Fractionated Treatment Planning of Radiation Therapy Considering Biological Response

A Dissertation

Presented to

the Faculty of the Department of Industrial Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

in Industrial Engineering

by

Nasrin Nouri

May 2019

Fractionated Treatment Planning of Radiation Therapy Considering Biological Response

Nasrin Nouri

Approved:

Chair of the Committee
Gino Lim, Professor,
Industrial Engineering

Committee Members:

Jiming Peng, Associate Professor,
Industrial Engineering

Taewoo Lee, Assistant Professor,
Industrial Engineering

Cumaraswamy Vipulanandan, Professor,
Civill and Environmental Engineering

Navin Varadarajan, Associate Professor,
Chemical and Biomolecular Engineering

Suresh K. Khator, Associate Dean
Cullen College of Engineering

Gino Lim, Professor and Chair,
Industrial Engineering

To my loving husband, Vahid Hejazi,

for his encouragement, care, support, and patience,

and to my parents, Nourieh and Ghasem,

who have never left my side throughout these years,

and to my brother and sister-in-law, Amin and Somaye,

for always believing in me.

Acknowledgements

I would like to formally thank:

Dr. Gino J. Lim, my research advisor, for his support, excellent guidance, and great advice during past four years. I have learned so much during the course of this PhD. Without his guidance and persistent help this dissertation would not have been possible.

Dr. Jiming Peng, Dr. Taewoo Lee, Dr. Cumaraswamy Vipulanandan, and Dr. Navin Varadarajan for serving on my dissertation committee. They really helped me with their valuable insight and guidance. My sincere thanks to Dr. Taewoo Lee for his gracious support.

Dr. Wenhua Cao, Dr. Xaiodong Zhang and Dr. Li Liao for their inspiring suggestions and kind help.

The staff of Industrial Engineering department, Mr Philip Abraham, Ms Sharon Hall, Ms Cyrena Edwards and Ms Andreeka Lewis, for their patience and kind support.

Fractionated Treatment Planning of Radiation Therapy Considering Biological Response

An Abstract

of a

Dissertation

Presented to

the Faculty of the Department of Industrial Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

in Industrial Engineering

by

Nasrin Nouri

2019

Abstract

The goal of radiation therapy for cancer patients is to kill tumor cells by damaging their DNA. For the majority of patients, the prescribed dose is divided into several treatment sessions (fractionated treatment plan) to avoid lethal damage to the surrounding healthy organs called organs at risk (OARs). In conventional practice, the treatment policy is to deliver an equal amount of radiation dose to the patient over multiple treatment sessions. Such an approach neglects different uncertainties associated with tumor dynamics, biological response to radiation, and organ motion that occur during radiation treatment. In this dissertation we are proposing methods to tackle the current challenges and shortcomings in radiotherapy treatment planning.

In the first part of this dissertation, a constrained Partially Observable Markov Decision Process (POMDP) approach is proposed based on an extended biological model of cell survival to incorporate the biological response from the patient in the fractionated radiotherapy plan. A Gompertzian growth function is used to explain dependence of tumor growth rate on its density and shape. The aim of our model is to maximize the expected biological equivalent dose (EBED) of tumor, while keeping the OAR survival under control. Because the condition of a tumor can change and it is not fully observable through CT images during the treatment horizon, POMDP enables us to consider the tumor symptoms through probabilistic belief and partial observation probabilities. We provide a control limit policy to investigate whether there is an advantage of using POMDP over the conventional plan in terms of tumor damage and OAR sparing. Numerical results showed potential impact of the POMDP policies to enhance tumor coverage compared to the conventional plan. The resulting policies suggested the use of a low dose at earlier sessions, and a higher dose at later sessions. This result reflects the impact of tumor density and shape on its growth and biological response. The POMDP policy was not recommended if the tumor was a late responding tissue and its corresponding OAR was an early responding tissue.

Unlike photon, the proton's linear energy transfer (LET) increases as it penetrates the

body. Therefore, proton therapy can be modulated to provide a better biological effectiveness. In the second part of this dissertation, we develop an LET-based IMPT optimization model that guarantees homogeneous biological effectiveness on the tumor structure and minimum damage to the OARs. The outcomes of this model serve as the action set in a constrained MDP framework developed to provide an optimal decision-making policy for dynamic and personalized fractionated proton therapy treatment plans. The tumor state is predicted using a random forest classification model built on radiomics data from CT images. The proposed model is implemented on the two cases of prostate cancer and pediatric ependymoma and compared to a regular IMPT model as the threshold. The results demonstrate that the LET-based IMPT model improves biological effectiveness and tumor control probability (TCP). Randomized MDP policies suggest a smaller dose target for a high tumor cell count where the tumor growth rate is at its lowest value. But as the tumor cell count decreases, a larger amount of dose is suggested to destroy faster growing tumor.

Proton's unique physical characteristics make proton therapy sensitive to organ motion such that a voxel can receive a nonuniform dose deposition between different fractions. Therefore, biological effectiveness of the treatment might deviate from the planned effectiveness. In the last part of this dissertation we develop a model to optimize the fractionation and IMPT problems at the same time. We use 4DCT data set for planning a 3D delivery technique to handle complex respiratory motion patterns while avoiding sophisticated 4D delivery systems. Two models are used to solve this problem; a statistical mean-variance model, and a robust worst-case model. The worst-case robust model provides a more robust dose distribution over all structures compared to statistical mean-variance model. Both models suggest larger amount of radiation dose in the first week of treatment and gradually decreasing the dose towards the last week. The resulting weekly mean BED is shown to be almost equal in all treatment weeks, compensating for the increased repair effect resulting from nonuniform voxel dose between fractions. Because of conservatism of worst-case robust model, a larger total dose has to be delivered in every treatment week to achieve the same biological effectiveness as statistical mean-variance model.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Tumor is a mass of cells with abnormal growth that do not contribute to the body functioning. A tumor is cancerous when it has the ability of invading the neighboring tissues or spreading to other body parts. Cancer is a class of over 100 diseases that affect humans and is the second leading cause of death worldwide. The American Cancer Society has estimated that over 1.7 million new cancer cases will be diagnosed in 2018 and the number of new cases are expected to rise by about 70% over the next two decades (*American Cancer Society Cancer Facts & Figures*, 2018).

There is no single cure for cancer, but there are three major treatments for different cancer types; surgery, chemotherapy and radiation therapy. The ultimate goal of the treatment is complete removal of cancer without damaging the healthy organs. For the isolated and solid tumors that do not metastasize surgery is the primary method of treatment. Since this is not the case for most of the cancers, surgery is also used for prognosis and staging before other methods of treatment. Chemotherapy uses specific drugs to affect rapidly dividing cells which are not limited to cancer cells only. Therefore, it has the potential to harm the healthy tissue as well. Radiation therapy uses ionizing radiation to kill cells by damaging the DNA which stops them from growing and dividing. Radiation therapy makes damages to the healthy tissue too, but the harm is confined to the treatment region in contrast with chemotherapy. About 60 percent of cancer patients in the US receive curative radiation

therapy despite many advances in other treatment modalities.

There are two types of radiation therapy based on the source of radiation; external beam radiation therapy is a type of treatment in which the radiation is directed to the tumor from outside the patient's body. In brachytherapy however, a sealed or unsealed source of radiation is placed in the organ under treatment or very close to it. Many studies have been dedicated to both brachytherapy (Rivard, Venselaar & Beaulieu, 2009; Ferrer et al., 2008) and external beam radiation therapy in terms of treatment planning, but the focus of this study is on external beam radiotherapy planning.

Any method aiming to treat a tumor requires a precise knowledge of the tumor volume. Therefore, initiating the treatment planning requires digital imaging of the tumor and internal anatomy of the patient. Several digital imaging modalities such as computed tomography (CT), magnet resonance imaging (MRI), or position emission tomography (PET) can be used to view and locate the tumor. Since it is necessary to differentiate the tumor from healthy organs, a physician contours the tumor in the digital images. After identifying the tumor structure, the healthy organs that fall in the treatment site and might be affected by radiation are defined as *organs at risk* (OAR), and the rest on healthy organs are defined as normal tissue.

A radiation oncologist defines the prescription dose based on cancer type and stage as well as the treatment types the patient is receiving. The radiation dose is measured in gray (Gy) and mostly varies from 60 to 80 Gy. Also, the radiation oncologist defines the dose-volume constraint (DVC) for OARs to spare them from lethal damage. For the majority of patients, the radiation dose required to eradicate the tumor is higher than the tolerance of OARs. For such cases, the prescribed dose is typically divided into several treatment sessions called fractionated treatment plan. Fractionation not only results in smaller dose tolerable to OARs, but also provides the OARs enough time to recover from the sublethal damage. The ideal process for treatment planning is to examine the patient right before each treatment session and update tumor contours indicating the shape and size of the tumor, and make an appropriate decision about radiation dose based on tumor response to treatment. However,

this process is not practical because updating tumor contours and finding corresponding optimal radiation dose may not be an option or it can be time consuming and prone to human errors. In current practice, the treatment policy is to deliver an equal amount of radiation dose in all treatment sessions for most cancer patients. But there are some unanswered questions in this point: Is the equal dose per fraction an efficient treatment plan? Can we apply the same treatment plan to all patients with different characteristics? What internal and external factors affect the treatment plan quality and efficiency?

1.2 Problem Description

The purpose of this thesis is to address some of the unanswered challenges in radiation therapy treatment planning aiming to improve the quality of treatment delivered to cancer patients. The followings are the challenges we try to resolve:

1.2.1 Treatment planning considering biological response of the tumor

There are two types of tissues based on intrinsic response to radiation therapy; the *early responding tissues* in which cells divide faster and have higher intrinsic radiosensitivity, and the *late responding tissues* with slowly dividing cells and lower intrinsic radiosensitivity. Aside from tissue type, the biological response of the cancer cells depends on many factors that affect the cell cycle. The radiosensitivity of the cell is the least when in synthesis phase (S) and is the most when in mitotic phase (M) as shown in Figure 1.1 (Pawlik & Keyomarsi, 2004), but there are conditions that make the cell radio-resistant. There is a restriction checkpoint in the first growth phase (G1) where the cell becomes ready to enter the synthesis phase. This checkpoint decides whether the DNA is ready to duplicate in the S phase or it has suffered a lethal damage which results in cell death (Apoptosis). Also, depending on internal and external conditions this checkpoint might delay the progress of the cell in the cycle, a condition called the resting phase (G0) in the cell cycle (Bruce Alberts & Hunt, 2015). A cell becomes radio-resistant when entered in the resting phase (McIlwrath et al., 1994). The resting phase occurs when DNA needs time to repair from sublethal damages, or when there is external stress because of overpopulation or lack of sufficient oxygen source, a condition called hypoxia.

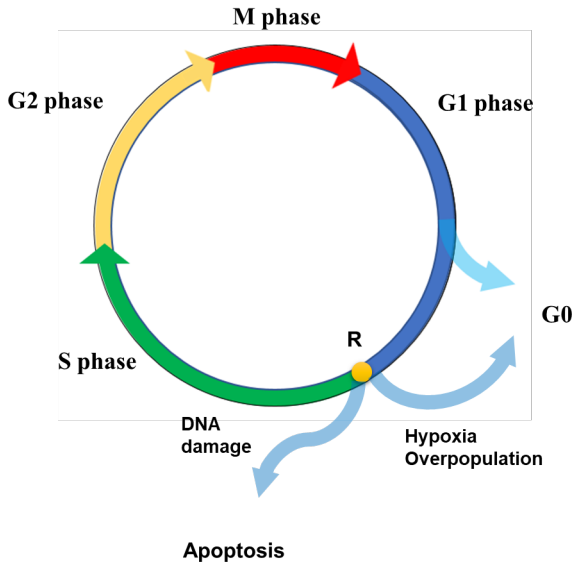


Figure 1.1: Schema of the cell cycle and cell phases

During several fractions of treatment, tumor cells that were in a relatively radio-resistant phase during one fraction might cycle into a sensitive phase before the next fraction is given (Ang, 1998). Therefore, the radiosensetivity of the tumor cells are changing during the treatment horizon and a monotone treatment plan might not be efficient. Also, it has become more clear in recent studies that one patient responds to radiation different than others (Bibault et al., 2013). This dissertation aims to consider the biological response of the patient body to provide a personalized treatment plan. Therefore, the mechanism of action of different radiations and their effect on biological response of the patient body must be considered.

At each treatment fraction, the radiation kills some cells by damaging their DNA, but causing only sublethal damage to some cells. The cells that are active in their cycle and have not been damaged divide into more cells. The rest of the cells that are in resting phase and therefore neither dividing nor radiosensitive, might come out of resting phase as the result of DNA repair, or extension of blood vessels as source of oxygen (reoxygenation), or relocation of cells to overcome overpopulation issue (redistribution). These phenomena are known as four R's of radiation therapy; repair, repopulation, reoxygenation and redistribution.

Various ways to include radiobiological aspects in therapy planning have been studied such as cumulative radiation effect, nominal standard dose, and time-dose factor (Brenner, 2008). These studies use the concepts of tumor control probability (TCP) and normal tissue complication probability (NTCP) and are based on past clinical data. However, before focusing on TCP/NTCP concepts, the biological response to radiation has to be determined. Biological response for both tumor and OAR is demonstrated by cell killing (Brenner, 2008; Gerweck, Zaidi & Zietman, 1994). The most common tool for quantitative prediction of biological response is the *linear-quadratic model* (LQ) (Fowler, 1989; Thames, 1985). In the LQ model the survival of tumor cells is defined as a function of dose and intrinsic radiosensetivity of the tumor tissue. Biologically effective dose (BED) is a measure to characterize the radiation effect based on the survival fraction from LQ model. For fractionated treatment plan, the LQ model is extended to include the four R's of biological factors as the time dependent evolution of tumor in response to treatment.

1.2.2 Treatment planning considering biological effectiveness of different radiations

Low energy photon (x-ray and gamma-ray) is the most commonly used radiation in cancer treatment. Intensity modulated radiation therapy (IMRT) is an image-guided technique that has greatly improved the dose shaping capability of external beam photon therapy (Bortfeld, 2006). IMRT provides a 3D dose shape conformal to the tumor volume by optimizing the intensity of incoming beams of radiation so that the intensity of beams going through sensitive OAR are kept low while the intensity of beams going to the target increase. However, the photon beams when entered human body deposit most of their energy in the beginning of their path, causing more destruction to organs that are close to surface while delivering less of their energy to the tumor(Figure 1.2).

Heavy ion radiotherapy or particle therapy on the other hand, uses the energy of charged particles to kill the cancer cells. The dose deposition of charged particles increases as the beam penetrates deeper in the tissue with a sharp longitudinal dose fall-off at the end of the particle range called *Bragg Peak* (Krämer & Scholz, 2006). This feature has made

charged particles like proton popular in increasing number of radiotherapy facilities recently. Intensity modulated proton therapy (IMPT) utilizes this feature and builds an optimal proton dose distribution in which a three-dimensional conformal dose shape is delivered to the tumor volume with maximal sparing of OARs. The protons are energized to specific velocities that indicate their stopping points in patient's body. Protons of different energies with Bragg peaks at different depths are applied to treat the entire tumor resulting in a total radiation dosage called the spread-out Bragg peak (SOBP). Therefore, the tissues behind (or deeper than) the tumor receive almost no radiation from proton therapy, while the tissues in front of (shallower than) the tumor receive radiation dosage based on the SOBP.

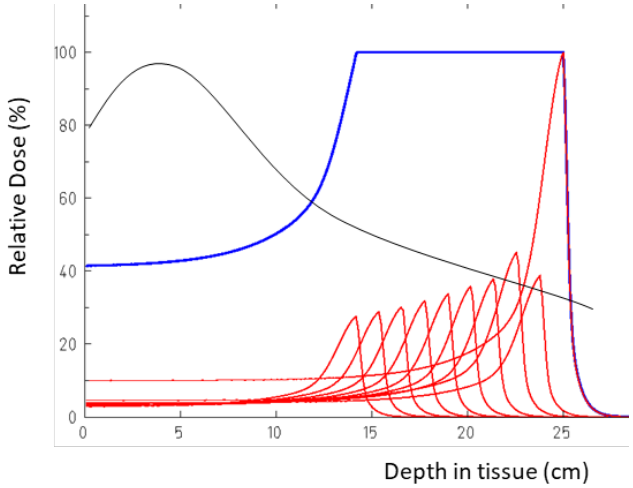


Figure 1.2: Comparison of energy deposition for photon (black line) and proton (red lines) and proton Bragg Peak (blue line)

As the protons move inside the body, their velocity decreases and their interaction with molecules increases which results in maximum energy deposition. The deposited energy, known as *linear energy transfer* (LET), can be controlled by the physician so that the tumor receives highest energy while the OARs receive the lowest energy. Therefore, the biological effectiveness of the proton dose is different than the same photon dose. Relative biological effectiveness (RBE) is a measure for converting a reference dose into photon dose to yield the same biological effect. For proton beam, an RBE of 1.1 is used in clinical practice, but in the distal edge of the SOBP the RBE increases and the range of values goes from 0.7 to 1.6 at

the middle of the SOBP (Paganetti et al., 2002). Paganetti (2014) explains that this average value was deduced (a) for the center of the target volume, (b) at 2 Gy, and (c) averaged over various endpoints and so neglects any dependency of RBE on dose, endpoint or proton beam properties. Several recent studies have provided IMPT plans based on variable RBE (Chen et al., 2018; Ödén, Eriksson & Toma-Dasu, 2017; Willers et al., 2018), but such plan can result in a heterogenous RBE distribution and cause undesired side effects (Ödén, Eriksson & Toma-Dasu, 2017). However, the variable RBE is calculated based on LET which can be controlled so that the tumor receives more energy while the OARs receive similar or less energy. Therefore, considering LET in IMPT planning can account for proton biological effectiveness while mitigating the risk of heterogenous RBE distribution (Unkelbach et al., 2016).

1.2.3 Treatment planning considering uncertainties

A radiation treatment normally consists of one planning session and multiple irradiation sessions. In the planning phase, the patient's internal organs are visualized using CT images. The visualized structures are the basis for construction of the treatment plan and the intention is to deliver this plan in all irradiation sessions. To track the tumor response to treatment, the ideal process is to examine the patient right before each treatment session and update tumor contours indicating the shape and size of the tumor so a decision about radiation dose can be made based on how tumor has responded so far to the treatment. However, observing the tumor and identifying the effect of treatment through digital imaging requires a major amount of time and money and is not practical since the cancer patient cannot wait for long before treatment session, and yet it is prone to human error. The linear-quadratic model assumes tumor response and growth are constant over time; however as explained earlier, there are experimental evidence suggesting the response of tumor cells to radiation and their growth rate are significantly affected by cell cycle regulation and the LQ model is not completely capable of calculating the survival fraction of the tumor in response to treatment. Therefore, we are facing with uncertainty of tumor biological response during fractionated radiotherapy plan.

Proton therapy has unique physical characteristics enabling a 3D dose and energy delivery conforming to the target volume with minimal damage to the healthy organs. If the position or shape of internal structures change during irradiation session due to geometric uncertainties, the outcome of the delivered proton therapy treatment might be different than the expected outcome. the International Commission on Radiation Units and Measurements (ICRU) considers three sources of geometrical uncertainty (Stroom & Heijmen, 2002): patient set-up variation, machine related errors, and organ motion. Patient set-up errors are due to variations in the daily positioning of the patient on the treatment couch. Machine related errors such as beam sizes and gantry angles, are generally considered small due to modern radiotherapy equipment. Organ motion can happen between fractions (known as inter-fractional motion) or within a fraction (called intra-fractional motion). Inter-fractional organ motion can occur due to weight change or variations in rectum or bladder filling, while intra-fractional organ motion is caused by cardiac action and respiration. Inter-fractional motions are easier to deal with because they need intervention only at the beginning of each fraction. But intra-fractional organ motion happen during treatment delivery and change organ position on a time-scale of seconds to minutes (Bert & Durante, 2011).

Because of proton's highly conformal dose deposition, the organ motion results in unequal dose delivered to a specific voxel in different fractions. In proton therapy, some voxels in PTV might face underdosage in some fractions. Therefore, the biological response to treatment for the tumor volume can be different from that of the planned treatment (Flampouri et al., 2006; Nøttrup et al., 2007; Yan & Lockman, 2001). Radiation causes double-strand breaks (DSB) in the DNA. If the radiation dose is not enough to produce lethal DSB, the DNA can repair itself by rejoining the pairs of DSB with a constant rate. The LQ model quantifies both lethal damage and repair effect. According to the LQ model, a reduced radiation dose can result in a reduced lethal damage and increased repair effect and vice versa. Therefore, it is important to investigate the effect of respiratory motion on the biological response in a fractionated treatment plan.

1.3 Objectives and contributions

Our goal in this thesis is to provide solutions to the above-mentioned challenges under the following subjects:

1.3.1 Personalized treatment planning considering biological response uncertainty

This subject attempts to answer if the equal dose fractionation is still optimal under the presence of tumor size and biological response uncertainty. If we understand better about patient specific characteristics with regards to radiation treatment, it is possible to deliver a lesser amount of radiation to achieve the desired treatment goal. Hence, our first goal is to develop a control limit policy as a general guide to determine if a dynamic treatment plan will be better choice than the equal-dose treatment scheme based on relative properties of tumor and OAR.

Furthermore, we aim to provide a general policy to determine how much dose to deliver at each decision epoch during the treatment. The decision can be made based on progression of tumor cells before each treatment. However, evaluation of tumor conditions requires a treatment planner to examine daily CT images, which can be an expensive process in practice. Therefore, we propose a *Partially Observable Markov Decision Process* (POMDP) framework to make such decisions using partially available information about tumor progression during the treatment. In this study, we use the extended LQ model to include the tumor biological response and develop an indefinite time and constrained POMDP to provide optimal treatment policy where the number of the tumor cells before each decision epoch is considered uncertain.

Value iteration is an exact algorithm widely used for solving POMDPs, however this algorithm is not tractable for non-small problems.The structure of our model allows us to use point-based approximation of value iteration as a solution algorithm to limit the size of the solution set at each iteration. Furthermore, a pruning algorithm is used for the constrained POMDP to maintain feasibility of the solution space. We also discus conditions specific to our model to guarantee monotonicity of value function and existence of threshold-type

optimal policy over the belief space.

1.3.2 Personalized and LET-based proton therapy treatment planning considering tumor biological response uncertainty

This subject aims to provide a decision-making policy for a dynamic and personalized fractionated proton therapy treatment plan that can improve the treatment quality (i.e. more damage to the tumor for less damage to OARs) by modifying the treatment based on patient characteristics and tumor response. We combine IMPT optimization into fractionation optimization by defining the set of actions as the dose distributions that provide optimal solution to the LET-based IMPT model. This way, we provide a comprehensive decision making policy that assures the highest quality of a personalized treatment in terms of both fractionation planning and radiation delivery.

Our policy chooses the optimal IMPT plan (action) to be delivered based on the tumor feedback (state). The IMPT optimization model is developed to include LET in calculation of biological response and consider potential effect of variable RBE on treatment quality. This model guarantees homogeneous biological effectiveness instead of common homogeneous dose distribution over the target volume while minimizing the damage to multiple OARs.

In clinical practice, CT images are obtained at the pre-treatment step for treatment planning, where a team of physicians contour the tumor and OAR structure in the CT image. Because this process is time consuming and expensive, it is not always practical to repeat it every week. However, the CT images contain important information âĂŞ *Radiomics*, that include quantitative image features and can offer potential aid in diagnosis, prognosis and prediction of response to treatment of cancer. The idea behind using radiomics in cancer treatment planning is that quantitative analysis of a medical images may improve the efficiency of the planning process while providing better information about the tumor (Gillies et al., 2016). Our objective is to employ a machine learning model to predict tumor state using radiomics data obtained from CT images of a certain cancer site. The trained model then can be applied to weekly CT images during the treatment horizon to predict the tumor size without contouring.

1.3.3 Dealing with respiratory motion in a fractionated IMPT treatment plan

We aim to address the effect of breathing motion on a fractionated proton therapy for non-small cell lung cancer. We consider the biologically effective dose in IMPT model to assure maintaining the required biological response at the presence of breathing motion.

A 4DCT dataset is used to account for organ motion during respiratory cycle. Therefore, our model can handle complex motion patterns and does not require simplifying assumptions about organs motion. In a 4D radiotherapy, the treatment plan is obtained for each 4DCT image and the radiation is delivered with a moving pencil beam synchronized with the breathing motion. However, this delivery method requires a sophisticated delivery system that is not available in all radiotherapy centers (Li et al., 2011). Also, the interplay effect caused by the lag between breathing motion and the pencil beam motion can result in degradation of the treatment quality. We consider a 3D treatment delivery technique so that this model does not require a 4D delivery system and there is no need of dealing with interplay effect.

Two methods are used to solve our proposed model; a statistical model that optimizes the mean and variance of dose delivered to each organ structure, and a robust method that optimizes the worst possible case of dose deposition to each organ structure. While the first method provides a realistic solution for this problem, the second model optimizes a lower bound of treatment quality resulting in a risk-averse treatment plan.

1.4 Organization

This thesis is organized as follows; **Chapter 2** is a comprehensive overview of the relevant literature on treatment plan optimization based on biological response uncertainties associated with radiation therapy. A review of the efforts made to address the biological response to radiation is provided in **Section 2.1**. **Section 2.2** reviews the studies on difference between photon and proton beams and proposed methods to address this difference. A review of some uncertainties of radiation therapy and the methods used to deal with them is presented in **Section 2.3**. A literature review POMDP is provided in **Section 2.4**.

In **Chapter 3**, we propose our stochastic approach to incorporate biological response uncertainties in photon therapy treatment planning optimization. First, a POMDP framework is developed based on extended LQ model to take into account the biological uncertainties, then theoretical control limits and policy structures are investigated depending on our model structure. Finally, the suggested policy is presented using a simulation method and the suggested policy is compared to the traditional treatment scheme.

In **Chapter 4**, we develop an LET-based IMPT optimization model to incorporate variable biologically effectiveness of the proton beams. We leverage radiomics data in predicting tumor state and build a MDP framework to provide a personalized and dynamic decision-making policy for a fractionated radiotherapy treatment planning. We implement our approach on two real cancer cases, a prostate cancer and a pediatric ependymoma and compare the performance of our LET-based IMPT model with the common fixed-RBE IMPT model.

In **Chapter 5**, we address the effect of respiratory motion on a fractionated proton therapy and develop an model to optimize the dose distribution in the IMPT problem and fraction size in the fractionation problem. We use a 4DCT dataset of a non-small lung cancer patient to include the entire raspiratory phases in our model. We solve our model using two different methods and compare their results.

In **Chapter 6**, we conclude the dissertation with a summary of our contributions and discuss the future researches that can be pursued.

Chapter 2

Literature review

2.1 Biologic Response in Fractionation

In considering biological response of tumor in treatment planning, Ribba et al. (2006) explained that before the synthesis phase in a cell cycle, a checkpoint controls internal and external conditions for the DNA to replicate. At this check point if local availability of oxygen is low and/or local density of cells is high, the cell goes into a resting phase which stops the cell from both repopulating and responding to radiation. Therefore, both sensitivity to radiation and growth of a cell depends on the cell cycle state when radiotherapy is happening.

In modification of fractionation pattern, Withers (1985) explained four possibilities: the increase or decrease of overall treatment time and the increase or decrease in dose per fraction. The modification of decreasing dose per fraction is called hyperfractionation. On the other hand, an attractive modification is called hypofractionation in which larger doses are delivered in a shorter treatment period. These modifications are used to answer the different tumor responses.

Fowler (1989) was the first one who used the LQ model in a fractionation problem. The LQ model was proposed to define the survival fraction of tumor based on radiation dose. Fowler (2001) later included tumor proliferation in the LQ model and claimed that the main factor in choosing the optimal total dose, fraction size and total treatment time in a fractionation problem is the proliferation status of the tumor. Several studies have used LQ

model and its extension in fractionation problem and showed that faster growing tumors need shorter radiation treatment plans (Saberian, Ghate & Kim, 2015; Yang & Xing, 2005; Brenner et al., 1995; Bortfeld et al., 2015). The dynamic biologically conformal radiation therapy (DBCRT) model presented by Kim, Ghate & Phillips (2012) was the only study that considered uncertainty in LQ model by assuming a probability distribution for tumor intrinsic radiosensetivity.

Knowing that some tumors possess specific growth curves, some works have considered several growth models for the tumor. Multiple studies have investigated exponential growth models (Wheldon & Kirk, 1976; Wheldon, Kirk & Orr, 1977; Armpilia, Dale & Jones, 2004) while others show efficiency of Gompertzian growth curves (Brunton & Wheldon, 1978; McAneney & O'Rourke, 2007; Usher, 1980; Bortfeld et al., 2015).

2.2 Radiobiological effectiveness of proton

The popularity of proton therapy has been increasing in recent years because of its physical properties, but there are still questions about the biological effect of protons. Experimental evidence has (Courdi et al., 1994) shown that the biological effect caused by proton beams does not depend on the physical dose only, but also on energy deposition (Belli et al., 1993; Paganetti et al., 2002). Currently most clinical proton centers use a constant RBE of 1.1 meaning that protons are 10% more effective than photons. But it has been shown that the RBE increases with depth in the spread-out Bragg peak (SOBP) and becomes about 10% higher than mid-SOBP RBE (Paganetti, 2003) and can reach values of 1.3 to 1.4 at the distal edge (Robertson et al., 1975; Courdi et al., 1994). Therefore, it is important to account for the RBE variations in proton beams (Kraft, 2000; Paganetti et al., 1997). Marshall et at (2016) investigated the effect of variable RBE in proton therapy fractionation by comparing the survival fraction on distal, central and proximal areas of the SOBP of different fractionation regimens with an interaction period of 24 hours. They outlined that the strong dependence of RBE on proton LET has a significant effect on treatment efficiency especially in a variable dose fractionation regimen. In general, the RBE of

protons is calculated based on equivalence of LQ model and for the endpoint and dose and energy range used clinically, RBE has a linear dependence on the local energy spectrum which is characterized by linear energy transfer (LET) (Paganetti et al., 2002; Carabe et al., 2012; Holloway & Dale, 2013).

Currently, the goal of IMPT planning is mostly delivering a homogenous physical dose to the target area, but the biological effectiveness of such plan is not necessarily homogenous. Wilkens & Oelfke (2005) incorporated three-dimensional RBE in inverse IMPT planning and provided more homogenous biological effectiveness based on distribution of the biological dose (RBE multiplied by the physical dose) in the target. Grassberger et al. (2011) showed different IMPT techniques yield equivalent dose distribution while resulting in considerably different LET distributions. Giantsoudi et al. (2013) utilized calculated LET distribution in a dose-based multi criteria optimization to enhance treatment plan.

2.3 Dealing with uncertainties

There are many biological studies on the other hand that believe the LQ model does not integrate the complex features of tumor growth. Instead, tumor growth behavior after radiation strongly depends on a tumor's response to radiation. Barazzuol et al. (2010) considered the LQ model in their investigation on the effects of chemotherapy after each radiation treatment session on glioblastoma cancer and indicated that chemotherapy actually increased tumor response to radiation. Rockne et al. (2009) have also investigated tumor response to radiotherapy, extending Swanson's reaction−diffusion model by LQ model in Gliomas. Kim, Ghate & Phillips (2012) considered the tumor specific constant α in an LQ model as a random variable dependent on time in order to capture tumor response variation during treatment sessions.

In photon therapy, the established tool to compensate for geometrical uncertainties is the planning target volume (PTV). The PTV assumes that the dose distribution is shape invariant with respect to geometrical uncertainties. In proton therapy, because of the motion uncertainties, the distal fall-off of protons is rarely used to spare an OAR that is within 12

cm to the target volume in the direction of the beam. Moreover, the concept of PTV is not suitable for proton therapy as geometrical uncertainties distort the dose distributions because of range uncertainties.

Specific incorporation of the temporal domain into treatment planning for radiotherapy is typically referred to as 4D treatment planning . For photon beam therapy, several studies have investigated the full 4D treatment plan optimization that takes all phases of the 4DCT into account (Nohadani, Seco & Bortfeld, 2010; Zhang et al., 2004). Similar methods have been proposed for the incorporation of range uncertainties by robust treatment planning for proton beams (Pflugfelder, Wilkens & Oelfke, 2008*a*; Unkelbach et al., 2009). However, these studies have not considered the energy deposition and variable RBE of proton in 4D treatment planning and thus the effect of these uncertainties on biological effectiveness of proton therapy has not been investigated.

2.4 Partially Observable Markov Decision Process

Partially observable Markov decision processes (POMDPs) are generalization to MDPs where the state of the system is completely observable. In POMDP instead, a partial observation of the system results in a probability distribution of the states which is referred to as belief state. The belief state converts the POMDP into a MDP with continuous state. Smallwood & Sondik (1973) formulated unconstrained POMDP and provided the properties and value iteration algorithm, as the first exact algorithm, for deriving optimal policy for a finite-state infinite-time POMDP.

The value iteration algorithm becomes intractable for non-small problems, therefore many approximation methods have been studied to overcome this issue (Cassandra, 1998; Poupart, 2005; Hansen, 1998). Grid-based approximation places a finite grid on the belief space and values are computed for points in the grid, and interpolation is used to evaluate all other points in the simplex (Lovejoy, 1987; Hauskrecht, 2000). Point-based algorithms also have made impressive progress solving POMDPs with hundreds of states (Hsu, Lee & Rong, 2008; Pineau et al., 2003). Although the solution theory and algorithm for unconstrained

POMDP is widely studied, the constrained POMDP is relatively less considered in the literature.

Chapter 3

Personalized treatment planning considering biological response uncertainty

3.1 Introduction

The goal of radiation therapy for cancer patients is to kill tumor cells by damaging their DNA. For the majority of patients, the radiation dose required to eradicate the tumor is higher than the tolerance of the surrounding healthy organs called *organs at risk* (OARs). For such cases, the prescribed dose is typically divided into several treatment sessions called fractionated treatment plan. The ideal process for treatment planning is to examine the patient right before each treatment session and update tumor contours indicating the shape and size of the tumor, and make an appropriate decision about radiation dose based on tumor response to treatment. However, this process is not practical because updating tumor contours and finding corresponding optimal radiation dose may not be an option or it can be time consuming and prone to human errors.

In current practice, the treatment policy is to deliver an equal amount of radiation dose in all treatment sessions for most cancer patients. However, it has become more clear in recent studies that one patient responds to radiation different than others (Bibault et al., 2013). Ribba et al. (2006) explained the resting mechanism of cell cycle in response to low availability of oxygen or high local density of cells. Therefore, the sensitivity to radiation and the growth of a cell (i.e., tumor proliferation) both are related to the cell cycle phase

and gene level activities (Scott et al., 2017) as radiation particles travel through the patient body.

Because the radiotherapy typically takes place over multiple fractions, it is possible to modify the amount of daily radiation dose if it can result in better treatment outcomes. Withers (1985) explained four possibilities in modifying the fractionation pattern: the increase or decrease of overall treatment time and the increase or decrease in dose per fraction. Among these four, hyperfractionation, where the prescribed radiation dose is delivered to the patient in multiple smaller fractions, is known to be promising to the current treatment scheme considering the biological characteristics of the tumor. In contrast to Withers's alternative modifications, hypofractionation reduces the treatment time by delivering greater than conventional doses per fraction. To account for such a change during the course of treatment, this paper aims to provide a dynamic treatment plan designed specifically for each patient and each tumor site.

One commonly used mathematical model in fractionated radiotherapy optimization is linear-quadratic (LQ) model (Douglas & Fowler, 1976) considering the biological effective dose (BED) measure to characterize the radiation interaction with tumor and surrounding organs. Fowler (1989) suggested using the LQ model in a fractionation problem, in which the survival fraction of tumor is defined as a function of the radiation dose. He then included tumor proliferation in the LQ model, and claimed the proliferation status of the tumor to be the primary factor in choosing the optimal total dose, fraction size (i.e. amount of dose per fraction) and total treatment time in a fractionation problem (Fowler, 2001). It is shown in several studies that faster growing tumors need to be treated in a shorter treatment length using a higher amount of radiation dose at each fraction. Saberian, Ghate & Kim (2015) considered tumor proliferation with a time lag before the proliferation starts in a fractionation problem, and observed that a shorter treatment course worked better for a fast proliferating tumor. Bortfeld et al. (2015) incorporated accelerated tumor proliferation in an LQ-based treatment planning problem. Knowing that some tumors possess

specific growth curves, multiple studies have investigated exponential growth models (Wheldon & Kirk, 1976; Wheldon, Kirk & Orr, 1977; Armpilia, Dale & Jones, 2004) while others used Gompertzian growth curves (Brunton & Wheldon, 1978; Usher, 1980; McAneney & O'Rourke, 2007).

Beyond the tumor growth factor, Brenner et al. (1995) proposed an LQR model considering four factors to address the biological response of tumor to radiation therapy. These are known as four R's of radiation therapy; *repair of sublethal damage*, *repopulation*, *redistribution*, and *re-oxygenation*. Yang & Xing (2005) developed a LQR-based radiation treatment planning model for a fractionation problem to minimize the BED of organs at risk to the BED of tumor. In the LQR model, the redistribution and reoxygenation effects are cast in terms of resensitization with an average resensitization time. The surviving fraction SF_t of tumor cells irradiated with an arbitrary fractionation scheme can be expressed as

$$SF_t = exp(\sum_{i=1}^{t}(-\alpha d_i - \beta G_i(\tau_{Rp})d_t^2 + \frac{1}{2}\sigma^2 G_i(\tau_{Rs})d_t^2)), \tag{3.1}$$

where, α and β are linear quadratic constants characterizing the intrinsic radio-sensitivity of tissue, σ^2 is the variance of Gaussian distribution of α, τ_{Rp} is average repair time, τ_{Rs} is average resensitization time, and d_t is the dose delivered at each treatment session t (Yang & Xing, 2005). The $G_i(\tau)$ corresponds to the generalized Lea-Catcheside function and is used to describe the time-dependent effects of repair and resensitization. Yang & Xing (2005) calculated this function using

$$G(\tau) = \frac{2}{D^2}\int_0^T R(u)[\int_0^u R(w)exp(\frac{-(u-w)}{\tau})dw]du. \tag{3.2}$$

The biological equivalent dose (Deasy & Fowler, 2005) is calculated based on the total surviving fraction, TF, resulting from the cell kill effect and the cell growth effect,

$$BED = -\frac{1}{\alpha}ln(TF). \tag{3.3}$$

The LQ model assumes that the growth of tumor and tumor response to radiation remain constant over time. However, experimental evidences suggest that the response of tumor cells to radiation and their growth rate are significantly affected by cell cycle regulation, and may change over time (Barazzuol et al., 2010).

The majority of studies on LQ-based optimization of the fractionated treatment planning problem have used non-linear programming as a solution method (Bertuzzi et al., 2013; Yang & Xing, 2005; Saberian, Ghate & Kim, 2015). Since the tumor is evolving based on the amount of radiation dose delivered, Kim, Ghate & Phillips (2009) utilized a Markov Decision Process (MDP) technique for fractionated radiotherapy treatment planning. A drawback of these studies is the assumption that tumor response to radiation can be predicted in advance without uncertainty. Therefore, we propose a *Partially Observable Markov Decision Process (POMDP)* framework to address this issue. POMDP is a tool for sequential decision-making under uncertainty with respect to both the current state and the future evolution of the system (Smallwood & Sondik, 1973). This technique deals with the uncertainty of the tumor state by considering a probabilistic state space called belief space. Smallwood & Sondik (1973) developed an exact algorithm, the value iteration algorithm, for deriving optimal policy for a finite-state infinite-time POMDP. However, because the value iteration algorithm becomes intractable for even small problems, many approximation methods have been proposed to overcome the issue (Cassandra, 1998; Poupart, 2005; Hansen, 1998; Pineau et al., 2003). Although the theory and algorithm for unconstrained POMDP is widely studied, only few studies (Isom, Meyn & Braatz, 2008) have developed algorithms for a constrained POMDP.

This paper attempts to answer if the equal dose fractionation is still optimal under the presence of tumor size and biological response uncertainty. If we understand better about patient specific characteristics with regard to radiation treatment, it is possible to deliver a lesser amount of radiation to achieve the desired treatment goal. Hence, our first goal is to develop a control limit policy within POMDP as a general guide to determine if a dynamic treatment plan will be better choice than the equal-dose treatment scheme based

on relative properties of tumor and OAR. Furthermore, we aim to provide a general policy to determine how much dose to deliver at each decision epoch during the treatment. The decision can be made based on progression of tumor cells before each treatment. However, evaluation of tumor conditions requires a treatment planner to examine daily CT images, which can be an expensive process in practice. Therefore, we propose to use POMDP to make such decisions using partially available information about tumor progression during the treatment. In this study, we develop an indefinite time and constrained POMDP to provide optimal treatment policy where the number of the tumor cells before each decision epoch is considered uncertain. The structure of our model allows us to use point-based approximation of value iteration as a solution algorithm to limit the size of the solution set at each iteration. Furthermore, a pruning algorithm is used for the constrained POMDP to maintain feasibility of the solution space.

The rest of the paper is organized as follows: Section 3.2 gives a brief introduction to POMDP. Section 3.3.1 provides the POMDP formulation for the fractionated radiotherapy treatment plan and the control limits and optimal policy structure for this problem are discussed in Section 3.3.2. The solution method is explained in Section 3.4 and numerical examples are provided in Section 3.5 followed by the conclusion and future work in section 3.6.

3.2 2. POMDP: A short overview

The POMDP is a sequential decision-making model for an agent who acts in a stochastic environment with only partial knowledge about the state of its environment available. Table 4.1 lists definitions of parameters and variables that are used to formulate our problem as a constrained POMDP.

A constrained POMDP can be defined as a tuple $\{\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{O}, \mathcal{Z}, \mathcal{C}\}$. At each time point, the system is in a particular state, $s \in \mathcal{S}$. The problem is to select a feasible action $a \in \mathcal{A}$ that optimizes the performance of the system for its remaining lifetime. Once the decision, (a), is made, the system evolves into a next state, $s' \in \mathcal{S}$, and the system produces an observation, $o \in \mathcal{O}$, that is correlated with the state transition. In an indefinite horizon

Table 3.1: Definition of notations

Notation	Description
$\mathcal{S}$	Set of possible states
$\mathcal{A}$	Set of possible actions
$\mathcal{T}$	Transition probability
$\mathcal{R}$	The immediate reward function
$\mathcal{O}$	Set of possible observations
$\mathcal{Z}$	Observation Probability
$\mathcal{C}$	The immediate cost function
S^t	State of the system at time t
A^t	Decision made at time t
O^t	Observation received at time t
q	Constraint upper bound
T	Terminal time
V_t	Value function for the POMDP policy at time t

POMDP, the process ends when the system enters a terminal state. The transition probability is defined as $\mathcal{T}(s' \mid s, a) = Pr(S^{t+1} = s' \mid S^t = s,\ A^t = a)$ for $a \in \mathcal{A}$ and $s, s' \in \mathcal{S}$, and the observation probability is defined as $\mathcal{Z}(o \mid s, a) = P(O^t = o \mid S^t = s,\ A^t = a)$ for $o \in \mathcal{O}$, $a \in \mathcal{A}$ and $s \in \mathcal{S}$. The decision in each state results in an immediate reward and cost. The constrained and indefinite time POMDP problem can be stated as

$$\begin{aligned} &\max \quad E(\sum_{t=1}^{N} \mathcal{R}(s,a)), \\ &s.t. \\ &E(\sum_{t=1}^{N} \mathcal{C}(s,a)) \leq q, \\ &a \in \mathcal{A}. \end{aligned} \tag{3.4}$$

Note that N is the time when the system enters the terminal state, and q is an input parameter specified by the agent. Since the current observation does not fully reveal the current state, the agent considers all previous observations and actions when choosing a future action to take. All these information is contained in the *belief vector*, (b), which is a probability distribution over the state space, $s \in \mathcal{S}$. For any possible state s, $b(s)$ is the probability associated with state s. The set of all possible belief vectors is called the belief space and it is denoted by $\mathcal{B}$. Given a belief vector b, selected action a, and corresponding

observation o, the resulting belief vector b_o^a can be computed for the $s' \in \mathcal{S}$ component of b as

$$b^{'}(s^{'} \mid a, o, b) = \frac{\mathcal{Z}(o \mid s^{'}, a) \sum_{s \in \mathcal{S}} \mathcal{T}(s^{'} \mid s, a) b(s)}{\sum_{s' \in \mathcal{S}} \mathcal{Z}(o \mid s^{'}, a) \sum_{s \in \mathcal{S}} \mathcal{T}(s^{'} \mid s, a) b(s)}. \tag{3.5}$$

A POMDP decision policy $\pi(b)$ is a mapping from belief space $\mathcal{B}$ to action set $\mathcal{A}$ which determines an optimal action for each possible belief vector.

The value of a POMDP policy π at time t, known as value function $V_t^\pi : \mathcal{B} \to \mathbb{R}$, is the expected remaining reward that can be gained at time t. The value function has two parts: (i) the immediate reward obtained at time t and (ii) the discounted maximum expected remaining reward of $t+1$ given belief vector b_o^a. Averaging over all possible states and observations, the following Bellman's equation is obtained as

$$V_t(b) = \max_a [\sum_s b(s) \mathcal{R}(s, a) + \gamma \sum_s \sum_{s'} \sum_o b(s) \mathcal{T}(s^{'} \mid s, a) \mathcal{Z}(o \mid s^{'}, a) V_{t+1}(b_o^a)]. \tag{3.6}$$

An optimal policy π^* gives the optimal action that maximizes V_t for each b. Here, V_t^* is the corresponding value of the optimal policy π^*.

3.3 Methodology

3.3.1 The Fractionation Problem

The state of the system, $s \in \mathcal{S}$, is defined as the normalized number of tumor cells, which is assumed to be partially observable. A healthy state is considered as the terminal state of the system in our model. The set of possible actions includes: (i) to deliver the conventional dose regime, d, (ii) to increase the dose by a specific amount, $d + \Delta_1$, and (iii) to decrease the dose, $d - \Delta_2$. Partial observations about the tumor are to be obtained from weekly CT images. At the beginning of each treatment week, the physician selects an action to take (i.e., keep the regime as is, increase, or decrease) for the current week. However, due to the effect of tumor growth after the treatment session in the previous week, followed by the weekend break, the observation is made and the belief vector is updated at the beginning of the next week treatment cycle.

The goal of treatment planning is then to maximize the BED of the tumor and minimize the BED of organs at risk, while keeping the tumor survival rate below the OAR survival rate. We define the reward and cost functions of the POMDP model as follows:

$$\mathcal{R}(s,a) = BED_T(s,a) - \lambda BED_{OAR}(s,a) \tag{3.7}$$

and

$$\mathcal{C}(s,a) = \frac{SF_T}{SF_{OAR}}. \tag{3.8}$$

Note that indices T and OAR refer to tumor and OAR, respectively, and λ denotes the weight factor of OAR in Equation 3.7.

The transition probability function, $\mathcal{T}(s' \mid s,a)$, and observation probability function, $\mathcal{Z}(o \mid s,a)$, can be obtained from an expert's opinion. The tumor growth is assumed to follow a Gompertzian Growth function (Bortfeld et al., 2015) as

$$\phi(x) = \left(X_t/X_\infty\right)^{(1-exp(pt))}. \tag{3.9}$$

Here, x is the number of tumor cells, X_∞ is the carrying capacity or the maximum number of tumor cells, and p is a parameter that controls the growth rate.

The BED for action $a \in \mathcal{A}$ in state $s \in \mathcal{S}$ with average number of tumor cells x_s is calculated as

$$BED(s,a) = \sum_{i=1}^{t} \left(a_i + \left(\frac{G_i(\tau_{Rp})}{(\frac{\alpha}{\beta})} - \frac{1}{2}\frac{\sigma^2}{\alpha} G_i(\tau_{Rs}) \right) a_i^2 \right) - \frac{\phi(x_s)}{\alpha}. \tag{3.10}$$

3.3.2 Structure of Optimal Solution

3.3.2.1 Optimal Control Limits without dose limitation

This section explains how one can use the optimal control limits for dose variation to determine the values of Δ_1 and Δ_2 in the action set. Let D^{pres} be the conventional

prescription dose for each session, typically, $2\,Gy$. Depending on the response of the patient to radiation, the treatment planner may increase the dose by a specific amount, incurring a proportional cost of c_r multiplied by the amount of the increase. Similarly, they may decrease the dose by a certain amount, incurring a proportional cost of c_l multiplied by the amount of the decrease. Henceforth, we use the term *pushing-right* as increasing the dose and *pushing-left* as decreasing the dose. Therefore, the delivered dose would be $d+R$ in the case of increasing the dose, $R \geq 0$, and $d-L$ in the case of decreasing the dose, $L \geq 0$.

The four R's of radiation therapy (Repair, Repopulation, Redistribution, and Reoxygenation) may not have a significant effect overnight on tumor biological response to radiation due to Equation 3.2 (Yang & Xing, 2005). Hence, we consider a decision horizon of one week with a constant amount of dose per week. Accordingly, the weekly delivered BED to the tumor with n days of treatment delivery per week can be written as

$$BED_T(s,d) = nd + \Big(\frac{G_{Rp}^T}{(\frac{\alpha}{\beta})_T} - \frac{1}{2}\frac{\sigma^2}{\alpha_T}G_{Rs}^T\Big)d^2 - \frac{\phi_T(x_s)}{\alpha_T}, \tag{3.11}$$

where, G_{Rp}^T and G_{Rs}^T denote the weekly cumulative effect of repair and resensitization of tumor cells, respectively, and $\phi_T(s)$ is the growth rate of tumor cells with x_s as the number of tumor cells. Let θ be the dose-volume parameter which indicates the ratio of the average dose received by the OAR volume to the average dose received by the tumor. Then, the average BED delivered to the OAR can be written as

$$BED_{OAR}(d) = n\theta d + \frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2 d^2 - \frac{\phi_{OAR}}{\alpha_{OAR}}. \tag{3.12}$$

Similarly, G_{Rp}^{OAR} is the cumulative effect of OAR cell repair within one week of treatment with constant dose, and ϕ_{OAR} is the growth rate of OAR cells.

We define the cost (c_r) of *pushing-right* as the excess BED exposed to the OAR when increasing the dose to $d+R$,

$$c_r = n\theta + \frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2(2d+R). \tag{3.13}$$

Similarly, the cost of *pushing-left* is defined as the loss of BED delivered to tumor cells when decreasing the dose to $d - L$,

$$c_l = n + (\frac{G_{Rp}^T}{(\frac{\alpha}{\beta})_T} - \frac{1}{2}\frac{\sigma^2}{\alpha_T}G_{Rs}^T)(2d - L). \tag{3.14}$$

Based on these definitions, a reward function can be defined as

$$\begin{aligned}
&\mathcal{R}(s, D^{pres} + R - L) = n(1 - \lambda\theta)(D^{pres} + R - L) + \\
&\left(\frac{G_{Rp}^T}{(\frac{\alpha}{\beta})_T} - \frac{1}{2}\frac{\sigma^2}{\alpha_T}G_{Rs}^T - \lambda\frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2\right)(D^{pres} + R - L)^2 - \\
&\frac{\phi_T(s)}{\alpha_T} + \lambda\frac{\phi_{OAR}}{\alpha_{OAR}} - c_r R - c_l L.
\end{aligned} \tag{3.15}$$

It is trivial to see that *pushing-right* and *pushing-left* cannot happen at the same time, i.e., $R \cdot L = 0$ at all times.

Theorem 1. Suppose α_T is constant and $\sigma = 0$. Let $A = \frac{G_{Rp}^T}{(\frac{\alpha}{\beta})_T}$ and $B = \frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2$ denote rations of cell repair to radio-sensitivity for tumor and OAR, respectively. Then, $Y = \frac{B}{A}$ is the ratio of OAR cell repair to tumor cell repair corresponding to their α/β ratios. Denoting $H_1 = \frac{n((1+\lambda)\theta - 1)}{2D}$, and $H_2 = \frac{n(\lambda\theta - 2)}{2D}$, the optimal control limits for radiation therapy fractionation are as

$$R^* = \begin{cases} +\infty & \text{if } Y < \frac{1}{\lambda+1} \text{ and } \theta \leq \frac{1}{\lambda+1} \\ 0 & \text{if } Y \geq \frac{1}{\lambda+1} \text{ and } \theta > \frac{1}{\lambda+1} \\ \frac{n\left(1-\theta(1+\lambda)\right)}{2\left(A-(1+\lambda)B\right)} - D^{pres} & \text{otherwise if } A - (1+\lambda)B > H_1 \end{cases}, \tag{3.16}$$

and

$$L^* = \begin{cases} D^{pres} - \frac{n(2-\lambda\theta)}{2(\lambda B - 2A)} & \text{if } Y > \frac{2}{\lambda} \text{ and } \lambda B - 2A > H_2 \\ +\infty & \text{otherwise} \end{cases}. \tag{3.17}$$

Proof. Since the reward function is a quadratic function, it is continuous and twice differentiable, therefore we can analyze its behavior with respect to *right-* and *left pushing*. We can obtain first derivative of the reward function regarding right shift as

$$\frac{\partial \mathcal{R}}{\partial R} = \big(1-(\lambda+1)\theta\big)n + \left(2\frac{G_{Rp}^{T}}{(\frac{\alpha}{\beta})_T} - \frac{\sigma^2}{\alpha_T}G_{Rs}^{T} - 2(\lambda+1)\frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2\right)D^{pres}$$
$$+\left(2\frac{G_{Rp}^{T}}{(\frac{\alpha}{\beta})_T} - \frac{\sigma^2}{\alpha_T}G_{Rs}^{T} - 2(\lambda+1)\frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2\right)R, \tag{3.18}$$

and the second derivative as

$$\frac{\partial^2 \mathcal{R}}{\partial R^2} = \left(2\frac{G_{Rp}^{T}}{(\frac{\alpha}{\beta})_T} - \frac{\sigma^2}{\alpha_T}G_{Rs}^{T} - 2(\lambda+1)\frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2\right)D^{pres}. \tag{3.19}$$

Considering α_T to be constant, if $\frac{\partial^2 \mathcal{R}}{\partial R^2}$ then the reward function is a convex function with respect to R and it would be optimal to increase R to infinity. The reward function is convex with respect to R if

$$2\frac{G_{Rp}^{T}}{(\frac{\alpha}{\beta})_T} - 2(\lambda+1)\frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2 > 0, \tag{3.20}$$

or equivalently

$$\frac{(\alpha/\beta)_T/(\alpha/\beta)_{OAR}}{G_{Rp}^{T}/G_{Rp}^{OAR}} < \frac{2}{2\lambda+1}. \tag{3.21}$$

If the above inequality doesn't hold, then the reward function is concave on R; therefore, we can conclude that if $\frac{\partial Rw}{\partial R}$ is positive on $R=0$ it is optimal to increase the dose by $R^* > 0$, otherwise $R^* = 0$. For $R = 0$,

$$\frac{\partial \mathcal{R}}{\partial R} = (1-(\lambda+1)\theta)n + (2\frac{G_{Rp}^{T}}{(\frac{\alpha}{\beta})_T} - 2(\lambda+1)\frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2)D^{pres}. \tag{3.22}$$

Since $Y = \frac{(\alpha/\beta)_T/(\alpha/\beta)_{OAR}}{G_{Rp}^{T}/G_{Rp}^{OAR}} < \frac{2}{2\lambda+1}$, therefore $2\frac{G_{Rp}^{T}}{(\frac{\alpha}{\beta})_T} - 2(\lambda+1)\frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2$ would be negative too. In that case if $\theta \geq 1/(\lambda+1)$ then reward function would be decreasing on $R=0$, therefore

$R^* = 0$. If $\theta < {}^1/_{(\lambda+1)}$ then $R^* > 0$ only if

$$(1-(\lambda+1)\theta)n + (2\frac{G_{Rp}^T}{(\frac{\alpha}{\beta})_T} - 2(\lambda+1)\frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2)D^{pres} > 0. \tag{3.23}$$

Defining $A = \frac{G_{Rp}^T}{(\frac{\alpha}{\beta})_T}$ and $B = \frac{G_{Rp}^{OAR}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2$, we can rewrite Equation 3.23 as $A - (1+\lambda)B > n((1+\lambda)\theta - 1)/2D$. Similarly, we can obtain first and second derivative of reward function with respect to left shift and explain the control limits. □

For different patients of the same cancer site, the tumor can possess different characteristics including response to radiation characteristics ${}^\alpha/_\beta$ and repair of sublethal damage characteristics (Rp). In an ideal condition where the tumor radio-sensitivity is higher than OAR radio-sensitivity and the tumor cell repair takes longer than OAR cell repair, the values of Y ratio is less than 1. In other words, smaller values of Y ratio indicate higher potential for damaging tumor and recovering OAR while larger values of Y ratio indicate potential of more OAR damaging and less tumor damaging. **Theorem 1** explains that the optimal control limits of dose change during the treatment horizon do not depend on tumor characteristics alone, but depend highly on Y ratio, dose-volume parameter, and the weight factor of OAR in Equation 3.15. Given the dose volume parameter, θ, and the weight factor of OAR, λ, one can use **Theorem 1** to find the optimal control limits based on the value of Y ratio.

3.3.2.2 Optimal Control Limits with dose limitation

In **Theorem 1**, the optimal right and left control limits are infinity in some conditions as in Equations 3.16 and 3.17. However, it is not allowed to deliver a radiation dose beyond a specific limit because it will create undesirable side effects. The physicians usually specify the bounds for the *right-* and *left-pushing* relative to the prescribed dose.

Corollary 1. Let v and u define the maximum acceptable *left-pushing* and the maximum acceptable *right-pushing*, respectively. The optimal control limits for radiation therapy

fractionation with dose limitation can be presented as follows:

$$R^* = \begin{cases} u & \text{if } Y < \frac{1}{\lambda+1} \text{ and } \theta \leq \frac{1}{\lambda+1} \\ 0 & \text{if } Y \geq \frac{1}{\lambda+1} \text{ and } \theta > \frac{1}{\lambda+1} \\ min\left\{u, \frac{n\left(1-\theta(1+\lambda)\right)}{2\left(A-(1+\lambda)B\right)} - D^{pres}\right\} & \text{otherwise if } A-(1+\lambda)B > H_1 \end{cases}, \tag{3.24}$$

and

$$L^* = \begin{cases} min\left\{v, D^{pres} - \frac{n(2-\lambda\theta)}{2(\lambda B-2A)}\right\} & \text{if } Y > \frac{2}{\lambda} \text{ and } \lambda B - 2A > H_2 \\ 0 & \text{otherwise} \end{cases}, \tag{3.25}$$

where $A = \frac{G^T_{Rp}}{(\frac{\alpha}{\beta})_T}$, $B = \frac{G^{OAR}_{Rp}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2$, $H_1 = \frac{n((1+\lambda)\theta-1)}{2D}$, $H_2 = \frac{n(\lambda\theta-2)}{2D}$ and $Y = \frac{B}{A}$.

Proof. The proof follows as of Theorem 1 except for the case when $R^* = \infty$ and $L^* = \infty$ at $Y < \frac{2}{2\lambda+1}$ where the reward function is convex with respect to right shift (R). In this case if $\frac{\partial\mathcal{R}}{\partial R} > 0$ then the reward function is increasing on R and therefore, the optimal limit is equal to the dose increase limit, $R^* = u$. One simple way of having $\frac{\partial\mathcal{R}}{\partial R} > 0$ is when all the components of $\frac{\partial\mathcal{R}}{\partial R}$ are positive. This happens when $1-(\lambda+1)\theta > 0$ and $\frac{G^T_{Rp}}{(\frac{\alpha}{\beta})_T} - (\lambda+1)\frac{G^{OAR}_{Rp}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2 > 0$. Therefore, when $\theta < \frac{1}{\lambda+1}$ and $Y = \frac{(\alpha/\beta)_T / (\alpha/\beta)_{OAR}}{G^T_{Rp}/G^{OAR}_{Rp}} < \frac{1}{\lambda+1}$, the reward is an increasing function of R and $R^* = u$.

When $\frac{1}{\lambda+1} < \frac{2}{2\lambda+1}$, one component of $\frac{\partial\mathcal{R}}{\partial R}$ is negative, but

$$\frac{\partial\mathcal{R}}{\partial R} \mid_{R=0} = (1-(\lambda+1)\theta)n + (2\frac{G^T_{Rp}}{(\frac{\alpha}{\beta})_T} - 2(\lambda+1)\frac{G^{OAR}_{Rp}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2)D^{pres}. \tag{3.26}$$

Therefore, if $2\frac{G^T_{Rp}}{(\frac{\alpha}{\beta})_T} - 2(\lambda+1)\frac{G^{OAR}_{Rp}}{(\frac{\alpha}{\beta})_{OAR}}\theta^2 < \frac{((\lambda+1)\theta-1)n}{2D^{pres}}$ then, $\frac{\partial\mathcal{R}}{\partial R} > 0$ and $R^* = u$. Otherwise, reward will be decreasing function of R at R=0 and will not improve with positive R and $R^* = 0$.

For $\theta > \frac{1}{\lambda+1}$ and $Y < \frac{1}{\lambda+1}$, one component of $\frac{\partial\mathcal{R}}{\partial R}$ is negative and the same condition as in above applies. For $\theta > \frac{1}{\lambda+1}$ and $Y > \frac{1}{\lambda+1}$, two components of $\frac{\partial\mathcal{R}}{\partial R}$ are negative and

$\frac{\partial \mathcal{R}}{\partial R} \mid_{R=0} < 0$, therefore it will not be beneficial to increase R and $R^* = 0$. Similar approach will conclude in conditions on left shift (L). □

Considering upper limits on *right-* and *left-pushing*, the maximum dose change is optimal only when Y ratio and θ dose-volume parameter are small. When the optimal *right-pushing* is positive, the exact R^* can be obtained by maximizing the reward function 3.15.

3.3.2.3 Optimal Policy Structure

In this section, we investigate the structure of the optimal policy based on the POMDP model explained in sections 3.3.1. Lovejoy (1987) provided sufficient conditions for the optimal value function to be monotone on the belief space using partial orders.

Definition 1. (Monotone likelihood ratio (MLR) ordering) (Krishnamurthy, 2016) Let $\mathcal{X}$ denote a finite completely ordered set with n elements. The set $\Pi(\mathcal{X})$,

$$\Pi(\mathcal{X}) = \left\{ \pi \in \mathbb{R}^n;\ \pi_i \geq 0 \text{ all } i; \sum_{i=1}^{n} \pi_i = 1 \right\},$$

has the MLR property when for $\pi, \pi' \in \Pi(\mathcal{X})$, π dominates π' with respect to the MLR order denoted as $\pi \geq_r \pi'$ if

$$\pi_i \pi'_{i'} \geq \pi_{i'} \pi'_i.$$

Definition 2. (Totally positive of order two (TP_2)) (Karlin, 1964) A function $K(x, y)$ of two real variables ranging over linearly ordered sets X and Y respectively, is said to be totally positive of order two if $K(x_1, y_1)K(x_2, y_2) \geq K(x_1, y_2)K(x_2, y_1)$ for $x_1 \geq x_2$ and $y_1 \geq y_2$.

Proposition 1. The reward function $\mathcal{R}$ in Equation 3.15 has an increasing order on that state set i.e. $\mathcal{R}(s, a) \geq \mathcal{R}(s', a)$ if $s > s'$.

Proof. The first derivative of reward function with respect to number of tumor cells is

$$\frac{\partial \mathcal{R}}{\partial x} = \frac{\partial(-\phi(x)/\alpha_T)}{\partial x} = \frac{b^{X_\infty}/x}{\alpha_T}, \tag{3.27}$$

which has always a positive value. Since the state is defined based on number of tumor cells, we can conclude that reward is increasing over the state set. □

Corollary 2. If **proposition 1** holds, and

i. The transition matrix is TP_2 for each action $a \in \mathcal{A}$,

ii. $\mathcal{Z}(j) \geq_r \mathcal{Z}(j^{'})$ for $j \geq j^{'}$ in $\mathcal{S}$, where $\mathcal{Z}(j)$ is the j^{th} column of the observation matrix,

then $b \geq_r b^{'}$ implies $V_t^*(b) \geq V_t^*(b^{'})$ (Lovejoy, 1987); In other words, the value function monotonically increasing in belief space and therefore, there exist two thresholds b_1^* and b_2^* such that

$$a^* = \begin{cases} d - L & \text{if } b \leq_r b_1^* \\ d - L & \text{if } b_1^* \leq_r b \leq_r b_2^* \\ d - L & \text{if } b \geq_r b_2^* \end{cases}. \tag{3.28}$$

3.4 Solution Method

We use the value iteration algorithm to solve the POMDP model 3.6. The value iteration algorithm starts with an initial value function and iteratively updates the estimate of the value function. In a finite POMDP problem the V_t^* will be piecewise linear and convex (PWLC) in the belief space (Smallwood & Sondik, 1973), while it can be approximated well by a PWLC function in an infinite POMDP problem (Spaan & Vlassis, 2005). The value function V_t at time t can be represented by a finite set of vectors $\{\mu_t^k\}$, $k = 1, 2, \ldots, |V_t|$, $\alpha_t^k \in \Re^{|S|}$. Each vector defines a hyperplane over the belief space and is associated with an action that is optimal under that hyperplane (Sondik, 1971). Given a set of vectors $\{\mu_t^k\}_{k=1}^{|V_t^\mu|}$ for policy μ, the value of a belief vector b is calculated by

$$V_t^\mu(b) = \max_{\{\mu_t^k\}} \big(\sum_s b(s)\ \mu_t^k(s) \big). \tag{3.29}$$

Given V_t as the solution set at the current stage, the size of solution set at the next stage $|V_{t-1}|$ will grow exponentially, being $|\mathcal{A}||V_t|^{|\mathcal{O}|}$ (Spaan & Vlassis, 2005). In addition, POMDP

models can become intractable because the optimal action must be calculated for all belief vectors in belief space $\mathcal{B}$. A point-based approximation of a value function involves computing an approximate solution for a finite set of belief vectors. For most POMDP problems, it is unlikely to reach all points in the belief space. Therefore, it is reasonable to update the value function for only reachable belief vectors (Vlassis, Spaan et al., 2004). In this study, the states are defined as the number of tumor cells in an increasing order, therefore we can assume that the reachable belief vectors must have a certain type of structure, i.e. monotone or concave. We define a finite set of belief vectors $B = \{b_0, b_1, \ldots, b_q\}$ by randomly sampling belief vectors following this structure.

the point-based approximation of the value iteration algorithm starts with an initial vector for each selected belief point, and performs the value backup only for the selected belief points. For updating the value function with the current solution set $V_{t+1} = \{\mu_{t+1}^{k'}\}$, $k' = 1, 2, \ldots, |V_{t+1}|$ we define the intermediate set $\Gamma^{a,o}$ as

$$\Gamma^{a,o} \leftarrow \alpha_i^{a,o}(s) = \sum_{s} \sum_{s'} \sum_{o} \mathcal{T}(s^{'} \mid s, a)\mathcal{Z}(o \mid s^{'}, a)\mu_{t+1}^{k'}(s^{'}), \forall \mu_{t+1}^{k'} \in V_{t+1}, \forall a \in \mathcal{A}, \forall o \in \mathcal{O}. \quad (3.30)$$

Then, we construct for each action in each belief point

$$V_t \leftarrow \mu_{t+1}^{k'} = \mathcal{R}(., a) + \gamma \sum_{o} \underset{\mu \in \Gamma^{a,o}}{\operatorname{argmax}}(\mu \cdot b), \quad \forall a \in \mathcal{A}, \forall b \in \mathcal{B}, k = 1, 2, \ldots . \quad (3.31)$$

For a constrained POMDP, the solution set V_t consists of pairs of μ for the reward and cost functions $(\mu_{\mathcal{R}_t}^k, \mu_{\mathcal{C}_t}^k) \in V_t$ where $\mu_{\mathcal{R}_t}^k$ and $\mu_{\mathcal{C}_t}^k$ are associated with reward and cost functions, respectively. The point-based algorithm is applied to both $\{\mu_{\mathcal{R}_t}^k\}$ and $\{\mu_{\mathcal{C}_t}^k\}$. The pairs that fail to satisfy condition $\mu_{\mathcal{C}_t}^k \cdot b \leq q$ are eliminated from the solution set. Therefore, the size of solution set will remain equal to or smaller than $|\mathcal{B}|$ in all iterations and the optimal solution will be obtained from Equation 3.29.

3.5 Numerical Examples

Experiments are made to compare the performance of two treatment planning schedules: the proposed dynamic treatment plan and the conventional treatment plan. For the dynamic treatment planning, we will illustrate how to use the POMDP policy considering tumor progression to determine optimal radiation dose per fraction at each decision epoch for the treatment. Two types of tissues are considered with regard to the response to radiation: *early responding* tissues having a faster proliferation rate, shorter repair time, and higher sensitivity to radiation, while *late responding* tissues having the opposite characteristics.

3.5.1 Performance metrics

Biological Equivalent Dose is used as the primary performance measure for the comparison purposes. Other commonly used measures to investigate effectiveness of a cancer treatment plan include tumor control probability (TCP) and normal tissue complication probability (NTCP). Hence, TCP and NTCP are also used as secondary performance measures. The TCP (Brenner, 1993) defines probability of killing tumor cells,

$$TCP = exp(x_0 SF). \tag{3.32}$$

The NTCP (Semenenko & Li, 2008) defines the probability of damage incurred to normal tissue surrounding the tumor and is calculated by

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx, \quad t = \frac{MD - TD_{50}}{mTD_{50}}. \tag{3.33}$$

Here, MD is the mean organ dose, TD_{50} is the uniform dose given to the entire organ volume that results in a 50% complication risk, and m is a measure of the slope of the sigmoid curve represented by the integral of the normal distribution.

3.5.2 Control limits

Table 2 provides a summary of the input parameters that several studies (Yang & Xing, 2005; Saberian, Ghate & Kim, 2015) have suggested for early responding and late responding tissues of tumor and OAR. Considering these input parameters and control limit rules

Table 3.2: Radiotherapy model input parameters

	Early Responding		Late Responding	
	Tumor	OAR	Tumor	OAR
α/β	$[1,4]Gy$	$3Gy$	$[4,10]Gy$	$10Gy$
Repair time	30 minutes	30 minutes	1.9 hours	4 hours

obtained in **Corollary 1**, we attempt to answer the question "Is it beneficial to alter the conventional treatment plan?" If so, how much change of dose should be considered in the dynamic treatment plan. Four cases are examined to answer this question.

Case 1: Early responding tumor with early responding OAR

The upper and lower bounds of Y ratio (see Section **??**) for dose-volume parameter $\theta \in [0,1]$ are presented in Figure 3.2.

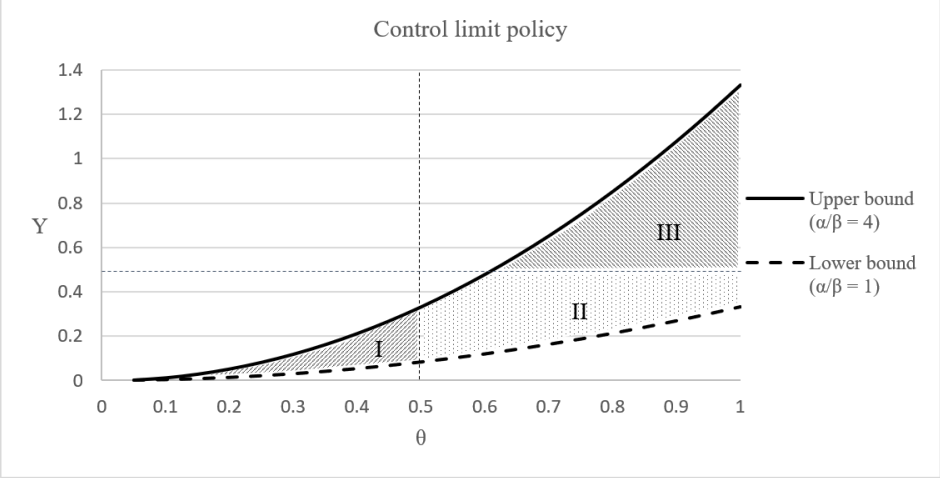


Figure 3.1: Control limit policy for early responding tumor and early responding OAR

The figure shows the control limit boundaries when equal weights are assigned to both the OAR and the tumor (i.e., $\lambda = 1$). Depending on the area in which Y ratio and θ fall, one can choose between the conventional planning scheme and the dynamic POMDP

planning. If they belong to Area I, it is recommended to use the dynamic POMDP planning with an increased dosage to D^{pres+u}. In Area II, if $A-(1+\lambda)B > H_1$ with $H_1 = n((1+\lambda)\theta-1)(2D^pres)$, then a POMDP plan with increased dosage to D^{pres+R}, $0 < R \leq u$, is recommended. If Y ratio and θ belong to Area III, the conventional plan is recommended over a POMDP plan. Similarly, recommended decisions can be derived if the importance factor of the OAR is smaller than that of the tumor, i.e., $\lambda < 1$.

Case 2: Early responding tumor with late responding OAR

Figure 3.2 shows the upper and lower bounds of Y ratio corresponding to different values of dose-volume parameter $\theta \in [0, 1]$.

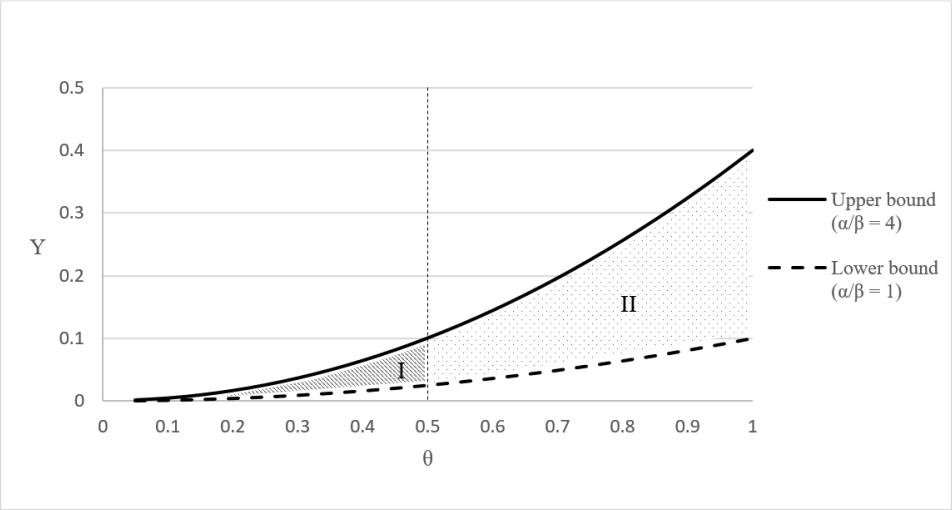


Figure 3.2: Control limit policy for early responding tumor and late responding OAR

Assuming $\lambda = 1$, the recommended planning scheme is: (1) the POMDP policy considering D^{pres+u} when Y ratio and θ belong to Area I; and (2) the POMDP policy considering D^{pres+R}, $0 < R \leq u$, in Area II if $A-(1+\lambda)B > H_1$ with $H_1 = n((1+\lambda)\theta-1)(2D^pres)$. This result is reasonable since when the radio-sensitivity of tumor is high and the radio-sensitivity of OAR is low. This enables us to increase the damage to the tumor while lowering damage to the OAR. Therefore, a higher amount of dose can be delivered.

Case 3: Late responding tumor with early responding OAR

Assuming $\lambda = 1$, the recommended planning scheme is: (1) the POMDP policy considering D^{pres+u} when Y ratio and θ belong to Area I; (2) the POMDP policy considering D^{pres+R}, $0 < R \leq u$ in Area II and Area III if $A - (1 + \lambda)B > H_1$, with $H_1 = n((1+\lambda)\theta - 1)(2D^p res)$; (3) the conventional plan in Area IV; and (4) the POMDP policy considering D^{pres-L}, $0 < L \leq v$ in Area V if $\lambda B - 2A > H_2$, with $H_2 = n(\lambda\theta - 2)(2D^p res)$ (Figure 3.3). This means when the radio-sensitivity of tumor is low and the radio-sensitivity of OAR is high, the optimal right shift is positive only if the dose-volume parameter, , is smaller.

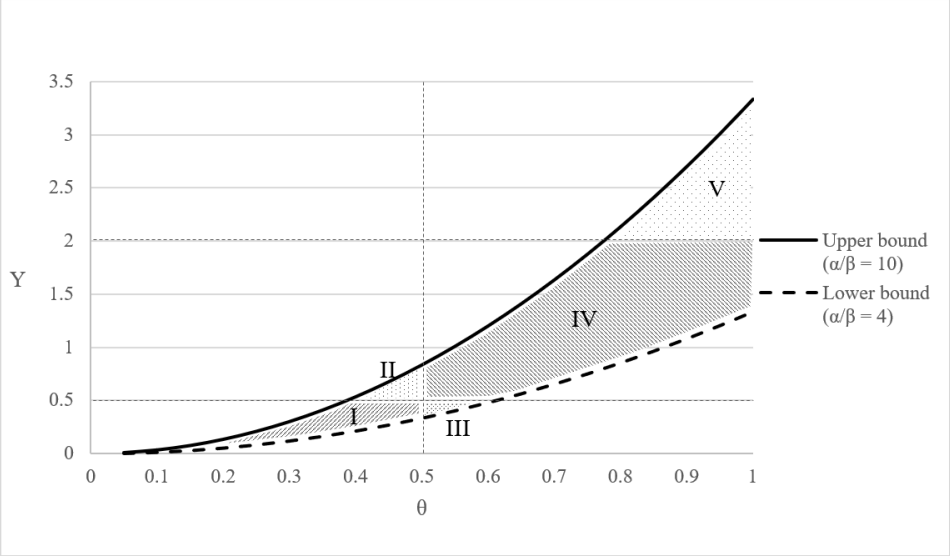


Figure 3.3: Control limit policy for late responding tumor and early responding OAR

Case 4: Late responding tumor with late responding OAR

Assuming $\lambda = 1$, the recommended planning scheme is: (1) the POMDP policy considering D^{pres+u} in Area I; (2) the POMDP policy considering D^{pres+R}, $0 < R \leq u$ in Area II if $A - (1 + \lambda)B > H_1$, with $H_1 = n((1 + \lambda)\theta - 1)(2D^p res)$; (3) The conventional plan in Area III; (Figure 3.4).

3.5.3 Optimal policy

In this section, we discuss numerical results based on given transition and observation probabilities as provided in Tables 3.3 and 3.4.

Table 3.3: Transition probability matrix

Transition Probabilities, Tr(S'|S,a)					
a=1	Ω	S(1)	S(2)	S(3)	S(4)
Ω	1	0	0	0	0
S(1)	0.3	0.5	0.2	0	0
S(2)	0.05	0.15	0.7	0.1	0
S(3)	0	0	0.14	0.71	0.05
S(4)	0	0	0	0.23	0.77

a=2	Ω	S(1)	S(2)	S(3)	S(4)
Ω	1	0	0	0	0
S(1)	0.4	0.3	0.2	0	0
S(2)	0.07	0.23	0.65	0.05	0
S(3)	0	0	0.24	0.75	0.01
S(4)	0	0	0	0.37	0.63

a=3	Ω	S(1)	S(2)	S(3)	S(4)
Ω	1	0	0	0	0
S(1)	0.5	0.35	0.1	0	0
S(2)	0.1	0.25	0.6	0.05	0
S(3)	0	0	0.34	0.66	0
S(4)	0	0	0	0.51	0.49

Table 3.4: Observation probability matrix

Observation probabilities, Z(O,S)					
	Ω	S(1)	S(2)	S(3)	S(4)
O(Ω)	1	0	0	0	0
O(1)	0	0.65	0.4	0.1	0.05
O(2)	0	0.3	0.4	0.45	0.25
O(3)	0	0.05	0.2	0.45	0.7

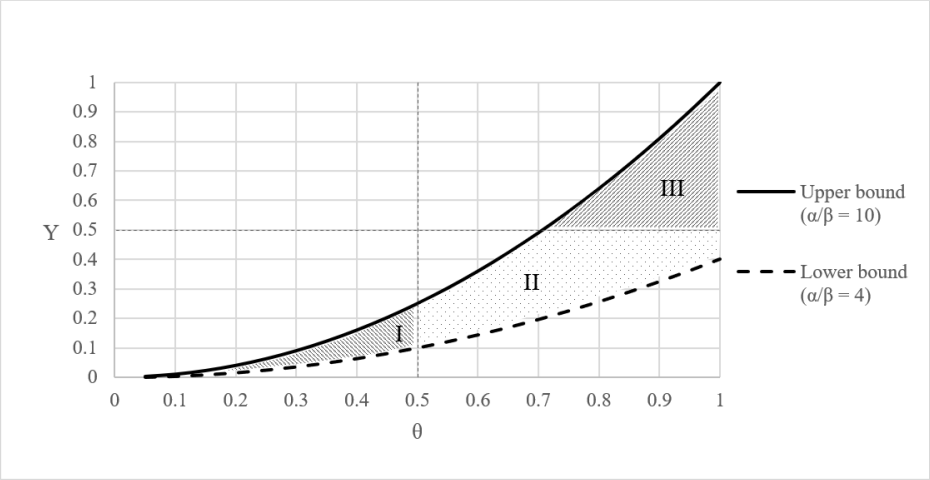


Figure 3.4: Control limit policy for late responding tumor and late responding OAR

The input parameters used in this section are presented in Table 3 (Bortfeld et al., 2015; Rew & Wilson, 2000; Yang & Xing, 2005). Using the these input parameters, we implemented a point-based value iteration algorithm on the belief space. The belief vectors in the proposed POMDP are subsets of the belief space. The elements of the vector are the probabilities of corresponding states. In our study, the probabilities in any vector can have three patterns: (1) monotonically increasing (e.g., [0.0, 0.2, 0.3, 0.5]), (2) monotonically decreasing, and (3) concave (e.g., [0.0 0.3 0.5 0.2]). Based on this structure, a point-based approximation of value iteration is used to solve the POMDP model. The solution consists of a set of $\{\mu_R^k\}$ which provides the optimal action using Equation 3.28. Since the belief vectors evolve with actions and observations, we present the policies by simulating observation scenarios. In this regard, a probability distribution over the observations set, $\mathcal{O}$, evolves at each decision epoch based on the history of actions and observation using Bayesian updates. Then, the observation scenarios are generated at each decision epoch based on the distribution over observations set at that epoch. The simulation is implemented 1,000 times for a treatment horizon of 8 weeks, and the optimal actions in all simulations over the treatment horizon are used to obtain the probability distributions over the set of actions for each week. Table 3.6 presents the probability distribution of optimal actions for different tumor-OAR combinations and different OAR weight factors, λ, with an initial belief state representing

Table 3.5: Radiotherapy model input parameters

		Early Responding	Late Responding
Tumor	α/β	3.5 G	10 G
	α	0.35 G^{-1}	0.1 G^{-1}
	τ_{Rp}	30 minutes	1.9 hours
	τ_{Rs}	1 day	2 days
	p	exp (−5.03)	exp (−6.92)
	X_0	6×10^{11}	4×10^{6}
	X_∞	5×10^{12}	
OAR	α/β	3 G	10 G
	α	0.315 G^{-1}	
	τ_{Rp}	30 minutes	4 hours
	ϕ_{OAR}	1.5	1.02

a worst degree of cancer. The possible actions include to deliver the prescribed dose, to change the prescribed dose by 0.2 Gy, or to terminate the treatment because it reached the terminal state.

Case 1: Early responding tumor with early responding OAR

For an early responding tumor and an early responding OAR, the optimal policy is to deliver a 2 Gy dose per fraction for all scenarios in the first five weeks. On week six, depending on the observation received from the tumor, 2 Gy dose was recommended in 71% of scenarios and 2.2 Gy dose was recommended in 23% of scenarios. Starting from week seven, 2.2 Gy dose was suggested for all scenarios except for the cases that reached the terminal state. As the treatment progresses, a trend of having a higher probability of entering the terminal phase was observed.

Case 2: Early responding tumor with late responding OAR

In this case, the optimal policy recommended a more aggressive treatment plan compared to Case 1. The 2 Gy dose was suggested for all scenarios for the first four weeks, but on

Table 3.6: Probability of optimal action with initial belief vector $[0, 0.01, 0.05, 0.14, 0.8]$

Tumor	OAR	λ	Opt. action	W = 1	W = 2	W = 3	W = 4	W = 5	W = 6	W = 7	W = 8
Early α/β = 3.5	Early α/β = 3	1	P(d = 1.8 Gy)	0	0	0	0	0	0	0	0
			P(d = 2 Gy)	1	1	1	0.99	0.98	0.71	0	0
			P(d = 2.2 Gy)	0	0	0	0	0	0.23	0.89	0.77
			P(TERM)	0	0	0	0.01	0.02	0.07	0.11	0.23
Early α/β = 3.5	Late α/β = 8.5	1	P(d = 1.8 Gy)	0	0	0	0	0	0	0	0
			P(d = 2 Gy)	1	1	1	0.99	0.35	0	0	0
			P(d = 2.2 Gy)	0	0	0	0	0.62	0.92	0.79	0.69
			P(TERM)	0	0	0	0.01	0.03	0.08	0.21	0.31
Late α/β = 10	Early α/β = 3	1	P(d = 1.8 Gy)	0	0	0	0	0	0	0	0
			P(d = 2 Gy)	1	1	1	0.99	0.98	0.93	0.9	0.85
			P(d = 2.2 Gy)	0	0	0	0	0	0	0	0
			P(TERM)	0	0	0	0.01	0.02	0.07	0.1	0.15
Late α/β = 10	Late α/β = 8.5	1	P(d = 1.8 Gy)	0	0	0	0	0	0	0	0
			P(d = 2 Gy)	1	1	1	1	1	1	0.7	0
			P(d = 2.2 Gy)	0	0	0	0	0	0	0.3	1
			P(TERM)	0	0	0	0	0	0	0	0
Early α/β = 3.5	Early α/β = 3	0.1	P(d = 1.8 Gy)	0	0	0	0	0	0	0	0
			P(d = 2 Gy)	1	1	1	0.99	0.68	0.2	0	0
			P(d = 2.2 Gy)	0	0	0	0	0.29	0.73	0.86	0.73
			P(TERM)	0	0	0	0.01	0.03	0.07	0.14	0.27
Early α/β = 3.5	Late α/β = 8.5	0.1	P(d = 1.8 Gy)	0	0	0	0	0	0	0	0
			P(d = 2 Gy)	1	1	1	0.99	0.61	0	0	0
			P(d = 2.2 Gy)	0	0	0	0	0.36	0.92	0.82	0.73
			P(TERM)	0	0	0	0.01	0.03	0.08	0.18	0.27
Late α/β = 10	Early α/β = 3	0.1	P(d = 1.8 Gy)	0	0	0	0	0	0	0	0
			P(d = 2 Gy)	1	1	1	1	0.98	0.75	0.32	0
			P(d = 2.2 Gy)	0	0	0	0	0	0.2	0.58	0.84
			P(TERM)	0	0	0	0	0.02	0.05	0.1	0.16
Late α/β = 10	Late α/β = 8.5	0.1	P(d = 1.8 Gy)	0	0	0	0	0	0	0	0
			P(d = 2 Gy)	1	1	1	0.99	0.97	0.74	0.29	0
			P(d = 2.2 Gy)	0	0	0	0	0	0.21	0.61	0.85
			P(TERM)	0	0	0	0.01	0.03	0.05	0.1	0.15

week five 2.2 Gy dose was suggested for 62% of scenarios and for the rest of treatment 2.2 Gy dose was suggested for all scenarios that did not reached the terminal state. Compared to Case 1, more scenarios reached the terminal phase to terminate the treatment towards the end of the treatment.

Case 3: Late responding tumor with early responding OAR

In the case, the OAR has higher sensitivity to radiotherapy than the tumor and a dose of 2 Gy was recommended for all scenarios in all treatment weeks. In this case, lower probability of termination can be observed compared to Cases 1 and 2, i.e. less scenarios reach the terminal state.

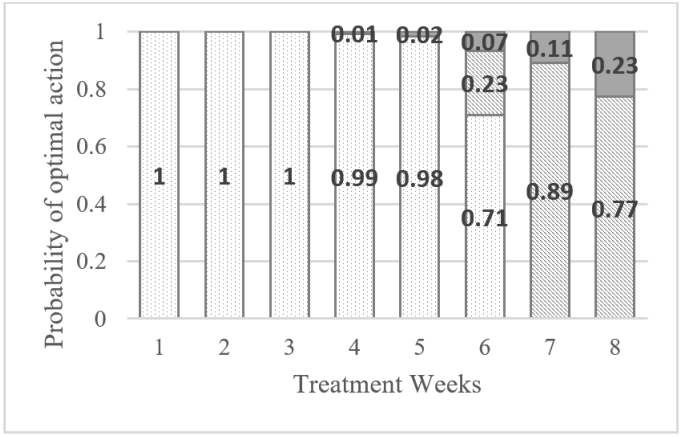
Case 4: Late responding tumor with late responding OAR

When the tumor and OAR both have low sensitivity to radiotherapy, increasing the dose from 2 Gy to 2.2 Gy was suggested on weeks seven (30%) and week eight (100%). However, the probability of treatment termination in this case is nearly zero during eight weeks of treatment, implying that more treatment sessions are required to reach the terminal state.

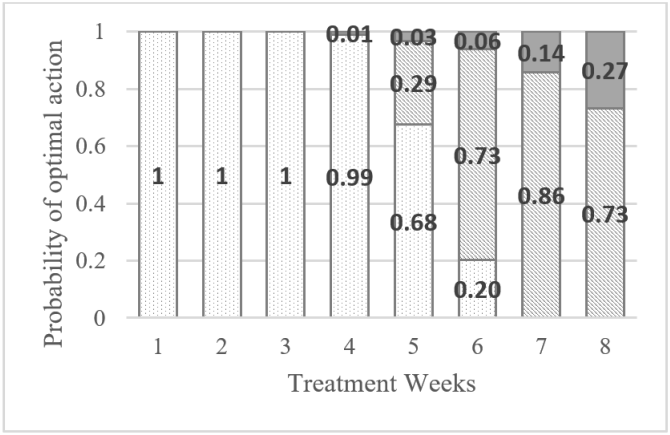
The effect of OAR weight factor

The importance factor of the organ as OAR in Equation 3.15 and the planner's preference on the OAR sparing relative to the tumor can influence the POMDP policy. Figure 2 shows the comparison of different OAR weight factors for the case of an early responding tumor and OAR. When the OAR weight factor is equal to that of tumor ($\lambda = 1$), it is recommended to administer a 2 Gy dose regime for the first five weeks; and increase the dose to 2.2 Gy for 23% of the scenarios on week six and continue with 2.2 Gy for all scenarios on weeks seven and eight (Figure 2(a)). When the OAR weight factor is less than that of tumor ($\lambda < 1$), a more aggressive plan was recommended, suggesting to increase the dose to 2.2 Gy for 29% of scenarios on week five and continue with high dose on the next weeks (Figure 3.5(b)). In addition, the termination probability of the latter was slightly improved in Figure 3.5(b) compared to Figure 3.5(a).

The effect of the initial belief vector



(a) OAR weight (λ) = 1



(b) OAR weight (λ) = 0.1

Figure 3.5: Probability of optimal action for an early responding tumor, early responding AOR for 8 weeks, with same initial belief vector; white bar: dose = 1.8 Gy, dotted bar: dose = 2 Gy, dashed bar: dose = 2.2 Gy, grey bar: Termination

Depending on the initial state of the tumor and the initial belief vector, the POMDP policy can recommend different actions. Figure 3 compares how different initial belief vectors influence optimal actions to take over the course of treatment. Two different examples of a belief vector are presented for an early responding tumor and early responding OAR for 8 weeks. Figure 3.6(a) is associated with a belief vector corresponding to a case with a high number of tumor cells, while Figure 3.6(b) is for the case with a low number of tumor cells.

When the belief vector corresponds to a high tumor cell count (Figure 3.6(a)), the POMDP policy is to start the treatment with 2 Gy dose and continue for five weeks, and

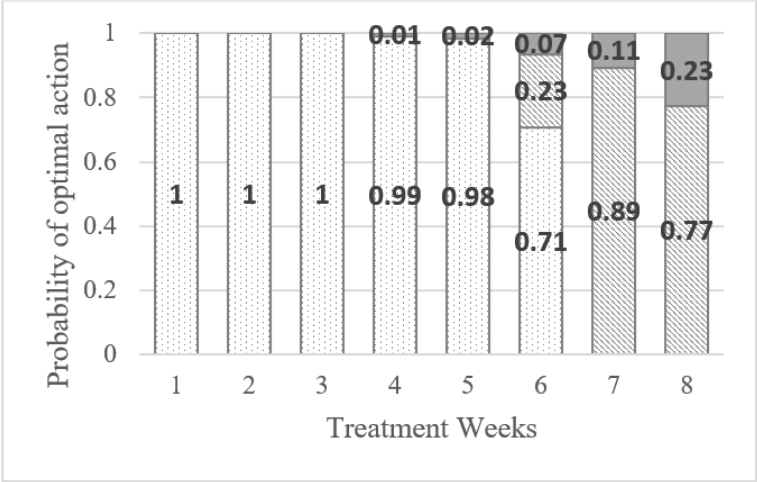
then increase the dose to 2.2 Gy for 23% of scenarios on week six and continue with high dose for all scenarios on weeks seven and eight. When the initial belief vector shows a low tumor cell count (Figure 3.6(b)), the optimal policy is to start with a reduced dose (i.e., 1.8 Gy) in the first week, then increase it to 2 Gy for weeks two to five for all scenarios and increase it to high dose (i.e., 2.2 Gy) starting from week six. In this case, the terminal state was reached for more scenarios starting from week four. For example, Figure 3.6(b) shows that 30% of the scenarios entered the terminal state on week four and treatment was terminated, while in Figure 3(a) only 1% of scenarios declared terminal state.

3.5.4 Comparison to the conventional scheme

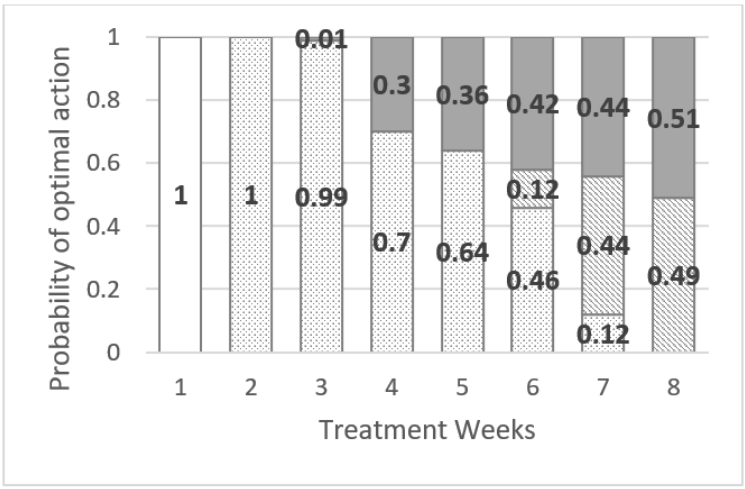
We now compare the optimal POMDP treatment plan with the conventional one. Figure 4.9 presents the expected BED for tumor and OAR for the two plans when starting from the same belief vector. The solid black and solid grey lines represent the expected reward for POMDP optimal plan and the conventional plan, respectively. On week six, the tumor expected BED (EBED) starts to decrease in the conventional plan, while it starts increasing for the POMDP optimal plan. This increase of tumor EBED on the POMDP plan is due to the increase of dose on week six, which also results in a slight increase of OAR EBED. However, because the reward calculated from Equation 3.7 is significantly improved compared to the conventional plan, the POMDP optimal plan is preferred.

Similarly, we compare the tumor control probability (TCP) and normal tissue complication probability (NTCP) for the two above mentioned scenarios. Figure 3.8 presents the results of this comparison.

The optimal scheme results in a significant improvement in cumulative TCP with a slight increase in cumulative NTCP. In other words, while keeping the damage to normal tissue under control, the optimal scheme has a significantly better performance in killing tumor cells.



(a) Initial belief P_1



(b) Initial belief P_2

Figure 3.6: Probability of optimal action for an early responding tumor and early responding OAR for 8 weeks with OAR weight factor (λ) = 1; white bar: dose = 1.8 Gy, dotted bar: dose = 2 Gy, dashed bar: dose = 2.2 Gy, and grey bar: Termination. $P_1 = [0, 0.01, 0.05, 0.14, 0.8]$, $P_2 = [0, 0.49, 0.27, 0.13, 0.11]$.

3.6 Conclusion

The tumor dynamics during the treatment horizon is uncertain because it depends on cell cycle regulations and the gene level activities within the cells. To incorporate this into an optimization model in a fractionated radiotherapy, we have utilized an extended biological model of cell survival for the tumor and OAR to include the four R's of radiation therapy.

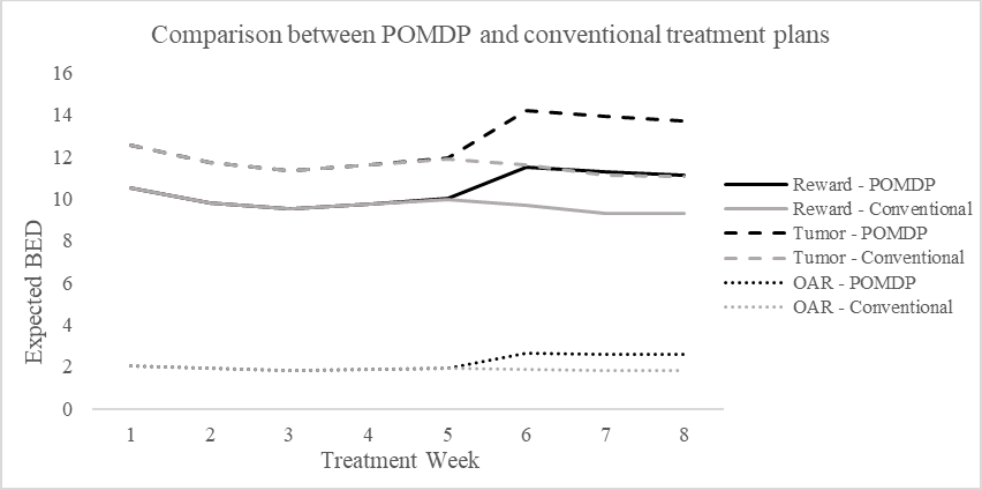
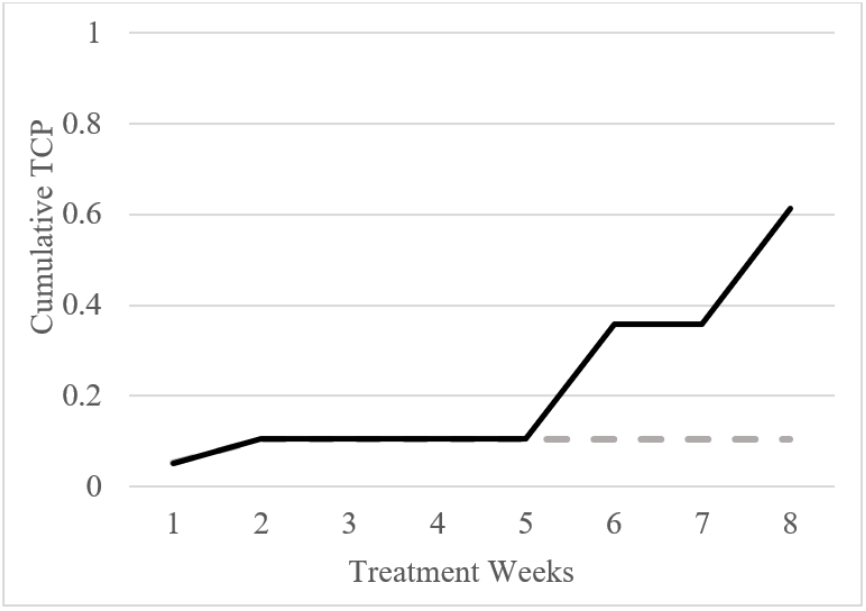
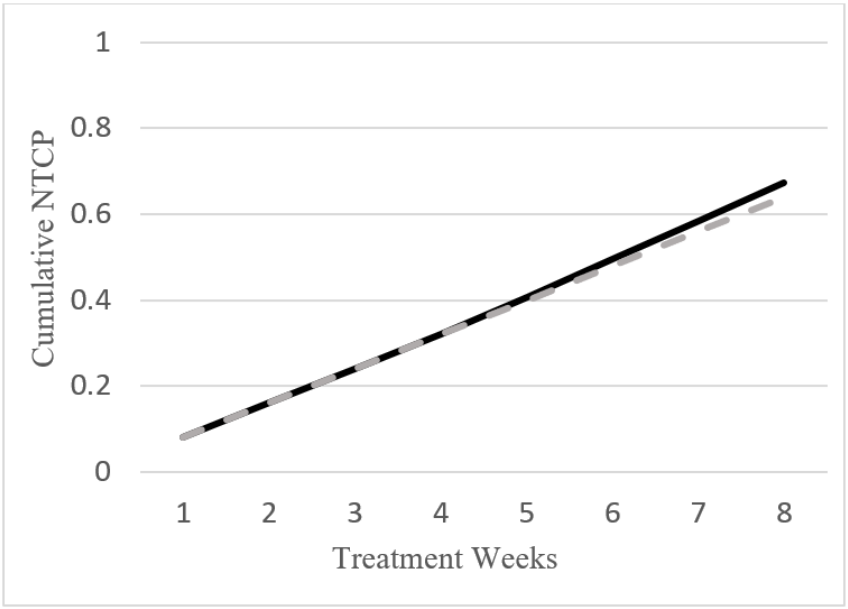


Figure 3.7: Comparison between POMDP and conventional treatment plans.

We have used a Gompertzian growth model for the tumor to explain the dependency of tumor growth rate on the gene level factors. First, we developed a control limit policy based on patient-specific tumor and OAR characteristics with respect to radiation treatment to determine whether a dynamic treatment plan can improve the treatment quality (i.e., tumor damage and OAR sparing) compared to the conventional one. Next, we developed a constrained POMDP framework based on the extended biological model to provide a decision-making policy that maximizes the expected biological equivalent dose (EBED) of tumor while keeping the OAR survival under control. POMDP enables us to incorporate the feedback from the patient into the treatment planning and provide a personalized dynamic treatment policy. Because observing the tumor condition during the treatment horizon is not practical, POMDP offers the probabilistic belief and partial observation probabilities to account for the uncertainty associated with a tumor's condition. Numerical results showed that using a dynamic treatment plan instead of the conventional one can result in higher tumor damage with slight increase in the OAR damage. However, when the tumor tissue was late responding (lower sensitivity to radiotherapy) and the OAR tissue is early responding (higher sensitivity to radiotherapy) the conventional treatment plan was recommended as a conservative plan.



(a) TCP



(b) NTCP

Figure 3.8: (a) Cumulative TCP during treatment weeks for optimal scheme (solid line) and conventional scheme (dashed line). (b) Cumulative NTCP during treatment weeks for optimal scheme (solid line) and conventional scheme (dashed line).

The optimal POMDP policies provided a trade-off between the tumor damage and OAR sparing depending on their tissue types, i.e., early responding tissues or late responding

tissues. The POMDP policy recommended optimal actions based on the updated belief vector at each decision epoch. Due to the Gompertzian tumor growth model, lesser number of tumor cells in the later weeks of treatment cause in an increase in the tumor growth rate. Therefore, the dose increase was recommended at the later weeks of treatment to increase the tumor damage and shorten the treatment length. This approach will be beneficial in OAR recovery as well, because the treatment ends shortly after the dose increase.

Comparing with the conventional treatment plan, we demonstrated that the POMDP policy results in significant improvement in expected biological equivalent dose (EBED) of tumor while causing only a slight increase in EBED of OAR. Therefore, the total expected reward of POMDP policy is significantly higher than that of conventional treatment plan. Similarly, the POMDP policy resulted in a significant improvement in TCP compared to the conventional plan, while increasing the NTCP for a small amount.

As a future work, one can incorporate reinforcement learning methods in POMDP to minimize the human error in identifying the patient feedbacks and utilizing them in the estimation of probabilistic tumor state, i.e., belief vector. In addition, image recognition methods can be used to estimate the exact tumor state and improve the precision of the decision process.

Chapter 4

Personalized and LET-based proton therapy treatment planning considering tumor biological response uncertainty and radiomics

4.1 Introduction

Radiation therapy uses high-energy particles such as photons or protons to destroy cancer cells (Baskar et al., 2012) and is a local treatment that only affects the organs located in the radiation field, commonly called organs at risk (OAR) (Cao et al., 2014; Bai et al., 2019). The aim of the treatment planning for radiation therapy is to reduce the OAR damage while killing the tumor cells.

Proton beams have interesting characteristics that make them popular in recent studies and clinical trials. The dose deposition of a proton beam increases as the beam penetrates deeper in the tissue with a sharp longitudinal dose fall-off at the end of the particle range called a *Bragg Peak* (Krämer & Scholz, 2006). Intensity modulated proton therapy (IMPT) takes advantage of this property to deliver a proton dose conformal to the tumor volume while delivering minimal dose to the OAR.

The biological effectiveness of the proton dose is higher than that of the same photon dose because the linear energy transfer (LET) of a proton particle is lower than that of a photon particle. In the current IMPT treatment planning, it is assumed that protons are 10% more

efficient than the proton particles, but this fixed value neglects any dependency of biological effectiveness on the amount of radiation dose and the endpoint of proton beam (Paganetti, 2014). Relative biological effectiveness (RBE) is a measure for converting a reference dose into the photon dose to yield the same biological effect. Several recent studies have provided IMPT plans based on variable RBE (Chen et al., 2018; Willers et al., 2018), but such plans can result in a heterogenous RBE distribution and cause undesired side effects (Ödén, Eriksson & Toma-Dasu, 2017). However, the proton LET can be controlled and optimized in an IMPT plan to account for proton biological effectiveness while mitigating the risk of having a heterogenous RBE distribution (Unkelbach et al., 2016).

The total prescribed radiation dose is usually delivered in a fractionated treatment plan that divides the prescribed radiation dose into several fractions to reduce the damage to OARs and provide enough time for them to recover from incurred damages. The conventional fractionated plan divides the prescribed dose equally in all treatment sessions. The effectiveness of this plan needs to be investigated because the biological response of each patient during the weeks of treatment can be different (Bibault et al., 2013; Ribba et al., 2006). Marshall et al. (2016) investigated the effect of variable RBE in a fractionated proton therapy plan and outlined that the strong dependence of RBE on proton LET has a significant effect on treatment efficiency, especially, in a variable dose fractionation regimen. Several clinical studies have investigated that shortening clinical fractionation schedules through hypofractionation might be beneficial due to unique characteristics of protons (Habl et al., 2014; Newhauser et al., 2015; Wang et al., 2013). Giantsoudi et al. (2013) utilized LET distribution in a dose-based multi-criteria optimization problem to enhance the fractionated treatment plan.

The biological effectiveness of a radiotherapy treatment plan is commonly characterized by the standard linear-quadratic (LQ) model (Douglas & Fowler, 1976). The extended LQ model, known as LQR (Brenner et al., 1995), was introduced to incorporate the time-dependent impact of four R's in radiation therapy, including: *repair of sublethal damage,*

repopulation, reoxygenation and *redistribution*. The survival fraction in LQR model is calculated by

$$S(D) = exp(-\alpha D - G(\tau_{Rp})D^2 + \frac{1}{2}\sigma^2 G(\tau_{Rp})D^2), \tag{4.1}$$

where D is the dose per fraction, α and β are radio-sensitivity parameters, σ^2 is the variance of Gaussian distribution of α, τ_{Rp} is an average repair time, τ_{Rs} is an average resensitization time which includes both reoxygenation and redistribution effects, and $G(\tau)$ is an expression for a generalized Lea-Catchside time factor (Sachs & Hlatky, 1990) that has the form

$$G(\tau) = \frac{2}{D^2}\int_0^T R(u)[\int_0^u R(w)exp(\frac{-(u-w)}{\tau})dw]du \tag{4.2}$$

and can be calculated for any fractionation plan (Brenner, 2008).

The instantaneous tumor growth rate increases when the number of remaining tumor cells decreases (Brunton & Wheldon, 1978; Ribba et al., 2006; Usher, 1980). Therefore, the tumor growth rate, $\Phi(.)$, is assumed to follow a Gompertzian growth function (Bortfeld et al., 2015) as

$$\Phi(x) = b\ln(\frac{X_\infty}{x}), \tag{4.3}$$

where x is the number of tumor cells, X_∞ is the carrying capacity or the maximum number of tumor cells, and b is a parameter that controls the rate of growth. The biologically equivalent dose (BED) is derived based on the survival fraction obtained from the LQ model and includes the biological effectiveness factors of the radiation dose (i.e. LET and variable RBE) on the organ;

$$BED = -\frac{\ln(S)}{\alpha}. \tag{4.4}$$

The tumor growth rate and biological response are assumed to be constant in the LQ model. However, these factors are shown to be significantly affected by cell cycle regulation,

and may change over time (Barazzuol et al., 2010; Ribba et al., 2006). The tumor specifications including density and size can be incorporated in treatment planning using Markov Decision Process (MDP). MDP is a stochastic dynamic programming technique commonly used for sequential decision-making. In the MDP framework, the system is in a particular state at each time point and a feasible action is selected to optimize the performance of the system for its remaining lifetime. Once the decision is made, the system evolves into a next state with a transition probability. The solution to an MDP is a policy. A *stationary Markovian policy* does not depend on time and the history of states and action in its choice of action. A *deterministic policy* prescribes a specific action in a state, while a *randomized policy* provides a probability distribution over the set of actions in a state.

To incorporate MDP framework in the problem of frationated radiotherapy planning, the tumor specifications (i.e., state) should be observed or predicted. In clinical practice, a team of physicians contour the tumor and OAR structures on the CT image. Because the contouring process is time consuming and expensive, repeating it frequently to observe tumor specifications is not practical. However, the CT images contain important information âĂŞ *Radiomics*, that can offer potential aid in diagnosis, prognosis and prediction of response to treatment of cancer (Gillies, Kinahan & Hricak, 2015; Lambin et al., 2012). Radiomics include quantitative image features based on image texture and intensity, and tumor shape and volume, and can be used in cancer treatment planning to improve the efficiency of the planning process while providing better information about the tumor (Gillies, Kinahan & Hricak, 2015).

Upon availability of pre-treatment CT images from a sufficient number of patients with the same cancer type, machine learning methods can be used to train a multinomial classification model that can predict tumor specifications using the radiomics data (Parmar et al., 2015; Lambin et al., 2012). The random forest method (Liaw, Wiener et al., 2002) is a supervised learning algorithm that can be used for both regression and classification in machine learning problems. This method selects random samples from a training set repeatedly (bagging) with replacement, and builds several trees from those samples. The

final output is obtained by aggregating over all the trees. Random forest is beneficial for relatively small to very large data sets.

This study aims to provide a decision-making policy for a dynamic and personalized fractionated proton therapy treatment plan. Such policy chooses the optimal IMPT plan and dose (action) to be delivered based on the tumor cell count (state). The tumor state at each decision epoch is predicted by applying a trained random forest classifier to the radiomics data from the new CT image. A constrained Markov Decision Process (MDP) framework is used to obtain the dynamic decision-making policy that deals with uncertainty of tumor response and growth rate and maximizes the biological effect to the tumor while controlling the total amount of dose delivered to the OAR. The MDP action set is obtained by optimizing the proposed LET-based IMPT model for different prescribed doses per fraction. Such a model incorporates the extended LQR model to account for the time dependent factors affecting the biological response, including repair and growth of cells. The LET-based IMPT model ensures the homogeneity of BED distribution over the tumor volume in all potential IMPT plans. This model may enable us to increase the tumor dose level without adding more damage to the OARs and, consequently, decrease the treatment duration. The results are provided and compared with that of a fixed RBE IMPT plan for two clinical cases of prostate cancer and pediatric ependymoma.

4.2 Methodology

The radio-sensitivity parameters α and β for protons depend not only on the tissue type, but also on the energy of irradiated particle deposited in the tissue which is referred to as LET. In-vitro studies on cell cultures have shown that α_p is a linear function of the dose-averaged LET (LET_d) and β_p is a constant (Wedenberg, Lind & Hårdemark, 2013),

$$\frac{\alpha_p}{\alpha_x} = 1 + h\frac{LET_d}{(\alpha/\beta)_x} \tag{4.5}$$

and

$$\frac{\beta_p}{\beta_x} \approx 1, \tag{4.6}$$

where, h is a free parameter of the expression. The value of parameter h that best fitted the experimental $\frac{\alpha_p}{\alpha_x}$ data was found to be 0.434 ± 0.7 Gy mm/keV mm/keV (Paganetti, 2014). This linear regression works fine in the LET region below 30 $keV/\mu m$, which is sufficient for the range of proton energy used in IMPT. This model provides the required dependency of biological response on LET. However, the LET is not constant along the central axis of a proton beam. In IMPT treatment planning, for a set of J beam spots irradiating the patient, D_{ij} denotes the dose deposition in voxel i per unit fluence of beamlet $j = 1, 2, \ldots, J$. the total dose delivered to voxel $i = 1, 2, \ldots, I$ is

$$D_i = \sum_{j=1}^{J} D_{ij} w_j, \tag{4.7}$$

where, w_j is the weight of beamlet j. Accordingly, the energy deposition (LET) in voxel i by beamlet j is denoted by L_{ij}. The dose-averaged LET is calculated as

$$LET_d = \frac{\sum_{j=1}^{J} L_i j D_i j w_j}{\sum_{j=1}^{J} D_{ij} w_j} \tag{4.8}$$

and, $LET_d = 0$ for $D_i = 0$. The BED for proton over n equal fractions is defined as (Jones & Dale, 2000)

$$BED_p = nD_p \frac{\alpha_p}{\alpha_x} + n \frac{\frac{\beta_p}{\beta_x}}{\frac{\alpha_p}{\alpha_x}} D^2. \tag{4.9}$$

Thus, proton BED for voxel i over n equal fractions can be obtained by by replacing Equations (4.5, 4.6, 4.7, 4.8) in Equation 4.9 as

$$BED_p^i(w_j) = n \sum_{j=1}^{J} D_{ij} w_j + \frac{nh}{(\alpha/\beta)_x} \sum_{j=1}^{J} L_{ij} D_{ij} w_j + n \frac{(\sum_{j=1}^{J} D_{ij} w_j)^2}{(\alpha/\beta)_x}. \tag{4.10}$$

The objective function for IMPT planning problem in this paper is to minimize deviation from prescribed BED on target while minimizing extra BED on the OAR and maintaining minimum and maximum dose constraints in the target. This objective function is shown in Equation (4.11) and results in a homogeneous BED and dose distribution in the target structure,

$$\min_{w} \sum_{i \in I_{Target}} |BED_p^i(w_j) - BED_{pres}^i| + \sum_{i \in I_{OAR}} BED_p^i(w_j). \tag{4.11}$$

To develop an MDP framework for optimizing the fractionated proton therapy treatment plan, we assume that the physician makes a decision on the amount of radiation dose to be delivered at the beginning of each treatment week. This decision is based on the tumor conditions such as cell count, volume or density. We incorporate tumor condition in the MDP framework as the MDP state, $s \in \mathcal{S}$, which is defined as the number of tumor cells. A healthy state is considered as the terminal state of the system in our model. The set of MDP actions, $\mathcal{A}$, in this model includes possible radiation doses: (i) the conventional regime dose, d, (ii) the dose increased by a specific amount, $d + \Delta_1$, and (iii) the dose decrease by a specific amount, $d - \Delta_2$. Definitions of parameters and variables used to formulate a constrained MDP are listed in Table 4.1.

Table 4.1: Definition of notations

Notation	Description
$\mathcal{S}$	Set of possible tumor states
$\mathcal{A}$	Set of possible radiation doses
$\mathcal{T}$	Transition probability
$\mathcal{R}$	The immediate reward function
$\mathcal{C}$	The immediate cost function
S^t	State of the tumor at time t
A^t	Decision made at time t
q	Constraint upper bound
T	End of treatment
V_t	Value function for the MDP policy at time t

The optimal solution of Equation (4.11), w^*, represents an optimal IMPT plan for a given prescription dose per fraction. The average BED, $\overline{BED}$, over the volume of a structure for

any w^* during n days of treatment delivery with the same amount of radiation dose, D, can be calculated as

$$\overline{BED}(s=x,a=D)=\frac{1}{|I|}\sum_{i}\Big[n\sum_{j=1}^{J}D_{ij}w_j^*+\frac{nh}{(\alpha/\beta)_x}\sum_{j=1}^{J}L_{ij}D_{ij}w_j^*+$$
$$n(\frac{G_{Rp}}{(\frac{\alpha}{\beta})_x}-\frac{1}{2}\frac{\sigma^2}{\alpha_x}G_{Rs})(\sum_{j=1}^{J}D_{ij}w_j^*)^2]-\frac{\phi(x)}{\alpha_x}. \tag{4.12}$$

Then, the objective of fractionated treatment plan is to maximize the target BED and minimize the OAR the BED by defining the MDP reward function as

$$\mathcal{R}(s,a)=\overline{BED}_{I_{Target}}(s,a)-\lambda.\overline{BED}_{I_{OAR}}(s,a), \tag{4.13}$$

where, λ is the OAR weight factor. To avoid delivering extra radiation dose to the OAR, a constraint is imposed to control the total $\overline{BED}$ delivered to the OAR as

$$\mathcal{C}(s,a)=\overline{BED}_{OAR}(s,a)\leq q. \tag{4.14}$$

Determination of the transition probability function, $\mathcal{T}(s'\mid s,a)=P(S^{t+1}=s'\,|S^t=s,A^t=a)$, is beyond the scope of this paper and can be obtained from an expert's opinion.

The constrained MDP problem can be stated as

$$\begin{aligned}&\max\quad E(\sum_{t=1}^{T}\gamma\mathcal{R}(s,a)),\\&s.t.\\&E\left[\sum_{t=1}^{T}\gamma\mathcal{C}(s,a)\right]\leq q,\\&a\in\mathcal{A},\end{aligned} \tag{4.15}$$

where, $\gamma\in[0,1]$ is the discount factor. We denote $b(s)$ as the initial probability that the current state is $s\in\mathcal{S}$. Defining $y(s,a)$ as the total expected discounted number of times action $a\in\mathcal{A}$ is executed in state $s\in\mathcal{S}$, a randomized stationary policy for a constrained

MDP can be obtained from the following LP model (Altman, 1999; Ross, 1989)

$$
\begin{aligned}
&\max \quad \sum_{s\in\mathcal{S}}\sum_{a\in\mathcal{A}} \mathcal{R}(s,a)y(s,a), \\
&s.t. \\
&\sum_{s\in\mathcal{S}}\sum_{a\in\mathcal{A}} \mathcal{C}(s,a)y(s,a) \leq q, \\
&\sum_{s\in\mathcal{S}}\sum_{a\in\mathcal{A}} (\delta(s,s') - \gamma\mathcal{T}(s'|s,a))y(s,a) = b(s'), s' \in \mathcal{S}, \\
&\sum_{s\in\mathcal{S}}\sum_{a\in\mathcal{A}} y(s,a) = 1, s \in \mathcal{S} \\
&y(s,a) \geq 0, s \in \mathcal{S}, a \in \mathcal{A}.
\end{aligned} \tag{4.16}
$$

The function $\delta(s,s')$ is one if $s = s'$ and is 0 otherwise. Let $y^*(s,a) > 0, a \in \mathcal{A}$. Then $\pi^*(s,a) = y^*(s,a)/\sum_{a\in\mathcal{A}} y^*(s,a)$.

An in-house program is developed Matlab R2018b (The Mathworks, Natick, MA) to extract radiomics data from the several CT images of the same cancer site. The radiomics data include the number of tumor voxels and other quantitative features of the CT images. The number of tumor voxels might not be the exact representation of the tumor cell count, but it can be used as an approximation of it. Therefore, each cancer case is labeled as belonging to one of the tumor states depending on its number of tumor voxels. The rest of the radiomics data are used as features of the cancer cases. The features data are pre-processed and scaled to avoid dominating feature values.

In machine learning, classification is a supervised learning algorithm that analyzes a labeled training data set and infers a function that can be used for predicting the labels of test data sets. The whole extracted radiomics data are randomly divided into a training set and a test set. The training feature set and labels are fed to a random forest trainer to build a classification model using 10 fold cross validation. The trained model is then applied to the test set to evaluate the accuracy. For each decision epoch, the trained random forest model can be used to predict the tumor state from a new raw CT image of the patient with the same type of cancer and the MDP framework can choose the optimal radiation dose and

IMPT plan for that epoch. The flowchart of the proposed framework is provided in Figure 4.1.

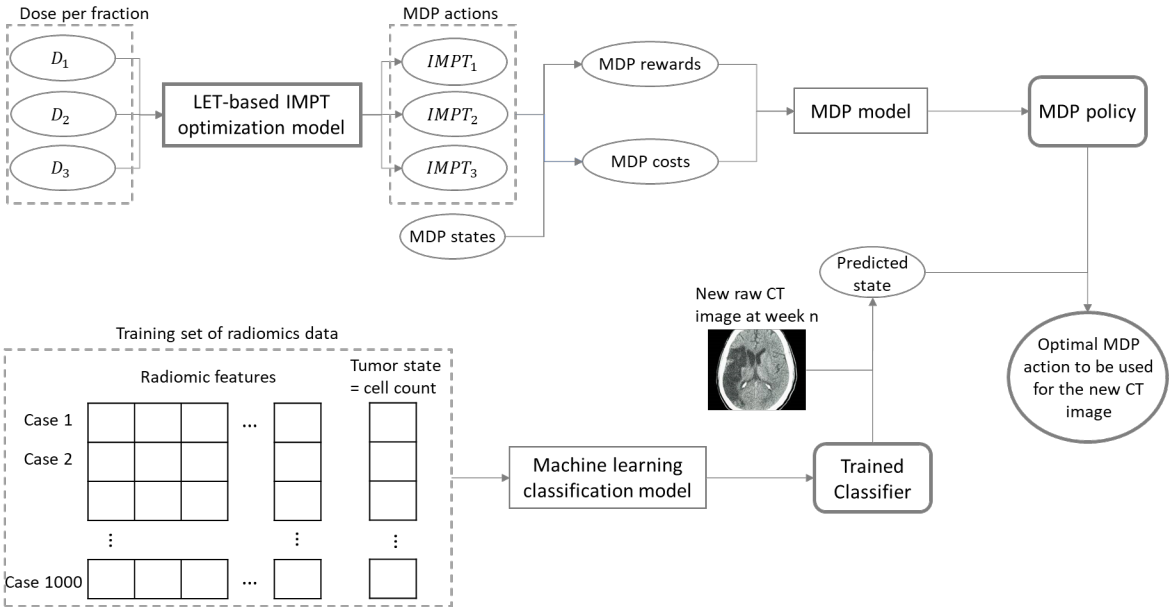


Figure 4.1: personalized LET-based IMPT and fractionation optimization framework

4.3 Numerical Results

In this section, we first present an example of using radiomics in tumor state prediction using random forest method. We then present the numerical results of our proposed LET-based fractionation model for two cancer cases: pediatric head and neck cancer and prostate cancer. The dose and LET deposition matrices and organ structure data are obtained from a cancer center in Houston, TX. The outcomes of our proposed LET-based IMPT model are compared to the benchmark model that optimizes the proton dose distribution based on constant RBE of 1.1. The dose-volume histogram (DVH) of the two models are compared for physical dose, RBE×Dose, and BED. For further evaluation of the treatment plan, tumor control probability (TCP) and normal tissue complication probability (NTCP) are calculated based on equivalent uniform dose (EUD) using the following equations (Gay &

Niemierko, 2007; Lyman, 1985; Warkentin et al., 2004)

$$TCP = \frac{1}{1 + \left(\frac{D_{50}}{EUD}\right)^{4\gamma 50}} \tag{4.17}$$

and

$$NTCP = P = \frac{1}{1 + \left(\frac{D_{50}}{EUD}\right)^{4\gamma 50}}, \tag{4.18}$$

where $EUD = \left(\sum_i v_i D_i^{1/u}\right)^u$ is the equivalent uniform dose, D_{50} and γ_{50} describe the dose and normalized slope at the 50% of the dose-response, D_i is the dose received by the partial volume v_i, and u is a parameter of dose-volume relationship. Table 4.2 provides the dose-response parameters for the PTV and OARs (Levegrün et al., 2002). All experiments were performed on a Linux computer with an Intel Quad 3.0 GHz processor and 396 GB RAM. The NLP model was solved using CONOPT3 version 3.15N and the LP model was solved using CPLEX version 12.6.0.0.

Table 4.2: Dose-response parameters

Cancer case	Organ	D_{50}	γ_{50}	u
Prostate	PTV	76.5 Gy	2.9	1
	Rectum	75 Gy	2.9	0.12
	Bladder	72 Gy	2.9	0.5
Pediatric ependymoma	PTV	64 Gy	3	1
	Brain	60 Gy	3	0.2
	Brainstem	65 Gy	3	0.14

4.3.1 Tumor state prediction using radiomics

The data for this section are obtained from a Medical Image Computing and Computer Assisted Intervention (MICCAI) grand challenge. The University of Texas MD Anderson Cancer Center (MDACC) provided a dataset of CT images of 315 patients with a total 995 head and neck cancer volumes. This dataset is used for training a classifier model that can predict the tumor state from any CT image of the same cancer type. Therefore, the

specifications of the treatment delivered to the patients will not have an impact on the model training. A total 89 radiomics features as well as the number of tumor voxels were obtained using an in-house program. The tumor voxels were categorized into four states based on the frequency of voxel count to maintain relevantly balanced class definition Figure 4.2.

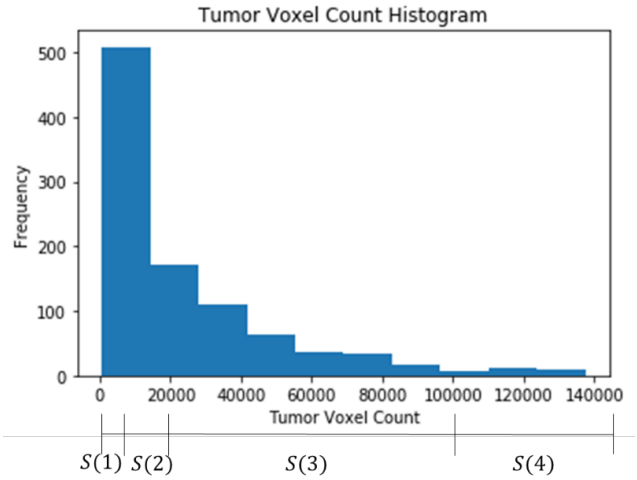
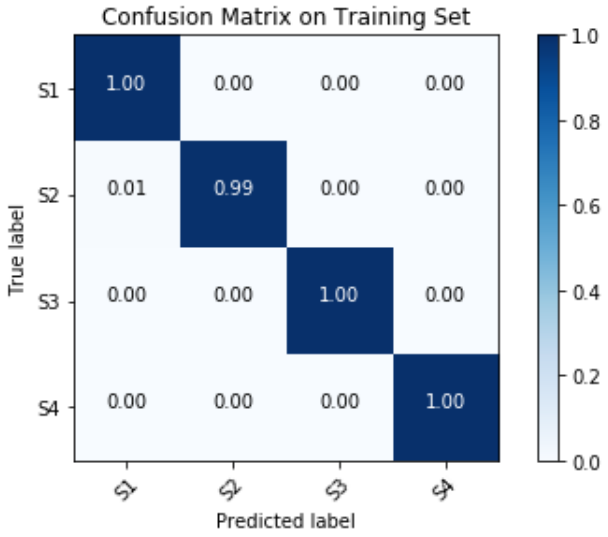
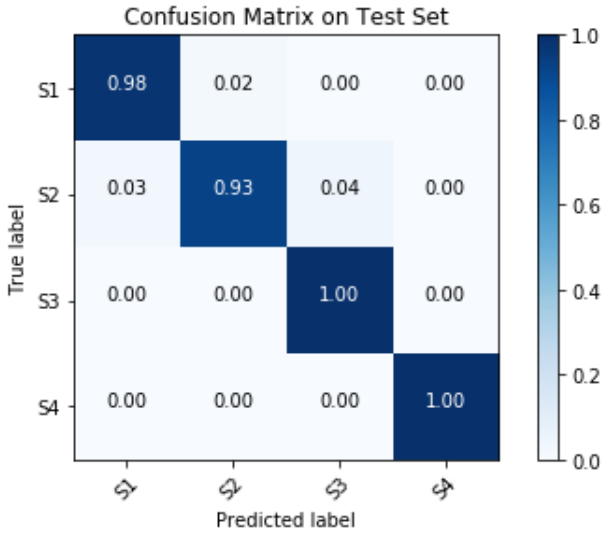


Figure 4.2: Voxel count frequency and tumor state

The whole data set was randomly divided into a training set (85%) and a test set (15%). A random forest model was trained to make a prediction on both the training set and the test set. A confusion matrix was used to show the prediction performance of the random forest model. A confusion matrix is a table that visualizes the percentage of predicted class versus actual class in the classification problems. Figure 4.3 presents the confusion matrices resulting from random forest on the training set and the test set. It can be observed that using radiomics data and machine learning methods, the tumor state can be predicted with a high accuracy for head and neck cancer. For the cases where the prediction accuracy is not good enough, radiomics can still be used to provide the confusion matrix as an observation probability function in a partially observable MDP.



(a) On the training set



(b) On the test set

Figure 4.3: Performance of the random forest model predicting tumor state using radiomics data.

4.3.2 Case study 1: Prostate cancer

Table 4.3 presents the treatment requirements and radiobiological characteristics of the structures for a prostate cancer IMPT plan. In this case, tumor and OARs have α/β of 3 Gy^{-1} so they are early responding tissues, i.e. they show an increased response to radiotherapy, repair the sub-lethal damage in a shorter period of time and grow faster.

Table 4.3: PTV and OAR characteristics for prostate cancer

Structure	Structure description	Dose constraint	α/β	τ_{Rp}
PTV	Target	Prescription: 70 Gy	3 Gy^{-1}	30 minutes
Rectum	OAR	Upper limit: 65 Gy	3 Gy^{-1}	30 minutes
Bladder	OAR	Upper limit: 65 Gy	3 Gy^{-1}	30 minutes

The prescribed total dose of 76 Gy is usually delivered in 38 fractions with a fraction size of 2 Gy. We investigate the possibility of altering this amount at the beginning of each week of treatment. As alternative actions, we consider changing this amount by ±0.2 Gy. The LET-based IMPT model was used to optimize the dose and energy distribution for three target doses 1.8, 2 and 2.2 Gy. The resulting physical dose and RBE×Dose distributions for the three target doses are provided in Figures 4.4, 4.5 and 4.6. The outcomes of the LET-based IMPT model are compared to those of a fixed-RBE IMPT model.

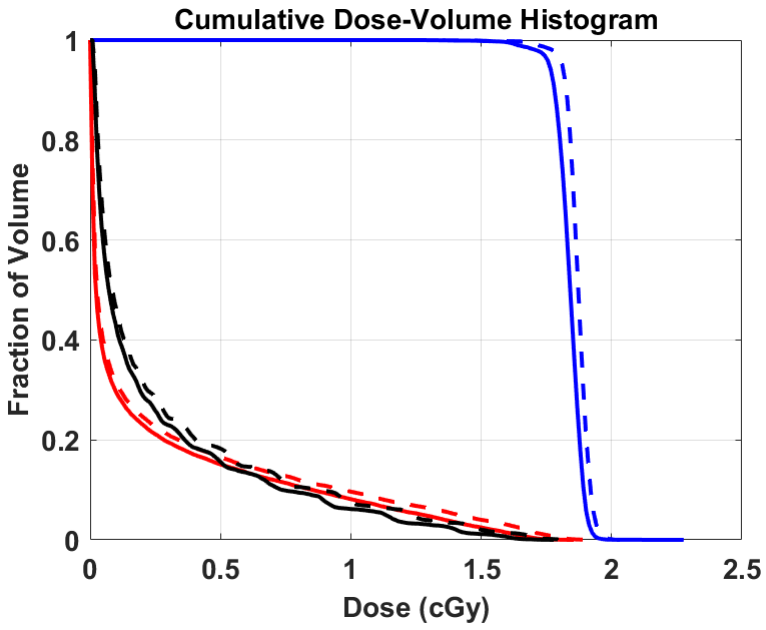
The figures show that the plan resulting from the LET-based IMPT model provides a more homogeneous RBE×Dose to the PTV with a similar RBE×Dose distribution over the OARs as the fixed-RBE IMPT model.

Figure 4.7 shows the comparison of BED distribution over structure for the three defined doses. The LET-based IMPT model provides a significant improvement on the tumor DVH. The three optimized IMPT models are considered as the possible actions in the MDP framework.

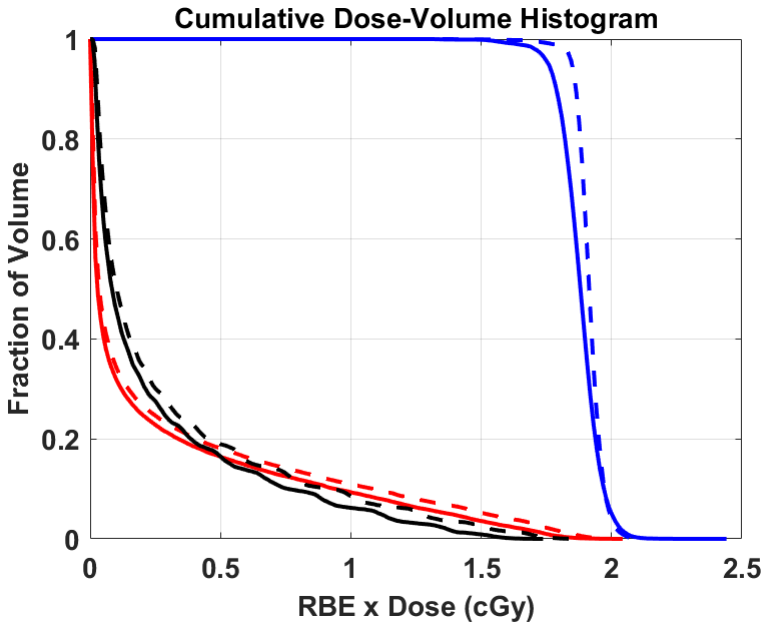
To evaluate the quality of the LET-based IMPT plan, we compute TCP and NTCP resulting from the optimal plan assuming the same plan will be used over 35 treatment fractions and compared it to that of the fixed-RBE IMPT plan. The TCP and NTCP measures for the tumor and OARs resulting from each model are presented in Table 4.4. It can be observed that with similar NTCP for OARs, the tumor TCP has been significantly improved with the LET-based IMPT model.

Table 4.4: Prostate tumor TCP and NTCP of OARs in Fixed-RBE IMPT and LET-based IMPT models

	TCP	Rectum NTCP	Bladder NTCP
Fixed-RBE IMPT	0.62	1.2E-3	1.5E-3
LET-based IMPT	0.67	3.6E-3	6.4E-3

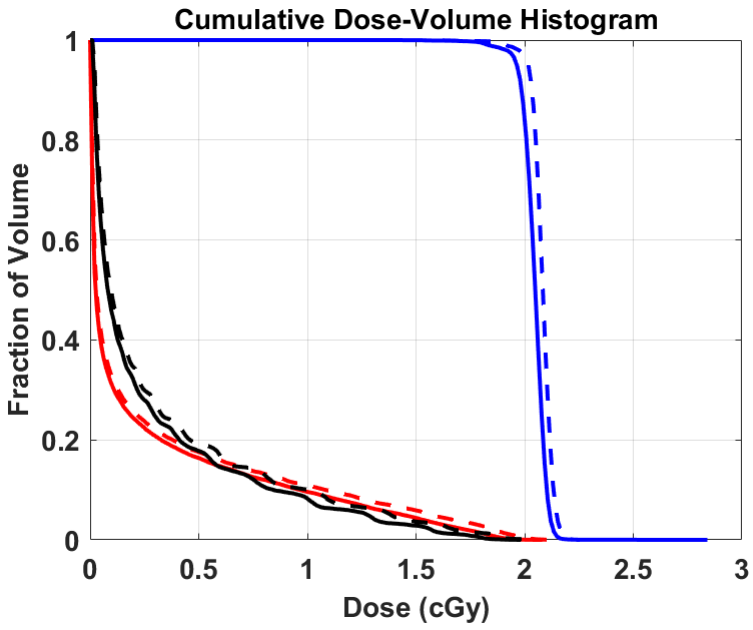


(a) Physical dose

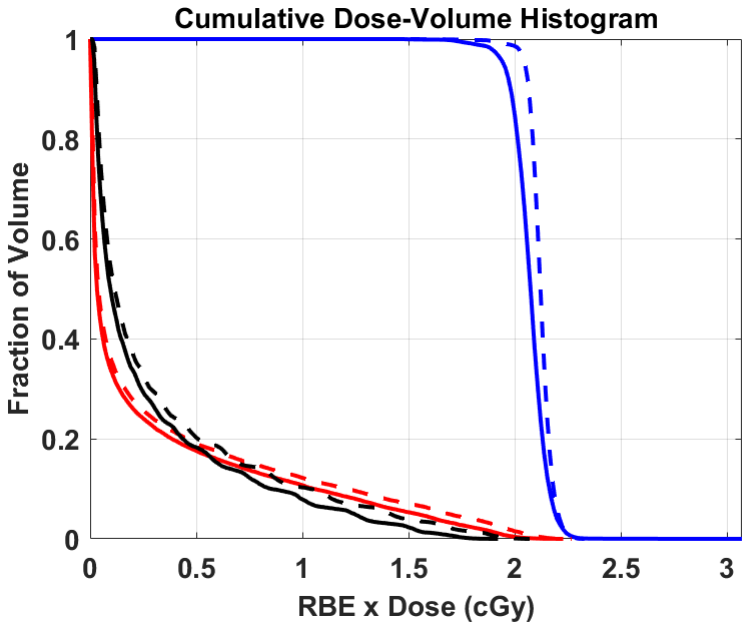


(b) Variable RBE×dose

Figure 4.4: Dose-volume histogram for (a) physical dose and (b) variable RBE×dose for target dose 1.8 Gy and target BED 2.88 Gy. Solid line: benchmark model, dashed line: LET-based IMPT model. Blue line: target volume, red line: rectum and black line: bladder.

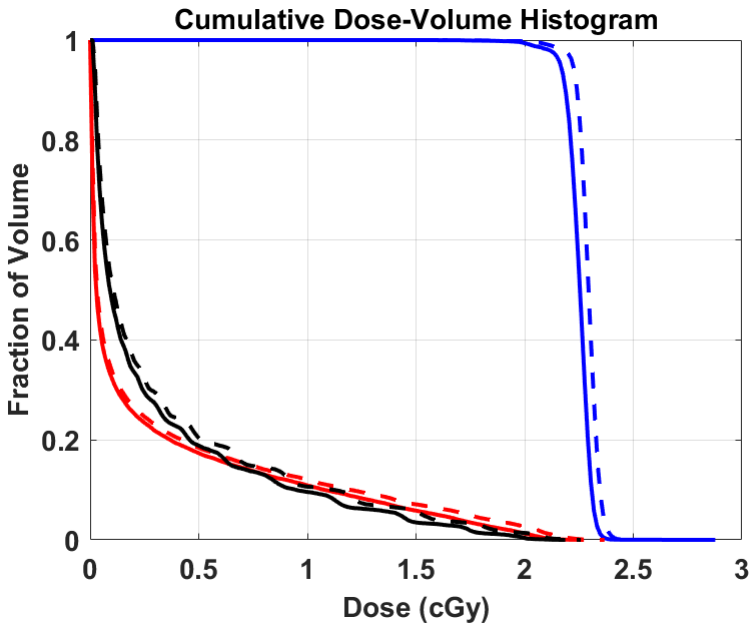


(a) Physical dose

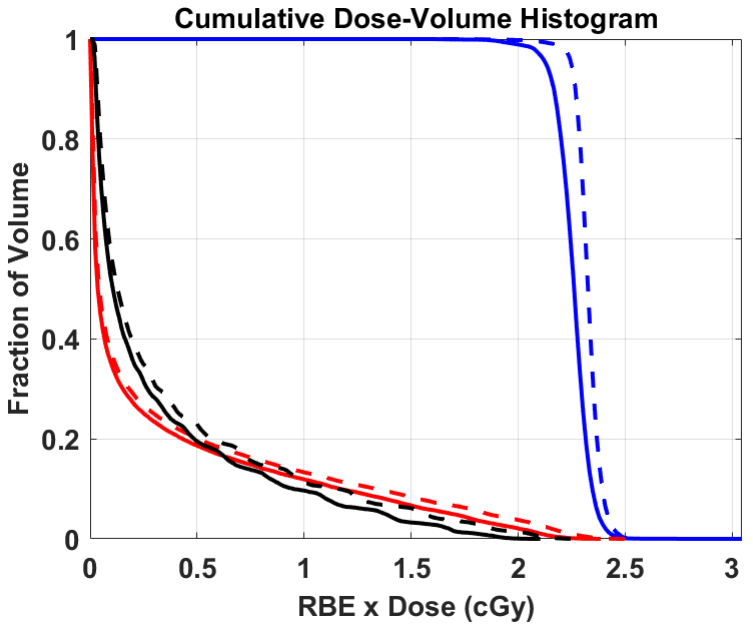


(b) Variable RBE×dose

Figure 4.5: Dose-volume histogram for (a) physical dose and (b) variable RBE×dose for target dose 2.0 Gy and target BED 3.35 Gy. Solid line: benchmark model, dashed line: LET-based IMPT model. Blue line: target volume, red line: rectum and black line: bladder.

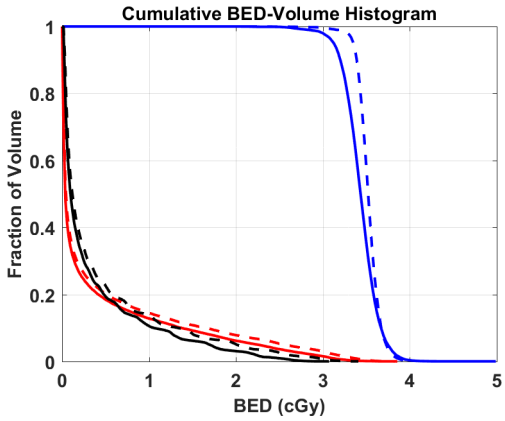


(a) Physical dose

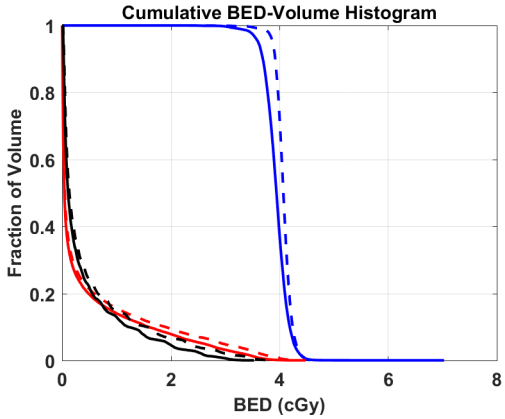


(b) Variable RBE×dose

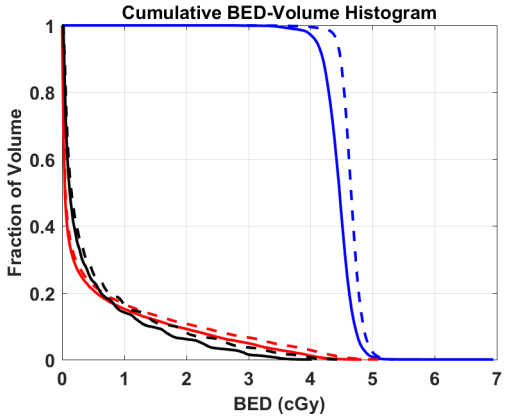
Figure 4.6: Dose-volume histogram for (a) physical dose and (b) variable RBE×dose for target dose 2.2 Gy and target BED 3.8 Gy. Solid line: benchmark model, dashed line: LET-based IMPT model. Blue line: target volume, red line: rectum and black line: bladder.



(a) Target dose 1.8 Gy and target BED 2.88 Gy



(b) Target dose 2.0 Gy and target BED 3.35 Gy



(c) Target dose 2.2 Gy and target BED 3.8 Gy

Figure 4.7: BED-volume histogram for (a) target dose 1.8 Gy and target BED 2.88 Gy, (b) target dose 2.0 Gy and target BED 3.35 Gy, and (c) target dose 2.2 Gy and target BED 3.8 Gy. Solid line: benchmark model, dashed line: LET-based IMPT model. Blue line: target volume, red line: rectum and black line: bladder.

The average BED delivered to tumor and OAR structures was calculated using Equation (4.12) and the MDP was solved using Equation (4.16). The optimal randomized MDP policy for fractionated treatment planning suggested to use a(3) for states one, two and three, but for state four which has the highest tumor cell count, a(1) was suggested in 79% of the cases. The reason for such results is that tumor growth rate is high when the tumor cell count is low, and vice versa. Therefore, in state four, the tumor growth rate is at its lowest and a lower dose can be delivered to the tumor. But in other states, the tumor growth rate increased and a higher dose was required to destroy the tumor.

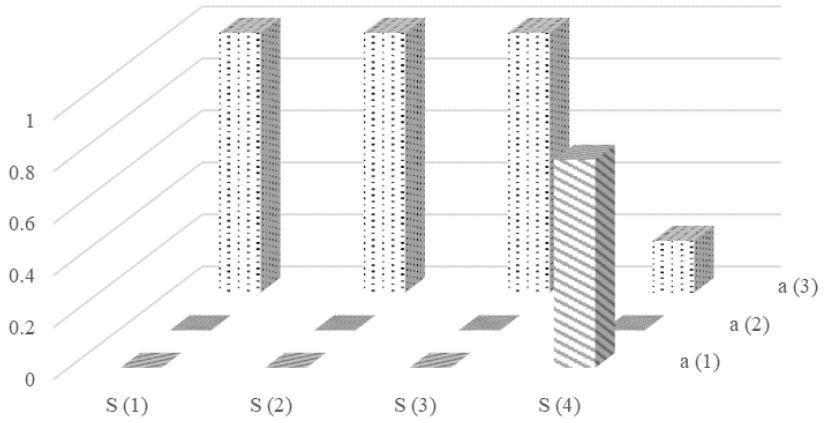


Figure 4.8: MDP randomized policy for fractionated treatment planning of prostate cancer

4.3.3 Case study 2: Pediatric ependymoma

In this section, the proposed model is applied on a pediatric ependymoma case with plan requirements and radiobiological characteristics as shown in Table 4. The tumor structure in this case has a relatively high α/β, causing the tumor to be less sensitive to radiotherapy compared to OARs. In other words, the tumor structure is late responding and has a longer average repair time and slower growth rate.

In a conventional fractionated plan, a prescribed dose of 54 Gy is delivered to the patient over 30 fractions with a fraction size of 1.8 Gy. Target doses 1.6 and 2.0 Gy are considered

Table 4.5: PTV and OAR characteristics for prostate cancer

Structure	Structure description	Dose constraint	α/β	τ_{Rp}
PTV	Target	Prescription: 54 Gy	10 Gy^{-1}	1.9 hours
Brain	OAR	Upper limit: 54 Gy	3 Gy^{-1}	30 minutes
Brainstem	OAR	Upper limit: 54 Gy	3 Gy^{-1}	30 minutes

as the potential alterations to the conventional plan. This case is different from the case before because the radio-sensitivity of OARs is greater than that of tumor. This means OAR has greater biological response than tumor for the same amount of dose. As shown below, brainstem receives less physical dose than tumor in Figure 4.9(a), but more RBE×Dose in Figure 4.9(b).

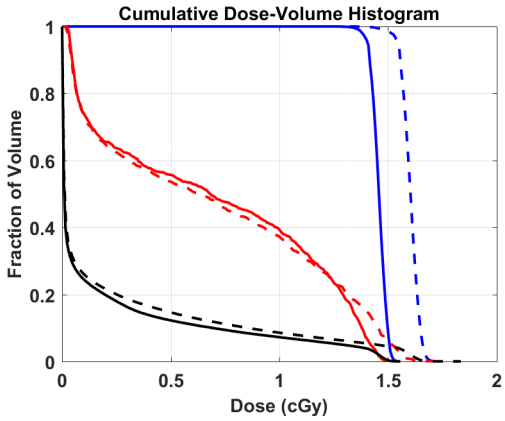
Similar results are shown in Figures 4.10 and 4.11 for target dose 1.8 and 2.0 Gy. The distribution of dose and RBE×Dose over tumor structure is improved significantly by our proposed LET-based IMPT model. Also, RBE×Dose distribution over brainstem has increased slightly in the new model. Figure 4.12 shows that brainstem BED is significantly exceeding tumor BED but given their radiobiological differences such result is inevitable. As in Figure 4.12 for a slight increase of brainstem BED, LET-based IMPT improves tumor BED.

Table 4.6 presents tumor TCP and OAR NTCP in our proposed model and the threshold model. TCP is significantly increased with LET-based IMPT while NTCP of brain is remaining similar and NTCP of brainstem slightly decreasing.

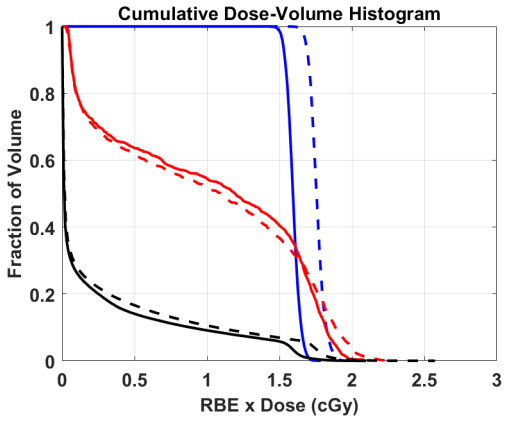
Table 4.6: Pediatric ependymoma tumor TCP and NTCP of OARs in Fixed-RBE IMPT and LET-based IMPT models

	TCP	Rectum NTCP	Bladder NTCP
Fixed-RBE IMPT	0.45	1.2E-3	1.5E-3
LET-based IMPT	0.72	3.6E-3	6.4E-3

Because OARs have greater radio-sensitivity compared to tumor, randomized MDP policy was obtained for two different weight of OAR, i.e. $\lambda = 1$ and $\lambda = 0.7$. If the decision-maker decides that OAR must have a weight factor equal to tumor, the optimal treatment policy suggests a(1) in states two, three and four, but a(2) in 82% of the cases only for state one to meet the total prescription amount of dose (Figure 4.13). However, if a smaller



(a) Physical dose



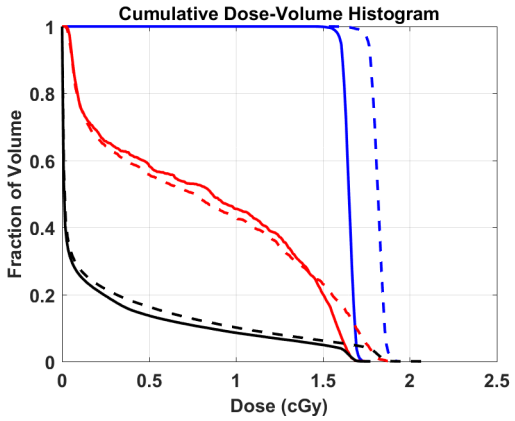
(b) Variable RBE×dose

Figure 4.9: Dose-volume histogram for (a) physical dose and (b) variable RBE×dose for target dose 1.6 Gy and target BED 1.86 Gy. Solid line: benchmark model, dashed line: LET-based IMPT model. Blue line: target volume, red line: brainstem and black line: brain.

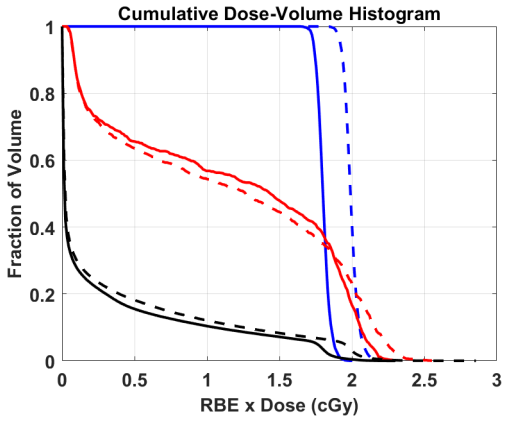
weight factor is considered for OARs, the optimal policy suggests a(1) for 82% of the cases and a(3) for 18% of the cases in state 4. As the tumor cell count decreases in states three to one, the suggested action eventually switched to a(3) to destroy the late responding tumor (Figure 4.14).

4.4 Conclusion

The goal of this paper was to improve the conventional fractionated proton therapy treatment plan by considering the variable biological effectiveness of proton beam and the



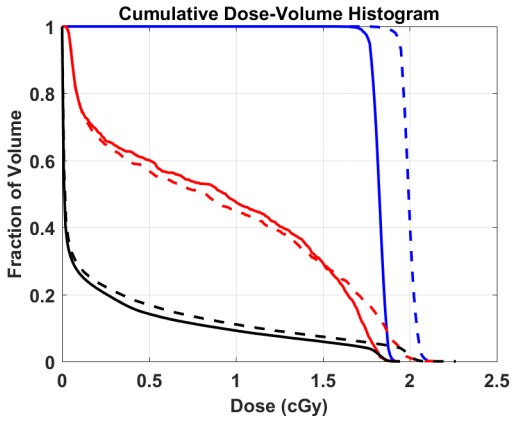
(a) Physical dose



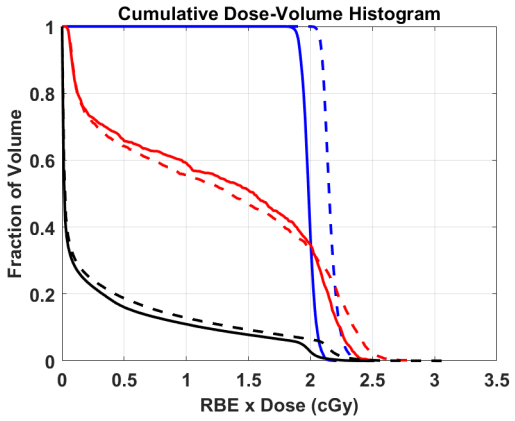
(b) Variable RBE×dose

Figure 4.10: Dose-volume histogram for (a) physical dose and (b) variable RBE×dose for target dose 1.8 Gy and target BED 2.12 Gy. Solid line: benchmark model, dashed line: LET-based IMPT model. Blue line: target volume, red line: brainstem and black line: brain.

uncertain biological response of the tumor to the radiation. An LET-based model was proposed for the problem of IMPT plan optimization. In this model, the BED was calculated based on tumor biological factors and proton LET. Because the tumor growth rate decreases when the number of tumor cells increases, we used a Gompertzian growth function for tumor. An MDP framework was utilized for the problem of fractionated treatment plan. The outcome of this framework was a randomized decision policy that can provide the optimal action for a given tumor state. The set of actions in this problem was obtained by solving the LET-based model for three amounts of target dose per fraction. The state of



(a) Physical dose

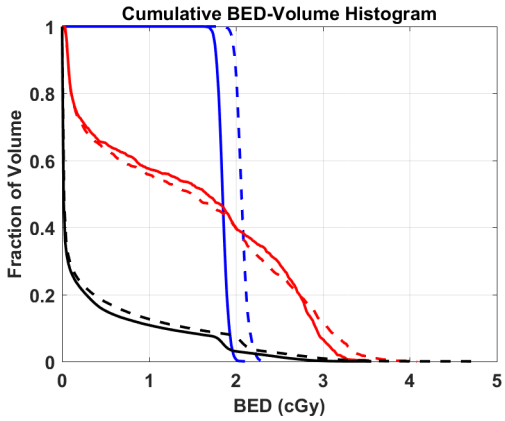


(b) Variable RBE×dose

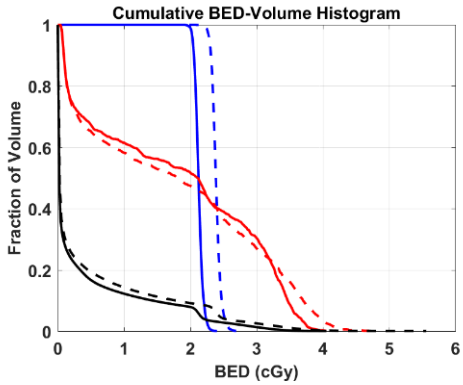
Figure 4.11: Dose-volume histogram for (a) physical dose and (b) variable RBE×dose for target dose 2.0 Gy and target BED 2.4 Gy. Solid line: benchmark model, dashed line: LET-based IMPT model. Blue line: target volume, red line: brainstem and black line: brain.

tumor was defined as the number of tumor cells. Tumor state at each point of time can be predicted from a new CT image by a random forest classifier that is trained based on radiomics data.

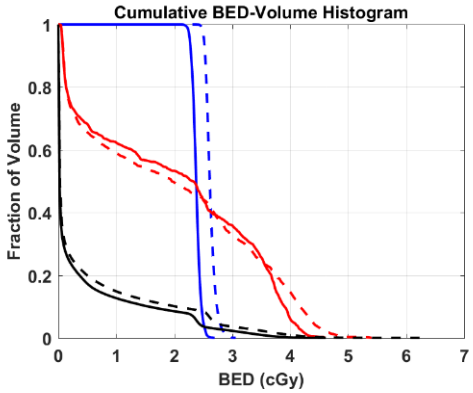
The proposed models were applied to two cases of prostate cancer and pediatric ependymoma. Optimal fractionated treatment policy and optimal IMPT plan were provided for each case. In the prostate cancer case, tumor and OARs had early responding tissue type. But in the pediatric ependymoma example, tumor was late responding, and OARs were



(a) Target dose 1.6 Gy and target BED 1.86 Gy



(b) Target dose 1.8 Gy and target BED 2.12 Gy



(c) Target dose 2.0 Gy and target BED 2.4 Gy

Figure 4.12: BED-volume histogram for (a) target dose 1.6 Gy and target BED 1.86 Gy, (b) target dose 1.8 Gy and target BED 2.12 Gy, and (c) target dose 2.0 Gy and target BED 2.4 Gy. Solid line: benchmark model, dashed line: LET-based IMPT model. Blue line: target volume, red line: brainstem and black line: brain.

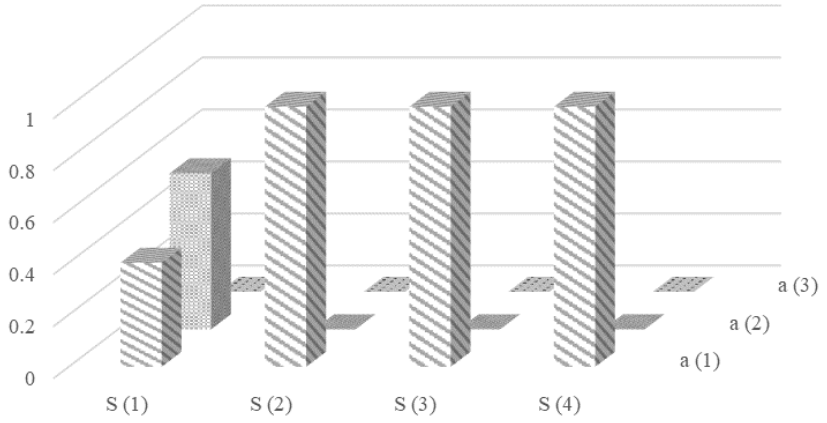


Figure 4.13: MDP randomized policy for fractionated treatment planning with $\lambda = 0.7$ for all OARs

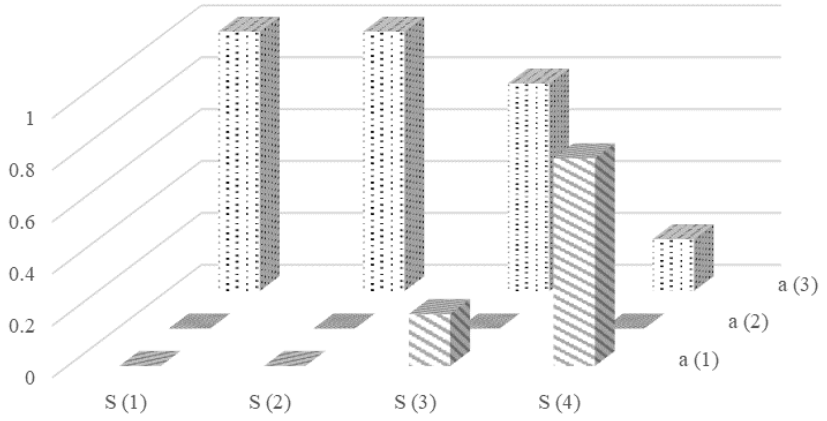


Figure 4.14: MDP randomized policy for fractionated treatment planning with $\lambda = 1$ for all OARs

early responding. Dose-volume histogram of physical dose, RBE×dose and biological equivalent dose compared the quality of IMPT plans obtained from our proposed model and a fixed-RBE threshold model. TCP and NTCP were used as additional quality measures to assess the proposed model. The LET-based IMPT model outperformed the threshold model in general, providing significantly increases TCP with similar NTCP.

Randomized MDP policies were provided for both cases. In the prostate cancer example,

smaller dose target was suggested for higher number of tumor cells because the tumor growth is at its lowest rate. But as the tumor cell count decreases, tumor growth rate increases and bigger dose target was suggested to destroy the tumor. For the pediatric ependymoma due to high radio-sensitivity of OARs and low radio-sensitivity of tumor, different OAR weight factors were used to investigate the impact of decision-maker preferences on treatment policy. It was observed that when larger weight factor is considered for OAR, the optimal policy acts conservative and suggests small dose per fraction but for smaller weight factors the optimal policy acts aggressive and suggests large dose per fraction specially when then number of tumor cells is smaller.

The random forest classifier showed a great accuracy in predicting the tumor state. However, even if a machine learning model did not provide high accuracy in predicting the tumor state based on the radiomics data, the resulting confusion matrix can be used as a partial observation in a POMDP framework, which is an extended form of MDP with partially observable states. Therefore, the radiomics data can provide possibility of tracking tumor condition in a personalized treatment planning regimen.

Chapter 5

Addressing respiratory motion in IMPT in a fractionated treatment planning problem

5.1 Introduction

Proton radiotherapy is one of the most advanced and increasingly popular treatment methods for various cancer types, offering significant clinical advantages over photon therapy. Proton particles have a finite range and increase their energy deposition as they move closer to the end of their range, resulting in a dose peak which is referred to as *Bragg peak* (Wilson, 1946). Such properties of proton particles allow for a radiation dose to better conform to the tumor shape and more effectively spare organs-at-risk (OARs) compared to the conventional photon therapy. In particular, in intensity modulated proton therapy (IMPT), the intensity of the proton beamlets can be individually modified to create an interval of the Bragg peaks, referred to as the spread-out Bragg peaks (SOBPs), which further improves sparing of OARs.

A proton therapy treatment is usually delivered via *fractionated treatment plan* i.e., the prescribed radiation dose is delivered over several fractions instead of a one-time treatment delivery. Fractionation helps reduce overall damage to OARs and provides enough time for them to recover from the damage. The quality of a fractionated radiotherapy plan depends on the resulting biological response from the tumor or OARs and is characterized by cell killing (Brenner, 2008; Gerweck, Zaidi & Zietman, 1994). The most common tool for

quantitative prediction of biological response is the linear-quadratic (LQ) model (Fowler, 1989; Thames, 1985). The conventional fractionation scheme divides the prescribed dose equally through all treatment fractions, but taking the biological response into account can improve the quality of a fractionated plan.

In the planning phase of an IMPT treatment, the patient's internal organs are visualized via CT images, which are then used for construction of the treatment plan to be delivered in all irradiation fractions. However, if the position or shape of the organ structures change during the session due to a geometric uncertainty, the outcome of the delivered treatment might be different from the expected outcome. The International Commission on Radiation Units and Measurements (ICRU) considers three sources of geometric uncertainty (Stroom & Heijmen, 2002): patient set-up variation, machine related errors, and organ motion. Patient set-up errors are due to variations in the daily positioning of the patient on the treatment couch; machine related errors, such as beam sizes and gantry angles, are generally considered negligible thanks to advancements in modern radiotherapy equipments; and organ motion can happen both between fractions (known as inter-fractional motion) and within a fraction (called intra-fractional motion). An inter-fractional organ motion can occur due to weight change or variations in organ size, while an intra-fractional organ motion is caused by cardiac action and respiration. Inter-fractional motions are easier to deal with because they need intervention only at the beginning of each fraction. On the other hand, intra-fractional organ motion happens during treatment delivery and changes organ position on a time-scale of seconds to minutes (Bert & Durante, 2011).

Several studies have proposed robust optimization techniques for treatment planning for IMRT and IMPT to address such geometric uncertainties. Bortfeld et al. (2002) simulated organ motion assuming a sinusoidal motion function and investigated the overall effect of respiratory motion on IMRT treatment quality based on mean and variance of dose on organs' volume. Later, Bortfeld et al. (2008) used a set of probability density functions over a set of pre-defined breathing phases to represent respiratory motion and provided a robust formulation of the IMRT treatment planning problem. **?** simulated cardiac motion by

constructing an average scan using all the phases of 4DCT datasets and developed a heart-sparing robust optimization approach for breast cancer IMRT treatment planning problem. Proton therapy is particularly sensitive to geometric uncertainties because the created SOBP might move over the tumor and its neighboring OAR structures, which can lead to dose hot spots in OAR and cold spots in target volume. In IMPT, the proton range and setup uncertainties are addressed by generating various scenarios. Unkelbach, Chan & Bortfeld (2007) used a probabilistic and a robust approach to deal with range uncertainty in the IMPT treatment planning problem. They optimized the expected value of objective function in the first approach and worst-case dose distribution in the second approach. They showed that both of the approaches provide promising outcomes. Pflugfelder, Wilkens & Oelfke (2008*b*) developed a worst-case optimization model accounting for both range and setup uncertainties in the IMPT treatment planning problem. Fredriksson, Forsgren & Hårdemark (2011) used a minmax optimization method to avoid the unnecessary conservatism of the worst-case method by only considering realizable scenarios. Chen et al. (2012) prespecified different scenarios of range and patient setup uncertainties and constructed a database of IMPT plans each optimizing different treatment objectives. By doing so, they explored the trade-off between treatment robustness and plan quality. The above-mentioned studies assumed scenarios of one or two dimensional organ movement with a specific distance (e.g. 5mm) to estimate geometric uncertainties in treatment planning problems. However, such assumptions are not based on the actual respiration-induced organ motion. Therefore, the resulting estimation cannot adequately account for organ deformation and complex motion patterns during breathing.

Four-dimensional CT imaging (4DCT) is widely used to address the respiration-induced organ motion (Rietzel et al., 2005; Underberg et al., 2004). A 4DCT is a sequence of three-dimensional CT (3DCT) images obtained over different respiration phases, which captures the changes in patient anatomy over time. Because 4DCT does not describe the tissue movements from a phase to the next, deformable registration tools have been developed (Hawkes et al., 2005; Lu et al., 2004) to match each point in one of the 3DCT scans with

the corresponding point in the 3DCT scan of another breathing phase. For photon beam therapy, several studies have developed four-dimensional treatment planning models that consider all phases of the 4DCT (**?**Nohadani, Seco & Bortfeld, 2010; Trofimov et al., 2005; Zhang et al., 2004). However, the 4DCTs have not been used in IMPT treatment planning yet.

The respiratory motion also affects the quality of fractionated treatment plan as it causes unequal dose delivery to a specific voxel across different fractions of proton therapy, i.e., some voxels in the PTV might face underdosage in some fractions. As a result, the biological response to treatment for the tumor volume can be different from that of the planned treatment (Flampouri et al., 2006; Nøttrup et al., 2007; Yan & Lockman, 2001). Radiation causes double-strand breaks (DSB) in the DNA. If the radiation dose is not sufficient to kill the cell, the DNA can repair itself by rejoining the pairs of DSB with a constant rate. Therefore, it is important to investigate the effect of respiratory motion on the biological response in a fractionated treatment plan.

In this study, we aim to address the negative impacts of breathing motion on a fractionated proton therapy for non-small cell lung cancer. We incorporate the biologically effective dose in the development of an IMPT treatment planning optimization model that combines several fractions into treatment epochs and corrects the dose per fraction during each epoch to maintain the required biological response in the face of breathing motion. The model incorporates a 4DCT dataset to account for organ motion over different respiratory phases to efficiently handle complex motion patterns without simplifying assumptions on the organ motion. We propose two different versions of this approach; (1) a mean-variance model that optimizes the mean and variance of dose delivered to each organ structure, and (2) a robust method that optimizes the worst possible case of dose deposition to each organ structure. While the first method provides a realistic solution for this problem, the second model optimizes the of treatment quality resulting in a risk-averse treatment plan. We also investigate the impact of the number of epochs and the length of epochs on the BED delivered by the treatment plan in all fractions.

5.2 Methodology

Let $R = \{1, , m\}$ denote the set of respiration phases obtained from the 4DCT dataset. In a fractionated treatment plan, the amount of radiation dose received by voxel v in respiratory phase $r \in R$ and fraction $k \in \{1, 2, , n\}$ is calculated by

$$D(v, r, k) = \sum_{b \in B} d(v, b, r) w(b, k), \tag{5.1}$$

where $d(v, b, r)$ is the dose deposited to voxel v from unit intensity of beamlet b in respiratory phase r, and $w(b, k)$ is the intensity of beamlet $b \in B$, $B = \{1, , l\}$, in fraction k. The linear-quadratic model measures biological effectiveness of a given radiation dose, D, by calculating the surviving fraction, S, of cells as (Oliver, 1964)

$$S = \exp\left(-\alpha D - \beta G D^2\right). \tag{5.2}$$

Here, α and β are intrinsic radiobiological parameters of the organ of interest, and G is the generalized Lea-Catchside time factor (Brenner, 2008) accounting for fractionation effects given a constant repair rate, λ. For irradiation with n equal dose and short fractions separated by a time $\mathcal{T}$, the time factor, G, is calculated by

$$G = [2\theta/(1-\theta)][n - (1-\theta^n)/(1-\theta)], \tag{5.3}$$

where $\theta = \exp\left(-\lambda \mathcal{T}\right)$ (Thames, 1985). According to Thames (1985), biologically equivalent dose (BED) is then defined based on logarithm of the surviving fraction, S, from Equation (5.2):

$$BED = D + \frac{G}{\alpha/\beta} D^2. \tag{5.4}$$

The physical dose and BED received by a voxel in each fraction might be different from the planned dose and BED due to organ motion. Therefore, a modification to the conventional fractionated treatment plan is necessary. The fractionated radiotherapy treatment typically

takes place every week and daily modification and optimization of a treatment plan is time consuming and not cost effective. Instead, we consider a modification of treatment plan periodically when it is needed. We divide the entire n treatment fractions into Q "epochs" with the equal number of fractions, t, in each epoch, i.e., $n = tQ$, and epoch q includes fractions $k = t(q-1)+1$ to $k = tq$ (Figure 5.1).

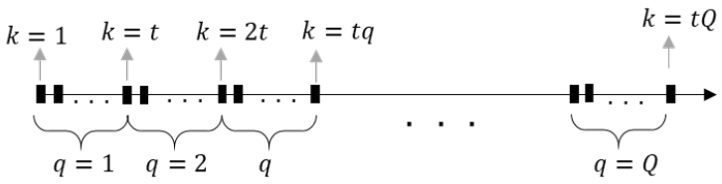


Figure 5.1: Example of dividing the treatment fractions into epochs

Denoting the repair time-factor at the end of first epoch by G_1, the repair time factor at the end of epoch $q \in \{1,,Q\}$ is approximated as follows (derivation is provided in the appendix)

$$G_q \approx G_1\big(1 + \frac{q-1}{1 - 1/t(1-\theta)}\big). \tag{5.5}$$

Proof. Assume fractionated treatment plan is split in Q epochs each having t fractions. Denote G_q as the repair time-factor at the end of epoch $q \in \{1,,Q\}$ and G_1 as repair-time factor at the end of epoch 1. According to Equation 5.3

$$G_1 = [2\theta(1-\theta)][t - (1-\theta^t)(1-\theta)]$$

and

$$G_q = [2\theta(1-\theta)][tq - (1-\theta^{tq})(1-\theta)].$$

Therefore, the ratio of $G_q G_1$ is

$$\frac{G_q}{G_1} = \frac{(q(1-\theta) - (1-\theta^{tq})t)}{((1-\theta) - (1-\theta^t)t)}.$$

Because $\theta = exp(-\lambda\mathcal{T})$ is always less than 1 ($\theta = 2.5E-3$ for a late responding tumor, and $\theta = 0.6$ for an early responding tumor) θ^{tq} and θ^t ($t \geq 5$) become very small and can be considered zero. Therefor the ratio of $G_q G_1$ can be approximated by

$$\frac{G_q}{G_1} \approx \frac{q(1-\theta) - \frac{1}{t}}{(1-\theta) - \frac{1}{t}} = \frac{q - \frac{1}{t(1-\theta)}}{1 - \frac{1}{t(1-\theta)}}.$$

So, the repair time factor at epoch q, G_q, is approximately calculated by

$$G_q \approx G_1 \left(1 + \frac{q-1}{1 - \frac{1}{t(1-\theta)}}\right).$$

□

The amount of dose per fraction is equal in all t fractions of epoch q, but it can be modified at the beginning of each epoch. Denote by $D_q(v)$ the amount of radiation dose received by voxel v using the dose intensity distributions $W(q)$ in each fraction of epoch q. The total physical dose received by voxel v in epoch q is $tD_q(v)$ and the corresponding BED is

$$BED(v) = \sum_{q=1}^{Q} \left(tD_q(v) + \frac{G_q}{\alpha/\beta} t(D_q(v))^2\right). \tag{5.6}$$

The treatment planning goal is to find an optimal dose intensity distributions, $W^*(q)$, for each epoch q that minimizes $f(W(q))$ while including the effect of motion uncertainty in fractionated treatment planning, where

$$\begin{aligned} f(W(q)) = &\sum_{v \in T} |BED(v) - BED^{presc}| + \\ &\sum_{v \in T} |D_q(v) - D^{presc}| + \sum_{v \in OAR} |D_q(v) - D^{max}|_+ \end{aligned} \tag{5.7}$$

and $[x]_+ = max\{x, 0\}$, D^{max} is the maximum amount of dose that OAR voxels can tolerate, and D^{presc} and BED^{presc} are the prescribed radiation dose and its corresponding BED, respectively. A mean-variance model and a worst-case robust model are used to solve this

problem.

5.2.1 Mean-variance model

Assume that the dose delivered to voxel v in fraction k, $D(v,k)$, is a random variable from an unknown probability distribution. Let $p(r)$ denote the proportion of time spent in respiratory phase r during radiation delivery, where $0 \leq p(r) \leq 1$ and $\sum_{r=1}^{m} p(r) = 1$. Assuming the same dose distribution, $w_q(b)$, in all fractions of epoch q, the mean dose delivered to voxel v in each fraction of epoch q is

$$E(D_q(v)) = \sum_{b \in B} \big(\sum_{r \in R} p(r) d(v,b,r) \big) \times w_q(b) \tag{5.8}$$

and the variance of dose delivered to voxel v in each fraction of epoch k is

$$Var(D_q(v)) = \sum_{r \in R} p(r) \Big(\sum_{b \in B} \big[d(v,b,r) w_q(b) \big] - E(D(v,k)) \Big)^2. \tag{5.9}$$

Therefore, the mean dose and variance of dose delivered to voxel v over t fractions of epoch q are $t.E(D_q(v))$ and $t.Var(D_q(v))$, respectively. Using Equations (5.8) and (5.9), Equation (5.6) can be rewritten as

$$BED(v) = \sum_{q=1}^{Q} \big(tE(D_q(v)) + \frac{G_q}{\alpha/\beta} t (E(D_q(v)))^2 \big), \tag{5.10}$$

and the objective function for the optimization model can be written as

$$\begin{aligned} f(W(q)) = & \sum_{v \in T} |BED(v) - BED^{presc}| + \sum_{v \in T} |D_q(v) - D^{presc}| \\ & + \sum_{v \in OAR} |D_q(v) - D^{max}|_+ + \sum_{v \in T} Var(D_q(v)) + \sum_{v \in OAR} Var(D_q(v)). \end{aligned} \tag{5.11}$$

In this study, we assume equal proportion of time spent in each respiratory phase to avoid computational burden in the inverse treatment planning based on 4DCT data (Nohadani,

Seco & Bortfeld, 2010; Suh, Murray & Keall, 2014). However, upon availability of experimental data obtained by measuring the time spent in each phase at the time of acquiring 4DCT images, an uncertainty set can be generated for $p(r)$ using the method provided by Bortfeld et al. (2008).

5.2.2 Worst-case robust model

As the patient breathes during the treatment delivery, we assume all of the respiratory phases will be visited. The mean-variance model described in Section (5.2.1) might not be ideal for such decision-makers. The worst-case robust model creates an unrealistic scenario that assumes the tumor receives the minimum dose from all respiratory phases and the OARs receive the maximum dose from all respiratory phases. In other words, the worst-case dose, $D^{wc}(v,k)$, for a voxel v in fraction k is defined by

$$D^{wc}(v,k) = \begin{cases} \min_r\{\sum_b D(v,r,k) \times w(b,k)\} & \text{if} \quad v \in T \\ \max_r\{\sum_b D(v,r,k) \times w(b,k)\} & \text{if} \quad v \notin T \end{cases} \tag{5.12}$$

This model does not depend on the probability of each respiratory phase, $p(r)$, as it focuses on the worst case, i.e., the lower bound of the treatment quality (Pflugfelder, Wilkens & Oelfke, 2008*b*). This lower bound is obtained by minimizing

$$\begin{aligned} f(W(q)) = &\sum_{v\in T} |BED^{wc}(v) - BED^{presc}| + \\ &\sum_{v\in T} |D_q^{wc}(v) - D^{presc}| + \sum_{v\in OAR} |D_q^{wc}(v) - D^{max}|_+, \end{aligned} \tag{5.13}$$

where, $D_q^{wc}(v)$ is the amount of radiation dose received by voxel v in one fraction of epoch q based on the worst-case scenario and $BED^{wc}(v)$ is the total BED received by voxel v based on the worst-case scenario.

5.3 Numerical Examples

5.3.1 Patient Data

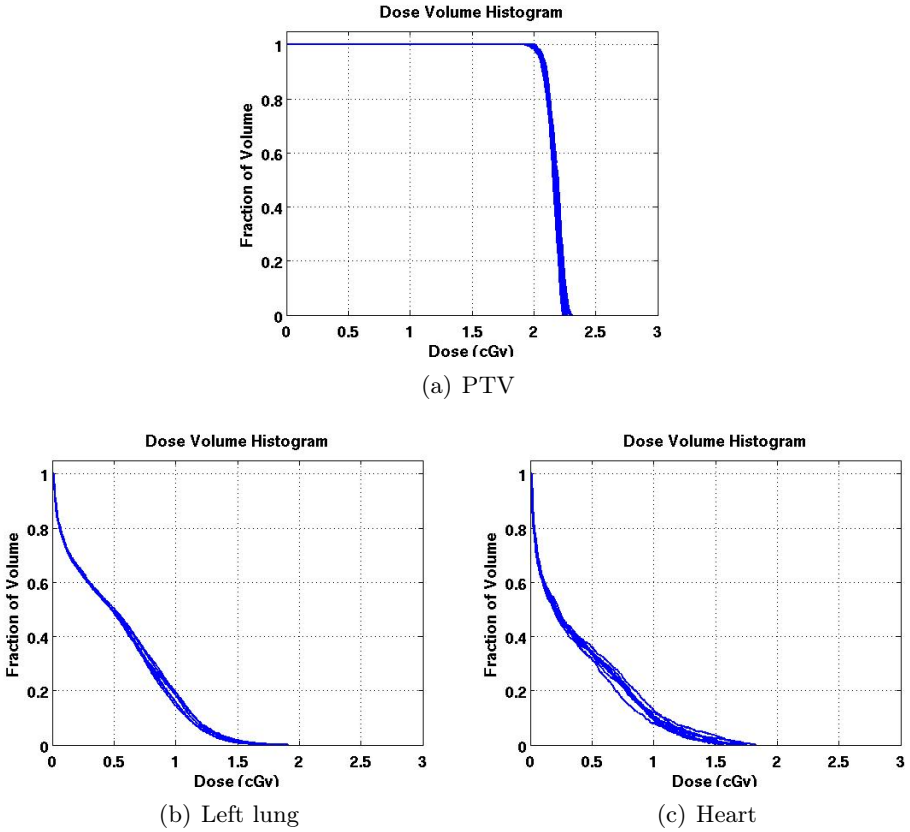
4DCT images of a stage IIIB non-small cell lung cancer patient were obtained from the Cancer Imaging Archive (TCIA) (Balik et al., 2013; Clark et al., 2013; Hugo et al., 2016; Roman et al., 2012). According to TCIA, 4DCT images were acquired using a 16-slice helical CT scanner as respiration-correlated CTs with 10 breathing phases (0 to 90%, phase-based binning) and 3mm slice thickness. The total prescribed dose was 70 Gy delivered through 35 fractions with 2 Gy radiation per fraction. The dose deposition matrices are obtained based on three gantry angles 0°, 180°, and 270° from 4DCT images using MatRad (Wieser et al., 2017). The dose deposition matrices obtained for each phase were registered to the T50 CT image (end of exhale phase), which is considered the reference CT. The registration is performed using OpenREGGUI (Di Perri et al., 2017; Lin et al., 2017; Veiga et al., 2017) based on Demons algorithm (Thirion, 1996).

5.3.2 Numerical Results

Both mean-variance and worst-case models were implemented on a Linux computer with an Intel Quad-core 3.0 GHz processor and 352 GB RAM and the optimization models were solved using CPLEX version 12.6.0.0. Given that a fractionated treatment is usually delivered to the patient on five weekdays, we define the epochs as weeks with five fractions in each epoch. The dose-volume histogram (DVH) is used to evaluate the treatment plan provided by each model. The tumor is located on the left lung, and therefore the left lung and heart are considered OARs. The dose limitations for the OARs indicate that total prescription dose (i.e., 2 Gy per fraction) can be received by less than 30% of the heart volume, while less than 35% of lung volume can receive more than 20 Gy (i.e., 0.57 Gy per fractionw) (Emami et al., 1991).

Figure 5.2 shows a set of ten DVHs for each of the PTV and OARs where each DVH

line corresponds to a respiratory phase resulting from the mean-variance optimization model (Section 5.2.1). The figure shows that the dose received by the left lung in all respiratory phases form a narrow cloud of DVH, but a relatively high variance is observed on the heart. This variance is clearly a consequence of cardiac motion, which is not necessarily consistent with the patientâĂŹs respiratory patterns. The DVHs for the PTV also show some variance on less than 40% of its volume.



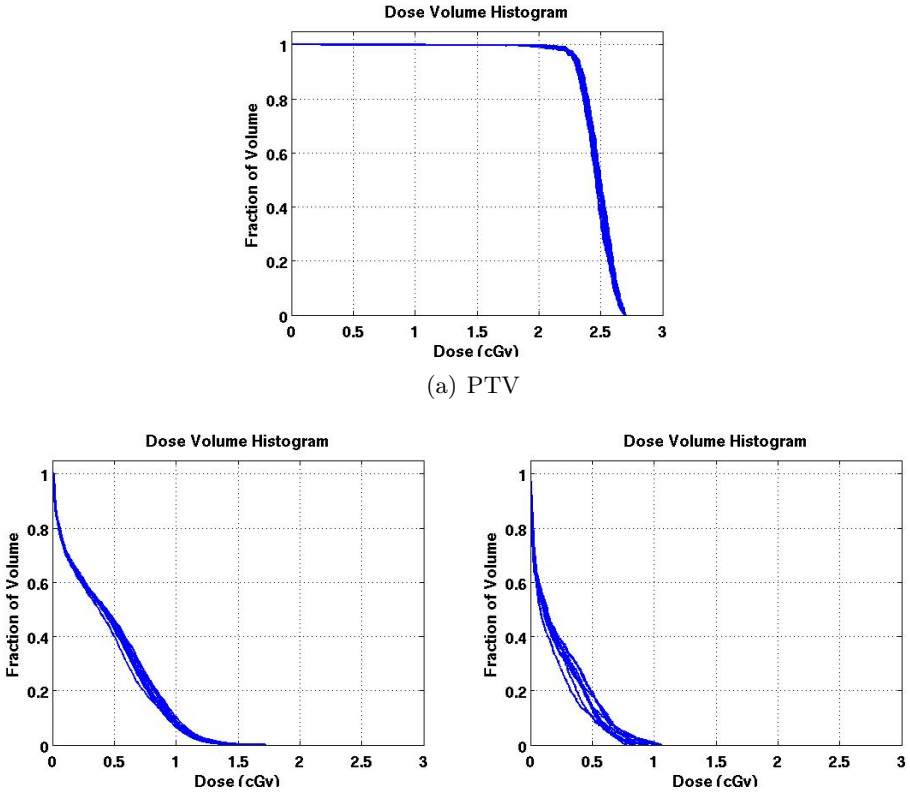
(a) PTV

(b) Left lung

(c) Heart

Figure 5.2: Dose-volume histogram from mean-variance optimization model for (a) PTV, (b) left lung, and (c) heart

Figure 5.3 shows the cloud of DVHs resulting from the worst-case method. The results show tighter dose bounds for all structures. Comparing with the results from the mean-variance model, the worst-case model provides better dose fall-off for the PTV, while increasing the dose delivered to the left lung. The worst-case model also improves on the DVHs for the heart.



(a) PTV

(b) Left lung

(c) Heart

Figure 5.3: Dose-volume histogram from worst-case robust model optimization for (a) PTV, (b) left lung, and (c) heart

For a detailed comparison between the two models, Table 5.1 shows the mean and standard deviation of different metrics for the structures over all respiratory phases in the first week of treatment. The comparison is made between mean-variance model, worst-case model, and a benchmark model that optimizes treatment based on the end of exhale CT image only. In a DVH graph, the dose received by more than x% of a volume is known as Dx%, and the percentage of the volume that receives at least y Gy amount of dose is known as VyGy. Table 5.1 provides D90%, D50% and D10% for all structures, minimum dose and V2Gy for the PTV, maximum dose and V0.57Gy for the left lung, and maximum dose and V1Gy for the heart. The average D10% and D90% from mean-variance model to the PTV varies from 2.23 Gy to 2.08 Gy and the standard deviation is significantly low. The worst-case model provides a range of 2.33 Gy to 2.62 Gy for average D90% and D10% with higher

standard deviations compared to the mean-variance model. The minimum dose received on average by the PTV is 2 Gy and 2.15 Gy with the mean-variance model and the worst-case model, respectively. The average max dose that the worst-case model delivers to the left lung is 1.52 Gy with 0.06 standard deviation, while with the mean-variance model, the max dose is on average 1.77 Gy with 0.03 standard deviation. Similarly, the maximum max dose delivered to the heart by the mean-variance and the worst-case models was 1.65 Gy and 0.9 Gy, respectively. Both models deliver more than 2 Gy dose to 100% of the PTV. The left lung V0.57Gy was 41% and 35% with the mean-variance model and the worst-case model, respectively. Similarly, the heart V1Gy was 9% and 2% with the mean-variance model and the worst-case model, respectively. Although the worst-case model results in lower dose to the OARs and higher dose to the PTV compared to the mean-variance model, the mean-variance model provides lower variance overall. Both models outperform the benchmark model.

Table 5.1: Comparison of mean and standard deviation of DVH metrics

Organ	Criterion	Mean-variance	Worst-case	Benchmark
PTV	D 90% [Gy]	2.08 ± 0.01	2.33 ± 0.01	1.4 ± 0.14
	D 50% [Gy]	2.17 ± 0.01	2.48 ± 0.01	1.9 ± 0.06
	D 10% [Gy]	2.23 ± 0.01	2.62 ± 0.01	2.6 ± 0.27
	Min dose [Gy]	2 ± 0.01	2.15 ± 0.01	1.19 ± 0.11
	V 2Gy [%]	100 ± 0.0	100 ± 0.0	47 ± 0.08
Left Lung	D 90% [Gy]	0.03 ± 0.0	0.02 ± 0.0	0.02 ± 0.04
	D 50% [Gy]	0.48 ± 0.01	0.37 ± 0.02	0.4 ± 0.03
	D 10% [Gy]	1.12 ± 0.02	0.86 ± 0.02	1.13 ± 0.05
	Max dose [Gy]	1.77 ± 0.03	1.52 ± 0.06	1.85 ± 0.09
	V 0.57Gy [%]	41 ± 0.3	35 ± 1.5	46 ± 1.8
Heart	D 90% [Gy]	0.02 ± 0.01	0.009 ± 0.0	0.02 ± 0.01
	D 50% [Gy]	0.21 ± 0.01	0.11 ± 0.01	0.18 ± 0.03
	D 10% [Gy]	1.01 ± 0.04	0.56 ± 0.04	1.14 ± 0.18
	Max dose [Gy]	1.65 ± 0.12	0.9 ± 0.15	1.78 ± 0.18
	V 1Gy [%]	9 ± 2	2 ± 0.3	13 ± 3.5

In order to demonstrate the effect of breathing motion on fractionated treatment plan, Figures 5.4 and 5.5 show the average D50% physical dose and the resulting BED delivered

in fractions of each epoch to the PTV. The mean dose and BED here are obtained by averaging over all respiratory phases. The figures recommend to start from a larger amount of radiation dose in the first week of treatment and gradually decrease it to the prescribed dose. This pattern is a result of breathing motion as the larger amount of radiation dose at the beginning compensates for BED inhomogeneities from breathing motion and maintains the BED approximately equal during the course of treatment. Although the proposed fractionation plan proposed by the mean-variance model is more aggressive compared to the benchmark model, the resulting BED is consistently close the desired BED. However, the BED resulting from the benchmark model is always under the desired BED.

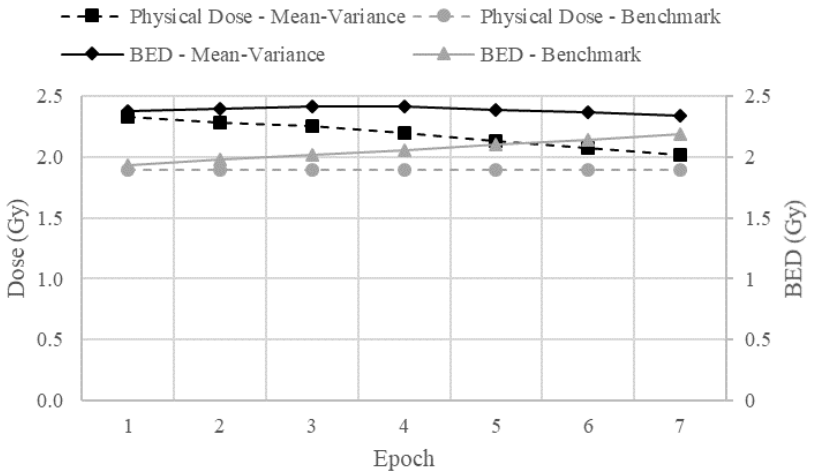


Figure 5.4: Average physical dose versus average BED delivered in fractions of each epoch to the 50% of the PTV in mean-variance model

The worst-case robust model recommends a larger amount of dose in each fraction compared to the mean-variance model (Figure 5.6 versus Figure 5.4). Consequently, BED delivered to 50% and 90% of the PTV by this model is slightly greater than that from the mean-variance model. The reason is that the worst-case model optimizes the treatment plan based on the worst-case (i.e., minimum) dose over all possible respiratory phases.

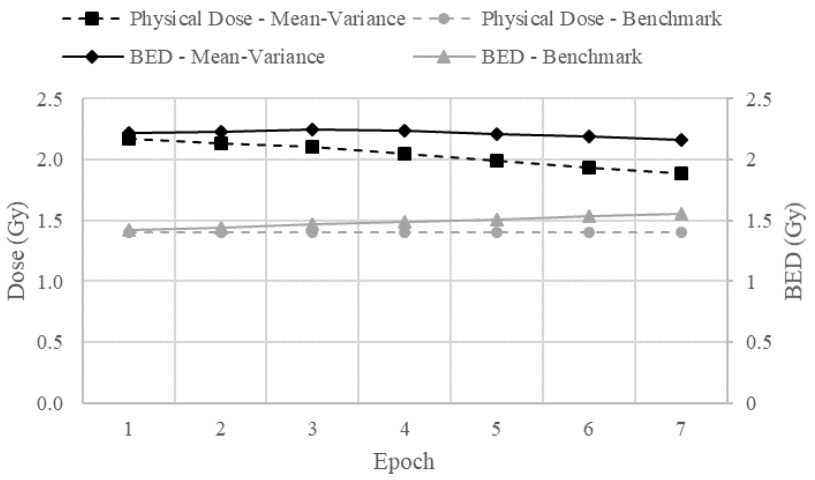


Figure 5.5: Average physical dose versus average BED delivered in fractions of each epoch to the 90% of the PTV in mean-variance model

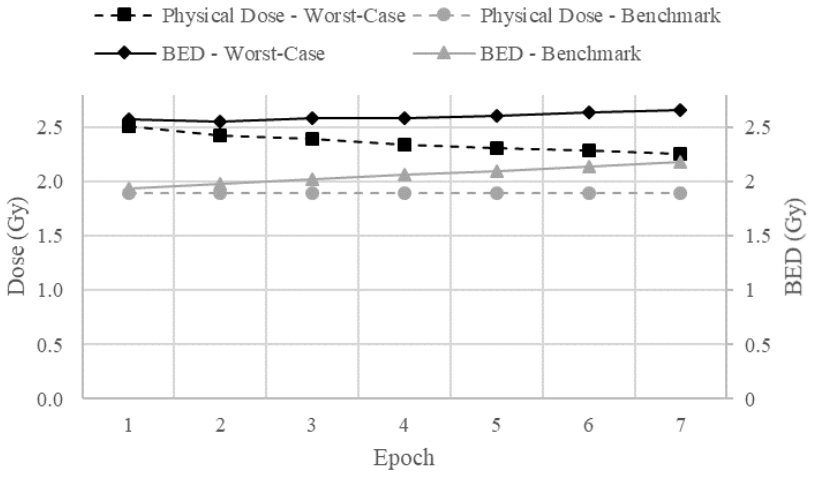


Figure 5.6: Average physical dose versus average BED delivered in fractions of each epoch to the 50% of the PTV in worst-case robust model

5.3.2.1 The impact of epoch frequency and length

In this study, we propose to combine several fractions into several epochs and apply possible modifications to the treatment plan in each epoch. Therefore, we investigate the effect of the number and length of epochs (i.e., the frequency of plan modification) on the treatment plan quality.

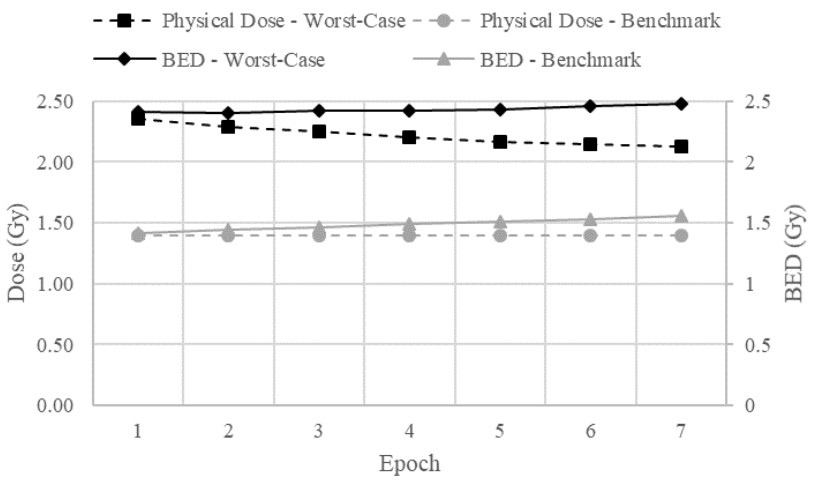
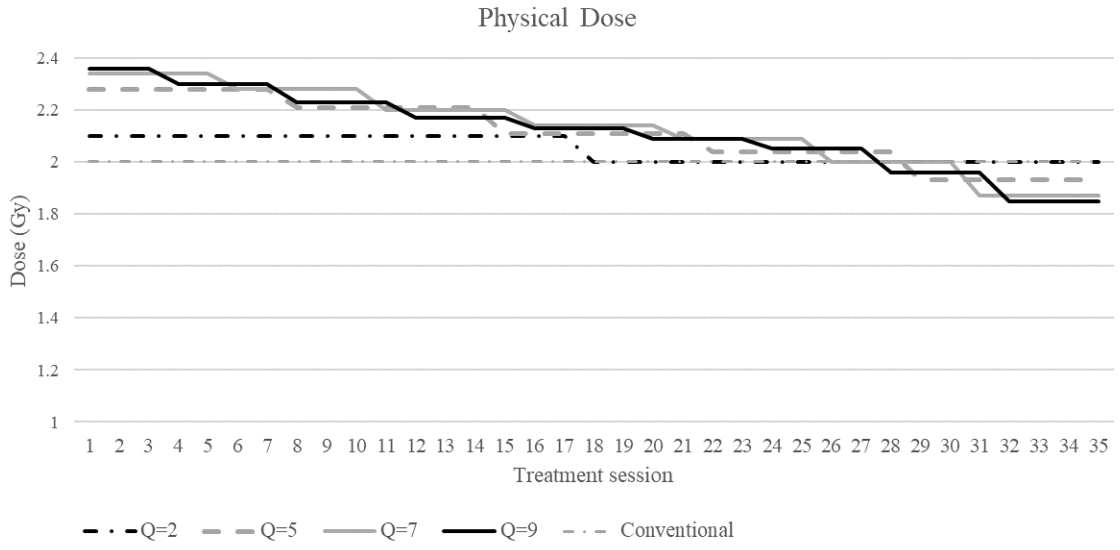


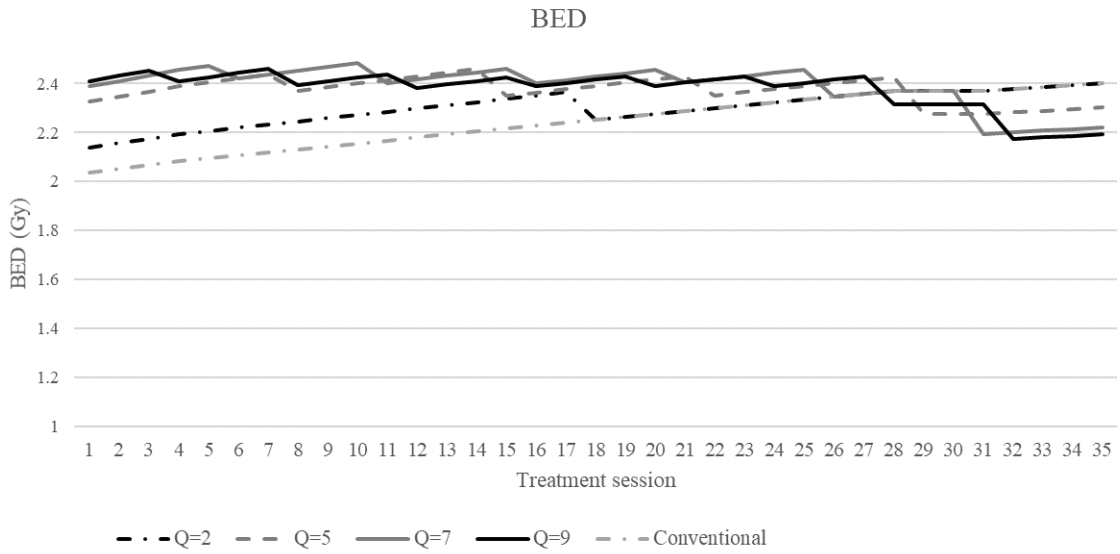
Figure 5.7: Average physical dose versus average BED delivered in fractions of each epoch to the 90% of the PTV in worst-case robust model

We selected the mean-variance model to implement for the following cases: 1) $(Q, t) = (2, 18)$, i.e. two epochs each with 18 treatment fractions, 2) $(Q, t) = (5, 7)$, 3) $(Q, t) = (7, 5)$, and 4) $(Q, t) = (9, 4)$. When $Q = 2$ and 9, the first epoch has less fractions than other epochs so that the total number of treatment fractions remains is 35. The average daily physical dose and BED resulting from the optimized treatment plan for these seven cases are presented in Figure 5.8.

When $Q = 2$, the average daily physical dose has changed once from 2.1 Gy in the first epoch to 2 Gy in the second epoch, but the average daily BED starts from 2.15 Gy in the first session and increases to 2.44 Gy in the last session. When $Q = 9$, fractionated plan is optimized more frequently and the physical dose is adjusted for shorter epochs to yield the desired BED. As a result, the average daily physical dose starts from 2.36 Gy in the first epoch and decreases to 1.85 Gy in the last epoch. In this case the average daily BED changes between 2.3 Gy and 2.45 Gy in the first 7 epochs, but drops to 2.2 Gy in the last epoch. The variation of average physical dose and BED for cases 2 to 6 is between these two extreme cases.



(a) Physical Dose



(b) BED

Figure 5.8: Comparison of effect of different epoch number and length on (a) daily average physical dose, (b) daily average BED

Figure 5.8(a) shows that as the frequency of treatment plan modification increase, treatment starts with a higher physical dose in the first session and ends with a lower physical dose in the last session. Figure 5.8(b) shows that as the frequency of treatment plan modification decreases, the average daily BED in the earlier fractions decreases, often not satisfying the required biological effectiveness. While the treatment plan with $Q = 2$ is not desired because of the wide range of average daily BED and lower average BED in the first epoch, the treatment plan with $Q = 9$ might not be practical as it can be time consuming. Therefore, the treatment plan with $Q = 5$ can provide a compromise: consistency in terms of both physical dose and BED while not requiring too many plan corrections.

5.4 Conclusion

The goal of this study was to optimize the fractionated proton therapy treatment plan for non-small cell lung cancer while taking into account the effect of respiratory motion. To this aim a fractionated IMPT optimization framework was developed to provide optimal solution for both IMPT and fractionation planning of the treatment and was solved using a mean-variance model and a worst-case robust model. The proposed framework combined several fractions into epochs and modified the physical radiation dose to maintain the total required BED at the presence of breathing motion. A 4DCT dataset was used to handle complex motion patterns.

Both mean-variance and worst-case robust models improved the IMPT plan compared to the benchmark model in terms of standard deviation of dose over all respiratory phases and variation of dose to different PTV volumes (Table 5.1). Because the worst-case model used worst-case dose scenario as described in Section 5.2.2, the average dose delivered to the PTV in all phases of respiration was greater than that of the mean-variance model and the average dose deilvered to the OARs was lower than that of the mean-variance model. However, the mean-variance model provided a lower standard deviation of dose over the respiratory phases. Based on the position of the tumor with respect to surrounding OARs and the radiobiological characteristics of the tumor and OARs, one can decide whether the

tumor can receive higher dose and tolerate the higher standard deviation through the worst-case IMPT plan or the mean-variance model with higher standard deviation is more suitable for a patient. Both models suggested to start from higher radiation dose in the first week of treatment and gradually decrease it in subsequent weeks to address the increased repair effect resulting from nonuniform voxel dose between fractions. The average BED resulting from this fractionation pattern was shown to be almost equal in all treatment weeks. A comparison between optimized treatment plans with different length of epochs showed that less frequent dose modification (i.e. fewer epochs each with more fractions) results in lower BED in the first fractions while more frequent dose modification (i.e. more epochs each with fewer fractions) results in higher physical dose in the beginning and lower amount of physical dose in the end of the treatment.

The current assumption in 4D planning of radiation therapy is that equal proportion of time is spent in each respiratory phase. This assumption can be investigated by measuring the average time spent in each respiratory phase at the time of acquiring 4DCTs. In addition, this study assumed a fixed relative biological effectiveness (RBE) for proton. While this assumption provides a good estimation of the treatment quality in terms of biological effectiveness, investigating the effect of variable RBE can provide a more accurate estimation in this regard.

Chapter 6

Summary

6.1 Conclusion

In this dissertation we aimed to address the unanswered questions about biological effectiveness of fractionated raditherapy treatment plan.

In Chapter 3 we utilized an extended biological model of cell survival for the tumor and OAR to include the four R's of radiation therapy and used a Gompertzian growth model for the tumor to explain the dependency of tumor growth rate on the gene level factors. First, we developed a control limit policy based on patient-specific tumor and OAR characteristics with respect to radiation treatment to determine whether a dynamic treatment plan can improve the treatment quality (i.e., tumor damage and OAR sparing) compared to the conventional one. Next, we developed a constrained POMDP framework based on the extended biological model to provide a decision-making policy that maximizes the expected biological equivalent dose (EBED) of tumor while keeping the OAR survival under control. Numerical results showed that using a dynamic treatment plan instead of the conventional one can result in higher tumor damage with slight increase in the OAR damage. However, when the tumor tissue was late responding (lower sensitivity to radiotherapy) and the OAR tissue is early responding (higher sensitivity to radiotherapy) the conventional treatment plan was recommended as a conservative plan. The POMDP policy recommended optimal actions based on the updated belief vector at each decision epoch. Due to the Gompertzian tumor growth model, lesser number of tumor cells in the later weeks of treatment cause in

an increase in the tumor growth rate. Therefore, the dose increase was recommended at the later weeks of treatment to increase the tumor damage and shorten the treatment length.

In Chapter 4, we provided decision policies for dynamic and personalized fractionated proton therapy treatment plans. A constrained MDP model was used to obtain such policy which deals with uncertainty of tumor response and growth rate. The state in our suggested MDP model was defined as tumor cell count estimated by applying random forest classification method on radiomics data of CT images. The actions in the MDP model were obtained from optimal LET-based IMPT plans. The extended LQR model was to include four R's of radiation therapy in treatment planning. To explain the tumor growth dynamics, Gompertzian growth function was incorporated in the LQR model. A LET-based IMPT model was developed to ensure the homogeneity of BED distribution over the tumor volume.

The proposed model was applied to two cases of prostate cancer and pediatric ependymoma. Optimal fractionated treatment policy and optimal IMPT plan were provided for each case. In the prostate cancer case, tumor and OARs had early responding tissue type. But in the pediatric ependymoma example, tumor was late responding, and OARs were early responding. Different dose targets were defined as alterations to the conventional treatment plan. The LET-based IMPT model outperformed the threshold model in general, providing significantly increases TCP with similar NTCP. Randomized MDP policies were provided for both cases. In the prostate cancer example, smaller dose target was suggested for high tumor cell count where the tumor growth is at its lowest rate. But as the tumor cell count decreases, tumor growth rate increases and bigger dose target was suggested to destroy tumor. For the pediatric ependymoma due to high radio-sensitivity of OARs and low radio-sensitivity of tumor, different OAR weight factors were used to investigate the impact of decision-maker preferences on treatment policy.

The goal of Chapter 5 was to investigate the effect of intra-fractional organ motion on a fractionated proton therapy for non-small cell lung cancer. The biologically effective dose was included in the IMPT model to maintain the total required biological response at the presence of breathing motion. A 4DCT dataset was used to handle complex motion patterns.

The aim of this study was to avoid sophisticated 4D delivery systems and interplay effect by using 4DCT dataset to plan a 3D delivery technique. A statistical model and a worst-case robust model were used to approach this problem from different angles. The statistical model focused on mean and variance of dose delivered to each organ structure, and the robust model focused on the worst possible case of dose deposition to each organ structure.

The proposed models were implemented on a 4DCT dataset from a non-small cell lung cancer patient. The dose distribution over PTV structure from worst-case robust model was better than that of statistical mean-variance model in terms of both DVH bound and dose homogeneity. Both models suggested to start from a larger amount of radiation dose in the first week of treatment and gradually decreasing it on next weeks. This pattern of fraction size aims to compensate for the increased repair effect resulting from nonuniform voxel dose between fractions. As an evidence, the weekly mean BED resulting from this fractionation pattern was shown to be almost equal in all treatment weeks. However, because of conservatism of worst-case robust model, a larger total dose has to be delivered in every treatment week to achieve the same biological effectiveness.

The methods used in this study serve different points of view. However, the statistical mean-variance model can provide more accurate and practical solution. The current assumption in 4D planning of radiation therapy is that equal proportion of time is spent in each respiratory phase. This assumption can be investigated by measuring the average time spent in each respiratory phase at the time of acquiring 4DCTs. In addition, this study assumed a fixed relative biological effectiveness (RBE) for proton. While this assumption provides a good estimation of the treatment quality in terms of biological effectiveness, investigating the effect of variable RBE can provide a more accurate estimation in this regard.

6.2 Outcomes

Journal Publication

- Nasrin Nouri and Gino J. Lim, "*Personalized treatment planning considering biological response uncertainty: A Partially Observable Markov Decision Process Approach*".
- Nasrin Nouri, Gino J. Lim and Wenhua Cao, "*Personalized and LET-based proton therapy treatment planning considering tumor biological response uncertainty*".
- Nasrin Nouri, Gino J. Lim and Taewoo Lee, "*Dealing with respiratory motion in a fractionated IMPT treatment plan*".

Conference Presentation

- Nasrin Nouri, Gino J. Lim, "*Optimal Radiotherapy Treatment Policy Based on Tumor Biological Response*," INFORMS Annual Meeting. Nashville, TN, November 2016.

References

Altman, Eitan. 1999. *Constrained Markov Decision Processes*. Vol. 7 CRC Press.

American Cancer Society Cancer Facts & Figures. 2018. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2018.html.

Ang, K Kian. 1998. "Altered fractionation trials in head and neck cancer." In *Seminars in radiation oncology*. Vol. 8 Elsevier pp. 230–236.

Armpilia, CI, RG Dale & B Jones. 2004. "Determination of the optimum dose per fraction in fractionated radiotherapy when there is delayed onset of tumour repopulation during treatment." *The British Journal of Radiology* 77(921):765–767.

Bai, Xuemin, Gino Lim, Hans-Peter Wieser, Mark Bangert, David Grosshans, Radhe Mohan & Wenhua Cao. 2019. "Robust optimization to reduce the impact of biological effect variation from physical uncertainties in intensity-modulated proton therapy." *Physics in Medicine & Biology* 64(2):025004.

Balik, Salim, Elisabeth Weiss, Nuzhat Jan, Nicholas Roman, William C Sleeman, Mirek Fatyga, Gary E Christensen, Cheng Zhang, Martin J Murphy, Jun Lu et al. 2013. "Evaluation of 4-dimensional computed tomography to 4-dimensional cone-beam computed tomography deformable image registration for lung cancer adaptive radiation therapy." *International Journal of Radiation Oncology* Biology* Physics* 86(2):372–379.

Barazzuol, Lara, Neil G Burnet, Raj Jena, Bleddyn Jones, Sarah J Jefferies & Norman F Kirkby. 2010. "A mathematical model of brain tumour response to radiotherapy

and chemotherapy considering radiobiological aspects." *Journal of Theoretical Biology* 262(3):553–565.

Baskar, Rajamanickam, Kuo Ann Lee, Richard Yeo & Kheng-Wei Yeoh. 2012. "Cancer and radiation therapy: current advances and future directions." *International Journal of Medical Sciences* 9(3):193.

Belli, M, F Cera, Ret al Cherubini, AMI Haque, F Ianzini, G Moschini, O Sapora, G Simone, MA Tabocchini & P Tiveron. 1993. "Inactivation and mutation induction in V79 cells by low energy protons: re-evaluation of the results at the LNL facility." *International journal of radiation biology* 63(3):331–337.

Bert, Christoph & Marco Durante. 2011. "Motion in radiotherapy: particle therapy." *Physics in Medicine and Biology* 56(16):R113.

Bertuzzi, A, C Bruni, F Papa & C Sinisgalli. 2013. "Optimal solution for a cancer radiotherapy problem." *Journal of mathematical biology* 66(1-2):311–349.

Bibault, Jean-Emmanuel, Ingrid Fumagalli, Charles Ferté, Cyrus Chargari, Jean-Charles Soria & Eric Deutsch. 2013. "Personalized radiation therapy and biomarker-driven treatment strategies: a systematic review." *Cancer and Metastasis Reviews* 32(3-4):479–492.

Bortfeld, Thomas. 2006. "IMRT: a review and preview." *Physics in Medicine and Biology* 51(13):R363.

Bortfeld, Thomas, Jagdish Ramakrishnan, John N Tsitsiklis & Jan Unkelbach. 2015. "Optimization of radiation therapy fractionation schedules in the presence of tumor repopulation." *INFORMS Journal on Computing* 27(4):788–803.

Bortfeld, Thomas, Kimmo Jokivarsi, Michael Goitein, Jong Kung & Steve B Jiang. 2002. "Effects of intra-fraction motion on IMRT dose delivery: statistical analysis and simulation." *Physics in Medicine & Biology* 47(13):2203.

Bortfeld, Thomas, Timothy CY Chan, Alexei Trofimov & John N Tsitsiklis. 2008. "Robust management of motion uncertainty in intensity-modulated radiation therapy." *Operations Research* 56(6):1461–1473.

Brenner, David J. 1993. "Dose, volume, and tumor-control predictions in radiotherapy." *International Journal of Radiation Oncology* Biology* Physics* 26(1):171–179.

Brenner, David J. 2008. "The linear-quadratic model is an appropriate methodology for determining isoeffective doses at large doses per fraction." In *Seminars in Radiation Oncology*. Vol. 18 Elsevier pp. 234–239.

Brenner, David J, Lynn R Hlatky, Philip J Hahnfeldt, Eric J Hall & Rainer K Sachs. 1995. "A convenient extension of the linear-quadratic model to include redistribution and reoxygenation." *International Journal of Radiation Oncology* Biology* Physics* 32(2):379–390.

Bruce Alberts, Alexander Johnson, Julian Lewis David Morgan Martin Raff Keith Roberts Peter Walter John Wilson & Tim Hunt. 2015. "Molecular Biology - 5th ed." *The Quarterly Review of Biology* 90(3):343–343.
URL: *https://doi.org/10.1086/682635*

Brunton, GF & TE Wheldon. 1978. "Characteristic species dependent growth patterns of mammalian neoplasms." *Cell Proliferation* 11(2):161–175.

Cao, Wenhua, Gino Lim, Li Liao, Yupeng Li, Shengpeng Jiang, Xiaoqiang Li, Heng Li, Kazumichi Suzuki, X Ronald Zhu, Daniel Gomez et al. 2014. "Proton energy optimization and reduction for intensity-modulated proton therapy." *Physics in Medicine & Biology* 59(21):6341.

Carabe, Alejandro, Maryam Moteabbed, Nicolas Depauw, Jan Schuemann & Harald Paganetti. 2012. "Range uncertainty in proton therapy due to variable biological effectiveness." *Physics in Medicine and Biology* 57(5):1159.

Cassandra, Anthony Rocco. 1998. "Exact and approximate algorithms for partially observable Markov decision processes.".

Chen, Wei, Jan Unkelbach, Alexei Trofimov, Thomas Madden, Hanne Kooy, Thomas Bortfeld & David Craft. 2012. "Including robustness in multi-criteria optimization for intensity-modulated proton therapy." *Physics in Medicine & Biology* 57(3):591.

Chen, Yizheng, Clemens Grassberger, Junli Li, Theodore S Hong & Harald Paganetti. 2018. "Impact of potentially variable RBE in liver proton therapy." *Physics in Medicine & Biology* 63(19):195001.

Clark, Kenneth, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle et al. 2013. "The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository." *Journal of digital imaging* 26(6):1045–1057.

Courdi, A, N Brassart, J Herault & P Chauvel. 1994. "The depth-dependent radiation response of human melanoma cells exposed to 65 MeV protons." *The British Journal of Radiology* 67(800):800–804.

Deasy, Joseph O & JACK F Fowler. 2005. "Radiobiology of IMRT." *Intensity Modulated Radiation Therapy. A Clinical Perspective* pp. 53–74.

Di Perri, Dario, John A Lee, Anne Bol, François-Xavier Hanin, Guillaume Janssens, Daniel Labar, Annie Robert, Edmond Sterpin & Xavier Geets. 2017. "Evolution of [18F] fluorodeoxyglucose and [18F] fluoroazomycin arabinoside PET uptake distributions in lung tumours during radiation therapy." *Acta Oncologica* 56(4):516–524.

Douglas, BG & JF Fowler. 1976. "The effect of multiple small doses of x rays on skin reactions in the mouse and a basic interpretation." *Radiation Research* 66(2):401–426.

Emami, Bahman, J Lyman, A Brown, L Cola, M Goitein, JE Munzenrider, B Shank, LJ Solin & M Wesson. 1991. "Tolerance of normal tissue to therapeutic irradiation." *International Journal of Radiation Oncology* Biology* Physics* 21(1):109–122.

Ferrer, Montserrat, José Francisco Suárez, Ferran Guedea, Pablo Fernández, Víctor Macías, Alfonso Mariño, Asunción Hervas, Ismael Herruzo, María José Ortiz, Humberto Villavicencio et al. 2008. "Health-related quality of life 2 years after treatment with radical prostatectomy, prostate brachytherapy, or external beam radiotherapy in patients with clinically localized prostate cancer." *International Journal of Radiation Oncology* Biology* Physics* 72(2):421–432.

Flampouri, Stella, Steve B Jiang, Greg C Sharp, John Wolfgang, Abhijit A Patel & Noah C Choi. 2006. "Estimation of the delivered patient dose in lung IMRT treatment based on deformable registration of 4D-CT data and Monte Carlo simulations." *Physics in Medicine & Biology* 51(11):2763.

Fowler, Jack F. 2001. "Biological factors influencing optimum fractionation in radiation therapy." *Acta Oncologica* 40(6):712–717.

Fowler, John F. 1989. "The linear-quadratic formula and progress in fractionated radiotherapy." *The British Journal of Radiology* 62(740):679–694.

Fredriksson, Albin, Anders Forsgren & Björn Hårdemark. 2011. "Minimax optimization for handling range and setup uncertainties in proton therapy." *Medical physics* 38(3):1672–1684.

Gay, Hiram A & Andrzej Niemierko. 2007. "A free program for calculating EUD-based NTCP and TCP in external beam radiotherapy." *Physica Medica* 23(3-4):115–125.

Gerweck, Leo E, Syed T Zaidi & Anthony Zietman. 1994. "Multivariate determinants of radiocurability I: prediction of single fraction tumor control doses." *International Journal of Radiation Oncology* Biology* Physics* 29(1):57–66.

Giantsoudi, Drosoula, Clemens Grassberger, David Craft, Andrzej Niemierko, Alexei Trofimov & Harald Paganetti. 2013. "Linear energy transfer-guided optimization in intensity modulated proton therapy: feasibility study and clinical potential." *International Journal of Radiation Oncology* Biology* Physics* 87(1):216–222.

Gillies, Robert J, Paul E Kinahan & Hedvig Hricak. 2015. "Radiomics: images are more than pictures, they are data." *Radiology* 278(2):563–577.

Grassberger, Clemens, Alexei Trofimov, Anthony Lomax & Harald Paganetti. 2011. "Variations in Linear Energy Transfer Within Clinical Proton Therapy Fields and the Potential for Biological Treatment Planning." *International Journal of Radiation Oncology*Biology*Physics* 80(5):1559 – 1566.
URL: *http://www.sciencedirect.com/science/article/pii/S0360301610034486*

Habl, Gregor, Gencay Hatiboglu, Lutz Edler, Matthias Uhl, Sonja Krause, Matthias Roethke, Heinz P Schlemmer, Boris Hadaschik, Juergen Debus & Klaus Herfarth. 2014. "Ion Prostate Irradiation (IPI)–a pilot study to establish the safety and feasibility of primary hypofractionated irradiation of the prostate with protons and carbon ions in a raster scan technique." *BMC Cancer* 14(1):202.

Hansen, Eric Anton. 1998. "Finite-memory control of partially observable systems.".

Hauskrecht, Milos. 2000. "Value-function approximations for partially observable Markov decision processes." *Journal of Artificial Intelligence Research* 13:33–94.

Hawkes, David J, D Barratt, Jane M Blackall, C Chan, Philip J Edwards, K Rhode, Graeme P Penney, Jamie McClelland & Derek LG Hill. 2005. "Tissue deformation and shape models in image-guided interventions: a discussion paper." *Medical Image Analysis* 9(2):163–175.

Holloway, RP & RG Dale. 2013. "Theoretical implications of incorporating relative biological effectiveness into radiobiological equivalence relationships." *The British Journal of Radiology* 86(1022):20120417.

Hsu, David, Wee Sun Lee & Nan Rong. 2008. "A point-based POMDP planner for target tracking." In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE pp. 2644–2650.

Hugo, Geoffrey D., Elisabeth Weiss, William C. Sleeman, Salim Balik, Paul J. Keall, Jun Lu & Jeffrey F. Williamson. 2016. "Data from 4D Lung Imaging of NSCLC Patients.".

Isom, Joshua D, Sean P Meyn & Richard D Braatz. 2008. "Piecewise Linear Dynamic Programming for Constrained POMDPs." In *AAAI*. pp. 291–296.

Jones, Bleddyn & Roger G Dale. 2000. "Estimation of optimum dose per fraction for high LET radiations: implications for proton radiotherapy." *International Journal of Radiation Oncology* Biology* Physics* 48(5):1549–1557.

Karlin, Samuel. 1964. "Total positivity, absorption probabilities and applications." *Transactions of the American Mathematical Society* 111(1):33–107.

Kim, Minsun, Archis Ghate & Mark H Phillips. 2012. "A stochastic control formalism for dynamic biologically conformal radiation therapy." *European Journal of Operational Research* 219(3):541–556.

Kim, Minsun, Archis Ghate & MH Phillips. 2009. "A Markov decision process approach to temporal modulation of dose fractions in radiation therapy planning." *Physics in Medicine & Biology* 54(14):4455.

Kraft, Gerhard. 2000. "Tumor therapy with heavy charged particles." *Progress in Particle and Nuclear Physics* 45:S473–S544.

Krämer, Michael & Michael Scholz. 2006. "Rapid calculation of biological effects in ion radiotherapy." *Physics in Medicine & Biology* 51(8):1959.

Krishnamurthy, Vikram. 2016. *Partially Observed Markov Decision Processes*. Cambridge University Press.

Lambin, Philippe, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud GPM van Stiphout, Patrick Granton, Catharina ML Zegers, Robert Gillies, Ronald Boellard, André Dekker et al. 2012. "Radiomics: extracting more information from medical images using advanced feature analysis." *European journal of cancer* 48(4):441–446.

Levegrün, Sabine, Andrew Jackson, Michael J Zelefsky, Ennapadam S Venkatraman, Mark W Skwarchuk, Wolfgang Schlegel, Zvi Fuks, Steven A Leibel & C Clifton Ling. 2002. "Risk group dependence of dose-response for biopsy outcome after three-dimensional conformal radiation therapy of prostate cancer." *Radiotherapy and Oncology* 63(1):11–26.

Li, Xiaoqiang, Xiaochun Wang, Yupeng Li & Xiaodong Zhang. 2011. "A 4D IMRT planning method using deformable image registration to improve normal tissue sparing with contemporary delivery techniques." *Radiation Oncology* 6(1):83.

Liaw, Andy, Matthew Wiener et al. 2002. "Classification and regression by randomForest." *R News* 2(3):18–22.

Lin, Liyong, Kevin Souris, Minglei Kang, Adam Glick, Haibo Lin, Sheng Huang, Kristin Stützer, Guillaume Janssens, Edmond Sterpin, John A Lee et al. 2017. "Evaluation of motion mitigation using abdominal compression in the clinical implementation of pencil beam scanning proton therapy of liver tumors." *Medical Physics* 44(2):703–712.

Lovejoy, William S. 1987. "Some monotonicity results for partially observed Markov decision processes." *Operations Research* 35(5):736–743.

Lu, Weiguo, Ming-Li Chen, Gustavo H Olivera, Kenneth J Ruchala & Thomas R Mackie. 2004. "Fast free-form deformable registration via calculus of variations." *Physics in Medicine & Biology* 49(14):3067.

Lyman, John T. 1985. "Complication probability as assessed from dose-volume histograms." *Radiation Research* 104(2s):S13–S19.

Marshall, Thomas I, Pankaj Chaudhary, Anna Michaelidesová, Jana Vachelová, Marie Davídková, Vladimir Vondráček, Giuseppe Schettino & Kevin M Prise. 2016. "Investigating the implications of a variable RBE on proton dose fractionation across a clinical pencil beam scanned spread-out Bragg peak." *International Journal of Radiation Oncology* Biology* Physics* 95(1):70–77.

McAneney, H & SFC O'Rourke. 2007. "Investigation of various growth mechanisms of solid tumour growth within the linear-quadratic model for radiotherapy." *Physics in Medicine and Biology* 52(4):1039.

McIlwrath, Amanda J, Paul A Vasey, Gillian M Ross & Robert Brown. 1994. "Cell cycle arrests and radiosensitivity of human tumor cell lines: dependence on wild-type p53 for radiosensitivity." *Cancer Research* 54(14):3718–3722.

Newhauser, Wayne D, Rui Zhang, Timothy G Jones, Annelise Giebeler, Phillip J Taddei, Robert D Stewart, Andrew Lee & Oleg Vassiliev. 2015. "Reducing the cost of proton radiation therapy: The feasibility of a streamlined treatment technique for prostate cancer." *Cancers* 7(2):688–705.

Nohadani, Omid, Joao Seco & Thomas Bortfeld. 2010. "Motion management with phase-adapted 4D-optimization." *Physics in Medicine and Biology* 55(17):5189.

Nøttrup, Trine Juhler, Stine Sofia Korreman, Anders Navrsted Pedersen, Lasse Rye Aarup, Håkan Nyström, Mikael Olsen & Lena Specht. 2007. "Intra-and interfraction breathing variations during curative radiotherapy for lung cancer." *Radiotherapy and oncology* 84(1):40–48.

Ödén, Jakob, Kjell Eriksson & Iuliana Toma-Dasu. 2017. "Inclusion of a variable RBE into proton and photon plan comparison for various fractionation schedules in prostate radiation therapy." *Medical physics* 44(3):810–822.

Oliver, R. 1964. "A comparison of the effects of acute and protracted gamma-radiation on the growth of seedlings of Vicia faba: part II. Theoretical calculations." *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine* 8(5):475–488.

Paganetti, Harald. 2003. "Significance and implementation of RBE variations in proton beam therapy." *Technology in cancer research & treatment* 2(5):413–426.

Paganetti, Harald. 2014. "Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer." *Physics in Medicine & Biology* 59(22):R419.

Paganetti, Harald, Andrzej Niemierko, Marek Ancukiewicz, Leo E Gerweck, Michael Goitein, Jay S Loeffler & Herman D Suit. 2002. "Relative biological effectiveness (RBE) values for proton beam therapy." *International Journal of Radiation Oncology* Biology* Physics* 53(2):407–421.

Paganetti, Harald, Pawel Olko, Hubert Kobus, Regina Becker, Thomas Schmitz, Michael PR Waligorski, Detlef Filges & Hans-Wilhelm Müller-Gärtner. 1997. "Calculation of relative biological effectiveness for proton beams using biological weighting functions." *International Journal of Radiation Oncology* Biology* Physics* 37(3):719–729.

Parmar, Chintan, Patrick Grossmann, Johan Bussink, Philippe Lambin & Hugo JWL Aerts. 2015. "Machine learning methods for quantitative radiomic biomarkers." *Scientific Reports* 5:13087.

Pawlik, Timothy M & Khandan Keyomarsi. 2004. "Role of cell cycle in mediating sensitivity to radiotherapy." *International Journal of Radiation Oncology* Biology* Physics* 59(4):928–942.

Pflugfelder, D, J J Wilkens & U Oelfke. 2008*a*. "Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy." *Physics in Medicine Biology* 53(6):1689.
URL: *http://stacks.iop.org/0031-9155/53/i=6/a=013*

Pflugfelder, D, JJ Wilkens & U Oelfke. 2008*b*. "Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy." *Physics in Medicine & Biology* 53(6):1689.

Pineau, Joelle, Geoff Gordon, Sebastian Thrun et al. 2003. "Point-based value iteration: An anytime algorithm for POMDPs." In *IJCAI*. Vol. 3 pp. 1025–1032.

Poupart, Pascal. 2005. *Exploiting structure to efficiently solve large scale partially observable Markov decision processes.* University of Toronto.

Rew, DA & GD Wilson. 2000. "Cell production rates in human tissues and tumours and their significance. Part II: clinical data." *European journal of surgical oncology* 26(4):405–417.

Ribba, Benjamin, Olivier Saut, Thierry Colin, Didier Bresch, Emmanuel Grenier & Jean-Pierre Boissel. 2006. "A multiscale mathematical model of avascular tumor growth to investigate the therapeutic benefit of anti-invasive agents." *Journal of Theoretical Biology* 243(4):532–541.

Rietzel, Eike, George TY Chen, Noah C Choi & Christopher G Willet. 2005. "Four-dimensional image-based treatment planning: Target volume segmentation and dose calculation in the presence of respiratory motion." *International Journal of Radiation Oncology* Biology* Physics* 61(5):1535–1550.

Rivard, Mark J, Jack LM Venselaar & Luc Beaulieu. 2009. "The evolution of brachytherapy treatment planning." *Medical Physics* 36(6):2136–2153.

Robertson, James B, Jerry R Williams, Robert A Schmidt, John B Little, Daniel F Flynn & Herman D Suit. 1975. "Radiobiological studies of a high-energy modulated proton beam utilizing cultured mammalian cells." *Cancer* 35(6):1664–1677.

Rockne, R, EC Alvord, JK Rockhill & KR Swanson. 2009. "A mathematical model for brain tumor response to radiation therapy." *Journal of Mathematical Biology* 58(4-5):561.

Roman, Nicholas O, Wes Shepherd, Nitai Mukhopadhyay, Geoffrey D Hugo & Elisabeth Weiss. 2012. "Interfractional positional variability of fiducial markers and primary tumors in locally advanced non-small-cell lung cancer during audiovisual biofeedback radiotherapy." *International Journal of Radiation Oncology* Biology* Physics* 83(5):1566–1572.

Ross, Keith W. 1989. "Randomized and past-dependent policies for Markov decision processes with multiple constraints." *Operations Research* 37(3):474–477.

Saberian, Fatemeh, Archis Ghate & Minsun Kim. 2015. "Optimal fractionation in radiotherapy with multiple normal tissues." *Mathematical medicine and biology: a journal of the IMA* 33(2):211–252.

Sachs, RK & LR Hlatky. 1990. "Dose-rate dependent stochastic effects in radiation cell-survival models." *Radiation and Environmental Biophysics* 29(3):169–184.

Scott, Jacob G, Anders Berglund, Michael J Schell, Ivaylo Mihaylov, William J Fulp, Binglin Yue, Eric Welsh, Jimmy J Caudell, Kamran Ahmed, Tobin S Strom et al. 2017. "A genome-based model for adjusting radiotherapy dose (GARD): a retrospective, cohort-based study." *The Lancet Oncology* 18(2):202–211.

Semenenko, VA & XA Li. 2008. "Lyman–Kutcher–Burman NTCP model parameters for radiation pneumonitis and xerostomia based on combined analysis of published clinical data." *Physics in Medicine and Biology* 53(3):737.

Smallwood, Richard D & Edward J Sondik. 1973. "The optimal control of partially observable Markov processes over a finite horizon." *Operations research* 21(5):1071–1088.

Sondik, Edward J. 1971. "The Optimal Control of Partially Observable Markov Decision Processes." *PhD the sis, Stanford University* .

Spaan, Matthijs TJ & Nikos Vlassis. 2005. "Perseus: Randomized point-based value iteration for POMDPs." *Journal of artificial intelligence research* 24:195–220.

Stroom, Joep C & Ben JM Heijmen. 2002. "Geometrical uncertainties, radiotherapy planning margins, and the ICRU-62 report." *Radiotherapy and Oncology* 64(1):75–83.

Suh, Yelin, Walter Murray & Paul J Keall. 2014. "Imrt treatment planning on 4d geometries for the era of dynamic mlc tracking." *Technology in Cancer Research & Treatment* 13(6):505–515.

Thames, Howard D. 1985. "An âĂŸincomplete-repairâĂŹmodel for survival after fractionated and continuous irradiations." *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine* 47(3):319–339.

Thirion, J-P. 1996. "Non-rigid matching using demons." In *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on.* IEEE pp. 245–251.

Trofimov, Alexei, Eike Rietzel, Hsiao-Ming Lu, Benjamin Martin, Steve Jiang, George TY Chen & Thomas Bortfeld. 2005. "Temporo-spatial IMRT optimization: concepts, implementation and initial results." *Physics in Medicine & Biology* 50(12):2779.

Underberg, René WM, Frank J Lagerwaard, Johan P Cuijpers, Ben J Slotman, John R Van Sörnsen De Koste & Suresh Senan. 2004. "Four-dimensional CT scans for treatment planning in stereotactic radiotherapy for stage I lung cancer." *International Journal of Radiation Oncology* Biology* Physics* 60(4):1283–1290.

Unkelbach, Jan, Pablo Botas, Drosoula Giantsoudi, Bram L Gorissen & Harald Paganetti. 2016. "Reoptimization of intensity modulated proton therapy plans based on linear energy transfer." *International Journal of Radiation Oncology* Biology* Physics* 96(5):1097–1106.

Unkelbach, Jan, Thomas Bortfeld, Benjamin C. Martin & Martin Soukup. 2009. "Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning." *Medical Physics* 36(1):149–163.
URL: *http://dx.doi.org/10.1118/1.3021139*

Unkelbach, Jan, Timothy CY Chan & Thomas Bortfeld. 2007. "Accounting for range uncertainties in the optimization of intensity modulated proton therapy." *Physics in Medicine & Biology* 52(10):2755.

Usher, JR. 1980. "Mathematical derivation of optimal uniform treatment schedules for the fractionated irradiation of human tumors." *Mathematical Biosciences* 49(3-4):157–184.

Veiga, Catarina, Guillaume Janssens, Thomas Baudier, Lucian Hotoiu, Sebastien Brousmiche, Jamie McClelland, Ching-Ling Teng, Lingshu Yin, Gary Royle & Boon-Keng Kevin Teo. 2017. "A comprehensive evaluation of the accuracy of CBCT and deformable registration based dose calculation in lung proton therapy." *BioMedical Physics & Engineering Express* 3(1):015003.

Vlassis, Nikos, Matthijs TJ Spaan et al. 2004. "A fast point-based algorithm for POMDPs." In *Benelearn 2004: Proceedings of the Annual Machine Learning Conference of Belgium and the Netherlands*. pp. 170–176.

Wang, Yi, Jason A Efstathiou, Hsiao-Ming Lu, Gregory C Sharp & Alexei Trofimov. 2013. "Hypofractionated proton therapy for prostate cancer: Dose delivery uncertainty due to interfractional motion." *Medical Physics* 40(7).

Warkentin, Brad, Pavel Stavrev, Nadia Stavreva, Colin Field & B Gino Fallone. 2004. "A TCP-NTCP estimation module using DVHs and known radiobiological models and parameter sets." *Journal of Applied Clinical Medical Physics* 5(1):50–63.

Wedenberg, Minna, Bengt K Lind & Björn Hårdemark. 2013. "A model for the relative biological effectiveness of protons: the tissue specific parameter α/β of photons is a predictor for the sensitivity to LET changes." *Acta Oncologica* 52(3):580–588.

Wheldon, TE & J Kirk. 1976. "Mathematical derivation of optimal treatment schedules for the radiotherapy of human tumours. Fractionated irradiation of exponentially growing tumours." *The British Journal of Radiology* 49(581):441–449.

Wheldon, TE, J Kirk & JS Orr. 1977. "Optimal radiotherapy of tumour cells following exponential-quadratic survival curves and exponential repopulation kinetics." *The British Journal of Radiology* 50(597):681–682.

Wieser, Hans-Peter, Eduardo Cisternas, Niklas Wahl, Silke Ulrich, Alexander Stadler, Henning Mescher, Lucas-Raphael Müller, Thomas Klinge, Hubert Gabrys, Lucas Burigo et al.

2017. “Development of the open-source dose calculation and optimization toolkit matRad.” *Medical physics* 44(6):2556–2568.

Wilkens, Jan J & Uwe Oelfke. 2005. “Optimization of radiobiological effects in intensity modulated proton therapy.” *Medical physics* 32(2):455–465.

Willers, Henning, Antino Allen, David Grosshans, Stephen J McMahon, Cläre von Neubeck, Claudia Wiese & Bhadrasain Vikram. 2018. “Toward A variable RBE for proton beam therapy.” *Radiotherapy and Oncology* .

Wilson, Robert R. 1946. “Radiological use of fast protons.” *Radiology* 47(5):487–491.

Withers, H Rodney. 1985. “Biologic basis for altered fractionation schemes.” *Cancer* 55(S9):2086–2095.

Yan, Di & David Lockman. 2001. “Organ/patient geometric variation in external beam radiotherapy and its effects.” *Medical Physics* 28(4):593–602.

Yang, Yong & Lei Xing. 2005. “Optimization of radiotherapy dose-time fractionation with consideration of tumor specific biology.” *Medical physics* 32(12):3666–3677.

Zhang, Tiezhi, Robert Jeraj, Harry Keller, Weiguo Lu, Gustavo H Olivera, Todd R McNutt, Thomas R Mackie & Bhudatt Paliwal. 2004. “Treatment plan optimization incorporating respiratory motion.” *Medical physics* 31(6):1576–1586.