intercept = -0.0095
- Brier score = 0.0061

Model generation repeated for delivered dose:

- Only these 2 parameters
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 0.912(chemo Y/N) + 0.019(CSMG_DA) -1.662

Overall:

- Nagel R2 = 0.137
- -2LL = 245.096

Discrimination:

- ROC AUC = 0.668 (SE=0.039, 95%CI 0.593 – 0.744)
- Discrimination slope = 0.104 (SE = 0.022)

Calibration:

- HL: chi-squared = 7.810, p = 0.452
- Cal slope = 1.0295
- Intercept = -0.014
- Brier score = 0.0084

Gr2+ CTCAE Dys

Univariate associations (regression)

Variable	Nature	Wald	P value	OR	95% CI
Gender (male)	binary	0.379	0.538	1.320	0.546 – 3.194
Age	cont	0.028	0.868	0.997	0.965 – 1.031
Baseline wt	cont	0.030	0.863	0.998	0.979 – 1.018
Weight loss	cont	0.001	0.969	1.001	0.931 – 1.077
Weight loss (%)	cont	0.013	0.909	0.996	0.931 – 1.066
T stage	ordinal	0.042	0.838	1.031	0.769 – 1.383
Bin T stage	binary	0.007	0.933	1.031	0.502 – 2.119
Bin N stage	**binary**	**3.167**	**0.075**	**2.002**	**0.932 – 4.300**
cetuximab	nominal	0.066	0.798	1.241	0.238 – 6.465
cisplatin	**nominal**	**5.263**	**0.022**	**2.606**	**1.150 – 5.905**
SACT Y/N	**binary**	**4.620**	**0.032**	**2.428**	**1.081 – 5.450**
Cisplatin Y/N	**binary**	**5.597**	**0.018**	**2.514**	**1.171 – 5.394**
Neck diss	binary	0.626	0.429	1.317	0.666 – 2.606
parotidectomy	binary	0.000	0.999		
SMGectomy	ordinal	1.025	0.311	1.418	0.721 – 2.790
Smoker (PY)	cont	0.285	0.593	1.005	0.988 – 1.022
Alcohol (total units)	**cont**	**3.599**	**0.058**	**1.006**	**1.000 – 1.012**
IPG_DP	**cont**	**7.248**	**0.007**	**1.060**	**1.016 – 1.105**
CPG_DP	**Cont**	**8.005**	**0.005**	**1.046**	**1.014 – 1.079**
ISMG_DP	Cont	2.343	0.126	1.083	0.978 – 1.198
CSMG_DP	**Cont**	**7.649**	**0.006**	**1.039**	**1.011 – 1.068**
SPC_DP	**Cont**	**10.474**	**0.001**	**1.129**	**1.049 – 1.215**
MPC_DP	**Cont**	**6.821**	**0.009**	**1.094**	**1.023 – 1.170**
OC_DP	**Cont**	**8.005**	**0.005**	**1.069**	**1.021 – 1.120**
SGL_DP	**Cont**	**7.121**	**0.008**	**1.082**	**1.021 – 1.147**
IPG_DA	**cont**	**7.520**	**0.006**	**1.060**	**1.017 – 1.106**
CPG_DA	**Cont**	**8.351**	**0.004**	**1.045**	**1.014 – 1.077**
ISMG_DA	Cont	3.047	0.081	1.105	0.988 – 1.237
CSMG_DA	**Cont**	**7.869**	**0.005**	**1.040**	**1.012 – 1.069**
SPC_DA	**Cont**	**10.584**	**0.001**	**1.130**	**1.050 – 1.217**
MPC_DA	**Cont**	**7.202**	**0.007**	**1.090**	**1.024 – 1.162**
OC_DA	**Cont**	**8.674**	**0.003**	**1.073**	**1.024 – 1.124**
SGL_DA	**Cont**	**7.295**	**0.007**	**1.084**	**1.022 – 1.148**
Baseline Dys	ordinal	0.733	0.392	1.278	0.729 – 2.239

Co-linearity

No further analysis needed

Parameters put into model for LR based forward selection:

- Bin N
- cisplatin Y/N
- Alcohol (total units)
- IPG_DA
- CPG_DA
- CSMG_DA
- SPC_DA
- OC_DA
- SGL_DA

NB. MPC_DA omitted due to colinearity with SGL_DA. (SGl_DA UVA associations stronger)

Model performance with this procedure (Dys_v3):

- Nagelkerke R2 = 0.105
- -2LL = 147.896
- HL: chi-squared = 4.998, p = 0.758

Significance of adding each parameter:

- **SPC p = 0.001 (Step 1)**
- **Alcohol total units p = 0.083 (Step 2)**
- **SGL p=0.089 (Step 2)**
- Bin N. P = 0.549
- Cisplatin Y/N p = 0.955
- IPG p =0.291
- CPG p = 0.874
- CSMG p=0.967
- OC p=0.314

Therefore stopped at step 2, with CPG and Bline_Xer as model parameters

Model generation repeated for planned dose (Dys_v3.1):

- Only these 3 parameters
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 0.004(total EtOH)+(0.086xSPC_DP)+(0.071xSGL_DP) – 10.401

Overall:

- Nagel R^2 = 0.195
- -2LL = 150.223

Discrimination:

- ROC AUC = 0.749 (SE 0.042, 95%CI 0.666 – 0.831)
- Discrimination slope = 0.0569589 (SE = 0.020)

Calibration:

- HL: chi-squared = 2.676, p = 0.953
- Cal slope = 1.0808
- Intercept = -0.0166
- Brier score = 0.0024

<u>Model generation repeated for delivered dose:</u>

- These 3 parameters
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 0.004(total EtOH)+(0.089xSPC_DA)+(0.068xSGL_DA) – 10.597

Overall:

- Nagel R2 = 0.203
- -2LL = 149.336

Discrimination:

- ROC AUC = 0.752 (SE 0.043, 95% CI 0.668 – 0.836)
- Discrimination slope = 0.0575966 (SE = 0.020)

Calibration:

- HL: chi-squared = 4.343, p = 0.825
- Cal slope = 1.0591
- Intercept = -0.0122
- Brier score = 0.0023

Gr3+ H&N35 – Q41 – Have you had a dry mouth?

Univariate associations (regression)

Variable	Nature	Wald	P value	OR	95% CI
Gender (male)	**binary**	**4.451**	**0.035**	**0.438**	**0.203 – 0.943**
Age	cont	0.000	0.984	1	0.974 – 1.026
Baseline wt	cont	1.608	0.205	0.990	0.974 – 1.006
Weight loss	**cont**	**4.205**	**0.040**	**1.067**	**1.003 – 1.134**
Weight loss (%)	**cont**	**5.650**	**0.017**	**1.072**	**1.012 – 1.136**
T stage	ordinal	0.000	0.991	1.001	0.792 – 1.265
Bin T stage	binary	0.089	0.766	1.091	0.614 – 1.939
Bin N stage	**binary**	**8.009**	**0.005**	**2.320**	**1.295 – 4.157**
cetuximab	nominal	0.056	0.812	1.150	0.363 – 3.647
cisplatin	**nominal**	**4.068**	**0.044**	**1.846**	**1.017 – 3.350**
SACT Y/N	**binary**	**3.548**	**0.060**	**1.749**	**0.978 – 3.127**
Cisplatin Y/N	**binary**	**4.178**	**0.041**	**1.806**	**1.025 – 3.184**
Neck diss	binary	1.474	0.225	1.425	0.804 – 2.525
parotidectomy	binary	0.499	0.480	0.660	0.208 – 2.092
SMGectomy	ordinal	0.362	0.547	1.194	0.670 – 2.126
Smoker (PY)	cont	0.606	0.436	0.994	0.980 – 1.009
Alcohol (total units)	cont	0.106	0.744	1.001	0.995 – 1.007
IPG_DP	**cont**	**6.689**	**0.01**	**1.039**	**1.009 – 1.069**
CPG_DP	**Cont**	**14.476**	**0.000**	**1.048**	**1.023 – 1.073**
ISMG_DP	**Cont**	**4.804**	**0.028**	**1.043**	**1.004 – 1.082**
CSMG_DP	**Cont**	**16.374**	**0.000**	**1.036**	**1.018 – 1.054**
SPC_DP	**Cont**	**17.782**	**0.000**	**1.079**	**1.041 – 1.117**
MPC_DP	**Cont**	**13.416**	**0.000**	**1.068**	**1.031 – 1.107**
OC_DP	**Cont**	**15.594**	**0.000**	**1.060**	**1.030 – 1.092**
SGL_DP	**Cont**	**12.466**	**0.000**	**1.055**	**1.024 – 1.087**
IPG_DA	**cont**	**7.754**	**0.005**	**1.041**	**1.012 – 1.071**
CPG_DA	**Cont**	**15.695**	**0.000**	**1.048**	**1.024 – 1.072**
ISMG_DA	**Cont**	**4.938**	**0.026**	**1.044**	**1.005 – 1.084**
CSMG_DA	**Cont**	**16.568**	**0.000**	**1.036**	**1.018 – 1.053**
SPC_DA	**Cont**	**18.353**	**0.000**	**1.079**	**1.042 – 1.118**
MPC_DA	**Cont**	**13.661**	**0.000**	**1.066**	**1.030 – 1.103**
OC_DA	**Cont**	**16.046**	**0.000**	**1.060**	**1.030 – 1.091**
SGL_DA	**Cont**	**12.549**	**0.000**	**1.055**	**1.024 – 1.086**
Baseline H&N_35	ordinal	2.375	0.123	1.411	0.911 – 2.187

Additional Co-linearity diagnostics

Test	Tolerance	VIF
Gender vs WL(%)	0.986	1.015
Gender vs Bin N	0.999	1.001
Gender vs cisplatin Y/N	0.995	1.005
Gender vs IPG	0.996	1.004
Gender vs CPG	0.999	1.001
Gender vs ISMG	0.996	1.004
Gender vs CSMG	0.998	1.002
Gender vs SPC	1	1
Gender vs MPC	0.998	1.002
Gender vs OC	1	1
Gender vs SGL	0.996	1.004
WL(%) vs Bin N	0.912	1.097
WL(%) vs cisplatin Y/N	0.880	1.136
WL(%) vs IPG	1	1
WL(%) vs CPG	0.943	1.060
WL(%) vs ISMG	0.980	1.021
WL(%) vs CSMG	0.919	1.088
WL(%) vs SPC	0.931	1.074
WL(%) vs MPC	0.953	1.049
WL(%) vs OC	0.923	1.083
WL(%) vs SGL	0.936	1.068

Parameters put into model for LR based forward selection:

- gender
- Weight loss (%)
- Bin N
- Cisplatin Y/N
- IPG_DP
- CPG_DP
- ISMG_DP
- CSMG_DP
- SPC_DP
- MPC_DP
- OC_DP
- SGL_DP

NB. SGL_DA omitted due to colinearity with MPC_DA. (MPC_DA UVA associations stronger)

Model performance with this procedure:

- Nagelkerke R2 = 0.196
- -2LL = 165.362
- HL: chi-squared = 11.498, p = 0.175

Significance of adding each parameter:

- **SPC_DP. p = 0.000 (step 1)**
- Weight loss (%) = 0.133
- Gender. P = 0.140
- Bin N. P = 0.691
- Cisplatin Y/N p = 0.896
- IPG p =0.380
- CPG p = 0.904
- ISMG p=0.376
- CSMG p=0.863
- MPC p = 0.405
- OC p=0.348

Decision taken that this effect could be due to confounding, or co-linearity in the data, not passing pre-defined threshold.

Therefore procedure repeated with only dose metrics from structures known to contain salivary glands;

- Nagelkerke R2 = 0.125
- -2LL = 173.554
- HL: chi-squared = 6.053, p = 0.641

Significance of adding each parameter:

- **OC_DP. p = 0.000 (step 1)**
- CSMG p=0.078
- **Weight loss (%) = 0.10**
- Gender. P = 0.202
- Bin N. P = 0.926
- Cisplatin Y/N p = 0.465
- IPG p =0.3484
- **CPG p = 0.126**
- ISMG p=0.617
- CSMG p=0.078

Therefore OC, CPG and WL put into model for bootstrapping procedure. Decision yo include CPG 'a-priori', and based on the underlying premise that models are used to test underlying hypothesis – not the other way round.

Model generation repeated for planned dose:

- Only these 3 parameters
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 0.037(WL%) + 0.052(OC_DP) + 0.023(CPG_DP) – 3.535

Overall:

- Nagel R^2 = 0.171
- -2LL = 237.089

Discrimination:

- ROC AUC = 0.690 (SE 0.038, 95%CI 0.616 – 0.764)
- Discrimination slope = 0.1245505 (SE = 0.024)

Calibration:

- HL: chi-squared = 10.216, p = 0.250
- Cal slope = 0.9635
- Intercept = 0.017
- Brier score = 0.1256344

Model generation repeated for delivered dose:

Model performance:

S = 0.036(WL%) + 0.051(OC_DA) + 0.025(CPG_DA) – 3.569

Overall:

- Nagel R2 = 0.180
- -2LL = 235.681

Discrimination:

- ROC AUC = 0.697 (SE0.038, 95% CI 0.624 – 0.771)
- Discrimination slope = 0.1281094 (SE = 0.024)

Calibration:

- HL: chi-squared = 4.355, p = 0.824
- Cal slope = 0.9381
- Intercept = 0.03
- Brier score = 0.0054661

Gr3+ H&N35 – Q42 – Have you had sticky saliva?

Univariate associations (regression)

Variable	Nature	Wald	P value	OR	95% CI
Gender (male)	**binary**	**3.061**	**0.08**	**1.965**	**0.922 – 4.189**
Age	cont	0.001	0.973	1	0.972 – 1.029
Baseline wt	cont	0.649	0.420	0.993	0.976 – 1.010
Weight loss	cont	2.080	0.149	1.047	0.984 – 1.115
Weight loss (%)	**cont**	**3.070**	**0.080**	**1.055**	**0.994 – 1.119**
T stage	ordinal	0.158	0.691	1.052	0.818 – 1.354
Bin T stage	binary	2.022	0.155	1.559	0.845 – 2.877
Bin N stage	binary	1.095	0.295	1.393	0.749 – 2.592
cetuximab	**nominal**	**2.884**	**0.089**	**2.812**	**0.853 – 9.278**
cisplatin	**nominal**	**5.265**	**0.022**	**2.206**	**1.122 – 4.336**
SACT Y/N	**binary**	**5.890**	**0.015**	**2.270**	**1.171 – 4.400**
Cisplatin Y/N	**binary**	**3.579**	**0.059**	**1.818**	**0.979 – 3.378**
Neck diss	binary	0.06	0.806	0.926	0.499 – 1.715
parotidectomy	binary	1.531	0.216	0.379	0.081 – 1.763
SMGectomy	ordinal	0.520	0.471	0.788	0.413 – 1.505
Smoker (PY)	cont	0.184	0.668	0.997	0.981 – 1.012
Alcohol (total units)	cont	0.002	0.964	1	0.994 – 1.006
IPG_DP	cont	1.222	0.269	1.017	0.987 – 1.047
CPG_DP	**Cont**	**11.185**	**0.001**	**1.046**	**1.019 – 1.074**
ISMG_DP	**Cont**	**2.829**	**0.093**	**1.035**	**0.994 – 1.077**
CSMG_DP	**Cont**	**12.270**	**0.000**	**1.039**	**1.017 – 1.061**
SPC_DP	**Cont**	**5.198**	**0.023**	**1.037**	**1.005 – 1.071**
MPC_DP	**Cont**	**7.268**	**0.007**	**1.054**	**1.014 – 1.095**
OC_DP	**Cont**	**3.353**	**0.067**	**1.027**	**0.998 – 1.056**
SGL_DP	**Cont**	**8.922**	**0.003**	**1.063**	**1.021 – 1.107**
IPG_DA	cont	1.304	0.253	1.017	0.988 – 1.046
CPG_DA	**Cont**	**12.479**	**0.000**	**1.047**	**1.021 – 1.074**
ISMG_DA	**Cont**	**2.752**	**0.097**	**1.033**	**0.994 – 1.073**
CSMG_DA	**Cont**	**12.361**	**0.000**	**1.039**	**1.017 – 1.061**
SPC_DA	**Cont**	**5.396**	**0.020**	**1.038**	**1.006 – 1.070**
MPC_DA	**Cont**	**7.406**	**0.007**	**1.052**	**1.014 – 1.091**
OC_DA	**Cont**	**3.606**	**0.058**	**1.027**	**0.999 – 1.056**
SGL_DA	**Cont**	**9.033**	**0.003**	**1.064**	**1.022 – 1.108**
Baseline H&N_35	ordinal	0.238	0.626	1.131	0.690 – 1.851

No additional co-linearity diagnostics required.

Parameters put into model for LR based forward selection:

- gender
- weight loss (%)
- SACT Y/N
- CPG_DP
- ISMG_DP
- CSMG_DP
- SPC_DP
- OC_DP
- SGL_DP

NB. MPC_DA omitted due to colinearity with SGL_DA. (SGl_DA UVA associations stronger)

Model performance with this procedure:

- Nagelkerke R2 = 0.159
- -2LL = 166.372
- HL: chi-squared = 5.076, p = 0.749

Significance of adding each parameter:

- **CSMG_DP. p = 0.000 (step 1)**
- Weight loss (%) = 0.747
- Gender. P = 0.140
- SACT Y/N p = 0.678
- CPG p = 0.483
- ISMG p=0.817
- SPC p = 0.896
- OC p=0.987
- SGL p = 0.668

Therefore CSMG dose only put into model for bootstrapping.

Model generation repeated for planned dose:

- Only this parameters
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 0.038(CSMG_DP) – 2.559

Overall:

- Nagel R^2 = 0.111
- -2LL = 224.130

Discrimination:

- ROC AUC = 0.636 (SE 0.040, 95%CI 0.557 – 0.715)
- Discrimination slope = 0.0709058 (SE = 0.016)

Calibration:

- HL: chi-squared = 8.013, p = 0.432
- Cal slope = 0.9456
- Intercept = 0.017
- Brier score = 0.0093349

Model generation repeated for delivered dose:

Model performance:

S = 0.038(CSMG_DA) – 2.596

Overall:

- Nagel R2 = 0.112
- -2LL = 223.893

Discrimination:

- ROC AUC = 0.641 (SE = 0.040, 95% CI 0.563 – 0.720)
- Discrimination slope = 0.0728533 (SE = 0.016)

Calibration:

- HL: chi-squared = 10.447, p = 0.235
- Cal slope = 0.9592
- Intercept = 0.0128
- Brier score = 0.01206961

Gr3+ H&N35 – Q44 – taste alteration?

Univariate associations (regression)

Variable	Nature	Wald	P value	OR	95% CI
Gender (male)	binary	0.253	0.615	1.213	0.571 – 2.577
Age	cont	0.613	0.434	1.011	0.984 – 1.039
Baseline wt	cont	0.223	0.637	0.996	0.980 – 1.013
Weight loss	cont	1.143	0.285	1.034	0.973 – 1.099
Weight loss (%)	cont	1.084	0.298	1.031	0.974 – 1.091
T stage	ordinal	1.275	0.259	0.869	0.681 – 1.109
Bin T stage	binary	0.378	0.539	0.829	0.456 – 1.507
Bin N stage	binary	2.466	0.116	1.614	0.888 – 2.934
cetuximab	nominal	1.817	0.178	2.250	0.692 – 7.315
cisplatin	**nominal**	**7.750**	**0.005**	**2.492**	**1.310 – 4.738**
SACT Y/N	**binary**	**7.864**	**0.005**	**2.463**	**1.312 – 4.623**
Cisplatin Y/N	**binary**	**6.390**	**0.011**	**2.159**	**1.189 – 3.922**
Neck diss	binary	0.181	0.670	0.879	0.486 – 1.590
parotidectomy	binary	1.170	0.279	0.482	0.128 – 1.810
SMGectomy	ordinal	0.189	0.663	0.874	0.478 – 1.600
Smoker (PY)	cont	0.258	0.612	0.996	0.981 – 1.011
Alcohol (total units)	cont	1.028	0.311	1.003	0.997 – 1.009
IPG_DP	cont	1.696	0.193	1.019	0.991 – 1.048
CPG_DP	**Cont**	**8.951**	**0.003**	**1.038**	**1.013 – 1.064**
ISMG_DP	Cont	2.525	0.112	1.026	0.994 – 1.059
CSMG_DP	**Cont**	**11.420**	**0.001**	**1.032**	**1.013 – 1.051**
SPC_DP	**Cont**	**10.669**	**0.001**	**1.059**	**1.023 – 1.097**
MPC_DP	**Cont**	**7.175**	**0.007**	**1.043**	**1.011 – 1.076**
OC_DP	**Cont**	**5.324**	**0.021**	**1.033**	**1.005 – 1.061**
SGL_DP	**Cont**	**7.254**	**0.007**	**1.039**	**1.011 – 1.069**
IPG_DA	cont	2.091	0.148	1.020	0.993 – 1.048
CPG_DA	**Cont**	**8.712**	**0.003**	**1.036**	**1.012 – 1.060**
ISMG_DA	Cont	2.452	0.117	1.025	0.994 – 1.057
CSMG_DA	**Cont**	**11.498**	**0.001**	**1.032**	**1.013 – 1.051**
SPC_DA	**Cont**	**10.674**	**0.001**	**1.058**	**1.028 – 1.094**
MPC_DA	**Cont**	**7.324**	**0.007**	**1.042**	**1.012 – 1.074**
OC_DA	**Cont**	**4.650**	**0.031**	**1.029**	**1.003 – 1.056**
SGL_DA	**Cont**	**6.861**	**0.009**	**1.037**	**1.009 – 1.065**
Baseline H&N_35	**ordinal**	**8.310**	**0.004**	**2.081**	**1.264 – 3.428**

No additional co-linearity diagnostics required.

Parameters put into model for LR based forward selection:

- SACT Y/N
- CPG_DP
- CSMG_DP
- SPC_DP
- MPC_DP
- OC_DP
- Bline HN35 Q44

NB. SGL_DA omitted due to colinearity with MPC_DA. (MPC_DA UVA associations stronger)

Model performance with this procedure:

- Nagelkerke R2 = 0.149
- -2LL = 233.459
- HL: chi-squared = 3.622, p = 0.890

Significance of adding each parameter:

- **CSMG_DP. p = 0.000 (step 1)**
- **Bline HN 44 p = 0.003**
- SACT Y/N p = 0.234
- CPG p = 0.685
- SPC p = 0.165
- MPC p = 0.775
- OC p=0.901

Therefore CSMG dose and Bline toxicity only into model for bootstrapping.

Model generation repeated for planned dose:

- Only these 2 parameters
- Bootstrapped – 1000 samples, simple sampling.

Model performance:

S = 0.815(Bline_HN35_Q44) + 0.033(CSMG_DP) – 3.072

Overall:

- Nagel R^2 = 0.149
- -2LL = 233.459

Discrimination:

- ROC AUC = 0.681 (SE 0.039, 95%CI 0.604 – 0.758)
- Discrimination slope = 0.1100199 (SE = 0.023)

Calibration:

- HL: chi-squared = 3.622, p = 0.890
- Cal slope = 0.9766
- Intercept = 0.0091
- Brier score = 0.00389225

<u>Model generation repeated for delivered dose:</u>

Model performance:

S = 0.815(Bline_HN35_Q44) + 0.033(CSMG_DP) – 3.104

Overall:

- Nagel R2 = 0.151
- -2LL = 233.197

Discrimination:

- ROC AUC = 0.682 (SE = 0.039, 95% CI 0.605 – 0.758)
- Discrimination slope = 0.1112617 (SE = 0.023)

Calibration:

- HL: chi-squared = 5.393, p = 0.715
- Cal slope = 0.9704
- Intercept = 0.0113
- Brier score = 0.00613201

LENT/SOM – Salivary gland (subjective) – Gr2+

Univariate associations (regression)

Variable	Nature	Wald	P value	OR	95% CI
Gender (male)	**binary**	**4.719**	**0.030**	**0.427**	**0.198 – 0.920**
Age	cont	0.013	0.909	0.998	0.972 – 1.025
Baseline wt	cont	2.011	0.156	0.988	0.972 – 1.004
Weight loss	**cont**	**3.996**	**0.046**	**1.065**	**1.001 – 1.132**
Weight loss (%)	**cont**	**5.725**	**0.017**	**1.073**	**1.013 – 1.137**
T stage	ordinal	0.038	0.845	0.977	0.773 – 1.235
Bin T stage	binary	0.013	0.908	1.034	0.582 – 1.838
Bin N stage	**binary**	**10.741**	**0.001**	**2.680**	**1.486 – 4.833**
cetuximab	nominal	0.110	0.741	1.216	0.383 – 3.859
cisplatin	**nominal**	**4.800**	**0.028**	**1.951**	**1.073 – 3.549**
SACT Y/N	**binary**	**4.253**	**0.039**	**1.848**	**1.031 – 3.313**
Cisplatin Y/N	**binary**	**4.835**	**0.028**	**1.892**	**1.072 – 3.341**
Neck diss	binary	1.030	0.310	1.343	0.760 – 2.375
parotidectomy	binary	0.448	0.503	0.674	0.213 – 2.138
SMGectomy	ordinal	0.149	0.699	1.120	0.630 – 1.993
Smoker (PY)	cont	0.699	0.403	0.994	0.980 – 1.008
Alcohol (total units)	cont	0.026	0.871	1	0.995 – 1.006
IPG_DP	**cont**	**8.500**	**0.004**	**1.046**	**1.015 – 1.077**
CPG_DP	**Cont**	**15.619**	**0.000**	**1.050**	**1.025 – 1.076**
ISMG_DP	**Cont**	**4.907**	**0.027**	**1.044**	**1.005 – 1.085**
CSMG_DP	**Cont**	**15.640**	**0.000**	**1.035**	**1.017 – 1.053**
SPC_DP	**Cont**	**17.276**	**0.000**	**1.077**	**1.040 – 1.115**
MPC_DP	**Cont**	**12.576**	**0.000**	**1.063**	**1.028 – 1.100**
OC_DP	**Cont**	**16.130**	**0.000**	**1.063**	**1.032 – 1.094**
SGL_DP	**Cont**	**11.849**	**0.001**	**1.052**	**1.022 – 1.083**
IPG_DA	**cont**	**8.922**	**0.003**	**1.045**	**1.015 – 1.076**
CPG_DA	**Cont**	**17.107**	**0.000**	**1.050**	**1.026 – 1.075**
ISMG_DA	**Cont**	**5.042**	**0.025**	**1.046**	**1.006 – 1.088**
CSMG_DA	**Cont**	**15.976**	**0.000**	**1.035**	**1.018 – 1.052**
SPC_DA	**Cont**	**17.702**	**0.000**	**1.077**	**1.040 – 1.115**
MPC_DA	**Cont**	**12.743**	**0.000**	**1.061**	**1.027 – 1.096**
OC_DA	**Cont**	**16.442**	**0.000**	**1.062**	**1.031 – 1.093**
SGL_DA	**Cont**	**12.079**	**0.001**	**1.053**	**1.023 – 1.083**
Baseline	**ordinal**	**2.762**	**0.097**	**1.464**	**0.934 – 2.295**

No further co-linearity diagnostics required.

Parameters put into model for LR based forward selection:

- Gender
- Weight loss (%)
- Bin N
- cisplatin Y/N
- IPG_DP
- CPG_DP
- ISMG_DP
- CSMG_DP
- SPC_DP
- MPC_DP
- OC_DP
- Baseline

NB. SGL_DA omitted due to colinearity with MPC_DA. (MPC_DA UVA associations stronger)

Model performance with this procedure:

- Nagelkerke R2 = 0.188
- -2LL = 166.475
- HL: chi-squared = 7.867, p = 0.447

Significance of adding each parameter:

- **SPC_DP p =0.000 (step 1)**
- **Baseline p = 0.065**
- **Weight loss p = 0.095**
- Gender p = 0.169
- Bin N p = 0.833
- Cisplatin Y/N p = 0.734
- IPG_DP p = 0.503
- CPG_DP p = 0.591
- ISMG_DP p = 0.550
- CSMG_DP. p = 0.742
- MPC p = 0.451
- OC p=0.560

A-priori rationale to suppose that SPC dose is co-linear to outcome and that doses to other structures are a better predictor. Therefore procedure repeated without SPC and MPC.

- Nagelkerke R2 = 0.130
- -2LL = 173.168
- HL: chi-squared = 5.757, p = 0.674

Significance of adding each parameter:

- **OC_DP p =0.000 (step 1)**
- Baseline p = 0.153
- **Weight loss p = 0.073**
- Gender p = 0.228
- Bin N p = 0.536
- Cisplatin Y/N p = 0.398
- IPG_DP p = 0.286
- **CPG_DP p = 0.076**
- ISMG_DP p = 0.572
- **CSMG_DP. p = 0.076**

Repeated with highlighted parameters (plus baseline – from iteration 1).

- Nagelkerke R2 = 0.147
- -2LL = 235.893
- **HL: chi-squared = 24.879, p = 0.002**

Significance of adding each parameter:

- **OC_DP p =0.000 (step 1)**
- **Baseline p = 0.054**
- **CPG_DP p = 0.076**
- **CSMG_DP p = 0.079**
- Weight loss p = 0.110

Four highlighted parameters added to model for bootstrapping.

- Nagelkerke R2 = 0.199
- -2LL = 237.087
- HL: chi-squared = 8.781, p = 0.361

CSMG has lowest coefficient so removed.

Planned dose model:

- Parameters: CPG_DP, OC_DP, baseline
- Bootstrapped – 1000 samples, simple sampling.

S = (0.696 x Baseline) + (0.041 x CPG_DP) + (0.043 x OC_DP) – 3.534

Model performance

Overall:

- Nagel R^2 = 0.200
- -2LL = 241.897

Discrimination:

- ROC AUC = 0.735 (SE 0.035, 95%CI 0.665 – 0.804)
- Discrimination slope = 0.1542364 (SE = 0.025)

Calibration:

- HL: chi-squared = 9.511, p = 0.301
- Cal slope = 1.0208
- Intercept = -0.0092
- Brier score = 0.01010344

Delivered dose model:

- Parameters: CPG_DA, OC_DA, baseline
- Bootstrapped – 1000 samples, simple sampling.

S = (0.714 x Baseline) + (0.043 x CPG_DP) + (0.041 x OC_DP) – 3.560

Model performance

Overall:

- Nagel R^2 = 0.210
- -2LL = 240.096

Discrimination:

- ROC AUC = 0.741 (SE 0.035, 95%CI 0.673 – 0.810)
- Discrimination slope = 0.1622482 (SE = 0.026)

Calibration:

- HL: chi-squared = 6.002, p = 0.647
- Cal slope = 1.0425
- Intercept = -0.0191
- Brier score = 0.00551808

LENT/SOM – Salivary gland (management) – Gr1+

Univariate associations (regression)

Variable	Nature	Wald	P value	OR	95% CI
Gender (male)	binary	0.124	0.725	0.875	0.415 – 1.843
Age	cont	0.849	0.357	0.987	0.961 – 1.014
Baseline wt	cont	1.636	0.201	0.989	0.973 – 1.006
Weight loss	cont	1.917	0.166	1.044	0.982 – 1.108
Weight loss (%)	cont	1.771	0.183	1.039	0.982 – 1.099
T stage	ordinal	1.879	0.170	0.845	0.665 – 1.075
Bin T stage	binary	0.827	0.363	0.761	0.423 – 1.370
Bin N stage	binary	7.697	0.006	2.335	1.283 – 4.250
cetuximab	nominal	0.253	0.615	1.388	0.386 – 4.987
cisplatin	**nominal**	**19.968**	**0.000**	**4.504**	**2.328 – 8.716**
SACT Y/N	**binary**	**17.410**	**0.000**	**3.958**	**2.074 – 7.552**
Cisplatin Y/N	**binary**	**21.080**	**0.000**	**4.264**	**2.296 – 7.920**
Neck diss	**binary**	**2.898**	**0.089**	**0.590**	**0.322 – 1.083**
parotidectomy	**binary**	**3.349**	**0.067**	**0.239**	**0.051 – 1.107**
SMGectomy	ordinal	2.666	0.103	0.594	0.318 – 1.110
Smoker (PY)	cont	0.433	0.510	0.995	0.981 – 1.010
Alcohol (total units)	cont	0.139	0.710	1.001	0.995 – 1.007
IPG_DP	**cont**	**3.310**	**0.069**	**1.026**	**0.998 – 1.056**
CPG_DP	**Cont**	**19.270**	**0.000**	**1.059**	**1.032 – 1.087**
ISMG_DP	**Cont**	**5.603**	**0.018**	**1.078**	**1.013 – 1.147**
CSMG_DP	**Cont**	**22.408**	**0.000**	**1.053**	**1.031 – 1.076**
SPC_DP	**Cont**	**17.656**	**0.000**	**1.092**	**1.048 – 1.137**
MPC_DP	**Cont**	**13.449**	**0.000**	**1.086**	**1.039 – 1.135**
OC_DP	**Cont**	**15.929**	**0.000**	**1.068**	**1.034 – 1.104**
SGL_DP	**Cont**	**11.285**	**0.001**	**1.059**	**1.024 – 1.096**
IPG_DA	**cont**	**3.240**	**0.072**	**1.025**	**0.998 – 1.053**
CPG_DA	**Cont**	**19.229**	**0.000**	**1.056**	**1.031 – 1.082**
ISMG_DA	**Cont**	**6.194**	**0.013**	**1.087**	**1.018 – 1.162**
CSMG_DA	**Cont**	**22.485**	**0.000**	**1.053**	**1.031 – 1.075**
SPC_DA	**Cont**	**18.230**	**0.000**	**1.093**	**1.049 – 1.139**
MPC_DA	**Cont**	**14.326**	**0.000**	**1.087**	**1.041 – 1.135**
OC_DA	**Cont**	**15.966**	**0.000**	**1.067**	**1.033 – 1.101**
SGL_DA	**Cont**	**11.819**	**0.001**	**1.063**	**1.027 – 1.101**
Baseline	ordinal	0	1		

Additional Co-linearity diagnostics

Test	Tolerance	VIF
Neck diss vs cisplatin Y/N	0.960	1.042
Neck diss vs parotidectomy	0.972	1.029
Neck diss vs IPG_DP	0.989	1.011
Neck diss vs CPG_DP	0.979	1.022
Neck diss vs ISMG_DP	0.997	1.003
Neck diss vs CSMG_DP	0.975	1.026
Neck diss vs SPC_DP	1	1
Neck diss vs MPC_DP	1	1
Neck diss vs OC_DP	0.983	1.017
Neck diss vs SGL_DP	1	1

Parameters put into model for LR based forward selection:

- cisplatin Y/N
- Neck diss
- parotidectomy
- IPG_DP
- CPG_DP
- ISMG_DP
- CSMG_DP
- SPC_DP
- MPC_DP
- OC_DP

NB. SGL_DA omitted due to colinearity with MPC_DA. (MPC_DA UVA associations stronger)

Model performance with this procedure:

- Nagelkerke R2 = 0.168
- -2LL = 171.165
- HL: chi-squared = 2.167, p = 0.975

Significance of adding each parameter:

- **CSMG_DP p =0.000 (step 1)**
- **Cisplatin Y/N p = 0.011**
- Neck diss p = 0.672
- Parotidectomy p = 0.248
- IPG_DP p = 0.430
- CPG_DP p = 0.453

- ISMG_DP p = 0.281
- SPC_DP p = 0.245
- MPC p = 0.320
- OC p = 0.178

Therefore CSMG_DP and cisplatin added to bootstrapped model

Planned dose model:

- Parameters: CSMG_DP, cisplatin Y/N
- Bootstrapped – 1000 samples, simple sampling.

S = (0.912 x cisplatin Y/N) + (0.040 x CSMG_DP) – 2.731

Model performance

Overall:

- Nagel R^2 = 0.240
- -2LL = 224.927

Discrimination:

- ROC AUC = 0.735 (SE 0.035, 95%CI 0.665 – 0.805)
- Discrimination slope = 0.1766224 (SE = 0.026)

Calibration:

- HL: chi-squared = 4.090, p = 0.849
- Cal slope = 0.9723
- Intercept = 0.0114
- Brier score = 0.0044942

Delivered dose model:

- Parameters: CSMG_DP, cisplatin Y/N
- Bootstrapped – 1000 samples, simple sampling.

S = (0.895 x cisplatin Y/N) + (0.040 x CSMG_DA) – 2.753

Model performance

Overall:

- Nagel R^2 = 0.240
- -2LL = 224.930

Discrimination:

- ROC AUC = 0.735 (SE 0.036, 95%CI 0.665 – 0.804)
- Discrimination slope = 0.1782203 (SE = 0.026)

Calibration:

- HL: chi-squared = 5.544, p = 0.698
- Cal slope = 0.9733
- Intercept = 0.0101
- Brier score = 0.00578686

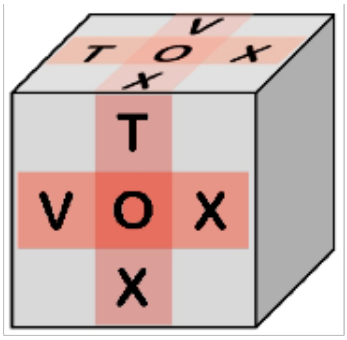
A6 Documents related to the Minot-OAR study

A6.1 VoxTox study protocol, including substantial amendment for Minot-OAR sub-study

A6.2 Patient information sheet – Minot-OAR sub-study

A6.3 Participant consent form – Minot-OAR sub-study

A6.4 GP letter – Minot-OAR sub-study



VoxTox

Linking radiation dose at the voxel level with toxicity

PROTOCOL

An observational study to collect comprehensive toxicity data for patients undergoing image-guided intensity-modulated radiotherapy to the head and neck, prostate, and central nervous system

Chief Investigator: Prof Neil Burnet
University of Cambridge/
Cambridge University Hospitals NHS Foundation Trust

A National Institute for Health Research Clinical Research Network (NIHR CRN) Portfolio Study

Study Sponsor: Cambridge University Hospitals NHS Foundation Trust/ University of Cambridge

REC Reference Number: 13/EE/0008

EudraCT Number: 2012-003367-22

IRAS ID: 109324

Protocol Number: VOXTOX Protocol Version: 2.0 23 February 2017

CANCER RESEARCH UK

PROTOCOL SIGNATURES

I give my approval for the attached protocol entitled 'VoxTox – Linking radiation dose at the voxel level with toxicity dated 23/02/2017.

Chief Investigator

Name: Neil Burnet

Signature:

Date: 23/02/2017

I agree to abide by all the provisions set forth in the attached protocol entitled 'VoxTox – Linking radiation dose at the voxel level with toxicity' dated 23/02/2017.

I agree to comply with the International Conference on Harmonisation Tripartite Guideline on Good Clinical Practice (ICH – GCP) and The European Clinical Trials Directive, 2001/20/EC and 2005/28/EC.

I agree to ensure that the confidential information contained in the document will not be used for any other purpose other than the evaluation or conduct of the clinical investigation without the prior written consent of the Sponsor.

Principal Investigator

Name: Neil Burnet

Signature:

Date: 23/02/3017

CLINICAL STUDY MANAGEMENT GROUP

Prof Neil Burnet, Consultant Oncologist, Professor of Radiation Oncology (Chair)

Dr Gillian Barnett, Clinical Research Training Fellow in Clinical Oncology

Amy Bates, Research Radiographer

Elizabeth Benns, Patient Representative

Dr Richard Benson, Consultant Oncologist

Dr Simon Bond, Senior Statistician, Cambridge Clinical Trials Unit

Dr Charlotte Coles, Consultant Oncologist

Dr Ferdia Gallagher, Clinical Scientist/Consultant Radiologist

Dr Mark Gurnell, Consultant Endocrinologist, University Lecturer

Dr Raj Jena, Clinician Scientist/Consultant Oncologist

Dr David Noble, Specialty Registrar in Clinical Oncology, Clinical Research Fellow

Dr Marina Romanchikova, Radiotherapy Physicist

Dr Simon Russell, Consultant Oncologist

Michael Simmons, VoxTox Programme Coordinator

Hilary Stobart, Patient Representative

TABLE OF CONTENTS

1.0 STUDY SUMMARY

Title	VoxTox - Linking radiation dose at the voxel level with toxicity An observational study to collect comprehensive toxicity data for patients undergoing image-guided intensity-modulated radiotherapy (IG-IMRT) to the head and neck, prostate and central nervous system
Aim	To establish the toxicity following IG-IMRT and how this relates to the delivered dose to the salivary glands, rectum and hypothalamic- pituitary axis (HPA)
Eligibility Criteria	**Inclusion criteria:** - Age ≥ 18 years - Malignant or benign tumour of the head and neck, prostate or central nervous system (CNS) - Already treated with, or suitable for treatment with, radical radiotherapy (RT) with daily image guidance using TomoTherapy or equivalent technology - Suitable for the 5 year follow up schedule - Adequate cognitive ability to participate in the interviews and complete the forms and questionnaires - Written informed consent. **Exclusion Criteria:** - Previous RT to the area being treated - CNS tumours without the cranium being part of the RT target
Study Design	Observational cohort study
Study Schema	**Discovery cohort – patients already treated:** Retrospective and prospective collection of clinical toxicity data until 5 years after radiotherapy. Archiving of imaging data. **Consolidation cohort – patients yet to be treated:** Prospective collection of clinical toxicity data before, during and for 5 years after radiotherapy. Optimisation of image guidance protocols and archiving of imaging data. • **MINOT-OAR sub-study:** Patients with head and neck cancer undergoing multiple timepoint MRI (before and during treatment), in addition to standard protocol image guidance and acquisition of toxicity data.
Objectives	**The primary objective** is to establish in detail the toxicity following IG-IMRT **Secondary objectives** are: - to optimise the image guidance protocols - to establish the delivered dose to the salivary glands, rectum and HPA - to assess whether the delivered dose correlates better with toxicity than the planned dose - to find imaging biomarkers of late toxicity early in a course of radiotherapy treatment.

2.0 BACKGROUND

Radiotherapy (RT) is one of the most potent and cost-effective curative treatments for cancer [Bentzen QUARTS 2005]. In the United Kingdom about 300,000 new cancer cases occur each year [Cancer Research UK CancerStats. 2011]. Around 50% of patients require radiotherapy at some time in their illness, and 60% are treated with curative intent [IARC 2011]. There is therefore a large potential benefit from both improvements in tumour control and reduction in toxicity.

The success of radiotherapy in eradicating tumours depends on the total radiation dose, delivered accurately. For most tumours, the higher the dose, the higher the chance of local tumour control and cure. There is a steep dose-cure relationship, both in experimental animal systems and in man, and a 5% increase in dose will typically achieve an increase in tumour cure in the range 5-10% [Suit 2002].

However, what limits RT dose is the tolerance of the normal tissues surrounding the tumour [Burnet et al 1996]. As the dose is increased so the incidence and severity of normal tissue damage also rises, and when severe, normal tissue damage can produce significant morbidity, which may even be life-threatening. Thus selection of the appropriate treatment is based on a balance between lowering the dose to keep the incidence of severe normal tissue complications at an acceptably low level, and raising the dose to increase the probability of local control and cure. For both head and neck cancer (HNC) and prostate cancer there is clinical evidence that dose escalation improves local tumour control rates. For many tumours, increasing local tumour control also increases cures, in the absence of metastatic disease. There is a relatively narrow therapeutic window, though it has been widened by modern treatment technologies, especially intensity-modulated radiotherapy (IMRT). In this way, normal tissue toxicity has a direct effect on the number of patients cured, as well as on morbidity.

In a given treatment setting, different patients experience different severity of toxicity. Some of this variation is the result of differences in anatomy, of both the tumour target and surrounding normal tissues, leading to variation in the doses delivered to the normal tissues. A component of this dose variation results from differences in position from one day to the next during the course of treatment; this is known as interfractional motion. Such differences may be due, for example, to patient positioning, internal organ movement, or progressive weight loss during the treatment course. Positional variation can be improved by the use of image-guided radiotherapy (IGRT).

Where the dose variation is reduced to a minimum, additional variation is seen which may be the result of differences in underlying tissue sensitivity to radiation, and which in turn may have a genetic basis [Barnett et al 2009]. Investigation of this genetic aspect requires the best possible knowledge and control of dose [Burnet et al 1998, Barnett et al 2009].

Toxicity has grown in importance as cure rates have risen, with earlier cancer detection and more effective treatment, becoming a major issue for individual patients and society alike. The financial cost of managing late effects of cancer treatment in survivors is high. Methods to reduce toxicity in cancer survivors will enhance quality of life and reduce the social and population burden from morbidity. Reducing toxicity will also allow development of protocols for both dose escalation and combination with conventional chemotherapy and newer molecular-targeted agents.

Radiotherapy technology

Intensity-modulated radiotherapy

IMRT allows reduction of dose-limiting toxicity, an important goal in itself, and through this mechanism, allows dose escalation to improve local control and cure [Cahlon et al 2008,

Cozzarini et al 2007, Dearnaley et al 2007, Donovan et al 2007, Nutting et al 2011, Veldeman et al 2008, Viani et al 2009, Zelefsky et al 2008]. The technique is especially valuable for patients with complex targets, especially concave shapes, including head and neck and prostate, but also other sites. It has been estimated that IMRT will benefit 30% of radiotherapy patients [Williams et al 2010].

Image-guided radiotherapy

IGRT uses imaging to localise the target. At its most interactive, IGRT uses daily computed tomography (CT) imaging, on the treatment couch, to adjust the patient's position prior to treatment, to improve the accuracy of dose delivery [Verellen et al 2007].

The prostate is a mobile internal structure that can move up to 2cm in the anterior- posterior direction relative to the pelvis, from one day to the next, and use of IGRT can improve targeting [Rimmer 2009, Verellen et al 2007, Xing et al 2006]. Patients with HNC typically lose weight during RT, and IGRT can correct resulting positioning inaccuracy. For patients with CNS tumours, IGRT ensures accurate treatment delivery. In turn, this allows the dose delivered to important normal structures, such as the hypothalamus and the pituitary gland, to be accurately known.

TomoTherapy – combining IGRT and IMRT

The TomoTherapy system is an integrated unit for the delivery of helical IMRT and IGRT, consisting of a compact in-line 6MV accelerator mounted on a CT ring gantry, with a detector array opposite (Fig 1). It provides volumetric, fan-beam megavoltage (MV) CT imaging on the treatment couch.

Figure 1. *The first of our 2 TomoTherapy units in Cambridge. Since clinical implementation in our department in 2007, we have used daily imaging with positional correction on all patients [Burnet et al 2010].*

TomoTherapy Image Guidance

CT scans are acquired with the linear accelerator slightly detuned to reduce the mean beam energy to 0.75 MeV (nominal accelerating voltage 3.5 MeV), and these are known as MV CTs, as opposed to kilovoltage (KV) scans used for diagnosis and RT planning. Imaging dose at depth is low; a typical value is 1.3 cGy for normal (4mm) slice thickness scan settings. Although the soft tissue definition is less good than diagnostic CT, it is easily good enough to allow registration with the planning CT, using the interfaces between fat and soft tissue, and soft tissue and bone [Burnet et al 2010].

The system can be set up to produce CT slices with a nominal equivalent thickness of 2, 4 or 6mm, representing 'fine', 'normal' and 'coarse' settings, by changing the pitch of the machine. The treatment radiographers determine the settings to use each day and may use a combination of different settings over the patient's treatment course, according to need. This decision is based on the spatial resolution of the image needed for matching, balanced against the dose from imaging and the speed required for the scanning procedure. The spatial resolution of the images is finer with smaller slice thicknesses; the imaging dose is less with larger slice thicknesses, and the acquisition speed is faster with larger slice thicknesses. Typical doses are 2.5 cGy for fine, 1.3 cGy for normal and 0.9 cGy for coarse settings. This compares to 200 cGy for a typical daily treatment dose. The average scan time is 3.5 minutes, with a standard deviation of 1.3 minutes [Burnet et al 2010].

In order to provide image-guided intensity-modulated (IG-IMRT) to the maximum possible number of patients at Addenbrooke's, the fastest possible scanning protocol was developed. The scan length (range) is kept as short as possible, to include only the length over which positional imaging is needed. 98% of scans in 2007 were taken using the fine setting. In 2008, 25% of patients were imaged with fine scans and 68% with the normal setting. Year on year, the average scan slice thickness has increased, and now the majority of patients are imaged with coarse scans (97% in 2011).

This policy exists to minimize scanning time and maximize machine capacity, and has been made possible by our radiographers' experience of matching. We sometimes still need to use the normal setting when a patient's prostate gland has shrunk following neoadjuvant hormones. Other centres have adopted a different approach, and are routinely using finer CT slices and longer scan ranges. These protocols include scanning the whole of the parotid glands in HNC patients [Han et al 2008], and using normal and fine settings to provide image guidance for patients having prostate radiotherapy [Langen et al 2005, Kupelian et al 2008, Shah et al 2008].

TomoTherapy archive and patient numbers

In Cambridge we have an archive of daily volumetric imaging of patients treated on our TomoTherapy units. This archive is unique within the UK and contains imaging data for over 1000 patients. We treat on average between 400 and 500 patients per annum on our 2 units, all of whom have daily imaging and positional correction. In 2010, 417 patients were treated: 48% for prostate cancer, 24% for HNC, and 7% for CNS tumours. In 2011, 473 patients were treated: 43% for prostate cancer, 24% for HNC and 9% for CNS tumours.

Toxicity from Radiotherapy

Dose-limiting toxicity varies with the site irradiated. Two of the most significant clinical toxicities are xerostomia (dry mouth) after treatment for HNC, and rectal toxicity after RT for prostate cancer. Both affect quality of life adversely. Good evidence exists for a benefit from RT dose escalation for these two sites.

Toxicity rates in patients treated with IGRT or IMRT appear lower than rates in those treated without using these technologies [Skala et al 2007, Nutting et al 2011]. Studies have tended to focus on hard end points, such as rectal bleeding, at limited time points

after RT, such as 2 and 5 years. Toxicity following treatment using both image guidance and intensity modulation has not been clearly documented. Investigation of toxicity in this group requires instruments able to describe more subtle symptoms of dysfunction than those traditionally used.

Xerostomia in HNC

Severe xerostomia (defined as long-term salivary function of <25% of baseline) is usually avoided if at least one parotid gland is spared to a mean dose below 20 Gy or both glands to less than 25 Gy [Deasy et al 2010]. This mean planned dose is only a static estimate of the dose to the parotid gland, which fails to account for changes in dosimetry due to shifts in patient position and weight loss during treatment.

A recent randomised controlled trial demonstrated reduced grade≥2 xerostomia with IMRT compared to conventional RT (29% vs 83%, p<0.0001) [Nutting et al 2011], despite the proximity of the tumour target to the parotid. The mean planned dose to the contralateral parotid gland was 25.4 Gy.

Our own dosimetry study yielded a similar mean planned dose to the contralateral parotid of 26.2 Gy, but re-calculation after repeat contouring on MVCTs yielded a higher mean dose of 33 Gy [Loo et al 2011]. This dose increase would lead to a 15% increase in probability of xerostomia ≥grade 2 (19.5% to 34.4%). In other words, 15% more patients would experience this side effect than originally planned, with consequent reduction in quality of life. This confirms the assertion that accurate dosimetry for normal tissue structures calculated for the entire treatment duration can yield significant differences in predicted treatment toxicity.

Prostate cancer - rectal and other toxicities

Injury to the rectum is the major toxicity associated with prostate RT, manifesting as diarrhoea, bleeding, urgency and stricture [Barnett et al 2011]. Damage to the anal sphincter, bladder and potency can also occur. RT planning is problematic because of the contiguity of the rectum with the posterior surface of the prostate gland, and the marked mobility of the prostate gland within the pelvis. The gland can move up to 2cm in the anterior-posterior direction relative to the bony pelvis, so a significant proportion of the rectum can move into, or out of, the high dose volume [Rimmer 2009] (Fig 2).

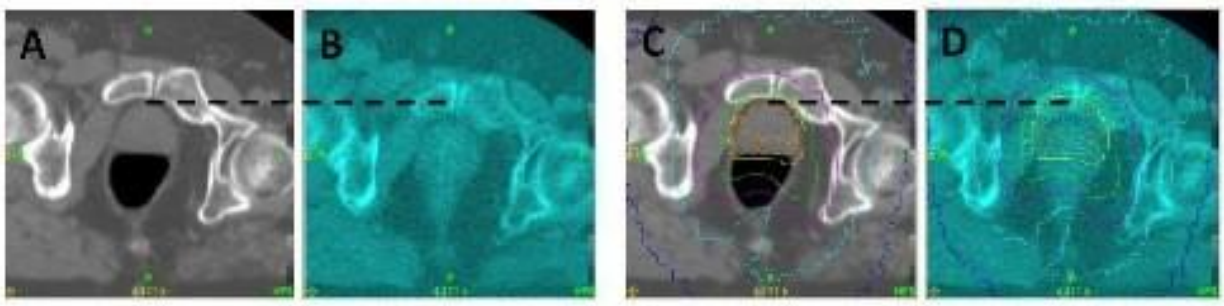


Figure 2. *Planning CT (A) compared to daily Tomo MVCT (B). The patient is more posterior on the treatment machine MVCT, and the rectum an entirely different shape. (C) shows the intended dose, and in (D) the dose overlay shows that dose to both target and rectum would be substantially lower. In this case, the position was corrected before treatment to deliver the dose correctly to the target. Without IGRT, rectal dose would have been lower than planned; after IGRT correcting target position, the rectal dose would have been higher. However, in both scenarios the distribution of dose to the rectum is entirely different from that planned.*

Studies of estimated rectal dose, typically based on weekly cone-beam CT scans, suggest that the majority of patients (60% [Hatton et al 2011], 75% [Chen et al 2010]) have worse rectal dose-volume histograms (DVHs) than shown on the treatment plan. This is consistent with our own data, an example of which is shown in Figure 3.

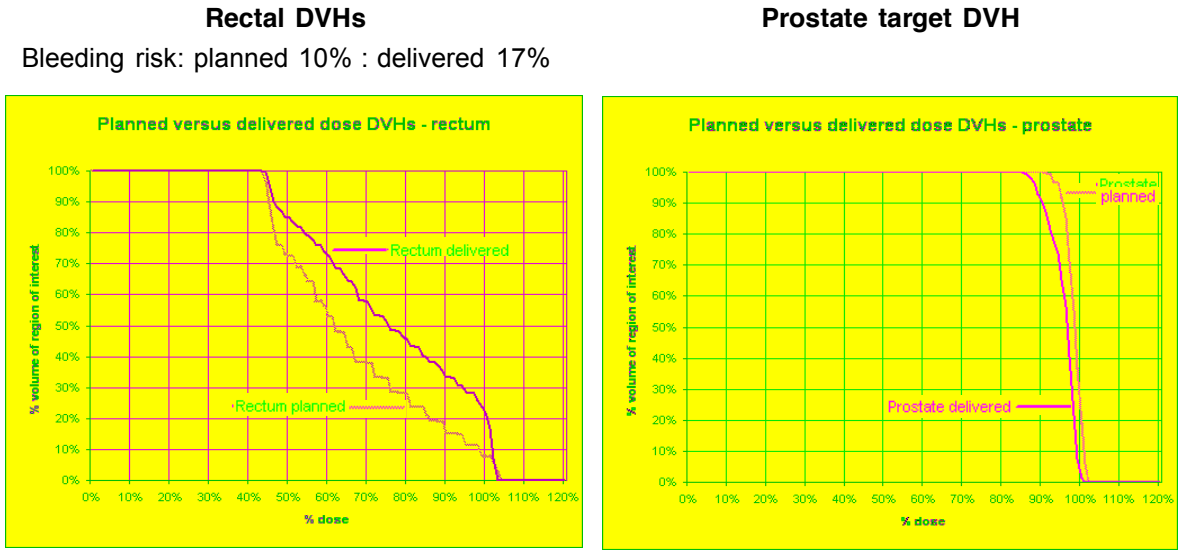


Figure 3. *Output from in-house 'SHERRI' dose evaluation software (written by Dr Raj Jena) showing change in DVH for the prostate target (red) and the rectum (green), for an example prostate patient. Both rectal and target DVHs deteriorate as a result of differences in prostate position during the treatment course. The predicted probability of rectal bleeding on the plan is 10%, but this rises to 17% in the delivered treatment.*

All such studies have used manual re-contouring on treatment machine scans to define the rectum, and identified that this is not practical on any large scale, so that development of automated tools is essential.

Toxicity can be modeled to first approximation as a solid structure using the Lyman-Kutcher-Burman model [Michalski et al 2010, Tudor 2007]. Although much of the existing dose-response data suggests that high doses are predominant in determining the risk of toxicity, there is also evidence of other subtle effects of dose. Analysis of the RT01 trial has shown that the number of DVH points which are violated [Gulliford et al 2010], and the shape of the dose distribution, are correlated with outcome [Buettner et al 2009].

Where rectal toxicity can be reduced through the use of IGRT, it is predicted that a 13% dose increase can be achieved [Ghilezan et al 2004], resulting in a 9% increase in tumour control. This is based on biochemical progression-free survival (bPFS) of 60% at 64 Gy and 71% at 74 Gy, from the RT01 prostate radiotherapy trial. This gives $_{D}50$ (dose required to achieve a 50% tumour control probability (TCP)) of 55 Gy, and gamma 50 (the change in TCP for a 1% change in dose at the 50% tumour control level) of 0.7, suggesting bPFS will rise from 71% to 80%.

This confirms the assertion that IGRT with daily imaging can reduce the effects of prostate motion, and that that improved modeling of rectal wall toxicity could be used to design future IGRT-based treatment protocols for prostate cancer.

Hypothalamic-pituitary axis dysfunction in CNS tumours

Hypothalamic-pituitary axis (HPA) dysfunction occurs in a high proportion of patients undergoing RT for pituitary adenomas and parasellar tumours. Failure of growth hormone secretion occurs in the majority of patients, and disturbance of gonadotroph (luteinising and

follicle-stimulating hormones (LH & FSH)), corticotroph (adrenocorticotrophic hormone (ACTH)) and thyrotroph (thyroid stimulating hormone (TSH)) function is observed in 30-60% [Anderson et al 1999, Darzy et al 2009]. Meticulous studies in children have demonstrated a dose and volume effect on hypothalamic function [Merchant et al 2002], but this requires quantification in adults.

Differential radiosensitivity exists in the different elements of the HPA, and intriguingly, there appears to be a particular sensitivity to dose per fraction [Anderson et al 1999, Darzy 2009]. In many cases an IMRT plan generates a dose gradient across the HPA. The gradient may reduce the dose *and* the dose per fraction, resulting in twice the benefit in reduced toxicity.

Currently we lack fine-grained data on the differential dose-volume response in adults. Image-guided IMRT can be used to reduce dose to the HPA, which may decrease the need for life-long replacement therapy and complications of endocrine dysfunction [Anderson et al 1999]. This data would provide evidence to endorse this approach, contribute to IMRT treatment optimisation, inform clinical risk stratification, and target early endocrine follow-up to those most at risk.

Anatomical changes during Radiotherapy

In many patients, the anatomy of the treated area changes significantly during treatment. This is particularly relevant for patients receiving radiotherapy for HNC. These tumours can shrink rapidly [Barker et al 2004], whilst patients often lose weight, and important organs at risk such as the salivary glands may change as a direct result of radiation damage [Vasquez Osorio et al 2008]. Therefore, there are substantial changes in size, shape and position of organs at risk during treatment, which can lead to significant differences between planned and delivered radiation dose, and impact on the risk of long-term toxicity. These changes can be seen and measured on MVCT images, but the high signal to noise ratio means that these measurements lack precision. It is well known that MRI gives better soft tissue definition than is possible with CT, and it has been used to describe changes in salivary gland anatomy pre to post radiotherapy [Houweling et al 2011]. Functional MRI also gives the possibility to assess the biology of changes occurring in soft tissue, and post-treatment MRI biomarkers have shown strong links to late radiotherapy toxicity [Juan et al 2015].

2.1 Rationale for Proposed Study

Current RT dose escalation strategies are based on static models of the patient anatomy, and do not take into account variation in patient position, and shape and location of mobile internal organs. Uncertainty in the dose actually delivered to normal tissues is recognised as a limitation in radiotherapy at the present time [Jaffray et al 2010].

The VoxTox Programme [Appendix 2] is bringing together a cross-disciplinary group of clinicians and clinically-orientated scientists in Cambridge to test the hypothesis that: "A better understanding of the dose received and the fate of normal tissues receiving RT will facilitate lower toxicity, and contribute to higher cancer cure rates". This will be achieved by implementing more sophisticated models of delivered dose than have previously been possible in clinical practice, thus linking dose at the voxel level with toxicity (VoxTox).

The goal of the Programme is to provide the tools that would allow:

1. A radiation oncologist to individualise treatment for a single patient, and
2. A clinical trials team to design and power the next generation of clinical trials in IMRT, IGRT and Adaptive Radiotherapy.

Within the Programme, sophisticated models of delivered dose, operating at the voxel level, will be developed. The discrepancy between RT doses calculated at treatment planning and those doses actually being delivered will be analysed. This will be achieved by developing systems capable of tracking the location of each point (voxel) within the patient, and updating the dose throughout the treatment course. Techniques from image processing, materials

modeling and high-energy physics will be applied to develop an integrated workflow for the radiation oncologist.

The VoxTox Study will collect and archive the clinical and imaging data required for the VoxTox Programme. The study will collect fine-grained clinical data for patients treated with IG-IMRT for HNC, prostate cancer and CNS tumours. It will provide a detailed database of clinical outcomes of acute and late toxicity in these patients, using established and validated scoring systems. This data will reflect patients' symptoms and functioning, their quality of life, and their own reported outcome measures. It will also document the effects of RT on pituitary function as measured by dynamic endocrine testing. The VoxTox Programme will compare this 'observed toxicity' with the expected toxicity for these patients. The relationships between these toxicities, and the planned and delivered RT doses will also be explored.

In the MINOT-OAR sub-study, patients with HNC will be invited to undergo a brief MRI scan at the beginning of the treatment course, and at 3 separate timepoints thereafter during their course of radiotherapy. MINOT-OAR stands for MultIple TimepoiNt MRI TO Track Organ At Risk (OAR) Changes In Patients Undergoing Radical Radiotherapy for Head and Neck Cancer. These images will be a powerful tool for the wider VoxTox programme. The nature and time course of changes in both the parotid and submandibular glands will be defined and will help to train tools for automated identification of these structures on MVCT images. This will improve our ability to accurately calculate delivered radiation dose and predict toxicity. Furthermore, MRI biomarkers measured at the end of treatment have been shown to predict toxicity, and we will look at the presence of these biomarkers earlier in the patient's course of treatment. By combining this data with information about radiotherapy dose we aim to select patients who will benefit from having their radiotherapy plan adapted during treatment, thereby reducing the risk of severe and lasting treatment side effects. Further details regarding the rationale and conduct of the MINOT-OAR sub-study are available in Appendix 3.

In order to be able to establish the delivered dose to each parotid gland over the course of treatment, then both glands need to be imaged regularly in their entirety. Currently, HNC patients at Addenbrooke's are imaged daily, but the scan length includes only the range over which positional imaging is needed. This means that the cranial part of one or both glands may not be included. We will therefore need to slightly lengthen the scans in some patients in order to fully cover them. Typically this will lead to an increase in the length scanned of around 1-2cm and is routine practice at other centres [Han et al 2008].

In the same way, in order to be able to establish the delivered dose to the rectum over the course of treatment, we wish to investigate how much of the rectum needs to be seen on the image guidance scans. This may mean extending the scan length by up to 5cm. The thickness of the CT determines the precision with which the top and bottom of critical structures can be identified and this is particularly important with prostate treatment. Most of these patients are imaged with coarse slices at Addenbrooke's; patients whose prostates have significantly reduced in size with neoadjuvant hormones often need to be imaged with normal slice thickness. We wish to investigate whether finer (smaller thickness) scans can improve imaging quality in the longitudinal direction. Normal and even fine slice thicknesses are routinely used at other centres for these patients [Langen et al 2005, Kupelian et al 2008, Shah et al 2008].

In these ways, our image guidance protocols will be optimised as part of the VoxTox Study. The image guidance data for each patient will be linked with their other data including planning CT scan, radiotherapy plan and clinical toxicity data. All data will then be archived in a suitable format for further analysis within the VoxTox Programme.

The discrepancy between expected and observed toxicity for the delivered dose should represent the difference in radiation sensitivity between individuals, hypothesized to result from underlying genetic factors. This will provide additional information for the RAPPER radiogenomics study of radiation toxicity (Radiogenomics: Assessment of Polymorphisms for Predicting the Effects of Radiotherapy CRUK Ref: C8857/A4764 and C1094/A11728). All patients eligible for the VoxTox Study will be invited to contribute a blood sample to the

RAPPER study, and if they consent then their clinical data collected within VoxTox will be passed to the RAPPER team.

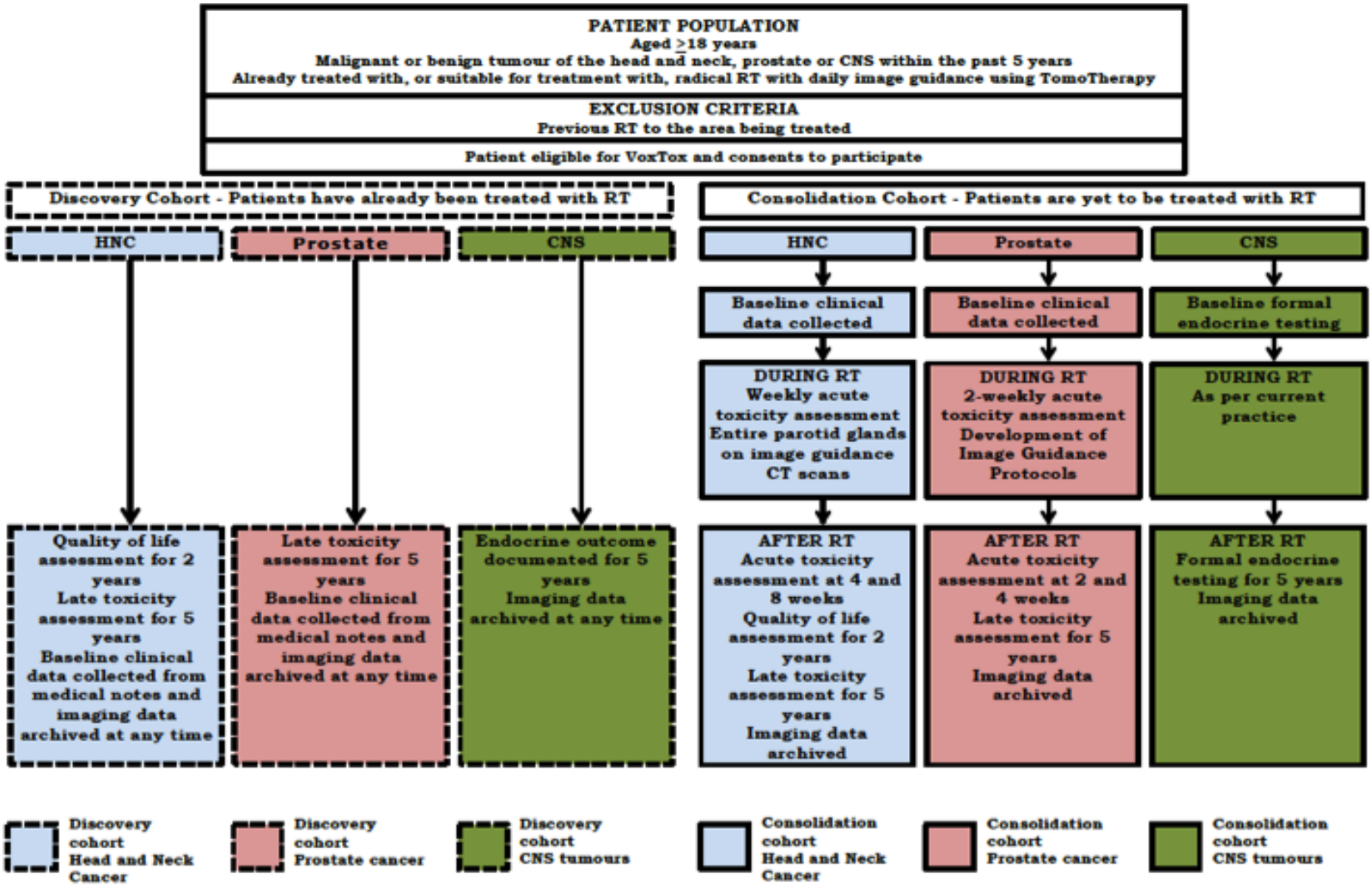
3.0 AIM

To establish the toxicity following IG-IMRT to the head and neck, prostate, and CNS, and how this relates to the delivered dose to the salivary glands, rectum and HPA.

4.0 STUDY DESIGN

This is an observational cohort study.

VOXTOX STUDY SCHEMA



5.0 PATIENT SELECTION AND ELIGIBILITY

5.1 Source of Patients

Patients who have already been treated with, or are suitable to be treated with, IG-IMRT for HNC, prostate cancer, or a CNS tumour at Addenbrooke's Hospital.

5.2 Number of Patients

There are 2 main cohorts of patients in the study, and an embedded sub-study:

- Discovery cohort. These patients have already been treated with RT, and were approached retrospectively. All patients eligible for the discovery cohort have now been approached. Two hundred and sixty nine (269) prostate patients and 95 HNC patients have been recruited to this cohort, which is now closed.

- Consolidation cohort. These patients are yet to be treated with RT. As of 21/2/2017, 255 prostate patients, 222 head and neck cancer patients and 24 CNS patients have been recruited. Based on current recruitment rates we anticipate a total of 260 prostate patients by 31/3/17, when the prostate consolidation cohort will close to recruitment. Recruitment to the standard head and neck consolidation cohort will continue until 31/3/17. Thereafter, head and neck patients may be recruited to the consolidation cohort within the MINOT-OAR sub-study. Based on current recruitment rates for the head and neck consolidation cohort, and anticipated numbers for the MINOT-OAR sub-study, we anticipate up to 250 head and neck consolidation patients. VoxTox CNS will continue to recruit patients, and we anticipate recruiting up to 50 patients in total.

- MINOT-OAR sub-study. The recruitment target for this sub-study is 15 patients. We anticipate that this recruitment will take 18 months to complete. Further details regarding patient numbers are available in the MINOT-OAR sub-study in Appendix 3.

There is no intervention in the study and no minimum sample size. These patient numbers have been produced as a guide for resource implications.

5.3 Inclusion Criteria

- Aged $\geq$18 years
- Malignant or benign tumour of the head and neck, prostate or CNS within the past 5 years
- Already treated with, or suitable for treatment with, radical RT with daily image guidance using TomoTherapy or equivalent technology
- Suitable for the planned 5 year follow up schedule
- Written informed consent

- For CNS patients, the RT target must include part of the cranium. This includes patients with gliomas, skull base tumours, those requiring cranio-spinal axis RT, and certain pituitary tumours.

5.3.1 Additional inclusion criteria – MINOT-OAR sub-study

- Histologically confirmed squamous cell carcinoma of the oropharynx, larynx or hypopharynx.
- Radiotherapy treatment volumes include lymph node regions on at least one side of the neck.
- Consented to receive concomitant systemic therapy.
- Baseline performance status 0 or 1.

5.4 Exclusion Criteria

- Previous RT to the area being treated
- RT treatment other than with daily image guidance using TomoTherapy or equivalent technology
- CNS tumours without the cranium being part of the RT target.

5.4.1 Additional exclusion criteria – MINOT-OAR sub-study

- Surgery prior to radiotherapy as part of definitive treatment for this malignancy
    - Specifically, neck dissection with or without laryngectomy or pharyngolaryngectomy prior to radiotherapy.
    - *Staging* surgery is not considered to be an exclusion criterion. Thus general anaesthetic with pan-endoscopy, tonsillectomy and/or debulking biopsy of the primary site, or biopsy of a lymph node would not exclude patients from this sub-study.
- Patients receiving radiotherapy to the primary site only (no neck irradiation).
- Standard MRI contra-indications. Specifically these are:
    - Heart pacemakers and implantable cardiac defibrillators
    - Insulin pumps
    - Implantable hearing aids
    - Implanted aneurysm clips
    - Neurostimulators
    - Metallic bodies in the eye.
- These points should be checked and discussed with patients otherwise suitable for recruitment. Some of the devices listed are now MRI compatible and safe, and if the patient has unambiguous written advice from the device manufacturer that it is MRI compatible and safe, they may still be considered for the study. If there is any doubt then the patient should not be invited to participate in the study.

6.0 IMAGE GUIDANCE PROTOCOLS

The image guidance scans from the discovery cohort will give initial information about the protocols. These will be optimised for the consolidation cohort in order to provide imaging data of sufficient quality to permit voxel tracking within the VoxTox Programme.

For the HNC patients, both parotid glands need to be imaged in their entirety. We will need to slightly lengthen the scans at the cranial end in some patients (estimated to be around 80% of patients) in order to fully cover them. This will lead to a maximal increase in the length scanned of 2cm. This amounts to an increase in imaging time of 0.2 minutes per patient per day, and a combined total of 2.9 minutes on both machines per day. The maximum number of fractions for HNC patients is 35 and the dose for each standard scan is estimated as 0.32 mSv. The estimated maximum dose using the standard protocol is therefore 11.2 mSv. The dose for the extended scans is estimated as 0.35 mSv and the estimated maximum dose using the study protocol is 12.2 mSv. This dose is 0.2% of the maximum dose from the radiotherapy. The combined additional total dose corresponds to a risk of radiation-induced cancer during the rest of the lifetime of the standard population of approximately 1 in 20,750 [ICRP 2007]. For clarity, this will be described as 1 in 20,000 in the participant information sheet.

For the prostate patients, there are two questions to resolve. Firstly, we wish to establish whether it is necessary to image the whole rectum on the image guidance scans in order for our software to be able to track it over a course of treatment. This would mean extending the scan length by up to 5cm in some patients. The entire rectum is already being imaged as standard for patients having their pelvic lymph nodes treated (13% of all prostate patients), and we will use this archived imaging to address this question in the first instance.

Secondly, we wish to investigate whether finer scans can improve imaging quality in the longitudinal direction. This would mean imaging patients with normal (4mm) rather than standard coarse (6mm) slices. We will initially study a cohort of 20 patients in whom half

the scans are done with coarse and half with normal slices. Assuming 5 cases on treatment being part of this sub study (out of ~32 per day treated), with half scans normal, this will add a total of 1.0 minute per day across both machines for the duration of the 20-patient study. The maximum number of fractions for prostate patients is 37. The dose for each coarse scan is estimated as 0.8 mSv and the estimated maximum dose using the standard protocol is 30.0 mSv. Changing to half the scans being coarse and half normal would result in an estimated maximum dose of 36.8 mSv. This dose is 0.5% of the maximum dose from the radiotherapy. The combined additional total dose corresponds to a risk of radiation-induced cancer during the rest of the lifetime of the standard population of approximately 1 in 2920.

Analysis of the imaging from the pelvic lymph node patients and the 20-patient cohort will give initial information about the benefits of imaging the whole rectum, and of using finer scan slices. We will use this information to guide the second tranche of imaging protocols for prostate patients, and this will be an ongoing process of development during the study. Table 1 illustrates the estimated doses and risks from the spectrum of imaging protocols that we may use for prostate patients in the study.

We will keep the length of the patient scanned (in cm) as short as possible, and use the thickest CT slices possible, in as many patients as possible. This will minimize imaging doses and acquisition times, ensuring minimum risk to patients, and minimum impact on the department. The risk quoted in the participant information sheet is that estimated from the imaging protocol that would deliver the highest dose, estimated at 1 in 720, and for clarity this will be phrased as 1 in 700..

Imaging Protocol	Estimated maximum dose (mSv)	Additional dose above standard (mSv)	Lifetime risk of malignancy from estimated maximum dose	Lifetime risk of malignancy from additional study dose
Standard length and all scans coarse (6mm) thickness	**30.0**	**0**	**1 in 670**	**N/A**
Standard length and half scans coarse and half normal (4mm) thickness	36.8	6.8	1 in 540	1 in 2920
Standard length and all scans normal thickness	43.3	13.3	1 in 460	1 in 1500
5cm longer and all scans coarse thickness	40.0	10.0	1 in 500	1 in 1990
5cm longer and all scans normal thickness	57.8	27.8	1 in 350	1 in 720

Table 1. *Image guidance dose estimates for potential prostate protocols in VoxTox. Risk of malignancy is calculated at 5% per Sv for the nominal adult population [ICRP 2007].*

Our estimates of risk of radiation-induced cancer are based on the current recommendations of the ICRP (International Commission on Radiological Protection). They estimate the risk of radiation-induced cancer over the remainder of the lifetime of the nominal adult population as 5% per Sv. In line with the ICRP recommendations we have used this figure to calculate the above estimates. However, the real risks in our population are likely to be substantially less. ICRP suggests that an adult population may have a lower risk of 4.1% per Sv. Our patients are older than a standard adult population: the mean age at randomization in the PARSPORT trial of RT in HNC patients was 58.4 (range 37.5-82.8) and median age in the RT01 trial of prostate RT was 67 (range 46-80) [Nutting et al 2011, Dearnaley et al 2007]. In addition, our patients have cancer and therefore, on average, their life expectancy is likely to be significantly less than in the nominal adult population.

We do not anticipate having to change the image guidance protocol for CNS patients, and this will continue to be as per current practice.

7.0 IMAGING PROTOCOL – MINOT-OAR SUBSTUDY

In addition to the standard VoxTox Image-Guidance protocol, patients recruited to the MINOT-OAR substudy will undergo four MRI examinations during the course of their treatment. Note that entry into this substudy will not affect their radiotherapy treatment in any way, and will be conducted in accordance with standard clinical practice and the rest of this protocol. The first scan will be conducted before the first fraction is delivered. Scan 2 will be conducted in week 2, scan 3 in week 4 and scan 4 in week 6. All MR imaging will be conducted in a scanner physically close to the radiotherapy department, to minimize disruption for patients. Scans will take no more than half an hour, and in most instances will be only 15-20 minutes in duration. No intravenous contrast is required.

MRI scans done within the MINOT-OAR substudy will be sent to the clinical PACS system, labeled with standard NHS identifiers. An NHS consultant radiologist will review them and produce a governance report. The MRI images will also be imported into the research database of the trust radiotherapy planning software system (Prosoma Research) for analysis. Further details of this imaging protocol including a schedule of events are available in the MINOT-OAR clinical work instruction in Appendix 3.

8.0 ARCHIVE OF IMAGING DATA

The image guidance scans for each patient (both cohorts) will be linked with their other radiotherapy data including their planning CT scan and RT plan. A process for anonymising the data, and inserting a unique patient identifier, or 'token', will be developed. The imaging data will then be passed from the NHS network to the University of Cambridge secure network, using a 'tunnel' through the hospital firewall, for which permission has been granted. Data transfer will use 128Bit SSL (Secure Sockets Layer) encryption, and the SFTP (secure file transfer protocol) protocol. All standard VoxTox imaging data will be stored on a named server in the data warehouse in the Cavendish Laboratory. The data from a particular patient will be linked to their clinical data, using their token. MR imaging data from the MINOT-OAR sub-study that is transferred to the Cavendish laboratory, will use the same procedure.

9.0 CLINICAL TOXICITY

The clinical toxicity evaluation depends on the patient's cohort, and on the site treated. All patients will be followed for 5 years, and on the whole the time points are identical to those already being used for routine follow up.

Acute and late toxicity will be assessed by interviews with radiographers. We will combine established clinical scoring tools in order to obtain as detailed information as possible without repetition. Quality of life will be assessed by validated questionnaires. We also wish to add a validated patient reported outcome measure (PROM), based on instruments used in clinical trials (e.g. Pivotal trial PROM).

For HNC patients, we will use components from the following tools:

- CTCAE v3 Common Toxicity Criteria for Adverse Events, version 3
- LENT SOM(A) Late Effects of Normal Tissue: Subjective, Objective, Management (Analytical)
- EORTC QLQ H+N35 European Organisation for Research and Treatment of Cancer
- Modified xerostomia questionnaire

For prostate patients, we will use components from the following tools:

- RTOG Radiation Therapy Oncology Group
- LENT SOM(A) Late Effects of Normal Tissue: Subjective, Objective, Management (Analytical)
- RMH Royal Marsden Hospital symptom scale
- UCLA PCI University of California, Los Angeles, Prostate Cancer Index

We wish to evaluate the potential added value of electronic collection of acute and late toxicity data using tablet devices. The Mosaiq system will be used to input the data and it will then be transferred in two directions:

1. Directly via secure wireless connection and stored within Mosaiq on the RT department desktop computers, where it will be available for use by clinical staff with access to Mosaiq

2. Anonymised and tokenized, sent via the secure wireless connection to the Cambridge clinical trials unit (CCTU) and stored as electronic case report forms (CRFs)

We will compare this electronic method for acute toxicity data collection with the paper version currently being used. Quantitative analysis will establish the benefit in terms of the completeness of data collected, the times involved, and the satisfaction of patients and staff.

We will use qualitative analysis to explore the views of patients and doctors regarding the electronic method of collecting late toxicity data. In particular we are interested in the experiences of patients entering their own data electronically before their clinic appointment and of doctors having access to this data when they see patients in clinic.

Formal endocrine testing will be undertaken within the endocrine investigation unit at the Wolfson Diabetes and Endocrine Clinic in the Institute of Metabolic Science at Addenbrooke's Hospital. Current practice is for CNS patients to have a simple screening blood test as part of annual follow up. If this is abnormal, or if they are symptomatic, then they are referred to an endocrinologist and formal endocrine testing is undertaken in order to establish the nature and extent of endocrine dysfunction. We want to implement this more detailed formal testing for all CNS patients in the consolidation cohort.

We will measure/perform:

- Serum insulin-like growth factor 1 (IGF-1) levels and the glucagon stimulation test
- Paired 0900hr serum ACTH/cortisol measurement combined with short synacthen® testing (SST)
- Free thyroid hormone (TH), thyroxine (T4), triiodothyronine (T3) and TSH levels and thyrotropin releasing hormone (TRH) testing
- Serum Prolactin
- 0900hr serum FSH, LH, sex hormone binding globulin (SHBG) and testosterone or oestradiol
- Electrolytes and paired serum and urine osmolalities

Samples will also be stored for future analysis related to this study.

9.1.0 Discovery cohort

Baseline data will be collected retrospectively from case records for HNC and prostate patients. This will include clinical factors that may impact on the development of late RT toxicity such as pre-existing morbidity, smoking status and medication [Barnett et al 2011].

The number of times a patient will be assessed as part of the study depends on the length of time between the end of treatment and recruitment (see VoxTox schedule of assessments on page 20).

9.1.1 Head and Neck Cancer

- Quality of life will be assessed every 6 months for 2 years after the end of RT.
- Late toxicity will be assessed every year for 5 years after the end of RT.

9.1.2 Prostate Cancer

- Late toxicity will be assessed every year for 5 years after the end of RT.

9.1.3 Central Nervous System Tumours

- Endocrine outcome will be documented every year for 5 years after the end of RT including:
  - results of endocrine blood tests performed as part of routine follow up
  - whether referral to an Endocrinologist has been required
  - results of any additional tests requested by Endocrine team

9.2.0 Consolidation cohort

Baseline data (as for discovery cohort) will be collected after recruitment and before RT for HNC and prostate patients. Quality of life will be assessed prior to RT in HNC patients. Formal endocrine testing will be performed in CNS patients before RT. (See VoxTox schedule of assessments on page 20).

9.2.1 Head and Neck Cancer

- Acute toxicity will be assessed every week for 6 weeks from the start of RT and then at 4 and 8 weeks after RT.
- Late toxicity will be assessed after RT at 3 and 6 months, 1 year, and then yearly until 5 years since RT.
- Quality of life will be assessed at 3 and 6 months, 1 year, 18 months and 2 years after RT.

9.2.2 Prostate Cancer

- Acute toxicity will be assessed every 2 weeks for 8 weeks from the start of RT and then at 2 and 4 weeks after RT.
- Late toxicity will be assessed after RT at 3 and 6 months, 1 year, and then yearly until 5 years since RT.

9.2.3 Central Nervous System Tumours

- Formal endocrine testing will be performed at 6 months, 1 year, and then yearly until 5 years since RT.
- A simple 0900hr basal blood sample will be sent by the GP for Free T4, Free T3, TSH, cortisol, LH, FSH, SHBG, testosterone or oestradiol, prolactin, IGF-1 and electrolyte levels at 18, 30, 42 and 54 months after RT.

10.0 ARCHIVE OF CLINICAL DATA

The anonymised, tokenized clinical toxicity data, as well as imaging and biomarker data from the MINOT-OAR sub-study, will populate a My SQL database within the Cambridge Clinical Trials Unit (CCTU). The quality of life, PROM and endocrine data will also be entered into this. The data from a particular patient will be linked to their imaging data using their token.

VOXTOX SCHEDULE OF ASSESSMENTS

Cohort	Site	Assessment	Before RT	During RT (W=Week)								After RT (W=Week and M=Month)													
				W1	W2	W3	W4	W5	W6	W7	W8	W2	W4	W8	M3	M6	M 12	M 18	M 24	M 30	M 36	M 42	M 48	M 54	M 60
Discovery	HNC	Quality of life														X	X	X	X						
		Late toxicity															X		X		X		X		X
	Prostate	Late toxicity															X		X		X		X		X
	CNS	Endocrine outcome															X		X		X		X		X
Consolidation	HNC	Baseline clinical data	X																						
		Acute toxicity		X	X	X	X	X	X				X	X												
		Late toxicity													X	X	X		X		X		X		X
		Quality of life	X												X	X	X	X	X						
	Prostate	Baseline clinical data	X																						
		Acute toxicity			X		X		X		X	X	X												
		Late toxicity													X	X	X		X		X		X		X
	CNS	Formal endocrine testing	X													X	X		X		X		X		X
		Basal blood sample																X		X		X		X	

Discovery cohort Head and Neck Cancer

Discovery cohort Prostate cancer

Discovery cohort CNS tumours

Consolidation cohort Head and Neck Cancer

Consolidation cohort Prostate cancer

Consolidation cohort CNS tumours

11.0 OBJECTIVES

The primary objective of the VoxTox Study is to establish the toxicity following IG-IMRT. We will produce a detailed toxicity database of patients who have been treated using this technology for tumours of the head and neck, prostate, and CNS. This will incorporate clinical outcomes of acute and late toxicity using established and validated scoring systems, endocrine testing, quality of life data and PROMs. Data will be collected at multiple time points (see Section 8). The database will contain at least 90% of toxicity data from at least 90% of patients recruited to the study. Evaluation of the database will include a description of the incidence, and time to onset, of toxicities.

There are 3 main secondary objectives:

- To optimise the image guidance protocols
  Development of the image guidance protocols for these patients will take into account the radiation dose to the patient, the length of time needed for scanning, and the image quality obtained. This will be evaluated in terms of the utility of the scans for the methodology within the VoxTox Programme, and in establishing the doses and times involved

- To establish the dose actually delivered to the salivary glands, rectum and HPA over a course of RT
  We will compare the delivered dose to these currently dose-limiting organs with the doses that were calculated at RT planning and will quantify this difference for the first time.

- To assess whether the delivered dose correlates better with toxicity than the planned dose
  The observed toxicity will be compared with the toxicity expected from the RT plan. The difference in toxicity will be linked to differences between the delivered and planned doses.

11.1.0 Linking to VoxTox Programme

The data generated by the VoxTox Study will be evaluated further within the VoxTox Programme [Appendix 2].

11.1.1 Data correlation

The difference between the planned dose and the dose actually delivered during the whole course of radiotherapy will be linked to differences between expected and observed toxicity. Observed toxicity will be correlated with parameters describing delivered dose, to determine the best predictors of toxicity. A number of parameters could correlate with clinical outcome, including percentages above a given dose, including median, maximum, and-near maximum dose, near maximum-dose with an absolute volume exclusion (e.g. $0.1cm^3$, $2cm^3$), and parameters involving the structure and shape of dose-surface maps.

11.1.2 Modeling Normal Tissue Complication Probability and Tumour Control Probability

We will model dose-volume and dose-surface relationships for rectal toxicity using standard methods, and compare our results with large clinical series [Burman et al 1991, Fiorino Review 2009, Gulliford et al 2010].

We will extend existing work on Tumour Control Probability (TCP) and Normal Tissue Complication Probability (NTCP), to model risks of both tumour control and toxicity for the whole population versus risks for individual patients. The effect of changing the planning target volume (PTV) margin on both TCP and NTCP will be studied in order to determine the effects of margin size and positional variation. The effect of imaging frequency will be tested

with a number of potential imaging strategies, examining the effect for both the population and individual patients. Although tumour control is not a primary objective of VoxTox, it will be considered here because it is linked with toxicity. We will develop secondary cancer induction models, which is an extension of existing work.

11.1.3 Development of tools for integration into the clinical workflow

We will develop a suite of tools designed to show accumulating dose, added after each daily treatment, compared to the original planning constraints. If predicted toxicity is less than originally expected, then individualised dose escalation could be achieved, up to the original toxicity risk. Where predicted toxicity is greater than initially estimated, it would be possible to re-plan treatment.

The concept is to develop automated systems to improve patient outcome, without additional staff workload or training. The exact design of such tools will require input from clinicians, physicists, engineers, and ultimately from manufacturers.

11.1.4 Development of hollow structure models using agent-based computing

We will expand existing work on computer-based modeling of hollow normal tissue structures, such as the rectum, for which dose-volume based analyses are not entirely satisfactory. We will take a portion of the clinical dataset and 'tune' the parameters of the model so as to provide a good fit to the clinical data. We will implement an approximate Bayesian computation approach to model fitting [see Beaumont 2010], building on our experience with the cellular Potts models [Sottoriva and Tavaré 2010], using the discovery dataset. Having fitted a model, it can be assessed on the consolidation dataset to see how well it reproduces the observed clinical outcomes.

11.2 Linking to RAPPER radiogenomics study

All patients eligible for the VoxTox Study will be invited to contribute a blood sample to the national RAPPER radiogenomics study (Ethics Ref: 05/Q0108/365).

Anonymous dose and outcome data will be passed to the RAPPER study. The greater refinement of estimates of delivered dose (as opposed to planned dose), together with NTCP modeling, is likely to allow greater insight into biological determinants of individual variation in toxicity.

12.0 STATISTICAL CONSIDERATIONS

The VoxTox study was designed to recruit as many patients as possible to maximize statistical power. An early analysis of toxicity data suggests that key event rates may differ from those originally anticipated; prostate – 17% for Gr 2 proctitis, head and neck - 25% for Gr 2 xerostomia. With total anticipated recruitment of 530 prostate patients, and 340 head and neck patients, these toxicity rates give 95% confidence intervals of ±3.5% (prostate) and ±5.0% (HNC), the same as those anticipated in the original design (±3.5% and ±4.9% respectively), and sufficiently accurate to make comparisons with historical data and plan any future studies. Long-term follow up is ongoing, and once data are sufficiently mature, multivariate analysis will be used as reported in the RT01 trial [Barnett et al 2011]. We will also use the Standardized Total Average Toxicity Score (STAT) as developed by our group [Barnett et al 2012].

The MINOT-OAR sub-study will recruit up to 15 patients. This will provide sufficient training data for segmentation tools developed within VoxTox, and is powered to detect differences between baseline and mid-treatment MRI ADC (Apparent Diffusion Coefficient) at 5% alpha and 90% power. This is based on data from Juan et al, who reported a difference in ADC of 0.21 (0.81 vs 1.02, standard deviation +/- 0.12) between patients with and without xerostomia.

13.0 RESEARCH GOVERNANCE

13.1.0 Trial Administration and Logistics

Cambridge University Hospitals NHS Foundation Trust and the University of Cambridge are jointly sponsoring the study, in accordance with the principles of Good Clinical Practice (GCP).

12.1.1 Responsibilities of Clinical Study Management Group

The Clinical Study Management Group (CSMG) has overall responsibility for facilitating and coordinating the conduct of the study. It is also responsible for monitoring and collating the data obtained and reporting the study outcomes. The CSMG will be required to contribute to the annual report submitted to Cancer Research UK.

13.2 Case Report Forms

Electronic CRFs will be completed at each assessment of acute and late toxicity and these will be stored in the CCTU. They will not be made available to third parties.

13.3 Protocol Compliance/Monitoring

The VoxTox Study is being conducted in accordance with the spirit and the letter of the declaration of Helsinki, the ICH-GCP Guidelines, the protocol and applicable local regulatory requirements. The Principal Investigator (PI) will ensure that:

- Sufficient data is recorded for all participating patients to enable accurate linkage between hospital records and CRFs
- Source data and all study related documentation are accurate, complete, maintained and accessible for monitoring and audit visits
- All study staff are trained and briefed appropriately
- All original Consent Forms are dated and signed by the patient and the person taking consent and kept in the Study Master File together with a copy of the Patient Information Sheet they were given at the time of consent
- Copies of CRFs are retained for 15 years to comply with international regulatory requirements
- Members of the CSMG monitor receipt of CRFs and check CRFs for compliance with the protocol and missing data
- Should a monitoring visit, audit or inspection be requested, that study documentation and source data are made available

13.4 Archiving

Essential documents are documents that individually and collectively permit evaluation of the conduct of the study and the quality of the data produced, for example CRFs. These documents will be maintained in the CCTU, Cambridge Cancer Trials Centre (CCTC) and Cavendish Laboratory in a way that will facilitate the management of the study, audit and inspection. They will be retained for a sufficient period (at least 15 years) for possible audit and inspection. Documents will be securely stored with security designed to meet the necessary regulatory requirements and access will be restricted to authorised personnel. An archive log will be maintained to track archived documents.

13.5 Financial Matters

The VoxTox Study is investigator designed and led and is being funded by Cancer Research UK. The study is recognised as being eligible for the NIHR CRN portfolio. The MINOT-OAR sub-study is funded by the consumables component of the Clinical Research Fellowship Grant of one of the study investigators (DJN).

13.6 End of Study

For the purposes of ethics approval, the study end date is deemed to be the date of the last data capture.

14.0 STUDY MANAGEMENT

A Clinical Study Management Group has been set up comprising:

Prof Neil Burnet, Consultant Oncologist, Professor of Radiation Oncology (Chair)

Dr Gillian Barnett, Clinical Research Training Fellow in Clinical Oncology

Amy Bates, Research Radiographer

Elizabeth Benns, Patient Representative

Dr Richard Benson, Consultant Oncologist

Dr Simon Bond, Senior Statistician, Cambridge Clinical Trials Unit

Dr Charlotte Coles, Consultant Oncologist

Dr Ferdia Gallagher, Clinical Scientist/Consultant Radiologist

Dr Mark Gurnell, Consultant Endocrinologist, University Lecturer

Dr Raj Jena, Clinician Scientist/Consultant Oncologist

Dr David Noble, Specialty Registrar in Clinical Oncology, Clinical Research Fellow

Dr Marina Romanchikova, Radiotherapy Physicist

Dr Simon Russell, Consultant Oncologist

Michael Simmons, VoxTox Programme Coordinator

Hilary Stobart, Patient Representative

Not withstanding the legal obligations of the Sponsor and Chief Investigator, the CSMG has operational responsibility for the conduct of the study.

15.0 PATIENT PROTECTION AND ETHICAL CONSIDERATIONS

15.1 Liability/Indemnity/Insurance

This study is an investigator-led study endorsed by Cancer Research UK. Cambridge University Hospitals NHS Foundation Trust and the University of Cambridge are jointly sponsoring the study. Cambridge University Hospitals NHS Foundation Trust, as a member of the NHS Clinical Negligence Scheme for Trusts, will accept full financial liability for harm caused to participants in the clinical study caused through the negligence of its employees and honorary contract holders. The University of Cambridge will arrange insurance for negligent harm caused as a result of protocol design and for non-negligent harm arising thorough participation in the clinical trial.

15.2 Patient Confidentiality

Patients will provide their full name, date of birth, hospital number, address, telephone number and GP details at registration and these details will be kept of behalf of the PI in the CCTC. The personal data recorded on all documents will be regarded as confidential, and to preserve each subject's anonymity, only their token will be recorded on CRFs. The PI must

maintain in strict confidence all study documents e.g. patients' written consent forms and must ensure the patient's confidentiality is maintained.
CCTC and CCTU will maintain the confidentiality of all subject data and will not reproduce or disclose any information by which subjects could be identified. In the case of competent authority queries, it is necessary to have access to the complete study records, provided that patient confidentiality is protected.

15.3 Ethical Considerations

It is the responsibility of the PI to obtain a favourable ethical opinion (main and site). It is the responsibility of the PI to give each patient, prior to inclusion in the study, full and adequate verbal and written information regarding the objective and procedures of the study and the possible risks involved. Sufficient time (a minimum of 24 hours) should be allowed for the patient to decide on study entry. Patients must be informed about their right to withdraw from the study at any time. Written patient information must be given to each patient before enrolment. The written patient information is an approved patient information sheet according to national guidelines. Patients will be encouraged to participate in associated studies including the RAPPER study, but if they decline, this will not exclude them from the main VoxTox Study.
It is the responsibility of the PI or designated representative, to obtain signed informed consent from all patients prior to inclusion in the study.
This study has been approved by the NRES Committee East of England - Essex Research Ethics Committee on 4 February 2013.

15.4 Patient Information

The importance of providing a high level of information to patients is recognised. Each patient invited into the VoxTox Study will receive a participant information sheet. They will be encouraged to participate in RAPPER, but if they decline, this will not exclude them from VoxTox.

16.0 WITHDRAWAL OF PATIENTS FROM STUDY

A subject may be withdrawn from the study at any time if they become unsuitable for the follow up required e.g. due to a decline in performance status. No further data will be collected, but information already received will be used for the original study purpose. This is fully explained in the patient information sheet and patients are asked to consent to this. A study deviation form should be completed for any patient who is withdrawn.

17.0 PROTOCOL AMENDMENTS

The PI will submit proposed protocol amendments to the CSMG. These will then be reviewed and agreed by the Sponsor prior to submission to the REC.

18.0 PUBLICATION POLICY

All publications and presentations relating to the study will be authorised by the CSMG. Authorship will be determined by the CSMG and will include the PI and appropriate members of the CSMG.

19.0 ASSOCIATED STUDIES

19.1 RAPPER radiogenomics study

As noted above (Section 10.2), all patients eligible for the VoxTox Study will be invited to contribute a blood sample to the RAPPER study, and if they consent then their data collected within VoxTox will be passed to the RAPPER team.

20.0 REFERENCES

Anderson JR, Antoun N, Burnet N, Chatterjee K, Edwards O, Pickard JD, Sarkies N. Neurology of the pituitary gland. *J Neurol Neurosurg Psych* 1999; 66 (6): 703-721

Barker JL, Jr., Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, Morrison WH, Rosenthal DI, Chao KS, Tucker SL, Mohan R, Dong L. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys. 2004;59(4):960-70.

Barnett GC, West CML, Dunning AM, Elliott RM, Coles CE, Pharaoh PD, Burnet NG. Normal Tissue Reactions to Radiotherapy: Towards Tailoring Treatment Dose by Genotype. *Nature Reviews Cancer* 2009; 9(2): 134-42

Barnett GC, De Meerleer G, Gulliford SL, Sydes MR, Elliott RM, Dearnaley DP. The Impact of Clinical Factors on the Development of Late Radiation Toxicity: Results from the Medical Research Council RT01 Trial (ISRCTN47772397). *Clin Oncol*. 2011; 23(9): 613-24

Barnett GC, West CML, Coles CE, Pharaoh PD, Talbot CJ, Elliott RM, Tanteles GA, Symonds RP, Wilkinson J, Dunning AM, Burnet NG, Bentzen SM. Standardized Total Average Toxicity (STAT) Score: A scale- and grade-independent measure of late radiotherapy toxicity to facilitate pooling of data from different studies. *Int J Radiat Oncol Biol Phys*. 2012; 82(3): 1065-74

Beaumont MA. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics* 2010; 41: 379–405

Bentzen SM, Heeren G, Cottier B, Slotman B, Glimelius B, Lievens Y, van den Bogaert W. Towards evidence-based guidelines for radiotherapy infrastructure and staffing needs in Europe: the ESTRO QUARTS project. *Radiother Oncol*. 2005; 75(3): 355-65

Burman C, Kutcher GJ, Emami B, et al. Fitting of normal tissue tolerance data to an analytic function. *Int J Radiat Oncol Biol Phys* 1991; 21:123–35

Burnet NG, Wurm R, Nyman J, Peacock JH. Normal tissue radiosensitivity - how important is it? *Clin Oncol* 1996; 8: 25-34

Burnet NG, Johansen J, Turesson I, Nyman J, Peacock JH. Describing patients‘ normal tissue reactions: Concerning the possibility of individualising radiotherapy dose prescriptions based on potential predictive assays of normal tissue radiosensitivity. Steering Committee of the BioMed2 European Union Concerted Action Programme on the Development of Predictive Tests of Normal Tissue Response to Radiation Therapy. *Int J Cancer* 1998: 79; 606-613

Burnet NG, Adams EJ, Fairfoul J, Tudor GSJ, Hoole ACF, Routsis, DS Dean JC, Kirby RD, Cowen M, Russell SG, Rimmer YL, Thomas SJ. Practical aspects of implementation of helical tomotherapy for intensity-modulated and image-guided radiotherapy. *Clin Oncol*. 2010; 22(4): 294-312

Buettner F, Gulliford SL, Webb S, Sydes MR, Dearnaley DP, Partridge M. Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: an analysis of data from the MRC RT01 trial (ISRCTN 47772397). *Phys Med Biol*. 2009; 54(21): 6535-48

Cahlon O, Hunt M, Zelefsky MJ. Intensity-modulated radiation therapy: supportive data for prostate cancer. *Semin Radiat Oncol*. 2008; 18(1): 48-57

Cancer Research UK CancerStats: Incidence 2008 - UK. 2011. http://info.cancerresearchuk.org/prod_consump/groups/cr_common/@nre/@sta/documents /gen eralcontent/cr_072111.pdf

Chen L, Paskalev K, Xu X, Zhu J, Wang L, Price RA, Hu W, Feigenberg SJ, Horwitz EM, Pollack A, Ma CM. Rectal dose variation during the course of image-guided radiation therapy of prostate cancer. *Radiother Oncol*. 2010; 95(2): 198-202

Cozzarini C, Fiorino C, Di Muzio N, Alongi F, Broggi S, Cattaneo M, Montorsi F, Rigatti P, Calandrino R, Fazio F. Significant reduction of acute toxicity following pelvic irradiation with helical tomotherapy in patients with localized prostate cancer. *Radiother Oncol*. 2007; 84(2): 164-70

Darzy KH. Radiation-induced hypopituitarism after cancer therapy: who, how and when to test. *Nat Clin Pract Endocrinol Metab*. 2009; 5(2): 88-99. Review.

Dearnaley DP, Sydes MR, Graham JD, Aird EG, Bottomley D, Cowan RA, Huddart RA, Jose CC, Matthews JH, Millar J, Moore AR, Morgan RC, Russell JM, Scrase CD, Stephens RJ, Syndikus I, Parmar MK; RT01 collaborators. Escalated-dose versus standard-dose conformal radiotherapy in prostate cancer: first results from the MRC RT01 randomised controlled trial. *Lancet Oncol*. 2007; 8(6): 475-87

Deasy JO, Moiseenko V, Marks L, Chao KS, Nam J, Eisbruch A. Radiotherapy dose-volume effects on salivary gland function. *Int J Radiat Oncol Biol Phys*. 2010; 76(3 Suppl): S58-63. Review.

Donovan E, Bleakley N, Denholm E, Evans P, Gothard L, Hanson J, Peckitt C, Reise S, Ross G, Sharp G, Symonds-Tayler R, Tait D, Yarnold J; Breast Technology Group. Randomised trial of standard 2D radiotherapy (RT) versus intensity-modulated radiotherapy (IMRT) in patients prescribed breast radiotherapy. *Radiother Oncol*. 2007; 82(3): 254-64

Fiorino C, Valdagni R, Rancati T, Sanguineti G. Dose-volume effects for normal tissues in external radiotherapy: pelvis. *Radiother Oncol*. 2009; 93(2): 153-67. Review.

Ghilezan M, Yan D, Liang J, Jaffray D, Wong J, Martinez A. Online image-guided intensity-modulated radiotherapy for prostate cancer: How much improvement can we expect? A theoretical assessment of clinical benefits and potential dose escalation by improving precision and accuracy of radiation delivery. *Int J Radiat Oncol Biol Phys.* 2004; 60(5): 1602-10

Gulliford SL, Foo K, Morgan RC, Aird EG, Bidmead AM, Critchley H, Evans PM, Gianolini S, Mayles WP, Moore AR, Sánchez-Nieto B, Partridge M, Sydes MR, Webb S, Dearnaley DP. Dose-volume constraints to reduce rectal side effects from prostate radiotherapy: evidence from MRC RT01 Trial ISRCTN 47772397. *Int J Radiat Oncol Biol Phys*. 2010; 76(3): 747-54

Han C, Chen YJ, Liu A, Schultheiss TE & Wong JYC. Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy. *Int J Radiat Oncol Biol Phys.* 2008; 70(4): 1256-62

Hatton JA, Greer PB, Tang C, Wright P, Capp A, Gupta S, Parker J, Wratten C, Denham JW. Does the planning dose-volume histogram represent treatment doses in image-guided prostate radiation therapy? Assessment with cone-beam computerised tomography scans. *Radiother Oncol.* 2011; 98(2): 162-8

Houweling AC, Schakel T, van den Berg CA, Philippens ME, Roesink JM, Terhaard CH, Raaijmakers CP. MRI to quantify early radiation-induced changes in the salivary glands. Radiother Oncol. 2011;100(3):386-9.

International Agency for Research on Cancer (IARC). 2011. GLOBOCAN 2008. http://globocan.iarc.fr/factsheets/cancers/all.asp

International Commission on Radiological Protection (ICRP). The 2007 Recommendations. ICRP publication 103. *Ann ICRP* 2007; 37(2-4): 1-332

Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tomé WA. Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. *Int J Radiat Oncol Biol Phys*. 2010; 76(3 Suppl): S135-9. Review.

Juan CJ, Cheng CC, Chiu SC, Jen YM, Liu YJ, Chiu HC, Kao HW, Wang CW, Chung HW, Huang GS, Hsu HH. Temporal Evolution of Parotid Volume and Parotid Apparent Diffusion Coefficient in Nasopharyngeal Carcinoma Patients Treated by Intensity-Modulated Radiotherapy Investigated by Magnetic Resonance Imaging: A Pilot Study. PLoS One. 2015;10(8):e0137073.

Kupelian PA, Lee C, Langen KM, Zeidan OA, Mañon RR, Willoughby TR & Meeks SL. Evaluation of image-guidance strategies in the treatment of localized prostate cancer. *Int J Radiat Oncol Biol Phys.* 2008*;* 70(4): 1151-7

Langen KM, Zhang Y, Andrews RD, Hurley ME, Meeks SL, Poole DO, Willoughby TR & Kupelian PA. Initial experience with megavoltage (MV) CT guidance for daily prostate alignments. *Int J Radiat Oncol Biol Phys.* 2005; 62(5): 1517-24

Loo H, Fairfoul J, Chakrabarti A, Dean JC, Benson RJ, Jefferies SJ, Burnet NG. Tumour shrinkage and contour change during radiotherapy increase the dose to organs at risk but not the target volumes for head and neck cancer patients treated on the TomoTherapy HiArtTM system. *Clin Oncol*. 2011; 23(1): 40-7

Merchant TE, Goloubeva O, Pritchard DL, Gaber MW, Xiong X, Danish RK, Lustig RH. Radiation dose-volume effects on growth hormone secretion. *Int J Radiat Oncol Biol Phys*. 2002; 52(5): 1264-70

Michalski JM, Gay H, Jackson A, Tucker SL, Deasy JO. Radiation dose-volume effects in radiation-induced rectal injury. Int J Radiat Oncol Biol Phys. 2010; 76(3 Suppl): S123-9. Review.

Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, Miles EA, Miah AB, Newbold K, Tanay M, Adab F, Jefferies SJ, Scrase C, Yap BK, A'Hern RP, Sydenham MA, Emson M, Hall E; PARSPORT trial management group. Parotid-sparing intensity-modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. *Lancet Oncol*. 2011; 12(2): 127-36

Rimmer YL. The Implementation and Optimisation of Image-Guided Radiotherapy in Prostate Cancer. Doctor of Medicine (M.D.) degree thesis, University of East Anglia, 2009

Robar JL, Day A, Clancey J, Kelly R, Yewondwossen M, Hollenhorst H, Rajaraman M, Wilke D. Spatial and dosimetric variability of organs at risk in head-and-neck intensity-modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2007;68(4):1121-30.

Shah AP, Langen KM, Ruchala KJ, Cox A, Kupelian PA & Meeks SL. Patient dose from megavoltage computed tomography imaging. *Int J Radiat Oncol Biol Phys*. 2008; 70(5): 1579-87

Skala M, Rosewall T, Dawson L, Divanbeigi L, Lockwood G, Thomas C, Crook J, Chung P, Warde P & Catton C. Patient-assessed late toxicity rates and principal component analysis after image-guided radiation therapy for prostate cancer. *Int J Radiat Oncol Biol Phys.* 2007; 68(3): 690-698

Sottoriva A & Tavaré S. Integrating approximate Bayesian computation with complex agent-based models for cancer research. In COMPSTAT 2010 -- Proceedings in Computational Statistics, Eds. Saporta G & Lechevallier Y. Springer, Physica Verlag. pp. 57-66

Suit H. The Gray Lecture 2001: coming technical advances in radiation oncology. *Int J Radiat Oncol Biol Phys.* 2002; 53(4): 798-809. Review.

Tudor GS. Comparison of rectal dose histograms under conditions of nonlinear isodoses. *Phys Med Biol*. 2007; 52(1): 275-89

Vasquez Osorio EM, Hoogeman MS, Al-Mamgani A, Teguh DN, Levendag PC, Heijmen BJ. Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration. Int J Radiat Oncol Biol Phys. 2008;70(3):875-82.

Veldeman L, Madani I, Hulstaert F, De Meerleer G, Mareel M, De Neve W. Evidence behind use of intensity-modulated radiotherapy: a systematic review of comparative clinical studies. *Lancet Oncol*. 2008; 9(4):367-75

Verellen D, De Ridder M, Linthout N, Tournel K, Soete G, Storme G. Innovations in image-guided radiotherapy. *Nat Rev Cancer*. 2007; 7(12):949-60

Viani GA, Stefano EJ, Afonso SL. Higher-than-conventional radiation doses in localized prostate cancer treatment: a meta-analysis of randomized, controlled trials. *Int J Radiat Oncol Biol Phys*. 2009; 74(5): 1405-18

Williams MV, Cooper T, Mackay R, Staffurth J, Routsis D, Burnet N. The implementation of intensity-modulated radiotherapy in the UK. *Clin Oncol*. 2010;22(8):623-8

Xing L, Thorndyke B, Schreibmann E, Yang Y, Li TF, Kim GY, Luxton G, Koong A. Overview of image-guided radiation therapy. *Med Dosim.* 2006; 31(2): 91-112

Zelefsky MJ, Levin EJ, Hunt M, Yamada Y, Shippy AM, Jackson A, Amols HI. Incidence of late rectal and urinary toxicities after three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for localized prostate cancer. *Int J Radiat Oncol Biol Phys*. 2008; 70(4): 1124-9

APPENDICES

Appendix 1: Abbreviations and Glossary

ACTH	Adrenocorticotrophic hormone
bPFS	biochemical Progression Free Survival
CCTC	Cambridge Cancer Trials Centre
CNS	Central Nervous System
CRF	Case Report Form
CSMG	Clinical study management group
CT	Computed tomography
CCTU	Cambridge Clinical Trials Unit
$_{D}50$	Dose required to achieve a 50% tumour control probability
DVH	dose-volume histogram. This describes the volume of a region of interest receiving a certain dose, or percentage dose. Its purpose is to summarise 3D dose distributions in a graphical 2D form
FSH	Follicle-stimulating hormone
gamma 50	the change in TCP for a 1% change in dose at the 50% tumour control level
GCP	Good clinical practice
HNC	head & neck cancer
HPA	Hypothalamic-Pituitary Axis (hormone production and control)
IGF-1	Insulin-like growth factor 1
ICH-GCP	International Conference on Harmonisation – Good Clinical Practice
IG-IMRT	Image-guided intensity-modulated radiotherapy
IGRT	Image-guided radiotherapy. Used to ensure that the dose distribution is correctly positioned over the tumour target, hence minimising dose to normal tissue structures. At its most interactive, uses daily CT imaging, with on-line image matching and positional correction before treatment.
IMRT	intensity-modulated radiotherapy. Allows dose distributions to be generated which achieve the best possible conformation to the tumour target; especially useful for complex targets, including those with concave shapes.
KV CT	kilovoltage CT (computerised tomography) scan. Kilovoltage CT is the diagnostic standard. Optimum difference between bone and soft tissue is achieved, because X-rays of this energy interact with matter by the photoelectric effect.
LH	Luteinising hormone
MACRO	A software application designed for data entry and analysis of clinical trial data

MVCT	megavoltage CT (computerised tomography) scan, performed with the TomoTherapy unit, using a mean beam energy to 0.75 MeV (nominal accelerating voltage 3.5 MeV). X-rays of this energy interact with matter by Compton scattering, resulting in relatively less absorption in bone than with a kV CT (but less artefact from metal implants and dental amalgam).
NTCP	normal tissue complication probability
PI	Principal Investigator
PROMs	Patient Reported Outcome Measures
PTV	Planning Target Volume. The safety margin applied to a tumour target to allow for uncertainties and positioning discrepancies in the process of treatment planning and delivery. The more precise the elements in the planning and treatment chain, the smaller the margin can be. The PTV is added outside gross tumour and tissues potentially infiltrated by microscopic disease, and thus encloses normal tissue. Reduction in the PTV margin, for example by using IGRT, should reduce the volume of normal tissue treated and consequently the toxicity, or NTCP.
RAPPER	Radiogenomics: Assessment of Polymorphisms for Predicting the Effects of Radiotherapy. The study is testing the hypothesis that variation in radiation toxicity amongst individuals treated with the same (tightly controlled) dose of radiotherapy is due to underlying genetic variation (polymorphisms). RAPPER is funded by Cancer Research UK.
REC	Research Ethics Committee
RT	Radiotherapy
SHBG	Sex hormone binding globulin
SST	Short synacthen® test
STAT	Standardized Total Average Toxicity Score, developed by the RAPPER group as a new metric of global radiation toxicity
T3	Triiodothyronine
T4	Thyroxine
TCP	Tumour Control Probability
TH	Thyroid hormone
Token	see tokenization
Tokenization	the process for anonymising the data, where a unique patient identifier, or 'token', is inserted into a data set to replace patient identifiable data. The token is used to link together all data from the same patient.
TRH	Thyrotropin releasing hormone
TSH	Thyroid stimulating hormone

Voxel	A voxel (volumetric pixel) is a volume element, representing a value on a regular grid in three-dimensional space. It is analogous to a pixel, which represents 2D image data in a bitmap.
Xerostomia	dry mouth

Appendix 2: The VoxTox Research Programme

VOXTOX PROGRAMME BOARD

Prof Neil Burnet, PI, Professor of Radiation Oncology

Dr Richard Benson, Consultant Oncologist

Dr Charlotte Coles, Consultant Oncologist

Dr Raj Jena, Clinician Scientist/Consultant Oncologist

Prof Andy Parker, Professor of High Energy Physics

Dr Simon Russell, Consultant Oncologist

Michael Simmons, Programme Coordinator

Dr Michael Sutcliffe, Reader in Mechanics of Materials (Engineering)

Prof Simon Tavaré, Professor of Cancer Research (Bioinformatics), and Professor,

Department of Applied Mathematics and Theoretical Physics

Dr Simon Thomas, Head of Radiotherapy Physics

Dr Mark Gurnell, Consultant Endocrinologist, University Lecturer

The VoxTox Research Programme, funded by Cancer Research UK, focuses on reducing toxicity from radiotherapy (RT). This would have benefits for patients and society. Additionally, individualised reduction of toxicity risk will allow dose escalation, and combination with chemotherapy. Since a strong radiotherapy dose-response relationship exists for many tumours, this will improve cancer cure rates.

Aims

- To quantify the differences between planned dose and the dose actually delivered during the whole course of radiotherapy, and to link these data to differences between expected and observed toxicity.
- To determine the relative contributions to individual variation in toxicity from underlying biological factors or physical dose variation.
- To develop a suite of integrated software tools for dose review during the treatment course, with the objective of individualising treatment, based on predicted toxicity.

Methods

We will analyse the discrepancy between RT doses which are planned, and those actually delivered, based on daily CT data acquired within our clinical programme of high precision image-guided radiotherapy. In Cambridge we have a unique archive of daily volumetric imaging of patients treated on our TomoTherapy units. We will develop systems to map the location of each point (voxel) within the patient outline, and then to re-compute the dose at that point each day during treatment. We will apply techniques from image processing, materials modeling, tumour & tissue modeling and radiation biology to construct models of cumulative dose during treatment that cannot be achieved using current techniques. To do this we will bring together a cross-disciplinary group of clinicians and clinically-orientated scientists from the University Department of Oncology, the NHS Oncology Centre, the Cavendish (Physics) Laboratory, the University Engineering Department and the CR UK Cambridge Research Institute.

How will the results be used?

The results will be used to improve radiotherapy based on individualised toxicity. More accurate data on the relative contributions to individual variation from underlying biological factors or physical dose variation will feed in to studies of genetic determinants of individual variation in toxicity (radiogenomics).

The development of an integrated suite of software tools for dose review during the treatment course will allow individual toxicity prediction and provide capability to re-plan treatment including toxicity as an optimisation parameter. Together, these tools will allow individualised treatment with reduced toxicity. The tools may be commercialisable.

Appendix 3: The MINOT-OAR sub-study

MultIple TimepoiNt MRI TO Track Organ At Risk (OAR) Changes In Patients Undergoing Radical Radiotherapy for Head and Neck Cancer

A3.1 Source of patients

Patients treated with radical radiotherapy and concomitant weekly systemic therapy for squamous cell carcinoma of the oropharynx, larynx or hypopharynx at Addenbrooke's Hospital.

A3.2 Number of patients

The MINOT-OAR sub-study will recruit up to 15 patients. This will provide sufficient training data for segmentation tools developed within VoxTox, and is powered to detect differences between baseline and mid-treatment MRI ADC at 5% alpha and 90% power. This is based on data from Juan et al, who reported a difference in ADC of 0.21 (0.81 vs 1.02, st deviation +/- 0.12) between patients with and without xerostomia.

According to inclusion and exclusion criteria, approximately 2 patients per month will be eligible for MINOT-OAR. Assuming a recruitment rate of 60% and an attrition rate of 33%, we would therefore expect target recruitment of 15 patients within 18 months.

A3.3 Inclusion Criteria

In addition to the general inclusion criteria for the VoxTox study, the following are additional inclusion criteria for the MINOT-OAR substudy

- Histologically confirmed squamous cell carcinoma of the oropharynx, larynx or hypopharynx.
- Radiotherapy treatment volumes include lymph node regions on at least one side of the neck.
- Consented to receive concomitant systemic therapy.
- Baseline performance status 0 or 1.

A3.4 Exclusion Criteria

In addition to the general exclusion criteria for the VoxTox study, the following are additional exclusion criteria for the MINOT-OAR substudy

- Surgery prior to radiotherapy as part of definitive treatment for this malignancy
    - Specifically, neck dissection with or without laryngectomy or pharyngolaryngectomy prior to radiotherapy.
    - *Staging* surgery is not considered to be an exclusion criterion. Thus general anaesthetic with pan-endoscopy, tonsillectomy and/or debulking biopsy of the primary site, or biopsy of a lymph node would not exclude patients from this sub-study.
- Patients receiving radiotherapy to the primary site only (no neck irradiation).
- Standard MRI contra-indications. Specifically these are:
    - Heart pacemakers and implantable cardiac defibrillators
    - Insulin pumps
    - Implantable hearing aids
    - Implanted aneurysm clips
    - Neurostimulators
    - Metallic bodies in the eye.
- These points should be checked and discussed with patients otherwise suitable for recruitment. Some of the devices listed are now MRI compatible and safe, and if the patient has unambiguous written advice from the device manufacturer that it is MRI compatible and safe, they may still be considered for the study. If there is any doubt then the patient should not be invited to participate in the study.

A3.5 Patient Recruitment

Eligible patients will initially be identified via the Addenbrooke's Hospital head and neck cancer MDT. Patients that require radical radiotherapy will be asked to consent and booked for this treatment as standard in the subsequent clinic (this takes place immediately after the MDT), but VoxTox (MINOT-OAR) will not be discussed at this stage. Eligible patients will initially be approached at their next visit, when they attend for their radiotherapy planning appointment. This will be done by either the study lead or a suitably trained healthcare professional from within the research group (as identified by the delegation log). The study will be discussed with patients and they will receive a copy of the patient information leaflet. When patients return for their standard pre-treatment visit, they will have a further opportunity to ask questions about the study, and give consent to participate if they wish to.

A3.6 Patient Pathway

Patients recruited to the MINOT-OAR sub-study will not be required to make additional hospital visits compared to standard clinical practice. On the days that MRIs are scheduled, patient will be asked to report to the radiotherapy department 1 hour before their scheduled treatment time. They will be met by a member of the research team, and escorted to the MRI scanner unit, which is a short walk from the radiotherapy reception. They will then undergo their study MRI, which will be undertaken by experienced MRI radiographers. Patients will then be escorted back to radiotherapy reception and treatment radiographers informed that the patient has returned from their study scan.

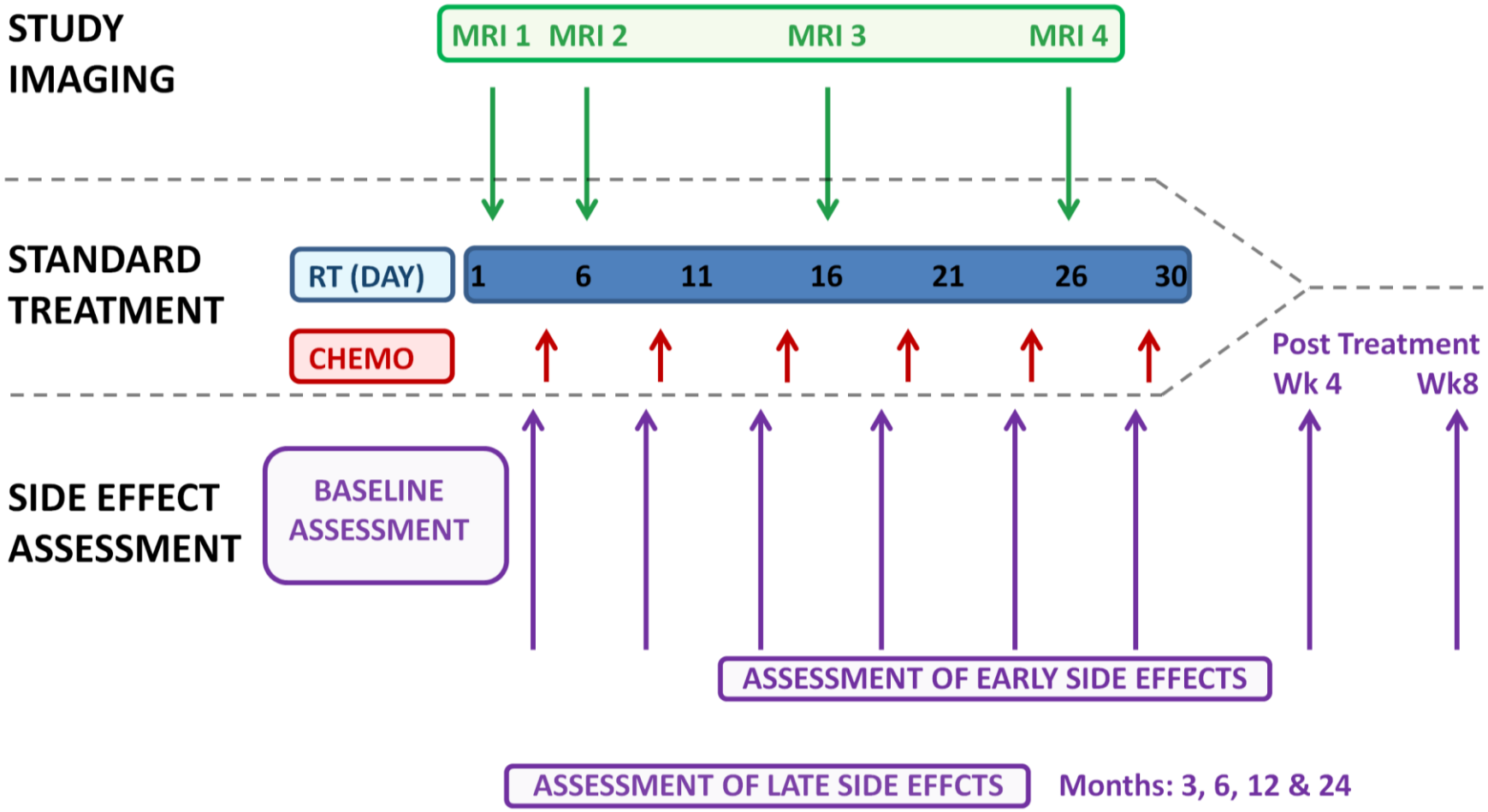
A3.7 Imaging Protocol

A3.7.1 Imaging schedule

- Patients recruited to the MINOT-OAR study will undergo 4 MRI examinations during the course of their treatment.
- The first will be done shortly before the first radiotherapy fraction is delivered.
- Scan 2 will be scheduled for week 2, scan 3 for week 4 and scan 4 for week 6.

A schema to show scheduling of imaging within the MINOT-OAR substudy is shown on the following page.

VoxTox MINOT-OAR substudy schema



A3.7.2 Imaging specifics

- Study MRI scans will initially be done in the WBIC.
- Patients will be positioned supine to match their position on the radiotherapy treatment unit as accurately as possible.
- The imaging protocol has been designed to minimise scan time to ensure that these scans cause minimal inconvenience and are well tolerated by patient recruited to MINOT-OAR. The imaging sequences will include:
    - Volumetric, axial T1 weighted images
    - Volumetric, axial T2 weighted images
    - Axial diffusion-weighted sequences
- The maximum scan time will be 30 minutes, but it is anticipated that most scans will be substantially shorter than this (15-20 minutes).
- This list is not exhaustive, and if, during the conduct of the research, new findings from our study (or others), indicate different sequences to improve our imaging protocol then these will be included. However, minimising scan time to ensure tolerability will remain a priority, and sequences requiring cannulation and/or injection will not be added except by amendment to the protocol.

A3.8 Data management

A3.8.1 Imaging data collection

- Completed scans will be uploaded to the main Addenbrooke's Hospital PACS system where they will be reviewed by a consultant radiologist specializing in head and neck disease and a governance report produced.
- Relevant data for the study will be measured directly from the PACS system, and images will be imported into the research database of the Trust radiotherapy planning system (Prosoma) for further analysis.
- If images are to be sent outside of the Trust, they will be fully anonymised and tokenized, following the procedure set out in the main protocol.

A3.8.2 MR Imaging data curation and storage

- MR images that are sent to PACS will be labeled with standard NHS identifiers.
- Each patient will have 3 study identifiers generated.
- Two of these will be VoxTox study identifiers as described in the main body of the protocol 1 specifically to MINOT-OAR (for example VT0001 and VT1_H_5E3821K1).
- A MINOT-OAR specific ID will also be generated. This will take the form MTxxx, and start sequentially from MT001, and will be allocated at consent.

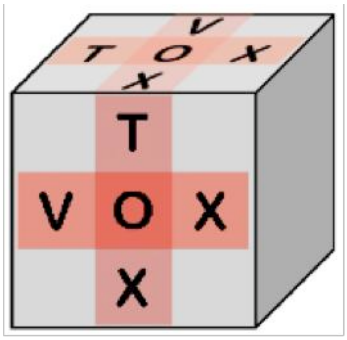
A3.9 Statistical considerations

The primary purpose of the MINOT-OAR sub-study is to augment existing techniques within the VoxTox. The anatomical MR imaging data will be used to further develop existing VoxTox tools. Volumes, distances and areas will be normalised to starting values to allow comparability between patients, and means, standard errors, standard deviations, and 95% confidence intervals of these changes will be calculated. Using lower quality CT imaging, Robar and colleagues [Robar et al. 2007] found a mean reduction in final gland volume of 25.6% with a standard deviation of 11%, and a rate of change of 0.6%/day, from a sample size of 15. Vasquez-Osorio and colleagues saw a mean reduction in parotid volume of 18% with a standard deviation of 7% from a sample size of 10. A sample size of 15 will therefore generate sufficient data to inform our model and augment the VoxTox methodology.

Secondary endpoints for the MINOT-OAR sub-study will be to establish the link between early MRI biomarkers and toxicity events. As described in sections 12.0 of the main protocol and section A3.2 of this appendix, a sample size of 15 is powered to detect previously reported differences in ADC with 5% alpha and 90% power.

VoxTox

Linking radiation at the Voxel level with Toxicity



MINOT-OAR SUB-STUDY

Using MRI scans to better understand how salivary glands respond to radiotherapy treatment

(MULTIPLE TIMEPOINT MRI TO TRACK ORGAN AT RISK (OAR) CHANGES IN PATIENTS UNDERGOING RADICAL RADIOTHERAPY FOR HEAD AND NECK CANCER)

Participant Information Sheet

- **You are being invited to take part in a research study.**
- Before deciding whether or not to take part, we want you understand why this research is being done and what it involves.
- Please take time to read the following information carefully, and talk to others about the study if you wish.
- Please ask us if anything is not clear, or if you would like more information.
- Please take time to consider whether or not you wish to take part, and discuss this family and friends if you wish.

Section 1: gives you a brief summary of of the study.

Section 2: tells you the purpose of the study and what will happen if you take part.

Section 3: gives more detailed information about the conduct of the study

CANCER RESEARCH UK | CAMBRIDGE CENTRE

UNIVERSITY OF CAMBRIDGE

Cambridge University Hospitals NHS Foundation Trust

Section 1: Study Summary

We are running this study to understand side effects from radiotherapy in more detail. The study does not involve any change to your treatment, which will be entirely standard.

We are interested in the effect that radiotherapy has on healthy tissues in the head and neck region, including the salivary glands. We would like to understand more clearly how much radiation has been delivered to these areas, and to look at side effects in detail. We will do this by studying the 'position-check' scans you have on every treatment day and collecting detailed information on any side effects that might occur. We would also like to do four short MRI scans, spaced out over treatment, to give us more information about how the radiotherapy treatment is affecting the salivary glands.

The information from this study may allow us to reduce side effects for patients treated with radiotherapy in the future. This could mean that we are able to give a higher dose of radiation to the cancer, leading to more patients being cured.

Section 2: Study Purpose, and what will happen

1. Why are we doing this research study?

- The main aim of this study is to improve our understanding of the way that salivary glands change during a course of radiotherapy.

- These glands make saliva, which makes the mouth feel moist and comfortable, and protects the teeth. Damage to these glands can cause a permanent dry mouth, thick or sticky saliva, dental and swallowing problems.

- Modern radiotherapy is carefully targeted at the tumour and lymph node regions. Although treatment is precise and accurate, the parotid and submandibular salivary glands often receive a dose of radiation that can lead to permanent damage. This complication can be uncomfortable for patients and lead to problems with the teeth, jaw bone, and swallowing.

- We think that in the future, adapting radiotherapy plans during a course of treatment may reduce the risk of these side effects. At present, we don't how to identify patients who will benefit from this approach, or when to do it. This is partly because we don't know the way in which important structures (including the salivary glands) change during treatment.

- In this study, patients will have 4 MRI scans alongside their normal radiotherapy treatment, spaced out over the 6 weeks of a normal treatment course. These scans give excellent pictures of the salivary glands, so we can make very accurate measurements of the changes that are happening to them. The scans may also give us some information of how the radiation might be affecting them, and when these changes start to occur.

- This information will help us link changes that happen during treatment with the side effects that patients suffer. We can use this information to identify people who need their radiotherapy plan adapting, and decide when to do it. We hope that this will lead to reduced side effects, and possible higher cure rates, for patients in the future who have radiotherapy for head and neck cancer.

2. Will my radiotherapy treatment be affected?

- Your treatment will be completely standard. You will receive the same radiotherapy treatment whether or not you join the study.

- You will have a CT scan before treatment called a planning scan. We use this to plan your radiotherapy, and the scan will completely standard whether you join the study or not. CT scans are built up from small 3-dimensional (3D) units called voxels. Each voxel represents a tiny cube of tissue in you body, similar to the way that photographs are made up of many 2-dimensional (2-D) pixels.

- Your treatment will be delivered using a technique called Intensity-Modulated Radiotherapy (IMRT). This uses radiation beams from lots of different angles to target the tumour accurately, and keep the radiation dose to important normal organs as low as possible.

- Your radiotherapy will be delivered on a treatment machine that can also perform a type of CT scan. These scans are used to check that you are in the correct position for treatment. This is called Image-Guided Radiotherapy, and you will have a positional CT scan every day whether or not you join the study.

- Patients in the study will have positional CT scans with a few more slices (images), then those who are not. We will use them to study the radiation dose delivered to each voxel in the scan.

- These improved positional CT scans do not increase overall appointment time. They will expose you to a slightly higher radiation dose. This dose is about 600 times less than the total radiotherapy dose (about 0.2%), and this slight increase is very unlikely to cause you harm. The size of positional CT scans we are using in this study is routine practice in many other hospitals.

3. What is an MRI Scan?

- MRI stands for Magnetic Resonance Imaging. MRI scanners use strong magnetic fields and radiowaves to excite hydrogen atoms. Sensitive detectors pick up differences in this effect between different structures within your body. An advanced computing system then produces images from these differences. MRI scans are very good at producing detailed images of soft tissue such as glands, muscle, fat, blood vessels and skin. This makes them extremely useful for this type of study.

4. What happens during the MRI?

- You will be taken from the radiotherapy department to the MRI scanner, which is nearby. You will be asked to lie down on the scanner couch and made comfortable with pads and pillows. It is extremely important to stay as still as possible during the scan, so tell the radiographers if you're not as comfortable as you would like.

- The scans being done in this study do not need contrast medium (dye), so you will not need any injections for these scans.

- Once you are comfortable on the couch, the radiographers will leave the scanner room and move next door to the control room. They will be watching you all the time, and you will be able to talk to them via the intercom.

- The couch will slowly move inside the scanner. During the scan, the machine may well be very noisy and this is completely normal. You will be given ear-plugs and/or headphones, but if you feel worried or uncomfortable, tell the radiographers immediately.

- You may have had an MRI scan as part of the investigation of your cancer. The procedure for the study scans will be similar, but these study scans are likely to be shorter.

- MRI scans in the study are permitted to take up to half an hour, but most will be much shorter than this (around 15-20 minutes).

- On the days that you have MRI scans, we recommend arriving at radiotherapy reception an hour before your scheduled treatment time. This is to make sure there is plenty of time to take you to the MRI scanner, and be back in time for your treatment.

- You are welcome to bring someone with you to the scan appointment (and your radiotherapy treatment), but for safety reasons, this person would not be allowed into either the scan room or the radiotherapy treatment room.

5. Are there any risks or side effects from MRI scans?

- As far as is known at present, MRI scans are extremely safe. They do not use X-Rays, so there is no radiation exposure from having an MRI scan.

- Because the scanner works by making a strong magnetic field, it is very important not to take any metal objects into the scanner room. The radiographers will discuss this with you in detail when you come for your scans.

- Patients who have hearing aids, pacemakers, metallic clips in the brain, metallic heart valves, or metal fragments in their eyes should not normally have MRI scans. If this applies to you, you should not enrol in the study. We will discuss this with you, so please ask if you're not sure, or have any questions.

- There are no side effects from MRI scans.

6. Why have I been invited to take part?

- You have been invited to participate in this study because you are about to start radiotherapy for a head and neck cancer at Addenbrooke's Hospital.

- We plan to recruit about 15 patients to this study over the next year to 18 months. To minimise disruption for patients participating in the study, and to get the best scientific information, only certain patients having radiotherapy for head and neck cancer are eligible.

7. Do I have to take part?

- No, participating in this study is completely voluntary. If you decide to participate, you will be asked to sign the Informed Consent Form at the end of the information sheet. However, you are still free to change your mind, and leave the study at any time, without giving a reason. If you choose not to participate, or to leave the study, your future medical treatment and standard of care will not be affected in any way.

8. When do the MRI scans happen, and what will happen if I decide to take part?

- Your radiotherapy treatment takes 6 weeks to deliver. You will have treatment on Monday to Friday, but not Saturday or Sunday. This schedule will not be affected if you decide to participate in the study.

- All patients invited to participate in this study have been offered low dose weekly chemotherapy to maximise the chance of being cured. This aspect of your treatment is standard care, and will not be affected by participating in the study. Chemotherapy treatments are normally delivered on Thursdays and Fridays.

- If you decide you would like to participate in the study, your MRI scans will normally be done on either a Monday, or a Tuesday to ensure that you will not need to have chemotherapy and an MRI scan on the same day.

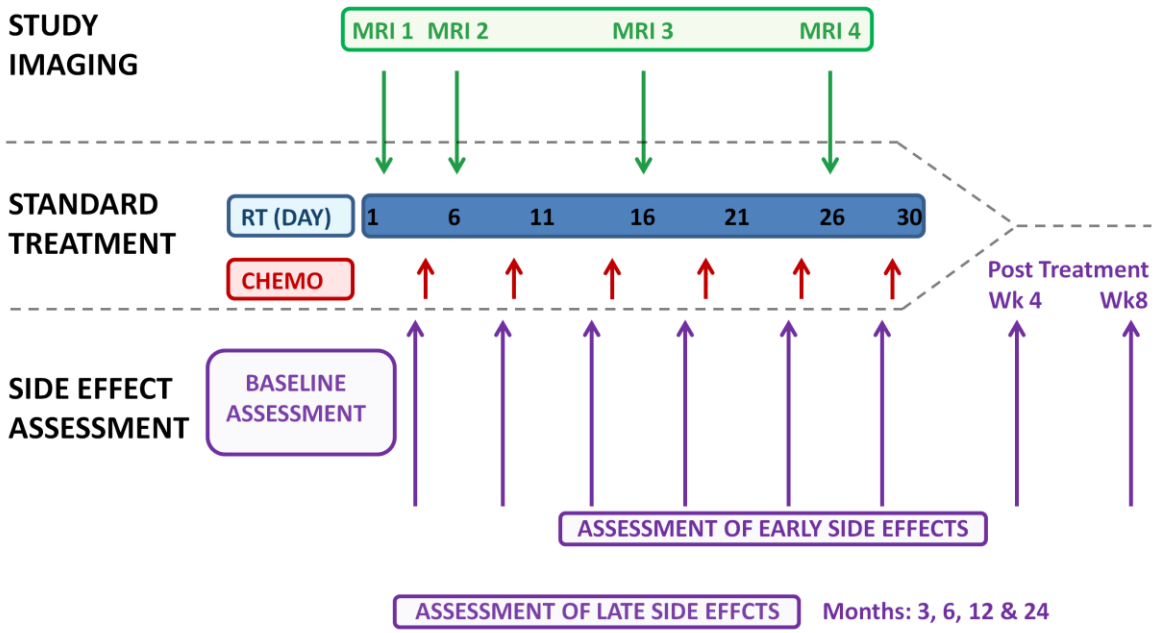
- MRI scans will be done at the beginning of the first, second, fourth and final week of treatment (weeks 1, 2, 4 and 6). This will mean the scans will be done at the following times;
    - Scan 1 – before treatment starts.
    - Scan 2 – week 2 (treatment days 6-10).
    - Scan 4 – week 4 (treatment days 16 to 20).
    - Scan 6 – week 6 (treatment days 26 to 30).

- It is important to look for a link between what we see on the scans, and the treatment side effects you experience. You will be seen in clinic before treatment starts, and once a week during treatment, whether you decide to participate in the study or not. If you decide to take part, the pre-treatment visit will take a little longer (10-15 minutes) as will need to make a careful assessment of whether or not you have problems with dry mouth or swallowing before we start treatment. At the weekly clinic visits during treatment, you will be asked detailed questions about treatment side effects, but these visits should not take any longer than normal. You will also be seen 4 and 8 weeks after treatment finishes, to check that early side effects are settling.

- Once you have completed radiotherapy treatment, your clinic follow up with surgeons and oncologists in clinic 10 will proceed as normal. If you decide to take part in the study, we will also make more detailed assessments of on-going side effects after treatment has finished. This will happen at months 3, 6, 12 and 24 after the end of treatment. These assessments will be done on days when you are coming for a follow-up visit to clinic 10 anyway. Assessments will mean an extra consultation on these clinic days, and we anticipate your total appointment time will take about half an hour longer than normal. If you are unable to come to the hospital when your follow up is due then we may ask you to take part in an interview over the phone.

- On the next page is a diagram, showing how the standard treatment schedule works, and where the MRI scans and side effect assessment schedule will fit into this.

STUDY DESIGN



9. What are the possible disadvantages and risks of taking part?

- Your hospital visits will be longer on the treatment days when you have MRI scans. We anticipate the visit being about an hour longer. You will not need to make any extra journeys to the hospital to have these scans.

- There are no side effects or risks to health associated with having MRI scans. The scanner is quite enclosed and can be noisy. Some patients can feel a little claustrophobic inside an MRI scanner, but most don't find it a problem. If you have a history of claustrophobia we will need to talk to you about this carefully before you decide to participate.

- The positional CT scans have a few more slices than normal scans, meaning a tiny increase in radiation dose. This means that there is a very rare risk of future cancers due to the slightly longer scan. The risk is unlikely to be very different from standard CT scans, but may affect 1 extra patient in 20,000 over the rest of their life.

- Interviews to collect information about side effects may take a little longer than standard clinic visits. Most of these interviews will be organised on days when you are attending the hospital anyway, for treatment or follow up. The assessment on week 4 after the end of treatment may not coincide with a standard follow up visit, and this would require an extra hospital visit.

10. What are the possible benefits of taking part?

- We hope that the information collected in this study will be of benefit to patients with head and neck cancer in the future.

- You are unlikely to benefit personally from being involved in the study. We may be able to see the tumour or cancerous lymph nodes on your MRI scans, but they will not give us information about your progress or prognosis. A radiologist will check these scans, but this is to ensure that there are no new problems that we didn't know about before.

11. What happens when the study stops?

- When the study stops, your appointments will be exactly the same as if you had not taken part. We may ask permission to contact you about taking part in future research.

12. Expenses and payment?

- We regret that we are unable to offer any payment for participating in this study.

Section 3: Study conduct

13. What if new information becomes available?

- Sometimes during the course of a study new information becomes available that might affect your decision to continue participating in the study. Your study doctor will contact you to discuss the new information and whether you wish to continue participating in the study. If you still wish to continue on the study, you will be asked to sign a new Informed Consent Form.

- The study sponsor or doctor may decide to stop the study at any time. If that happens we will tell you why the study has been stopped.

14. What if I decide I no longer wish to participate in the study?

- You are free to stop taking part in this study at any time without giving a reason and without affecting your future care or medical treatment. If you decide not to participate any further then your care will go back to being as it would have been if you had not joined the study. Any information we have already gathered will continue to be used, but no further information will be collected.

- The study doctor may also choose to withdraw you from the study if they feel it is in your best interests or if you have been unable to comply with the requirements of the study. If this happens we will tell you why the study has been stopped.

15. What if there is a problem?

- **I**f something does go wrong and taking part in the research harms you and this is due to someone's negligence then you may have grounds for legal action for compensation against Addenbrooke's Hospital or the University of Cambridge. The normal National Health Service complaints mechanism will still be available to you (if appropriate). In addition, the University has insurance that provides no-fault compensation for non-negligent harm: you may be entitled to make a claim for this.

- If you wish to complain or have concerns about any aspect of the way you have been approached or treated during this study, you use the NHS complaints procedure. In the first instance it may be helpful to contact the Patient Advice and Liaison Service (PALS) on 01223216756.

16. Will my taking part in this study be kept confidential?

- All information collected about you as a result of your participation in the study will be kept strictly confidential. Your personal and medical information will be kept in a secure file and will be treated in the strictest confidence. You may ask to see your personal information at any time and correct any errors if necessary.

- We would like to inform your GP of your participation in this study.

- Authorised staff, who work for, or with, the sponsor of the study, or the hospital R&D Department, may require access to your personal information and/or medical records to verify the data for this study. All the information will be treated in the strictest confidence during the review process.

- If in the future we think that our knowledge and understanding of cancer could be advanced by combining information from this study with information collected from other studies, we would like to ask your permission to do this. If this happens, information about you may be passed on to other researchers, but they would not be able to identify you from that information.

17. What will happen to the results of the study?

- The results of the study will be anonymous. It will not be possible to identify you from any of the data that are produced. When the results of this study are available they may be published in a peer-reviewed medical journals and used for presentations at medical conferences.

- If you would like to obtain a copy of the published results please contact you study doctor who will be able to arrange this for you.

18. Who is organising (sponsoring) and funding the study?

- Cambridge University Hospitals NHS Foundation Trust and the University of Cambridge are jointly sponsoring this study.

- Cancer Research UK is funding the study.

19. Who has reviewed this study?

- All research within the NHS is reviewed by an independent group of people called a Research Ethics Committee, to protect your interests. The NRES Committee East of England – Essex Research Ethics Committee has approved this study. Cancer Research UK has also reviewed the study.

20. Further information and contact details

Clinical Research Fellow: David Noble, Dept. of Oncology
Working Hours: 0900 – 1800, Monday to Friday
Telephone Number: 01223769306
Email: david.noble@addenbrookes.nhs.uk

Research Radiographer: Amy Bates, Cambridge Cancer Trials Centre
Working hours: 0830-1630
Telephone Number: 01223216083
Email: cctc@addenbrookes.nhs.uk

You may find the following websites helpful:

www.compRT.org/minotoar

www.macmillan.org.uk/information-and-support/treating/radiotherapy/external-beam-radiotherapy-explained/types-of-external-beam-radiotherapy.html

https://be.macmillan.org.uk/be/p-23359-having-an-mri-scan.aspx

VoxTox Head & Neck (MINOT-OAR sub-study)

INFORMED CONSENT FORM

Study Title: VoxTox - linking radiation dose at the Voxel level with Toxicity
Study IRAS ID: 109324
Principal Investigator: Professor Neil Burnet

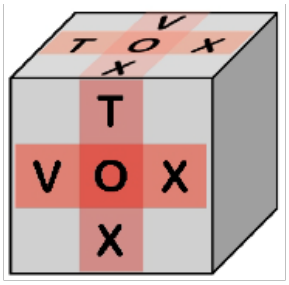
Participant Number:

If you agree with each statement below, please initial the box.	**INITIALS**
1. I have read and understood the VoxTox Head and Neck (MINOT-OAR sub-study) Patient Information Sheet Version 2.0 for the above study and I confirm that the study procedures and information have been explained to me. I have had the opportunity to ask questions and I am satisfied with the answers and explanations provided.	
2. I understand that my participation in this study is voluntary and that I am free to withdraw at any time, without giving a reason and without my medical care or legal rights being affected.	
3. I understand that sections of my medical notes or information related directly to my participation in this study may be looked at by responsible individuals from the sponsor, regulatory authorities and research personnel where it is relevant to my taking part in research. I give permission for these individuals to have access to my records.	
4. I give permission for my GP to be informed of my participation in this study.	
5. I understand that the doctors in charge of this study may close the study, or stop my participation in it at any time.	
6. I agree to participate in this study.	
7. I understand that the sponsors of this study may make my data available to other researchers for future ethically approved research. Any data transferred to a third party for future research will be anonymous and will not contain my personal information. I give permission for these individuals to have access to my data.	

Name of patient	Signature	Date
Name of person taking consent	Signature	Date
Name of researcher if different from above	Signature	Date

1 copy for the patient, 1 copy for the study team, 1 copy to be retained in the hospital notes.

Cambridge University Hospitals NHS
NHS Foundation Trust



VoxTox
Cambridge Cancer Trials Centre
S4, Box 279
Cambridge University Hospitals NHS Foundation Trust
Addenbrooke's Hospital, Hills Road
Cambridge
CB2 0QQ

Date:

VoxTox (MINOT-OAR substudy)

Linking radiation dose at the Voxel level with Toxicity

An observational study to collect detailed information about side effects for patients undergoing image guided intensity-modulated radiotherapy to the head and neck, prostate, and central nervous system

Dear Dr

VoxTox : linking radiation dose at the voxel level with toxicity

Name:
Date of birth:
Address:

Your patient has been invited to participate in an observational radiotherapy study which aims to collect comprehensive toxicity data for patients who have undergone image guided intensity-modulated radiotherapy for cancer of the head & neck. In addition to standard treatment, this will involve the collection of detailed toxicity information, and 4 additional non-diagnostic MRI scans taken during their course of radiotherapy.

This is a study at Addenbrooke's Hospital. I enclose a copy of the information sheet given to the patient. If you want further details about this study please contact me at the above address, or view the following website: www.voxtox.org.

If any of the contact details shown above are incorrect, please inform us.

Yours sincerely,

Professor Neil Burnet
Consultant Oncologist and Principal Investigator