Dosimetric and Radiobiological Fitting of Xerostomia and Dysphagia 12 Months After Treatment for Head and Neck Tumors

by

Alexander Aronoff Kubli

Graduate Program in Medical Physics
Duke University

Date:_____

Approved:

_____
Anuj Kapadia, Chair

_____
Shiva Das, Supervisor

_____
Panayiotis Mavroidis

_____
Yvonne Mowery

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Graduate Program in Medical Physics in the Graduate School
of Duke University

2018

ABSTRACT

Dosimetric and Radiobiological Fitting of Xerostomia and Dysphagia 12 Months After Treatment for Head and Neck Tumors

by

Alexander Aronoff Kubli

Graduate Program in Medical Physics
Duke University

Date:_____
Approved:

_____
Anuj Kapadia, Chair

_____
Shiva Das, Supervisor

_____
Panayiotis Mavroidis

_____
Yvonne Mowery

An abstract of a thesis submitted in partial
fulfillment of the requirements for the degree
of Master of Science in the Graduate Program of Medical Physics
in the Graduate School in
Duke University

2018

Abstract

Oropharyngeal Squamous Cell Carcinoma (OPSCC) is by far the most predominant form of head and neck cancer in the United States. The survival rate for OPSCC is very high, which, while fortunate, yields many patients who are left with the late term toxicities consequent of their treatment. This thesis aimed to use patient-reported outcome (PRO) data from two sources – the PRO-CTCAE and the QLQ-C30 – along with the dosimetric data of patients that have already been treated, in order to characterize retrospectively a relationship between patient dosimetric data and the severity of response of PRO data. In particular, PRO data was used as a way to characterize the severity of patient-experienced xerostomia and dysphagia. Additionally, this data was used to fit the radiobiological parameters for two normal tissue complication probability (NTCP) models: the Lyman-Kutcher-Burman (LKB) model, and the Relative Seriality (RS) model. Overall, it was found that the PRO-CTCAE data was more robust than the QLQ-C30 data in its characterization. Based on the PRO-CTCAE data, the V_{52} (volume which receives at least 52 Gy) of the combined constrictors and the V_{59} of the superior pharyngeal constrictor show the strongest relationship with patient-reported dysphagia. Additionally, the V_{27} of the contralaterals and the V_{12} of the contralateral parotid show the strongest relationship with patient-reported xerostomia. Furthermore, it was found that the dose response curves for both NTCP models fit the data with similar accuracy.

Contents

List of Tables

List of Figures

1. INTRODUCTION

1.1 Background of Head and Neck Cancer

Head and neck cancer accounts for 3-4% of all cancers in the United States (Cancer.net, 2017). In 2014, there were about 15,000,000 people in the United States with cancer (seer.cancer.gov, 2014). Using that figure as an estimate of the cancer incidence across the country, there are about 500,000 people in the country with head and neck cancer. The vast majority of head and neck cancer is oropharyngeal squamous cell carcinoma (OPSCC). The 5-year survival rate for OPSCC has been steadily increasing over time; from 1992 to 2002 the survival rate increased from 55% to 66% (Ringash, 2015); since then, the survival rate has increased to over 72% (seer.cancer.gov, 2013). Using these rough numbers with a conservative estimate (the survival rate from 2013), roughly 350,000 of the 500,000 people suffering from OPSCC in the United States will survive their treatment. Beyond this, the incidence of OPSCC in the country has also been increasing over time (seer.cancer.gov, 2013). Since 2000, the rate of diagnosed OPSCC has increased by about 66% (seer.cancer.gov, 2013). Overall, such a high survival rate combined with the high number of number of OPSCC patients yields a large number of people who will have to persist with the late term complications they develop post-treatment. Consequently, it is important to establish the late term toxicities most commonly associated with head and neck cancer treatment.

1.2 Late Term Complications

There are a handful of common side effects for head and neck cancer patients that have been treated with radiation therapy. These side effects include dry mouth, mouth sores, difficulty swallowing, jaw stiffness, nausea, and tooth decay (cancer.net, 2018). These side effects are typically much more severe than the normal onset of these effects in people who have not been treated with radiation therapy; for example, patient dry mouth in OPSCC patients often leads to all teeth falling out (Florida, 2018). In pursuit of a way to characterize the late term complications experienced by head and neck cancer patients, several publications will be highlighted. A study looking at 230 patients across multiple Radiation Therapy Oncology Group (RTOG) trials found that 43% of patients suffered from severe late pharyngeal/laryngeal toxicity post-treatment (Machtay et al., 2008). For this study, severe late toxicity was defined as chronic grade 3 or 4 toxicity, using the RTOG late toxicity scoring system. Physiologically speaking, this means that patients required a feeding tube or worse (Machtay et al., 2008). A publication put out by the MD Anderson Cancer Center (MDACC) looked at the last follow-up for 29 patients (an average of 10 months post-treatment). At the last follow-up, 19/29 (66%) of them were gastrostomy-dependent, and 23/29 (79%) of them suffered from persistent aspiration, which is a condition where foreign objects (food, liquids, etc.) enter the windpipe and induce a coughing or wheezing effect (Hutcheson et al., 2012). A multi-institutional study between the Memorial Sloan Kettering Cancer Center, Princess Margaret Hospital, and MDACC looked at 2315 patients being treated for OPSCC. Of

those patients, 63% needed a gastrostomy tube during treatment. After one year, that number dropped to 7%. After two years, it was at 3.2% (Setton et al., 2015). Although there is a steep reduction in the percentage of patients that need a gastrostomy tube as time since treatment increases, the rate of tube use starts very high and after a couple of years remains necessary for a non-trivial number of patients. Therefore, there are significant issues for patients that are treated for their OPSCC -- the two most common of which are dysphagia (difficulty swallowing) and xerostomia (dry mouth) (Cancer.net, 2017).

Xerostomia is most largely associated with the radiation dose to the salivary glands. High amounts of radiation to the salivary glands damage their function, and lead to reduced salivary production. Sparing the parotid glands and the submandibular glands (SMG) can have a large effect on the rate of xerostomia post-treatment. Additionally, the mean parotid dose is thought to be most closely associated with salivary production. By sparing at least one parotid gland and at least one SMG – the contralateral salivary glands – the risk of xerostomia can be significantly decreased. In particular, at least one parotid gland should receive a mean dose of less than 20 Gy, or both glands should receive a mean dose of less than 25 Gy (Deasy et al., 2010).

Dysphagia is a large issue with patients being treated for head and neck cancer. A publication from the Journal of Surgical Oncology looked at patients who were treated for squamous cell cancer (Belcher, Hayes, Fedewa, & Chen, 2014). All patients were given a gastrostomy tube pre-treatment as a pre-emptive measure for inevitable

dysphagia brought on by the therapy treatment. The publication looked at the relationship between the pharyngeal constrictors and dysphagia, and found that there was a relationship between the inferior pharyngeal constrictor (IPC) and post-treatment dysphagia. Notably, the V_{65} for the IPC should be less than 15%, the V_{60} should be less than 40%, and the mean dose to the IPC should be less than 55 Gy (Belcher et al., 2014). According to the data used for this publication, adhering to these IPC dose constraints yields the largest reduction of dysphasia post-treatment.

1.2 Traditional OPSCC Treatment

Traditional OPSCC treatment is typically treated to a prescription of about 70 Gy at around 2 Gy per fraction (Joseph O. Deasy, 2010). To reduce the rate of severe xerostomia in patients, the recommended dose volume limits for a traditional treatment are: (1) at least one parotid should receive a mean dose of less than 20 Gy, or (2) both glands should be spared to a mean dose of less than 25 Gy (Joseph O. Deasy, 2010). To reduce the rate of dysphagia, the volumes of the pharyngeal constrictors receiving less than 60 Gy and less than 50 Gy should be minimized (Tiziana Rancati, 2010).

1.3 HPV-Mediated OPSCC

Traditionally, the patient demographics most closely associated with head and neck cancer are older, white men with a large amount of smoking and alcohol exposure (Deschler, Richmon, Khariwala, Ferris, & Wang, 2014). Patients with head and neck cancer are typically treated with 70 Gy of intensity-modulated radiation therapy (IMRT) (Deschler et al., 2014). However, with the decrease in smoking rates across the country

and the overall increase in HPV transmission – particularly with younger people – a new type of head and neck cancer is becoming the predominant "face" of the disease.

HPV-mediated OPSCC differs from traditional head and neck cancer in two ways: first, the underlying physiological mechanisms for HPV-mediated OPSCC are different than they are for traditional OPSCC; second, the patient demographics for this HPV-mediated cancer are different – HPV-mediated OPSCC tends to affect younger white men with little tobacco or alcohol exposure (Deschler et al., 2014). A consequence of the younger patient population is an overall increase in survival rate, as younger people tend to live longer post-treatment than older people. Fortunately, HPV-mediated OPSCC has seen better outcomes than traditional OPSCC with regards to tumor control. Consequently, de-intensification of treatment for HPV-mediated OPSCC has been proposed. A fundamental question being asked is, "Because HPV-mediated OPSCC is more responsive to radiation treatment, could the total treatment prescription be reduced in order to reduce some of the late term effects that patients experience after being treated?" (Lewis, Kang, Levine, & Maghami, 2015).

1.4 De-escalation of Treatment

Research involving the de-escalation of HPV-mediated OPSCC typically follows two aims: (1) with the implementation of a de-escalated treatment schema, determine the best dose constraints for critical structures at risk; (2) in pursuit of new dose constraints, determine which types of data should serve as markers for the clinical endpoints of dysphagia and xerostomia.

A publication looking at the psychological needs of 127 head and neck cancer patients post-treatment found that 68% of the patients had unmet psychological needs after treatment. Additionally, a quarter of the patients had a clinically significant Hospital Anxiety and Depression Scale (HADS) score. The HADS is patient-reported questionnaire aimed at scoring patient anxiety and depression. Questions on the HADS include, "I still enjoy the things I used to enjoy," and, "I feel as if I am slowed down." Overall, the publication found the "overwhelming presence of unmet psychological needs" for patients after their treatment (Henry et al., 2014). Patient-reported scoring systems like the HADS highlight the large unmet needs that patients see in themselves post-treatment. Consequently, using patient-reported scoring as a metric for the severity of dysphagia and xerostomia may provide a robust way to determine dose constraints to critical organs and to meet patient needs after treatment.

Toward these aims, a publication from the University of North Carolina – Chapel Hill (UNC-CH) looked at the dosimetric data for 45 patients. Using this dosimetric data in conjunction with patient-reported outcome data, the group found that the V_{15} (the volume of the region that receives 15 Gy) of the combined contralaterals was most closely associated with the onset of patient-reported xerostomia, and that the V_{55} – V_{60} of the superior pharyngeal constrictor (SPC) was most closely associated with the onset of patient-reported dysphagia (Chera, 2017). A second publication from UNC-CH used the same data in order to fit the radiobiological parameters for different common normal tissue complication probability (NTCP) models. This publication also found that the

combined contralaterals were associated with xerostomia and that the SPC was associated with dysphagia. Additionally, it found that all of the tested NTCP models fit the data equally well (Mavroidis et al., 2017).

1.5 Purpose of This Research

The aim of this research is to extend the previous UNC-CH publications using a larger patient cohort (initial cohort = 157). This provides a way to validate the results of the previous publications with more patients in order to provide more statistical power to the results. The dominant questions that are addressed through this work are: (1) Are the results found for this thesis consistent with the results of the previous publications? (2) Are there specific structures or combination of structures that are closely associated with the patient reported endpoints of dysphagia and xerostomia? (3) Are there dose constraints that are compatible to be used in the treatment plan optimization of de-intensified head & neck RT in conjunction with expressing treatment outcome using the PRO-CTCAE scoring system? (4) How well do different NTCP models – functions that describe the complication probability for a given tissue as a result of dose given to the tissue – fit the data, and how do the fits compare between models? Fundamentally, many patients are treated for OPSCC through radiation therapy. Most of these patients survive, and typically have to exist with very severe late-term complications as a result of their radiation therapy treatment. The overarching purpose of this work is to help determine better ways to perform IMRT treatment on these patients in order to reduce the rate and severity of some of the most common side effects that patients experience.

2. MATERIALS AND METHODS

2.1 Patient Profile

The thesis consisted of an initial cohort of 157 patients across three studies. These patients had been previously treated at three different sites: University of North Carolina (Chapel Hill, NC), University of Florida (Gainesville, FL), and University of North Carolina-Rex hospital (Raleigh, NC). All patients were favorable risk (T0 to T3, N0 to N2c, M0, i.e. cancers with variable tumor size and regional node involvement, but no distant metastasis), HPV positive, and were treated with a de-intensified 60 Gy intensity modulated radiation therapy (IMRT) plan. Patients were treated over a concurrent 6-week chemo-radiation therapy schema at 2 Gy per fraction over 30 fractions, with a weekly cisplatin chemotherapy schema of 30 mg/m^2. 54 Gy was delivered to regions at risk for subclinical disease (e.g. lymph node basins in the head and neck area). Patients were at least 18 years of age, with minimal smoking history.

For each patient, four types of clinical outcome data were used as ways to describe the severity of dysphagia and/or xerostomia for each patient: (1) the patient-reported version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE), (2) the EORTC questionnaire to assess the quality of life for cancer patients (QLQ-C30), (3) the physician reported CTCAE, and (4) a dysphagia assessment questionnaire (EAT-10). For each patient, three time points of data were acquired: (1) baseline data taken pre-treatment, (2) 6-month follow-up data, and (3) 12-month follow-up data.

2.2 Contouring in Raystation

Using Raystation (Raysearch Laboratories, Stockholm, Sweden) seven critical structures were contoured for each patient: right and left parotids, right and left submandibular glands (SMG), and the superior, medial, and inferior pharyngeal constrictors (SPC, MPC, and IPC) as separate structures. The contouring boundaries were determined by training done under a certified clinical medical physicist; contours were performed predominantly based on training rather than on specific anatomical borders. Regions that were contoured during treatment were preferentially used (i.e. they were not re-contoured). Some patients were not contoured because their data had not been uploaded to the UNC-CH Raystation data servers. Patients were initially contoured manually. However, an atlas of 30 manually contoured patients was created in order to contour the majority of the structures. The automatic segmentations were then checked and corrected by hand. The accuracy was spot checked by a certified clinical medical physicist.

Once all of the contouring was completed, relative dose volume histograms (DVHs) for each patient were exported through Raystation. Because patients had already been treated, they had pre-existing plans, and dose calculations had already been performed. Consequently, calculation of the DVHs for the freshly-contoured regions was a straightforward sum of the dose to each voxel for the new regions. Of the 157 patients that had been contoured, 125 had a usable, pre-existing plan for which the dose calculations existed and for which the DVHs could be generated. The remaining

patients had patient data uploaded to the UNC-CH data servers, but did not have full treatment plans with dose calculations uploaded to those servers.

2.3 Absolute Volumes

In order to create new regions after DVHs were exported, absolute volumes were pulled for each region and for each patient. A previously written Raystation script was modified in order to acquire these volumes (in cubic centimeters) for each patient.

2.4 Managing and Organizing DVH Data

The R programming language was used for the majority of data handling for this thesis. First, the DVH file for each patient was imported into a common dataframe. Because the Raystation 'Export DVH' function pulls the DVHs for every contoured region (providing many more than the seven regions of interest), a subset of that dataframe was filtered, removing the excess regions.

Then, the relative volumes for each DVH were scaled by the absolute volume for each region, producing absolute DVHs for each region for each patient. Using these absolute values, several new DVHs were generated by summing the absolute volumes for a given dose value across multiple regions. Specifically: the parotids were combined into a single structure; the contralateral parotid and contralateral SMG were combined into a combined contralateral region; both parotids and SMGs were combined into a salivary glands region; the SPC and MPC were combined into a joint SPC-MPC region; the entire pharyngeal constrictor was created by joining the SPC, MPC, and IPC. The anatomical definitions of the three sections of pharyngeal constrictors are described. A)

SPC: Cranial: caudal tips of pterygoid plates; Caudal: superior end of hyoid bone, B) MPC: Cranial: upper edge of hyoid bone; Caudal: lower edge of hyoid bone, and C) IPC: Cranial: lower edge of hyoid bone; Caudal: lower edge of cricoid cartilage.

Additionally, bilateral regions were designated as either contralateral or ipsilateral (rather than as left or right). The mean dose was used as an indication of which designation a region should receive – a higher mean dose indicated the ipsilateral side, and a lower mean dose indicated the contralateral side. Physically, the ipsilateral side corresponded to the side of the primary tumor, and the contralateral side corresponded to the side opposite the primary tumor.

2.5 Managing and Organizing Clinical Outcome Data

Of the four sets of clinical outcome data, this thesis looked predominantly at the PRO-CTCAE and the QLQ-C30 as indicators of the severity of our clinical end points.

The original format of the clinical outcome data did not have follow-up time points designated. Each row corresponded to a given follow-up examination for a patient, where a patient had a row for each time he or she provided follow-up data. From this data, the baseline, 6-month, and 12-month time points were determined. The baseline data was defined as the closest report of data prior to the treatment start. The 6 and 12-month follow-up data was defined as the time point closest to 6 or 12-months after the treatment end date for the patient. Typically, a 90-day threshold was applied in order to ensure that reasonable time points were selected. However, a small number of

follow-up time points were selected manually, for cases where there was no follow up within a 90-day window, but for which a reasonable choice still existed.

Although the clinical outcome data included many parameters, specific parameters were selected based on their association with either xerostomia or dysphagia. From the PRO-CTCAE data, "difficulty swallowing" was associated with dysphagia and "severity of your dry mouth" was associated with xerostomia. For the QLQ-C30 data, dysphagia scoring was performed by combining four columns: "Have you had problems swallowing liquids," "Have you had problems swallowing pureed food," "Have you had problems swallowing solid food," "Have you choked when swallowing?" These four scores were normalized based on the standard QLQ-C30 scoring methodology, where Score = [(average – 1)/range] * 100. The xerostomia score for the QLQ-C30 was simply the result from the column, "Have you had a dry mouth?"

Once the baseline, 6, and 12-month scores for each patient were determined, a threshold was applied for each set of clinical outcome data in order to categorize patients into groups, based on whether or not they exceeded the threshold toward our clinical endpoints. Thresholding was predominantly determined by the overall scoring range for a given questionnaire, and was set in order to provide a good delineation between less severe and more severe outcomes post-treatment. For the PRO-CTCAE data and the xerostomia score for the QLQ-C30, this threshold was set to 2. For the dysphagia score in the QLQ-C30, this threshold was set to 10. Then, for a given time point, if the net difference between a patient's baseline score and the score for that time

point was ≥ the threshold, that patient was considered to have a clinically relevant response to treatment.

2.6 Analysis

Of the initial 176 patients, 83 fit the criteria for our analysis. The loss of patients in our cohort was dominated by two factors: (1) the lack of dosimetric data for patients whose plans had not been imported into Raystation; (2) a lack of follow-up data for one of the studies, due to the cohort for that study being treated too recently to have much follow-up data. The 12-month time point for patient-reported outcome data was used for the analysis.

2.6.1 Dosimetric Data vs. PRO Data

A relationship between dosimetric data and patient reported outcome data was determined by performing many receiver operating characteristic (ROC) analyses. For each set of patient reported outcome data and for each region, an ROC curve was generated for dose-volume points (V_D) ranging from 5 Gy to 80 Gy (at 1 Gy increments). Generally, an ROC analysis is a way to characterize how well a given test can predict a given outcome. As a figure of merit for this ROC information, the area under the curve (AUC) score for a given V_D was used. This was calculated for a given V_D by taking the area under the corresponding ROC curve.

2.6.2 NTCP Analysis

Two normal tissue complication probability (NTCP) models were fit for this thesis: (1) the Lyman-Kutcher-Burman (LKB) model (Appendix A); (2) the Relative Seriality (RS) model (Appendix B). The purpose of a NTCP model is to provide dose response curves for critical structures at risk. At a given dose amount, an NTCP model provides the probability that the critical structure will experience complications post-irradiation. For this thesis, 'complication' was defined as exceeding the threshold toward a clinical endpoint, based on the patient-reported outcome data.

Three parameters were fit for each model (Equations 1-3 for the LKB model, and equations 4-5 for the RS model). For the LKB model: (1) TD50, or the dose given to the organ volume that results in 50% complication probability; (2) m, a measure of the steepness of the dose response curve; (3) n, the volume effect of the equivalent uniform dose across the volume. For the RS model: D50, the dose given to the organ volume that results in 50% complication probability; (2) γ, the steepness of the sigmoid part of the dose response curve; (3) s, a measure of the relative seriality of the structure. For reference, seriality is important because it implies what kinds of dose damage are important to a structure. Damage to a single point of a serial structure can lead to a loss in overall function of the structure, because serial structures have a very linear function. Contrastingly, damage to a single point in a parallel structure typically does not lead to overall loss of function for that structure; parallel structures are much more susceptible to overall damage to the structure because they have a much less linear function than

serial structures. As a consequence, parallel structures are more susceptible to the mean dose to the structure.

After the radiobiological parameters were fit for each model, the area under the curve was taken for each dose response curve in order to determine the AUC value for each model. Then, the odds ratio (OR) was determined for each model as a way to associate how much more likely a given patient would be to meet our clinical endpoint, given that the dose to a structure is above a given dose threshold for the model.

Finally, a linear regression was performed on these dose response curves in order to determine dose constraints for the dose-volume points that best corresponded to whether or not the thresholds toward our clinical endpoints were exceeded. In this way, dose constraints were determined for 10, 15, and 20% complication probability based both on the best V_D points for relevant structures, and based on the mean dose to those structures.

3. RESULTS

3.1 Dysphagia

3.1.1 Dysphagia DVHs

Figures 1 and 2 show DVHs that were delineated using the PRO-CTCAE data.

3.1.1.1 Combined Constrictors

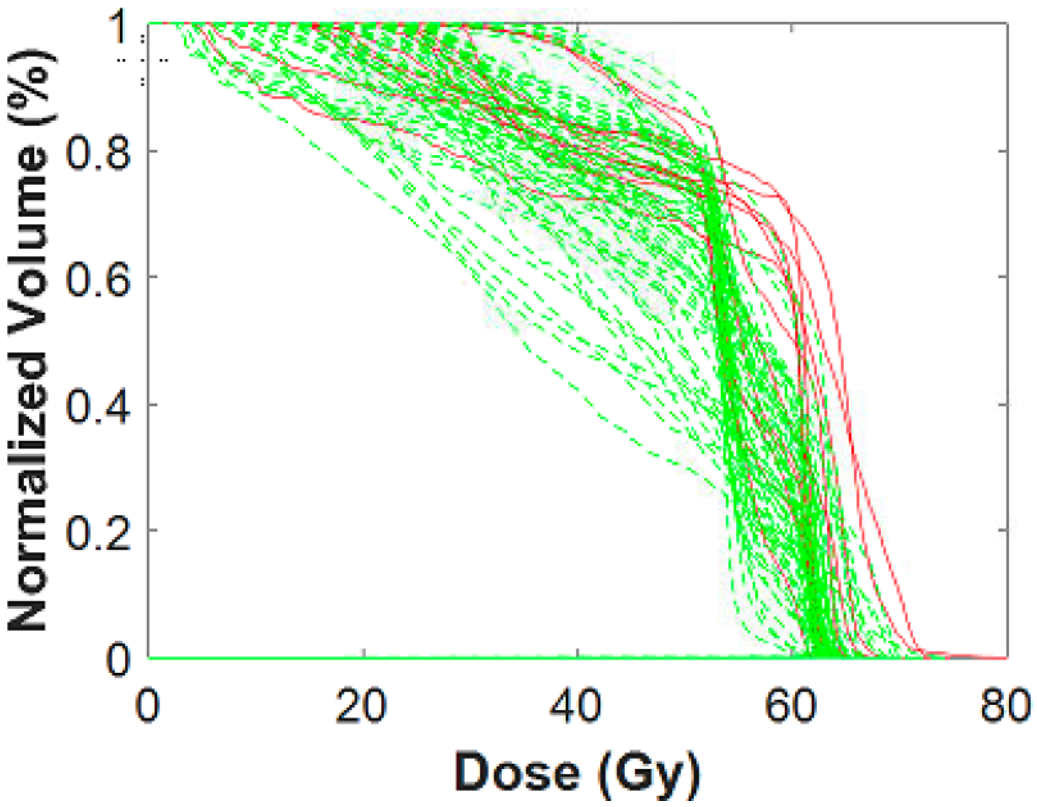

Figure 1: DVHs for the combined constrictors for each patient. Red lines indicate patients whose 12-month scores exceeded the threshold toward our clinical endpoint for dysphagia at 12 months post-treatment. The x-axis indicates dose (in Gy), and the y-axis indicates the relative volume (as a percentage of the absolute volume) receiving at least a given dose.

3.1.1.2 Superior Pharyngeal Constrictor (SPC)

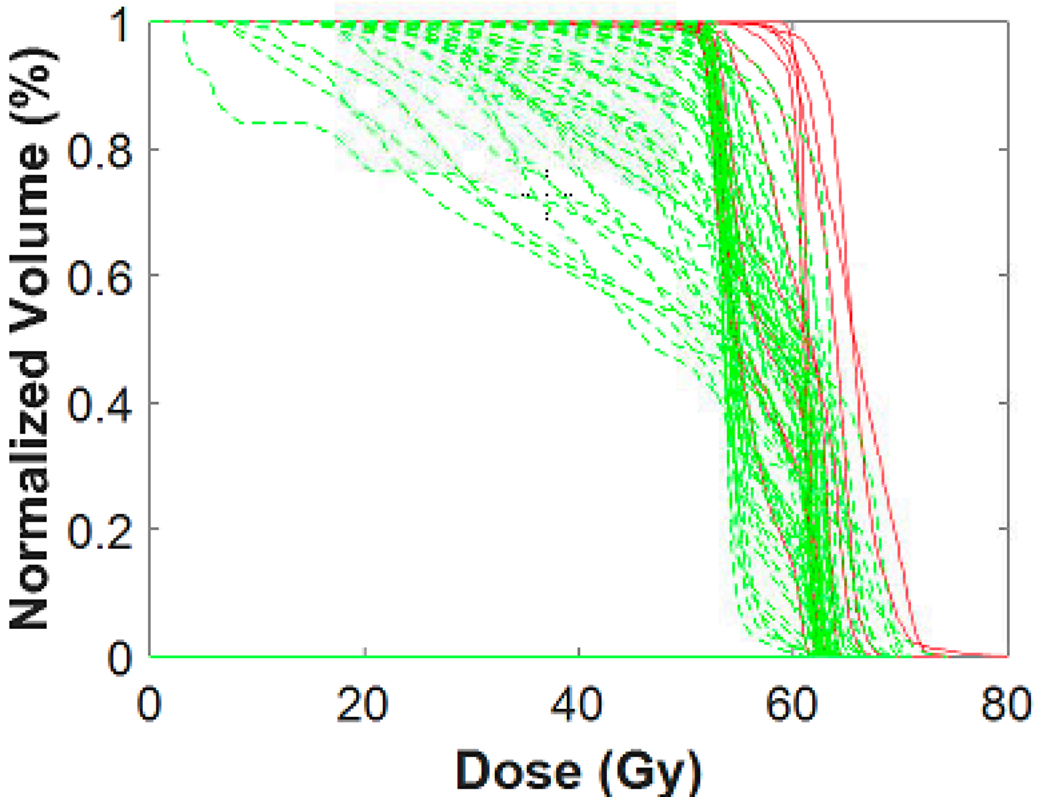

Figure 2: DVHs for the superior pharyngeal constrictor (SPC) for each patient. Red lines indicate patients whose 12-month scores exceeded the threshold toward our clinical endpoint for dysphagia at 12 months post-treatment. The x-axis indicates dose (in Gy), and the y-axis indicates the relative volume (as a percentage of the absolute volume) receiving at least a given dose.

3.1.2 Dysphagia AUC Plots

Figures 3 and 4 show the AUC plots for both the PRO-CTCAE data and the QLQ-C30 data for dysphagia. Note that an AUC value of less than 0.5 implies that higher dose corresponds to lower toxicity. Because patients were treated to a uniform prescription, higher dose to a substructure implies lower dose to the rest of that structure.

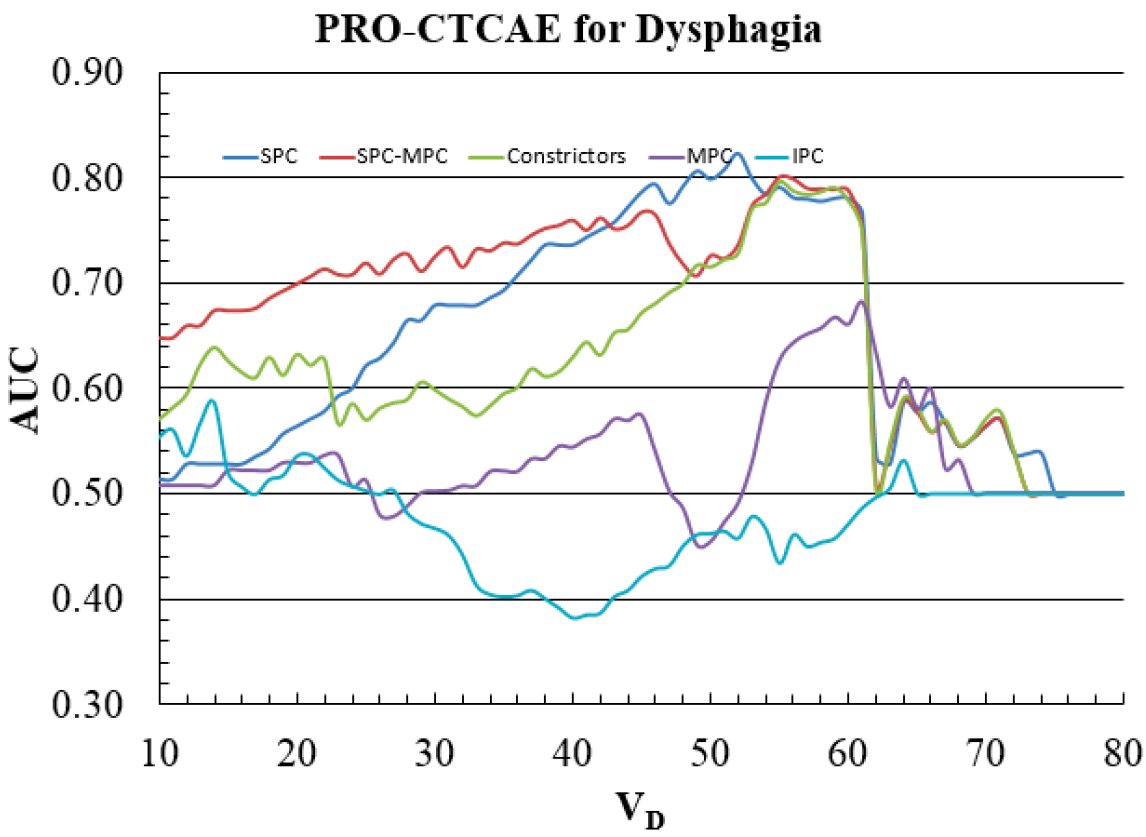

Figure 3: AUC plots for the PRO-CTCAE data for dysphagia. The x-axis indicates the dose-volume points (VD) from 5 to 80 Gy. The y-axis indicates the AUC score for a given dose-volume point. The SPC, SPC-MPC, combined constrictors, MPC, and IPC are indicated in blue, red, green, purple, and sky blue, respectively.

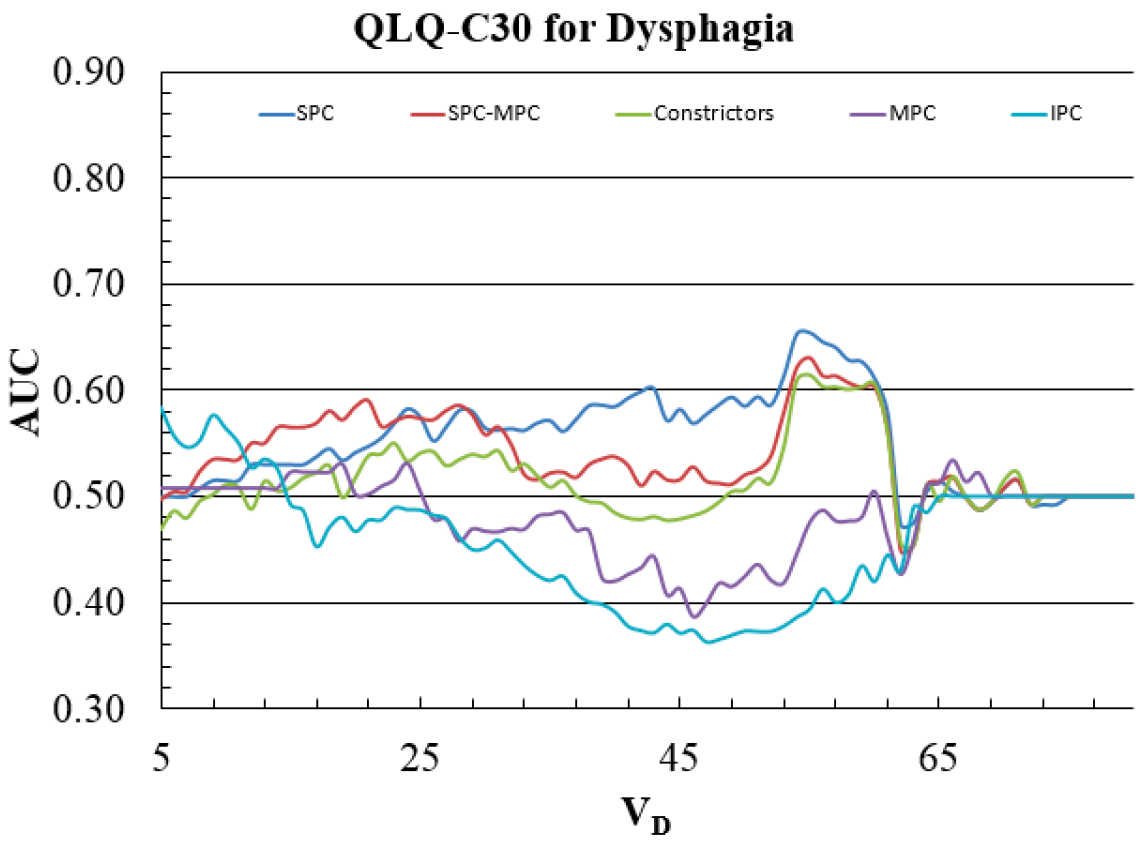

Figure 4:AUC plots for the QLQ-C30 data for dysphagia. The x-axis indicates the dose-volume points (VD) from 5 to 80 Gy. The y-axis indicates the AUC score for a given dose-volume point. The SPC, SPC-MPC, combined constrictors, MPC, and IPC are indicated in blue, red, green, purple, and sky blue, respectively.

3.1.3 Dysphagia Dose Response Curves

Figure 5 shows the dose response curves for both the PRO-CTCAE data and the QLQ-C30 data.

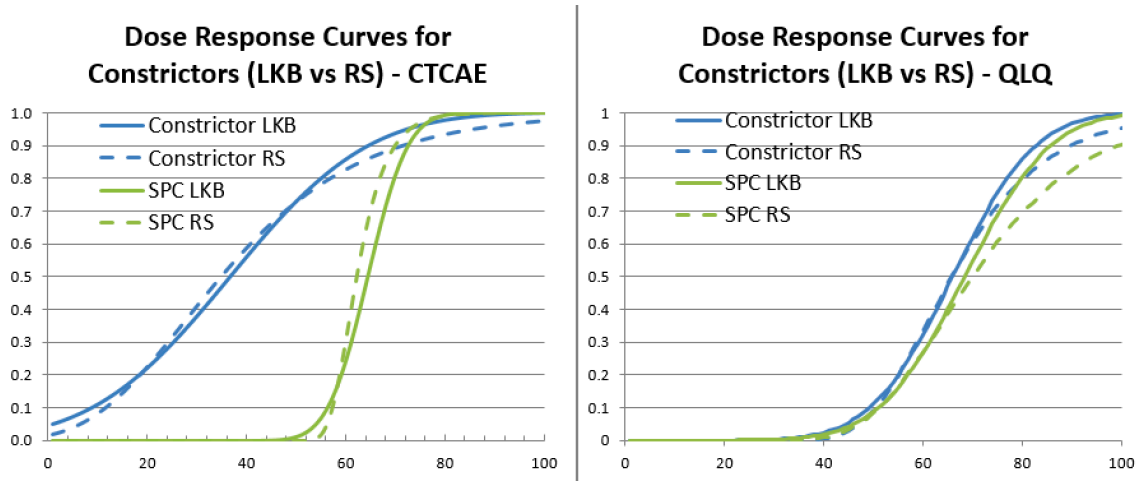

Figure 5: dose response curves for both the PRO-CTCAE data (left) and the QLQ-C30 data (right) for dysphagia. Solid lines indicate the LKB model, and dashed lines indicate the RS model. The blue curves correspond to the combined constrictors, and the green curves correspond to the superior constrictor alone. The x-axis refers to the dose (in Gy) to the region, and the y-axis refers to the complication probability for the tissue at that dose.

3.1.4 Dysphagia NTCP Fit Parameters

3.1.4.1 LKB Model

Table 1: Radiobiological parameters fit for the LKB model for both the PRO-CTCAE and the QLQ-C30 data. D_{50}, m, and n are each of the radiobiological parameters fit for the model.

Dataset	*Region*	D_{50}	*m*	*n*
PRO-CTCAE	Constrictors	58.71	0.10	0.45
	SPC	63.42	0.10	1.00
QLQ-C30	Constrictors	65.03	0.20	0.24
	SPC	67.44	0.20	1.00

3.1.4.2 RS Model

Table 2: radiobiological parameters fit for the RS model for both the PRO-CTCAE and the QLQ-C30 data. D_{50}, γ, and s are each of the radiobiological parameters fit for the model.

Dataset	*Region*	D_{50}	γ	*s*
PRO-CTCAE	Constrictors	58.43	2.68	10^{-4}
	SPC	61.20	5.37	10^{-4}

QLQ-C30	Constrictors	68.84	1.74	0.62
	SPC	69.00	1.49	10^{-4}

3.1.5 Dysphagia AUC and OR Results

Table 3: AUC results for the PRO-CTCAE and QLQ-C30 data, for the combined constrictors and the superior constrictor alone. D_{mean} refers to the mean dose across the structure, and V_{best} refers to the dose-volume point (VD) on the AUC plot with the strongest relationship between exposure and outcome.

Dataset	*Region*	D_{mean}	V_{best}
PRO-CTCAE	Constrictors	0.81	0.79 (V_{52})
	SPC	0.86	0.82 (V_{59})
QLQ-C30	Constrictors	0.52	0.61 (V_{55})
	SPC	0.62	0.65 (V_{55})

Table 4: Dose constraints for the combined constrictors and the superior constrictor alone. The constraints for a 10, 15, and 20% complication probability are listed based both on the best dose-volume point for a given structure, and on the mean dose to that structure.

		Constrictors (V_{52})	*SPC (V_{59})*
10% complication	V_D	< 64%	< 42%
	Mean dose	< 50 Gy	< 55 Gy
15% complication	V_D	< 68%	< 50%
	Mean dose	< 51 Gy	< 57 Gy
20% complication	V_D	< 71%	< 57 %
	Mean dose	< 52 Gy	< 58 Gy

Table 5: Strength of the relationship between the NTCP models for each set of patient-reported outcome data for dysphagia. Shown are the AUC and OR results for

each set of patient-reported outcome data, for both the combined constrictors and the superior constrictor. The AUC values are listed for both the LKB model and for the RS model, along with their strongest associated odds ratios (with a 95% confidence interval).

Dataset	*Region*	*LKB AUC*	*LKB OR [threshold] (95% CI)*	*RS AUC*	*RS OR [threshold] (95% CI)*
PRO-CTCAE	Constrictors	0.823	10.200 [32 Gy] (1.513 – 68.771)	0.794	10.200 [36 Gy] (1.513 – 68.771)
	SPC	0.836	19.143 [32 Gy] (3.905 – 93.837)	0.840	19.250 [41 Gy] (4.357 – 85.058)
QLQ-C30	Constrictors	0.574	2.067 [19 Gy] (0.655 – 6.517)	0.572	5.878 [16 Gy] (0.726 – 47.581)
	SPC	0.627	4.259 [20 Gy] (1.257 – 14.431)	0.622	3.343 [21 Gy] (1.058 – 10.565)

3.2 Xerostomia

3.2.1 Xerostomia DVHs

Figures 6 and 7 show DVHs that were delineated using the PRO-CTCAE data.

3.2.1.1 Combined Constrictors

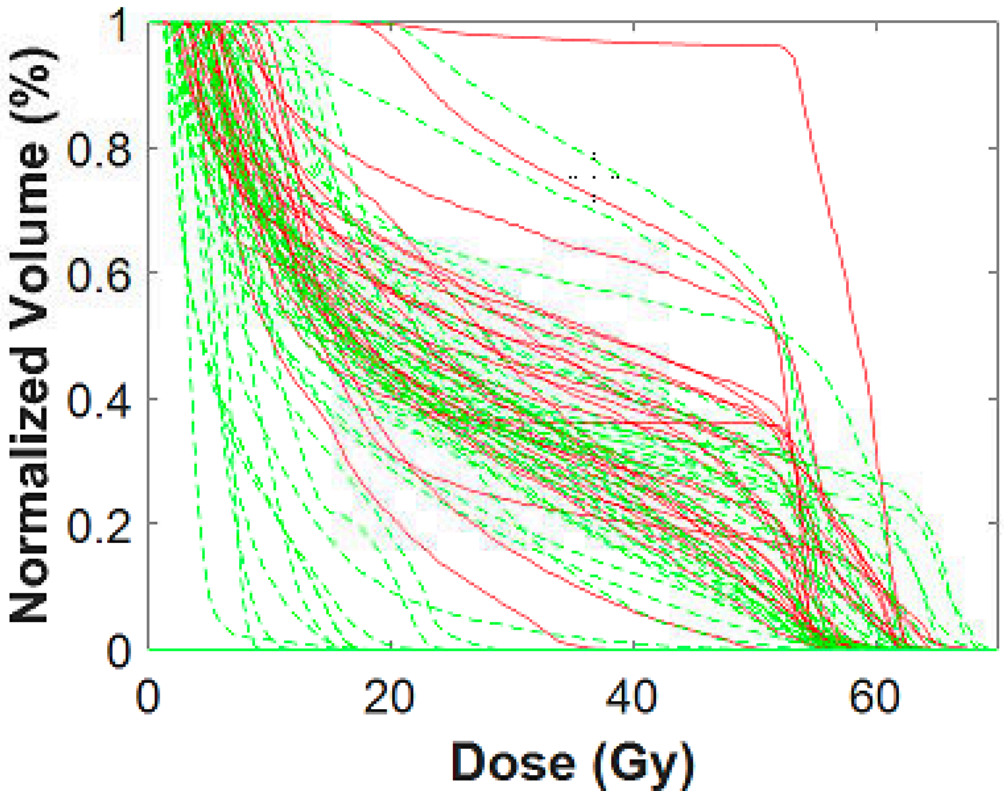

Figure 6: DVHs for the combined contralaterals for each patient. Red lines indicate patients whose 12-month scores exceeded the threshold toward our clinical endpoint for xerostomia at 12 months post-treatment. The x-axis indicates dose (in Gy), and the y-axis indicates the relative volume (as a percentage of the absolute volume) receiving at least a given dose.

3.2.1.2 Contralateral Parotid

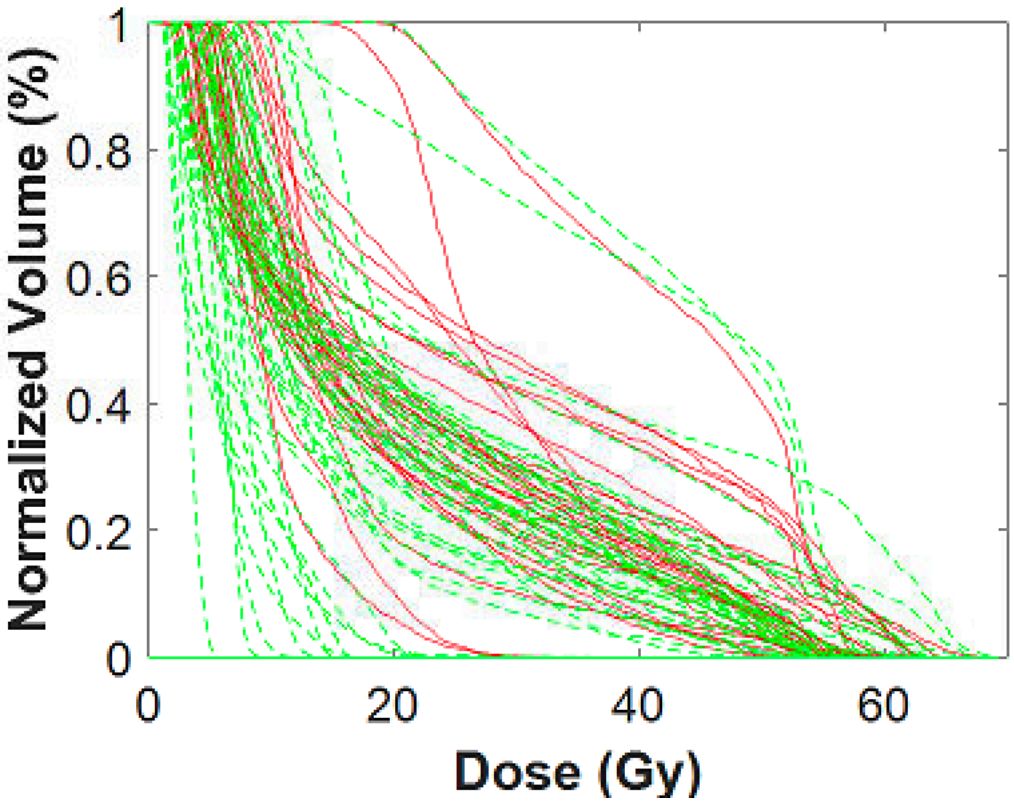

Figure 7: DVHs for the contralateral parotid for each patient. Red lines indicate patients whose 12-month scores exceeded the threshold toward our clinical endpoint for xerostomia at 12 months post-treatment. The x-axis indicates dose (in Gy), and the y-axis indicates the relative volume (as a percentage of the absolute volume) receiving at least a given dose.

3.2.2 Xerostomia AUC Plots

Figures 8 and 9 are the AUC plots for both the PRO-CTCAE data and the QLQ-C30 data for xerostomia.

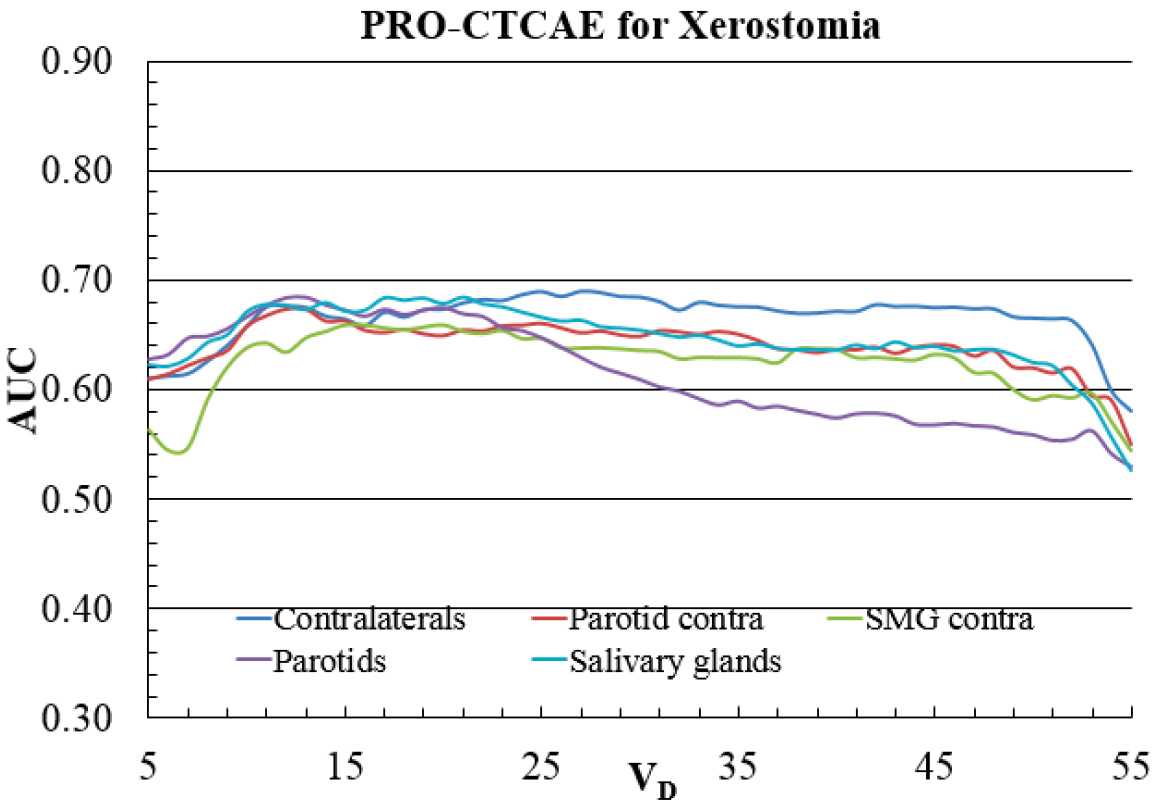

Figure 8: AUC plots for the QLQ-C30 data for xerostomia. The x-axis indicates the dose-volume points (VD) from 5 to 55 Gy. The y-axis indicates the AUC score for a given dose-volume point. The combined contralaterals, contralateral parotid, contralateral SMG, combined parotids, and combined salivary glands are indicated in blue, red, green, purple, and sky blue, respectively.

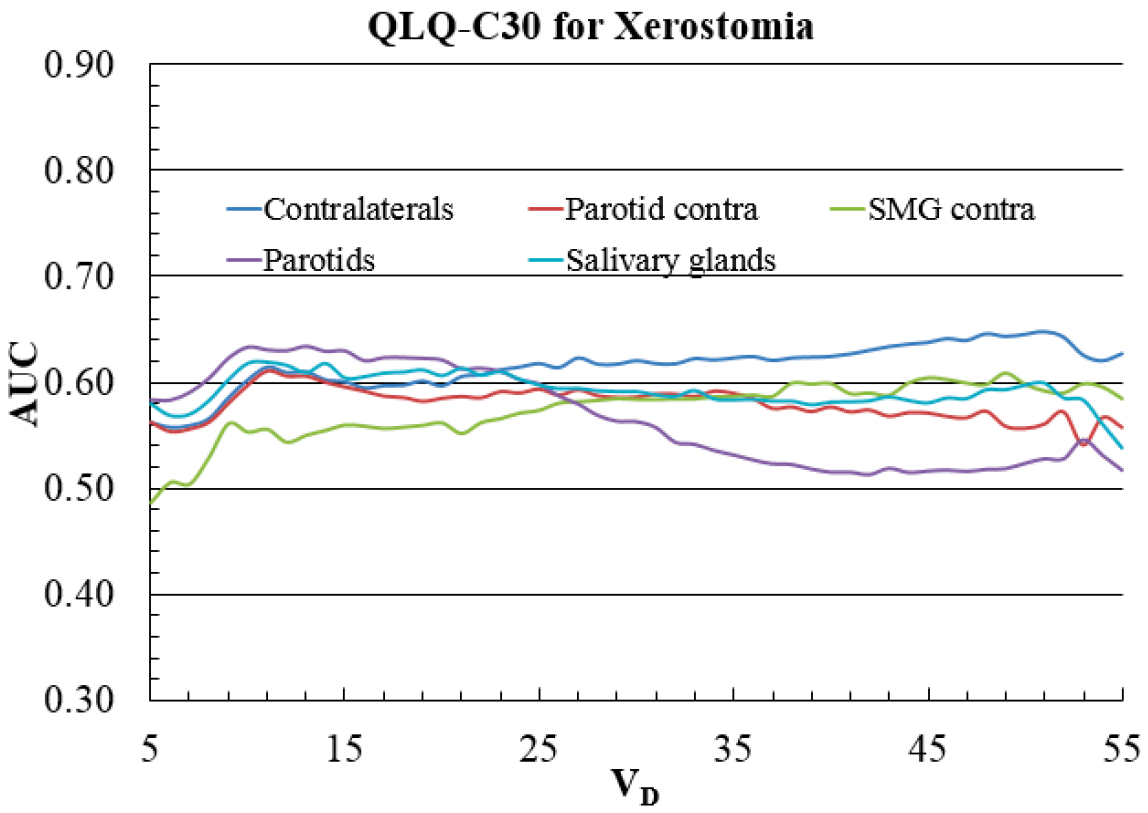

Figure 9: AUC plots for the QLQ-C30 data for xerostomia. The x-axis indicates the dose-volume points (VD) from 5 to 55 Gy. The y-axis indicates the AUC score for a given dose-volume point. The combined contralaterals, contralateral parotid, contralateral SMG, combined parotids, and combined salivary glands are indicated in blue, red, green, purple, and sky blue, respectively.

3.2.3 Xerostomia Dose Response Curves

The dose response curves for both the PRO-CTCAE data and the QLQ-C30 data are shown in Figure 10.

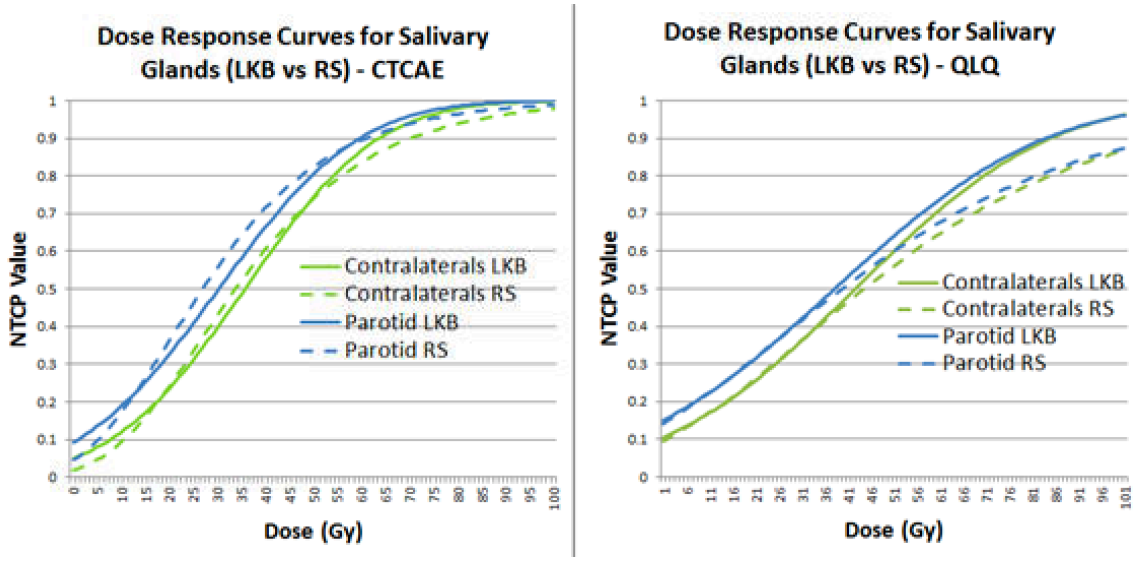

Figure 10: Dose response curves for both the PRO-CTCAE data (left) and the QLQ-C30 data (right) for xerostomia. Solid lines indicate the LKB model, and dashed lines indicate the RS model. The green curves correspond to the combined contralaterals, and the blue curves correspond to the contralateral parotid alone. The x-axis refers to the dose (in Gy) to the region, and the y-axis refers to the complication probability for the tissue at that dose.

3.2.4 Xerostomia NTCP Fit Parameters

3.2.4.1 LKB Model

Table 6: Radiobiological parameters fit for the LKB model for both the PRO-CTCAE and the QLQ-C30 data. D_{50}, m, and n are each of the radiobiological parameters fit for the model.

Dataset	*Region*	*D50*	*m*	*n*
PRO-CTCAE	Contralaterals	35.75	0.61	0.68
	Contra Parotid	30.33	0.76	0.64
QLQ-C30	Contralaterals	41.49	0.80	1.00
	Contra Parotid	37.00	0.97	1.00

3.2.4.2 RS Model

Table 7: Radiobiological parameters fit for the RS model for both the PRO-CTCAE and the QLQ-C30 data. D_{50}, γ, and s are each of the radiobiological parameters fit for the model.

Dataset	*Region*	*D50*	γ	*s*
PRO-CTCAE	Contralaterals	33.89	0.51	10^{-4}
	Contra Parotid	26.95	0.41	10^{-4}
QLQ-C30	Contralaterals	43.36	0.31	10^{-4}
	Contra Parotid	38.54	0.25	10^{-4}

3.2.5 Xerostomia AUC and OR Results

Table 8: AUC results for the PRO-CTCAE and QLQ-C30 data, for the combined contralaterals and the contralateral parotid alone. D_{mean} refers to the mean dose across the structure, and V_{best} refers to the dose-volume point (VD) on the AUC plot with the strongest relationship between exposure and outcome.

Dataset	*Region*	D_{mean}	V_{best}
PRO-CTCAE	Contralaterals	0.69	0.69 (V_{27})
	Contra Parotid	0.69	0.67 (V_{12})
QLQ-C30	Contralaterals	0.64	0.65 (V_{51})
	Contra Parotid	0.61	0.61 (V_{11})

Table 9: Dose constraints for the combined contralaterals and the contralateral parotid alone. The constraints for a 10, 15, and 20% complication probability are listed based both on the best dose-volume point for a given structure, and on the mean dose to that structure.

		Contralaterals (V_{27})	*Contra Parotid (V_{12})*
10% complication	V_D	0%	0%
	Mean dose	< 7 Gy	--
15% complication	V_D	< 7%	< 11%
	Mean dose	< 12 Gy	< 6 Gy
20% complication	V_D	< 17%	< 24 %
	Mean dose	< 17 Gy	< 10 Gy

Table 10: Strength of the relationship between the NTCP models for each set of patient-reported outcome data for xerostomia. Shown are the AUC and OR results for each set of patient-reported outcome data, for both the combined constrictors and the superior constrictor. The AUC values are listed for both the LKB model and for the RS model, along with their strongest associated odds ratios (with a 95% confidence interval).

Dataset	*Region*	*LKB AUC*	*LKB OR [threshold] (95% CI)*	*RS AUC*	*RS OR [threshold] (95% CI)*
PRO-CTCAE	Contralaterals	0.664	3.257 [45 Gy] (1.132 – 9.370)	0.666	3.341 [41 Gy] (1.257 – 8.878)
	Contra Parotid	0.650	4.114 [43 Gy] (1.230 – 13.763)	0.662	3.917 [43 Gy] (1.253 – 12.239)
QLQ-C30	Contralaterals	0.623	3.735 [28 Gy] (0.996 – 14.013)	0.617	2.881 [35 Gy] (1.030 – 7.676)
	Contra Parotid	0.601	3.852 [35 Gy] (1.176 – 12.617)	0.601	3.238 [35 Gy] (1.027 – 10.205)

4. DISCUSSION

4.1 Dysphagia

Looking at the DVHs for both the combined constrictors and the superior constrictor (Figures 1, 2) there is a pretty clear delineation between patients that did and did not exceed the threshold toward our clinical endpoint for dysphagia. In particular, the two patient groupings seem separated at around the 60 Gy dose point. This grouping can be seen quite clearly in the AUC plots for both sets of outcome data. The best relationship between the PRO-CTCAE thresholding and our clinical endpoint seems to be at around the V_{55}-V_{60} for the SPC and for the combined constrictors. The AUC values for the SPC-MPC at these dose-volume points are also quite high. However, this region is likely dominated by the SPC; the relationship between dosimetric data and outcome is roughly as good as random chance for the MPC alone. The results for the QLQ-C30 data are consistent with the results of the PRO-CTCAE data. However, the relationship between dosimetric and PRO data is shown to be much stronger for the PRO-CTCAE data than for the QLQ-C30 data, yielding max values over 0.8 for the PRO-CTCAE data but only around 0.65 for the QLQ-C30 data. Indeed, looking at the data more numerically (Table 3) shows that the results for the PRO-CTCAE data, in particular for the mean dose, show a very strong relationship between dosimetric data and the thresholding for our clinical endpoint. Additionally, the V_{52} of the constrictors and the V_{59} of the SPC are great indicators toward this endpoint. Overall, the QLO-C30 data is

not much better than random chance, though there is a solid relationship for the V_{55} of the SPC. It is worth noting that while previous publications (Li et al., 2009) found that the IPC was a good predictor of dysphagia post-treatment, the results of this thesis found that the IPC was not a great indicator toward that endpoint.

Looking at the dose response curves for dysphagia (Figure 5), both models seem to fit the data equally well. The PRO-CTCAE data appears to yield more consistency than the QLQ-C30 data overall. In particular, the models deviate more dramatically at higher doses for the QLQ-C30 data, but remain quite consistent for the PRO-CTCAE data. Additionally, the steepness of the dose response curve for the superior pharyngeal constrictor for the PRO-CTCAE data can provide a great way to separate patients who exceed our clinical threshold from those that do not. Looking at the results more numerically (Table 4), the AUC values for the PRO-CTCAE data are much more robust than they are for the QLQ-C30 data. Furthermore, the odds ratios for the PRO-CTCAE models are significant and show a strong relationship between dosimetry and outcome. In contrast, the results for the QLQ-C30 are not much better than random chance, though the odds ratios for the SPC are significant. Both models also are consistent with respect to the radiobiological characterization for each region. In particular, the D_{50} values for a given region and a given set of PRO data are consistent between models. Additionally, the volume effect of the LKB model is inversely proportional to the seriality of the RS model. This provides further evidence that both models fit the data

similarly, i.e. one would expect that a parallel organ has a high s value in the RS model and a low n value in the LKB model, or vice versa (Appendix A).

4.2 Xerostomia

Looking at the DVHs for both the combined contralaterals and the contralateral parotid (Figures 6, 7), the delineation between exceeding and not exceeding the threshold toward our clinical endpoint for xerostomia is much less clear than it is for dysphagia. This makes sense in terms of the physical function of the salivary glands compared to that of the pharyngeal constrictors. Notably, the salivary glands have a much more variable volume throughout treatment than the pharyngeal constrictors. Furthermore, salivary function is overall more well-distributed across the head and neck, while swallowing function is well-localized to the constrictors in the throat. Still, there is a clear weighting on one side or the other for each group, in particular around the V_{15} for each plot. The AUC values for the PRO-CTCAE data are relatively high, especially for the contralaterals and the contralateral parotid. Additionally, most regions seem to be a good indicator toward our clinical endpoint at around the V_{15}. Overall, the results for the QLQ-C30 data are not particularly strong, and are generally not much better than random chance. Looking at this information numerically (Table 7), the mean dose seems to be a good indicator of our clinical endpoint. Furthermore, the V_{27} of the contralaterals and the V_{12} of the contralateral parotid show a strong relationship.

The dose response curves for xerostomia (Figure 10) show that for both sets of clinical outcome data, both models have similar fits. However, the fitting for the PRO-

CTCAE data is relatively consistent between models across all doses, while the fitting for the QLQ-C30 data deviates more at higher doses. Still, the AUC values for each model for each region are relatively consistent and show a positive relationship between the exposure and outcome. Additionally, the odds ratios for each NTCP model are significant, with the exception of the QLQ-C30 contralaterals model. As with dysphagia, the radiobiological parameters fit for xerostomia are consistent between models. In particular: the D_{50} values for a given region and set of PRO data are relatively consistent, and the n parameter for the LKB results is inversely proportional to the s value for the RS results – indication that both models predict the same amount of seriality for a given structure.

4.3 Summary of Results

Overall, it was found that the PRO-CTCAE data seems to be more sensitive than the QLQ-C30 data when used as a way to characterize the significance of our clinical endpoints. Based on the PRO-CTCAE data, the V_{52} of the combined constrictors and the V_{59} of the superior pharyngeal constrictor show the strongest relationship with patient-reported dysphagia. Additionally, the V_{27} of the contralaterals and the V_{12} of the contralateral parotid show the strongest relationship with patient-reported xerostomia. Furthermore, it was found that both NTCP models fit the data with similar accuracy. These dosimetric constraint results are consistent with the previous UNC-CH publications. Going forward, these results can be beneficial in IMRT treatment for OPSCC, both in an attempt to use these more refined dose constraints as a way to reduce

the onset of xerostomia and dysphasia for patients post-treatment, and also as a way to identify patients that will likely need additional focused care (e.g. speech and swallowing training) post-treatment. As a final note, about 45 patients are remaining in order to validate the results of this thesis as an independent patient cohort.

4.4 Clinical Application of Work

The results found through this thesis can be applied clinically in order to reduce the rate of xerostomia and dysphagia for patients treated with OPSCC. This can be done in several ways. First, the NTCP models that were generated are a great way to characterize the dose response relationships for the combined contralaterals, contralateral parotid, combined constrictors, and the superior pharyngeal constrictor. A clinic can use this information during treatment planning for defining the dose distributions that will limit the respective risk of complications to the levels that are acceptable by that clinic. These dose response curves are based on the shape of the full dose distribution to that organ. Alternatively, one can pick specific dose metrics in plan optimization. For example, if one wants to use only the mean dose as a metric toward the onset of tissue complication, then the results of the regression can be used, which provides specific thresholds for those metrics for different levels of risk (complication rates). Then, the clinic can determine an "acceptable" complication probability threshold in order to determine its own dose limits. Second, it is apparent through the results of this thesis that some types of patient-reported outcome data can provide predictions for whether or not any given patient will end up with the onset of clinically significant

outcomes post-treatment. In particular, the PRO-CTCAE seems like a robust tool toward characterizing these results. Going forward, a clinic has the option to mirror the analysis that was performed in this thesis in order to figure out which types of patient reported outcome data correspond well to significant clinical endpoints, in order to determine in-house dose constraints for their own treatment schemas.

5. CONCLUSIONS

The use of patient-reported outcome data as a way to determine the severity of a clinical endpoint has strong justification – a patient reporting how he or she finds his or her quality of life can be a direct way to determine patient response post-treatment. Naturally, the quality and reliability of patient-reported outcome data must be verified. Looking at the results of this thesis, it seems that PRO-CTCAE responses can be a great tool to use when determining dose constraints for de-intensified treatment of OPSCC, in particular to reduce the rate of dysphagia post-treatment.

There are two important questions to ask going forward. First, how reliable are the results found for this thesis? Validation of these results is very important. In particular, looking at the entire project chain with more scrutiny could be a strong first step. It is possible that something as simple as choice of threshold for clinically relevant response, or the quality of the contouring, could have an effect on what is shown. Second, what other ways can PRO data be used in order to improve the ways in which we treat head and neck cancer patients? Perhaps the PRO-CTCAE is most robust in characterizing dysphagia, but other questionnaires would provide a robust

characterization of other late term effects of treatment. Ideally, future work would help determine, overarching, which kinds of patient-reported outcome data are good characterizations of clinical endpoints.

Appendix A: Lyman-Kutcher-Burman (LKB) Model

The Lyman-Kutcher-Burman (LKB) model is characterized by the equation

[EQN 1]

$$\mathrm{NTCP} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx$$

where t is described by

[EQN 2]

$$t = \frac{\mathrm{gEUD} - \mathrm{TD50}}{m \times \mathrm{TD50}}$$

and the generalized equivalent uniform dose (gEUD) is described by

[EQN 3]

$$\mathrm{gEUD} = \left(\sum_i v_i d_i^{1/n} \right)^n$$

The three radiobiological parameters that are fit for this model are:

1. TD50, the dose to the tissue that results in a 50% complication probability.

2. m, a fitting parameter that is reflected in the slope of the dose response curve, i.e. it dictates the steepness of the response of the tissue to increasing dose.
3. n, the volume effect for the gEUD. A value of 1 implies the average dose to the volume; a value much less than 1 implies that the higher dose regions in the volume are being weighted higher than the lower dose regions.

The value of n has strong implications of the seriality of the structure. Typically, parallel structures are very sensitive to high dose regions, because a loss in function at one point in the structure can cause a loss in function across the structure. In contrast, serial structures are more sensitive to the mean dose across the structure. This is because high dose to a single point will not cause as much of an overall loss in function to the structure.

Appendix B: Relative Seriality (RS) Model

The Relative Seriality (RS) model is characterized by the equation

[EQN 4]

$$NTCP_{RS} = \left\{ 1 - \prod_i [1 - P(D_i)^s]^{v_i} \right\}^{1/s}$$

where v_i indicates the volume (as a percent of the total volume) receiving a given dose D_i. $P(D_i)$ refers to the complication probability at that dose for that volume, and is described by

[EQN 5]

$$P(D_i) = 2^{-\exp\left[e\gamma_s\left(1-\frac{D_i}{D_{50}}\right)\right]}$$

The three parameters being fit by this model are:

1. D50, the dose that results in a 50% complication probability for the tissue.
2. γ, a measure of the steepness of the sigmoid part of the dose response curve, i.e. how rapidly the tissue responds as the dose is increased.
3. s, a measure of the relative seriality of the structure. A value of 1 indicates a parallel structure, i.e. the dose response for the tissue is more dominated by the high dose regions, because function loss at a single point can cause overall function loss. A value much less than 1 implies a parallel organ, where the average dose is a more significant indicator of function.

References

Belcher, R., Hayes, K., Fedewa, S., & Chen, A. Y. (2014). Current treatment of head and neck squamous cell cancer. *J Surg Oncol, 110*(5), 551-574. doi:10.1002/jso.23724

Cancer.net. (2017). Head and Neck Cancer: Statistics. Retrieved from https://www.cancer.net/cancer-types/head-and-neck-cancer/statistics

cancer.net. (2018). Side Effects of Radiation Therapy. Retrieved from https://www.cancer.net/navigating-cancer-care/how-cancer-treated/radiation-therapy/side-effects-radiation-therapy

Chera, B. (2017). Dosimetric predictors of patient reported xerostomia and dysphasia with de-intensified chemoradiotherapy for HPV-associated oropharyngeal squamous cell carcinoma.

Deasy, J. O., Moiseenko, V., Marks, L., Chao, K. S., Nam, J., & Eisbruch, A. (2010). Radiotherapy dose-volume effects on salivary gland function. *Int J Radiat Oncol Biol Phys, 76*(3 Suppl), S58-63. doi:10.1016/j.ijrobp.2009.06.090

Deschler, D. G., Richmon, J. D., Khariwala, S. S., Ferris, R. L., & Wang, M. B. (2014). The "new" head and neck cancer patient-young, nonsmoker, nondrinker, and HPV positive: evaluation. *Otolaryngol Head Neck Surg, 151*(3), 375-380. doi:10.1177/0194599814538605

Florida, U. o. (2018). Guidelines for Dental Extractions Before Head & Neck Radiation Therapy. Retrieved from https://dental.ufl.edu/patient-care/patient-information/special-care-instructions-for-cancer-patients/guidelines-for-dental-extractions-before-head-neck-radiation-therapy/

Henry, M., Habib, L. A., Morrison, M., Yang, J. W., Li, X. J., Lin, S., . . . Hier, M. (2014). Head and neck cancer patients want us to support them psychologically in the posttreatment period: Survey results. *Palliat Support Care, 12*(6), 481-493. doi:10.1017/S1478951513000771

Hutcheson, K. A., Lewin, J. S., Barringer, D. A., Lisec, A., Gunn, B., Moore, M. W. S., & Holsinger, F. C. (2012). Late Dysphagia after Radiotherapy-Based Treatment of Head and Neck Cancer. *Cancer, 118*(23), 5793-5799. doi:10.1002/cncr.27631

Joseph O. Deasy, P. D., Vitali Moiseenko, Ph.D., Lawrence Marks, M.D., K.S. Clifford Chao, M.D., Jiho Nam, Ph.D., Avraham Eisbruch, M.D. (2010). Radiotherapy Dose–Volume Effects on Salivary Gland Function. *International Journal of Radiation Oncology, 76*(3), S58-S63.

Lewis, A., Kang, R., Levine, A., & Maghami, E. (2015). The New Face of Head and Neck Cancer: The HPV Epidemic. *Oncology (Williston Park), 29*(9), 616-626.

Li, B., Li, D., Lau, D. H., Farwell, D. G., Luu, Q., Rocke, D. M., . . . Chen, A. M. (2009). 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, 4*, 52. doi:10.1186/1748-717X-4-52

Machtay, M., Moughan, J., Trotti, A., Garden, A. S., Weber, R. S., Cooper, J. S., . . . Ang, K. K. (2008). Factors Associated With Severe Late Toxicity After Concurrent Chemoradiation for Locally Advanced Head and Neck Cancer: An RTOG Analysis. *Journal of Clinical Oncology, 26*(21), 3582-3589. doi:10.1200/JCO.2007.14.8841

Mavroidis, P., Price, A., Fried, D., Kostich, M., Amdur, R., Mendenhall, W., . . . Chera, B. (2017). Dose-volume toxicity modeling for de-intensified chemo-radiation therapy for HPV-positive oropharynx cancer. *Radiother Oncol, 124*(2), 240-247. doi:10.1016/j.radonc.2017.06.020

Ringash, J. (2015). Survivorship and Quality of Life in Head and Neck Cancer. *J Clin Oncol, 33*(29), 3322-3327. doi:10.1200/jco.2015.61.4115

seer.cancer.gov. (2013). 5-Year Relative Survival By Year Dx, By Cancer Site, All Ages, All Races, Both Sexes, 1975-2013. Retrieved from https://seer.cancer.gov/faststats/selections.php?series=cancer

seer.cancer.gov. (2014). Cancer Stat Facts: Cancer of Any Site. Retrieved from https://seer.cancer.gov/statfacts/html/all.html

Setton, J., Lee, N. Y., Riaz, N., Huang, S. H., Waldron, J., O'Sullivan, B., . . . Garden, A. S. (2015). A multi-institution pooled analysis of gastrostomy tube dependence in patients with oropharyngeal cancer treated with definitive intensity-modulated radiotherapy. *Cancer, 121*(2), 294-301. doi:10.1002/cncr.29022

Tiziana Rancati, P. D., Marco Schwarz, Ph.D., Aaron M. Allen, M.D., Felix Feng, M.D., Aron Popovtzer, M.D., Bharat Mittal, M.D., Avraham Eisbruch, M.D. (2010). Radiation Dose–Volume Effects in the Larynx and Pharynx. *International Journal of Radiation Oncology, 73*(3), S64-S69.