

Impact of Biomechanical Modeling of Anatomical Variations on the Uncertainties of the
Delivery and Understanding of Radiation Therapy

by

Molly Margaret McCulloch

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Nuclear Engineering and Radiological Sciences)
In The University of Michigan
2019

Doctoral Committee:

Professor Kristy K. Brock, Co-Chair
Associate Professor Martha M. Matuszak, Co-Chair
Assistant Professor Joan M. Greve
Professor Kimberlee J. Kearfott
Professor Edward J. Larsen

Molly M. McCulloch

mollyma@umich.edu

ORCID iD: [0000-0002-7531-1263](https://orcid.org/0000-0002-7531-1263)

© Molly M. McCulloch 2019

Dedication

For my parents.

Acknowledgements

First, I would like to thank my advisor and greatest mentor, Dr. Kristy Brock. Her love for research and dedication to her field are both inspiring and contagious. It was such an honor to earn my Ph.D. under the guidance of such a leader in medical physics. As her student, I was challenged and encouraged every day. She provided me with opportunities I never would have thought possible for a graduate student. I truly owe my love for research to her.

Dr. Martha Matuszak first introduced me to medical physics and the research possibilities in radiation oncology. Martha was instrumental in the success of my graduate school career. I left all our meetings feeling supported and determined, ready to tackle my next obstacle. I also had the pleasure of having Dr. Kim Kearfott, Dr. Joan Greve, and Dr. Ed Larsen on my doctoral committee. I was incredibly lucky to have their expertise while writing my dissertation. All of them put in so much time and effort into ensuring my work was clear and impactful.

The Morfeus lab was such a fantastic environment to complete my research. Guillaume Cazoulat was an incredible teacher throughout the past four years. His patience, support, and willingness to help me through the complex problems I encountered were remarkably valuable. Dan Polan made my transition into the lab seamless by sharing his expertise and experiences. Dr. Bastien Rigaud and Brian Anderson were the best lab mates I could have imagined. The friendships I have developed with my research group make life in the lab (and outside) even more enjoyable.

I am appreciative of the help and consultations from my collaborators at the MD Anderson Cancer Center. Dr. Caroline Chung and Dr. Dave Fuller have been a blast to work with over the last couple of years, and I look forward to working together post-Ph.D.

Finally I wish to thank my loving family and friends. My parents, Louise and Joe McCulloch, never ceased to provide me with anything I needed to succeed. They nurtured my love for learning for as long as I can remember. They picked up the phone every single time I called and were willing to proofread any paper no matter how late it was. They listened to my insecurities and provided the love and support I needed every single day. They will forever be my biggest cheerleaders and I could not ask for better parents. My sister, Maggie, was always available for a pep talk or comedic relief. My friends from back home, undergrad, and graduate school have made this ride a complete joy.

Table of Contents

Dedication	ii
Acknowledgements	iii
List of Figures	viii
List of Tables	x
List of Acronyms	xi
Abstract	xiii
Chapter 1. Introduction	1
1.1 Cancers in the liver and head and neck	1
1.2 Treatment modalities	2
1.2.1 Liver cancer	2
1.2.2 Head and neck cancer	3
1.3 Historical review of dose accumulation, adaptive radiation therapy, and toxicity modeling in radiation therapy	4
1.3.1 Liver cancer	4
1.3.2 Head and neck cancer	5
1.3.3 Toxicity from radiation therapy	6
1.3.4 Liver cancer toxicities	9
1.3.5 Head and neck cancer toxicities	11
1.3.6 Dose accumulation for liver	13
1.3.7 Dose accumulation and adaptive radiation therapy for head and neck cancer	16
1.4 Goals and Organization of this thesis	19
Chapter 2. Background	22
2.1 Radiation Therapy Workflow	22
2.1.1 Computed Tomography	23
2.1.2 Radiation Treatment Planning	24
2.1.3 In-Room Imaging/Cone-Beam Computed Tomography	29
2.1.4 Rigid Registration	31
2.1.5 Radiation Dose Calculation	32
2.1.6 Dose calculation algorithms	33

2.2	Accounting for Anatomical Variations	37
2.2.1	Deformable Image Registration	39
2.2.2	Biomechanical Model-Based DIR Algorithms	42
2.2.3	Dose Accumulation	48
2.3	Adaptive Radiation Therapy	49
Chapter 3. A Simulation Study to Assess the Potential Impact of Developing Normal Tissue Complication Probability Models with Accumulated Dose		51
3.1	Abstract	52
3.2	Introduction	53
3.3	Methods and Materials	54
3.3.1	Patient Data	54
3.3.2	Differences between planned and accumulated doses	56
3.3.3	NTCP Models – Change in probability of toxicity	57
3.3.4	Simulated Toxicities for Duodenum and Stomach	57
3.3.5	NTCP Models – Change in probability of toxicity	60
3.4	Results	60
3.4.1	NTCP Models for Duodenum and Stomach	62
3.5	Discussion	69
3.6	Conclusion	73
Chapter 4. Predictive Models to Determine Clinically Relevant Deviations in Delivered Dose for Head and Neck Cancer		75
4.1	Abstract	75
4.2	Introduction	77
4.3	Methods and Materials	78
4.3.1	Patient Data	78
4.3.2	Image Guided Radiation Therapy	80
4.3.3	Calculation and Evaluation of Delivered Dose	80
4.3.4	Model Development	83
4.3.5	Data Inclusion	85
4.3.6	DIR Evaluation and Validation	86
4.3.7	Independent Model Validation	87
4.4	Results	88
4.4.1	Contour Reproducibility and DIR Validation	89
4.4.2	Dose Deviation Threshold	89
4.4.3	Independent Model Validation	91

4.5	Discussion	92
4.6	Conclusion.....	95
Chapter 5. Biomechanical Modeling of Neck Flexion for Deformable Alignment of the Salivary Glands in Head and Neck Cancer Images		96
5.1	Abstract	97
5.2	Introduction	98
5.3	Methods.....	100
5.4	Patient Data	100
5.4.1	Image registration	101
5.4.2	Registration evaluation	104
5.5	Results	105
5.6	Discussion	109
5.7	Conclusion.....	111
Chapter 6. Biomechanical Modeling of Dose-Induced Volumetric Changes of the Parotid Glands for Deformable Image Registration.....		112
6.1	Abstract	113
6.2	Introduction	113
6.3	Materials and Methods.....	116
6.3.1	Patient Data.....	116
6.3.2	Biomechanical Model-based Deformable Image Registration.....	118
6.3.3	Neck flexion alignment.....	118
6.3.4	Standard Morfeus.....	118
6.3.5	Modeling the dose-induced shrinkage	122
6.3.6	Displacement vector fields analysis.....	123
6.4	Results	124
6.4.1	Population Model.....	124
6.4.2	Target Registration Errors and Dice Similarity Coefficients.....	125
6.4.3	Jacobian determinant distribution and impact on dose mapping	126
6.5	Discussion	130
6.6	Conclusion.....	132
Chapter 7. Conclusions and Future Work		133
7.1	Conclusions	133
7.2	Summary and Future Work	135
References.....		140

List of Figures

Figure 1-1. NTCP/TCP curves.....	8
Figure 2-1. Radiation therapy workflow.....	23
Figure 2-2. Immobilization mask for head and neck (HN) radiation therapy.	24
Figure 2-3. Depiction of gross tumor volume (GTV), clinical target volume (CTV), and planning target volume (PTV).	26
Figure 2-4. Optimized dose distribution on a HN CT scan.	29
Figure 2-5. Head and neck (HN) 3D cone-beam CT (CBCT).....	30
Figure 2-6. Linear accelerator with cone-beam computed tomography (CBCT).....	31
Figure 2-7. Rigid alignment of CT images.	32
Figure 2-10. Delivered dose distribution on a head and neck (HN) cone-beam computed tomography (CBCT).	37
Figure 2-11. Planning CT (left) and mid-treatment CBCT (right) with the same contours overlaid.	38
Figure 2-12. Deformation field obtained with a DIR algorithm that can be used to spatially align images.	40
Figure 2-13. Depiction of Morfeus process.	48
Figure 2-14. Follow-up CT scan with accumulated dose overlaid.	49
Figure 3-1. Planned Vs. Accumulated Dose Plots.....	59
Figure 3-2. Percent Change in Dose and NTCP.	62
Figure 3-3. Probability of Duodenal and Gastric (with Cirrhosis) Toxicity.....	64

Figure 3-4. Probability of Duodenal Toxicity [%] (Based on Accumulated Model) for Duodenum and Stomach.....	68
Figure 4-1. Contour examples.....	81
Figure 4-2. Dose Distribution Examples.	82
Figure 5-1. Examples of rigid registration performance for three cases.....	101
Figure 5-2. PG DSC Histogram.	107
Figure 5-3. DSC Boxplots.....	109
Figure 6-1. Depiction of standard Morfeus process.....	121
Figure 6-2. Depiction of dose-based boundary condition application to Morfeus.	123
Figure 6-3. Population Model.	125
Figure 6-4. Resulting TRE for each DIR Method.....	126
Figure 6-5. Jacobian Maps.....	128
Figure 6-6. DVF Differences.	129
Figure 6-7. Dose Distribution Differences and DVH.	130
Figure 7-1. Examples of sialography images.....	137

List of Tables

Table 2-1. Duodenal NTCP results.....	64
Table 2-2. Stomach NTCP results:	66
Table 3-1. Patient characteristics for all patients included in the study.....	78
Table 3-2. Planning directives for normal tissue structures used in the initial planning.....	79
Table 3-3. Planning and dose deviation thresholds for the organs at risk and the target volumes.	84
Table 3-4. Results of DIR validation study and CBCT contour reproducibility (mean and standard deviation).....	87
Table 3-5. Number of organs with clinically relevant dose deviations and values of those deviations.....	89
Table 3-6. Model based on submandibular glands (SG).....	90
Table 4-1. Registration method comparison according to PG DSC.....	106
Table 4-2. Registration method comparison according to PG DSC.....	108
Table 5-1. Treatment details for patients included in the study.....	116

List of Acronyms

Acronym	Definition
ANACONDA	ANAtomically CONstrained Deformation Algorithm
CBCT	Cone-beam computed tomography
CT	Computed tomography
CTV	Clinical target volume
DIR	Deformable image registration
DSC	Dice similarity coefficient
DVF	Deformation vector field
DVH	Dose volume histogram
EBRT	External beam radiation therapy
EPL	Equivalent path length
ETAR	Equivalent tissue-air ratio
FEA	Finite element analysis
FEM	Finite element model
FOV	Field of view
GBM	Glioblastoma multiforme
gEUD	Generalized equivalent uniform dose
GTV	Gross tumor volume
HCC	Hepatocellular carcinoma
HN	Head and neck
HPV	Human papillomavirus
HU	Hounsfield units
ICRU	International Commission on Radiation Units and Measurements
IGRT	Image guided radiation therapy
IMRT	Intensity modulated radiation therapy
ITV	Internal target volume
LKB	Lyman-Kutcher-Burman
LQED2	Linear quadratic equivalent dose at 2Gy per fraction
MI	Mutual information
MRI	Medical resonance imaging
NPV	Negative predictive value
NTCP	Normal tissue complication probability

OAR	Organs at risk
PC	Percent change
PET	Positron-emission tomography
PG	Parotid gland
PPV	Positive predictive value
PTV	Planning target volume
QA	Quality assurance
QoL	Quality of life
QUANTEC	Quantitative Analyses of Normal Tissue Effects in the Clinic
RILD	Radiation induced liver disease
ROI	Region of interest
RT	Radiation therapy
SBRT	Stereotactic body radiation therapy
SD	Standard deviation
SSD	Sum of squared differences
TACE	Trans-arterial chemoembolization
TE	Echo time
TPS	Treatment planning system
TR	Repetition time
TRE	Target registration error
VMAT	Volumetric modulated arc therapy

Abstract

External beam radiation therapy is an effective and widely used focal cancer therapy. However, due to anatomical changes during radiation therapy, both in the tumor and in the normal tissue, the delivered radiation dose can deviate from the planned radiation dose. These responses may compromise the delivery of the most effective treatment and lead to an increased risk of complications in normal tissues. The ability to estimate the delivered radiation to the tumor and normal tissues with high accuracy requires modeling the patient response to dose. Modern medical imaging, such as computed tomography (CT) and medical resonance imaging (MRI), provides a method to evaluate spatial and functional changes of the tumor and normal tissue over the course of radiation therapy. A comprehensive evaluation of these changes requires identification of the tumor and normal tissue, through image segmentation, and accurate alignment of images, through image registration. In the head and neck region, varying angles of neck flexion, rapid tumor response and weight loss cause early changes in healthy tissue. In the abdominal region, motion due to breathing and digestion cause changes in the tumor position and normal structures. When the deviations between delivered and planned dose are great enough, the radiation treatment plan should be reoptimized, in order to ensure that the tumor is adequately treated and that the normal tissue is maximally avoided. Estimating the delivered dose to sufficient accuracy is therefore an important requirement for effective adaptive replanning. This dissertation work develops different techniques based on biomechanical models

of the anatomical changes to improve estimates of delivered dose, which can ultimately lead to improvements in treatment adaptation strategies as well as a better understanding of toxicity. A series of experiments based on finite element modeling were conducted to model the uncertainties between planned and delivered dose, as well as the potential impact of modeling on different organ sites. Abdominal normal tissue complication probability models were developed based on estimated delivered dose and their accuracy compared to traditional models based on planned dose. Following this study, a predictive model was developed for the head and neck site, in order to find how early in treatment significant deviations in planned and delivered dose could be predicted. After seeing the large potential deviations between planned and delivered dose in the head and neck site, a comprehensive study was conducted to model the changes that potentially cause these large deviations. This comprehensive head and neck model was developed in two steps; first, the positional changes due to flexion were resolved and second, the dose response to the parotid glands was modeled using finite element modeling. Each clinical site poses different challenges, and this dissertation work highlights two areas in which modeling the deviations between planned and delivered dose will improve advanced adaptive radiation therapy.

Chapter 1. Introduction

The majority of cancer patients are treated with radiation therapy at some point during their treatment process. Radiation therapy is a comprehensive part of a treatment plan and is used to target cancer and spare normal tissue. The traditional approach has for decades consisted of planning a dose distribution from a single image of the anatomy and then delivering the treatment during several daily dose delivery fractions. However, the anatomy of the patient constantly changes throughout the duration of treatment, potentially causing variation in the dose to the tumor and other organs. In order to ensure proper delivery of the prescribed dose to the target, image-guided radiotherapy has been developed based on the acquisition of images before each treatment fraction in order to reduce uncertainties in the localization of the target (tumor) during dose delivery. More recently, adaptive treatment strategies taking advantage of these images have been proposed, consisting of evaluating the delivered dose, identifying variations from planning, updating treatment decisions, and modifying the treatment plan accordingly. These novel patient-specific treatment strategies lead to more accurate radiation therapy, but there still exists the need to evaluate the impact of these strategies.

1.1 Cancers in the liver and head and neck

Liver cancer is the sixth most common cancer worldwide, and accounts for 841,080 cases and 5% of all cancers diagnosed globally in 2018 [1]. The most common primary liver cancer is hepatocellular carcinoma (HCC). The highest incidence of HCC occurs in East Asia and sub-

Saharan Africa, where cases are mainly due to the hepatitis B virus [2]. However, HCC is the fastest growing cause of cancer deaths in the United States, which is linked to the hepatitis C virus. Additionally, non-alcoholic fatty liver disease (NAFLD) as well as obesity are also causes linked to the increased prevalence of HCC in the United States. HCC has a 5-year survival rate below 12% [2].

Cancers in the head and neck (HN) region include lip and oral cavity, nasopharynx, oropharynx, hypopharynx, and salivary glands. These cancers account for 710,237 cases and 4% of all cancers diagnosed worldwide in 2018 [1]. HN cancer is also more common in eastern than in western countries. The highest incidence of HN cancers occur in Southeast Asia [3], where risk factors include the areca nut and smokeless tobacco [4]. Incidence of HN cancer is relatively low in the United States (3%) [5], and is primarily due to smoking and alcohol use [6]. Similar to liver cancer, HN cancer can also be caused by viruses; there has recently been a rise in the incidence of oropharyngeal cancers due to the human papillomavirus (HPV) [7]. The overall 5-year survival rate for HN cancers is estimated between 56% and 75% [8-10].

1.2 Treatment modalities

1.2.1 Liver cancer

While surgery to remove the tumor is the best option for cure of liver cancer, it isn't always possible based on the location and size of the tumor [11]. Most chemotherapy drugs are not effective in the treatment of liver cancer, although hepatic artery infusion, which involves the

hepatic artery directly receiving the drugs, has been successful in early studies [12]. Small tumors can be treated with ablation therapy, which is performed by inserting a needle into the tumor to apply heat, cold, radio waves, or alcohol to kill cancer cells. This method is the most effective in tumors that are within 3cm in diameter [13]. For tumors between 3cm and 5cm in diameter, ablation therapy is most effective when used with embolization [13], which aims to block the blood supply to cancer cells. Embolization can also be combined with chemotherapy for trans-arterial chemoembolization (TACE). External beam radiation therapy (EBRT) is considered for liver cancer in cases where the other treatment modalities are not appropriate, however non-randomized studies show it may as effective as ablation [14]. Radiation therapy can be used for primary liver cancer as well as metastases in the liver, and can be delivered to tumors of various sizes. Additionally, radiation therapy has been used as a bridge to transplantation in cases of HCC [15].

1.2.2 Head and neck cancer

Similarly to liver cancer, surgery is commonly used for the treatment of HN cancer.

Chemotherapy, in contrast to the treatment of liver cancer, is effective in the treatment of HN cancer. Surgeries performed for the treatment of HN cancer include laser surgery, which is used for early-stage tumors, excision, dissection (of the neck or lymph nodes), and reconstructive surgery. These surgeries can lead to side effects such as difficulty breathing due to swelling, hearing loss, voice alterations, facial disfigurement, and impaired speech. Additionally, patients can develop problems with chewing and swallowing, which can lead to the need for a feeding tube [16]. Chemotherapy is another method used to treat HN cancer, and involves the delivery of

drugs to the bloodstream to kill cancer cells. Chemotherapy can lead to gastro-intestinal (GI) problems, hair loss, risk of infection, and fatigue [17]. More than half of HN cancer patients are prescribed external beam radiation therapy (EBRT) [18] alone or in conjunction with surgery (in the case of advanced HN cancers). Head and neck cancer patients can also be treated concurrently with EBRT and chemotherapy (chemoradiation).

1.3 Historical review of dose accumulation, adaptive radiation therapy, and toxicity modeling in radiation therapy

1.3.1 Liver cancer

SBRT is a type of EBRT that delivers precise high dose radiation beams over 1-5 fractions and can be used to treat liver cancer and spare surrounding normal tissue. Several studies have demonstrated treatment efficacy using SBRT for liver cancer [19-21]. Recently, Scorsetti et al [19] analyzed 5-year local control, overall survival, progression free survival, and toxicity rates for 61 patients treated with SBRT for inoperable liver metastases. The long term results of this study suggested the safety and efficacy of SBRT for these patients. Dawson et al [20] studied 79 patients with liver cancer treated with individualized SBRT. Dose-limiting toxicity for these patients was not observed, indicating that individualized, image-guided SBRT for liver is a feasible treatment method. Kwon et al [21] evaluated the long-term effects of 42 patients treated with SBRT for inoperable primary small HCC. The long term results of this study found complete response for the in-field lesion in 60% of the patients and overall 3-year survival of 59%, indicating that SBRT is promising for inoperable small HCC.

Delivering high doses of radiation to tumors in the liver is challenging due to the position of liver in the abdomen near radiosensitive gastrointestinal organs. The stomach and duodenum develop toxicity (e.g. nausea, vomiting and bleeding) from radiation treatment and are commonly the dose-limiting organs during treatment [22]. Studies have shown that toxicity of these luminal GI structures are predicted by maximum dose to a small volume [22]. As these organs are situated so close to the liver, the dose to the tumor may need to be limited. Substantial deformation can occur in this region due to breathing and stomach filling [23,24]. While weight loss and tumor response are typically not observed over the course of treatment of liver cancer, random variations occur potentially causing the normal tissue to receive more dose than what was planned [25] increasing the risk of toxicity, which makes it an interesting area to study the uncertainties between planned and delivered radiation dose.

1.3.2 Head and neck cancer

The current most common radiation therapy delivery method for HN cancer, usually post-surgery and often concurrently with chemotherapy, is intensity modulated radiation therapy (IMRT). IMRT employs non-uniform beam intensities, and the plans are typically developed with inverse planning and optimization algorithms. IMRT, like SBRT, can precisely target the tumor while sparing surrounding normal tissue, which is important for parallel organs. This treatment method allows the elevation of dose to the tumor volume and the reduction of dose to surrounding normal tissues. HN IMRT usually delivers 60-70Gy over 30-35 fractions. This

method has been shown to improve target coverage from the traditional radiation therapy methods for HN cancer treatment [26,27], with reports of local control exceeding 90% [28,29].

Treating HN cancer patients can be challenging due to the anatomical changes of the patient during radiation therapy, due to weight loss, patient positioning, and dose-response of the tumor. During radiation therapy to HN cancers, there is often a rapid response of normal tissues near the target volume [30,31]. Studies have shown a rapid change in volume for the parotid glands over the course of radiation treatment, and this volumetric response has been correlated with radiation dose [32]. These rapid changes observed during radiation therapy make the HN region an important area to study the uncertainties between planned and delivered radiation dose.

1.3.3 Toxicity from radiation therapy

As the dose increases during radiation therapy, the probability of tumor control increases, but so does the likelihood of complications for healthy tissue. For optimal radiation therapy, the goal is to minimize normal tissue complication probability (NTCP) and maximize tumor control probability (TCP) [33]. NTCP models aim to quantify the risk of complications based on radiation dose to the patient using a simplified metric [34]. An optimal dose is one that gives the best separation of tumor control and reasonable risk of complication for normal tissue (Fig 1-1). Dose constraints placed on the OAR during radiation treatment planning are defined based on population models such as NTCP. The use of these types of models are critical in optimizing the radiation dose. Hence, it is crucial that these models are as accurate as possible.

The most popular NTCP model has been proposed by Lyman [35] in 1985. The model follows a sigmoidal relationship between dose and probability as defined by equation (1-1):

$$NTCP = (2\pi)^{-1/2} \int_{-\infty}^{[D-TD50(v)]/[mTD50(v)]} \exp\left(-\frac{t^2}{2}\right) dt \quad (1-1)$$

where D is the dose, TD50 is the whole organ dose where NTCP is 50%, v is the fractional volume, and m is the slope of the curve. The model for TCP also follows a sigmoidal relationship (Eq. 1-2):

$$TCP = e^{-N(SF_2)^{\frac{D}{D_{ref}} \frac{\alpha + \frac{D}{n}}{\beta + D_{ref}}}} \quad (1-2)$$

where N represents the amount of clonogenic cells that reside in the tumor, SF₂ represents the fraction of cells that survived the reference radiation dose (D_{ref}), D represents the dose, n represents the amount of treatment fractions, and α/β represents the dose at which equal numbers of cells are killed from the linear and quadratic components. Figure 1-1 represents an example of an NTCP function and the associated TCP function. There exists the need to determine a dose that will ensure a good compromise between control probability and complication probability.

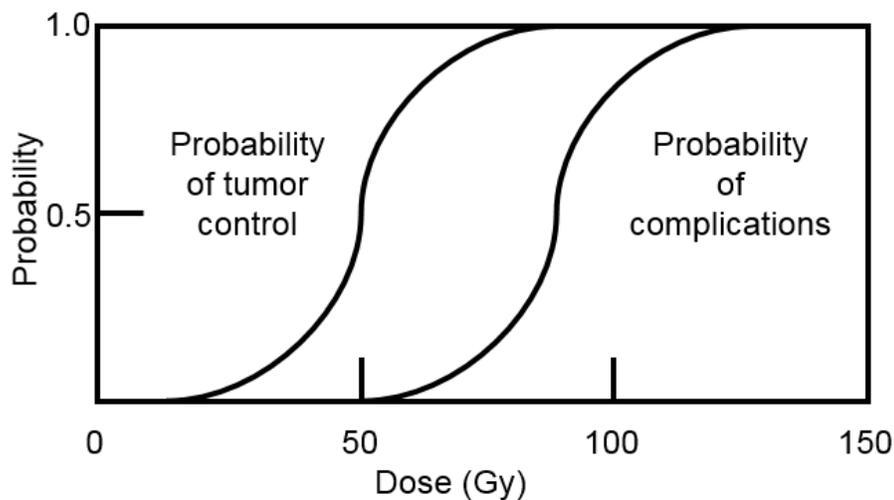


Figure 1-1. NTCP/TCP curves.

Curves for probability of tumor control and probability of complication, with dose [Gy] on the x-axis and probability on the y-axis. Both curves show a sigmoidal relationship between the probability of toxicity and the dose. Optimal radiation therapy compromises between these two curves.

Currently, these toxicity models are based on planned dose. However, Velec et al (2012), Wong et al (2017), and Reese et al (2011) showed that delivered dose is not equivalent to planned dose [25,36,37]. The field needs to accumulate radiation dose during treatment and develop new toxicity models based on accumulated dose. With the recent advances in image-guided radiation therapy, this is now theoretically possible.

There are a limited number of studies that describe toxicity to the stomach due to radiation, separately from the small bowel. One study [38] found that 40.5Gy after chemotherapy and steroids led to 4% of patients developing acute nausea of grade 3 or higher. A randomized study [39] found that a single fraction of 8Gy to the lower hemi-body led to toxicity rates of 6-66% for moderate to severe nausea, depending on the type of chemotherapy. Complications in the small bowel as a response to radiation therapy include diarrhea, fistula, perforation, ulcers, obstruction, and constriction. One study [40] showed that cervical cancer patients had an increase in toxicity

of the small bowel with the addition of chemotherapy to 45Gy radiation therapy. A 5% toxicity rate was observed for those treated with radiation therapy alone, and a 14% toxicity rate was observed for those also treated with chemotherapy. The limited data for toxicity of gastrointestinal organs receiving dose during radiation therapy leads to the need to develop new, more specific, and more accurate toxicity models for these organs.

Deasy et al [32] indicated in a Qualitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC) Organ-Specific Paper that in order to improve HN patient toxicity due to radiation therapy, it must be determined if anatomical variations occur in the salivary glands during radiation therapy, and if shrinkage of the parotid glands should be incorporated in prediction models. They stated that the main goal is to validate accurate prediction models for the function of the salivary glands.

The importance of the accurate accumulation of dose to normal tissues for improved understanding of toxicity was highlighted in a QUANTEC Vision paper by Jaffray et al [41] in 2010, stating that creating accurate distributions of the accumulated dose for the tumor and normal tissue was clinically feasible. In order to make this process regular in radiation therapy, deformable image registration, dose accumulation, and techniques to estimate accumulated dose uncertainties must be further developed.

1.3.4 Liver cancer toxicities

Radiation-induced liver disease (RILD) can occur just weeks after completion of radiation therapy, or months to years later. RILD is the most common significant toxicity that occurs due to EBRT and is a major limitation due to its association with high mortality rates. Symptoms of RILD include abdominal pain, hepatomegaly, fatigue, anicteric ascites, and abdominal swelling [42]. Although the characteristics of RILD are well understood, the pathogenesis remains unclear. RILD can lead to liver failure and has a 10-20% mortality rate [43].

Before technological advances of the early 2000s, radiation treatments for liver cancer consisted of irradiation of the entire liver. In these treatments, RILD rates were 5% for doses up to 30Gy, 10-13% for doses of 30-35Gy, and 44% for doses above 35Gy [44,45]. For daily 2Gy/fraction radiation therapy, the TD5/5, or the dose at which there is a 5% risk at 5 years, for RILD is ≤ 28 Gy for primary liver cancer and ≤ 30 Gy for liver metastases, which are the tolerance doses for whole liver irradiation [44]. These doses are considered palliative, which means that the dose aims to slow tumor growth, shrink the tumor, or control symptoms, and not necessarily to cure the cancer.

The liver is a parallel organ, meaning that the damage to one region of the liver does not indicate the loss of function to the entire organ. NTCP models have been developed to describe the partial volume tolerance of the liver to radiation, where the toxicity being modeled is RILD. These models are used to design radiation treatments for the liver that minimize risk of RILD. The Lyman NTCP curve (1985) [35] models the relationship between NTCP risk and the dose of uniform partial liver irradiation as sigmoidal. The effective volume, V_{eff} , is typically used in the Lyman NTCP model and refers to the uniformly irradiated volume (Eq. 1-3):

$$V_{eff} = \sum_i v_i (D_i / D_{ref})^{1/n} \quad (1-3)$$

where v_i is the volume, D_i is the dose, and D_{ref} is the reference dose, and n determines volume dependence. For normal tissue, $n > 0$. For a parallel organ, where the dose metric is the mean dose, $n = 1$. For a serial organ, where the dose metric is the maximum dose, $1 > n > 0$. A Lyman NTCP model has been established [46] based on the correlation between mean liver dose and the risk of RILD for a series of 203 patients. The study also demonstrated that doses over 100Gy could potentially be delivered to small regions of the liver ($V_{eff} < 30\%$) with a resulting NTCP risk under 5%.

1.3.5 Head and neck cancer toxicities

Toxicity due to radiation treatment of HN cancer includes mucositis, nausea, vomiting, cough, dehydration, fatigue, stomatitis, weight loss, infection, hemorrhage and many other adverse effects [47]. A common side effect that can severely decrease the quality of life (QoL) for HN cancer patients is xerostomia (dry mouth). A major indicator of xerostomia during HN radiation therapy is severe radiation damage to the salivary glands [48].

Xerostomia risk can be decreased with the reduction of parotid gland volume receiving high dose. Using IMRT, at least one parotid gland can be spared in certain patients. The high dose is only delivered to the region of the parotid gland in close proximity to the tumor, and the remaining parotid gland receives a much lower dose [48,49]. Several studies [50-54] have shown

that salivary function in the parotid glands can be conserved when IMRT is used. Additionally, several prospective trials have shown that IMRT can reduce xerostomia risk while still controlling the tumor [29,55-58].

Several uncertainties in toxicity modeling exist. NTCP models are historic population models based on toxicity data from thousands of patients. Because biological differences exist between patients, there is some error in the models. Xerostomia is determined by salivary flow, which is assessed based on the collection of patient saliva. One method consists of having the patient swallow, instructing them not to move or swallow, and then collecting and measuring their saliva [59]. Another method consists of having the patient chew gum at a certain pace, and instructing them to “Spit out, keep chewing” every minute [59]. These methods are also not exact and can potentially add uncertainty to the models. Additionally, toxicity grading is not exact and potentially adds uncertainty to the models. Xerostomia is rated from grade 1 through grade 3, where grade 1 xerostomia indicates that the patient is symptomatic with dry/thick saliva and without significant alteration to the diet (unstimulated salivary flow > 0.2 ml/min), grade 2 indicates that the patient is symptomatic and that their oral intake is altered (unstimulated salivary flow 0.1 to 0.2 ml/min), and grade 3 indicates that the patients cannot adequately ingest orally (unstimulated salivary flow < 0.1 ml/min) [60]. With the advent of functional imaging, more quantitative data can be generated for more accurate grading. Uncertainty in the dose also adds uncertainty to NTCP models. Resolving these uncertainties using accumulated dose may lead to more accurate toxicity modeling, and for the correction of biological uncertainties in these models.

1.3.6 Dose accumulation for liver

Lujan [61] applied convolutional methods to account for uncertainties from patient setup and organ motion of the liver due to breathing. Eq. 1-4 shows his derived dose distribution that includes breathing motion.

$$\bar{D}(x, y, z) = \int D_0(x, y, z')p_{om}(z' - z)dz' \quad (1-4)$$

Where $\bar{D}(x, y, z)$ represents, for point x, y, z , the new dose distribution, $D_0(x, y, z')$ represents, for point x, y, z' , the original dose distribution, and $p_{om}(z' - z)$ is a probability distribution function describing breathing motion. Breathing motion was also fit to a mathematical model,

$$z(t) = z_0 - b\cos^{2n}\left(\frac{\pi t}{\tau} - \phi\right) \quad (1-5)$$

Where z_0 represents the exhale position, b represents the motion amplitude, $z_0 - b$ represents the inhale position, τ represents the breathing period cycle, n represents the model shape, and ϕ represents the breathing cycle starting phase. Eq. 1-6 shows his derived dose distribution that includes patient setup uncertainties.

$$\bar{D}(x, y, z) = \iiint D_0(x', y', z')N(x' - x, y' - y, z' - z)dx'dy'dz' \quad (1-6)$$

Where $\bar{D}(x, y, z)$ represents, for point x, y, z , the average dose incorporating uncertainties, $D_0(x', y', z')$ represents, for point x', y', z' , the static dose, $N(x' - x, y' - x, z' - x)$ represents the normalized PDF for each direction. His work demonstrated that convolutional methods could predict the delivered dose distribution for any amount of radiation therapy fractions. His work also demonstrated that for an infinite number of fractions, the delivered dose distribution could be calculated with convolutional techniques. Due to the finite number of fractions during true radiation treatment, the patient may receive a different dose distribution than the average. A method was then developed to calculate the standard deviation for any point in the patient. The finite-fractioned treatment was then compared to the convolutional method. The convolutional method included breathing motion and patient setup error in a dose distribution and results were validated with Monte Carlo methods. This work demonstrated that breathing motion and patient setup uncertainties should be convolved with dose functions for liver radiation therapy.

Brock [62] studied two deformable image registration methods to describe deformation of the liver during radiation therapy. The first method involved finite element analysis (FEA), which was based on a linear-elastic mechanical liver model. The second method involved mutual-information based registration, and consisted of a geometric liver model. The accuracy of the mutual-information based alignment was 1.0, 1.2, and 1.44 mm for the left-right, anterior-posterior, and inferior-superior directions. One inhale CT and one exhale CT was used to include breathing deformation in dose calculations. The dose delivered to the liver with breathing deformation was summed using the transformation from the deformable alignment. Eq. 1-7 shows the total dose delivered for a voxel with initial position x, y, z during breathing.

$$D(x, y, z) = \sum_{\phi=0}^1 \left[D_{\phi}^E(x + \phi\Delta_x, y + \phi\Delta_y, z + \phi\Delta_z)(1 - \phi) + D_{\phi}^I(x + \phi\Delta_x, y + \phi\Delta_y, z + \phi\Delta_z)\phi \right] T_{\phi} \quad (1-7)$$

Where Δ_x , Δ_y , and Δ_z represent the transformation that resulted from the mutual-information based alignment in each direction, D^E and D^I represent the doses from the exhale and inhale dose grids, respectively, ϕ represents the breathing phase, and T_{ϕ} represents the time weighting based on ϕ . Static dose calculations were compared to the dose calculations that included deformation, in order to determine the effect of deformation. Based on this comparison, 33% of the patients included in the study had prescribed dose changes exceeding one treatment fraction. For tumors residing in the inferior and superior regions of the liver, the effect of deformation was the greatest.

Later work by Brock et al [63] involved the development of a biomechanical model-based deformation image registration algorithm, Morfeus. This method models the deformation of individual organs by assigning linear elastic properties to each of them and applying boundary conditions on their surfaces. This method was evaluated for the lung, liver, spleen, external surface, stomach, and kidneys. Velec [64] used Morfeus to track the tumor and surrounding normal tissue during radiation therapy for liver cancer. Dose accumulation for the liver cancer patients was enabled by modeling their daily anatomic variations. Estimates of delivered dose were improved by dose accumulation. Delivered dose reconstruction indicated substantial deviations (>5%) between planned and delivered dose for 70% of the patients included in the study. These uncertainties ranged between -42% and 8% for normal tissue and -15% and 5% for

tumors. In order to minimize the error in targeting the tumor as well as deviations between planned and delivered dose, image-guidance with respiratory-correlated imaging was used. This work also included the development of a treatment strategy that reduced breathing motion impact and employed a margin that accounted for deformation of the liver. The applied margin decreased the clinical margin by 38% and allowed for a 9% dose increase to the tumor while maintaining the toxicity risk to the surrounding normal tissue. The substantial deviations between planned and accumulated dose observed in this study indicate the need for improved toxicity models, based on accumulated dose.

1.3.7 Dose accumulation and adaptive radiation therapy for head and neck cancer

Treating HN patients with IMRT leads to sharp dose gradients, which can potentially cause greater inaccuracies in the dose distribution than standard techniques in the case of patient positioning errors and anatomical changes. Image-guided radiation therapy (IGRT) techniques can provide 3D images of the patient during radiation therapy, allowing for corrections in tumor position for more accurate radiation therapy.

Several studies have investigated the volumetric changes for HN patients undergoing radiation therapy, exploring tumor shrinkage and weight loss. Barker et al [31] studied 14 patients with in-room imaging (CT-on-rails) acquired three times per week. Rigid image registration was used to compare the daily images with the planning image. The tumor and parotid glands shrunk 1.8% and 0.6%, respectively, for each day of treatment. Geets et al [65] investigated the volumetric change for 10 patients using CT scans taken four times throughout treatment. The per-treatment

images were rigidly registered to the planning image. The tumor shrunk 65.5% after delivery of 45Gy (mean dose). Han et al [66] used daily imaging (helical megavoltage CT) rigidly registered to the planning image for five patients and found 1.1% shrinkage in the parotid glands per day of treatment, with an overall average change from 20.5cm³ to 13.2cm³. Hansen et al [67] acquired CT scans after delivery of 38Gy for 13 patients, rigidly aligned the images to the planning images, and found no volumetric change in the tumor, but 15.6% shrinkage in the right parotid gland and 21.5% shrinkage in the left parotid gland. Robar et al [68] acquired weekly CTs for 15 patients, rigidly aligned them to the planning images, and found shrinkage in the superficial lobe of the parotid glands of 4.9% per week.

Osorio et al [69] acquired CT images after 46Gy was delivered to 10 patients, registered the per-treatment CT to the planning image using deformable image registration, and found 25% shrinkage in the tumor, 17% shrinkage in the parotid glands, and 20% shrinkage in the submandibular glands after 46Gy. Castadot [70] acquired CT scans four times during treatment for 10 patients, used deformable image registration to align them with the planning images, and found tumor shrinkage of 3.2% per treatment day and parotid gland shrinkage of 0.9% per treatment day.

Several studies have investigated the impact that these anatomical variations have on the planned radiation dose, based on image-guided radiation therapy (IGRT) methods. Hansen et al [67] acquired CT scans after 38Gy was delivered to 13 patients and rigidly registered the per-treatment CTs to the planning images. The parotid gland V_{26Gy} (volume of the parotid glands that received at least 26Gy) increased by 10.9%, and the mandible V_{60Gy} increased by 7.2%. This

study also found that replanning lead to a significant improvement in the maximum dose to the spinal cord and brainstem, mean dose and V_{26Gy} to the right parotid gland, and maximum dose and V_{60Gy} to the mandible. Robar et al [68] studied 15 patients with weekly CT scans and found that the mean dose and V_{26Gy} increased 2.6% and 3.5% in the left parotid gland and 0.2% and 0.3% in the right parotid gland. Han et al [66] used daily helical megavoltage CTs and rigid registration for 5 patients and found an increase in the mean dose to the parotid gland from 0.83 to 1.42Gy.

O'Daniel et al [71] studied 11 patients with in-room imaging (CT-on-rails) acquired two times per week. Deformable image registration was used to register the per-treatment images to the planning CT. A median dose increase of 1Gy to the parotid glands was observed. This study also found that the cumulative dose to the parotid gland was higher than the planned dose in the case that IGRT was not used during daily patient setup. Lee et al [72] used daily helical megavoltage CT and deformable image registration and found that the daily mean dose delivered to the parotid gland deviated from the planned dose by 15%. The total accumulated mean dose to the parotid gland was 32.7Gy, while the total planned mean dose was 29.7Gy. This study also found a relationship between weight loss of the patient and greater mean dose to their parotid glands. Castadot [70] used CT scans taken four times throughout treatment for 10 patients and aligned them to the planning image using deformable image registration. This study found that the parotid glands were planned to a mean dose of 17.9Gy, but in actuality received 18.7Gy, and that the submandibular glands were planned to a mean dose of 51.9Gy, but in actuality received 52.8Gy.

Kong et al [73] showed the relationship between salivary flow in HN cancer patients and the mean dose to their parotid glands. The study demonstrated that even for parotid glands that received no dose, there was still no salivary flow for some patients. One potential explanation for this observation is that deviations between the planned and delivered dose existed for these patients. While their parotid glands were planned to receive no dose, they received some dose due to anatomical changes that may have occurred during radiation, leading to xerostomia. Toxicity models for radiation treatment are based on the planned mean dose to the entire gland, and may not be accurate. To date, there is a need to develop new models, as well as to assess the uncertainties and increase the accuracy of dose accumulation in these models.

1.4 Goals and Organization of this thesis

Adaptive radiation therapy aims to limit the discrepancies between planned and delivered radiation dose with replanning. Adaptive radiation therapy relies on accurate dose accumulation, which is based on accurate anatomical modeling. Dose accumulation based on accurate anatomical modeling can potentially translate to more accurate toxicity models. This thesis focuses on modeling the anatomical changes of patients undergoing radiation therapy and the potential clinical impact of the discrepancies between planned and delivered dose that these changes cause, for the abdominal region and the HN region.

There are three major aims included in this thesis work: i) to improve the understanding of toxicity by quantifying deviations between accumulated and planned dose in NTCP models, ii) to develop predictive models for treatment replanning decisions, and iii) to develop biomechanical

models of dose response in order to improve toxicity prediction models. This thesis includes four main chapters, based on the major aims of this dissertation.

Chapter 3 investigates the potential clinical impacts of building luminal NTCP models based on accumulated dose rather than planned dose. NTCP models were simulated based on accumulated dose using statistical modeling for patients treated with liver SBRT. It was hypothesized that using accumulated dose data from DIR and statistical modeling to simulate new toxicity models for the duodenum and stomach will result in potentially clinically significant differences from the standard model. The goal of the first aim was to quantify the potential clinical impact of developing NTCP curves based on accumulated dose rather than the standard, planned dose model. Luminal (tubular) GI structures were chosen for this study because these are often radiation dose-limiting organs.

Chapter 4 describes the development of predictive models for treatment replanning decisions in HN treatments. The need to replan is frequent in the HN site, as weight loss and tumor response can cause potentially clinically significant deviations between delivered and planned dose. Using these findings, we hypothesized that a dose deviation threshold could be applied to build a model that predicts, with 100% sensitivity, the need to replan a patient to avoid clinically relevant dose deviations. The goal of this aim was to develop and validate a predictive model to determine selection criteria where HN patients will need adaptive replanning by mid-treatment to avoid substantial deviations in their accumulated dose.

Chapter 5 investigates an optimal technique to resolve the varying angles of neck flexion

between images acquired during HN radiation therapy. A biomechanical model-based DIR algorithm with boundary conditions on bony anatomy is proposed to accurately account for neck flexion of HN patients.

Chapter 6 examines the development of biomechanical models of dose response of the parotid glands to improve toxicity prediction models. The dose metric used for assessment of the parotid glands is mean dose, and this metric is used as the basis for toxicity modeling. However, recent small animal studies show that there may be a stronger correlation between dose to substructures of the parotid glands and toxicity [74]. In order to establish the relationship between the substructure dose and parotid gland toxicity, the dose to the substructures must be evaluated. The spatial response of the parotid glands to radiation is expressed using thermal expansion coefficients to apply dose-based boundary conditions. The volumetric change in the parotid glands during radiation treatment is dramatic compared to other organs, making this an exciting site to evaluate the dose-response.

Chapter 7 is the concluding chapter that discusses the potential clinical impact and possibility for future studies based on this work.

Chapter 2. Background

Approximately 40% of the US population will develop cancer during their lives[75], and 50% of these patients are treated with radiation therapy as part of their cancer treatment. Radiation therapy is part of a comprehensive treatment plan that may also include surgery or chemotherapy. Radiation is used to treat cancer by delivering high doses of radiation to the tumor in the form of x-rays, gamma rays, or charged particles. Radiation destroys cancer cells through DNA damage, which can be direct or indirect. Once DNA is damaged, those cells can no longer divide or survive. It is therefore crucial to limit the dose to normal tissue. For biological reasons the treatment is often fractionated, e.g. delivered over several treatment sessions.

2.1 Radiation Therapy Workflow

Once patients are diagnosed with cancer and referred for radiation therapy, they undergo a treatment workflow that includes imaging in radiation oncology, radiation treatment planning, dose delivery over a series of fractions, and imaging for patient follow up (Fig. 1.1). Although radiation therapy encompasses external beam radiation therapy (EBRT) and brachytherapy, this dissertation will focus on EBRT, which is the most common method of radiation treatment.

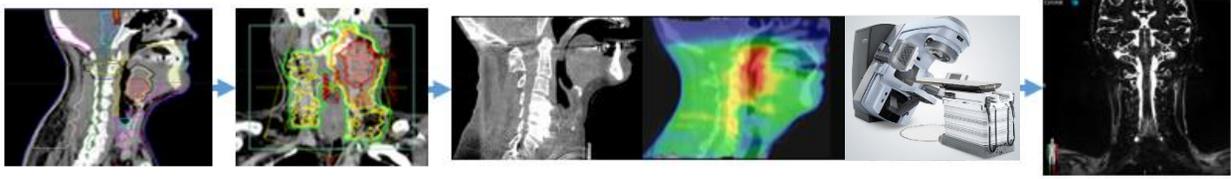


Figure 2-1. Radiation therapy workflow.

Images from left to right: treatment imaging, treatment planning, dose delivery, and follow-up imaging. Linear accelerator: Clinac iX System, Varian Medical Systems, Palo Alto, CA.

2.1.1 Computed Tomography

The standard imaging modality for treatment planning is computed tomography (CT). CT scanners consist of a detector opposite of a kV x-ray tube, which rotates around the patient. Measurements are taken as the x-ray and detector rotates, which returns data for multiple x-ray projections through the patient. A cross-sectional image can then be reconstructed mathematically to represent tissue visually based on its estimated attenuation coefficients, known as CT numbers. CT numbers are referred to as Hounsfield Units (HU), which are calculated using Eq. 2-1,

$$HU = 1000 * \frac{\mu - \mu_{water}}{\mu_{water}} \quad (2-1)$$

where μ represents the linear attenuation coefficient. Hounsfield units are dependent on the energy of the CT beam but generally range from -1,000 for air to -100 for fat to 0 for most tissue to 1,000 for bone. Electron density (electrons/cm³) is needed in order to perform radiation dose calculations. Electron density is calculated using the CT number based on the linear correlation between CT numbers and attenuation coefficients. The calibration is performed by obtaining CT

images of phantoms with known electron densities, representing lung, muscle, and bone to enable the relationship between CT number and electron density to be established.

The planning CT, or CT simulation as it is often referred to as, is the acquisition of a diagnostic CT image with the patient in the treatment position using an immobilization device. An example of an immobilization device used for brain and head and neck treatments is a mask (Fig. 1.2) that is designed to individually fit to the patient. The immobilization mask is a plastic mesh that is heated and formed to the patient's head. Once the plastic returns to room temperature, it becomes rigid and is used to reproduce the position at CT simulation during each treatment delivery and aid in reducing patient movement during radiation delivery.



Figure 2-2. Immobilization mask for head and neck (HN) radiation therapy.

The material is heated, molded to the individual patients face, and cooled. The patient wears the mask during each fraction of radiation therapy to ensure that they are in the same position as they were during their planning scan.

2.1.2 Radiation Treatment Planning

Radiation treatment planning starts with the outlining of normal tissue and tumor on the planning CT. This outlining process is called contouring or tissue delineation. The radiation oncologist

contours the tumor (target volume) and then the organs at risk (OAR), which are normal tissues near the tumor. The target volumes defined in radiation oncology and recommended by the International Commission on Radiation Units and Measurements (ICRU) are the gross tumor volume (GTV), clinical target volume (CTV), internal target volume (ITV), and planning target volume (PTV). An example of these target volumes for a head and neck cancer patient can be seen in Fig. 1.3. The GTV consists of the tumor, including primary or metastatic disease, that is visible on all available images. The CTV includes the region around the visible tumor that may have microscopic disease, not visible on imaging. The ITV accounts for the CTV as well as size, shape, and positional uncertainties in the CTV. The PTV encompasses internal variations, which include CTV position, shape, and size, and external variations, which include beam and patient positioning [76]. Van Herk et al [77] developed a formula (Eq. 2-2) to calculate the PTV margin, M , necessary for the CTV to be covered by a minimum of 95% of the prescription dose for 90% of the population. Systematic errors, which are typically introduced during treatment planning and carried through treatment, cause shifts in the dose distribution (relative to the CTV). Random errors, which occur because of variations in day-to-day treatment, cause dose distribution blurring.

$$M = 2.5 \Sigma + 1.64(\sigma - \sigma_p) \quad (2-2)$$

where Σ represents the standard deviation for systematic errors, σ_p represents the width of the radiation beam's penumbra, and σ is the sum of σ_p^2 and the squared standard deviation of the random errors.

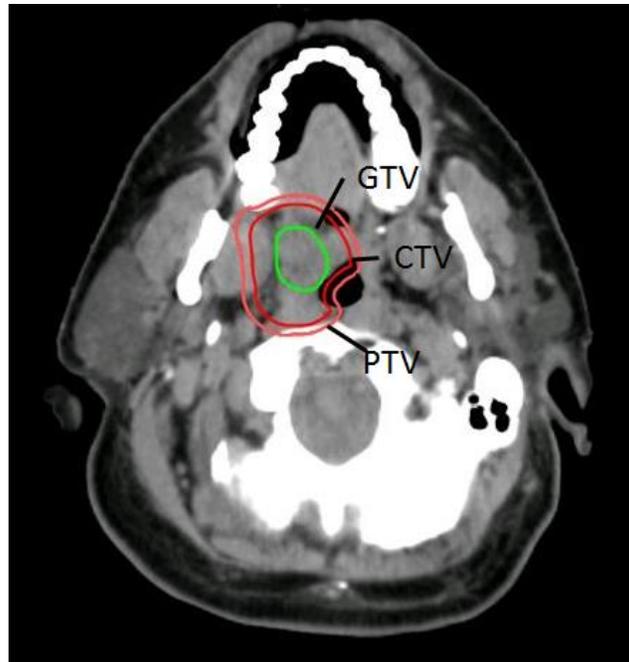


Figure 2-3. Depiction of gross tumor volume (GTV), clinical target volume (CTV), and planning target volume (PTV).

Following contour delineation, radiation parameters corresponding to the function of the linear accelerator, for example the number of beams, beam angle and fluence through a multileaf collimator, are optimized by the treatment planning software in order to obtain a 3D dose distribution. This dose distribution (Fig. 1.4) will aim for the target to receive the prescribed dose and for the dose to other structures to be as low as possible and not exceed given constraints. Advanced treatment delivery techniques have been developed to modulate the radiation, enabling improved sparing of normal tissue and targeting of the tumor. These techniques include volumetric modulated arc therapy (VMAT), which is a radiation therapy technique in which delivery of radiation is entirely dynamic, and moves linearly between positions of the multi-leaf collimator (device composed of Tungsten leaves that move individually in and out of the particle beam path to block it) while the radiation beam is on with variable gantry speed and dose rate. .

VMAT also has the potential to offer additional advantages, such as reduced treatment delivery time compared with conventional static field intensity modulated radiotherapy (IMRT).

IMRT is a traditional technique which involves non-uniform beam intensities and is normally delivered with treatment delivery equipment controlled by a computer. With IMRT, the patient can be treated from different directions with nonuniform beams that are optimized to spare normal tissue while delivering high radiation dose to the tumor. The TPS converts the beams into many beamlets and optimizes their weights or fluences. The intensities and weights of the beamlets can be adjusted to meet the requirements of the dose plan (inverse planning). Optimal intensity profiles are calculated using analytic or iterative methods on a computer. Analytic methods are done with mathematical techniques that invert the dose distribution with back projection algorithms. Iterative methods consist of iteratively adjusting the weights of the beamlets to minimize cost function value. The standard cost function is the least square function (Eq. 2-3).

$$C_n = \left[\left(\frac{1}{N} \right) \sum_r W(\vec{r}) (D_0(\vec{r}) - D_n(\vec{r}))^2 \right]^{0.5} \quad (2-3)$$

where C_n represents the cost at iteration n, \vec{r} is a point in the patient, $D_0(\vec{r})$ represents the optimal dose, $D_n(\vec{r})$ represents the calculated dose at that point, $W(\vec{r})$ represents the weight of the structures, and N represents the number of dose points. For the tumor, the root mean squared difference between the true dose and the prescription dose equals the cost. For the normal tissue, the root mean squared difference between the true dose and zero dose is the cost. The sum of the

normal tissue and tumor costs, based on their weights, determines the total cost. The goal of the optimization algorithm is to minimize the total cost for every iteration until the goal radiation dose is reached.

In order to prevent errors during the radiation therapy process, quality assurance (QA) is performed. QA is a process completed by a team of medical physicists in order to ensure the safety of patients, and includes measuring dose with a phantom, which is made of a synthetic material that mimics human tissue and comparing the measurements with the calculations obtained from the treatment planning system. Before delivering VMAT to a patient, measurement-based QA is performed. This is often done using a cylindrical acrylic phantom accompanied by an array of diode detectors that compare the measured and calculated doses [78]. The measurements are then assessed using both distance and accuracy metrics, referred to as a gamma analysis [79,80]. Most institutions use a 3%, 3 mm pass rate for 95% of the points [78]. When this cannot be achieved, the plan must be evaluated by the medical physicist to determine its suitability for patient treatment.

To deliver radiation, the linear accelerator accelerates electrons, which then collide with a Tungsten target to generate high-energy x-rays through bremsstrahlung. A multileaf collimator, situated inside the head of the linear accelerator, shapes the radiation beam. Finally, the beam emerges from the gantry, which rotates around the patient accordingly. The patient lies on the treatment couch, which is also moveable in 3 translational directions and in special cases rotational axes as well. Based on gantry rotation and treatment couch movement, the radiation beam can deliver dose at any angle. Before treating the patient, the aforementioned parameters

must be designed in process called ‘treatment planning’, which is typically performed by a board certified dosimetrist, or in some clinics or in especially complex cases, a board-certified medical physicist.

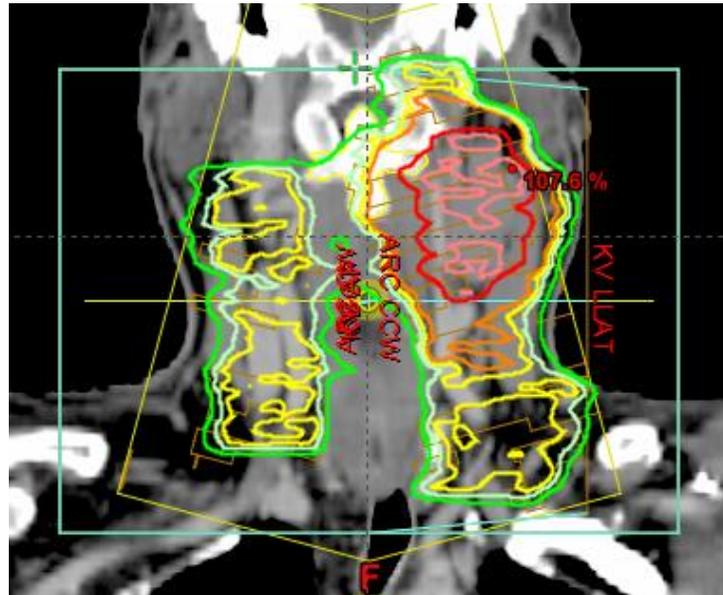


Figure 2-4. Optimized dose distribution on a HN CT scan.

The red contour surrounds the target tissue (tumor) and refers to the high-dose region. The green contour surrounds the low-dose region, and the yellow and orange contours represent the medium-dose region. Dose distributions are optimized to limit dose to surrounding tissue while distributing the maximum dose to the tumor.

2.1.3 In-Room Imaging/Cone-Beam Computed Tomography

The radiation dose prescribed by the radiation oncologist is typically delivered over a series of days, or fractions of smaller doses, which are delivered over the course of radiation therapy. The number of fractions varies depending on the treatment strategy. Before dose delivery at each fraction, the patient is positioned in the immobilization device, with the goal to reproduce their position from the CT simulation. Traditionally, the patient is repositioned according to external

landmarks such as skin markers. Portal imaging, which acquires images with the MV radiation beam from the linear accelerator, allows for the alignment of bony anatomy and some soft tissue structures (e.g. the lung diaphragm interface). While this technique is more reliable than skin markers for positioning, it offers limited soft tissue targeting. A recent and widely adopted advancement is the addition of a kV x-ray tube mounted 90 degrees from the MV radiation beam, enabling cone-beam CT (CBCT) acquisition in the treatment room (Fig. 1.6). The CBCT, as its name indicates, consists of a cone beam of x-rays contrary to a fan beam for the CT. CBCT imaging allows the visualization of bony anatomy and soft tissue in 3D, and can be useful for anatomical localization in the head and neck, abdomen, and pelvis regions (Fig 1.5).



Figure 2-5. Head and neck (HN) 3D cone-beam CT (CBCT).

These images are acquired at each fraction of radiation therapy to align the patient with their initial position.

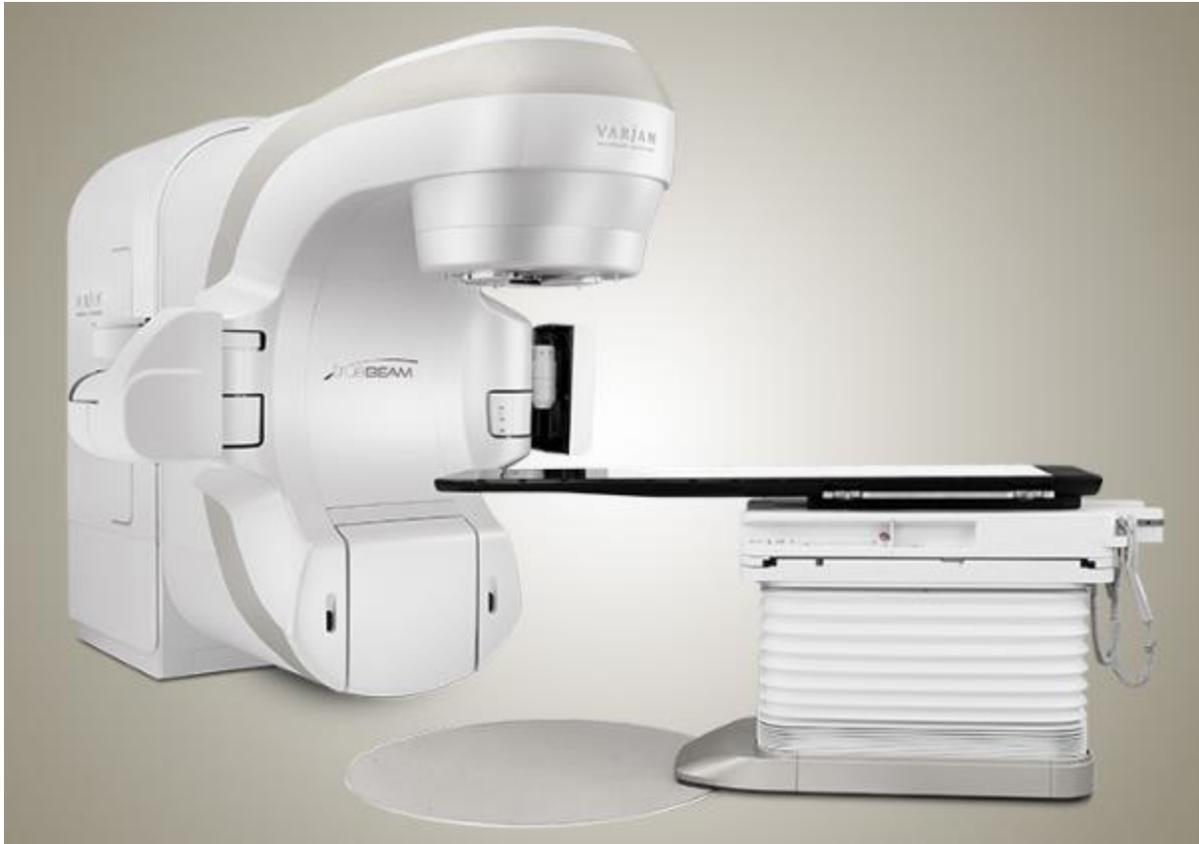


Figure 2-6. Linear accelerator with cone-beam computed tomography (CBCT).

True Beam, Varian Medical Systems, Palo Alto, CA.

2.1.4 Rigid Registration

Image registration involves finding the geometric transformation between corresponding anatomical points in two images. These two images are referred to as the reference dataset and the target dataset. Image registration in radiation oncology is used for image-guided treatment, contour propagation, dose accumulation, and integrating multi-modality images such as CT, magnetic resonance imaging (MRI), and positron-emission tomography (PET). The CBCT acquired just before radiation therapy allows for proper repositioning of the patient to the radiation beam by shifting and possibly slightly rotating the treatment couch. The applied

translation and rotation can be visualized with image fusion tools after the rigid image registration is calculated between the CBCT and the CT. Rigid image registration aims to optimize six transformation parameters: a translation in the superior-inferior, left-right, and anterior-posterior directions, and a rotation around each axis relative to the coordinate system origin. These three translations and three rotations are applied to the target image in order to align it to the reference image. This rigid transformation can then be applied to the patient to be accurately positioned prior to treatment. The optimal alignment prioritizes the alignment of a reference structure, such as a specific normal tissue or the tumor, or as a compromise between the alignments of different structures. An example alignment between the two images can be shown in a fused image (Fig. 1.7).



Figure 2-7. Rigid alignment of CT images.

A planning CT (left) and pre-fraction CBCT (middle) are rigidly aligned. A fused image (right) can be generated based on rigid registration. The purple represents the CT and the green represents the CBCT.

2.1.5 Radiation Dose Calculation

The International Commission on Radiation Units and Measurements (ICRU) recommends a combined dose uncertainty below 5% [76]. As dose calculations from the TPS, machine calibration, and patient setup carry additional uncertainty, the dose calculation algorithm itself must estimate the dose distribution within an accuracy of 3%. The most accurate dose calculation

method is Monte Carlo, which simulates photon and electron transport to model the beam [81] . This method involves determining the history of each particle using a random number generator. The dose distribution is generated by combining the energy deposited by each particle. Monte Carlo methods for dose calculations are computationally expensive and currently not practical in the clinic, however they are used as benchmark calculations for other methods [82] . The other common dose calculation algorithms can be classified under two types. One type involves models based on equivalent tissue-air ratio (ETAR) [83] or the scaling of equivalent path length (EPL) [84] for inhomogeneity corrections. These techniques do not model the differentiations in lateral electron transport. The second type is based on convolution techniques, where inhomogeneities are managed with EPL correction or scaled kernels. The lateral transport of electrons is approximated for these techniques. The first type does not result in accurate dose distribution calculations in patients, but can be used in some TPSs to give an estimate of the absorbed dose. The second type results in dose calculations near the accuracy of Monte Carlo simulations, but have the benefit of being less time-consuming.

2.1.6 Dose calculation algorithms

The dose calculation algorithms used in this thesis were based on the treatment planning system used for those patients. One treatment planning system was Pinnacle (Philips, U.S.A). The dose calculation algorithm implemented in Pinnacle is the collapsed cone convolution (CCC) method. The CCC method employs angular discretization of the kernel to allow energy transport and deposition. The following equation describes the energy deposition per radial distance based on

this method where the dose fractions are scaled for the point kernel h_{ρ_o} , for a homogeneous medium with a mass density of ρ_o :

$$\iint \frac{h_{\rho_o}}{\rho}(r, \Omega)r^2 d^2\Omega = A_{\Omega_i}e^{-a_{\Omega_i}r} + B_{\Omega_i}e^{-b_{\Omega_i}r} \quad (2-4)$$

where Ω_i is a discrete angular sector (cone), ρ is the density, r is the radius, and A and B are elements. Due to interactions as the lower cone's vertex, energy meant to be deposited in voxel B' is deposited in voxel B, as a consequence of using the CCC approximation. Although the displacement increases with distance, the first scatter fraction decreases as the distance increases, which allows for the conservation of total energy deposition. Displacement errors occur within voxels, as most of the energy is deposited where it is released. It can be noted from this equation that due to the cone's increasing cross section, the radius's inverse square cancels as the radius increases. The terma distribution, where terma can be defined as the total energy released per mass [85], is convolved with the angular discretized kernel. The cones are essentially collapsed into their axes, as we approximate the energy release to be rectilinearly deposited in the volume elements of the axis. To represent the cone axes, a transport line lattice is created to incorporate the irradiate volume so every calculation voxel is intersected. For all discrete directions of the collapsed kernel, there must be a parallel subset of lines. We can express analytically the energy transport along a line, because the kernel is described exponentially. The resulting recursive formulas need to be evaluated just once for each voxel on the line. During this recursion, kernel scaling for heterogeneities is completed for scatter and primary dose kernels, and the recursion pass each voxel a minimum of one time for each direction. In the case that the points are individually calculated, the amount of operations is proportional to MN^4 , where M represents the

number of angular bins. The CCC method has been compared with dose distributions generated by Monte Carlo simulations with generally satisfactory results. However, in off-axis dose profiles, the differences between CCC and Monte Carlo increase. In low density regions, differences of up to 10% have been observed [86].

Another treatment planning system used for the studies described in this thesis was Eclipse (Varian Medical Systems, Palo Alto, CA). The dose calculation algorithm implemented in Eclipse is the analytic anisotropic algorithm (AAA). This superposition algorithm is similar to a pencil beam algorithm and represents patient scatter and beam properties with a multiple-source model. Small beamlets (β) make up the clinical beams. The calculation grid size is related to the beta cross sections. The AAA method assumes that the dose at a point is the sum of the doses of a depth and lateral part, perpendicular to the beamlet direction. The depth dependent functions of the convolution are calculated on the beamlet's central fan line. The scatter kernels that model energy transport are discretized in 16 directions (lateral). Equation (Eq. 2-5) shows the absorbed energy ($E_{i,\beta}$) for a particle (i) and area β ,

$$E_{i,\beta}(\xi, z) = \Phi_{i,\beta} I_{i,\beta}(z) \int K_{i,\beta}(u - x', v - y', z') du dv \quad (2-5)$$

where an equivalent path is represented by ξ , the depth is represented by z , the fluence is represented by $\Phi_{i,\beta}$, the attenuation function is represented by $I_{i,\beta}(z)$, and the lateral energy transport is represented by $K_{i,\beta}$. Scaling the beamlet's attenuation function allows for the heterogeneities of the tissue to be incorporated. The AAA algorithm reduces computational time with analytical convolution, which is made possible by using analytical functions only [87]. Two

previous studies [87,88] showed that the deviations between CCC, AAA, and Monte Carlo algorithms are within 5% of Dmax for clinical scenarios.

As these algorithms are not 100% accurate, the estimated deviations due to patient motion explored in this thesis will be affected. The greatest errors within errors exist near the photon beam edges, as the spatial derivatives are largest in these regions. While Monte Carlo dose calculation algorithms are known to perform the most accurate dose calculations, they are time consuming and computationally expensive, and the CCC and AAA methods are known to result in dose calculations within an acceptable margin of error in the clinic. Dose calculation errors can be greatest in regions of high geometric complexity, such as the head and neck region. One study [89] compared the accuracy of a finite size pencil beam algorithm with EPL inhomogeneity correction with a Monte Carlo algorithm. Statistical significance was not observed for the mean and maximum doses to the GTVs when comparing the pencil beam and Monte Carlo method. Differences in the D50, mean dose, and maximum dose were not statistically significant between the two methods for the spinal cord and mandible. Differences in the D50, mean dose, and maximum dose were 4%, 4%, and 1%, respectively for the parotid glands. Additionally, a 6% difference was observed in the volume that received over 25 Gy, V25. Another study [90] used ArcCHECK to test the 3D gamma passing rate for three dose calculation algorithms, using the 3 mm/3% criteria. The planned dose was compared to the measured dose for a pencil beam algorithm, a CCC algorithm, and a Monte Carlo algorithm. For the pencil beam algorithm, a $91.75 \pm 9.12\%$ passing rate was observed. For the CCC algorithm, a $93.12 \pm 7.75\%$ passing rate was observed. For the Monte Carlo algorithm, a $94.52 \pm 5.85\%$ passing rate was observed.

Recalculation of the daily dose distribution (Fig. 1.8) can be done using CBCT. This dose calculation is more challenging than on the planning CT due to the limited field of view (FOV) and to artifacts. CBCT artifacts are caused by inconsistencies between the physical imaging process and mathematical modeling. Artifacts are evident when the CT numbers in the reconstructed image are inconsistent with the attenuation coefficients in the tissue being imaged. Examples of artifacts that occur for CBCT are shadowing and scattering effects. Artifacts cause contour uncertainty as tissue boundaries can be unclear depending on the location of the artifact. However, previous work [91] has demonstrated that dose can be calculated on the CBCT within an accuracy of 2% (within a specified region) compared to calculations on the CT. The dose distribution allows for the delivered dose assessment during the treatment fraction to the target and OAR.

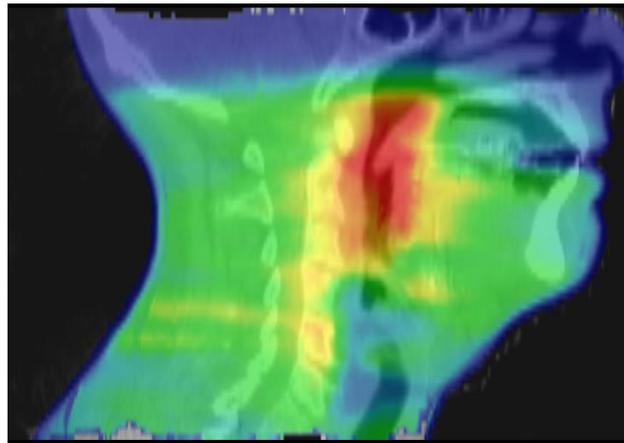


Figure 2-8. Delivered dose distribution on a head and neck (HN) cone-beam computed tomography (CBCT).

The red refers to the high dose region, the yellow refers to a moderate dose region, and the green and blue refer to low dose regions. The high dose region is sculpted around the tumor.

2.2 Accounting for Anatomical Variations

A leading cause in the uncertainty in the delivered dose to the patient is anatomical variation between fractions that occur over the course of treatment. Anatomical variation can be due to patient positioning, physiological processes (e.g. digestion, breathing, etc) weight loss, or volumetric changes of normal and tumor tissue due to radiation. When tumor shrinkage and other anatomical variations occur, it becomes challenging to determine if the dose delivered to that structure is consistent with the planned dose (Fig. 1.9). Velec et al [25] demonstrated that dose differences to the liver could exceed 5%. Recent studies have also been conducted to show the differences between planned and delivered dose for head and neck cancer, including Heukelom et al [92], who demonstrated that deviations of $>3\%$ in normal tissue complication probability (NTCP) between planned and delivered dose occur in 25% of head and neck patients. Deformable image registration (DIR) improves our ability to quantify discrepancies between planned and delivered dose. Quantification of these discrepancies allows investigation into new replanning strategies.

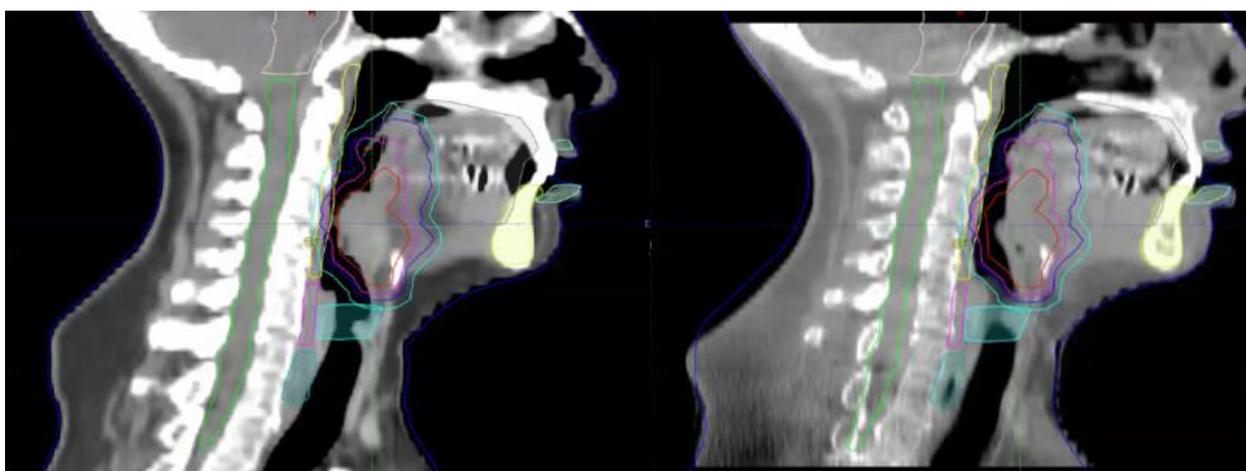


Figure 2-9. Planning CT (left) and mid-treatment CBCT (right) with the same contours overlaid.

The tumor contour and high dose region is delineated in red and no longer fits around the target tissue on the mid-treatment CT.

2.2.1 Deformable Image Registration

While rigid registration involves translation and rotation in all directions, DIR allows more degrees of freedom by allowing a spatially variant transformation. DIR algorithms are considered parametric when the transform is defined by a limited number of parameters, for example, when using B-spline functions [93]. DIR algorithms are considered non-parametric when the voxel displacements are independently calculated and regularized in an iterative process, for example, in the Demons algorithm [94]. All DIR algorithms provide a dense deformation vector field (DVF), i.e., a displacement for each voxel of the reference image. Since there exists a displacement vector for each voxel in the source image, the number of degrees of freedom can be up to three times the number of voxels in the source image. A regularization function constrains the spatially variant vector fields to establish that the transformation is realistic anatomically. The regularization function allows for a smooth deformation field.

The deformation between the CBCT and the planning CT must be calculated in order to align images. Accurate alignment of images allows for the calculation of dose accumulated by different structures and eventually to decide if re-planning is necessary. Calculating the deformation between the CBCT and CT can be done with DIR which provides a DVF that maps each voxel of the planning CT onto the CBCT.

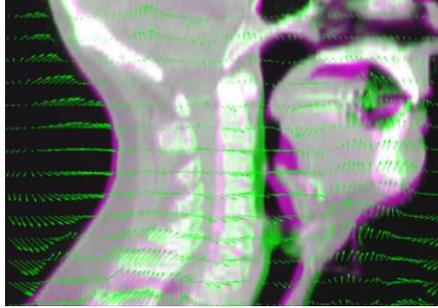


Figure 2-10. Deformation field obtained with a DIR algorithm that can be used to spatially align images. The purple and green each represent different images.

Registration metrics are used to assess how well two images are aligned. DIR is the optimization process that provides the transformation (DVF) that will maximize or minimize the registration metric. Two main categories of DIR are used in the clinic and for research. The most common DIR techniques used clinically are intensity-based, which use functions of the image intensities as similarity metrics such as the following methods.

Sum of Squared Differences

In the sum of squared differences (SSD) approach, the secondary image is transformed to the primary image to minimize the SSD (Eq. 2-6):

$$SSD(I_1, I_2) = \sum_{x \in \Omega} (I_1(x) - I_2(x))^2 \quad (2-6)$$

where all voxels x within an image grid Ω are considered. The method is based on the intensity of a voxel in the primary image (I_1) corresponding to the intensity of a voxel in the secondary image (I_2). This method is inexpensive in computing time, and is only applicable in cases where both images are of the same modality.

Cross-Correlation

Cross-correlation (CC) is applicable for the case that there is a linear relationship between the two images [95]:

$$CC(I_1, I_2) = \frac{\sum_{x \in \Omega} (I_1(x) - I_2(x))}{\sqrt{\sum_{x \in \Omega} I_1(x)^2 * \sum_{x \in \Omega} I_2(x)^2}} \quad (2-7)$$

Mutual Information

Mutual information (MI) is a useful method for when each image is of a different modality:

$$MI(I_1, I_2) = H(I_1) + H(I_2) - H(I_1, I_2) \quad (2-8)$$

where $H(I_1)$ and $H(I_2)$ (Eq. 2-9) represent the marginal entropies and $H(I_1, I_2)$ (Eq. 2-10) represents joint entropy.

$$H(I) = - \sum_i p(i) \log p(i) \quad (2-9)$$

where i spans all intensities in I .

$$H(I_1, I_2) = - \sum_{\mathbf{i}} p(\mathbf{i}) \log p(\mathbf{i}) \quad (2-10)$$

Where i spans all intensity pairs between I_1 and I_2 . The probabilities are estimated from the images.

Another technique is geometry-based DIR, which does not directly use voxel intensities but uses features extracted from the images such as landmarks or anatomical surfaces. Some approaches have been proposed to combine intensity-based and geometry-based DIR. One such approach used in this Ph.D. and implemented in the treatment planning system RayStation is the ANAtomically CONstrained Deformation Algorithm (ANACONDA) [96]. This hybrid approach combines intensity information as well as contours defined manually on anatomical surfaces. For the ANACONDA method, a smooth DVF is optimized using the quasi-Newton algorithm. Similarity between the images is determined by correlation coefficients, which guide the DVF. Regularization of the DVF is done by the weighted Dirichlet energy, which measures variability of the function. Regularization with this method involves first resolving the DVF smoothness and invertibility, and then penalizing large deviations in the regions of interest (ROI). Controlling ROIs can be chosen by the user to guide the DIR, which in addition to the intensity-based part, makes it a hybrid algorithm. Deformation of each ROI surface is computed by registering distance maps representing the surface.

2.2.2 Biomechanical Model-Based DIR Algorithms

Biomechanical model-based DIR uses material properties of the tissue in the image to deform the image. These material properties define the stiffness and flexibility of the tissue. The biomechanical model-based DIR discussed in this dissertation is based on finite element models

(FEM) of the structures depicted in the images. In a FEM-based DIR algorithm, each ROI is represented by a series of joined nodes that create tetrahedrons to approximate the surface. FEM analysis solves the geometric relationship between tetrahedrons. For biomechanical model-based DIR, boundary conditions define motion and deformation of the tissue.

FEM-based DIR algorithms are particularly useful compared to intensity-based algorithms for multi-modality applications and in cases where anatomical deformation is complex. FEM-based DIR algorithms allow for the implementation of dose-based boundary conditions to describe the volumetric response of tissue and have been validated for several organ sites. Polan et al [97] applied dose-driven volumetric response to an FEM-based DIR algorithm to improve the accuracy of modeling anatomical changes in the liver. Velec et al [98] demonstrated that the accuracy of FEM-based DIR was near image voxel resolution accuracy for multiple image modalities for the thorax, abdomen, and prostate. Another advantage of using FEM-based DIR algorithms is the ability to apply internal boundary conditions to organs. Cazoulat et al [99] added boundary conditions to lung vasculature and significantly improved the accuracy in modeling lung and tumor response during radiation therapy. Yan et al [100] developed a biomechanical model of an elastic body in order to quantify the organ motion of radiation therapy patients.

The biomechanical model-based DIR algorithm used for the studies described in this thesis was Morfeus [63], which is based on organ segmentations. With Morfeus, we can create a dense DVF between the reference image and target image. Based on contours, triangular surfaces meshes are

generated to represent the organs of interest. Material properties (Young's Modulus and Poisson's Ratio) are assigned to the organs of interest. Soft tissues are modeled as linear-elastic materials, which restore to their initial state after the load is removed. We characterize the material using the elastic modulus (E), which is the slope of the plot of stress over strain, and Poisson's ratio (ν), which is the ratio of the transverse strain to the axial strain. The shear modulus, G, is defined by shear stress divided by angular deformation. The bulk modulus, K, is defined by the load divided by the volumetric change. The aforementioned variables are related in the following equations.

$$G = \frac{E}{2(1 + \nu)} \quad (2-11)$$

$$K = \frac{E}{3(1 - 2\nu)} \quad (2-12)$$

Hooke's law represents the relationship linking stress and strain for elastic models, and is defined in the equation

$$\sigma = C\varepsilon \quad (2-13)$$

where the stress and strain are represented by σ and ε , and the elastic moduli is represented by C.

A tetrahedral FEM model of all organs is generated. For each structure, a surface projection algorithm determines the displacement of the surface nodes. Those displacements are used as boundary conditions in the model. Two different surface projection algorithms have been

previously implemented in Morfeus, in order to determine boundary conditions. The first method is HYPERMORPH, which is an algorithm in HYPERMESH, which is a FEM pre-processor from Altair Engineering in Troy, MI. HYPERMORPH relies on the FEM surface meshes generated based on the organs. The secondary mesh is converted to a surface. Next, an initial center-of-gravity (COG) registration is completed between the two meshes. The surface tetrahedrals are grouped into domains that depend on the FEM's curvature. The nodes are then guided in each domain by a "handle". The base model domains are mapped to the secondary model domains, and the deformable alignment is guided by the handle and then an orthogonal node projection. This alignment is performed from node to surface. Finally, boundary conditions are based on the vector difference of each node, before and after alignment. Another surface projection algorithm has been recently added to handle organs presenting large deformations. This algorithm is based on the accelerated Demons algorithm. The Demons algorithm [94] is a greyscale image-, or intensity-based method. Because of the similar intensities between images, this method is used for registration between CT images in RT. The forces of the Demons method can be described using the optical flow equation. The following equation describes the displacement \vec{u} of a point from the reference image, R, to the point corresponding from the target image, T:

$$\vec{u} = \frac{(t - r)\vec{\nabla}r}{|\vec{\nabla}r|^2 + (t - r)^2} \quad (2-14)$$

where \vec{u} represents the displacement in each direction, t and r represent the intensities of the target and references images, and $\vec{\nabla}r$ represents the reference image gradient. Iteratively, the

Demons algorithm solves the equation (Eq. 2-15) for each voxel and composes the displacement field \vec{u} with the global displacement field \vec{U} :

$$\vec{U}_n = \vec{u}_n + \vec{U}_{n-1} \quad (2-15)$$

In order to minimize noise and ensure smoothness of the transformation the global DVF is filtered with a Gaussian kernel. Since the deformation is driven only by the gradient of the reference image, the method was improved in [101] with the addition of an active force to the deformed (target) image. This active force, \vec{f}_t , can be described in the following equation.

$$\vec{f}_t = - \frac{(r - t)\vec{\nabla}t}{|\vec{\nabla}t|^2 + (r - t)^2} \quad (2-16)$$

Finally, the total force of a point can be described with the equation

$$\vec{f} = \vec{f}_r + \vec{f}_t = (t - r) * \left(\frac{\vec{\nabla}r}{|\vec{\nabla}r|^2 + (r - t)^2} + \frac{\vec{\nabla}t}{|\vec{\nabla}t|^2 + (r - t)^2} \right) \quad (2-17)$$

where \vec{f}_r represents the force from the gradient of the reference image and \vec{f}_t represents the force from the gradient of the target image. These force calculations are appropriate with relatively small deformations. However, there are often large deformations in clinical cases, so multiresolution method [102] was applied. This method applies an iterative demons diffusion

approach, based on low-resolution images, and increases calculation speed by improving the convergence. An additional normalization factor (a) was also applied. With this incorporation of a , adjustments to the force strength at each iteration is possible. This factor is applied as shown in the following equation

$$\vec{d} = (t - r) * \left(\frac{\vec{\nabla}r}{|\vec{\nabla}r|^2 + a^2(r - t)^2} + \frac{\vec{\nabla}t}{|\vec{\nabla}t|^2 + a^2(r - t)^2} \right) \quad (2-18)$$

where larger deformations require smaller a values.

Morfeus was used in this work to align the parotid glands after both positional changes and dose response. Using Morfeus to describe the changes during radiation therapy is important because FEM-based deformable image registration algorithms can be used to show not just the changes in the surface of the organs, but also to model what is happening internally. The Morfeus process is outlined in Fig. 2-13.

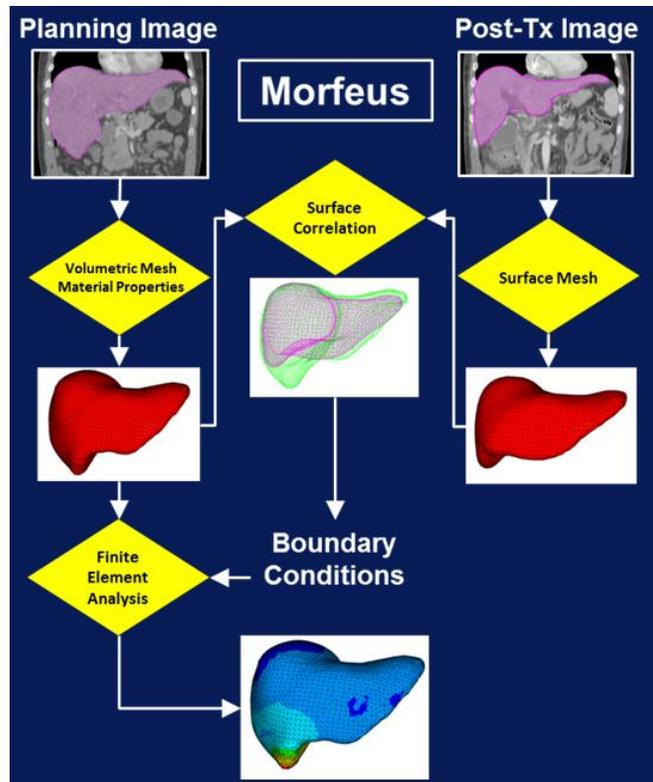


Figure 2-11. Depiction of Morfeus process.

Using a pre-treatment (initial) and a post-treatment (secondary) image, Morfeus aligns the initial image to the corresponding secondary image. A volumetric mesh is generated and material properties are assigned to the initial image. Then, a surface mesh is generated from the secondary image. The surface correlation is calculated and applied as boundary conditions to the initial model. Finally, the finite element analysis (FEA) solves the deformation between the initial image and the secondary image.

2.2.3 Dose Accumulation

Upon completion of any fraction during radiation therapy, the dose (d) can be accurately accumulated using DIR. A CBCT from a daily treatment fraction is deformed to the planning CT. The DVF (U) is then used to deform the dose from the CBCT onto the planning CT, which allows the comparison between the planning dose and delivered dose. Ultimately, accumulating the dose using DIR can indicate the need for re-planning the patient. After completion of treatment, assessment of the treatment efficacy is performed based on the acquisition of follow-

up images. These images need to be registered to the image on which the accumulated dose has been calculated (Fig. 1.11) to better understand the correlation between delivered dose and outcome. The ultimate goal is to optimize treatment strategies for future patients. The following equation describes the daily accumulated dose:

$$d(x) = D(x + U(x)) \quad (2-19)$$

where d is the daily accumulated dose at the voxel coordinate x , and D is the deformed daily dose distribution based on the DVF U . This daily accumulated dose is calculated for each fraction and all doses are summed to calculate the total accumulated dose.



Figure 2-12. Follow-up CT scan with accumulated dose overlaid.

2.3 Adaptive Radiation Therapy

As patients progress through radiation therapy, the tumor and normal tissue may change due to weight loss and tumor response. As tissues change during radiation therapy, the radiation

treatment plan may no longer correspond to the patient anatomy. Continuing to use the original treatment plan may lead to an unacceptable dose deviation. In this case, creating a new treatment plan that is optimal for the new patient geometry is necessary. This process, ‘adaptive radiotherapy’, is the process where a new radiation treatment plan is triggered based on new information acquired during the radiation therapy course. For this, a new CT scan is usually acquired on which the contours are delineated to optimize a new treatment plan, based on the new anatomy of the patient.

Chapter 3. A Simulation Study to Assess the Potential Impact of Developing Normal Tissue Complication Probability Models with Accumulated Dose

Current practice in radiation oncology involves the use of historic population models for predicting toxicity. These models are based on the planned dose to a specific organ. Recent studies explored the deviations between planned and delivered dose to the liver [62,64].

When NTCP models were first developed, accurate dose accumulation to estimate the true delivered dose to the organ over treatment was not possible. However, accurate dose accumulation is now possible, and the same classic, planned dose-based models are still being used in the clinic. The radiation dose to liver tumors is often limited by the surrounding normal tissue. As the normal tissue in the abdominal region consists mainly of serial structures, which lose function completely when any part of the organ loses function, the priority of radiation treatment for liver cancer is to spare the gastro-intestinal organs. Limiting the dose to normal tissue can potentially limit the dose to the tumor and therefore tumor control. As there are potential deviations between the planned and delivered dose, the planned dose-based model is not necessarily accurate, and with the possibility of dose accumulation in clinical trials, it is now possible to generate more accurate toxicity models. This chapter explores the potential clinical impact of generating new toxicity models based on the accumulated dose to the patient.

3.1 Abstract

The purpose of this study is to analyze the potential clinical impact of differences between planned and accumulated dose on the development and use of normal tissue complication probability (NTCP) models. Thirty patients previously treated with SBRT for liver cancer and for whom accumulated dose was computed, were assessed retrospectively. The linear quadratic equivalent dose at 2Gy per fraction (LQED2) and generalized equivalent uniform dose (gEUD) were calculated for planned and accumulated dose. Stomach and duodenal Lyman-Kutcher-Burman (LKB) NTCP models ($\alpha/\beta=2.5$, $n=0.09$) were developed based on planned and accumulated gEUD and the differences between the models assessed. In addition, the error in determining the probability of toxicity based on the planned dose was evaluated by comparing planned doses into the NTCP model created from accumulated dose. The standard, planned-dose NTCP model overestimates toxicity risk for both the duodenal and stomach models at doses below approximately 20Gy (6-fractions), and underestimates toxicity risk for doses above approximately 20Gy (6-fractions). Building NTCP models using accumulated dose rather than planned dose changes the predicted risk by up to 16 % (Mean=6%,SD=7%) for duodenal toxicity and 6% (Mean=2%,SD=2%) for stomach toxicity. For a protocol planning a 10% iso-toxicity risk to the duodenum, a 15.7Gy (6-fractions) max dose constraint would be necessary when using standard NTCP models based on planned dose, while a 17.6Gy (6-fractions) max dose would be allowed when using NTCP models based on accumulated dose. Assuming that accumulated dose is a more accurate representation of the true delivered dose than the planned dose, this simulation study indicates the need for prospective clinical trials to evaluate the impact of building NTCP models based on accumulated dose.

3.2 Introduction

Stereotactic body radiation therapy (SBRT) delivers high precision external beam radiation therapy (EBRT) treatment in 2-5 fractions. The goal of using SBRT is to deliver an ablative dose to the tumor while sparing normal tissue leading to lower toxicity [103]. While SBRT has been demonstrated to increase local control of liver cancer [104], toxicity risks must still be acknowledged for liver and luminal gastrointestinal (GI) structures [105]. Dose escalation has the potential to improve local control [106]. However, increase in dose is often limited by normal tissue toxicity risk. Based on previous studies [25], approximately 30% of patients are dose limited based on GI toxicity. Overestimating toxicity risk can lead to conservative treatment for the patient, potentially leading to lower chances of tumor control. Underestimating toxicity risk can subject the patient to unplanned risks. It is therefore critical to have an accurate understanding of toxicity to normal tissue. Significant efforts have been made in the development and validation of accurate Normal Tissue Complication Probability (NTCP) models [107-110], which aim to characterize the correlation between dose and the likelihood of side effects [111]. Specifically, Lyman-Kutcher-Burman (LKB) NTCP models have been used to investigate the dose-volume response for liver cancer [112].

LKB NTCP models have been previously developed for duodenal toxicity [108] demonstrating that the model can predict outcomes after SBRT. In addition, a separate investigation showed that the Lyman NTCP model can predict gastric bleeding [107], demonstrating that patients with cirrhosis are at increased risk. Both studies built the NTCP models based on the planned dose to each patient. Recent retrospective studies have shown that the planned dose differs from the

accumulated dose [25,113,114], especially when evaluating the maximum dose to luminal GI structures, such as the duodenum and stomach, for which the differences could reach 42% and 14%, respectively, of the prescribed dose [25]. It is hypothesized that building NTCP models based on accumulated dose will lead to a substantially different model, which may have clinical impact. The goal of this work is to evaluate how the known uncertainty between planned and delivered dose limits translates into potential uncertainties in NTCP modeling, and to determine the differences in the model if accumulated dose was used in the derivation of the model parameters.

Previous work supports the hypothesis that accumulated dose can improve understanding of clinical outcomes of SBRT, demonstrating the accumulated dose to the GTV more strongly predicts for total time to local progression [114]. The current study builds on this work, assessing the impact of accumulated dose on NTCP models by comparing LKB NTCP models based on planned dose with those based on accumulated dose. Preceding clinical trials, the aim of this study is to assess the potential differences in the development of the NTCP models using accumulated dose versus planned dose. The second aim is to apply these models for patient specific assessment.

3.3 Methods and Materials

3.3.1 Patient Data

A previous study retrospectively evaluated the deviations between planned and accumulated dose to tumors and normal tissues in liver SBRT in 30 patients [25]. The patients were treated on an institutional review board-approved trial under free-breathing conditions, with 6 fractions to an individualized risk-based dose of 27-30Gy. Daily cone-beam CT (CBCT) guidance was performed. For each patient, the dose was accumulated by registering the daily inhale/exhale CBCT to the planning CT scan with Morfeus, an in-house biomechanical model-based deformable registration algorithm (DIR) [63]. For each of the thirty patients, planning dose and accumulated dose for stomach and duodenum were used in the Lyman NTCP model.

Dose Accumulation

The dose accumulation was previously reported [25] and will be briefly described here for completeness. Morfeus was used to accumulate the dose for each patient. The dose distribution from the static radiation treatment plan was calculated in the treatment planning system on the exhale planning CT and the inhale planning CT. Both dose distribution files were imported into Morfeus. In Morfeus, the organs are described using finite element models, which represents substructure within the organs through tetrahedral elements. Through the DIR, the locations of the tetrahedral elements can be tracked between the exhale and inhale images of 4D scans. Interpolation of dose matrices onto the position of each element at exhale, inhale, and four intermediate phases is performed to accumulate the dose. The weighting of each phase was determined by the time spent at that phase and the elemental position in the breathing cycle. Finally, a summation of elemental dose over the breathing cycle was calculated to determine the accumulated dose.

The planned dose is defined as the static clinical planned dose, which was found by interpolating the dose matrix from the exhale CT onto the initial tetrahedral mesh constructed from the anatomy on the exhale CT. The calculation of this dose does not include any changes due to breathing motion or setup error.

Accumulated dose refers to the dose accumulated over SBRT accounting for residual setup errors (e.g. those errors still present following daily image guidance), respiratory motion, shifting of the liver, and deformation. To account for setup errors and organ deformation present at each fraction, the exhale CT is deformed to the exhale CBCT of each fraction. Next, to account for daily breathing motion, the exhale CT is deformed from the exhale CBCT to the inhale CBCT of each fraction. To accumulate the dose, dose matrices from the exhale CT and inhale CT are interpolated onto the deformation map from the exhale to inhale CBCT of each fraction. Finally, the doses from the 6 fractions of treatment are summed. In order to accumulate the dose, DIR tracked anatomical motion and deformation in the dose matrices of the initial planning 4D CT.

3.3.2 Differences between planned and accumulated doses

Percent change (PC) from planned to accumulated mean dose was analyzed for duodenum (N=30), stomach for primary liver cancer (non-cirrhotic) patients (N=15), and stomach for cirrhotic patients (N=15), calculated by the equation

$$PC_{dose} = \frac{Accumulated\ Dose - Planned\ Dose}{Planned\ Dose} * 100 \quad (3-1)$$

3.3.3 NTCP Models – Change in probability of toxicity

The delivered and accumulated doses (in the form of tabular DVH) were biologically corrected to the Linear Quadratic Equivalent Dose at 2Gy per fraction (LQED2), using an α/β ratio of 2.5 [107]. The gEUD was calculated with $n=0.09$. The Lyman NTCP model, shown in equation 3-2, was used for toxicity modeling in this study.

$$NTCP = \Phi\left(\frac{gEUD - TD_{50}}{m \cdot TD_{50}}\right) \quad (3-2)$$

The function Φ represents the NTCP model [115] where gEUD is evaluated using equation [105], $TD_{50}=24.6$ and $m=0.23$ for duodenum [108], $TD_{50}=22$ and $m=0.21$ for stomach and cirrhosis [107], and $TD_{50}=56$ and $m=0.21$ for stomach and no cirrhosis [107].

The percent change (PC) in NTCP from planned to accumulated dose was calculated using the equation

$$PC_{NTCP} = \frac{NTCP_{Accumulated} - NTCP_{Planned}}{NTCP_{Planned}} * 100 \quad (3-3)$$

3.3.4 Simulated Toxicities for Duodenum and Stomach

Toxicity models were simulated for duodenum and stomach with cirrhosis. Stomach without cirrhosis was not modeled, as there were no patients with an absolute change in NTCP of greater than 5% between the accumulated and planned dose. In order to simulate toxicity models for a larger cohort of patients than was used in the study, a resampling process was developed.

Figure 1 depicts the planned dose over the accumulated dose for the original 30 patients for duodenum, and the 15 cirrhosis patients for stomach. Based on the trends seen in Figure 1, a model was fit to predict planned dose given accumulated dose in order to resample data points.

$$d_{plan} \sim N(d_{acc} + \beta d_{acc}, \sigma^2 d_{acc}) \quad (3-4)$$

Where d_{plan} is planned dose, d_{acc} is the accumulated dose, and β (bias function of dose) and σ^2 are parameters estimated by the linear regression model shown in Figure 1. These parameters were estimated via maximum likelihood.

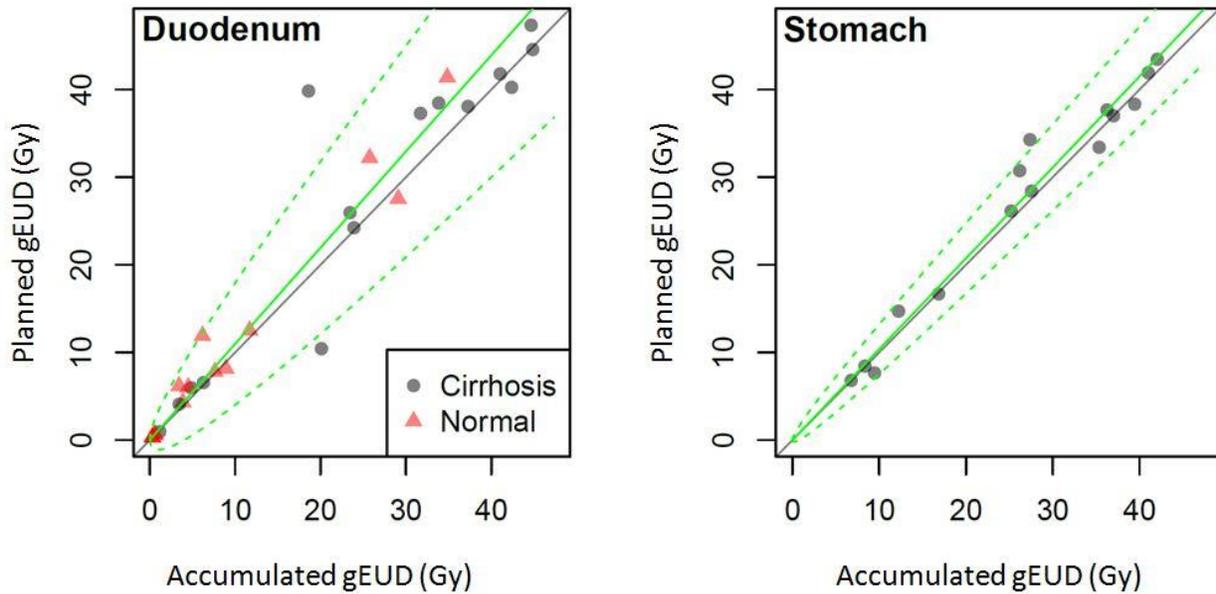


Figure 3-1. Planned Vs. Accumulated Dose Plots.

Planned Vs Accumulated Dose for Duodenum (a) and Stomach with Cirrhosis (b). The solid green line depicts the mean, the dashed line shows the 95% confidence limits for regression, and the solid gray line represents the case that planned dose and accumulated dose are equal.

NTCP models are typically estimated using planned dose values. To quantify the effect of differences between planned and accumulated doses on model parameter estimation, the following steps were performed:

1. The distribution of planned dose was estimated as a function of accumulated dose, which involved both estimation of systematic bias (i.e. $\beta \neq 1$ in Eq. 6) and random variation of planned dose about accumulated dose. In the case that the two dose terms agreed perfectly, there would be neither bias (i.e. $\beta=1$) nor variation (i.e. $\sigma=0$).
2. Both toxicity values were simulated (using the accumulated dose NTCP model) and planned dose values (using Eq. 6) from the observed delivered dose values in the study ($n=30$). Larger sample sizes can overcome variation between planned and accumulated dose. However, larger sample sizes do not overcome systematic bias.

Thus, studies of size 30, 150, or 600 patients were simulated by using each observed patients delivered dose value 1, 5, or 20 times. Because a random toxicity outcome and random planned dose value was simulated using the relationship established in Eq. 6, the multiple outcomes simulated from a single patient in the 30 patient dataset were distinct.

3. For each of the simulated studies, the NTCP model parameters were estimated based on the planned dose values.
4. Steps 2 and 3 were repeated 2,000 times. Finally, the mean NTCP curve was plotted as well as the 10th and 90th percentiles of the curves.

3.3.5 NTCP Models – Change in probability of toxicity

Finally, the deviations in the probabilities of toxicity between the standard and accumulated NTCP models were quantified. As the true probability of toxicity risk is unknown, this was estimated with the probability of toxicity risk of the d_{acc} derived from the accumulated dose model (NTCP₁). The d_{plan} and d_{acc} (in maximum dose to 0.5cc) of each patient in the study were correlated to the probability of toxicity risk using each NTCP model. NTCP₂ was defined as the probability of toxicity risk of d_{plan} derived from the standard model and NTCP₃ was defined as the probability of toxicity risk of d_{plan} derived from the accumulated model. The error in the standard model was calculated as the difference between NTCP₂ and NTCP₁. The error in the accumulated model was calculated as the difference between NTCP₃ and NTCP₁.

3.4 Results

The previously reported deviations of 5% or greater in 70% of the patients [25] translates to a deviation of greater than 5% NTCP for 57% of the patients for duodenum and 60% of patients for stomach with cirrhosis in the current study.

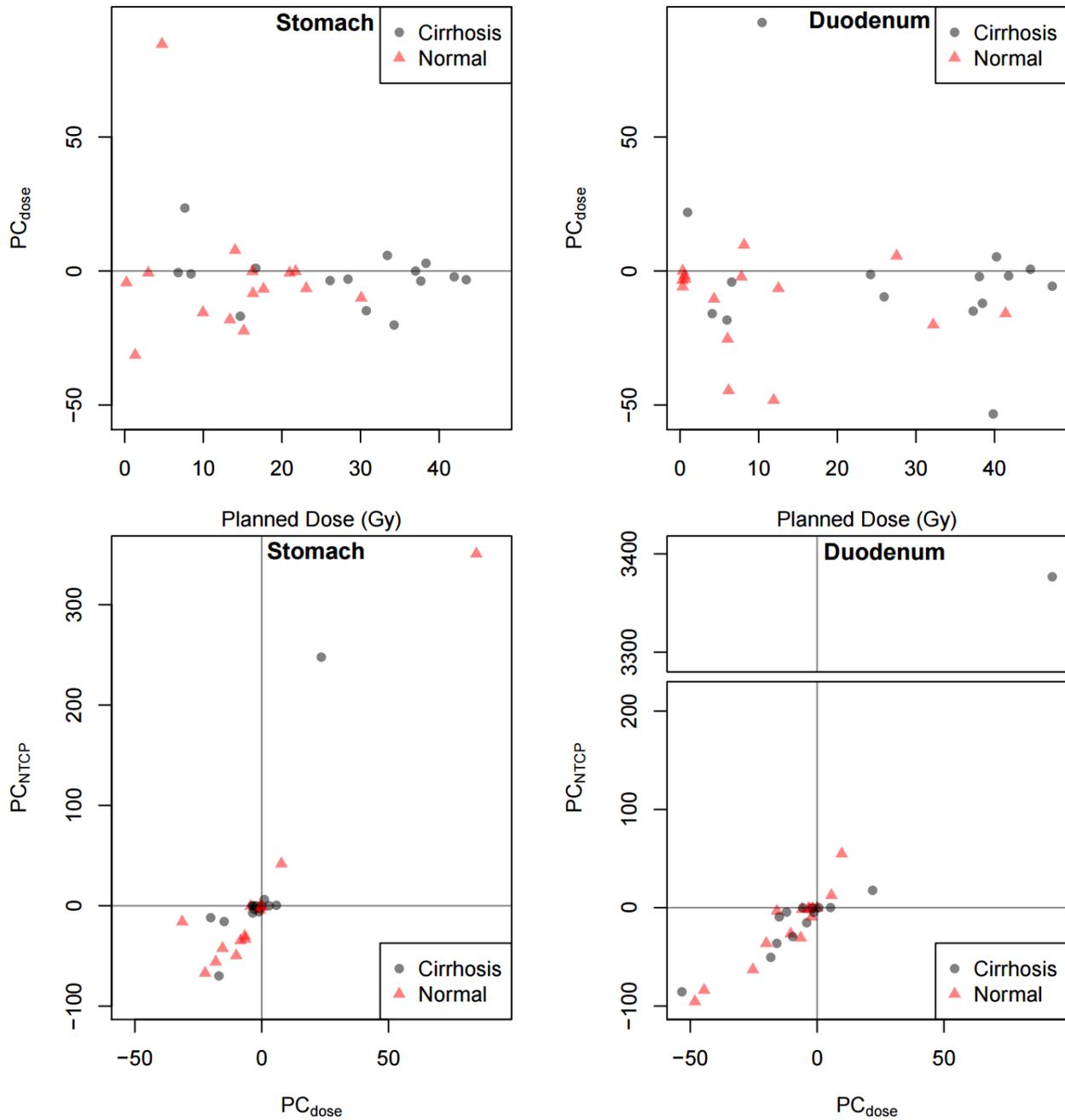


Figure 3-2. Percent Change in Dose and NTCP.

Percent Change from Planned to Accumulated Dose (PC_{dose}) by Planned Dose (Gy) for Stomach and Duodenum and Percent Change in NTCP by Percent Change in Dose (PC_{NTCP}) for Stomach and Duodenum. The solid gray line represents a zero percent change in dose.

3.4.1 NTCP Models for Duodenum and Stomach

For duodenal toxicity, the difference between NTCP based on accumulated dose and NTCP based on planned dose is substantial (Figure 3). Table 1 shows the results from deriving error from the toxicity models for duodenum for the 30 patients with dose accumulation. It was assumed that the toxicity risk using the accumulated dose value and the probability of toxicity risk derived from the accumulated model was the most accurate assessment of the toxicity risk. The error in using the planned dose with the standard model was 6% (SD=7%), where the error in using the planned dose with the accumulated model was 4% (SD=7%).

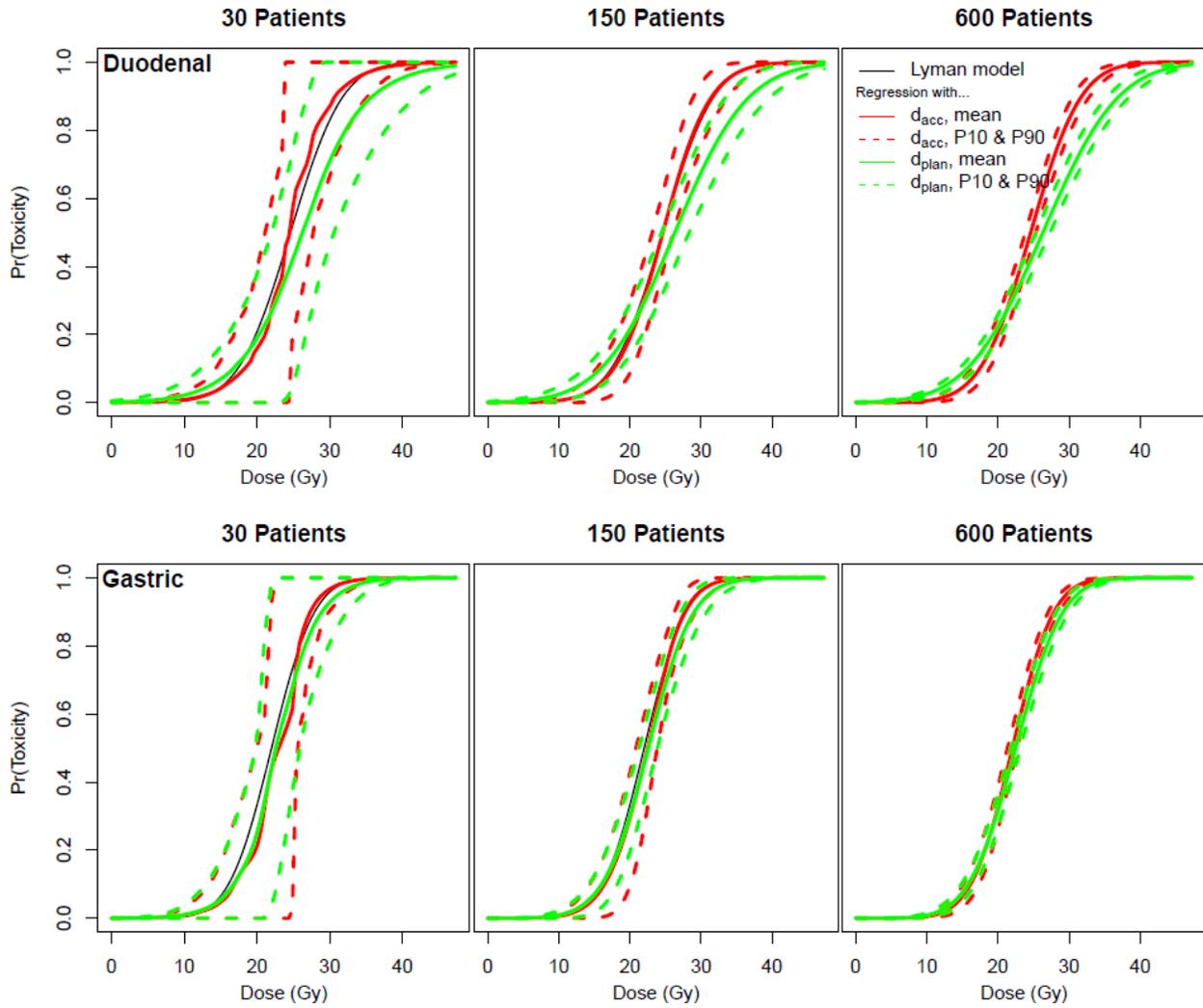


Figure 3-3. Probability of Duodenal and Gastric (with Cirrhosis) Toxicity.

Probability for simulations of 30, 150, and 600 patients. The black line represents the Lyman NTCP by dose. The solid red line represents the mean NTCP based on simulations with accumulated dose (d_{acc}). The green lines represent the mean NTCP based on simulations with planned dose (d_{plan}). The dashed lines represent the 10th and 90th percentiles.

Table 3-1. Duodenal NTCP results.

*Assumption that this model is closest to the true risk based on the most accurate measurement of delivered dose.

Patient	Dose _{planned} max dose to 0.5cc [Gy]	NTCP ₂ Dose _{planned} , Model _{std} [%]	NTCP ₃ Dose _{planned} , Model _{acc} [%]	Dose _{acc} max dose to 0.5cc [Gy]	NTCP ₁ * Dose _{acc} , Model _{acc} [%]	Difference NTCP ₂ – NTCP ₁	Difference NTCP ₃ – NTCP ₁
---------	---	---	---	---	---	--	--

1	29.91	66.61	82.93	29.41	80.55	13.94	2.38
2	25.44	45.72	55.86	20.91	25.18	20.54	30.68
3	8.63	1.74	0.28	6.67	0.10	1.63	0.18
4	6.67	0.96	0.10	5.86	0.07	0.90	0.04
5	15.32	9.34	4.91	15.12	4.58	4.76	0.33
7	29.70	65.66	81.94	29.74	82.14	16.48	0.19
8	0.79	0.13	0.00	0.77	0.00	0.12	0.00
9	10.42	2.87	0.66	9.04	0.34	2.53	0.32
11	0.69	0.12	0.00	0.64	0.00	0.12	0.00
12	9.77	2.40	0.49	9.69	0.47	1.93	0.02
13	8.70	1.77	0.29	5.65	0.06	1.71	0.23
14	0.64	0.12	0.00	0.63	0.00	0.12	0.00
15	1.94	0.19	0.01	1.59	0.01	0.19	0.00
16	9.61	2.29	0.45	9.29	0.39	1.90	0.06
17	19.07	19.16	15.94	12.36	1.55	17.61	14.39
18	32.28	76.18	91.47	27.76	71.44	4.74	20.03
19	31.22	72.08	88.16	30.92	87.08	15.00	1.08
20	32.35	76.42	91.64	30.06	83.59	7.17	8.06
21	29.59	65.20	81.45	29.52	81.07	15.88	0.37
22	30.08	67.35	83.70	28.75	77.15	9.81	6.55
23	1.32	0.16	0.00	1.26	0.00	0.15	0.00
24	6.56	0.93	0.10	5.79	0.06	0.87	0.03
25	1.47	0.16	0.00	1.30	0.00	0.16	0.00
26	22.39	31.78	34.34	20.75	24.26	7.51	10.07
27	23.26	35.59	40.27	21.12	26.39	9.20	13.88
28	24.19	39.80	46.83	22.88	37.62	2.18	9.21
29	33.22	79.52	93.76	33.82	94.94	15.43	1.18
30	31.25	72.20	88.26	30.20	84.19	12.00	4.07
31	10.15	2.66	0.58	7.00	0.12	2.54	0.46
32	9.57	2.27	0.44	8.49	0.26	2.01	0.18
					Average	6.30	4.13
					SD	6.53	7.31

For stomach toxicity, the difference between NTCP based on accumulated dose and NTCP based on planned dose is less substantial (Figure 3). Table 2 shows the results from deriving error from the toxicity models for stomach for the 30 patients with dose accumulation. The error in using

the planned dose with the standard model was 3% (SD=3%), where the error in using the planned dose with the accumulated model was 3% (SD=4 %).

Table 3-2. Stomach NTCP results:

*Assumption that this model is closest to the true risk based on the most accurate measurement of delivered dose.

Patient	Dose _{planned} max dose to 0.5cc [Gy]	NTCP ₂ Dose _{planned,} Model _{std} [%]	NTCP ₃ Dose _{planned,} Model _{acc} [%]	Dose _{acc} max dose to 0.5cc [Gy]	NTCP ₁ * Dose _{acc,} Model _{acc} [%]	Difference NTCP ₂ – NTCP ₁	Difference NTCP ₃ – NTCP ₁
1	29.21	90.23	94.19	29.24	94.26	4.03	0.08
2	7.10	0.19	0.09	9.08	0.31	0.11	0.22
3	10.31	1.06	0.64	10.09	0.56	0.51	0.08
4	17.63	17.65	16.85	17.88	18.28	0.63	1.43
5	2.86	0.01	0.00	3.22	0.01	0.01	0.00
7	20.57	35.83	37.46	19.51	29.06	6.77	8.40
8	13.94	5.17	4.05	12.36	1.91	3.26	2.15
9	19.89	31.05	31.97	19.67	30.24	0.81	1.73
11	0.61	0.00	0.00	0.61	0.00	0.00	0.00
12	12.31	2.66	1.86	11.96	1.55	1.11	0.31
13	23.55	58.30	63.03	22.20	51.50	6.80	11.53
14	23.05	54.58	58.89	22.91	57.63	3.05	1.27
15	26.44	77.80	83.33	23.63	63.77	14.03	19.56
16	19.46	28.22	28.72	19.96	32.47	4.25	3.75
17	28.18	86.43	91.14	27.78	89.65	3.22	1.48
18	28.59	88.03	92.45	26.32	82.73	5.30	9.72
19	5.62	0.08	0.03	5.56	0.03	0.05	0.00
20	31.28	95.40	97.79	29.89	95.68	0.28	2.10
21	31.07	95.02	97.56	31.65	98.20	3.17	0.64
22	32.56	97.30	98.88	31.98	98.47	1.17	0.42
23	29.23	90.30	94.24	30.09	96.09	5.79	1.85
24	31.74	96.18	98.25	31.15	97.65	1.48	0.60
25	18.63	23.06	22.86	17.96	18.72	4.34	4.14
26	27.12	81.49	86.79	26.54	83.89	2.40	2.90
27	16.79	13.84	12.74	15.43	7.66	6.17	5.07
28	22.86	53.06	57.19	22.80	56.70	3.64	0.49
29	11.14	1.58	1.01	12.15	1.72	0.14	0.71
30	30.28	93.27	96.40	28.59	92.45	0.83	3.95

31	17.66	17.82	17.04	16.96	13.49	4.33	3.54
32	17.12	15.28	14.28	15.68	8.45	6.83	5.83
					Average	3.15	3.13
					SD	3.08	4.30

In an iso-toxicity protocol with a 10% risk to the duodenum, 17.6Gy (6-fractions) would be allowed if the accumulated model was applied, while 15.7Gy (6-fractions) would be allowed if the standard model was applied. While the stomach model shows only slight differences between the standard and accumulated NTCP models at this probability of toxicity risk, the standard model still overestimates toxicity risk. With a 10% limit on toxicity risk to the stomach, the accumulated model would allow 16.1Gy (6-fractions), while the standard model would allow 15.8Gy (6-fractions).

The bias in the planned doses in Figure 3 is reflected in the differences between the solid red and green lines. Increasing sample size does not diminish this difference. The dashed green lines indicate variation in the fitted NTCP curves between trials. As the sample size increases from 30 to 600, the variation between hypothetical trials becomes very small so that almost any 600 patient study will result in nearly the same biased NTCP curve (solid green line).

The greatest deviation between the probability of toxicity risk based on the accumulated model and the probability of toxicity based on the planned model occurs at a planned dose of approximately 30Gy (6-fractions) for duodenum (Figure 4a) and 25Gy (6-fractions) for stomach (Figure 4b). This deviation is greater for duodenum (16% greater probability of toxicity risk based on the accumulated model) than for stomach (6% greater probability of toxicity risk based

on the accumulated model). For both the duodenum and stomach models, the planned model overestimates the probability of toxicity risk for planned doses until the crossover point of 21Gy (6-fractions) for duodenum and 18Gy (6-fractions) for stomach, and overestimates the probability of toxicity after the crossover point.

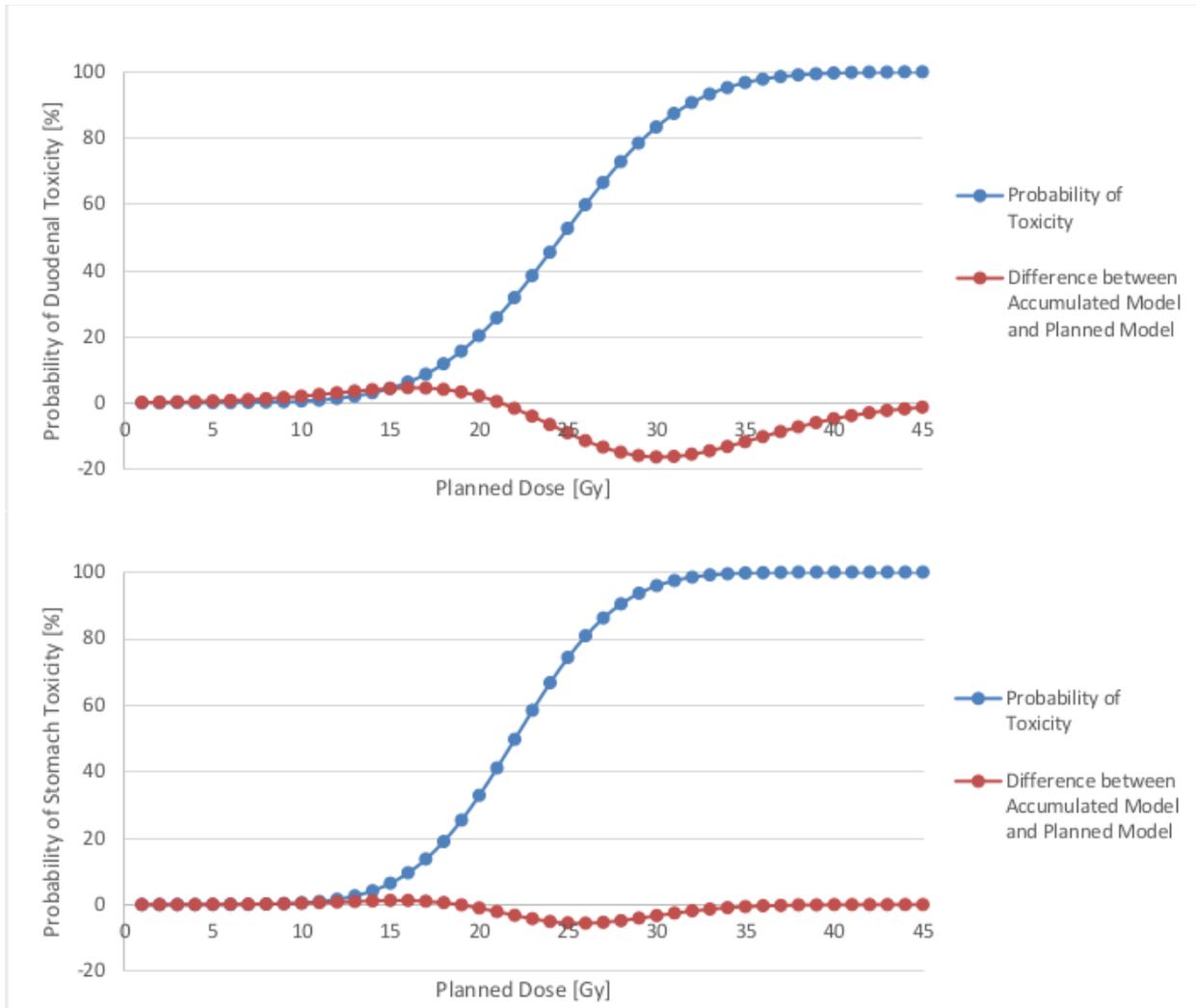


Figure 3-4. Probability of Duodenal Toxicity [%] (Based on Accumulated Model) for Duodenum and Stomach.

The blue line represents the probability of toxicity derived from the accumulated dose NTCP model versus planned dose. The orange line represents the difference between probability of toxicity using the accumulated model and the planned model. The largest difference for duodenum occurs at planned dose of approximately 30Gy (toxicity risk of approximately 80%). The largest difference for stomach occurs at planned dose of approximately 26Gy, (toxicity risk of approximately 80%).

3.5 Discussion

The potential impact of the deviation between the planned and delivered dose on the development of NTCP models was investigated for liver SBRT patients. Deviations were observed between the NTCP curves based on accumulated dose and standard NTCP curves for duodenal toxicity, indicating the potential impact that updated NTCP curves could have on patient treatment. For gastric toxicity, only modest deviations were observed, with a maximum deviation of 14%. However, more substantial and potentially clinically significant deviations were seen for the duodenum, with an average deviation of 6% and a maximum deviation of 21%. For both duodenal toxicity and gastric toxicity, standard NTCP curves overestimated toxicity risk at lower doses (for doses up to 21.5Gy based on 6-fraction dose for duodenum and up to 19Gy based on 6-fraction dose for stomach). Deriving probability of toxicity risk for planned dose using simulations of NTCP curves based on accumulated dose yields a smaller error than deriving probability of toxicity risk for planned dose using the standard, planned dose NTCP model.

This study was based on 30 original patients that were resampled to simulate a cohort of 150 and 600 patients to build statistical Lyman NTCP models. Due to the limited number of patients in the original analysis [25], true models could vary from the simulations. In addition, these 30 patients were treated at one institution, indicating the possibility of dependence of the results to the treatment setup and treatment planning tendencies. As with all models, there is some uncertainty and this NTCP model has been shown to potentially overestimate the toxicity risk

[116]. However, the focus of this paper is to investigate the potential difference between NTCP models given the known differences between accumulated and planned dose, to determine if there is a strong need to gather multi-institutional data to ensure the accuracy of these models. The results of this simulation study requires confirmation through prospective studies. Further affirmations are necessary for the impact assessment on normal tissue complications of the stomach and duodenum. In an image guided radiation therapy (IGRT) protocol, the goals are normally targeting tumor alignment and avoiding normal tissue. A complicated relationship exists between the planned and accumulated dose due to dose distribution and high dose regions, and a study with many more than 30 patients is necessary to understand this relationship.

The 30 patients in the current study were originally selected for a secondary analysis of SBRT dose accumulation [25], from the patients treated on phase I/II clinical trials [106,117,118] based on having available CBCT imaging and breathing motion $> 5\text{mm}$. They were not specifically selected based on any clinical outcome including toxicity. There were no grade 3 or higher GI toxicities in this limited sample of 30 patients. There were 2 patients with Grade 3 platelet counts, which are likely not related to duodenum or stomach toxicity, however are related to liver function and possibly non-classic radiation induced liver disease (RILD).

Of the 30 patients included in this study, there were 9 patients with a dose-limiting normal tissue (i.e. PTV coverage or prescription dose was limited due to normal tissue toxicity risks) [64]. The toxicity models developed for this study showed that the planned-dose model overestimates toxicity risk in the average dose range, implying that the dose-limiting organs could potentially receive a higher dose with the same risk of toxicity. The 30 patients in this study did not have GI

toxicity, however this may have been the result of a conservative tumor dose. Understanding the dose-toxicity relationship is important especially in escalation studies because under-dosing due to misinterpreted normal tissue complication risk can be detrimental to local control, as the tumor will not receive as high a dose as possible. The QUANTEC report [116] on stomach toxicity acknowledged that there is limited data for GI toxicity, indicating a need for a better understanding of GI toxicity models, especially with the introduction of molecular agents in radiation therapy. The understanding of the dose being delivered is crucial so that the models used for radiation therapy are as accurate as possible and enable the highest therapeutic ratio possible.

This study demonstrates the potential clinical importance of including accumulated dose in the development of NTCP models in future clinical trials. Since existing NTCP models are based on patient data that preceded volumetric daily imaging, they are based on only planned dose. Currently, equipment and software for volumetric daily imaging are available in the clinic, and therefore new DIR tools can be applied to perform dose accumulation. Future clinical trials should prospectively record delivered (accumulated) doses and toxicity outcomes in an effort to improve the understanding of the dose-response relationship. For example, in a phase III study [119] with 368 patients, more accurate toxicity data could be developed if accumulated dose was collected along with planned dose. This would require the collection of daily images or all patients to correlate true delivered dose with toxicity. This study supports the need for these trials to include the collection of imaging obtained at treatment delivery so that accumulated dose can be calculated. This simulation study suggests that deriving new NTCP models based on

accumulated dose will yield clinically significant results compared to the current models based on planned dose.

In terms of trial design, one possibility is an observational design in which patients are treated per standard of care and both planned and accumulated dose values are recorded along with eventual toxicity outcome. How many patients are required will depend on many factors including the overall toxicity rate which is not always low. In primary liver cancer (HCC) patients treated with radiation, approximately 20% will experience toxicity defined as a change in Child-Pugh score of 2 points or more. Demonstrating improved predictive performance of accumulated dose models would be one aim of a trial. However, the primary aim would more likely be characterizing differences between planned and accumulated dose values including estimating how many patients had large deviations and seeking to identify in advance which patients were likely to have large dose deviations. The size of the trial could be selected to achieve a desired level of precision in estimating the proportion of patients with large deviations. As an example, if the true proportion of patients with large deviations were 0.10, a trial of 100 patients would result in a standard error of the estimated proportion equal to 0.03.

Possible reasons for the numerical differences between the NTCP curves of the duodenum and stomach include variation of patients breathing magnitude and breathing trajectory. Regions of high dose gradient can differ based on the area of the organ and how it is affected by breathing motion. NTCP parameters, radiation treatment planning, and sensitivity of motion can all influence the final NTCP results for each organ, potentially causing the duodenum and stomach to have different numerical results. In addition, due to the limited number of patients, it may be

that potential deviations for the stomach were not observed, which may be observed in a larger population.

This study focused on liver cancer and the organs at risk for toxicity during radiation treatment. However, the results may reflect the impact of incorporating accumulated dose in the development of luminal GI toxicity models in other abdominal treatment sites such as stomach, pancreas, or pelvic sites, where toxicity to GI structures are of clinical concern. Because accumulated dose is a more accurate prediction of delivered dose than planned dose, NTCP models for all sites, e.g. head and neck, could be improved if based on accumulated dose instead of planned dose.

3.6 Conclusion

This work shows the importance of building new toxicity models based on accumulated dose. Accumulated dose provides an estimate of the radiation dose that the patient is actually delivered, rather than the dose that was planned before the start of their radiation treatment. Current toxicity models overestimate the toxicity risk to the duodenum and stomach at doses up to 21.5Gy and 19Gy, respectively, and underestimate the toxicity risk for higher doses. These findings indicate that the tumor could be delivered a higher dose, with the same probability of toxicity to the duodenum and stomach, which are both dose-limiting structures. Delivering a higher dose to the tumor potentially allows for higher tumor control, which could ultimately lead to a higher success rate for radiation treatment of liver cancer and lower rates of recurrence. This

work indicates the need for dose accumulation in clinical trials for liver cancer, as more accurate dose accumulation models could be developed based on this data.

Chapter 4. Predictive Models to Determine Clinically Relevant Deviations in Delivered Dose for Head and Neck Cancer

Currently, the need to replan a HN cancer patient undergoing radiation therapy is based on the decision of the physician, based on factors such as weight loss. Rapid response from the tumor, weight loss, changes in the patient setup, and shrinkage of normal tissue can all occur during HN radiation therapy, causing drastically different anatomy than in the planning CT. In some cases, replanning is necessary, however, there does not yet exist a model to aide in the decision to replan a patient in radiation oncology. However, with the possibility of accumulating the dose to the patient, deviations between planned and accumulated dose can be calculated at every fraction during treatment, allowing for the possibility to develop such a model.

This work aims to develop a model that can aid physicians in the decision to replan HN patients, with an explicit dose deviation threshold between planned and accumulated dose.

4.1 Abstract

To improve the understanding of deviations between planned and accumulated dose and establish metrics to predict for clinically significant dosimetric deviations mid-way through

treatment to evaluate the potential need to replan during fractionated radiation therapy (RT). 100 head and neck cancer patients were retrospectively evaluated. Contours were mapped from the planning CT to each fraction CBCT via deformable image registration (DIR). The dose was calculated on each CBCT and evaluated based on the mapped contours. The mean dose at each fraction was averaged to approximate the accumulated dose for structures with mean dose constraints and the daily max dose was summed to approximate the accumulated dose for structures with max dose constraints. A threshold deviation value was calculated to predict for patients needing mid-treatment replanning. This predictive model was applied to 52 patients treated at a separate institution. Dose was accumulated on 10 organs over 100 patients. To generate a threshold deviation that predicted the need to replan with 100% sensitivity, the submandibular glands required replanning if the delivered dose was at least 3.5Gy higher than planned by fraction 15. This model predicts the need to replan the submandibular glands with 98.7% specificity. On the independent evaluation cohort, this model predicts the need to replan the submandibular glands with 100% sensitivity and 98.0% specificity. The oral cavity, intermediate CTV, left parotid, and inferior constrictor patients each had one patient exceeding the threshold deviation by the end of RT. By fraction 15 of 30-35 total fractions, the left parotid gland, inferior constrictor, and intermediate CTV had a dose deviation of 3.1Gy, 5.9Gy, and 4.8Gy, respectively. When a DIR failure was observed, the dose deviation exceeded the threshold for at least one organ, demonstrating that an automated DIR-based dose assessment process could be developed with user evaluation for cases resulting in dose deviations. A mid-treatment threshold deviation was determined to predict the need to replan for the submandibular glands by fraction 15 of 30-35 total fractions of RT.

4.2 Introduction

Approximately 50,000 patients are diagnosed with head and neck (HN) cancer in the United States every year, with the incidence of human papillomavirus (HPV) – related cases rising [120]. Standard management includes chemotherapy plus radiation therapy. The combined chemoradiation is often so effective, especially for HPV+ patients, that their tumor often responds in volume over the course of treatment [121]. In addition, due to toxicities seen during treatment (i.e. mucositis), patients often lose weight during the course of chemoradiotherapy. These anatomical changes lead to deviations between the prescription (planned) dose and delivered dose, which are well documented for radiation therapy; however the majority of these studies used research software not immediately available for widespread clinical use [122-124]. Deviations call for plan modification during the fractionated treatment course for many head and neck cancer radiation treatment plans in order to ensure proper coverage of the tumor, to avoid overdosing of healthy tissue, and ultimately to ensure that the best treatment is delivered to the patient.

Although some deviation between the planned and delivered dose is expected and not likely clinically significant, deviations that are clinically significant need to be identified early enough in the treatment delivery process to ensure that a plan adaptation can be safely performed to correct for the deviation. The need for adaptive replanning is often left to the discretion of the physician, and a quantitative model could aid in the decision to replan. Currently, there is no existing model to predict the need to replan and uncertainty exists as to what levels of deviation from the planned dose would result in a significant deviation from the initial expectations of tumor control and normal tissue toxicity risk. The goal of this work was to perform a quality

control study, using tools commercially available to determine the best estimate of the delivered dose across a large number of patients and then use this data to design a model to identify, by fraction 15 (out of 30-35 total fractions), which patients would complete treatment with a dose deviation exceeding the identified tolerance.

4.3 Methods and Materials

4.3.1 Patient Data

Data from 100 HN cancer patients treated with 30-35 fractions between November 2013 and October 2016 were retrospectively evaluated for this quality control IRB exempt evaluation. All patients were planned with volumetric modulated arc therapy (VMAT) using a commercial treatment planning system (Eclipse v13.6, Varian Medical Systems, Palo Alto, CA) and treated on a linear accelerator with daily kilovoltage cone beam CT (CBCT) imaging for pre-treatment positioning. Patient and treatment planning information is detailed in Table 1. CT-based treatment planning was performed based on the planning directive structures and goals listed in Table 2. The study focused on soft tissue organs at risk for toxicity that were within the field of view (FOV) of the CBCT.

Table 4-1. Patient characteristics for all patients included in the study.

Human papillomavirus is denoted as HPV. Clinical target volume is denoted as CTV. Planning target volume is denoted as PTV. Gross tumor volume is denoted as GTV.

Table 1a. Patient Characteristics – Original patient cohort	
Patients (n)	100
Gender (male/female) (n)	80/20
Diagnosis site (n)	
Pharynx	58
Larynx	21

Lip and Oral Cavity	16
Other	5
HPV status [positive/negative/unrecorded] (n)	59/13/28
CTV [mean (range)] (cm ³)	154 (1-725)
CTV to PTV margin (mm)	3
Prescribed dose [mean (range)] (Gy)	68 (60-72)
Fractions [35/30/other] (n)	66/17/17
Patients with boost (n)	10
Patients with replans (n)	16
Elapsed Tx Time [median (range)] (days)	49 (39-74)
IV contrast used in CT (Y/N)	78/22
≥10% Weight Loss over entire Tx (n)	25
≥5% Weight Loss over entire Tx (n)	67

Table 1b. Patient Characteristics – Independent Evaluation set	
Patients (n)	52
Gender (male/female) (n)	36/16
Diagnosis site (n)	
Pharynx	27
Sinonasal	12
Oral Cavity	5
Larynx	1
Other	7
HPV status [positive/negative/unrecorded] (n)	10/10/32
GTV [mean (range)] (cm ³)	45(4-171)
Prescribed dose [mean (range)] (Gy)	67(60-70)
Fractions [35/33/32/30] (n)	14/27/1/10
Patients with boost (n)	13
Patients with replans (n)	13

Table 4-2. Planning directives for normal tissue structures used in the initial planning.

The mean and max dose refers to the dose metric used for that organ.

Table 2. Normal Tissue Structures and Goals	
Structure	Target Value (Gy)
Left/Right Submandibular Gland	Mean Dose < 30
Larynx	Mean Dose < 20
Inferior Constrictor	Mean Dose < 20
Superior Constrictor	Mean Dose < 50
Oral Cavity	Mean Dose < 30
Left/Right Parotid Gland	Mean Dose < 24

Spinal Canal	Max Dose < 45
--------------	---------------

4.3.2 Image Guided Radiation Therapy

Daily image guided radiation therapy (IGRT) was delivered using automatic registration of the CBCT to the planning CT using a multiple iteration technique, with the initial iterations focusing on automated intensity based alignment of the entire planning CT and the CBCT and the final iterations focusing on the custom region of interest structure that was made from the union of the spinal canal and high-risk planning target volume (PTV). Manual adjustments were permitted following the automatic registration by the treating radiation therapist. The isocenter was defined on the planning CT to be at chin level to ensure the FOV of the CBCT scan covered the anatomy of interest.

4.3.3 Calculation and Evaluation of Delivered Dose

The daily CBCTs were directly accessible in the commercial treatment management software for DIR using a Demon's-based contour propagation [101] and dose calculations (SmartAdapt, Varian Medical Systems, Palo Alto, CA). Figure 1 shows contours on the planning CT and the same contours propagated on the CBCT of a different fraction. Figure 2 shows the planned dose distribution and the recalculated dose on a daily CBCT. The calibrations between the CBCT and planning CT were matched using monthly measurements (QA) of standard materials with known electron density (Catphan®). The delivered dose at each fraction was calculated on the daily CBCT, and the dose was evaluated based on the propagated contours.

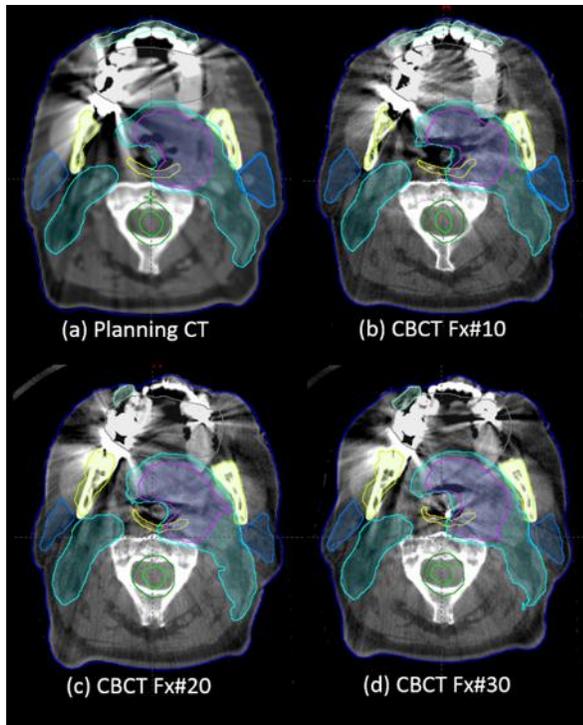


Figure 4-1. Contour examples.

Example of contours at planning, fraction 10, fraction 20, and fraction 30 of one patient, showing the deformed structures.

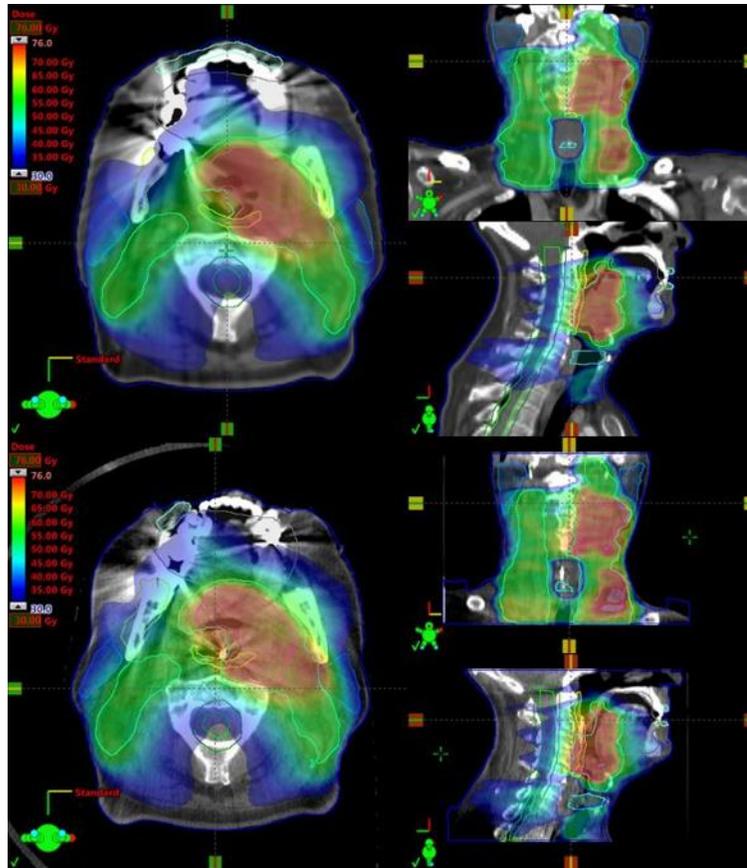


Figure 4-2. Dose Distribution Examples.

Example of the planning dose on the planning CT (top) and the dose distribution of the same plan calculated on the fraction 30 CBCT (bottom).

The following dose metrics were calculated:

Planned dose (D_{Plan}) is the clinical prescription planned dose, based on the planning CT and original contours, calculated in the treatment planning software and exported for analysis and comparison with delivered dose.

Delivered (Fraction) dose (D_{Frac}) for each fraction is the calculation of the clinical metric (e.g. mean or max dose) based on the dose calculated on the daily CBCT evaluated for the contours propagated from the planning CT onto the CBCT with DIR.

Accumulated dose (D_{Acc}) is the accumulation of clinical metric (D_{Frac}) over all treatment fractions. As dose accumulation is not currently available in the commercial treatment planning system, the following approximation was made to estimate the accumulated dose. For mean dose metrics, the accumulation is performed by averaging the daily calculation of the mean dose. For max dose metrics, the accumulation is performed by summing the daily calculation of the maximum dose, therefore representing the most conservative estimate (e.g. the highest max dose possible to a voxel).

4.3.4 Model Development

D_{Plan} , D_{Frac} for each fraction, and D_{Acc} were all calculated for each patient's treatment course.

D_{Acc} were calculated for each patient for fractions 1-5, 1-10, and 1-15, representing the accumulated dose to date up to fraction 5, 10, and 15, respectively, but scaled to the total number of fractions. That is, the D_{acc} at 1-5, 1-10, and 1-15 indicate the total dose that would be delivered if the entire treatment course was delivered as indicated by the first 5, 10, or 15 fractions. This scaling allows for direct comparison of dose deviations at each time point. Recently published data [125] suggests that a deviation between planned and delivered dose of less than 3.6Gy (15% of the 24Gy planning constraint) for parotid glands would not have a significant impact on the observed toxicity in a patient population. This 15% deviation criteria was then applied as a threshold for other HN organs with a mean dose constraint. This adaptive triggering model was

developed to identify dose deviations of 15% of the planning constraint of each organ at the completion of treatment early enough in treatment to enable corrective action. For CTVs, the dose deviation threshold was conservatively set at 7%, as the ICRU guideline [126] criteria is 95% - 107% relative to the prescribed dose. Table 3 depicts the dose deviation threshold between planned and delivered dose for each organ.

Table 4-3. Planning and dose deviation thresholds for the organs at risk and the target volumes.

Prescription dose is denoted as Rx. The dose deviation threshold is based on 15% of the planning constraint for all tissue with the exception of the target volume and the spinal canal. The threshold for spine is based on the maximum clinical constraint being 48Gy [127].

Organ	Dose Metric	Planning Constraint [Gy]	Dose Deviation Threshold [Gy]
Inferior Constrictor	Mean	20	3
Superior Constrictor	Mean	50	7.5
Spinal Canal	Max	45	3
High risk CTV	D95%	$\pm 7\%$ of Rx	4.10-5.18
Intermediate risk CTV	D95%	$\pm 7\%$ of Rx	2.83-5.17
Larynx	Mean	20	3
Oral Cavity	Mean	30	4.5
Left/Right Parotid glands	Mean	24	3.6
Left/Right Submandibular	Mean	30	4.5

Cases with a completed dose deviation to any structure that exceeded the structure-specific threshold were used to build a model to predict, with 100% sensitivity, whether a patient needs replanning, specifically looking at three time points: at the completion of fraction 5, fraction 10, and fraction 15. Each completed dose was compared to the planned dose to assess if the dose deviation threshold was exceeded for each organ for each patient. For every organ with a D_{Acc} that exceeded the dose deviation threshold, the minimum deviations between the D_{Plan} and D_{Frac} for fraction 1, the average of fractions 1-5, 1-10, and 1-15 were assessed to find the minimum

deviation in delivered dose at that time point to indicate that replanning was necessary to avoid completing treatment with a deviation that exceeded the specified tolerance. In some cases, due to proximity of the disease, normal tissue could not be spared at planning, dose constraints were substantially exceeded, and replanning when dose is increased would not improve toxicity risk. Based on 5-year normal tissue complication probability (NTCP) curves for the parotid glands [128], toxicity risk was evaluated at the parotid gland planning constraint (a mean dose of 24Gy), at 34Gy, and 54Gy. A mean dose of 24Gy corresponded to 20% toxicity risk, 34Gy to 30%, and 54Gy to 64%. As a 10Gy increase in mean dose corresponded to a 10% increase in toxicity, the model included organs with planned mean dose within 10Gy above the dose constraint, as we assumed that these patients could still benefit from replanning if the deviation in delivered dose exceeded the planned dose by more than the tolerance.

4.3.5 Data Inclusion

FOV and HU in CBCT was previously investigated demonstrating a 2% accuracy could be achieved with the exclusion of the superior 7mm and tissue in the shoulder region. The referenced study [91] compared dose calculations for the CBCTs and planning CTs of five oropharyngeal cancer patients undergoing radiation therapy. Rigid registration was performed between each CBCT and corresponding planning CT. CBCTs from the first three fractions of radiation treatment for each of the five patients included in the study were used for dose comparisons to measure the accuracy of dose calculation on CBCT. Three CBCT regions were assessed for dose calculation accuracy. Errors up to 80% were identified on the superior and inferior boundary of the CBCT. Errors were less than 2% after excluding the superior $7.1\text{mm} \pm 1.1\text{mm}$ extent of the CBCT and excluding the shoulder region in the inferior boundary. In the

current study, data was included only for organs on CBCT slices inferior to the superior-most 7mm and for organs that were not missing material in the trapezius/shoulder region due to field-of-view limitations. For the spinal canal, where Dmax was the metric of interest, the dose was evaluated only within the evaluable region of the CBCT.

4.3.6 DIR Evaluation and Validation

All DIR results were qualitatively evaluated using image fusion, evaluation of the deformation vector field (DVF), and visual assessment of the propagated structures onto the CBCT anatomy. For cases with DIR results identified as failing due to these qualitative metrics, the images were evaluated for artifacts. Images were re-registered to determine if a successful DIR result could be obtained. If a successful DIR could not be obtained, the treatment fraction was excluded from analysis and the dose accumulation was scaled based on the number of fractions with accurate DIR.

The assessment of DIR accuracy was based on contour propagation and comparison to manually drawn contours, therefore a baseline variation in the manual segmentation of normal tissues was established. Five HN patients were randomly selected for a reproducibility study. Segmentation of targets and critical organs was performed on four of the CBCTs by a radiation oncologist experienced in treating HN cancer and then a repeat segmentation was performed following a greater than two week time period, blinded to the original contours. The Dice similarity coefficient (DSC) was computed for the repeat contours for structures listed in Table 4.

The DSC calculates the overlap between two ROIs as defined by Eq. 4-1, and was used to assess the accuracy of the DIR based on the overlap between deformed and original contours.

$$DSC = 2 \frac{|A \cap B|}{|A| + |B|} \quad (4-1)$$

Where A is the primary ROI, and B is the secondary ROI. Based on this equation, a DSC of 1 indicates complete overlap, and a DSC of 0 indicates no overlap. The DSC was computed between the CT contours and CBCT contours propagated based on rigid registration alone and DIR.

Table 4-4. Results of DIR validation study and CBCT contour reproducibility (mean and standard deviation).

The p-value is based on a two-tailed paired Student’s T-test. Manual contour reproducibility was based on results of a resident radiation oncologist manually delineating structures on a CBCT, then repeating at a later date without reference to the original delineation. The submandibular glands are abbreviated to Sub. Glands in this table.

Structure	Registration Evaluation			Manual Contour Reproducibility
	Rigid Registration DSC	DIR DSC	T-Test P-Value	DSC
Spinal Canal	0.81±0.04	0.85±0.03	< 0.01	0.93±0.02
Sub. Glands	0.70±0.08	0.76±0.11	< 0.01	0.87±0.02
Larynx	0.65±0.10	0.74±0.10	< 0.01	0.90±0.03
Parotid Glands	0.64±0.09	0.71±0.08	< 0.01	0.78±0.05
Sup. Constrictor	0.53±0.11	0.61±0.08	< 0.01	0.78±0.03
Esophagus	0.41±0.10	0.58±0.09	< 0.01	0.63±0.18
Inf. Constrictor	0.40±0.19	0.60±0.10	< 0.01	0.64±0.16

4.3.7 Independent Model Validation

An independent validation was performed using a cohort of 52 patients treated at the MD Anderson Cancer Center whose dose was accumulated in a collaboration between the MD Anderson Cancer Center and the Netherlands Cancer Institute. Patient and treatment planning information is detailed in Table 1b. The original treatment plan was recalculated on the daily CT scans for each day of treatment, in order to calculate the delivered dose. The treatment isocenter was adjusted accordingly prior to dose recalculation. Using a deformation vector field (DVF) from DIR (Admire 1.04, Elekta AB, Stockholm, Sweden), the new dose distribution was mapped to the planning CT. The daily dose to each organ was calculated using the original contours on the planning CT. In the case of missing daily CTs, linear interpolation was performed on the CTs of the two closest fractions in order to create the missing fractionated doses. Nearest neighbor extrapolation was performed in the case that the daily CT from the first or last fraction was missing. The predictive model was applied to the delivered dose to 104 submandibular glands at fraction 15. The validation of the predictive model using this external data set was evaluated based on the dose deviation threshold set at 15% of the planning constraint (4.5Gy).

4.4 Results

Of the 3291 DIRs performed on 100 patients, 1% failed. These failures were due to metal artifact, beam hardening artifact, shadowing, and lack of contrast on the CBCTs. The maximum number of DIR failures for a single patient was 3. In the cases with DIR failures a deviation from the planned dose of greater than 15% was observed for at least one organ. DIR failures led to the re-registration of images, however that did not result in DIR improvement. Data due to errors in dose calculations and DIR were excluded from further analysis.

4.4.1 Contour Reproducibility and DIR Validation

Table 4 shows the results of the DIR validation and CBCT contour reproducibility. DIR improved DSC over rigid registration for 100% of inferior constrictors, 94% of esophagi and parotid glands, 90% of submandibular glands, and 75% of superior constrictors, spinal canals, and larynges included in the study. Based on a two-tailed paired Student’s T-test, improvement in DSC was statistically significant from rigid registration to DIR in all organs. The average DSC over all CBCTs included in the study is included in the table for both the rigid registration and the DIR.

4.4.2 Dose Deviation Threshold

The results of setting a 15% dose deviation threshold at treatment completion are shown in Table 5. There was one patient exceeding the threshold for oral cavity, left parotid gland, intermediate CTV, and inferior constrictor, and the average deviation between D_{plan} and D_{frac} is reported for the completion of treatment and fraction 15. For the submandibular glands, where 7 patients exceeded the threshold, the maximum deviation at completion of treatment and minimum deviation by fraction 15 are shown.

Table 4-5. Number of organs with clinically relevant dose deviations and values of those deviations.

The total number of organs were those that originally had dose calculations done for the study. The organs included in the model were those within the evaluable region of the image and not on images for which DIR failures were observed.

Organ	Planning Constraint [Gy]	Dose Deviation Threshold [Gy]	Total Number of Organs	Organs Included in Model	Organs Exceeding Deviation	Maximum Deviation at Completion of Tx [Gy]	Minimum Deviation by Fx15 [Gy]
Spinal Canal	45	3	99	99	1	3.1	NA

Left/Right	30	4.5	176	85	7	8.22	3.5
Sup.	50	7.5	100	60	0		
Oral Cavity	30	4.5	100	56	1	5.18	0.81
High CTV	±7%*Rx	4.10-5.18	103	43	0		
Left Parotid	24	3.6	100	37	1	3.77	3.08
Right Parotid	24	3.6	99	34	0		
Int. CTV	±7%*Rx	2.83-5.17	101	17	1	-6.65	-4.84
Inf. Const	20	3	97	12	1	5.62	5.86

The left and right submandibular glands were combined into a single model to account for all patients with completed dose deviations exceeding the threshold. Eighty-five total submandibular glands were without DIR failures and in the evaluable region of CBCTs. With a dose deviation threshold set at 15% of the planning constraint (4.5Gy), seven submandibular glands exceeded the threshold. By fraction 15, the minimum deviation of these seven glands, using the D_{acc} at fraction 15 scaled to the total delivered dose, was 3.5Gy (maximum = 7.1Gy, median = 5.1Gy, mean = 5.3Gy). This minimum deviation threshold was set for each time point of treatment to evaluate predictive power for final dose deviation. The resulting false positives, sensitivities, specificities, and positive and negative predictive values (PPV and NPV) are shown in Table 6. Out of the 12 patients with evaluable inferior constrictor dose deviations, one had a potentially clinically significant dose deviation of 5.6Gy at completion of treatment. This deviation, and the need to replan, was predicted by fraction 15, with a dose deviation of 5.86Gy.

Table 4-6. Model based on submandibular glands (SG).

The minimum dose deviation for Fx 1-15, 1-10, 1-5, and 1 is based on the minimum dose deviation by fraction 15 that leads to a dose deviation exceeding the dose deviation threshold at completion of treatment. CI refers to confidence interval. Dev. refers to deviation.

Organ	Total SGs (N)	Dose Deviation Threshold [Gy]	Organs Exceeding Dose Deviation	Test	Fx 1-15 Dose Dev. [Gy]	95% CI	Fx 1-10 Dose Dev. [Gy]	Fx 1-5 Dose Dev. [Gy]	Fx 1 Dose Dev. [Gy]
-------	---------------	-------------------------------	---------------------------------	------	------------------------	--------	------------------------	-----------------------	---------------------

			Threshold						
L/R SGs	85	4.5	7	Min. [Gy]	3.5		3.5	3.5	3.5
				False Pos.	1		3	2	10
				Sens.	1.00	0.59, 1.00	100	57.1	28.6
				Spec.	.99	0.93, 1.00	96.2	97.4	87.2
				PPV	.88	0.47, 1.00	70	66.7	16.7
				NPV	1.00	0.95, 1.00	100	96.2	93.2

Of the 17 patients with evaluable intermediate CTV dose deviations, one had a potentially clinically significant dose deviation of -6.7Gy at the completion of treatment. This deviation, and the need to replan, was predicted by fraction 15, with a dose deviation of -4.8Gy.

Of the 56 patients with evaluable oral cavity dose deviations, one had a potentially clinically significant deviation of 5.2Gy. However, the early treatment fractions did not show clinically relevant dose deviations, and therefore did not predict the final dose deviation. By fraction 15, the dose deviation was only 0.81Gy.

Of the 99 patients with evaluable spinal canal dose deviations, one had a potentially clinically significant deviation of 4.3Gy, leading to a D_{acc} of 48.1Gy, which exceeds the maximum planning constraint. Therefore, trends could not be evaluated for this organ.

4.4.3 Independent Model Validation

On the independent validation data set, four submandibular glands exceeded the dose deviation threshold by the end of treatment. By fraction 15, a total of six glands exceeded the 3.5Gy minimum deviation based on the original dataset; four of these glands were those that exceeded

the 4.5Gy deviation threshold by the end of treatment. Independent validation predicted the dose deviation threshold by the end of treatment with 100% sensitivity (95% CI: 40%, 100%), 98% specificity (95% CI: 93%, 100%), 67% PPV (95% CI: 22%, 96%) and 100% NPV (95% CI: 96%, 100%).

4.5 Discussion

A deviation threshold was developed from a large cohort of patients to aid physicians in the decision to replan HN cancer patients. This replanning strategy has the potential to impact HN cancer patients by automatically identifying cases that could benefit from replanning by the midpoint of treatment. Treatment planning and CBCT image data from 100 HN patients were retrospectively analyzed to build a model to predict clinically relevant deviations between planned and delivered dose. For these 100 patients, 9 HN organs of interest were evaluated over 30-35 fractions of VMAT. If a mean dose deviation threshold of 3.5Gy is set by fraction 15 for the submandibular glands, the need for replanning can be predicted with 98.7% specificity providing 87.5% PPV, and 100% NPV. Applying this model to an independent validation set from a different institution using a different DIR algorithm predicted the need for replanning with 100% sensitivity, 98% specificity, 67% PPV and 100% NPV. The inferior constrictor, intermediate CTV, oral cavity, and left parotid glands only had one patient exceeding the corresponding dose deviation threshold by the end of treatment. With the exception of the oral cavity, all of these patients would be flagged by fraction 15 of treatment using this model. Trends could not be evaluated for spinal canal as only one patient had a substantial deviation. The spinal canal was used for initial alignment of the CBCTs to CTs, likely influencing consistent delivery of planned dose.

As demonstrated in Table 6, the confidence intervals for specificity are relatively narrow (resulting from the majority of subjects not requiring replanning), while the confidence intervals for sensitivity are relatively wide (since the number of positives, i.e., subjects requiring replanning, is small in both cohorts). Although the model achieves 100% sensitivity, it's difficult to make strong statistical claims due to the limited number of cases that require replanning, indicating the need to further validate this model in a prospective clinical trial. Assuming that the true proportion of observations requiring treatment replanning is 8%, a minimum sample size of 100 subjects will be required to achieve $\geq 80\%$ power to reject the null hypothesis that the sensitivity is $\leq 70\%$ under the alternative that the sensitivity is $\geq 98\%$, using an exact test for a single proportion with one-sided significance level 0.05. This sample size is also sufficient to detect a change in the value of specificity from 85% to 95% (assuming 92 plans do not require replanning), also using a one-sided test with level 0.05.

Although true dose accumulation using an accurate DIR method is ideal, it has been shown [129] that the population differences are small (the difference between the population average of mean PG doses and the population average of DIR mean PG doses was 0.3Gy) and did not impact the assessment of replanning strategies. This study used manual delineations on CT and used those contours for DIR and to propagate the dose. The current study used DIR between CT and CBCT to propagate the contours, which could potentially increase uncertainties in the estimated delivered dose. The impact of uncertainties in the DIR algorithm on the assessment of delivered dose requires further investigation. Limited data exists in the literature on how uncertainties in DSC translate to uncertainties in dose. A recent study reported on the impact of the accumulated

minimum dose as a function of variation in DIR [130]. This study demonstrated that, in order to impact the minimum dose, a combination of low contour DSC in a high dose gradient would have to occur. A maximum difference of 2.5 Gy in the minimum dose was observed. The current study reported on the mean dose to each organ and would therefore be less sensitive to low DSC and high dose gradients.

Substantial DIR error occurred in a limited number of cases, 1%, and was attributed to artifact, shadowing effects, beam hardening artifact, shadowing, and lack of contrast on the CBCTs. The potential to develop a completely automated dose accumulation process is limited by the possibility of DIR errors and the potential need to evaluate the results in real time. Dose deviations from all cases with substantial DIR failure were evaluated and all of them were found to lead to reported dose deviations of greater than 15% between planned and delivered dose for at least one organ. This data suggest that it may be possible to develop an automated process and perform DIR evaluation only on those cases that result in a dose deviation greater than 15% (10% of cases). This requires further testing on additional patients.

A previous study [91] determined that FOV differences between CT and CBCT lead to dose calculation error up to 80% on the superior and inferior edge inside the CBCT FOV. For dose calculations to not exceed 2% error, organs in the superior 7mm and the shoulder regions of the CBCTs were excluded in evaluations. Future work must address this limitation to ensure a robust method is available to calculate the daily dose to all structures of interest. Augmenting the missing data from the CT or deforming the CT to the CBCT are potential methods to improve

these dose calculations.

The dose deviation threshold was based on research that stated that dose changes of less than 3.6Gy (15% of the planning constraint) to parotid glands would not have a significant impact in a population. As similar data for other soft tissue structures with mean dose constraints does not exist, this 15% was applied as the dose deviation threshold for the other HN organs with mean dose planning constraints evaluated in this study. However, consensus on a clinically meaningful dose deviation has not been reached across the radiation oncology community.

4.6 Conclusion

In this study, a predictive model was developed to aide physicians in the decision to replan HN patients undergoing radiation therapy. This is the first model developed for this purpose, and was validated on an external validation patient cohort. While this model was developed with 100% sensitivity (e.g. to catch all patients exceeding a clinically significant dose deviation threshold) on the original patient cohort, it also predicted the need to replan with 100% sensitivity on the external validation patient cohort. Using this model in the clinic provides physicians with the first predictive model to aide in the decision to replan HN patients undergoing radiation therapy.

Chapter 5. Biomechanical Modeling of Neck Flexion for Deformable Alignment of the Salivary Glands in Head and Neck Cancer Images

Patients undergoing radiation therapy for HN cancer often receive imaging in different positions. The patient may be setup differently in the diagnostic image, planning image, treatment images, and follow-up images. In order to map contours from one image to another, proper alignment must be completed between the two images. Currently, rigid registration and intensity-based deformable image registration is used clinically, in order to register two images and map the contours from one image to the other. However, rigid registration alone may not account for the changes in the tissue, and intensity-based registration methods may not allow us to see the changes inside the individual organs. Additionally, intensity-based registration methods do not allow for the extrapolation of contours outside of the image FOV, as investigated in the previous chapter. This work investigates a biomechanical model-based deformable image registration method to align the salivary glands, based on the varying angles of neck flexion of patients undergoing HN radiation treatment.

5.1 Abstract

During head and neck (HN) cancer radiation therapy, analysis of the dose-response relationship for the parotid glands (PG) relies on the ability to accurately align soft tissue organs between longitudinal images. In order to isolate the response of the salivary glands to delivered dose, from deformation due to patient position, it is important to resolve the patient postural changes, mainly due to neck flexion. In this study we evaluate the use of a biomechanical model-based deformable image registration (DIR) algorithm to estimate the displacements and deformations of the salivary glands due to postural changes. A total of 82 pairs of CT images of HN cancer patients with varying angles of neck flexion were retrospectively obtained. The pairs of CTs of each patient were aligned using bone-based rigid registration. The images were then deformed using biomechanical model-based DIR method that focused on the mandible, C1 vertebrae, C3 vertebrae, and external contour. For comparison, an intensity-based DIR was also performed. The accuracy of the biomechanical model-based DIR was assessed using Dice similarity coefficient (DSC) for all images and for the subset of images where the PGs had a volume change within 20%. The accuracy was compared to the intensity-based DIR. The PG mean (STD) DSC were 0.63 (0.18), 0.80 (0.08), and 0.82 (0.15) for the rigid registration, biomechanical model-based DIR, and intensity based DIR, respectively, for patients with a PG volume change up to 20%. For the entire cohort of patients, where the PG volume change was up 57%, the PG mean (STD) DSC were 0.60 (0.18), 0.78 (0.09), and 0.81 (0.14) for the rigid registration, biomechanical model-based DIR, and intensity based DIR, respectively. The difference in DSC of the intensity and biomechanical model-based DIR methods were not

statistically significant when the volume change was less than 20%. When the volume change was large, there was a significant difference, although the magnitude was small. These results demonstrate that the proposed biomechanical model with boundary conditions on the bony anatomy can serve to describe the varying angles of neck flexion appearing in images during radiation treatment and to align the salivary glands for proper analysis of dose-response relationships. It also motivates the need for dose response modeling following neck flexion for cases where parotid gland response is noted.

5.2 Introduction

During external beam radiation therapy (EBRT) of head and neck (HN) cancer, dose to the PG can lead to toxicity [32]. Studies have shown that limiting the mean dose to the salivary glands during HN radiation therapy leads to lower toxicity to the patient [131]. However, preclinical studies show that considering specific subregions of the glands could improve dose response modeling [74]. Understanding the effect of the dose to the subregions of the glands over the course of radiotherapy is challenging due to the volumetric response combined with the sharp dose gradient within the glands. Determining the dose to subregions over the course of treatment requires spatial alignment of longitudinal images. Deformable registration of the gland is challenging due to the volumetric response, deformation, and the uniform contrast of the glands on CT images which remain the standard imaging modality in the radiation therapy workflow.

The majority of techniques proposed for DIR of longitudinal images of HN cancer patients use intensity-driven algorithms to deform one image so that the structure boundaries match those in

the other image [132-135]. Using these conventional DIR approaches, the estimated deformations inside PGs will only depend on the displacement vector field (DVF) regularization model of the DIR algorithm. However, the shrinkage scheme is likely more complex, correlating to the heterogeneity of the dose inside the organ and potentially the sensitivity of the local structure within the gland. Biomechanical model-based DIR algorithms allows exploration of such complex deformations and volume change as the PGs respond to radiation therapy. Early work by Al-Mayah demonstrates that the use of dose-based boundary conditions in a biomechanical models can simulate the radiation dose response during HN radiation therapy [136].

Based on the above study, biomechanical model-based DIR provides the potential to investigate the PGs response to radiation. However, in order to understand the PG response to radiation dose other sources of displacements and deformations of the glands should be resolved first. In particular, flexion of the neck or movement of the mandible are often observed between images acquired at different time points because of the difficulty in reproducing the patient position in the presence of weight loss or when acquisition is performed by different imaging devices, limiting the consistency of immobilization devices. In order to accurately model the volumetric changes of the PG subregions due to dose-response during radiation therapy, an initial alignment of the PGs is necessary to first resolve the changes due to neck flexion. A previous feasibility study [137] investigated a biomechanical model of the HN with various boundary conditions. This study found that the highest accuracy, based on the DSC of the tumor and PGs, was found when placing boundary conditions on the vertebrae and mandible. However, the evaluation was tested on a small dataset of four pairs of images and requires further validation.

The goal of this study is to perform a comprehensive investigation and validation of the use of biomechanical model-based registration to resolve the misalignment of the salivary glands only due to those postural changes and compare to an intensity-based DIR method for a large cohort of patients demonstrating a range of levels of volumetric response of the parotid gland. The use of a commercially available biomechanical model of the patient anatomy with boundary conditions on relevant bones and external contours is proposed. For evaluation of the model, a cohort of 82 patients was retrospectively evaluated and for each patient a pair of CT scans presenting noticeable differences in the angle of neck flexion was selected. The overlap between mapped deformed gland volumes and original gland volumes were measured after solving the neck flexion using the biomechanical model-based DIR algorithm. In order to assess the accuracy of the biomechanical approach to resolve patient positioning, the performances of the biomechanical approach were compared to the performances of the rigid and intensity-based DIR methods for PGs with volume change within 20%, to evaluate the accuracy of alignment when little response is noted, as well as the whole cohort of patients, with volume change up to 57%, to evaluate the need of further dose-based boundary conditions on the glands when response is noted.

5.3 Methods

5.4 Patient Data

A retrospective IRB approved evaluation was performed on 164 PG from 82 oropharynx cancer patients who underwent EBRT. For each patient, two non-contrast enhanced CTs showing neck

flexion acquired on different dates were selected. The normal tissue structures were previously auto-segmented (Admire ABAS, Elekta, Stockholm, Sweden) on all images. For this study, all images and contours were imported into a TPS with biomechanical model-based deformable registration capabilities (RayStation v6.99, RaySearch Laboratories, Stockholm, Sweden).

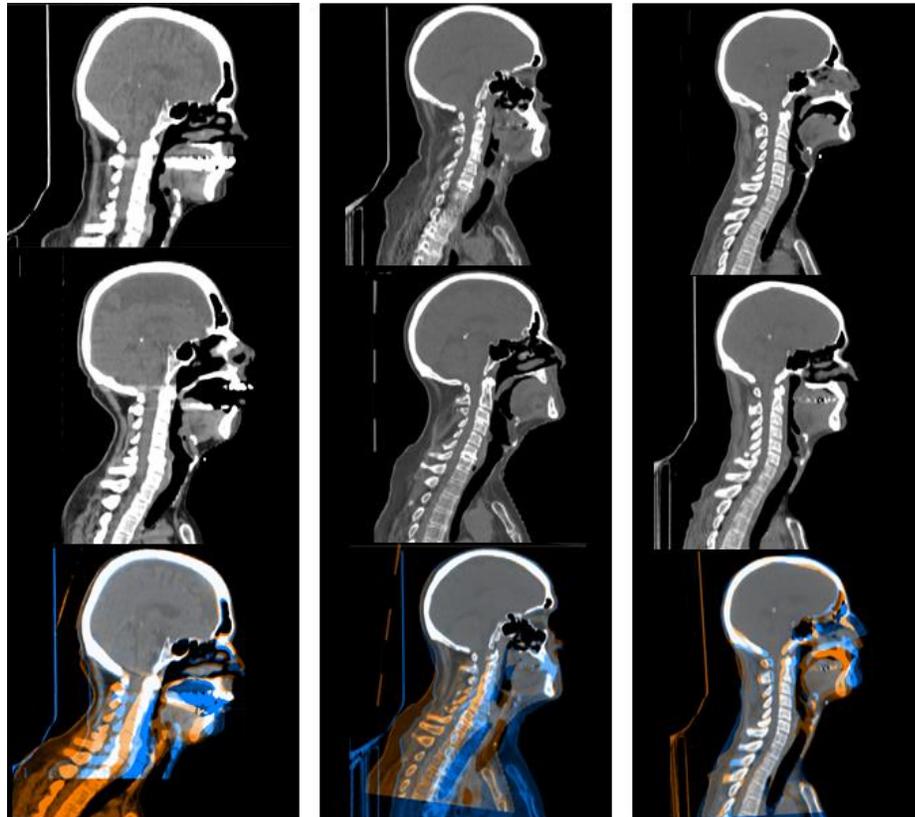


Figure 5-1. Examples of rigid registration performance for three cases.

The top row shows the primary images, the middle row shows the secondary images, and the bottom row shows the rigid registration results (with focus on skeletal anatomy) between the primary and secondary images.

5.4.1 Image registration

The following registration and evaluation steps were performed automatically for all 82 patients using a Python-based script in the TPS.

Rigid Registration. For each patient, gray level-based rigid registration was performed in the TPS between the two CTs displaying neck flexion using a focus on the skeletal anatomy. This registration step served as the initial registration for the DIR that followed. Examples of the rigid registrations are illustrated in Fig. 1. The bottom row of Figure 3 shows the overlay of the images after a global rigid registration based on all bony anatomy. On the left column both the vertebra and mandible were still misaligned. For the case in the middle column only the vertebra were still misaligned while for the third case only the mandible was still misaligned. It is expected that a biomechanical model driven by boundary conditions on the vertebrae and mandible would allow a better global alignment particularly for the PG.

Deformable Image Registration. Following the initial rigid registration, both intensity and biomechanical-based DIRs were performed between the CTs of each patient. The intensity-based DIR algorithm used in this study was the ANAtomically CONstrained Deformation Algorithm (ANACONDA) implemented in the commercial TPS (RayStation v6.99, RaySearch Laboratories, Stockholm, Sweden). For the ANACONDA method, the registration problem is represented as a nonlinear optimization problem, where the objective function is defined as $f : \mathbb{R}^n \rightarrow \mathbb{R}$, and R is there reference image. The equation for the nonlinear optimization problem is

$$f(v) = \alpha C(v) + (\beta H(v) + \gamma S(v)) + \delta D(v) \quad (5-1)$$

where non-negative weights are represented by α , γ , and δ , $C(v)$ represents the correlation coefficient, the term $\beta H(v)$ allows coordinate functions to minimize (approximately) the Dirichlet energy and causes the invertibility and smoothness of the DVF, $\gamma S(v)$ penalizes large ROI deviations (together, $\beta H(v) + \gamma S(v)$ control regularization of the deformation grid), and $D(v)$ includes ROIs that guide the deformation (in the case that the user has selected any). A smooth DVF is optimized using the quasi-Newton algorithm. Similarity between the images is determined by correlation coefficients, which guide the DVF. Regularization is attained with minimization of the weighted Dirichlet energy for coordinate functions of the DVF, and involves first resolving the DVF smoothness and invertibility, and then penalizing large deviations in the regions of interest (ROI). Controlling ROIs can be chosen by the user to guide the DIR, but were not selected in this study when using the algorithm. .

The biomechanical model-based DIR algorithm used for this study was the commercial implementation of Morfeus [63] in the TPS. Briefly, Morfeus creates tetrahedral meshes from the contours of the body and organs included in the model and assigns elastic properties to each of them. For each organ, a surface projection method between the organ surface on the reference and secondary images determines the displacement of the surface nodes of the tetrahedral meshes. Those displacements are used as boundary conditions in the model to solve the displacement of all the internal mesh nodes in a finite-element analysis. In the proposed model, boundary conditions were applied on the mandible, C1 vertebrae, and C3 vertebrae, as well as the patient external contour. These boundary conditions were chosen for their proximity to the PG, and their capacity to describe most possible postural changes.

5.4.2 Registration evaluation

Based on the rigid registration and two DIR methods (intensity and biomechanical), the left and right PG segmentations were propagated from the CT0 (the earlier dated CT) to CT1 (the later dated CT). The performance of each method to accurately propagate the PG structures were reported using the following metric:

Dice Similarity Coefficient. The DSC, which calculates the overlap between two ROIs as defined by Eq. 1, was used to assess the accuracy of the DIR based on the overlap between deformed and original contours.

$$DSC = 2 \frac{|A \cap B|}{|A| + |B|} \quad (5-2)$$

Where A is the primary ROI, and B is the secondary ROI. Based on this equation, a DSC of 1 indicates complete overlap, and a DSC of 0 indicates no overlap. The DSC was calculated in the TPS for the left and right PG, individually, based on the rigid registration, intensity-based DIR method, and the biomechanical model-based DIR method. For each case, the DSC was calculated between the deformed ROI mapped on the second CT and the original ROI on that CT. The mean DSC for PG variability (based on multiple observers) is 0.76 [138].

Volume Change. The volume of the left and right PG from each CT were calculated in the TPS in order to understand the analysis of the PG DSC. The DSC from the intensity-based DIR and

the biomechanical model-based DIR were evaluated in cohorts based on absolute percent volume change of the PG. A two-tailed paired Student's T-test was used to assess the statistical significance between the PG DSC of each DIR method for cases with volume change within 20% and for the entire cohort. The DSC as a function of volume change is important as the intensity based registration should account for volume change, whereas the biomechanical based registration does not have boundary conditions on the PGs and therefore will not account for volume change (e.g. when the 2 images represent both neck flexion and PG volume change, the biomechanical model, by definition, will only account for the neck flexion and the volume difference should be noted by a DSC less than 1).

DSC for cases with less than 20% volume change using biomechanical-based DIR will evaluate the accuracy and robustness of the algorithm to resolve the deformation and positional changes due to neck flexion. The DSC for cases with more than 20% volume change will evaluate the accuracy and robustness of the algorithm to partially resolve complete deformation, but highlight the potential need for further boundary constraints to resolve the volumetric response. The DSC for the intensity-based DIR will indicate the accuracy and robustness of the algorithm to resolve this complex deformation.

5.5 Results

Table 1 shows the mean, standard deviation (STD), maximum, and minimum DSC of the PG over all patients with volume change within 20% (N=63). The intensity-based DIR method

resulted in a slightly higher mean DSC average (0.02) and slightly higher standard deviation (0.07) at (0.82±0.15) than the biomechanical model-based DIR method (0.80±0.08), and both DIR methods resulted in a higher mean DSC and lower standard deviation than rigid registration (0.63±0.18). No significant differences were observed ($p = 0.13$) between the DIR methods and both methods had a DSC greater than the DSC reported for inter-observer variation (0.76). The minimum PG DSC were 0.17 and 0.53 for the intensity-based and biomechanical model-based DIR methods, respectively, indicating the potential for more ‘catastrophic’ registration errors with the intensity-based algorithm. The maximum PG DSC were 0.96 and 0.93 for intensity-based and biomechanical model-based, respectively.

Table 5-1. Registration method comparison according to PG DSC.

For each method, the mean, STD, max, and min DSC values of the PG are reported. The percentage of patients with a DSC superior to 0.75 are also reported.

DIR method	Mean	STD	Min	Median	Max	% of PG with DSC > 0.75
Rigid	0.63	0.18	0.07	0.66	0.94	20%
Intensity-	0.82	0.15	0.17	0.87	0.96	80%
Biomechan	0.80	0.08	0.53	0.82	0.93	76%

Figure 2 is a histogram depicting the number of PG with DSC greater than values ranging from 0 to 1. All PG (n=87) resulted in DSC greater than 0.1 for both methods, and 0 PG resulted in DSC greater than 0.95. Twenty-four (28%) PGs had a DSC that exceeded 0.9 for the intensity-based method, while only 8 (9%) PG exceeded 0.9 for the biomechanical model-based method. However, all of the biomechanical model-based registrations had a DSC of 0.5 or greater, whereas 84 of the intensity-based registration had a DSC of 0.5 or greater.

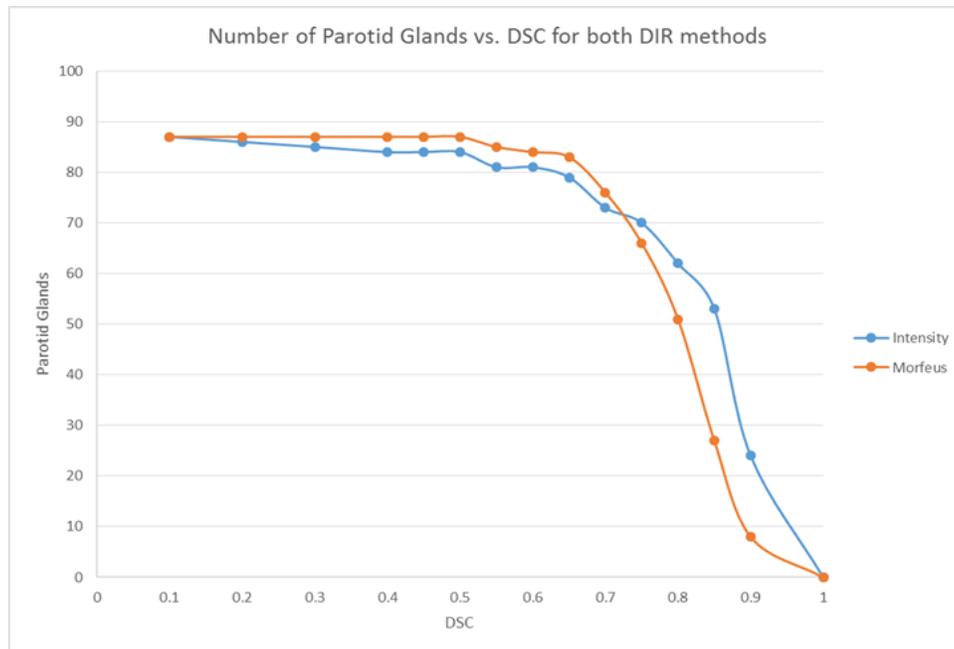


Figure 5-2. PG DSC Histogram.

Histogram depicting the number of PG with DSC greater than the corresponding value, ranging from 0 to 1. The results of the intensity-based DIR method is depicted in blue and the results of Morfeus are depicted in orange, for patients with <20% volume change.

Table 2 shows the mean, standard deviation (STD), maximum, and minimum DSC of the PG over all patients. The intensity-based DIR method resulted in a slightly higher mean DSC average (0.03) and slightly higher standard deviation (0.05) at (0.81 ± 0.14) than the biomechanical model-based DIR method (0.78 ± 0.09) , and both DIR methods resulted in a higher mean DSC and lower standard deviation than rigid registration (0.60 ± 0.18) . Similar to the subset evaluated above, both DIR algorithms had an average DSC greater than the inter-observer variability (0.76).

Table 5-2. Registration method comparison according to PG DSC.

For each method, the mean, STD, max, and min DSC values of the PG are reported. The percentage of patients with a DSC superior to 0.75 are also reported.

DIR method	Mean	STD	Max	Min	% of PG with DSC > 0.75
Rigid	0.60	0.18	0.94	0.07	21%
Intensity-based	0.81	0.14	0.96	0.17	77%
Biomechanical model-based	0.78	0.09	0.93	0.53	67%

Including cases where the volume change exceeded 20%, the differences in DSC between the intensity-based DIR and biomechanical model-based DIR becomes statistically significant. The significant difference between the PG DSC in cases with a 20% volume change demonstrates the need for additional boundary conditions to describe this volumetric response. The larger STD and lower minimum DSC of the intensity-based registration demonstrates the potential advantage of the biomechanical model-based approach when potentially combined with a dose-based boundary condition on the PGs. Figure 3 illustrates the range of DSC for each method as a function of volume change.

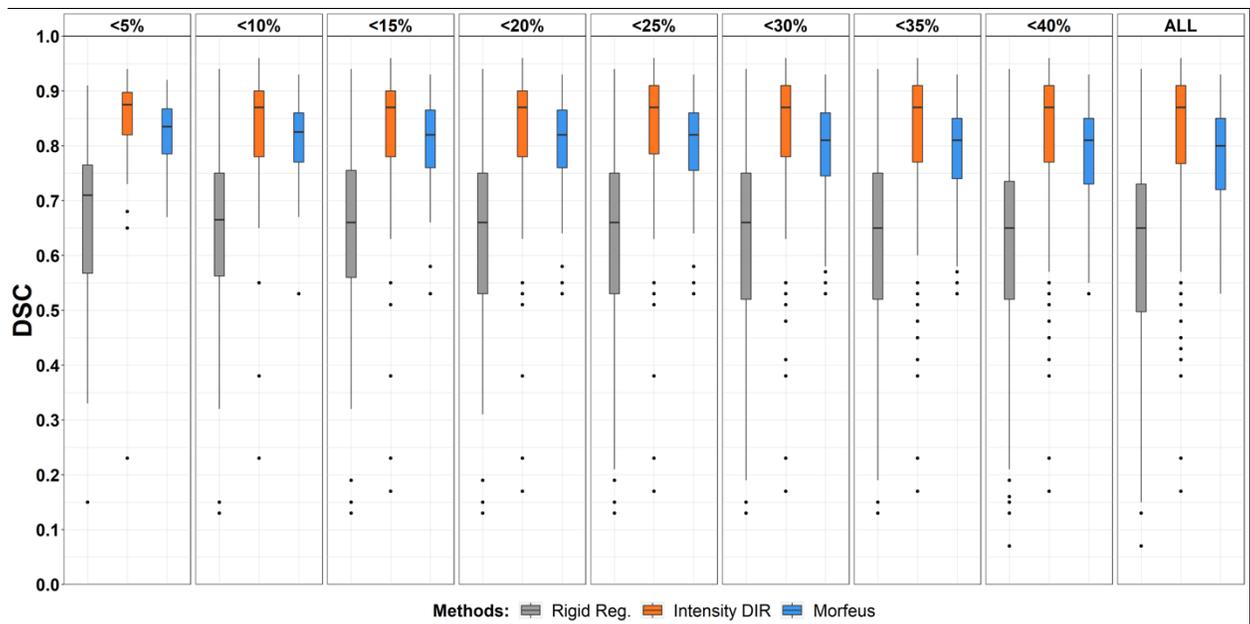


Figure 5-3. DSC Boxplots.

Boxplots of the DSC obtained for rigid registration (gray), intensity-based DIR (orange), and biomechanical model-based DIR (blue) for different volume changes of the DSC obtained for the propagated PG contours based on each DIR method. PG 5 represents volume change of 5%, and the same for other values. Outliers are shown in the bubbles below each box plot. Median DSCs are represented by the horizontal lines, the lower whiskers represent the minimum, the upper whiskers represent the maximum, the base of the box represents the lower quartile, and the ceiling of each box represents the upper quartile. Intensity-based and biomechanical model-based results are shown in pairs for each volume change threshold.

5.6 Discussion

In this study, we retrospectively evaluated the registration of 82 repeat CT scans for oropharyngeal cancer using the DSC between the deformed and original delineations of the PGs. For each set of CT scans, DIR was performed using intensity-based registration for full image DIR and a biomechanical model-based methods focused only on resolving the neck flexion, and therefore PGs with minimal volume change were evaluated in an isolated cohort.

There was no statistically significant difference observed between the DIR methods for cases with a PG volume change within 20%, and both DIR methods had a median DSC of greater than the inter-observer variation. Additionally, the clinical significance of a 0.02 decrease in mean DSC is arguably minimal. The standard deviation for the intensity-based method was larger than for the biomechanical model-based method and the minimum DSC was lower, indicating more potential failures in the results of the intensity-based method, which was further demonstrated by 3 PGs that had a DSC of less than 0.5 following intensity-based DIR.

This data demonstrates that modeling neck flexion alone using a biomechanical model-based registration algorithm aligns the PGs as accurately as an intensity-based registration and within the expected contour variation. For a volume change exceeding 20%, the results of each method were statistically significant, as expected given the lack of boundary conditions driving the alignment of the PGs in the biomechanical model. Future work will include applying dose-based boundary conditions to the biomechanical model-based algorithm [136] in order to build a comprehensive model that involves accurate neck flexion as well as a deformation model function of the dose distribution in the PG.

Evaluation of the DSC trends showed that there are 10 outliers in the total 164 PGs for the intensity-based method and 0 for the biomechanical model-based method. This indicates more failures in the intensity-based method. Table 2 shows the minimum DSC over all patients for the intensity-based method was 0.17, while the minimum DSC for biomechanical model-based method was 0.53. This data suggests that for challenging cases where intensity-based methods fail, even a simple biomechanical model-based method may give reasonable results.

5.7 Conclusion

In this study, a biomechanical model was developed to align the parotid glands based on changes only due to the varying angles of neck flexion of HN patients undergoing radiation therapy. For patients with minimal volume change in the parotid glands, and therefore changes only due to the position of the patient, the biomechanical model-based deformable image registration algorithm performed alignment of the parotid glands with indistinguishable results from the intensity-based method. Which the results of this study, it is known that biomechanical modeling with minimal boundary conditions can be used to resolve the neck flexion in images of patients undergoing radiation therapy for HN cancer. This model can be applied to patients with follow-up imaging done in different positions than during planning, allowing for the propagation of contours to map the tumor and normal tissue. Additionally, this model could be applied when the need arises to map contours outside of the field of view of the image, as in the work in the previous chapter.

Chapter 6. Biomechanical Modeling of Dose-Induced Volumetric Changes of the Parotid Glands for Deformable Image Registration

Current toxicity models for head and neck cancer are based on the mean dose to the entire gland, however recent studies [74] have shown that toxicity to the parotid gland may be more closely related to the dose to the substructures. Normal tissue structures deform through radiation therapy, due to the varying angles of neck flexion, weight loss, and dose-response. The last study resolved the deformation due to the varying angles of neck flexion, using a biomechanical model-based deformable image registration algorithm with optimal bony anatomy boundary conditions (C1 vertebra, C3 vertebra, and mandible). With the application of this initial registration, the addition of dose-based boundary conditions should resolve the remaining deformation. There first exists the need to determine the volumetric response of the parotid glands to radiation therapy, and study the response of the substructures of the glands. In this study, biomechanical models of the parotid gland response to radiation therapy are developed, by employing thermal expansion coefficients modified to describe the dose delivered and the resulting volumetric response.

6.1 Abstract

A major side effect of radiation treatment for head and neck cancer is xerostomia, which leads to a poor quality of life. Current toxicity models for the parotid glands are based on mean dose to the entire gland, however recent small animal studies show that there may be a better correlation between the radiation dose to the sub-regions of the parotid glands and the risk of toxicity.

Delivered dose to the parotid gland sub-regions can be estimated with dose accumulation using deformable image registration. The purpose of this study was to build a population model describing the relationship between radiation dose and parotid gland shrinkage and to use this model to aide in the improvement of deformable image registration accuracy of head and neck images during radiation therapy. A modified thermal-expansion coefficient was determined based on the population model and served to apply shrinkage to each element of the finite element model based on the planned dose. Based on the thermal-expansion coefficient, dose-based boundary conditions were applied to an in-house biomechanical model-based deformable image registration algorithm. Assessment of the accuracy of the model was based on target registration error (TRE). The accuracy was improved significantly ($p = 0.01$) from a TRE of 1.6 ± 0.9 mm for the standard model to a TRE of 1.4 ± 0.8 mm for the proposed model. Application of these models may allow better estimation of the delivered dose during radiation therapy and aide in the development of improved toxicity prediction models.

6.2 Introduction

Over 550,000 new cases of head and neck (HN) cancer are diagnosed leading to 380,000 deaths per year in the United States [139]. Most HN cancers recommended management include radiation

therapy. While treatment planning aims to avoid irradiation of the salivary gland to preserve their function, a total sparing is rarely achievable. A major side effect of the parotid gland (PG) irradiation is Xerostomia, which is defined as oral dryness [140]. Xerostomia leads to dysphagia, or difficult swallowing, which lowers the quality of life (QoL) substantially. A better sparing of the PG may improve the QoL of HN cancer patients [131]. Currently, prediction of toxicity for the PG is based on the mean dose to the entire gland. However, small animal studies show that the relationship between delivered dose and the sub-regions of the glands may be more meaningful [74], indicating the need for accurate estimation of the dose delivered to the PG sub-regions.

HN patients lose 6-10% of their body weight on average over cancer treatment [141], leading to anatomical changes and deformation. Additionally, tumors shrink by an average of 70% of their initial volume by the end of radiation treatment [31]. Due to those complex anatomical changes throughout RT, the precise delivery of the planned dose is compromised. Especially, as the tumor can be situated near a PG, a sharp dose gradient over the PG can exist, where even a small shift or volumetric change could potentially lead to a dramatic difference in the dose. The delivered dose to the PG has been reported as high as 30% above the planned dose distribution [142]. This increase can be clinically significant, as a 1 Gy dose increase correlates with a 5% decrease in PG function [143]. Understanding the response of the PGs to the dose is then hampered by the differences between planned and delivered dose.

Dose accumulation using deformable image registration (DIR) can be used to estimate delivered dose to the PG. Many studies investigated deformable image registration between the planning CT and followup CT or Cone-Beam CTs (CBCT) to accumulate the dose in the head and neck region

[125,144-146]. However, since only the boundaries of the gland are visible with these imaging modalities, the deformation estimated inside the gland is the result of a regularization model of the displacement vector field (DVF). Since a correlation exists between the mean dose and the volume change of the PG [147], the internal deformations of the gland may actually be more complex than an interpolation from the gland surface deformation, and may depend on the heterogeneity of the dose inside the gland and the variable sensitivity of the tissue in the gland.

Previous studies modeled the dose response of organs by using dose-based boundary conditions into a biomechanical model-based deformable image registration algorithm, Morfeus [63]. Al-Mayah *et al* demonstrated the feasibility of simulating a dose-induced shrinkage of the PGs for 5 patients by assigning a negative thermal expansion coefficient to the tetrahedral elements of the PG in Morfeus and by applying to each element a temperature corresponding to the locally planned dose [136]. More recently, Polan *et al* investigated the use of additional dose-based boundary conditions on the liver [97] versus the use of boundary conditions on the liver surface only. The use of dose BCs, based on a model established from 33 prior radiation therapy patients, was evaluated for the DIR of pre- and post-RT CT scans for 7 patients. The target registration error (TRE) was significantly improved when using dose BCs suggesting the model better described the spatial volume change distribution inside the liver.

The goal of this study was to establish the relationship between dose and PG shrinkage for a population of patients and to use that model to improve the DIR accuracy of longitudinal images of the PG during RT. MR images acquired at the time of treatment planning and mid-treatment were collected for a series of patients to visualize anatomical details inside the PG and allow

quantitative evaluation of the proposed method accuracy. The DVF differences in the PG, when considering dose-boundary conditions or not, were analyzed.

6.3 Materials and Methods

6.3.1 Patient Data

Nineteen patients treated at the MD Anderson Cancer Center with RT for oropharyngeal cancer in 31-35 fractions were retrospectively evaluated. Detailed patient data are shown in Table 1. For each patient, the following images were collected and imported in the treatment planning system RayStation: the treatment planning CT, planned dose distribution, a treatment planning MR and a MR acquired approximately at mid-treatment between fractions 15 and 25.

Table 6-1. Treatment details for patients included in the study.

BOT stands for base of tongue.

Characteristic	Value
RT modality (IMRT/Proton)	15/4
Total dose (range, median)	63-72,70
Total fractions (range, median)	31-35,33
Treatment duration (range, median)	39-50,44
Gender (M/F)	19/0
Age (range, median)	32-78,67
Tumor subsite (tonsil/BOT/neck)	11/6/2
HPV/p16 status (+/-)	18/1
Concurrent chemotherapy (yes/no)	13/6

1.1.1. Image data

For each patient, all PGs were delineated on the MR images. In order to ensure consistency between the contours on the planning MR and mid-treatment MR, the PGs were first delineated on the planning MR and automatically propagated onto the mid-treatment MR using the intensity-based DIR method available in RayStation. Manual edits were made on the propagated contour when necessary. The mandible, C1 vertebrae and C3 vertebrae were manually delineated on the planning CT. The contours of these bony structures were copied on the MR images following the global rigid registration and translated or rotated when necessary. The dose grid, rigidly registered MR images and all the contours for ten patients were exported from the treatment planning software for processing with an in-house implementation of Morfeus.

For each PG, three corresponding anatomical landmarks were placed on the planning and mid-treatment MR, with an effort to have them spatially well distributed. The landmarks were used to assess the registration accuracy by measuring, in each PG, the target registration errors (TRE) defined as the mean Euclidean distance between each pair of landmarks.

1.1.2. Analysis of the dose volume relationship

For the 38 PGs of the 19 patients in the optimization cohort, the volume of the gland was measured on both MR images and the planned mean dose was scaled to the mean dose delivered at the time of the mid-treatment MR. The relationship between volume change and mean delivered dose was analyzed.

6.3.2 Biomechanical Model-based Deformable Image Registration

6.3.3 Neck flexion alignment

Our previous work demonstrated that a global rigid registration of the HN anatomy could yield to a poor alignment of the PGs because of varying angles of neck flexion. The use of the biomechanical model-based DIR implemented in RayStation based on boundary conditions on only the mandible, C1 and C3 vertebrae allowed to correct for that neck flexion and improve the alignment of the PGs. In this study, this approach was systematically applied to register the mid-treatment MR onto the planning MR. The obtained $DVF_{flexion}$ were exported from RayStation and used to deform the contour of the PG corresponding to the mid-treatment MR.

6.3.4 Standard Morfeus

The in-house implementation of Morfeus [63] was used for this study. The workflow of the method is illustrated Figure 1. For each PG corresponding to the planning, Morfeus creates a tetrahedral mesh of the gland and surrounding anatomy. Based on previous studies, the Young's Modulus, defined as the ratio of the stress to the strain (Eq. 6-1), was set to 7.8 kPa and the Poisson's Ratio, defined as the ratio between the transverse and axial strain (Eq. 6-2), was set to 0.45. These properties were assigned to the PG elements.

$$E = \frac{\sigma}{\epsilon} \quad (6-1)$$

Young's Modulus is defined by the ratio of the uniaxial stress, σ , to the strain, ϵ .

$$\nu = -\frac{\epsilon_{trans}}{\epsilon_{axial}} \quad (6-2)$$

Poisson's Ratio is the transverse strain divided by the axial strain.

A surface projection algorithm is applied between the PG surfaces of the planning and mid-treatment after resolving the neck flexion. The displacements determined on the surface nodes are used as boundary conditions in the model to solve the displacement of all the internal mesh nodes in a finite-element analysis. The surface projection algorithm used for this study was based on the accelerated Demons algorithm [101]. The forces of the Demons method can be described using the optical flow equation. The following equation describes the displacement \vec{d} of a point from the reference PG, R, to the point corresponding from the target PG, T:

$$\vec{d} = \frac{(t - r)\vec{\nabla}r}{|\vec{\nabla}r|^2 + (t - r)^2} \quad (6-3)$$

where \vec{d} represents the displacement in each direction, t and r represent the intensities of the target and references PGs, and $\vec{\nabla}r$ represents the reference PG gradient. The Demons algorithm

solves the equation iteratively using optical flow and regularization of the DVF. The regularization component smooths, minimizes noise, and keeps the deformed image continuous. Since the deformation is driven only by the gradient of the reference PG, the method was improved [101] with the addition of an active force to the deformed (target) image. This active force, \vec{f}_t , can be described in the following equation.

$$\vec{f}_t = - \frac{(r - t)\vec{\nabla}t}{|\vec{\nabla}t|^2 + (r - t)^2} \quad (6-4)$$

Finally, the total force of a point can be described with the equation

$$\vec{f} = \vec{f}_r + \vec{f}_t = (t - r) * \left(\frac{\vec{\nabla}r}{|\vec{\nabla}r|^2 + (r - t)^2} + \frac{\vec{\nabla}t}{|\vec{\nabla}t|^2 + (r - t)^2} \right) \quad (6-5)$$

where \vec{f}_r represents the force from the gradient of the reference PG and \vec{f}_t represents the force from the gradient of the target PG. These force calculations are appropriate with relatively small deformations. To account for larger deformations, the multiresolution method [102] was applied. This method applies an iterative demons diffusion approach, based on low-resolution images, and increases calculation speed by improving the convergence. An additional normalization factor (α) was also applied. With this incorporation of α , adjustments to the force strength at each iteration is possible. This factor is applied as shown in the following equation

$$\vec{d} = (t - r) * \left(\frac{\vec{\nabla}r}{|\vec{\nabla}r|^2 + a^2(r - t)^2} + \frac{\vec{\nabla}t}{|\vec{\nabla}t|^2 + a^2(r - t)^2} \right) \quad (6-6)$$

where larger deformations require smaller a values.

Finally, the calculated displacements are resampled on the grid of the planning MR image to provide a DVF mapping the PG voxels of the planning image to the mid-treatment image.

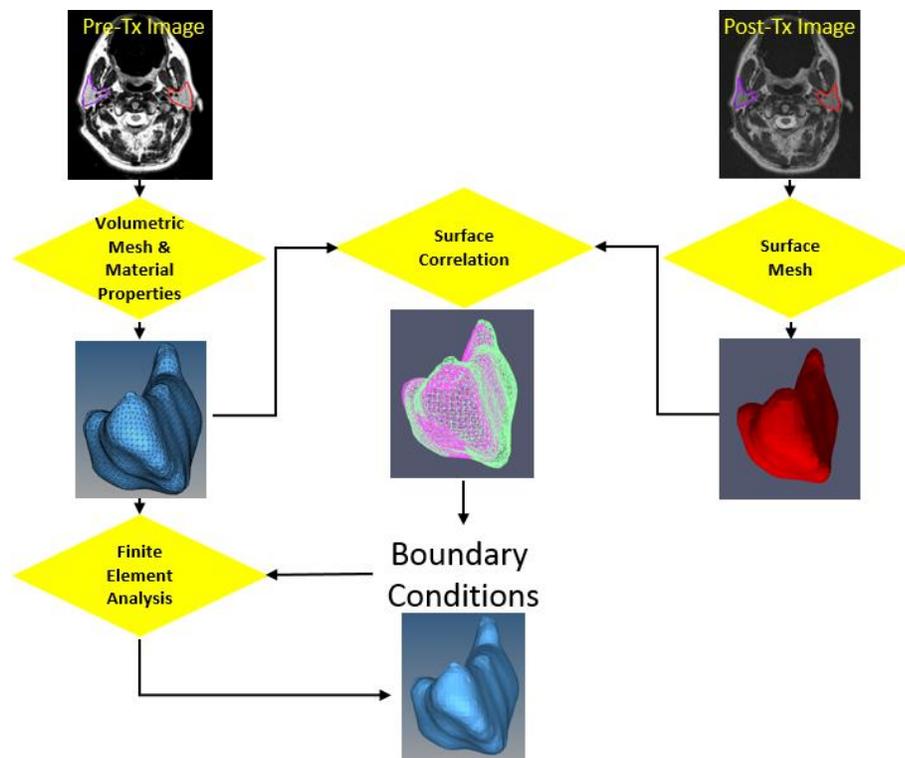


Figure 6-1. Depiction of standard Morfeus process.

The initial image is aligned to a corresponding image. A volumetric mesh is generated and material properties are assigned to the pre-treatment image. A surface mesh is generated from the mid-treatment image. A surface correlation is calculated and applied as boundary conditions to the pre-treatment model. FEA solves the deformation between the pre-treatment to mid-treatment PG image.

6.3.5 Modeling the dose-induced shrinkage

Morfeus was expanded to simulate a shrinkage of the PG as a response to a dose distribution as illustrated Figure 2. For this, each tetrahedral element of the PG was assigned an additional material property a thermal expansion coefficient. Where the classic thermal expansion coefficient involves the change in temperature, the modified version α applied in this study involves the dose to the PG:

$$\alpha = \frac{(V_m - V_i)}{3V_i D} \quad (6-7)$$

Where V_m the volume of the gland at mid-treatment is, V_i is the volume at planning, and D is the mean dose to the PG.

Optimization experiments showed that for the FEM solver to provide a stable solution considering the large volume variations to simulate, more compressibility of the tetrahedral elements had to be allowed by decreasing the Poisson's ratio to 0.3 compared to a value 0.45 for standard Morfeus.

After application of the dose-induced shrinkage as a first step (Figure 2), the post-response PG was used in a second step of a standard run of Morfeus (Figure 1.).

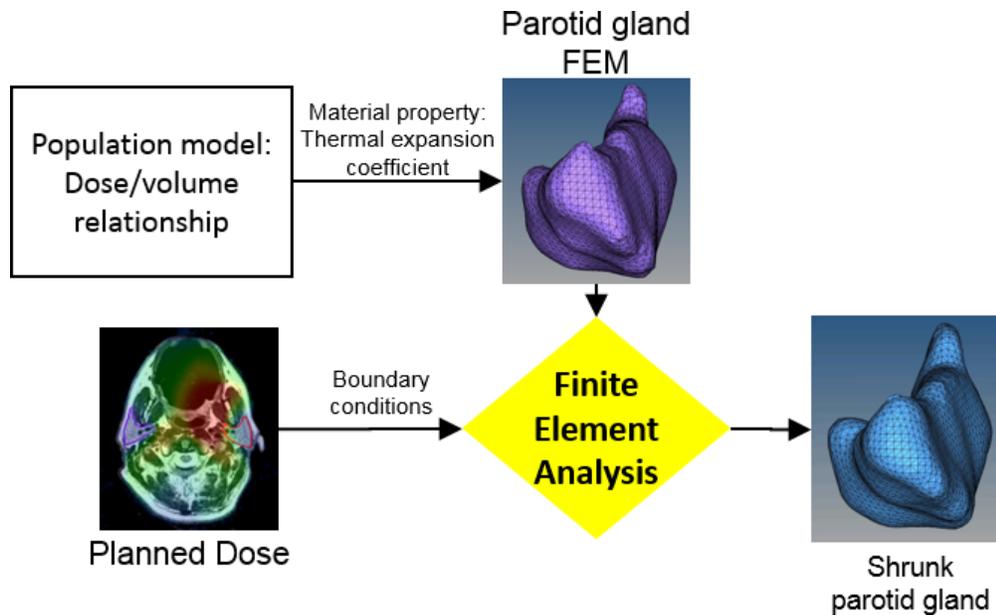


Figure 6-2. Depiction of dose-based boundary condition application to Morfeus.

The modified thermal expansion coefficient is developed based on the population model. A FEM of the PG is created. Boundary conditions are applied to the FEM of the PG and the PG responds with the application of the boundary conditions from FEA.

6.3.6 Displacement vector fields analysis

After DIR using standard Morfeus ($Morfeus_{std}$) or Morfeus with additional dose boundary conditions ($Morfeus_{dbc}$), the DVF was composed with the initial $DVF_{flexion}$ and the resulting DVF imported in RayStation for accuracy evaluation.

The Jacobian matrix is computed based on the DVF. The determinant of the Jacobian matrix is then calculated to represent the volumetric change of each voxel based on the DVF. The Jacobian determinant is described with the following equation, where T is the deformation, x is the voxel coordinates at which the local expansion or compression the volume change is being calculated.

$$J_T(x) = \det \begin{pmatrix} \frac{\partial T_x(x)}{\partial x} & \frac{\partial T_x(x)}{\partial y} & \frac{\partial T_x(x)}{\partial z} \\ \frac{\partial T_y(x)}{\partial x} & \frac{\partial T_y(x)}{\partial y} & \frac{\partial T_y(x)}{\partial z} \\ \frac{\partial T_z(x)}{\partial x} & \frac{\partial T_z(x)}{\partial y} & \frac{\partial T_z(x)}{\partial z} \end{pmatrix} \quad (6-8)$$

In the case that the Jacobian determinant is greater than one, volumetric expansion has occurred.

In the case that the Jacobian determinant is less than one, volumetric compression has occurred.

In order to compare the distribution of the volume changes inside the gland between Morfeus_{Std} and Morfeus_{dbc} and to verify the correlation with the dose distribution when considering Morfeus_{dbc}, the determinant of the Jacobian was computed for one PG example.

6.4 Results

6.4.1 Population Model

Figure 3 represents the measured volume changes versus the mean dose delivered at mid-treatment for the 38 PGs included in the study. A linear regression with an intercept set at the origin yielded to an expansion coefficient $\alpha = -0.0033$ following Eq 1.

However, experimentation showed that to achieve the final desired volume changes in

Morfeus_{dbc}, α had to be scaled by a factor of 1.3. $\alpha' = \alpha * 1.3 = -0.0043$.

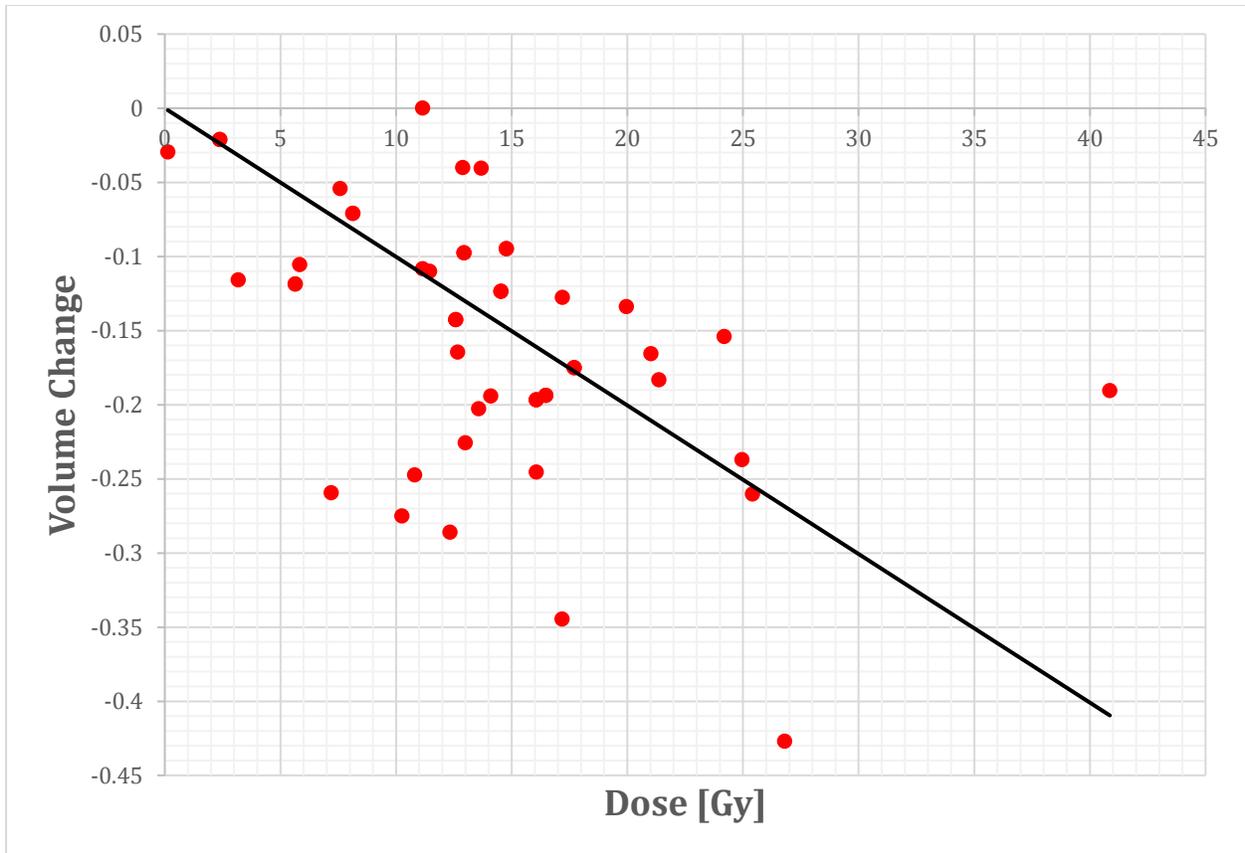


Figure 6-3. Population Model.

Population model of the volume change (absolute) over the dose [Gy] for calculation of the modified thermal expansion coefficient. Each point represents a different PG.

6.4.2 Target Registration Errors and Dice Similarity Coefficients

Box plots of the TRE for the initial alignment based on the vertebrae and mandible, $Morfeus_{std}$, and $Morfeus_{dbc}$ are shown in Fig. 4. The initial alignment resulted in a mean TRE of 3.1 ± 1.6 mm. $Morfeus_{std}$ resulted in an average TRE of 1.6 ± 0.9 mm. $Morfeus_{dbc}$ resulted in an average TRE of 1.4 ± 0.8 mm, a statistically significant improvement ($p = 0.01$) over $Morfeus_{std}$, based on a 2-Tailed paired Student's T-Test.

Since Morfeus uses the contours of the PG, the Dice Similarity Coefficient (DSC) for the PG was always measured above 0.93 and no significant differences were found between Morfeus_{std} and Morfeus_{dbc}.

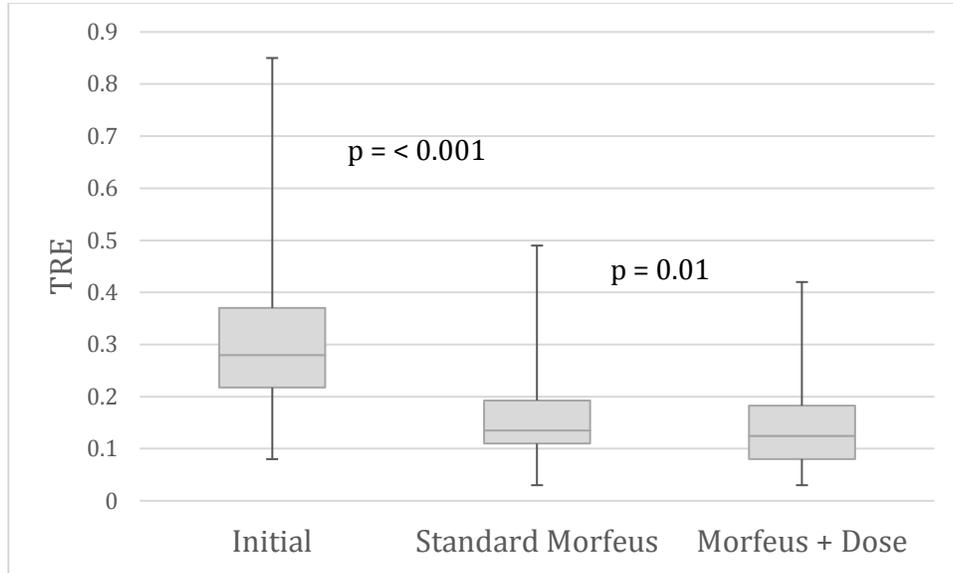


Figure 6-4. Resulting TRE for each DIR Method.

The whiskers represent the maximum and minimum values. The middle line represents the median values. The bottom horizontal lines represent the first quartiles. The top horizontal lines represent the third quartiles.

6.4.3 Jacobian determinant distribution and impact on dose mapping

1.1.3. Volume change maps

For comparison of the spatial volume change distributions, the Jacobian determinant maps were computed from the DVFs obtained after Morfeus_{std}, from the DVFs obtained after applying only dose BCs and from the DVFs obtained after Morfeus_{dbc}.

Figure 5 represents for an example PG, the dose distribution on a coronal slice and the histogram of the doses in the whole gland as well the different Jacobian maps, their histogram and the correlation with dose distribution.

A very low correlation between the dose distribution and the volume changes after $Morfeus_{std}$ was found with a Spearman correlation coefficient $r=-0.10$. A very high correlation with $r=-0.96$ was measured after the first step of $Morfeus_{dbc}$ when only dose BC were applied, indicating that our method worked as intended.

The correlation decreased after applying boundary conditions on the surface of the parotid gland in the second step of $Morfeus_{dbc}$ but remained relatively high ($r=-0.62$).

For this patient, the volume change was 20%. The DSC was the same (0.96) for $Morfeus_{std}$ and $Morfeus_{dbc}$ but the mean TRE improved when using $Morfeus_{dbc}$ from 1.6 mm to 1.0 mm, which supports our hypothesis of a correlation between dose and volume change distributions.

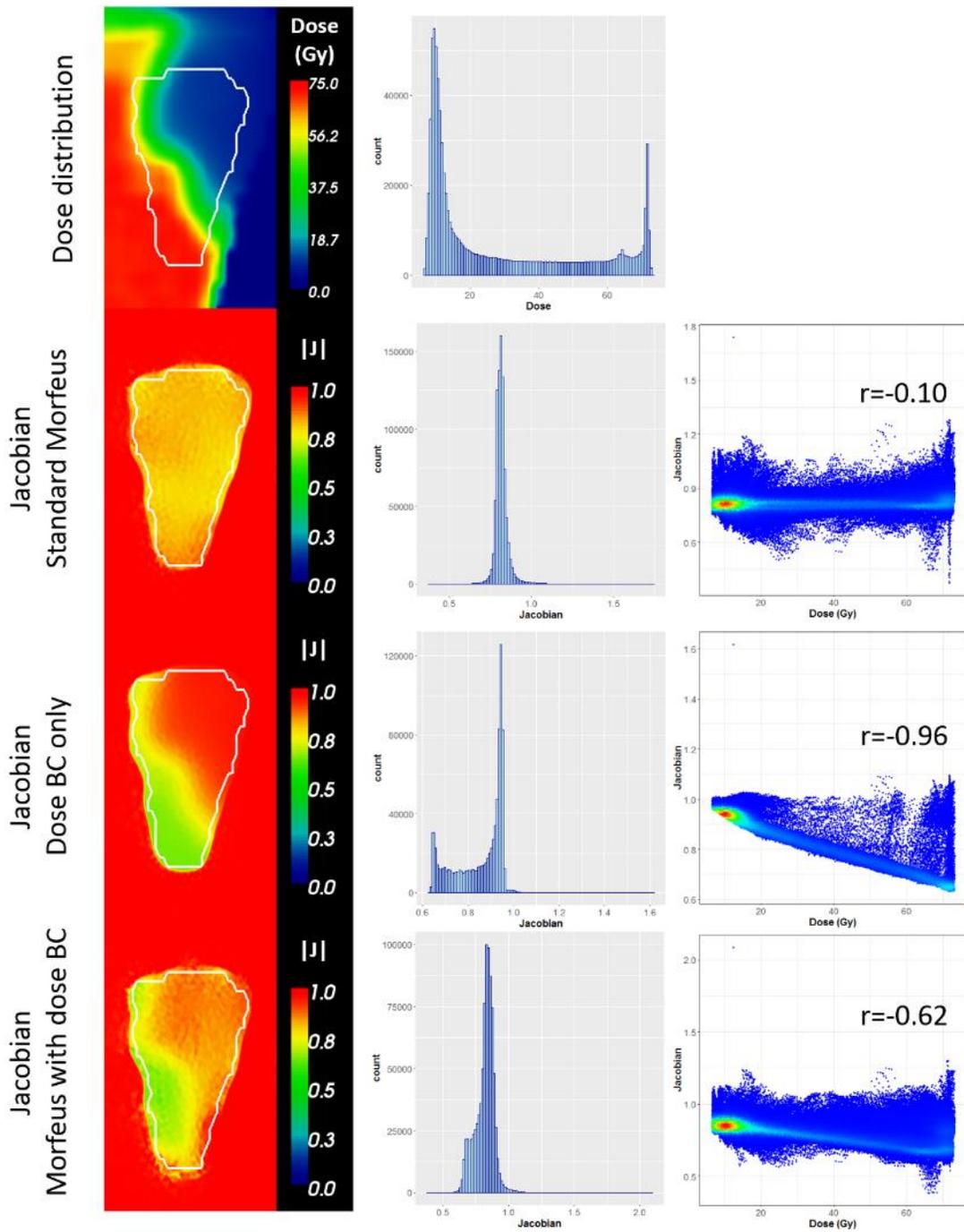


Figure 6-5. Jacobian Maps.

Left column: Images of the dose and Jacobian maps for the same coronal slice of a PG. The white contour represents the PG. Middle column: Histograms of the corresponding images. Right column: scatter plots representing the correlation with the Spearman correlation coefficient r between the Jacobian map and the dose distribution inside the PG.

1.1.4. DVF differences

The mean absolute DVF difference between $\text{Morfeus}_{\text{std}}$ and $\text{Morfeus}_{\text{dbc}}$ inside the example PG (Figure 6) was 0.7 ± 0.3 mm and the 95th percentile was 1.4 mm. Small differences were found on the boundaries of the PG, as the surface projection method in each DIR method resulted in similar surface point-to-point correspondences.

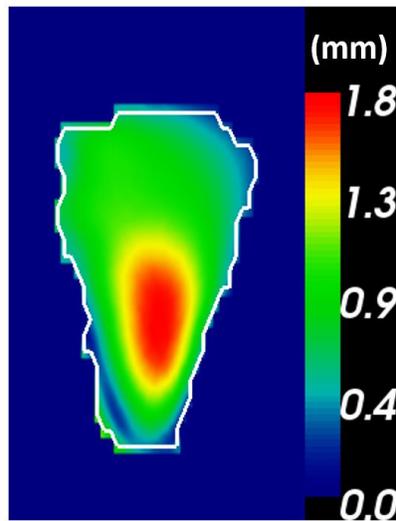


Figure 6-6. DVF Differences.

Norm in millimeters of the vector differences between the DVFs obtained with standard Morfeus and Morfeus with dose BC on the same coronal slice as for Figure 6.

1.1.5. Dosimetric impact

To assess the impact of the DVF differences for dose mapping applications, the deformation fields obtained after $\text{Morfeus}_{\text{std}}$ and $\text{Morfeus}_{\text{dbc}}$ were inverted and used to map the planned dose distribution from the planning MR to the mid-treatment MR. Figure 7 represents the absolute difference between the mapped dose distributions on the same coronal slice as previously and the

dose-volume histograms. While the DVF differences could be considered small, these differences translated for this example into impactful dosimetric differences. For example, D50, the minimum dose received by 50% of the volume was 17 Gy and 20 Gy for Morfeus_{std} and Morfeus_{dbc} respectively.

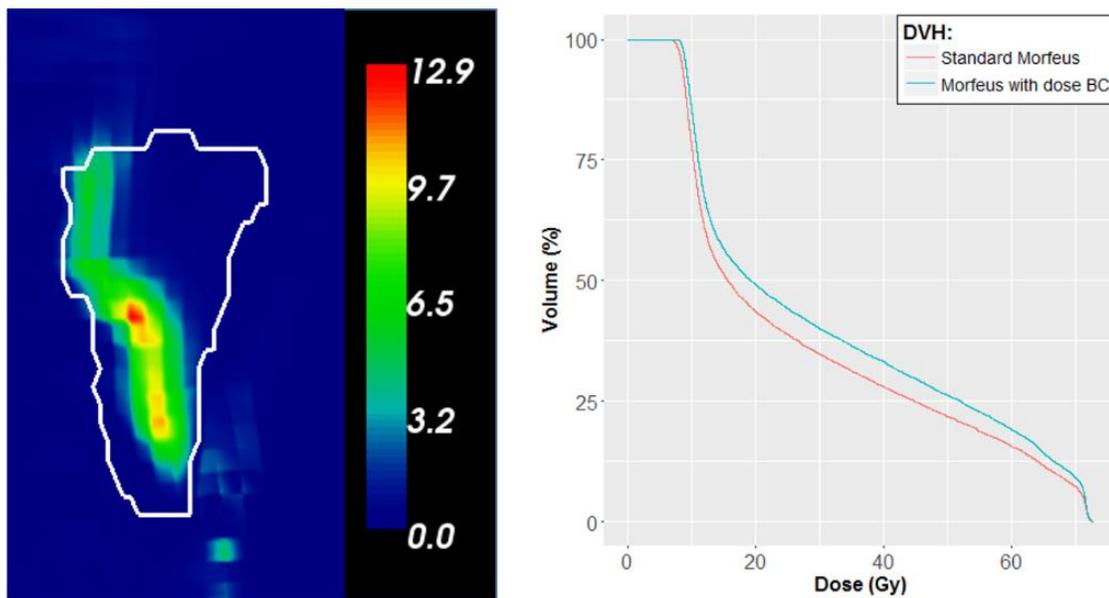


Figure 6-7. Dose Distribution Differences and DVH.

Left: Image of the absolute differences in Gy between the dose distributions mapped from the planning image to the mid-treatment image considering standard Morfeus or Morfeus with dose BCs. The white contour represents the contour of PG from the mid-treatment MR. Right: Comparison between the dose-volume histograms.

6.5 Discussion

A biomechanical model-based DIR method has been expanded to describe the dose-induced shrinkage of the PGs due to RT. Following initial alignment of the PGs, the first step of this DIR method involved applying the dose-based boundary conditions to the PG of the planning MR.

The second step consisted of applying boundary conditions on the PG surface to refine the deformation. The calculation of the Jacobian map of the final deformation field demonstrated that our numerical implementation provides the desired volumetric response behavior of the PG. The accuracy of the method was assessed by measuring the TRE. However, in MR images identifying landmarks can be challenging, which limits the comprehensiveness of this evaluation. The addition of MR sialography, which improves the conspicuity of the vessels in the PG, would aid in this effort. Work is ongoing to develop these sequences and acquire them in future clinical trials.

The hypothesis of this study was that the volume change inside the PGs depends on the dose distribution, with the tissues receiving higher dose being more likely to shrink than tissues in low dose regions. To test this assumption, the accuracy of the proposed DIR method using dose boundary conditions was compared to the accuracy of the standard DIR method based only on the surface boundary conditions. Both DIR methods resulted in similar DSC, but the statistically significant improvement in TRE using the expanded DIR method confirms our hypothesis, even in this limited cohort of patients. Consequently, traditional DIR methods based only on the PG boundaries may lead to inaccurate estimations of deformation inside the gland, and therefore inaccurate estimates of the delivered dose.

This study used MR images to enable the identification of internal landmarks for TRE evaluation. However, MR images are not always available for patients and therefore DIR must be performed on CT images, where internal landmarks are not visible. The method proposed here will maintain a consistent level of accuracy regardless of imaging modality, as long as the PG boundaries can be

identified as it only relies on the boundaries of the PG and the planned dose distribution. Having a robust algorithm is critical to enable widespread deployment of dose accumulation in the PG.

Although the average TRE improvement was modest, 1.6 mm to 1.4 mm, the new deformation model can yield to clinically meaningful dose differences, as illustrated in the highlighted patient. Future work will also investigate if the dose estimated with the proposed deformation model will lead to better predictors of the toxicity.

6.6 Conclusion

A model was developed to describe the volumetric response of the parotid glands to radiation dose, using dose-based boundary conditions, applying shrinkage based on radiation dose to each tetrahedral element in the model. This model had improved results over the standard biomechanical model-based deformable image registration algorithm, as well as a higher correlation between the volumetric change inside the gland and the radiation dose. The use of this model, applied to images with clear ductal anatomy, could allow for dose accumulation to the substructures of the parotid glands. Dose accumulation to substructures of the parotid gland can be correlated with predicted toxicity for the glands. Ultimately, more accurate toxicity models could be developed based on this work.

Chapter 7. Conclusions and Future Work

7.1 Conclusions

In chapter 2, the need for prospective clinical trials to evaluate the clinical impact of developing NTCP models based on accumulated dose, rather than planned dose, was demonstrated. The study used simulated data for 600 patients to develop toxicity models and indicated that historical NTCP models, based on planned dose, may overestimate the toxicity risk for lower doses, but underestimate the risk of toxicity for higher doses with errors up to 21%. The differences between NTCP models based on accumulated dose compared to NTCP models based on planned dose is greater for duodenum than for stomach. However, with the availability of deformable registration-based dose accumulation using volumetric daily imaging, improved NTCP models are possible and should be included in the development of future clinical trials.

With the use of this model developed in chapter 3, HN cases that would benefit from replanning could be identified. For submandibular glands, a dose deviation threshold of 3.5Gy at fraction 15 can predict the need to replan a patient. Therefore, a model was developed to aid physicians in the decision to replan HN patients based on the submandibular glands by fraction 15 of radiation treatment. At least one organ exceeded the dose deviation threshold for all CBCTs in which a DIR failure was identified, indicating the possibility of an automated process to perform DIR evaluation only for cases that would result in a dose deviation. Additionally, the accuracy of the

DIR algorithm in the commercial treatment management software was evaluated using contour propagation.

This retrospective study in chapter 4 demonstrated that PG deformation due to neck flexion can be modeled using the biomechanical model-based method with indistinguishable results from the intensity-based method ($p=0.13$ for volume change up to 20%). The minimum DSC based on the biomechanical-based method is 0.53, while the minimum DSC based on the intensity-based method is 0.17. This combined with the many outliers from the intensity-based method shows that while the intensity-based method can slightly outperform the biomechanical model-based method on average, there can be substantial failures using the intensity-based method, which is not observed with the biomechanical model-based method. The biomechanical model can be an initial step in a comprehensive model to describe the anatomical, patient positioning, and volumetric changes to the salivary glands during HN radiation therapy, and eventually aid in the development of toxicity models for this region.

In chapter 5, a biomechanical model-based DIR method was developed to describe the volumetric response of the PG to RT. A modified thermal expansion coefficient was applied to an existing in-house biomechanical model-based DIR algorithm. The adaptation to the existing DIR method improved the average TRE significantly ($p = 0.01$) from $1.6 \text{ mm} \pm 0.9 \text{ mm}$ to $1.4 \text{ mm} \pm 0.8 \text{ mm}$, which should translate into better estimations of the delivered dose through the course of radiation therapy and potentially in the development of more accurate toxicity prediction models.

7.2 Summary and Future Work

This dissertation consisted of a series of experiments based on finite element modeling that lead to better understanding of the uncertainty between planned and delivered radiation dose and the development of improved statistical models of toxicity. This work aimed to improve the understanding of toxicity during liver cancer radiotherapy by quantifying deviations between accumulated and planned dose in luminal NTCP models, develop predictive models for treatment replanning decisions in head and neck treatments, resolve the anatomical changes due to neck flexion using biomechanical modeling, and develop biomechanical models of dose response of the parotid glands in order to improve toxicity prediction models. The studies described were novel and will hopefully lead to continued development and advancement. Three paths of interest are detailed below. Briefly, future directions include applying the models previously described throughout this manuscript to additional anatomical sites. Another path involves the use of imaging modalities that were not explored throughout this dissertation. Finally, a long-term goal for future work related to this thesis includes clinical trials to expand the patient cohorts studied in the previous chapters.

The models developed for this dissertation were based on the abdominal region and the HN region. Further studies should explore other organ regions to improve toxicity risk of patients affected by those cancers, as well as to enhance the robustness of biomechanical models. One such organ site is the brain. Common brain cancers include glioblastoma multiforme (GBM), which is a type of glioma. Gliomas are cancers of the glial cells in the brain. GBM is an infiltrative tumor, and tumor cells can be found far beyond the margins visible on medical images. The brain is comprised of a large number of critical structures which are typically not

visualized on standard imaging. Determining the correlation between the dose to these critical structures and toxicity is essential for the QoL of the patient. However, it would not be feasible for a radiation oncologist to contour all of these structures for each patient during radiation therapy. There exists the need for an atlas of normal tissue contours that can be mapped onto each patient that undergoes radiation therapy for brain cancer, to spare these critical brain structures and preserve the patient's QoL. Biomechanical model-based DIR could aid in the development of such an atlas. Preliminary work of applying biomechanical model-based DIR to glioma patient images included post-surgery/pre-radiation therapy and post-radiation therapy images. Boundary conditions were placed on the GTV and brainstem in order to align normal structures. For one patient with substantial normal tissue shifts due to a reduction in swelling from the pre-radiation therapy image to the post-radiation therapy image, boundary conditions on the GTV alone yielded a DSC improvement from 0.56 to 0.74, 0.36 to 0.48, and 0.68 to 0.71 for rigid registration to DIR in the ventricles, left hippocampus, and right hippocampus, respectively. Applying the brain stem and GTV as boundary conditions for the other patients in the study slightly improved the alignment of the normal tissues. Boundary conditions on the surgical resection cavity are also being explored for this study.

The work presented in this dissertation was based on CT and MR images. However, several other imaging modalities are used in the diagnosis and treatment of cancer. Namely, MR-sialography (Figure 1), which is an imaging technique that relies on the bright signal produced from fluids on heavily T2-weighted images ($TR = 3500$ and $TE = 500$ ms). This imaging modality can be useful in the imaging of HN cancer, as the salivary flow can be visualized. Patients receiving imaging of this type are given vitamin C to stimulate salivary output and

improve the visibility of the salivary ducts. As the parotid and submandibular glands produce saliva, they can be successfully imaged using this technique, as shown in Figure 1. Future work involves using this imaging modality to obtain a detailed representation of the parotid gland anatomy. The segmentation of the salivary duct tree will allow a more thorough evaluation of the accuracy of the parotid gland shrinkage model proposed in this Ph.D. work by measuring how well the biomechanical model matches the whole duct tree and not only a limited number of manually picked landmarks. Using the duct segmentation to place additional internal boundary conditions in the proposed biomechanical model will also be investigated to evaluate the accuracy of the image alignment. The dose response assessment could be improved considering this advanced detailed image modality.

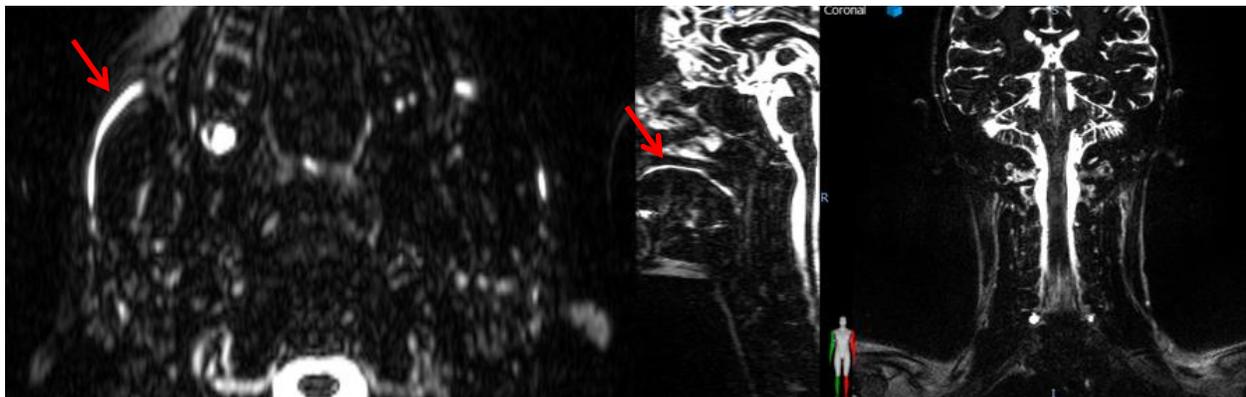


Figure 7-1. Examples of sialography images.

On the left is an axial slice depicting a bright parotid gland duct. In the middle is a sagittal view. On the right is a coronal view. The red arrow points to the location of the parotid gland duct.

Ultimately, applying the models developed throughout this Ph.D. to larger patient cohorts will be instrumental in validating these models for use in the clinic. This thesis work included the first model of normal tissue complication probability based on accumulated dose, which is now possible on a large scale due to recent technological advancements. This work involved a cohort of thirty patients, which were resampled to project the potential effect of using accumulated dose

for NTCP curves in place of the standard planned dose. The study demonstrated the potential clinical importance of including accumulated dose in the development of NTCP models in future clinical trials, in which the delivered (accumulated) doses and toxicity outcomes would be prospectively recorded to improve the understanding of the dose-response relationship. Such a trial could characterize the differences between planned and accumulated dose values and the true impact of these differences on NTCP models. Clinical trials for cancer research are conducted by a non-profit research organization, for example NRG Oncology. A phase III trial can include hundreds of patients assigned to either the standard or new treatment, with the goal of determining if the new treatment is an improvement over the standard. If daily imaging was collected for each patient undergoing a phase III trial for a new fractionated radiation treatment method, dose accumulation could be performed in order to correlate the accumulated dose with toxicity. With accumulated dose tabulated for hundreds or thousands of patients, more accurate toxicity models could be developed. One ongoing NRG trial for the treatment of liver cancer is a phase III randomized trial to study the efficacy of photons versus protons in overall survival for HCC patients. Another current phase III randomized trial is comparing Sorafenib (chemotherapy) with SBRT followed by Sorafenib in HCC patients to investigate if overall survival is improved when SBRT is used in addition to Sorafenib. In addition to NTCP modeling, this thesis also includes the first ever predictive model to aid in the decision to replan head and neck patients, where the current decision to adapt is based on the physician's 'best guess' and has no standardized metrics or guidelines to assist in this complex task. These HN predictive models, described in the second chapter of this thesis, require prospective validation in a larger, multi-institutional patient cohort. The model was developed based on 100 HN cancer patients treated at the University of Michigan and externally validated on 52 patients treated at

the MD Anderson Cancer Center. The development of adaptive clinical trials utilizing these predictive models are currently under development.

Finally, this work included the first dose response study of the parotid glands that included internal validation inside the organs. First, 100 patients were aligned based on bony anatomy, and next the volumetric response was added to the model. Accurate modeling of the parotid gland during radiation therapy, which is complex due to the dose functions, will allow a better estimate of the delivered dose distribution inside the parotid gland. Ultimately, these models will yield better understanding of the toxicity and dose relationship and will allow the development of better predictive models to be used in advanced adaptive radiation therapy strategies. No other approach currently accounts for the dose distribution inside the gland. These studies also require validation on a larger cohort of patients to be used clinically. A linear relationship was found between the volumetric change and mean dose to 40 parotid glands. However, more data points could show a more complex relationship. One patient in this cohort had a very high dose of 40Gy to the left parotid gland. The addition of more patients could provide more data points with very high dose. With a larger population model, the dose-response model would be more robust.

To summarize, this dissertation proposed original approaches to describe the uncertainties between planned and delivered dose during radiation therapy and demonstrated their utility for advanced adaptive radiotherapy techniques. There will be more detailed imaging involved in the management of cancer treatment, as in the case with the MR-linear accelerator, which provides real-time high-resolution imaging during radiation therapy, providing the opportunity to further demonstrate the refinement and use of these models.

References

1. American Institute for Cancer Research WCRF. Worldwide cancer data, global cancer statistics for the most common cancers. In: Editor, editor^editors. Book Worldwide cancer data, global cancer statistics for the most common cancers.
2. Mittal S, El-Serag HB. Epidemiology of hcc: Consider the population. *Journal of clinical gastroenterology* 2013;47:S2.
3. Francis D. Trends in incidence of head and neck cancers in india. *European Journal of Cancer* 2018;92:S23.
4. Joshi P, Dutta S, Chaturvedi P, Nair S. Head and neck cancers in developing countries. *Rambam Maimonides medical journal* 2014;5.
5. Ragin C, Modugno F, Gollin S. The epidemiology and risk factors of head and neck cancer: A focus on human papillomavirus. *Journal of dental research* 2007;86:104-114.
6. Mourad M, Jetmore T, Jategaonkar AA, Moubayed S, Moshier E, Urken ML. Epidemiological trends of head and neck cancer in the united states: A seer population study. *Journal of Oral and Maxillofacial Surgery* 2017;75:2562-2572.
7. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, Jiang B, Goodman MT, Sibug-Saber M, Cozen W. Human papillomavirus and rising oropharyngeal cancer incidence in the united states. *Journal of clinical oncology* 2011;29:4294.
8. Ridge JA, Glisson BS, Lango MN, Feigenberg S. Head and neck tumors. *Cancer management: a multidisciplinary approach* 2013;11.
9. Seattle Cancer Care Alliance UoWM. Head and neck cancer survival rates.
10. Pulte D, Brenner H. Changes in survival in head and neck cancers in the late 20th and early 21st century: A period analysis. *The oncologist* 2010;15:994-1001.
11. Society AC. Treatment of liver cancer, by stage, potentially resectable or transplantable liver cancers (stage i and some stage ii tumors) 2019.
12. Kanat O, Gewirtz A, Kemeny N. What is the potential role of hepatic arterial infusion chemo-therapy in the current armamentarium against colorectal cancer. *Journal of gastrointestinal oncology* 2012;3:130.
13. Society AC. Tumor ablation for liver cancer 2019.
14. Parikh ND, Marshall VD, Green M, Lawrence TS, Razumilava N, Owen D, Singal AG, Feng M. Effectiveness and cost of radiofrequency ablation and stereotactic body radiotherapy for treatment of early-stage hepatocellular carcinoma: An analysis of seer-medicare. *Journal of medical imaging and radiation oncology* 2018.
15. Sandroussi C, Dawson LA, Lee M, Guindi M, Fischer S, Ghanekar A, Cattral MS, McGilvray ID, Levy GA, Renner E. Radiotherapy as a bridge to liver transplantation for hepatocellular carcinoma. *Transplant International* 2010;23:299-306.
16. Cancer.Net. Head and neck cancer: Treatment options 2019.
17. Society AC. Chemotherapy for oral cavity and oropharyngeal cancer, how is chemo used to treat oral cavity and oropharyngeal cancers? 2019.
- 18.

- Song S. General principles of radiation therapy for head and neck cancer. *Up to Date, last update in November 2012*;26.
19. Scorsetti M, Comito T, Clerici E, Franzese C, Tozzi A, Iftode C, Di Brina L, Navarria P, Mancosu P, Reggiori G. Phase ii trial on sbrt for unresectable liver metastases: Long-term outcome and prognostic factors of survival after 5 years of follow-up. *Radiation Oncology* 2018;13:234.
 20. Dawson LA, Eccles C, Craig T. Individualized image guided iso-ntcp based liver cancer sbrt. *Acta Oncologica* 2006;45:856-864.
 21. Kwon JH, Bae SH, Kim JY, Choi BO, Jang HS, Jang JW, Choi JY, Yoon SK, Chung KW. Long-term effect of stereotactic body radiation therapy for primary hepatocellular carcinoma ineligible for local ablation therapy or surgical resection. Stereotactic radiotherapy for liver cancer. *BMC cancer* 2010;10:475.
 22. Kavanagh BD, Pan CC, Dawson LA, Das SK, Li XA, Ten Haken RK, Miften M. Radiation dose–volume effects in the stomach and small bowel. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S101-S107.
 23. Lujan AE, Balter JM, Ten Haken RK. A method for incorporating organ motion due to breathing into 3d dose calculations in the liver: Sensitivity to variations in motion. *Medical physics* 2003;30:2643-2649.
 24. Kumagai M, Hara R, Mori S, Yanagi T, Asakura H, Kishimoto R, Kato H, Yamada S, Kandatsu S, Kamada T. Impact of intrafractional bowel gas movement on carbon ion beam dose distribution in pancreatic radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* 2009;73:1276-1281.
 25. Velec M, Moseley JL, Craig T, Dawson LA, Brock KK. Accumulated dose in liver stereotactic body radiotherapy: Positioning, breathing, and deformation effects. *International Journal of Radiation Oncology* Biology* Physics* 2012;83:1132-1140.
 26. Hsiung C-Y, Hunt MA, Yorke ED, Chui C-S, Hu J, Xiong J-P, Ling CC, Lo S-K, Wang C-J, Huang E-Y. Intensity-modulated radiotherapy as the boost or salvage treatment of nasopharyngeal carcinoma: The appropriate parameters in the inverse planning and the effect of patient's anatomic factors on the planning results. *Radiotherapy and oncology* 2005;77:53-57.
 27. Hunt MA, Zelefsky MJ, Wolden S, Chui C-S, LoSasso T, Rosenzweig K, Chong L, Spirou SV, Fromme L, Lumley M. Treatment planning and delivery of intensity-modulated radiation therapy for primary nasopharynx cancer. *International Journal of Radiation Oncology* Biology* Physics* 2001;49:623-632.
 28. Lee N, Xia P, Quivey JM, Sultanem K, Poon I, Akazawa C, Akazawa P, Weinberg V, Fu KK. Intensity-modulated radiotherapy in the treatment of nasopharyngeal carcinoma: An update of the ucsf experience. *International Journal of Radiation Oncology* Biology* Physics* 2002;53:12-22.
 29. Wolden SL, Chen WC, Pfister DG, Kraus DH, Berry SL, Zelefsky MJ. Intensity-modulated radiation therapy (imrt) for nasopharynx cancer: Update of the memorial sloan-kettering experience. *International Journal of Radiation Oncology* Biology* Physics* 2006;64:57-62.
 30. 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. *Radiotherapy and Oncology* 2013;106:85-89.

31. Barker Jr JL, Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, Morrison WH, Rosenthal DI, Chao KC, Tucker SL. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated ct/linear accelerator system. *International Journal of Radiation Oncology* Biology* Physics* 2004;59:960-970.
32. Deasy JO, Moiseenko V, Marks L, Chao KC, Nam J, Eisbruch A. Radiotherapy dose–volume effects on salivary gland function. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S58-S63.
33. Baumann M, Petersen C. Tc_p and ntc_p: A basic introduction. *Rays* 2005;30:99-104.
34. Bentzen SM, Constine LS, Deasy JO, Eisbruch A, Jackson A, Marks LB, Ten Haken RK, Yorke ED. Quantitative analyses of normal tissue effects in the clinic (quantec): An introduction to the scientific issues. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S3-S9.
35. Lyman JT. Complication probability as assessed from dose-volume histograms. *Radiation Research* 1985;104:S13-S19.
36. Wong O, McNiven A, Chan B, Moseley J, Lee J, Le L, Ren C, Waldron J, Bissonnette J, Giuliani M. Evaluation of differences between estimated delivered dose and planned dose in nasopharynx patients using deformable image registration and dose accumulation. *International Journal of Radiation Oncology* Biology* Physics* 2017;99:E382.
37. Reese A, Hou J, Suntharalingam M, D'Souza W. Predicting differences in planned and delivered dose for patients with head and neck cancer treated with intensity modulated radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* 2011;81:S800.
38. Ishikura S, Tobinai K, Ohtsu A, Nakamura S, Yoshino T, Oda I, Takagi T, Mera K, Kagami Y, Itoh K. Japanese multicenter phase ii study of chop followed by radiotherapy in stage i–ii1, diffuse large b-cell lymphoma of the stomach. *Cancer science* 2005;96:349-352.
39. Sykes AJ, Kiltie AE, Stewart AL. Ondansetron versus a chlorpromazine and dexamethasone combination for the prevention of nausea and vomiting: A prospective, randomised study to assess efficacy, cost effectiveness and quality of life following single-fraction radiotherapy. *Supportive care in cancer* 1997;5:500-503.
40. Keys HM, Bundy BN, Stehman FB, Muderspach LI, Chafe WE, Suggs CL, Walker JL, Gersell D. Cisplatin, radiation, and adjuvant hysterectomy compared with radiation and adjuvant hysterectomy for bulky stage ib cervical carcinoma. *New England Journal of Medicine* 1999;340:1154-1161.
41. Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tomé WA. Accurate accumulation of dose for improved understanding of radiation effects in normal tissue. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S135-S139.
42. Kim J, Jung Y. Radiation-induced liver disease: Current understanding and future perspectives. *Experimental & molecular medicine* 2017;49:e359.
43. Lawrence TS, Robertson JM, Anscher MS, Jirtle RL, Ensminger WD, Fajardo LF. Hepatic toxicity resulting from cancer treatment. *International Journal of Radiation Oncology* Biology* Physics* 1995;31:1237-1248.
44. Pan CC, Kavanagh BD, Dawson LA, Li XA, Das SK, Miften M, Ten Haken RK. Radiation-associated liver injury. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S94-S100.

45. Russell AH, Clyde C, Wasserman TH, Turner SS, Rotman M. Accelerated hyperfractionated hepatic irradiation in the management of patients with liver metastases: Results of the rtog dose escalating protocol. *International Journal of Radiation Oncology* Biology* Physics* 1993;27:117-123.
46. Dawson LA, Normolle D, Balter JM, McGinn CJ, Lawrence TS, Ten Haken RK. Analysis of radiation-induced liver disease using the lyman ntcp model. *International Journal of Radiation Oncology* Biology* Physics* 2002;53:810-821.
47. Trotti A. Toxicity in head and neck cancer: A review of trends and issues. *International Journal of Radiation Oncology* Biology* Physics* 2000;47:1-12.
48. Dirix P, Nuyts S, Van den Bogaert W. Radiation-induced xerostomia in patients with head and neck cancer: A literature review. *Cancer: Interdisciplinary International Journal of the American Cancer Society* 2006;107:2525-2534.
49. Eisbruch A, Ship JA, Martel MK, Ten Haken RK, Marsh LH, Wolf GT, Esclamado RM, Bradford CR, Terrell JE, Gebarski SS. Parotid gland sparing in patients undergoing bilateral head and neck irradiation: Techniques and early results. *International Journal of Radiation Oncology* Biology* Physics* 1996;36:469-480.
50. Münter MW, Karger CP, Hoffner SG, Hof H, Thilmann C, Rudat V, Nill S, Wannemacher M, Debus J. Evaluation of salivary gland function after treatment of head-and-neck tumors with intensity-modulated radiotherapy by quantitative pertechnetate scintigraphy. *International Journal of Radiation Oncology* Biology* Physics* 2004;58:175-184.
51. Parliament MB, Scrimger RA, Anderson SG, Kurien EC, Thompson HK, Field GC, Hanson J. Preservation of oral health-related quality of life and salivary flow rates after inverse-planned intensity-modulated radiotherapy (imrt) for head-and-neck cancer. *International Journal of Radiation Oncology* Biology* Physics* 2004;58:663-673.
52. Saarilahti K, Kouri M, Collan J, Hämäläinen T, Atula T, Joensuu H, Tenhunen M. Intensity modulated radiotherapy for head and neck cancer: Evidence for preserved salivary gland function. *Radiotherapy and oncology* 2005;74:251-258.
53. Scrimger R, Kanji A, Parliament M, Warkentin H, Field C, Jha N, Hanson J. Correlation between saliva production and quality of life measurements in head and neck cancer patients treated with intensity-modulated radiotherapy. *American journal of clinical oncology* 2007;30:271-277.
54. Eisbruch A, Harris J, Garden AS, Chao CK, Straube W, Harari PM, Sanguineti G, Jones CU, Bosch WR, Ang KK. Multi-institutional trial of accelerated hypofractionated intensity-modulated radiation therapy for early-stage oropharyngeal cancer (rtog 00-22). *International Journal of Radiation Oncology* Biology* Physics* 2010;76:1333-1338.
55. Lee N, Harris J, Garden AS, Straube W, Glisson B, Xia P, Bosch W, Morrison WH, Quivey J, Thorstad W. Intensity-modulated radiation therapy with or without chemotherapy for nasopharyngeal carcinoma: Radiation therapy oncology group phase ii trial 0225. *Journal of clinical oncology* 2009;27:3684.
56. Tham IW-K, Hee SW, Yeo RM-C, Salleh PB, Lee J, Tan TW-K, Fong KW, Chua ET, Wee JT-S. Treatment of nasopharyngeal carcinoma using intensity-modulated radiotherapy—the national cancer centre singapore experience. *International Journal of Radiation Oncology* Biology* Physics* 2009;75:1481-1486.
57. Lin S, Pan J, Han L, Zhang X, Liao X, Lu JJ. Nasopharyngeal carcinoma treated with reduced-volume intensity-modulated radiation therapy: Report on the 3-year outcome of

- a prospective series. *International Journal of Radiation Oncology* Biology* Physics* 2009;75:1071-1078.
58. Lin S, Lu JJ, Han L, Chen Q, Pan J. Sequential chemotherapy and intensity-modulated radiation therapy in the management of locoregionally advanced nasopharyngeal carcinoma: Experience of 370 consecutive cases. *BMC cancer* 2010;10:39.
 59. Navazesh M, Kumar SK. Measuring salivary flow: Challenges and opportunities. *The Journal of the American Dental Association* 2008;139:35S-40S.
 60. ctep.cancer.gov. Common terminology criteria for adverse events v3.0 (ctcae) August 9, 2006.
 61. Lujan AE. Incorporation of uncertainties from setup errors and organ motion in 3-d dose calculations. 1999.
 62. Brock KK. The influence of liver deformation on 4-d dose calculations for radiotherapy 2003.
 63. Brock K, Sharpe M, Dawson L, Kim S, Jaffray D. Accuracy of finite element model-based multi-organ deformable image registration. *Medical physics* 2005;32:1647-1659.
 64. Velec M. Deformable dose reconstruction to optimize the planning and delivery of liver cancer radiotherapy. In: Editor, editor^editors. Book Deformable dose reconstruction to optimize the planning and delivery of liver cancer radiotherapy; 2014.
 65. Geets X, Tomsej M, Lee JA, Duprez T, Coche E, Cosnard G, Lonneux M, Grégoire V. 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. *Radiotherapy and Oncology* 2007;85:105-115.
 66. Han C, Chen Y-J, Liu A, Schultheiss TE, Wong JY. Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* 2008;70:1256-1262.
 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. *International Journal of Radiation Oncology* Biology* Physics* 2006;64:355-362.
 68. 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. *International Journal of Radiation Oncology* Biology* Physics* 2007;68:1121-1130.
 69. Osorio EMV, 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. *International Journal of Radiation Oncology* Biology* Physics* 2008;70:875-882.
 70. Castadot P, Lee JA, Geets X, Grégoire V. Adaptive radiotherapy of head and neck cancer. *Seminars in radiation oncology*. Elsevier. 2010. pp. 84-93.
 71. O'Daniel JC, Garden AS, Schwartz DL, Wang H, Ang KK, Ahamad A, Rosenthal DI, Morrison WH, Asper JA, Zhang L. Parotid gland dose in intensity-modulated radiotherapy for head and neck cancer: Is what you plan what you get? *International Journal of Radiation Oncology* Biology* Physics* 2007;69:1290-1296.
 72. Lee C, Langen KM, Lu W, Haimerl J, Schnarr E, Ruchala KJ, Olivera GH, Meeks SL, Kupelian PA, Shellenberger TD. Assessment of parotid gland dose changes during head and neck cancer radiotherapy using daily megavoltage computed tomography and

- deformable image registration. *International Journal of Radiation Oncology* Biology* Physics* 2008;71:1563-1571.
73. Pan C, Eisbruch A, Ten Haken RK. Physical models and simpler dosimetric descriptors of radiation late toxicity. *Seminars in radiation oncology*. Elsevier. 2007. pp. 108-120.
 74. Konings AW, Cotteleer F, Faber H, van Luijk P, Meertens H, Coppes RP. Volume effects and region-dependent radiosensitivity of the parotid gland. *Int J Radiat Oncol Biol Phys* 2005;62:1090-1095.
 75. Society AC. Lifetime risk of developing or dying from cancer In: Editor, editor^editors. Book Lifetime risk of developing or dying from cancer 2018.
 76. **ICRU**. ICRU report 83 prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (imrt)-journal of the icru-vol 10 no 1 2010: Oxford University Press; 2010.
 77. van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: Dose-population histograms for deriving treatment margins in radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* 2000;47:1121-1135.
 78. Aristophanous M, Suh Y, Chi PC, Whittlesey LJ, LaNeave S, Martel MK. Initial clinical experience with arccheck for imrt/vmat qa. *Journal of applied clinical medical physics* 2016;17:20-33.
 79. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Medical physics* 1998;25:656-661.
 80. Low DA, Dempsey JF. Evaluation of the gamma dose distribution comparison method. *Medical physics* 2003;30:2455-2464.
 81. Reynaert N, Van der Marck S, Schaart D, Van der Zee W, Tomsej M, van Vliet-Vroegindewij C, Jansen J, Coghe M, De Wagter C, Heijmen B. Monte carlo treatment planning: An introduction. *Report NCS-16 (Netherlands Commission on Radiation Dosimetry, Delft)* 2006.
 82. Chen W-Z, Xiao Y, Li J. Impact of dose calculation algorithm on radiation therapy. *World journal of radiology* 2014;6:874.
 83. Sontag MR, Cunningham JR. The equivalent tissue-air ratio method for making absorbed dose calculations in a heterogeneous medium. *Radiology* 1978;129:787-794.
 84. Milan J, Bentley R. The storage and manipulation of radiation dose data in a small digital computer. *The British journal of radiology* 1974;47:115-121.
 85. Ahnesjö A, Andreo P, Brahme A. Calculation and application of point spread functions for treatment planning with high energy photon beams. *Acta Oncologica* 1987;26:49-56.
 86. Krieger T, Sauer OA. Monte carlo-versus pencil-beam-/collapsed-cone-dose calculation in a heterogeneous multi-layer phantom. *Physics in Medicine & Biology* 2005;50:859.
 87. Hasenbalg F, Neuenschwander H, Mini R, Born EJ. Collapsed cone convolution and analytical anisotropic algorithm dose calculations compared to vmc++ monte carlo simulations in clinical cases. *Physics in Medicine & Biology* 2007;52:3679.
 88. Knöös T, Wieslander E, Cozzi L, Brink C, Fogliata A, Albers D, Nyström H, Lassen S. Comparison of dose calculation algorithms for treatment planning in external photon beam therapy for clinical situations. *Physics in Medicine & Biology* 2006;51:5785.
 89. Boudreau C, Heath E, Seuntjens J, Ballivy O, Parker W. Imrt head and neck treatment planning with a commercially available monte carlo based planning system. *Physics in Medicine & Biology* 2005;50:879.

90. Kim SJ, Kim SK, Kim DH. Comparison of pencil-beam, collapsed-cone and monte-carlo algorithms in radiotherapy treatment planning for 6-mv photons. *Journal of the Korean Physical Society* 2015;67:153-158.
91. B Rosen CL, K Brock The use of cbct for dose calculations in h/n patients: Impact of limited fov. In: Editor, editor^editors. Book The use of cbct for dose calculations in h/n patients: Impact of limited fov. Washington, DC; 2016.
92. Heukelom J, Fuller C, Kantor M, Kauwelo K, Rasch C, Sonke J. Po-0906: Ntcp differences between planned and delivered dose in treatment for head and neck cancer. *Radiotherapy and Oncology* 2016;119:S436-S437.
93. Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: Application to breast mr images. *Medical Imaging, IEEE Transactions on* 1999;18:712-721.
94. Thirion J-P. Image matching as a diffusion process: An analogy with maxwell's demons. *Medical image analysis* 1998;2:243-260.
95. Svedlow M, McGillem C, Anuta PE. Experimental examination of similarity measures and preprocessing methods used for image registration. LARS Symposia. 1976. pp. 150.
96. Weistrand O, Svensson S. The anaconda algorithm for deformable image registration in radiotherapy. *Medical physics* 2015;42:40-53.
97. Polan DF, Feng M, Lawrence TS, Ten Haken RK, Brock KK. Implementing radiation dose-volume liver response in biomechanical deformable image registration. *International Journal of Radiation Oncology* Biology* Physics* 2017.
98. Velec M, Moseley JL, Svensson S, Hårdemark B, Jaffray DA, Brock KK. Validation of biomechanical deformable image registration in the abdomen, thorax and pelvis in a commercial radiotherapy treatment planning system. *Medical physics* 2017.
99. Cazoulat G, Owen D, Matuszak MM, Balter JM, Brock KK. Biomechanical deformable image registration of longitudinal lung ct images using vessel information. *Physics in medicine and biology* 2016;61:4826-4839.
100. Yan D, Jaffray D, Wong J. A model to accumulate fractionated dose in a deforming organ. *International Journal of Radiation Oncology* Biology* Physics* 1999;44:665-675.
101. Wang H, Dong L, O'Daniel J, Mohan R, Garden AS, Ang KK, Kuban DA, Bonnen M, Chang JY, Cheung R. Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. *Physics in medicine and biology* 2005;50:2887.
102. Kostelec PJ, Weaver JB, Healy DM. Multiresolution elastic image registration. *Medical physics* 1998;25:1593-1604.
103. Scorsetti M, Clerici E, Comito T. Stereotactic body radiation therapy for liver metastases. *Journal of gastrointestinal oncology* 2014;5:190.
104. Nouhaud E, Créhange G, Cuffe A, Quivrin M, Rouffiac-Thouant M, Mineur L, Garcia R, Chamois J, Maingon P. Stereotactic body radiation therapy for liver tumors with or without rotational intensity modulated radiation therapy. *BMC research notes* 2013;6:492.
105. Thomas TO, Hasan S, Small Jr W, Herman JM, Lock M, Kim EY, Mayr NA, Teh BS, Lo SS. The tolerance of gastrointestinal organs to stereotactic body radiation therapy: What do we know so far? *Journal of gastrointestinal oncology* 2014;5:236.
106. Bujold A, Massey CA, Kim JJ, Brierley J, Cho C, Wong RK, Dinniwell RE, Kassam Z, Ringash J, Cummings B. Sequential phase i and ii trials of stereotactic body radiotherapy

- for locally advanced hepatocellular carcinoma. *Journal of Clinical Oncology* 2013;31:1631-1639.
107. Feng M, Normolle D, Pan CC, Dawson LA, Amarnath S, Ensminger WD, Lawrence TS, Ten Haken RK. Dosimetric analysis of radiation-induced gastric bleeding. *International Journal of Radiation Oncology* Biology* Physics* 2012;84:e1-e6.
 108. Murphy JD, Christman-Skieller C, Kim J, Dieterich S, Chang DT, Koong AC. A dosimetric model of duodenal toxicity after stereotactic body radiotherapy for pancreatic cancer. *International Journal of Radiation Oncology* Biology* Physics* 2010;78:1420-1426.
 109. Dean JA, Wong KH, Welsh LC, Jones A-B, Schick U, Newbold KL, Bhide SA, Harrington KJ, Nutting CM, Gulliford SL. 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. *Radiotherapy and Oncology* 2016;120:21-27.
 110. Holyoake DL, Aznar M, Mukherjee S, Partridge M, Hawkins MA. Modelling duodenum radiotherapy toxicity using cohort dose-volume-histogram data. *Radiotherapy and Oncology* 2017;123:431-437.
 111. Xu C-J, van der Schaaf A, Van't Veld AA, Langendijk JA, Schilstra C. Statistical validation of normal tissue complication probability models. *International Journal of Radiation Oncology* Biology* Physics* 2012;84:e123-e129.
 112. Dawson L, Balter J, McGinn C, Lawrence T, Ten Haken R. Description of radiation induced liver disease using the Lyman and local damage-organ injury ntcp models. *International Journal of Radiation Oncology* Biology* Physics* 2001;51:33-34.
 113. Scaife JE, Thomas SJ, Harrison K, Romanchikova M, Sutcliffe MP, Forman JR, Bates AM, Jena R, Parker MA, Burnet NG. Accumulated dose to the rectum, measured using dose-volume histograms and dose-surface maps, is different from planned dose in all patients treated with radiotherapy for prostate cancer. *The British journal of radiology* 2015;88:20150243.
 114. Swaminath A, Massey C, Brierley JD, Dinniwell R, Wong R, Kim JJ, Velec M, Brock KK, Dawson LA. Accumulated delivered dose response of stereotactic body radiation therapy for liver metastases. *International Journal of Radiation Oncology* Biology* Physics* 2015;93:639-648.
 115. Yorke ED. Parameters for and use of ntcp models in the clinic. th Annual AAPM Meeting, Minneapolis, Minnesota. 2007.
 116. Michalski JM, Gay H, Jackson A, Tucker SL, Deasy JO. Radiation dose-volume effects in radiation-induced rectal injury. *International Journal of Radiation Oncology* Biology* Physics* 2010;76:S123-S129.
 117. Lee MT, Kim JJ, Dinniwell R, Brierley J, Lockwood G, Wong R, Cummings B, Ringash J, Tse RV, Knox JJ. Phase I study of individualized stereotactic body radiotherapy of liver metastases. *Journal of clinical oncology* 2009;27:1585-1591.
 118. Regina VT, Hawkins M, Lockwood G, Kim JJ, Cummings B, Knox J, Sherman M, Dawson LA. Phase I study of individualized stereotactic body radiotherapy for hepatocellular carcinoma and intrahepatic cholangiocarcinoma. *J Clin Oncol* 2008;26:657-664.
 119. Oncology N. Nrg oncology protocols. In: Editor, editor^editors. Book Nrg oncology protocols; 2018.
 120. Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of human papillomavirus-positive head and neck squamous cell carcinoma. *Journal of Clinical Oncology* 2015;33:3235.

121. Salama JK, Haddad RI, Kies MS, Busse PM, Dong L, Brizel DM, Eisbruch A, Tishler RB, Trotti AM, Garden AS. Clinical practice guidance for radiotherapy planning after induction chemotherapy in locoregionally advanced head-and-neck cancer. *International Journal of Radiation Oncology* Biology* Physics* 2009;75:725-733.
122. Risholm P, Balter J, Wells WM. Estimation of delivered dose in radiotherapy: The influence of registration uncertainty. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2011. pp. 548-555.
123. Vineberg K, Hunter K, Schipper M, McShan D, Antonuk A, Feng M, Balter J, Eisbruch A. The influence of rotation on delivered dose in oropharyngeal cancer. *International Journal of Radiation Oncology* Biology* Physics* 2012;84:S81.
124. Orlandini LC, Coppola M, Fulcheri C, Cernusco L, Wang P, Cionini L. Dose tracking assessment for image-guided radiotherapy of the prostate bed and the impact on clinical workflow. *Radiation Oncology* 2017;12:78.
125. Hunter KU, Fernandes LL, Vineberg KA, McShan D, Antonuk AE, Cornwall C, Feng M, Schipper MJ, Balter JM, Eisbruch A. 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;87:676-682.
126. Purdy JA. Current icru definitions of volumes: Limitations and future directions. *Seminars in radiation oncology*. Elsevier. 2004. pp. 27-40.
127. Geng H, Giaddui T, Zhong H, Rosenthal D, Galvin J, Xiao Y, Linnemann N. Knowledge engineering-based quality evaluation of nrg oncology rtog 0522 treatment plans. *International Journal of Radiation Oncology* Biology* Physics* 2017;99:S174.
128. Braam PM, Roesink JM, Moerland MA, Raaijmakers CP, Schipper M, Terhaard CH. Long-term parotid gland function after radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* 2005;62:659-664.
129. Zhang P, Simon A, Rigaud B, Castelli J, Arango JDO, Nassef M, Henry O, Zhu J, Haigron P, Li B. Optimal adaptive imrt strategy to spare the parotid glands in oropharyngeal cancer. *Radiotherapy and Oncology* 2016;120:41-47.
130. Samavati N, Velec M, Brock KK. Effect of deformable registration uncertainty on lung sbrt dose accumulation. *Medical physics* 2016;43:233-240.
131. Vainshtein JM, Moon DH, Feng FY, Chepeha DB, Eisbruch A, Stenmark MH. Long-term quality of life after swallowing and salivary-sparing chemo-intensity modulated radiation therapy in survivors of human papillomavirus-related oropharyngeal cancer. *International Journal of Radiation Oncology* Biology* Physics* 2015;91:925-933.
132. Nix MG, Prestwich RJ, Speight R. Automated, reference-free local error assessment of multimodal deformable image registration for radiotherapy in the head and neck. *Radiotherapy and Oncology* 2017;125:478-484.
133. Kierkels RG, den Otter LA, Korevaar EW, Langendijk JA, van der Schaaf A, Knopf A-C, Sijtsma NM. An automated, quantitative, and case-specific evaluation of deformable image registration in computed tomography images. *Physics in medicine and biology* 2017.
134. Liu C, Kim J, Kumarasiri A, Mayyas E, Brown SL, Wen N, Siddiqui F, Chetty IJ. An automated dose tracking system for adaptive radiation therapy. *Computer methods and programs in biomedicine* 2018;154:1-8.
135. Loi G, Fusella M, Lanzi E, Cagni E, Garibaldi C, Iacoviello G, Lucio F, Menghi E, Miceli R, Orlandini LC. Performance of commercially available deformable image registration

- platforms for contour propagation using patient-based computational phantoms: A multi-institutional study. *Medical physics* 2017.
136. Al-Mayah A, Moseley J, Hunter S, Brock K. Radiation dose response simulation for biomechanical-based deformable image registration of head and neck cancer treatment. *Physics in Medicine & Biology* 2015;60:8481.
 137. Al-Mayah A, Moseley J, Hunter S, Velec M, Chau L, Breen S, Brock K. Biomechanical-based image registration for head and neck radiation treatment. *Physics in medicine and biology* 2010;55:6491.
 138. Nelms BE, Tomé WA, Robinson G, Wheeler J. Variations in the contouring of organs at risk: Test case from a patient with oropharyngeal cancer. *International Journal of Radiation Oncology* Biology* Physics* 2012;82:368-378.
 139. Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, Dicker DJ, Chimed-Orchir O, Dandona R, Dandona L. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: A systematic analysis for the global burden of disease study. *JAMA oncology* 2017;3:524-548.
 140. Delli K, Spijkervet FK, Kroese FG, Bootsma H, Vissink A. Xerostomia Saliva: Secretion and functions: Karger Publishers; 2014. pp. 109-125.
 141. Lee H, Havrila C, Bravo V, Shantz K, Diaz K, Larner J, Read P. Effect of oral nutritional supplementation on weight loss and percutaneous endoscopic gastrostomy tube rates in patients treated with radiotherapy for oropharyngeal carcinoma. *Supportive care in cancer* 2008;16:285-289.
 142. Sharpe M, Brock K, Rehbinder H, Forsgren C, Lundin A, Dawson L, Studer G, O'Sullivan B, McNutt T, Kaus M. Adaptive planning and delivery to account for anatomical changes induced by radiation therapy of head and neck cancer. *International Journal of Radiation Oncology* Biology* Physics* 2005;63:S3.
 143. Blanco AI, Chao KC, El Naqa I, Franklin GE, Zakarian K, Vicic M, Deasy JO. Dose-volume modeling of salivary function in patients with head-and-neck cancer receiving radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* 2005;62:1055-1069.
 144. Mencarelli A, van Kranen SR, Hamming-Vrieze O, van Beek S, Rasch CRN, van Herk M, Sonke J-J. Deformable image registration for adaptive radiation therapy of head and neck cancer: Accuracy and precision in the presence of tumor changes. *International Journal of Radiation Oncology* Biology* Physics* 2014;90:680-687.
 145. Li X, Zhang Y, Shi Y, Wu S, Xiao Y, Gu X, Zhen X, Zhou L. 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.
 146. Lou J, Huang P, Ma C, Zheng Y, Chen J, Liang Y, Li H, Yin Y, Liu D, Yu G. Parotid gland radiation dose-xerostomia relationships based on actual delivered dose for nasopharyngeal carcinoma. *Journal of applied clinical medical physics* 2018;19:251-260.
 147. Wang ZH, Yan C, Zhang ZY, Zhang CP, Hu HS, Kirwan J, Mendenhall WM. Radiation-induced volume changes in parotid and submandibular glands in patients with head and neck cancer receiving postoperative radiotherapy: A longitudinal study. *Laryngoscope* 2009;119:1966-1974.