Analytical Models and Data-Driven Methods for Radiation Therapy Treatment Planning

by
Saba Ebrahimi

A Dissertation submitted to the Department of Industrial Engineering,

Cullen College of Engineering

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Industrial Engineering

Chair of Committee: Gino Lim

Committee Member: Taewoo Lee

Committee Member: Ying Lin

Committee Member: David Mayerich

Committee Member: Wenhua Cao

University of Houston
August 2021

Dedicated to

my dear parents Afsaneh and Reza,

my sweet brother Amir Mahdi,

and all student victims of Flight 752 tragedy

with all my love

Acknowledgements

It is a pleasure to thank those who made this dissertation possible. First of all, my sincere gratitude goes to my advisor, Dr. Gino Lim, for providing all the guidance, mentorship, warm encouragement, kind support, and patience throughout this program that have given me more power and spirit to excel in the research. Many thanks to the dissertation committee members, Dr. Taewoo Lee, Dr. Ying Lin, Dr. David Mayerich, and Dr. Wenhua Cao for their valuable comments that improved the contents of the work.

I am extremely grateful to Dr. Radhe Mohan and Dr. Wenhua Cao for giving me the opportunity to work with one of the greatest research groups at one of the best cancer centers. It has been a pleasant experience working with them and I am sincerely thankful for their gracious support and inspiring suggestions. Also, I would like to thank Dr. Brian Hobbs and Dr. Steven Lin for their invaluable ideas and helpful comments on my research.

Big thank you to my wonderful friends, Saeed, Hedieh, Hamid, Ali, Mohsen, and Maryam who made the past five years very enjoyable. I am also thankful to my dear friends Azin, Zahed, Navid, Anahita, Xuemin, Mohammad, and Saeedeh for their warm hearts and supports.

Last and most importantly, my warmest gratitude goes to my beautiful mother and my kind father for all of their sacrifices, caring, encouragement, unconditional love, and support. I am also thankful to my sweet brother for the love and joy he has given to me during my life. The love and encouragement from my family has made it possible for me to complete the Ph.D. journey. Finally, I would like to thank my best friend, Saeed, who was there to support me and cheer me up during difficult days with his priceless love.

Abstract

The clinical goal of radiation therapy (RT) is to maximize the tumor damage and kill all the cancerous cells while minimizing toxic effects on surrounding healthy tissues during the course of treatment. Adaptive radiation therapy (ART) has been widely used to adjust the radiation dose in response to potential changes in tumor volume during the treatment to reduce the radiation toxicity in healthy organs. One of the key challenges in ART is to determine the best time to adapt the plan in response to uncertain tumor biological responses to radiation during the treatment. Tumor biological response change dynamically over time and can be different from one patient to another. Therefore, considering tumor biological responses to radiation in ART treatment planning is challenging due to the high levels of uncertainty in biological factors. Determining the possibility of treatment side-effects for each patient before starting the treatment is another challenge in radiation therapy treatment planning. This dissertation focuses on a combination of optimization, deep learning, and statistical methods to address the aforementioned challenges in this field and improve the survival of cancer patients treated with radiation therapy. We will tackle this problem from two different perspectives: (1) developing effective personalized radiation therapy treatment plans and (2) predicting possible critical side-effects of the treatment for each patient before the treatment.

First, we propose an automated radiation therapy treatment planning framework using reinforcement learning (RL) which incorporates uncertainty in tumor biological responses during treatment to find the optimal policy for ART. We also provide a novel tumor response model to estimate tumor volume changes and radiation responses during the treatment. This approach helps the decision-maker to control both biological and physical aspects of the treatment and achieve a robust solution under biological uncertainties without dealing with complex optimization models. The presented method provides much-needed flexibility in which a plan can be customized based on the patient case, cancer type, and the decision-maker's preference on treatment outcomes.

Second, we address one of the critical radiation therapy treatment side-effects known as radiation-induced lymphopenia (RIL). RIL occurs due to a severe reduction in the absolute lymphocyte count (ALC) after radiation exposure and can seriously affect patient survival. Therefore, we aim to assess the role of radiation therapy in ALC depletion to determine high-risk patients. To accomplish this goal, two mathematical models are proposed to approximate lymphocyte depletion based on radiation dose distributions and the ALC baseline for radiation therapy patients. Finally, we compare the potential post-treatment lymphocyte survival outcomes in cancer patients for photon and proton-based RT modalities.

Third, we develop a hybrid deep learning model in a stacked structure to predict the ALC depletion trend throughout radiation therapy treatment for cancer patients based on the pretreatment clinical information. Then, we extend the model to account for making predictions after the initial phase of treatment (e.g., at the end of week 1). A discriminative kernel is also developed to extract and evaluate the importance of temporal features. The presented deep learning structure can efficiently use information from different groups of clinical features to predict ALC depletion without requiring a large amount of data to process too many features while reducing bias and generalization error. This approach helps the physicians to identify patients at risk of severe RIL who might benefit from modified treatment approaches which ultimately improve survival of the patients.

In the last part of this dissertation, we provide an approach to estimate prediction intervals for ALC values. The proposed approach enables practical implications of predictive models in clinical decision-making by estimating the individualized predictive uncertainties. Finally, a comprehensive hybrid decision-making framework is proposed to assess RIL risk for a given patient based on a given treatment plan and its predicted post-treatment lymphocyte survival outcome. This decision-making framework can be used as a guide for physicians to take advantage of advanced deep learning models and make appropriate decisions in selecting the safest treatment plan for an individual patient in clinics.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background & Motivation

Cancer is one of the primary health problems in the world, and it is the second leading cause of death in the United States, accounting for 21% of all deaths. In 2020, about 1.8 million new cases of cancer are projected to occur in the United States, according to the American Cancer Society [1]. There are several cancer treatment options available that can be selected based on the type of cancer and the patient's health condition. Among these methods, radiation therapy (RT) is a common treatment modality for many cancer patients. In most cases, patients receive radiotherapy in addition to surgery or chemotherapy, but in some cases, it can be used alone as a primary treatment.

1.1.1 Radiation Therapy Treatment

Radiation therapy delivers high-energy ionizing radiation to the tumor to kill cancerous cells by damaging the tumor cells' DNA and causing a double-strand break. The DNA damage to the tumor cells stops them from growing and dividing. As a result, the tumor cells will die and will cease to regenerate. It takes days or weeks of treatment to sufficiently damage the tumor cells' DNA enough to kill the cells. After the treatment, cancer cells continuously proceed to die for weeks or months in order to be completely eradicated.

The two main types of radiation therapy based on the source of radiation are internal radiation therapy and external beam radiation therapy. This study focuses on external beam radiation therapy, which is the most widely used type of radiation therapy treatment. In external beam radiotherapy, the ionizing radiation (e.g., photon, proton, etc.) is delivered from a machine outside of the patient's body targeted towards the cancerous area and goes through a particular part of the patient's body to eradicate tumor cells. The external beam radiation therapy machine can move around and deliver the radiation from many different angles through different beams. There are different types of beam modalities used for external beam radiation treatment, such as photon beams and proton beams. In a photon therapy treatment (e.g., Intensity modulated radiation therapy (IMRT)), photon beams are used to treat tumor cells, whereas proton therapy treatment (e.g., Intensity modulated proton therapy (IMPT)) uses beams of proton particles for the operation. Each beam of radiation is partitioned into a large set of "beamlets" with individually adjustable intensities to deliver different doses of radiation across the tumor. Our work in this study spans both of these treatment modalities.

During radiation therapy treatment, the radiation can also damage healthy cells in the surrounding organs. Although the normal cells are less sensitive to radiation and can heal from radiation damage, radiation can cause serious health problems in some critical organs. A schematic illustration of clinical targets including the gross tumor volume (GTV), clinical target volume (CTV), planning target volume (PTV), as well as the organs at risk (OAR) are denoted in Figure 1.1. High doses of radiation are required to eradicate the tumor for the majority of cancer patients, which is usually higher than the OAR tolerance. Thus, the surrounding OARs require maximal protection. Therefore, the main clinical goal of radiation therapy is to maximize the tumor damage and kill all cancerous cells while minimizing the toxic effects on surrounding healthy tissues during treatment. Thus, such treatment must be carefully planned to achieve this goal and improve the patient's survival.

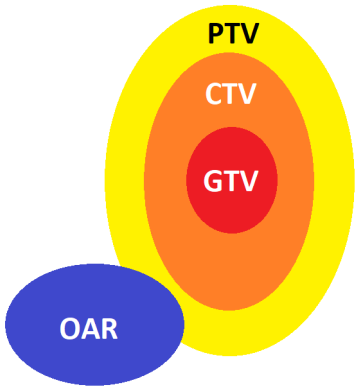


Figure 1.1: Schematic illustration of clinical targets (the gross tumor volume (GTV), clinical target volume (CTV), planning target volume (PTV)) and the organ at risk (OAR) volumes

1.1.2 Radiation Therapy Treatment Planning Procedure

Generally, the starting point to create a treatment plan is to acquire digital images of the patient's internal anatomy using medical imaging techniques, such as Computed Tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET). The images will be used to identify the structures and determine the geometry information, such as organs and tumor locations, organ sizes, and the isocenter. Then, target volumes and normal structures are manually contoured on the axial slices of the planning CT scan by a physician. Based on the extracted information, a treatment protocol will be defined for the patient by a radiation oncologist, which includes prescription doses to the target volumes and constraints to OARs, known as the maximum OAR dose tolerance. A prescription dose is the dose level necessary to eradicate target cells, while a tolerance dose is the level above at which complications for healthy tissues may occur. The radiation dose is measured in Gray (Gy) and mostly varies from 60 to 80 Gy.

Next, the planners usually use mathematical modeling and optimization techniques to deliver enough doses to the tumor while sparing the healthy tissue around the tumor as

much as possible. Finally, after the physician's approval, the generated plan and the extracted details will be utilized to set the treatment machine and begin the treatment. In general, the treatment region will be divided into three parts: (1) target(s) or planning target volume, (2) the healthy organs located close to the targets, which are called critical structures or organs at risk (OARs), and (3) normal healthy tissues. Figure 1.2 showcases the summary of the radiation therapy treatment process.

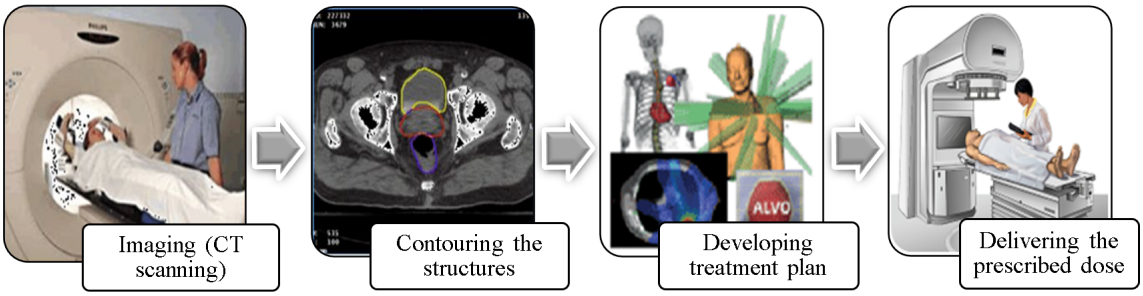


Figure 1.2: The steps of the radiation therapy planning workflow

1.1.3 Radiation Therapy Treatment Fractionation

The prescribed RT dose is usually divided into several sessions called fractions to achieve tumor control while enabling OAR cells to recover. This procedure in radiation therapy treatment planning is known as "fractionation". Fractionation not only provides enough time for OARs to recover and reduce the damage to healthy cells, but it also allows us to deliver the radiation dose during the radiation-sensitive phase of the tumor cell cycle and avoid the radio-resistant phase of the cycle. However, there are still some challenges and unanswered questions at this point, What is the best fractionation schedule for each patient? When is the optimal time to change the dose per fraction? How can we improve the quality of the treatment plan under uncertainties?

1.2 Problem Description

The primary goal of this dissertation research is to address some of the unanswered challenges in radiation therapy treatment planning to enhance the quality of the treatment and ultimately improve patients' survival. The following introduce the challenges we attempt to resolve:

1.2.1 Radiobiological Effects of Radiation Therapy and Tumor Cell Dynamics

Recent studies showed that both tumor and OAR cells are changing dynamically, and their biological responses are different over time [2, 3]. Other than the dose of radiation absorbed, there are some biological factors that affect the tumor's biological response to a given radiation dose. The most important biological factors in determining the tumor biological response are known as the five R's of radiation therapy, which include repair, repopulation, radiosensitivity, redistribution, and re-oxygenation [2, 4]. Tumor and healthy tissue cells can repair the sublethal damage of radiation. The radiation dose and the available time to repair could significantly affect the repair capacity of the cells. Therefore, a repair can be considered an important factor in fractionated radiation therapy to reduce the OAR toxicity, because healthy tissue cells are less sensitive to radiation, they are able to repair some of the radiation damage between fractions. Both tumor and healthy cells can proliferate during the course of treatment. Tumor cells can proliferate at different rates during the treatment due to the different amounts of damage and cell death that occur during the radiation. Tumors with higher rates of proliferation may need a larger amounts of radiation dose to be completely eradicated. Radiosensitivity indicates the relative susceptibility of the cells to radiation. Some tissues are highly radiosensitive (i.e., early responding tissues), and some tissues have lower radiosensitivity (i.e., late responding).

Furthermore, there are two more biological factors that address the biological response

of a tumor to radiation therapy and affect the total dose required for treatment, including redistribution and reoxygenation. Redistribution refers to the fact that the radiosensitivity of the cells can vary while in the different phases of the cell cycle (see Figure 1.3). Basically, the radiosensitivity of the cells is lowest during the synthesis phase (S), and it is highest when in the mitotic phase (M). During several radiation therapy treatment fractions, the tumor cells become radio-resistant after DNA damage caused by radiation and will enter into the resting phase to repair the sublethal damages. Variable radiosensitivity of the cells during the RT treatment can affect the treatment efficiency and outcome. Moreover, reoxygenation of hypoxic cells happens during fractionated treatment as a result of tumor shrinkage. Thus, the cells will be more radio-resistant and a larger dose of radiation will be needed to eradicate them.

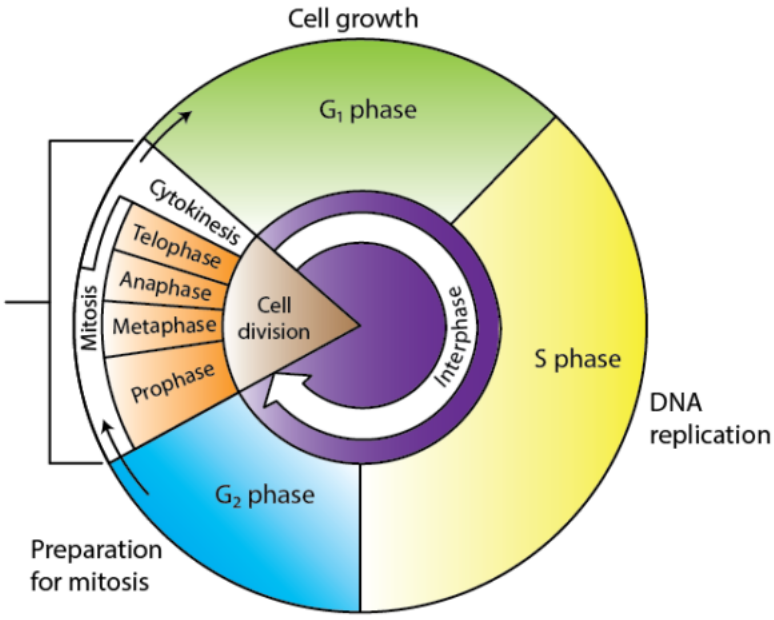


Figure 1.3: Schema of tumor cell cycle phases

Hence, biological factors vary based on cell cycle phase and gene level activities during

the radiation therapy treatment, which can result in different biological responses. Additionally, the biological response of a tumor is different from patient to patient [5]. Therefore, a general monotone treatment plan might not be efficient for all patients. However, conventional plans mostly ignore the dynamic nature of biological processes. In this regard, developing a personalized treatment plan that considers variable tumor biological response can significantly improve the quality of treatment. Aside from biological factors and patient characteristics, variable dose fractionation can also affect the tumor's biological response during radiation therapy treatment. Better treatment outcomes can be achieved by modifying the amount of dose per fraction in fractionated radiation therapy based on the tumor's biological response. Thus, understanding the cell dynamics and variable biological response of a tumor to the radiation can have a significant role in finding the optimal fractionation treatment plan.

1.2.2 Adaptive Radiation Therapy Considering Biological Response of a Tumor

The dynamic biological response of the tumor cells signifies the importance of using fractionation and adapting the treatment plan to tumor volume changes during the radiation therapy treatment. Adaptive radiation therapy (ART) is an iterative process that adjusts radiation dosage based on the information acquired between fractions. This type of treatment modification allows for customized fractionated dose delivery to mitigate treatment variations and re-optimize the treatment plan early during therapy. Figure 1.4 represents the workflow for ART treatment planning.

The ideal approach for considering dynamic tumor changes in determining the next treatment session dosage is to image patients during each visit and use updated tumor contours from the patient's image to adapt the treatment plan and find the fraction dose. This approach is not clinically practical because imaging the patient during each visit amidst

the treatment period can be costly, time-consuming, and prone to human error. In practice, trade-offs must be made considering costs, timing, and the recommended number of adaptive plans. Therefore, finding the optimal timing and policy for adaptation is necessary to improve the clinical feasibility of ART.

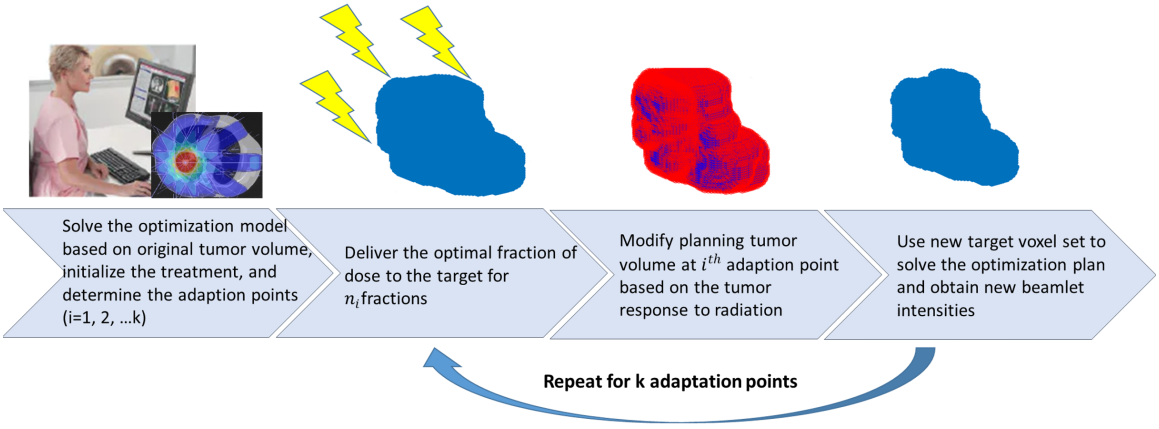


Figure 1.4: Adaptive radiation therapy workflow

Considering biological effects in ART treatment planning is challenging due to the inherent uncertainty of biological factors, the quality of treatment can be compromised. Also, the complexities of tumor biological responses and uncertainty inherent in biological factors make it difficult to determine how a treatment plan should be adapted. Therefore, one of the key challenges in ART is determining the best time to adapt the plan in response to tumor biological response and volume changes during the course of treatment to improve the final treatment outcome.

1.2.3 Radiation-Induced Immunosuppression

The effectiveness of radiation therapy treatment relies on the body's immune system response to the treatment. However, radiation therapy also suppresses the immune system through the killing of circulating lymphocytes in the radiation field. Lymphocytes have a significant role in the body's anticancer immune response, and they are highly radiosensitive even at low doses. So, lymphocytes can be killed as a result of low and moderate doses

of radiation exposure. Severe reduction in absolute lymphocyte count (ALC) will happen which will cause radiation-induced lymphopenia (RIL). Severe RIL can be diagnosed by measuring the ALC in the patient's blood. Grade 3 and grade 4 lymphopenia, according to the National Cancer Institute's Common Toxicity Criteria for Adverse Events, version 5.0, are defined as an ALC less than 200 $cells/\mu L$ and 500 $cells/\mu L$ during and immediately following the course of RT, respectively.

ALC has been shown to be an independent predictor of survival from cancer; and High-grade RIL (i.e., grade 3 with $ALC \leq 500\ cells/\mu L$ or grade 4 with $ALC \leq 200\ cells/\mu L$) has been highly associated with reduced overall patient survival for most of the tumor types, such as lung, esophagus, pancreas, breast, and bladder cancer. Also, preservation of the lymphocytes will increase treatment effectiveness and improve overall patient survival. Thus, it is critical to analyze the role of radiation therapy in ALC depletion and understand the clinical features such as dosimetric factors, patient- and treatment-specific characteristics affecting RIL risk for RT patients. Dosimetric factors have been considered as one of the most important factors affecting RIL severity and incidents. Therefore, RIL risk could be further minimized by using a different radiation modality and changing the treatment plan. Nevertheless, the true impact of pretreatment clinical factors in determining RIL risk is still unclear and requires further study and elaborations.

Moreover, predicting RIL before treatment based on pretreatment clinical information would improve RT treatment planning and help physicians identify patients with a high-risk of severe lymphopenia occurrence and avoid RT treatment for these patients. As a result, mitigation strategies could be developed, or modified treatment approaches could be used for the identified patients that may ultimately improve their survival. However, comprehensive models or explicit approaches that can forecast the kinetics of lymphocyte loss after radiation exposures in order to identify patients at a high risk of developing severe RIL have not been explored yet.

It is critical to develop efficient prediction models based on pretreatment clinical information to forecast ALC depletion during RT treatment and identify patients who are at high risk of severe RIL. However, any prediction has uncertainties and reporting the uncertainty of a prediction is very important in clinical decision-making problems because poor performance of prediction models in clinical practice may have serious adverse consequences for patients. Therefore, it is crucial for clinicians to have some sense of prediction uncertainty for a given individual patient.

In summary, one of the key challenges in RT treatment planning is predicting severe radiation-induced ALC depletion and the associated predictive uncertainty for individual patients. Also, most of the comprehensive and advanced prediction models are too complex for use in clinics. So, providing a straightforward and flexible decision-making framework based on the prediction models could be very helpful to make the best use of advanced prediction models in clinical practices.

1.3 Objectives & Contributions

This thesis aims to address the aforementioned radiation therapy treatment planning challenges in three research studies by presenting the following contributions:

- In the first work, we develop a novel comprehensive biological response model that incorporates tumor cell death, repopulation, reoxygenation, radiosensitivity for tumor cells, and repair of healthy tissues cells to predict the radiation response of the tumor and OARs during the course of treatment without taking a significant amount of time and effort to collect large-scale data sets and the necessity for expensive CT images. Then, we propose an automated personalized radiation therapy treatment planning framework by combining Reinforcement Learning (RL) and an optimization method to find the optimal adaptation points for ART and dynamically adapt the

plan to the uncertain tumor biological responses over time. The proposed approach not only controls the biological aspect of the treatment and incorporates tumor biological response uncertainty, but it also ensures the dose-volume requirements and clinical limits of the treatment with no need for dealing with complex optimization models. Moreover, the reinforcement learning framework for ART planning helps the decision-maker achieve a robust solution under high levels of uncertainty in the biological parameters while reducing the variability in the solution and improving control on the worst-cases. Furthermore, this approach provides much needed flexibility in which a plan can be customized based on the patient case and the planner's preference on treatment outcomes and extended for wider applications.

- The second work studies radiation-induced immunosuppression challenges. We develop two mathematical prediction models to approximate lymphocyte depletion based on dose distributions. In the first model, we use a piecewise-linear relationship between lymphocyte survival and radiation dose. The second model assumes an exponential function for ALC depletion, and it uses a non-linear regression to estimate post-treatment ALC. Moreover, we investigate the relationship between severe lymphopenia and poor treatment outcomes and compare the potential post-treatment lymphocyte survival outcomes in esophageal cancer patients for photon and proton-based modalities.

- For the third work, we propose a hybrid deep learning model in a stacked structure to predict the ALC depletion trend during RT treatment for cancer patients based on pretreatment clinical information. The proposed model consists of four channels, one channel based on long short-term memory (LSTM) network and three channels based on a deep neural networks, to process four categories of features followed by a fully connected neural network to integrate the outputs of the four channels and predict the

weekly ALC values. The hybrid deep-stacked structure can efficiently use information from different groups of features with different characteristics to predict weekly ALC without requiring a large amount of data to process too many features while reducing bias and the adverse effects of any noise in the data. The developed model is flexible and can be extended easily to account for early-treatment predictions (i.e., at the end of week 1 or 2). So, a discriminative kernel was developed to extract temporal features and assign different weights to each part of the input sequence which enables the model to focus on the most relevant parts. This approach helps the physicians to identify high-risk patients and select them for modified treatment approaches or mitigation strategies.

- In the last part of this study, we develop a deep learning model to predict the weekly ALC values and their associated uncertainties in form of prediction intervals for individual patients. Estimation of prediction intervals for a given individual patient enables practical implications of predictive models in clinical decision-making by considering individualized prediction risks. Moreover, different groups of patients with different pretreatment characteristics can be assessed in terms of ALC prediction uncertainties based on results from the proposed deep learning model. We also propose a comprehensive hybrid decision-making framework to select patients for RT treatment using the predicted values of ALC and their associated risk to be used in clinical practice with the goal of avoiding severe RIL for cancer patients. This decision-making framework is flexible, straightforward, easy to interpret by clinicians, and can be modified to account for different levels of risk. This approach enables physicians to easily take the advantage of complex advanced deep learning models in their decisions and identify high-risk patients who may benefit from treatment modifications. Also, the proposed decision-making framework can be used to evaluate the effect of any treatment modifications on RIL risk for a given patient; so,

the safest treatment plan can be chosen for the patient.

1.4 List of Outcomes

Journal Publications

- Gino J. Lim, Laleh Kardar, Saba Ebrahimi, Wenhua Cao. "*A Risk-Based Modeling Approach for Radiation Therapy Treatment Planning Under Tumor Shrinkage Uncertainty*," European Journal of Operational Research, 2019.

- Saba Ebrahimi, Gino Lim, Amy Liu, Steven H. Lin, Susannah G. Ellsworth, Clemens Grassberger, Radhe Mohan, Wenhua Cao. "*Radiation-Induced Lymphopenia Risks of Photon Versus Proton Therapy for Esophageal Cancer Patients*," International Journal of Particle Therapy, 2021.

- Saba Ebrahimi, Gino J. Lim. "*A Reinforcement Learning Approach for Finding Optimal Policy of Adaptive Radiation Therapy Considering Uncertain Tumor Biological Response*," **Submitted to** Artificial Intelligence In Medicine, 2021.

- Saba Ebrahimi, Gino J. Lim, Brian Hobbs, Steven H. Lin, Radhe Mohan, Wenhua Cao. "*A Hybrid Deep Learning Model for Forecasting Lymphocyte Depletion During Radiation Therapy*," **Submitted to** Medical Physics, 2021.

- Saba Ebrahimi, Gino J. Lim, Brian Hobbs, Steven H. Lin, Radhe Mohan, Wenhua Cao. "*A hybrid decision-making framework to assess lymphopenia risk for radiation therapy treatment plans based on the ALC predictions and its predictive uncertainties*," **To be submitted**, 2021.

Conference Proceedings

- Saba Ebrahimi, Gino J. Lim,"*Robust Adaptive Approach Incorporating Tumor Shrinkage in Radiation Treatment Planning*," In proceedings of the 2020 IISE Annual Conference, New Orleans, LA, October 2020.

Conference Presentations

- Saba Ebrahimi, Wenhua Cao, Amy Liu, Gino Lim, Steven H. Lin, Radhe Mohan, "*Assessment of Radiation-Induced Lymphopenia Risks for Esophageal Patients-Planning Study Comparing Proton and Photon Therapy*," 2019 AAPM Annual Meeting, San Antonio, TX, July 2019.

- Gino J. Lim, Laleh Kardar, Saba Ebrahimi, Wenhua Cao "*A Risk-based Modeling Approach For Radiation Therapy Treatment Planning Under Tumor Shrinkage Uncertainty*," INFORMS Annual Meeting, Seattle, WA, October 2019.

- Saba Ebrahimi, Gino J. Lim "*An Automated Framework for Adaptive Radiation Therapy Considering Biological Uncertainties*," Virtual INFORMS Annual Meeting, Washington DC, November 2020.

- Saba Ebrahimi, Gino J. Lim, Brian Hobbs, Steven H. Lin, Radhe Mohan, Wenhua Cao. "*Forecasting Absolute Lymphocyte Count Depletion During and After Radiation Therapy Using Deep Learning*," 2021 Virtual AAPM Annual Meeting, Columbus, OH, July 2021.

- Saba Ebrahimi, Gino J. Lim "*Predicting Radiation-induced Lymphopenia Risk In Esophageal Patients Treated By Proton And Photon Therapy*," INFORMS Annual Meeting, Anaheim, CA, October 2021.

1.5 Organization

The remainder of this dissertation is organized as follows. In Chapter 2, we review the relevant literature on the tumor biological response model, biological-based treatment planning, ART, and radiation-induced immunosuppression. In Chapter 3, we develop our automated ART framework based on the predicted tumor biological response using reinforcement learning. Solution approaches are provided, and the models are tested and evaluated. In Chapter 4, we propose our mathematical prediction models to estimate ALC during radiation therapy treatment based on dose distributions. The relationship between severe lymphopenia and poor treatment outcomes are investigated for proton and photon radiation modalities in esophageal cancer patients. In Chapter 5, we develop a hybrid deep learning model to predict weekly ALC depletion trends during the course of RT treatment based on pretreatment and early treatment clinical information. The training procedure, data descriptions are provided, and the performance of the model is evaluated and tested for a cohort of esophageal cancer patients. In Chapter 6, we provide an approach to estimate prediction intervals for estimated ALC values based on our proposed hybrid deep learning model. The effect of different patient- and treatment-specific factors on RIL risk are analyzed. Moreover, a comprehensive hybrid decision-making framework is proposed to assess RIL risk for a given patient based on a given treatment plan. The decision-making framework is evaluated using real data under different scenarios. Finally, we conclude this dissertation and discuss the directions of the future research in Chapter 7.

Chapter 2

Literature Review

2.1 Tumor Biological Response Models

Recent studies have shown that both tumor and OAR cells change dynamically, and their biological responses differ over time. The repairing of healthy cells, reoxygenation of tumor cells, repopulation of tumor cells, and radiosensitivity are important biological factors in determining the tumor and OAR response to radiation therapy treatment [2, 3]. Biological factors have been addressed in modeling tumor response to radiation during the treatment to determine fractionation dose and evaluate treatment plans.

A linear-quadratic (LQ) model [6, 7] is one of the most common radiation response models in fractionated radiation therapy. In this model, the tumor or OAR survival fraction is defined as a function of the radiation dose and can be used to find the fractionation dose [7]. More comprehensive tumor response models were developed to consider different biological factors, such as tumor repopulation [8–10].

Several studies have explored the effect of tumor repopulation status on the amount of dose per fraction and the total number of fractions [8, 9]. They suggested that faster-growing tumors need to be treated in a shorter treatment length using a higher radiation dose at each fraction. Further, Bortfeld et al. [10] included tumor repopulation in the LQ model for radiation therapy treatment planning to find the total number of treatment

days and fraction dose. Also, multiple studies have investigated tumor growth models considering exponential growth models [11–13] or Gompertzian growth curves [14, 15]. The advantages of using Bayesian modeling approaches to develop a patient-specific tumor growth model have been explored in several studies [16–19].

Besides tumor repopulation, other important factors, such as redistribution, repair of sublethal damage, and reoxygenation were considered in several studies to address the biological response of tumors [20, 21]. Yang and Xing [21] developed an LQR-based radiation therapy fractionation planning framework to minimize the ratio of the OAR's biological effective dose to the biological effective dose of the tumor. Furthermore, Jeong et al. [22] developed a tumor response model to assimilate hypoxia and proliferation interplay by considering three cell compartments including proliferating, intermediate, and hypoxic based on different levels of oxygen and glucose availability. As a result of mitotic death, they assumed that intermediate and hypoxic cells move into the proliferative and the intermediate compartments, respectively. They also proposed an approach to determine the initial compartmental cell distribution based on local growth fraction (GF) and volume doubling time (DT) values. Moreover, OAR repair is a major biological factor in healthy tissue response, which is higher than tumor cells' repair capacity and can be considered in the radiation therapy treatment planning [4].

None of the above studies considered all of the main biological factors. In this study, we develop a novel biological response model that incorporates tumor cell death, repopulation, reoxygenation, radiosensitivity, and OAR repair.

2.2 Biological-Based Treatment Planning

Several studies have shown the possibility of achieving better treatment outcomes by modifying the dose of radiation per fraction in fractionated radiation therapy based on the tumor biological response. Information from biological images has been used in several

studies to find biologically conformal nonuniform doses [23–25].

Various models have been proposed to incorporate biological response in radiation therapy treatment planning. Dynamic programming has been used in several studies to account for dynamic tumor response during radiation therapy and to find the optimal tumor fractionation dose [26], OAR repair [27], tumor repopulation [10, 28] and tumor shrinkage [29].

Kim et al. [27] used an LQ based response model through a finite-horizon Markov Decision Process model to find the tumor fractionation dose. Ghate et al. [28] reviewed a stochastic control framework based on biological images and response models to achieve personalized treatment plans that dynamically adapt to a tumor's uncertain biological response over time. Another study by Bortfeld et al. [10] included tumor repopulation in the response model in a dynamic programming framework for radiation therapy treatment planning to find the total number of treatment days and the optimal dose per fraction.

A recent study by Nohadani and Roy [30] showed that utilizing robust optimization techniques to account for cell oxygenation during radiation therapy can improve tumor control in a prostate cancer case. Unkelbach et al. [29] proposed a dynamic model to account for tumor shrinkage and tumor cell repopulation in liver cancer. In this dynamic model, the number of viable tumor cells at each stage is calculated by multiplying two exponential functions based on cell kill and cell repopulation in the previous stage. However, a comprehensive model to incorporate all important biological factors in a dynamic framework has not been studied yet.

Biological response to radiation varies from one patient to another [5]. Also, radiosensitivity and tumor proliferation are associated with tumor cell cycle phases and gene level activities [31]. Huang et al. [32] showed the advantages of personalized modeling on biological mechanisms by developing a volume-based tumor response model to predict the clinical outcome of radiation therapy for cervical cancer patients. Therefore, personalized

radiation therapy treatment planning has attracted researchers' attention in biological radiation therapy planning, which claims the importance of developing a different treatment plan for each patient as a response to the varied tumor response. Therefore, in this study, we focus on developing planning approaches that enable personalization of RT treatment plans.

2.3 Adaptive Radiation Therapy Treatment Planning

The literature shows that adapting a fractionated radiation therapy treatment in response to changes that occur during the course of treatment, known as adaptive radiation therapy (ART) method, improves treatment quality in terms of normal-tissue sparing and tumor cell reduction [33–36] as well as treatment cost and time [37, 38].

Veresezan et al. [38] recommended that error calculations and imaging studies should be repeated to verify treatment accuracy. So, an ideal approach to consider dynamic tumor changes will be to take images of the patient at every visit, update tumor contours, and revise the treatment plan if a significant change was observed in the tumor geometry. However, daily imaging may not be useful in practice because any changes over the span of a day may not be significant enough to modify the existing treatment plan. More importantly, imaging the patient at every visit during the treatment period can be costly, time-consuming and prone to human errors. Therefore, a compromise must be made considering tumor geometry change, costs, timing, and the recommended number of adaptive plans. This is the primary motivation for finding the optimal timing for adaptation to improve the clinical feasibility of ART.

Several approaches have been proposed to optimally determine how often to conduct adaptation during the treatment based on the latest tumor geometry information, focusing on target-volume reduction [35, 39, 40] and the dose per volume received in the tumor [19, 41–43].

Saka et al. [39] developed an image-based adaptive IMRT optimization approach in which an adaptation was suggested once before the 25^{th} fraction and once after the 25^{th} fraction on the basis of the latest tumor geometry information. Guckenberger et al. [40] proposed adapting the plan once or twice in week 3 or week 5 for a non-small cell lung cancer (NSCLC) patient. Zheng et al. [41] proposed that the plan adaptation for lung cancer treatment should occur at the 15^{th} fraction. They also showed that the adaptation point should be before the 31^{st} fraction to provide the most clinical benefit. Berkovic et al. [42] demonstrated that an adaptation performed around the 15^{th} fraction was most beneficial in IMRT for lung cancer patients.

Several approaches have been proposed to optimally determine the frequency of adaptation during the treatment based on the latest tumor geometry information, focusing on target-volume reduction [35, 39, 40] and the dose per volume received in the tumor [19, 41–43]. Most studies suggest that the optimal time for adaptation is when an adequate target volume reduction was observed and maintained [38–40]. However, there are some conflicting reports regarding the time of the largest tumor volume reduction and the best time to adapt the plan to the tumor volume changes for different patients with different biological response characteristics. Therefore, more studies must be done towards customization of an ART treatment plan for each patient based on biological response to the treatment.

2.4 Reinforcement Learning

Recent studies have proposed machine learning (ML) techniques to predict radiation therapy outcomes [44, 45], classify patients who would benefit from ART [46], and determine the ideal time for adaptation in ART [42, 47]. However, most of these approaches require large data sets which are not available for different types of cancer.

Reinforcement learning (RL) is one of the modern machine learning algorithms that features modeling of sequential data based on the interactions between an agent and an

environment. RL can be a good alternative to dynamic programming when there is a high level of uncertainty in a sequential decision-making problem because RL can generate a robust and risk-averse solution by incorporating the uncertainty in its environment.

Deep reinforcement learning (DRL) algorithms have been applied to find the best policy (a sequence of decisions) in many diverse fields such as robotics [48], computer vision [49], energy [50, 51], and healthcare [52, 53]. DRL approaches have been successfully used in many applications in healthcare domain such as treatment regime development [54–56], automated medical diagnosis [57, 58], resource scheduling and healthcare management systems [59, 60]. Among all of these applications, developing dynamic treatment regimes and sequential clinical decision-makings are becoming increasingly attractive for researchers [61]. Several DRL models have been developed in several studies to select the best treatment policy for some critical diseases such as cancer [53], sepsis [62, 63], diabetes [54, 55], and human immunodeficiency virus (HIV) [56] with goal of improving the long-term treatment outcome for the patients.

El Naqa et al. (2016) investigated the feasibility of RL for two stage adaptive radiation therapy using a simplified Q-learning algorithm with linear regression considering clinical covariant history as states and tumor control probability as the reward function. Their results demonstrated a promising feasibility of RL models in adaptive radiation therapy. However, more advanced nonlinear models are needed to be able to address biological aspects of multi-stage ART planning. Later, Jalalimanesh et. al. (2017) developed an agent-based model to simulate the tumor growth during radiation therapy and used a tabular Q-learning algorithm to find the optimal RT plan. Their results suggested that agent-based approach combined with RL is useful for simulating and optimizing Rt plans. However, they did not consider the uncertain biological response of the tumor and OAR cells in their model. Moreover, the tabular Q-learning cannot map high-dimensional state space due to the complexity. Also, finding the optimal action based on Bellman's equation is hard

when we have stochastic and non-linear dynamics in the decision-making environment [64]. Alternatively, using a neural network to map input states to (action, Q-value) pairs can help handling high-dimensional state space, uncertainty in tumor response dynamics and nonlinear rewards [49].

Tseng et al. [53] explored the feasibility of using deep reinforcement learning (DRL) based on historical treatment plans for automated knowledge-based ART for NSCLC patients. They developed a three-component neural network framework that includes a generative adversarial network (GAN) to learn patients' characteristics, a deep neural network (DNN) to estimate transition probabilities, and a deep Q-network (DQN) to find the optimal action. Their results showed that DRL can be used to achieve clinically acceptable results for knowledge-based ART while maximizing tumor local control. Their proposed approach is only useful in situations in which we have access to large-scale historical patient data and the certain value of tumor and OAR dosimetric and biological parameters. Collecting such a large-scale treatment plan data set for each cancer site needs a significant amount time and effort and is not accessible for everyone and all cancer cases. Moreover, the accuracy of the data set may not be guaranteed and can be prone to human errors. On the other hand, extracting the patients' characteristics from a large-scale data set is a time-consuming process and can be prone to overfitting. Therefore, in this dissertation, we introduce an automated decision-making framework by combining RL and optimization methods to find the optimal adaptation time for ART and dynamically adapt the plan to the tumor's uncertain biological response over time.

2.5 Radiation-Induced Lymphopenia

Significant lymphocyte count depletion (i.e., lymphopenia) is a common toxicity of radiation therapy and is associated with worse-off disease control in a number of solid tumors [65–68]. Because lymphocytes have a substantial role in the body's anticancer

immune response, severe lymphopenia can reduce patient survival even in the early stages of tumor progression [68–72].

Multiple recent studies have shown that severe lymphopenia is strongly associated with poor treatment outcomes in a number of solid tumors such as cervical [72, 73], pancreatic [74, 75], rectal [76], lung [69, 77], and esophageal [66, 78] cancers. Thus, preservation of the lymphocytes from radiation damage is crucial for the effectiveness of radiation therapy, and it is critical to understand the clinical and dosimetric factors affecting the severity and incidence of radiation-induced lymphopenia (RIL) and develop strategies for its mitigation.

Moreover, RIL risk likely varies by treatment modality. Recent studies have reported greater lymphocyte depletion in patients treated with photon therapy than with proton therapy [66, 68, 77–79]. For example, Shiraishi et al. [66] reported that proton beam therapy was associated with a lower risk of grade 4 lymphopenia compared with IMRT in esophageal cancer patients receiving neoadjuvant chemoradiotherapy. RIL commonly occurs in conventional photon radiation therapy, presumably due to the high radiosensitivity of lymphocytes and the large low and medium dose bath of photon therapy. Dose distribution patterns from protons and photons can differ greatly, and the dosimetric advantages of state-of-the-art proton therapy over photon therapy in terms of sparing organs at risk and normal tissue have been demonstrated in several clinical studies [80, 81]. Also, intensity-modulated proton therapy (IMPT) performs better than intensity-modulated radiation therapy (IMRT) in terms of dose sparing and robustness towards common anatomical changes in esophageal cancer patients [82]. Nevertheless, the true impact of dosimetric factors in determining RIL risk is still unclear and requires elaboration.

2.6 Prediction of Radiation-Induced Lymphopenia

Preservation of the lymphocytes from radiation damage is crucial for the effectiveness of RT [77, 78, 83]. Therefore, the ability to reliably predict RIL based on pretreatment factors (i.e., dosimetric factors, clinical, and patient-specific characteristics) would improve RT treatment planning.

Several studies have reported the strong associations between pretreatment factors including treatment-related characteristics (e.g., treatment modality, dose distribution patterns, fractionation regimens, etc.) [72, 77, 78, 84] and patient-specific factors (e.g., age, BMI, total blood volume, ALC baseline, etc.) [66, 71, 84–86]. Some of these studies have attempted to predict RIL based on different sets of pretreatment parameters.

Van Rossum et al. [85] showed the significance of age, planning target volume, body mass index, radiation modality, and baseline ALC in relation to grade 4 RIL for esophageal cancer patients and developed a pretreatment clinical nomogram based on these factors to determine the risk of grade 4 RIL for new patients. Later, Zhu et al. [86] developed a hybrid deep learning model to classify patients with grade 4 RIL based on patient characteristics and dosimetric features but they did not investigate the ALC kinetics during RT treatment for individual patients. However, comprehensive models that can forecast the kinetics of lymphocyte loss after fractionated radiation exposures in order to identify high-risk patients have not been explored yet. It is critical to fill this gap and provide a comprehensive prediction model that can forecast the ALC regressions during RT to develop RIL mitigation strategies at the right time and improve the effectiveness of RT for cancer patients.

2.7 Deep Learning Approaches for Treatment Outcome Prediction

The application of artificial intelligence and machine learning methods to extract insights from data is becoming increasingly attractive in many fields, including healthcare.

Although many healthcare applications have been developed, those that can predict disease progression [87], [88], treatment outcomes [89], or potential side effects [90], [91] play an important role in improving patient care.

Deep learning models have been developed to extract information from various kinds of data and for many tasks [92], [93]. Recurrent neural networks (RNNs) achieved significant results in extracting temporal information from sequential data such as text, audio, video, and time series [92]. The main advantage of RNNs is that they can maintain memory of recent events and update their current state based on both past states and current input data [94], [95]. Hochreiter and Schmidhuber [96] proposed the long short-term memory (LSTM) network as an improved variant of the RNN to handle the long-term dependency and vanishing gradient issues of RNNs. LSTM networks have been widely used for various kinds of tasks, including speech recognition [97], [98], image captioning [99], [100], trajectory prediction [101], [102], and text embedding [103], [104]. However, an LSTM network cannot be used alone for the current problem because the significant features that may predict RIL do not have uniform characteristics. A potential solution for this issue is to develop a stacked structure.

In a stacked structure, the algorithm nonlinearly integrates predictors in order to achieve higher prediction accuracy and reduce generalization error. Deep-stacked models can outperform state-of-the-art deep learning and machine learning models such as tree-based ensemble models and extreme gradient boosting algorithms [105], [106]. Therefore, we can take advantage of stacked deep structures to improve the deep learning model performance.

2.8 Predictive Uncertainty Quantification

Despite the success of standard deep learning methods in solving various healthcare problems, they may not be able to provide information about the reliability of their predictions [107]. Since decision-making in medical applications are mostly life-and-death decisions, quantifying reliability of predictions is crucial. However, most efforts in development of prediction models in this field have focused on improving the average accuracy of the algorithm, with little consideration for risk management.

The ability to identify patients who are at high risk of grade 4 RIL based on prediction models is very helpful to improve patient survival. However, poor performance from prediction models in clinical practice can have adverse consequences for patients. Therefore, it is important for clinicians to have some sense of when they can trust the prediction model results. Also, improving the management of risk and uncertainties in clinical decisions is reported as a potential opportunity to enhance the treatment outcome in a medical practice [108]. Nevertheless, model evaluation only on the basis of model performance measures (e.g., accuracy, mean squared error) cannot guarantee whether an individual prediction on a given patient should be trusted in clinical practice [109].

There are two main uncertainties in prediction results: (1) aleatoric uncertainty which is inherent noise and randomness in the real data due to the measurements and data collection errors. (2) epistemic uncertainty or model uncertainty due to inductive assumptions or an inadequate model, knowledge, and data [110, 111]. These two uncertainties need to be estimated sufficiently by uncertainty quantification models to account for prediction uncertainty. Consequently, in this study, we aimed to develop a hybrid deep learning structure that can estimate the uncertainty associated with each predicted ALC value for a given individual patient.

Several methods have been developed to quantify the uncertainty of linear/nonlinear

regression models and more complex prediction models such as neural networks in different applications by calculating prediction intervals [112–115] or determining a trust score [116]. A prediction interval can be computed for a neural network model with the assumption of normally distributed error for the neural network to account for aleatoric uncertainties [112, 114]. Neural network-based prediction intervals have been widely used in predicting health conditions and detecting diseases [116–119]. However, most of these methods require a large data set and long training. Moreover, Bayesian inference has shown good performance in quantifying uncertainty of predictions for traditional machine learning models such as random forest [120], support vector machines (SVM) [121], and variety of deep learning models, including neural network [122–124], long short-term memory (LSTM) network [125], convolutional neural network (CNN) [126–128], and recurrent neural network (RNN) [129]. However, these methods are computationally expensive and require significant modifications in the training process. Furthermore, deep Gaussian processes are effective alternative way to model the uncertainty of predictions as a non-parametric Bayesian approach by considering a Gaussian distribution over latent variables with respect to the input samples [130–132]. Unlike the Bayesian methods, this approach is easy to implement and achieves high quality predictive uncertainty estimates. In this study, we take advantage of this method in an ensembled neural network structure to estimate the uncertainty of ALC prediction.

Additionally, it is important to make the best use of prediction models and their estimated uncertainties to improve patient survival in clinical practice. However, there is no straightforward decision-making framework that can help physicians to easily use the results of complex deep learning models in their clinical decisions for each individual patient. Therefore, to fill this gap, we propose a hybrid decision-making framework to select patients for RT treatment using the predicted values of ALC and their associated risk to be used in clinical practice with the goal of avoiding grade 4 RIL for cancer patients. Such

a decision-making framework enables physicians to identify patients who are at high risk of grade 4 RIL and who may stand to benefit from treatment replanning, use of a different modality, or a pharmacological intervention and ultimately improve survival outcomes.

Chapter 3

A Reinforcement Learning Approach for Finding Optimal Policy of Adaptive Radiation Therapy Considering Uncertain Tumor Biological Response

3.1 Introduction

Cancer is one of the primary health problems in the world, and it is the second leading cause of death in the United States [133]. Radiation therapy (RT) is a common treatment modality for cancer patients. External beam radiotherapy is one of the most commonly used types of RT, in which ionizing radiation (e.g., photon, proton, etc.) goes through a particular part of patient's body to eradicate tumor cells. Unfortunately, it also damages healthy organs around the tumor called organs at risk (OARs). A high dose of radiation is required to control tumor cells from growing, while a maximal protection on the surrounding OARs must be ensured. The prescribed RT dose is usually delivered in multiple fractions to achieve tumor control while enabling OARs cells repair. In conventional RT, an equal amount of radiation dose is delivered to the patient in each fraction based on the computed

tomography (CT) images used to develop the treatment plan [134].

Recent studies showed that both tumor and OARs cells are changing dynamically during the treatment, and their biological responses to radiation also changes over time. Yet, the conventional plans mostly does not fully consider the dynamic nature of biological processes. Healthy cells repair, reoxygenation and repopulation of tumor cells and radiosensitivity are important biological factors in controlling efficiency of fractionated RT [135]. These factors play a significant role in the tumor and/or OAR response to RT treatment [2, 3]. Several studies have shown the possibility of achieving better treatment outcomes by modifying the amount of dose per fraction in fractionated RT based on the tumor's biological response. Information from biological images have been used in several studies to find biologically conformal nonuniform dose [23–25].

Most researchers have considered biological factors in modeling tumor response to radiation during the treatment. The linear-quadratic (LQ) model [6, 7] is one of the common radiation response models in fractionated RT. More comprehensive tumor response models have also been developed considering different biological factors, such as tumor repopulation [8–10]. Other factors in RT planning include redistribution, repair of sublethal damage, and reoxygenation [20, 21]. Furthermore, [22] developed a tumor response model to assimilate hypoxia and proliferation interplay by considering three cell compartments (i.e., proliferating, intermediate, and hypoxic) based on different levels of available oxygen and glucose. Similarly, OAR repair and radiosensitivity should also be considered to measure the OAR's biological response to radiation in the RT treatment planning [4].

The biological response to radiation varies from one patient to another [5]. Also, radiosensitivity and tumor proliferation are associated with the tumor cell cycle phase and gene level activities [31]. Therefore, personalized RT treatment planning has attracted researcher's attention in biological RT planning. Various models have been proposed to incorporate the biological response in RT treatment planning. Dynamic programming has

been used in several studies to account for dynamic tumor response during RT and finding tumor fractionation dose [26], OAR repair [27], tumor repopulation [10, 28] and tumor shrinkage [29]. However, a comprehensive model incorporating important biological factors in a dynamic framework has not been well studied.

The literature shows that the adaptive radiation therapy (ART) method improves treatment quality in terms of normal-tissue sparing and tumor cell reduction [33–36] as well as treatment cost and time [37, 38]. An ideal approach to consider dynamic tumor changes will be to take images of the patient at every visit, update tumor contours, and revise the treatment plan if a significant change was observed in the tumor geometry. However, daily imaging may not be useful in practice because any changes over the span of a day may not be significant enough to modify the existing treatment plan. More importantly, imaging the patient at every visit during the treatment period can be costly, time-consuming and prone to human errors. Therefore, a trade-off must be made considering tumor geometry change, costs, timing, and the recommended number of adaptive plans. This is the primary motivation for finding the optimal timing for adaptation to improve the clinical feasibility of ART.

Several approaches have been proposed to optimally determine the frequency of adaptation during the treatment based on the latest tumor geometry information, focusing on target-volume reduction [35, 39, 40] and the amount of dose per volume received in the tumor [19, 41–43]. Most studies suggest that the optimal time for adaption is when an adequate target volume reduction was observed and maintained [38–40]. However, there are some conflicting reports regarding the time of the largest tumor volume reduction occurred and the best time to adapt the plan to the tumor volume changes for different patients with different biological response characteristics. Therefore, a treatment should be customized for each patient.

Recent studies have proposed machine learning (ML) techniques to predict radiation

therapy outcomes [44, 45], identify patients who would benefit from ART [46], and determine the ideal time for adaptation in ART [42, 47]. ML techniques can help identify patients who will have high tumor volume reduction during RT and select them for ART by predicting the tumor regression during the course of treatment. Reinforcement learning (RL) is a machine learning algorithm that features modeling of sequential data based on the interactions between an agent and an environment. RL can be a good alternative to dynamic programming when there is a high level of uncertainty in a sequential decision-making problem because RL can generate a robust and risk-averse solution by incorporating the uncertainty in its environment. Deep reinforcement learning (DRL) algorithms have been applied to find the best policy (a sequence of decisions) in many diverse fields such as robotics [48], computer vision [49], energy [50, 51], and healthcare [52, 53].

Tseng et al. [53] explored the feasibility of using deep reinforcement learning (DRL) based on historical treatment plans for automated knowledge-based ART for NSCLC patients. They proposed a three component neural networks framework consisting of a generative adversarial network (GAN) to learn patients' characteristics, a deep neural network (DNN) to estimate transition probabilities, and a deep Q-network (DQN) to find the optimal action. The results of their study shows that DRL can be used to achieve clinically acceptable results for knowledge-based ART while maximizing tumor local control. Their proposed approach can be useful if one has access to large-scale historical patients' data and the certain value of tumor and OAR dosimetric and biological parameters are known. However, collecting such large-scale data for each cancer site needs significant time and effort, and it is not accessible for everyone and all cancer cases. Moreover, the accuracy of the data set may not be guaranteed and prone to human errors. On the other hand, extracting the patients' characteristics from a large-scale data set is a time consuming process and prone to overfitting [89].

Therefore, this chapter introduces a novel biological response model that incorporates

tumor cell death, repopulation, reoxygenation, radiosensitivity for tumor cells, and healthy tissues cell repair to predict radiation response of the tumor and OARs during the course of treatment. Using the model, an automated optimization framework is proposed by combining Reinforcement Learning (RL) and optimization method to find the optimal adaptation points for ART and dynamically adapt the plan to the tumor's uncertain biological response over time.

The contributions of this study are stated as (1) A biological-based treatment planning framework is proposed such that it not only controls the biological aspect of the treatment and incorporates the tumor biological response uncertainty, but also ensures the dose-volume requirements and clinical limits of the treatment without the need of dealing with complex optimization models; (2) The proposed reinforcement learning framework for ART planning can help the decision maker to achieve a robust solution under high levels of uncertainty in the biological parameters while reducing the variability in the solution and improving the control on the worst-cases which minimizes the undesirable effects of worst-cases on the treatment outcome; (3) Using the proposed comprehensive biological response model, the tumor volume regressions can be estimated without taking significant time and effort to collect large-scale data sets and avoid the need for expensive CT images. Also, an ART treatment plan can be determined in a shorter time compared to employing imaging information for the clinical implementation of ART considering the patient wait time and data collection time; (4) This approach enables the physicians to find an appropriate personalized ART policy in terms of fraction dose and timing of the adaptations using the volumetric and biological information to adapt the plan to the updated patient anatomy. It also can be used to identify patients who would benefit from ART as an alternative to the conventional equal-dose plan; and (5) The proposed approach is flexible enough to support a wide range of treatment objectives and preferences based on different decision makers for various cancer types.

The rest of this chapter is organized as follows. Section 2 explains how the temporal evolution of the tumor due to the radiation response is modeled. We then develop an RL framework for ART decision-making and discuss the associated mathematical formulations. Section 3 provides the sensitivity analysis of the model and the results from our experimental study using clinical lung cancer patient data. We conclude the chapter in Section 4.

3.2 Methodology

3.2.1 Problem Description

The goal of this study is to find the optimal policy for ART (i.e., the optimal timing of adaptation and the associated dose at each adaptation point) considering biological uncertainties that can improve the quality of treatment in terms of tumor control and OAR sparing. To accomplish this goal, first, we introduce a novel biological response model to estimate the tumor volume regressions with zero or minimal imaging during the treatment. Second, we propose an automated framework that combines RL and optimization methods. In this algorithm, the adaptation points are found based on an RL framework that encounters both tumor and OAR biological responses to radiation based on the response model considering the uncertainties in the values of biological parameters. We aim to achieve a maximum final tumor control while minimizing or maintaining the OARs toxicity levels by finding the actions to maximize the RL reward function. Third, to achieve the ultimate goal of RT treatment planning (i.e., maximizing the dose to the tumor while minimizing the dose to the OARs), beamlet intensities are optimized to satisfy dose-volume requirements and clinical limits for the patient based on the current predicted tumor volume and the proposed fraction dose from RL optimal policy at each adaptation point. As a result, we will find a robust optimal ART treatment plan that is biologically and clinically acceptable. The

list of sets, parameters, and variables used in the proposed tumor response model and RL algorithm are summarized in Table 3.1 and Table 3.2, respectively.

Table 3.1: Notations used in the dynamic biological response model and biological metrics

Notation	Description
Sets	
I	Set of treatment sessions (decision epochs)
T	Tumor structure
ϕ	OAR structure
Dynamic variables	
v_i	Tumor volume after delivering fraction i
u_i	Number of viable tumor cells after delivering fraction i
w_i	Number of dead tumor cells after delivering fraction i
m_i	Number of doomed tumor cells after delivering fraction i
d_i	Amount of dose in fraction i
Parameters	
N	Total number of treatment sessions
t_i	Time gap between fraction i and i-1
τ_g	The repopulation parameter
τ_d	Tumor decay parameter
τ_r^ϕ	OAR repair parameter
τ_g^ϕ	OAR repopulation parameter
OER	Reoxygenation parameter (Oxygen Enhancement Ratio)
α_p^T	Linear tumor radiosensitivity parameter in proliferating phase
α_h^T	Linear tumor radiosensitivity parameter in hypoxic phase
β_p^T	Quadratic tumor radiosensitivity parameter in proliferating phase
β_h^T	Quadratic tumor radiosensitivity parameter in hypoxic phase
α^ϕ	Linear radiosensitivity parameter of OAR
ρ	Ratio of dead cells at each stage
Biological metrics	
BED_i^T	Cumulative biological effective dose of the tumor after delivering fraction i
BED_i^ϕ	Cumulative biological effective dose of the OAR after delivering fraction i
SF_i	Total surviving fraction of the tumor after delivering fraction i

3.2.2 Dynamic Tumor Response Model

In this section, we introduce a model that incorporates the temporal evolution of the tumor due to radiation response. In this study, we assumed to have two cell compartments: (1) Proliferating, and (2) Hypoxic. This classification is consistent with the study by Jeong et al. [22] with one exception that we merged intermediate and hypoxic compartments

Table 3.2: RL algorithm notations

Notation	Description
Sets	
S	Set of possible states
A	Set of possible actions
RL components	
R	Immediate reward function
P	Transition probability
Q	Q-value function
Q^*	Optimal Q-value function
s_i	State of the system at step i
a_i	Decision set at step i
r_i	Immediate reward at step i
γ	discount factor for long-term reward
DDQN components	
w	DQN weights
w^-	Target network weights
L	Loss function
$N_{\mathcal{D}}$	Replay memory capacity
N_e	The total number of episodes
N_e	The total number of steps (treatment sessions)
ξ	Exploration decay rate

to one called hypoxic to be account for reoxygenation and hypoxia. The proliferating compartment contains cells that have adequate principle nutrients (i.e., glucose and oxygen) and they are in proliferating phase. Hypoxic compartment comprises cells without enough nutrients. Most of the cells in hypoxic component are starving and extremely hypoxic. Presumably, cells in hypoxic compartment cannot proliferate and the starving cells can die without RT exposure (necrotic cell death due to starving [136]). Only a fraction of cells in proliferating compartment are in cell-cycle and can proliferate. Therefore, we also incorporated three sub-compartments into each compartment to track different cell conditions (i.e., reoxygenation, cell-kill, cell decay caused by starving, cell cycle effect) during the course of treatment: (1) viable cells that are in the cell cycle, (2) doomed cells that are not in cell cycle, and (3) dead cells that are hypoxic and they are in decay process caused by cellular necrosis, apoptosis, metastasis, and cell migration. Figure 3.1(a) shows

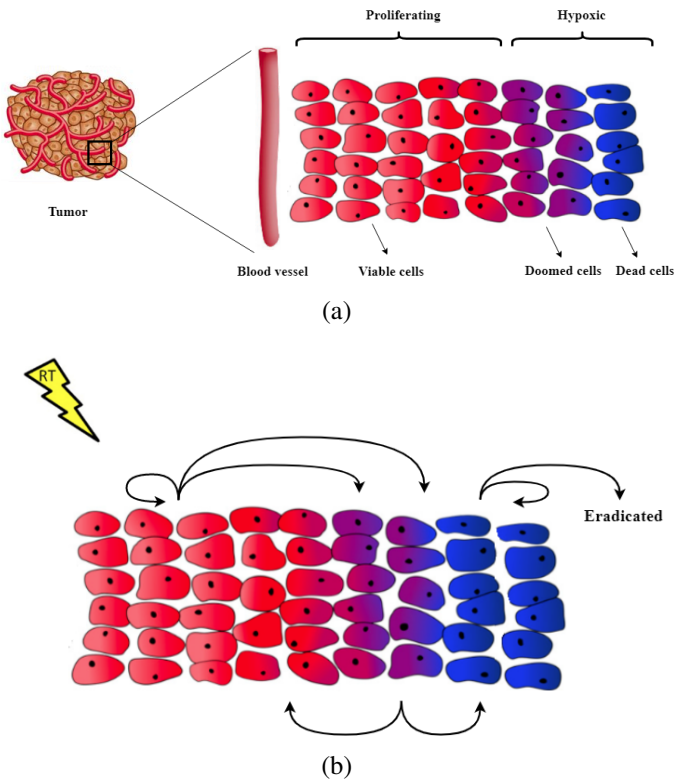


Figure 3.1: (a) Visualization of two tumor cells compartments and three sub-compartments (i.e., viable, doomed, and dead) (b) Redistribution of tumor cells sub-compartments after receiving radiation dose

the visualization of assumed tumor cell compartments and sub-compartments.

Doomed cells are the middle sub-compartment between metabolically active cells (intermediate) and hypoxic cells. Because of the increased mitotic cell death, proliferation of doomed cells do not have significant effect on number of cells in hypoxic compartment [22]. So, we assumed that doomed cells cannot proliferate as long as they get enough oxygen (i.e., move to viable sub-compartment). In contrast, doomed cells can move to dead sub-compartment if they receive enough radiation. At each stage of treatment, a proportion of viable cells can be moved into doomed and/or dead sub-compartment as a result of RT exposure. Figure 3.1(b) shows the redistribution of tumor cells sub-compartments after receiving the radiation dose.

Our proposed model considers proliferation and hypoxia as two important factors in

tumor radiation response and repopulation. The traditional L-Q model does not account for the necrotic cell death and exponential decay of the dead cells. Based on our model, only a fraction of lethally damaged cells are truly dead. The remaining cells are doomed, which means they are metabolically active without having enough oxygen to proliferate and are radioresistant [22]. As the doomed cells are hypoxic, higher doses of radiation are needed to kill these cells. This characteristic of hypoxic cells can be incorporated into the response model by adding the Oxygen Enhancement Ratio (OER). The OER indicates the required extra dose to achieve the same level of cell-kill/cell-survival of the hypoxic cell compared to the non-hypoxic (normoxic) cells [137]. The oxygen level can vary based on the distance from the tumor to blood vessels and blood vessel damage during radiation therapy [138]. We assume that surviving doomed cells can receive oxygen and return back to the proliferating phase as alive cells (reoxygenation of hypoxic cells). Necrotic cell loss is assumed to follow an exponential decay with parameter τ_d. Also, we consider an exponential tumor growth with parameter τ_g. Based on these assumptions, the number of cells in each sub-compartment at each time epoch can be calculated for viable (u_i), dead (w_i), and doomed (m_i) cells as

$$u_{i+1} = u_i \cdot exp(-\alpha_p^T d_i - \beta_p^T d_i^2) \cdot exp(\frac{t_i}{\tau_g}) + m_i \cdot exp(-\alpha_p^T \frac{d_i}{OER} - \beta_p^T \frac{d_i^2}{OER^2}), \tag{3.1}$$

$$\begin{aligned} w_{i+1} = \rho\, u_i \cdot (1 - exp(-\alpha_p^T d_i - \beta_p^T d_i^2)) \cdot exp(-\frac{t_i}{\tau_d}) \\ + m_i \cdot (1 - exp(-\alpha_p^T \frac{d_i}{OER} - \beta_p^T \frac{d_i^2}{OER^2})) \cdot exp(-\frac{t_i}{\tau_d}) + w_i \cdot exp(-\frac{t_i}{\tau_d}), \end{aligned} \tag{3.2}$$

$$\text{and} \quad m_{i+1} = (1-\rho)\, u_i \cdot (1 - exp(-\alpha_p^T d_i - \beta_p^T d_i^2)). \tag{3.3}$$

Where d_i is the dose for fraction i and t_i is the time gap between fraction $i-1$ and i. Tumor volume changes during RT can be estimated using cell compartment distribution at each time epoch and based on the initial total number of tumor cells. In addition to the

viable cells, doomed and dead cells (hypoxic cells) were also included in the total tumor volume calculation [29, 139]. Therefore, the tumor volume at each time epoch can be estimated as

$$v_{i+1} = u_{i+1} + w_{i+1} + m_{i+1}. \tag{3.4}$$

Moreover, we consider cell cycle radiosensitivity variations as described by Jeong et al. [139] for proliferating cells. So, the surviving fraction of proliferating (viable) cells can be calculated as

$$\begin{aligned} SF_i^p = {} & f_{G1}\, exp(-\alpha_{G1} d_i - \beta_{G1} d_i^2) + f_S\, exp(-\alpha_S d_i - \beta_S d_i^2) \\ & + f_{G2/M}\, exp(-\alpha_{G2/M} d_i - \beta_{G2/M} d_i^2) = exp(-\alpha_{eff} d_i - \beta_{eff} d_i^2), \end{aligned} \tag{3.5}$$

where α_{eff} and β_{eff} are effective α and β for proliferating phase, d_i is the fractional dose at stage i, and f_X, α_X and β_X are the fraction of cells, linear parameter, and quadratic parameter for a given cell cycle X (G1, S, or G2–M).

The increased radioresistance of hypoxic cells compared to proliferating cells can be quantified as $\alpha_h^T = \alpha_p^T / OER$ and $\beta_h^T = \beta_p^T / OER^2$ [137]. Hence, we propose the response models for the three cells types as

$$u_{i+1} = u_i\, SF_i^p\, .exp(\frac{t_i}{\tau_g}) + m_i\, exp(-\alpha_h^T d_i - \beta_h^T d_i^2), \tag{3.6}$$

$$\begin{aligned} w_{i+1} = {} & \rho\, u_i (1 - SF_i^p).\, exp(-\frac{t_i}{\tau_d}) \\ & + m_i\, (1 - exp(-\alpha_h^T d_i - \beta_h^T d_i^2)).exp(-\frac{t_i}{\tau_d}) + w_i\, exp(-\frac{t_i}{\tau_d}), \end{aligned} \tag{3.7}$$

and $$m_{i+1} = (1 - \rho)\, u_i (1 - SF_i^p). \tag{3.8}$$

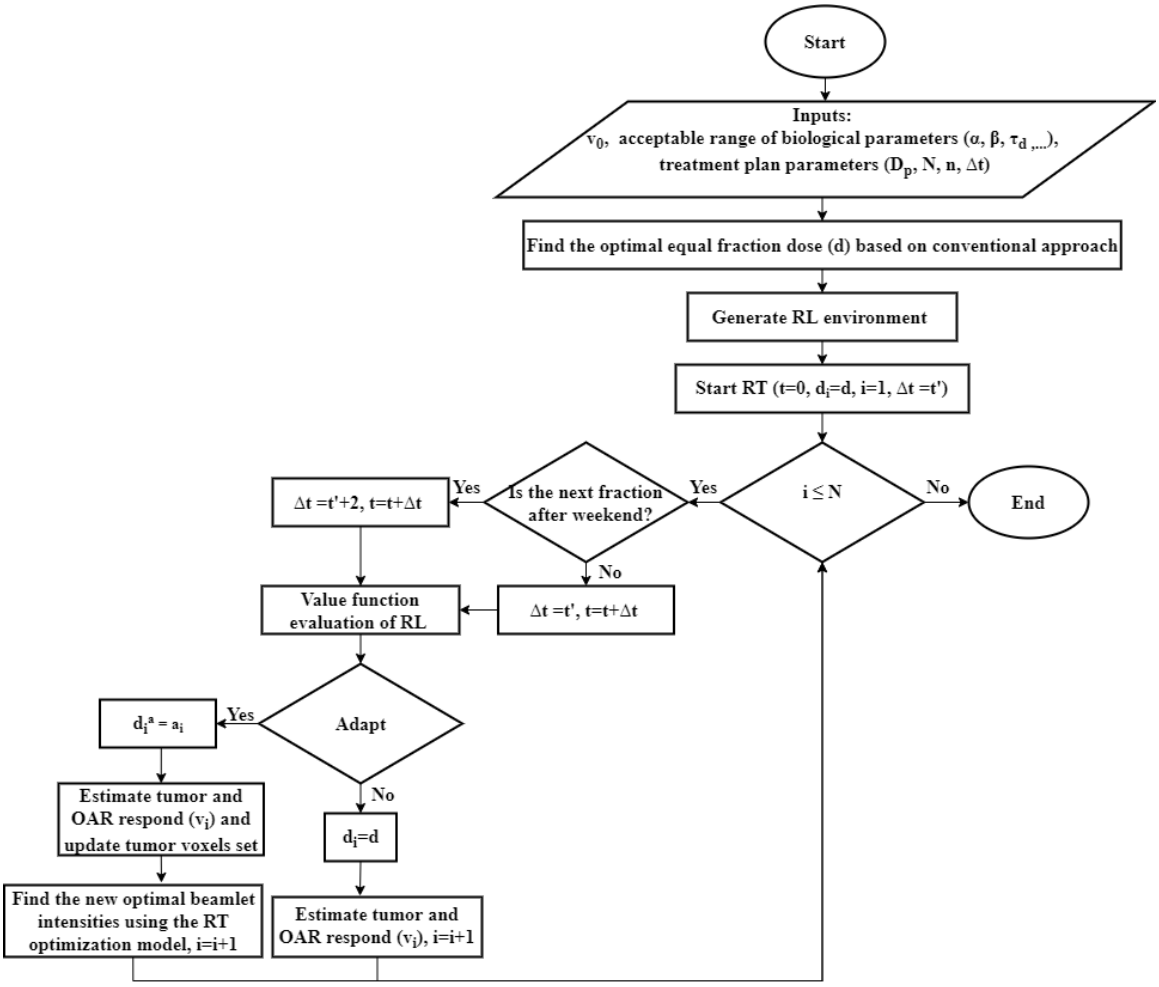


Figure 3.2: Flowchart of the proposed algorithm

3.2.3 Reinforcement Learning Framework for ART Problem

Our aim is to develop an automated framework that combines an RL and optimization methods, in which the adaptation points are found using the RL based on the biological effects and responses of the tumor and OAR to radiation using the proposed response model. Figure 3.2 shows the entire process of the proposed algorithm.

The goal of the agent in RL is to take actions that maximize the expected value of a predefined reward function. The RL environment can be described by the various states. The agent receives a reward (r_t) according to the selected decision being made under a specific

state (s_t), which leads to the next state (s_{t+1}). Using this feedback mechanism between the state and its corresponding reward, the agent can optimize its subsequent strategy for future actions. Figure 3.3 shows the RL procedure and its components.

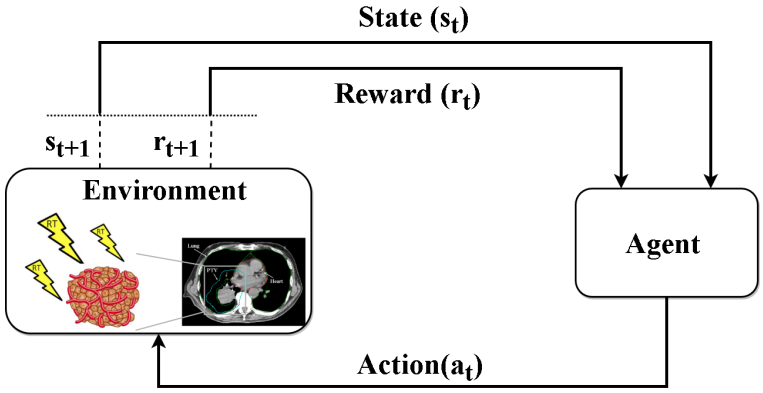


Figure 3.3: Reinforcement Learning Procedure

We propose a general RL framework based on a Markovian environment generated by the dynamic tumor response model to find the optimal policy of ART in which the environment can be customized based on the tumor biological parameters and patient information. Each time this framework is used for a specific cancer patient, the agent is learning so that its ability to find optimal actions gets improved for future patients with the same cancer type. Instead of taking images frequently to detect tumor volume changes during the treatment, our approach enables us to estimate the tumor volume regressions based on the proposed tumor response model used in the RL environment quantification. These estimates can be validated and/or corrected using a minimum number of imaging. Our approach is a continuous adaptation protocol on a weekly/daily basis, and we aim to find the optimal number of adaptations and the corresponding time as well as the radiation dose to be used until the next adaptation point.

3.2.3.1 Components of RL Framework For ART Problem

The environment in the proposed RL algorithm is a virtual environment to simulate ART treatment planning considering tumor volume changes using the proposed dynamic tumor response model. After the execution of the selected action, the agent obtains information on the next state and its corresponding reward value.

Algorithm 3.1 Action set construction method

Input: d, $\underline{d}$, $\bar{d}$ and Δ
if $0 < \bar{d} - d \leq \Delta \ \ and \ \ 0 < d - \underline{d} \leq \Delta$ **then**
$a \in \left\{\underline{d}, d, \bar{d}\right\}$;
else if $\bar{d} - d > \Delta \ \ and \ \ 0 \leq d - \underline{d} \leq \Delta$ **then**
$a \in \left\{\underline{d}, d, d + \frac{|\bar{d}-d|}{p}, ..., d + i(\frac{|\bar{d}-d|}{p})\right\}, \forall i = 1, 2, ..., p$ and $p = \left\lceil \frac{|\bar{d}-d|}{\Delta} \right\rceil$;
else if $d - \underline{d} > \Delta \ \ and \ \ 0 \leq \bar{d} - d \leq \Delta$ **then**
$a \in \left\{d - i(\frac{|\underline{d}-d|}{p}), ..., d - (\frac{|\underline{d}-d|}{p}), d, \bar{d}\right\}, \forall i = 1, 2, ..., p$ and $p = \left\lceil \frac{|\underline{d}-d|}{\Delta} \right\rceil$;
else if $\underline{d} - d > \Delta \ \ and \ \ \bar{d} - d > \Delta$ **then**
$a \in \left\{d - i(\frac{|\underline{d}-d|}{\underline{p}}), ..., d - (\frac{|\underline{d}-d|}{\underline{p}}), d, d + \frac{|\bar{d}-d|}{\bar{p}}, ..., d + j(\frac{|\bar{d}-d|}{\bar{p}})\right\}$,
$\forall i = 1, 2, ..., \underline{p}, \ \forall j = 1, 2, ..., \bar{p}, \ \underline{p} = \left\lceil \frac{|\underline{d}-d|}{\Delta} \right\rceil$ and $\bar{p} = \left\lceil \frac{|\bar{d}-d|}{\Delta} \right\rceil$;
else
$a \in \{d - \Delta, d, d + \Delta\}$.

At each time stage, a number of scenarios for tumor biological factors are generated. For each scenario, the tumor response to radiation and its immediate reward are calculated based on the current state and the action taken to determine the best action and next state. The set of actions includes possible decisions such as dose *increase ($+\Delta$), maintain, decrease ($-\Delta$)* as a result of the plan adaptation, where Δ is an amount of dose deviation from the conventional prescription dose d. For a given cancer type and its clinical protocols for treatment, some key input parameter values for reinforcement learning method can be given to the planner such as daily fractional dose d, Δ, dose lower-bound ($\underline{d}$), and dose upper-bound ($\bar{d}$). In some cases, the action sets may include more than the three actions, which could be multiples of Δ. Therefore, Algorithm 3.1 is developed to facilitate

constructing an appropriate action set for a given cancer case.

Our aim here is to find the optimal action at each time epoch (e.g., beginning of each week) to determine optimal adaptation points. We first show in Theorem 3.1 the existence of an optimal dose d_i^* of each fraction for a concave reward function on the closed interval $[\underline{d}, \bar{d}]$.

Theorem 3.1. For a concave reward function at each fraction, $R(d_i)$, there exists an optimal fractional dose $d_i^* = \operatorname{argmax} R(d_i)$ for $d_i \in [\underline{d},\ \bar{d}]$ as

$$\begin{cases} d_i^* = \underline{d}, & \text{if } \frac{\partial R(d_i^*)}{\partial d_i} < 0 \\ d_i^* \in [\underline{d},\ \bar{d}], & \text{if } \frac{\partial R(d_i^*)}{\partial d_i} = 0 \\ d_i^* = \bar{d}, & \text{if } \frac{\partial R(d_i^*)}{\partial d_i} > 0 \end{cases} \cdot$$

Proof. See Bolzano-Weierstrass Theorem [140]. □

Using Theorem 3.1, one can find the relation between the optimal fractional dose d_i^* and the equal fraction dose $d \in [\underline{d},\ \bar{d}]$ of the conventional reference plan. Hence, the corresponding action can be *increase, maintain, or decrease*, and the range of d_i^* can be determined by Corollary 3.1.1.

Corollary 3.1.1. The relation between the optimal fractional dose, d_i^*, to the conventional reference dose, $d \in [\underline{d},\ \bar{d}]$, can be determined by the gradient, $\frac{\partial R(d_i)}{\partial d_i}$, of the reward function at $d_i = d$ as

$$\begin{cases} d_i^* \in [\underline{d},\ d), & \text{if } \frac{\partial R(d_i)}{\partial d_i}|_{d_i=d} < 0 \\ d_i^* = d, & \text{if } \frac{\partial R(d_i)}{\partial d_i}|_{d_i=d} = 0 \\ d_i^* \in (d,\ \bar{d}], & \text{if } \frac{\partial R(d_i)}{\partial d_i}|_{d_i=d} > 0 \end{cases} \cdot$$

Proof. The proof is trivial following Theorem 3.1. □

Few performance measures are commonly used to evaluate an RT treatment plan including BED, tumor control probability (TCP) and normal tissue control probability (NTCP). BED is a measure to estimate the amount of radiation damage received in any structure. A higher tumor BED is known to give a better tumor control. In contrast, a lower OAR BED is desirable to have lower OAR toxicity. Therefore, we developed a multi-stage optimization model, in which the biological response of the treatment is defined based on the BED of the tumor and OAR. At each adaptation point determined by the RL (i.e., at stage k), an optimization problem is used to find the optimal fraction dose of the stage (d_k) as follows

$$\text{max} \quad \sum_{i=1}^{k} BED_i^T(d_i), \tag{3.9}$$

s.t.

$$\sum_{i=1}^{k} BED_i^{\phi}(\gamma\, d_i) \leq BED_k^{\phi_{Ref}}(\frac{k}{N}\, \gamma\, D_{pres}), \tag{3.10}$$

$$\sum_{i=1}^{k} BED_i^{T}(d_i) \geq BED_k^{T_{Ref}}((1-\frac{k}{N})\, D_{pres}^{L}), \tag{3.11}$$

$$SF_k\, SF_{N-k}^{Ref} \leq \varepsilon, \tag{3.12}$$

$$SF_{N-k}^{Ref} = SF((1-\frac{k}{N})\, D_{pres}), \tag{3.13}$$

$$SF_k = SF(d_{k-1}) \cdot SF(d_k), \tag{3.14}$$

$$\text{and} \quad d_l \leq d_k \leq d_u. \tag{3.15}$$

The objective function (3.9) maximizes the total BED on the tumor (T) by delivering $d_1, d_2, ...,$ and d_k at stages $1, 2, ..., k$. Constraint (3.10) controls the BED deviations from the conventional reference plan (i.e., an equal fraction dose $d_i = d, \forall i$) for the OAR biological tolerance to achieve the same or better OAR (ϕ) toxicity. Constraint (3.11) sets a lower-bound for tumor BED based on a set of biological parameters and delivered dose using the lower bound of the prescription dose to be account for the required clinical tumor BED.

Constraint (3.12) is to ensure that tumor cells will be completely eradicated at the end of treatment even if we continue the rest of treatment with the conventional plan (i.e., $d_i = d, i = N-K, \cdots, N$). Where, SF_k is the total surviving fraction at the end of stage k by delivering d_1, d_2, ..., and d_k at stages $1, 2$, ..., k, and SF_{N-k}^{Ref} is the the total surviving fraction after $N-k$ fractions based on the reference plan (labeled as Ref). Constraints (3.13) and (3.14) represent the calculation of SF_{N-k}^{Ref} and SF_k, respectively. Finally, constraint (3.15) ensures that the amount of the fraction dose is within its lower and upper bounds.

We can formulate a reward function based on the proposed biological optimization model by relaxing constraints and penalizing weighted constraint violations. Our proposed reward function of RL is defined as follows

$$\begin{aligned} R(s_i, a_i) = \lambda_1 \, BED_i^T(s_i, a_i) - \lambda_2 \, \left(BED_i^{T_{Ref}}(s_i) - BED_i^T(s_i, a_i) \right)^+ \\ - \lambda_3 \, \left(SF_i(s_i, a_i) \, SF_{N-i}^{Ref}(s_i, a_i) - \varepsilon \right)^+ - \lambda_4 \, \left(BED_i^{\phi}(s_i, a_i) - BED_i^{\phi_{Ref}}(s_i) \right)^+, \end{aligned} \tag{3.16}$$

where $BED_i^T(s_i, a_i)$ and $BED_i^{\phi}(s_i, a_i)$ are the cumulative BED after taking action $a_i \in \{A\}$ (delivering d_k dose) in state $s_i \in \{S\}$ at each time stage i and $(.)^+$ is a sign function defined as

$$(x)^+ = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases}. \tag{3.17}$$

The BED estimate of the tumor is calculated considering the surviving fraction of viable and doomed cells, tumor repopulation of viable cells, and decay of dead cells. Hence, the

proposed formulation for the BED of the tumor at each time stage is

$$BED_i^T(d_i) = -\left(\frac{u_{i-1}}{v_{i-1}}\right)\frac{\Delta t_i}{\alpha\ \tau_g} + \left(\frac{w_{i-1}}{v_{i-1}}\right)\frac{\Delta t_i}{\alpha\ \tau_d} + \left(\frac{u_{i-1}}{v_{i-1}}\right)\left(d_i + \frac{d_i^2}{\alpha/\beta}\right) + \left(\frac{m_{i-1}}{v_{i-1}}\right)\left(\frac{d_i}{OER} + \frac{d_i^2}{OER^2\ \alpha/\beta}\right). \tag{3.18}$$

Theorem 3.2. There exists a lower bound (d_l) on d_i such that $BED_i^T(d_i)$ is a non-negative and monotonically increasing function for all $d_i \geq d_l, \forall i = 1, \cdots, N$.

- If $\frac{u_{i-1}}{w_{i-1}} \leq \frac{\tau_g}{\tau_d}$, then $d_l = 0\ and\ BED_i^T(d_i) \geq 0$.
- If $\frac{u_{i-1}}{w_{i-1}} > \frac{\tau_g}{\tau_d}$, then $d_l = \frac{-b+\sqrt{\delta}}{2a} > 0, and\ BED_i^T(d_i) > 0$

  where,

  $a = \frac{OER^2\ u_{i-1}+m_{i-1}}{OER^2\ \alpha/\beta\ v_{i-1}}$,

  $b = \frac{OER\ u_{i-1}+m_{i-1}}{OER\ v_{i-1}}$,

  $c = -\left(\frac{u_{i-1}}{v_{i-1}}\right)\frac{\Delta t_i}{\alpha\ \tau_g} + \left(\frac{w_{i-1}}{v_{i-1}}\right)\frac{\Delta t_i}{\alpha\ \tau_d}$,

  $\delta = b^2 - 4ac$.

Proof. The BED function in (3.18) is a quadratic function of d_i. It is continuous and twice differentiable. The first derivative of $BEDi^T(d_i)$ with respect to d_i is strictly positive for all $d_i \geq 0$ as

$$\frac{\partial BEDi^T(d_i)}{\partial d_i} = \frac{OER\ u_{i-1} + m_{i-1}}{OER\ v_{i-1}} + 2\left(\frac{OER^2\ u_{i-1} + m_{i-1}}{OER^2\ \alpha/\beta\ v_{i-1}}\right)d_i > 0. \tag{3.19}$$

Also, the second derivative of $BEDi^T(d_i)$ with respect to d_i is strictly positive for all d_i as

$$\frac{\partial^2 BEDi^T(d_i)}{\partial d_i^2} = 2\left(\frac{OER^2\ u_{i-1} + m_{i-1}}{OER^2\ \alpha/\beta\ v_{i-1}}\right) > 0, \tag{3.20}$$

hence, $BEDi^T(d_i)$ is strictly convex and it is an increasing function for $d_i \geq 0$.

Next, we show the existence of a lower bound (d_l) on dose d_i. For the notational

convenience, we use the following abstract form of $BEDi^T(d_i)$ function, $ad_i^2 + bd_i + c$, where a, b, and c are defined as

$$a = \frac{OER^2\ u_{i-1} + m_{i-1}}{OER^2\ \alpha/\beta\ v_{i-1}}, \tag{3.21}$$

$$b = \frac{OER\ u_{i-1} + m_{i-1}}{OER\ v_{i-1}}, \tag{3.22}$$

and

$$c = -\left(\frac{u_{i-1}}{v_{i-1}}\right)\frac{\Delta t_i}{\alpha\ \tau_g} + \left(\frac{w_{i-1}}{v_{i-1}}\right)\frac{\Delta t_i}{\alpha\ \tau_d}. \tag{3.23}$$

The lower bound d_l can be found by examining the roots of $ad_i^2 + bd_i + c = 0$. The roots of a quadratic function are given by $d_i = \frac{-b \pm \sqrt{\delta}}{2a}$, where $\delta = b^2 - 4ac$. If $\delta \leq 0$, then the $BED_i^T(d_i)$ is non-negative for any $d_i \geq 0$. If $\delta > 0$, then there are two roots for $BED_i^T(d_i) = 0$. Since the first derivative of the function is equal to zero ($\frac{\partial BED_i^T(d_i)}{\partial d_i} = 0$) for a negative d_i, at least one of the roots must be negative. For a positive δ, there are two possibilities: (1) no positive root if $BED_i^T(0) \geq 0$ or (2) a positive root if $BED_i^T(0) < 0$. In the former case, the c value is positive and we have $\frac{u_{i-1}}{w_{i-1}} \leq \frac{\tau_g}{\tau_d}$. If we consider two negative roots as d_1 and d_2 where $d_1 > d_2$, the $BED_i^T(d_i)$ is non-negative for any $d_i \geq d_1$; since $d_1 < 0$ we conclude that $BED_i^T(d_i)$ is non-negative for any $d_i \geq 0$. In case (2), the c value is negative and we have $\frac{u_{i-1}}{w_{i-1}} > \frac{\tau_g}{\tau_d}$. For the positive root as in d_1, $BED_i^T(d_i)$ is non-negative for any $d_i \geq d_1$. Therefore, the lower bound is $d_l = \frac{-b+\sqrt{\delta}}{2a}$ for this case. □

We consider OAR repair and repopulation, which are major biological factors affecting the response to radiation on healthy tissues. The following equation is used to capture the OAR's biological response to radiation during the RT treatment:

$$v_{i+1}^{\phi} = v_i^{\phi}\ exp(-\alpha^{\phi} d_i^{\phi} - \beta^{\phi} {d_i^{\phi}}^2)\ exp(\frac{t_i}{\tau_g^{\phi}})\ exp(\frac{t_i}{\tau_r^{\phi}}), \tag{3.24}$$

using the general BED formulation, the BED of an OAR can be calculated as

$$BED_i^{\phi} = \left(d_i^{\phi} + \frac{{d_i^{\phi}}^2}{\alpha^{\phi}/\beta^{\phi}} \right) - \frac{\Delta t_i}{\tau_g^{\phi}\alpha^{\phi}} - \frac{\Delta t_i}{\tau_r^{\phi}\alpha^{\phi}}. \tag{3.25}$$

We assume that the OAR receives a heterogeneous dose with a sparing factor of γ, which indicates the ratio of the average dose received by the OAR to the average dose received by the tumor ($d_k^{\phi} = \gamma\ d_k$). We can also consider a variable γ at each stage of the treatment as γ_i. Unkelbach et al. [29] proposed that γ_i can be calculated as $\gamma_i = \gamma.v_i^{2/3}$ in which the sparing factor (γ_i) decreases as the area of radiation field needed to treat tumor decreases. As it is explained in section 3.2.2, we can estimate the tumor volume based on the number of tumor cells. The tumor volume at the beginning of stage i can be calculated by multiplying the number of tumor cells at the beginning of this stage (v_{i-1}) by the volume of each tumor cell (V_{cell}). If we assume to have spherical tumor volume, then we have

$$v_{i-1} \times V_{cell} = \frac{4}{3}\pi r^3 \tag{3.26}$$

$$\text{and} \quad r = \left(\frac{3}{4\pi}\ v_{i-1} \times V_{cell} \right)^{1/3}, \tag{3.27}$$

where r is radius of the estimated disk. So, the area of the area of radiation field needed to treat tumor can be estimated as

$$\pi r^2 = \left(\frac{3\sqrt{\pi}}{4}\ v_{i-1} \times V_{cell} \right)^{2/3}, \tag{3.28}$$

as a result, the sparing factor (γ_i) can be found as

$$\gamma_i = \gamma \left(\frac{3\sqrt{\pi}}{4}\ v_{i-1} \times V_{cell} \right)^{2/3}. \tag{3.29}$$

3.2.3.2 Deep Double Q-learning Network for Learning Process

The training data is a tuple of $\{S, A, R\}$ in a finite horizon (treatment duration) and the goal is to develop an optimal policy (sequence of decision rules) for ART to maximize the long-term reward which is defined based on RT performance metrics (e.g., BED, SF) for the treatment outcome. Therefore, the effect of the actions is evaluated not only based on the immediate reward but also the long-term or subsequent rewards. The value function $V(s)$ presents the value of a state which is defined as the total expected reward starting from the state expressed as *Q-value function* $Q(s, a)$ which can be calculated as follows

$$Q^{\pi}(s,a) = E\left[R_t + \gamma R_{t+1} + \gamma^2 R_{t+2} + ...|s,a\right] = E_{s'}\left[R_t + \gamma Q^{\pi}(s',a')|s,a\right]. \tag{3.30}$$

A Q-value function can be used to find the optimal value function $Q^*(s, a)$ as

$$Q^*(s,a) = E_{s'}\left[R_t + \gamma\, max\; Q^*(s',a')|s,a\right]. \tag{3.31}$$

Q-learning is a common approach to find the optimal Q-values in RL. Recent studies by Google DeepMind have shown that the Q-function can be evaluated efficiently using deep Q-network (DQN) that provides a stable solution to deep value-based RL [49, 141]. Several studies showed that DQN algorithm achieved better computational performance than Q-learning algorithm in complex RL environments [64, 142]. The time complexity is sublinear in the length of state period (i.e., the number of steps per episode multiplied by the total number of episodes), and the space complexity is sublinear in the number of state space, action space, and steps per episode (Leem et al., 2020, Liu et al., 2021). Since the same weights are using for estimating both target and Q-values in the DQN algorithm, both Q-values and target values are shifting and there is a big correlation between the target network and the output weights that are changing for training. So, we consider the idea

of fixed Q-targets introduced by DeepMind and employ a separate network with a fixed parameter (w^-) for estimating the target values, and update the target network at every τ steps based on the current DQN parameters. Therefore, we will have more stable learning process. An over-estimation of Q-values at the early stages of the training can be an issue in Q-learning. Hence, we use the Double DQN structure ([143]) to handle the problem. We use two networks to separate the action selection from the target Q value generation during the learning process to reduce a false positive error that resulted from a noisy Q-value. Therefore, the DQN network selects the best action first to take for the next state (i.e, the action with the highest Q value). Then, the target network calculates the target Q value according to the action taken at the next state. This results in a faster training and more stable learning process. The value function can be represented using a DQN with weights w as $Q(s, a, w)$ $Q^{\pi}(s, a)$ and the loss function is

$$L(w) = E\left[(R + \gamma\ max\ Q(s', a', w^-) - Q(s, a, w))^2\right]. \tag{3.32}$$

The entire learning process is summarized in Algorithm 3.2.

Algorithm 3.2 Reinforcement Learning Process

Initialization: Initialize replay memory buffer $\mathcal{D}$ to capacity $N_{\mathcal{D}}$, initialize network weights (w_1), initialize target network ($w_1^- \leftarrow w_1$), initialize the environment, initialize decay rate ξ, episode counter $e = 0$;
for $e \leq N_e$ **do**
  Reset the environment, $t = 0$, observe the first state
  for step $t \leq N$ **do**
    Increase decay rate ($\xi \leftarrow \xi + \Delta_\xi$)
    Use Epsilon Greedy Strategy, with probability ϵ select a random action a_t
    Otherwise select action $a_t = \text{argmax}_a \, Q(s_t, a; w)$
    Execute action a_t, calculate reward r_t and observe next state s_{t+1}
    Store the transition (s_t, a_t, r_t, s_{t+1}) in the replay memory $\mathcal{D}$
    Sample a random minibatch of the transitions (sj, a_j, r_j, s_{j+1}) from $\mathcal{D}$
    if If the episode ends at next state $(j + 1)$ **then**
      Set target $\hat{Q} = r_j$
    else
      Set $\hat{Q} = r_j + \gamma \, Q(s_{j+1}, \text{argmax}_{a'} \, Q(s_{j+1}, a'; w), w^-)$
    Perform a gradient descent step with loss $\left(\hat{Q} - Q(s_j, a_j; w)\right)^2$
    Every τ steps ($(e.N) + t > \tau$) reset the target network weights ($w_t^- \leftarrow w_t$)

3.2.4 RT Optimization Model

Using the proposed RL approach, we can find the optimal adaptation points to improve the biological response of the tumor and OARs by maximizing the reward function. The agent's action is based on the biological-based reward function, but the treatment plan also needs to meet physical dose requirements for the clinical purpose. Since it is far too complicated to consider all aspects of an RT treatment (i.e., biological and physical) in one reward function, an optimization model is proposed to control the dose-volume clinical requirements. Once the adaptation points are determined using the RL approach, a beamlet optimization model is solved to satisfy the dose-volume constraints.

This optimization model will be adapted at each adaptation point based on the corresponding predicted tumor volume. For this purpose, the tumor response model can be used to estimate tumor volume changes. The tumor volume change ratio at each stage $r_k^T(d_k)$ is a function of a delivered dose at the stage (d_k) and can be calculated as $r_k^T(d_k) = \frac{v_k(d_k)}{v_{k-1}(d_{k-1})}$.

Therefore, at each adaptation point determined by RL (i.e., at stage $i = k$), an optimization model is solved to find the optimal beamlet intensities as follows

$$\min \quad \sum_{s\in\{T\cup S\}} \frac{C_s}{|V_s^k|} \sum_{v\in V_s^k} D_v^k, \tag{3.33}$$

s.t.

$$D_v^k \leq U_v^k, \qquad \forall\, v \in V_s,\ s \in \{T \cup S\}\,, \tag{3.34}$$

$$D_v^k \geq L_v^k, \qquad \forall\, v \in V_s,\ s \in \{T\}\,, \tag{3.35}$$

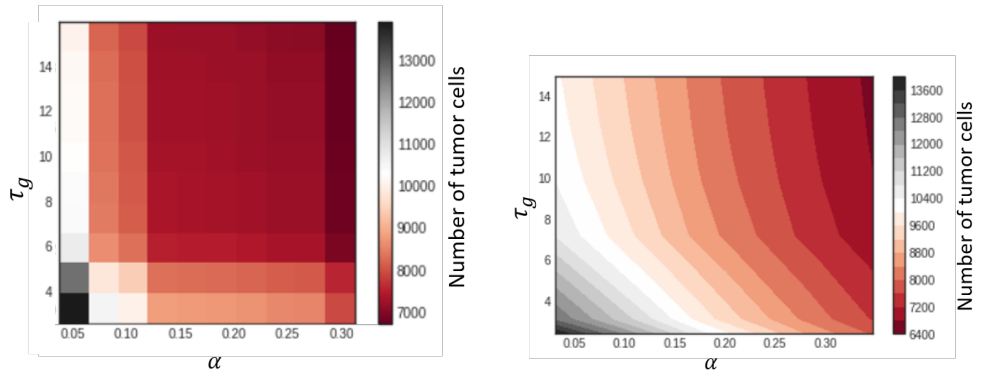
$$D_v^k = \sum_{b\in B} \Delta_{v,k,b}\, w_b, \qquad \forall\, v \in V_s,\ s \in \{T \cup S\}\,, \tag{3.36}$$

and $\quad w_b \geq 0, \qquad \forall\, b \in B. \tag{3.37}$

3.3 Numerical Experiments and Results

3.3.1 Experiment Setup

We evaluate the proposed tumor response model using simulation and compare the result with the conventional LQ response model. Then, a sensitivity analysis is performed to explore the effect of variability in the corresponding parameters and evaluate the observations based on clinical practices. Since tumor growth and radiosensitivity are two main biological factors in determining tumor radiation response, we consider four types of tumors based on the range of radiosensitivity parameter (α) and tumor growth factor (τ_g). Figure 3.4 shows tumor volume after delivering 2 Gy radiation dose to a tumor with the initial volume of 10,000 voxels considering $\alpha \in [0.03, 0.0365]\ Gy^{-1}$ and $\tau_g \in [0.5, 15]\ days^{-1}$. We categorize the tumor response types into four groups and their corresponding ranges of model parameters as summarized in Table 3.3 [9, 21, 144, 145]. Case I and Case II refer to the fast-growing tumors with low levels of radio sensitivity, and high radiosensitivity



(a) (b)

Figure 3.4: Number of tumor cells after delivering 2 Gy radiation to a tumor with the initial volume of 10000 voxels using $\alpha \in [0.03, 0.0365](Gy^{-1})$ and $\tau_g \in [0.5, 15](days^{-1})$

(i.e., Early responding tumors), respectively. Case III and Case IV are for slow-growing tumors with low radiosensitivity (i.e., Late responding tumors) and high radiosensitivity (i.e., Intermediate to late responding tumors), respectively.

Table 3.3: Tumor volume cases used in the sensitivity analysis of tumor response model

	Low radiosensitivity ($\alpha \in [0.03, 0.25]\ Gy^{-1}$)	**High radiosensitivity** ($\alpha \in [0.25, 0.365]\ Gy^{-1}$)
Fast growing tumor ($\tau_g \in [0.5, 7]\ days^{-1}$)	Case I	Case II
Slow growing tumor ($\tau_g \in [7, 60]\ days^{-1}$)	Case III	Case IV

3.3.2 Sensitivity Analysis of Tumor Response Model

3.3.2.1 Tumor Response during The Course of RT Treatment

To see the effect of radiation on tumor volume changes, we have simulated the tumor volume and tumor cells distribution in each sub-compartment. We used an specific set of parameters as $T_d = 60\ (days),\ \tau_d = 3\ (days^{-1}),\ \rho = 0.8,\ \alpha = 0.282\ (Gy^{-1}),\ \alpha/\beta =$

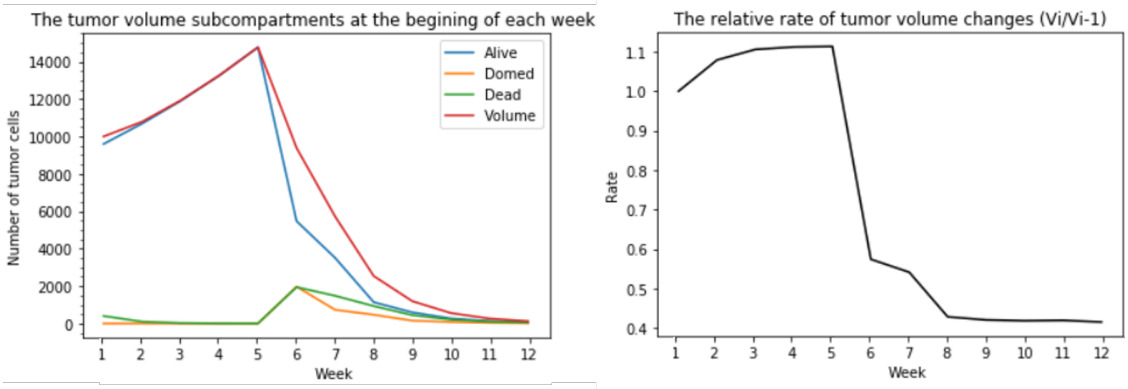


Figure 3.5: Illustration of dose response curves for the tumor volume, three sub-compartments, and plot of the relative tumor volume changes rate (v_i/v_{i-1})

$10(Gy),\ OER = 1.8$. We also assumed that the total number of tumor cells at the beginning of the time frame is 10^4 cells. First, we simulate the tumor volume before RT for a period of 4 weeks to see the effect of tumor growth. Then, we assumed that the treatment was started at the beginning of week 5.

Dose response curves for the whole tumor volume and three tumor cell sub-compartments and relative tumor volume changes rate (v_i/v_{i-1}) are shown in Figure 3.5. As it is shown in the figure, the model estimates that the number of alive tumor cells and tumor volume is increasing before RT due to tumor growth and it is decreasing after starting RT treatment.The number of doomed and dead tumor cells increase after at the early stages of RT treatment then it starts to decrease due to the decay of dead cells and mitotic cell death. The relative rate of tumor volume changes shows the tumor volume changes between two consecutive weeks and it is also increasing before RT and starts to decrease after beginning of RT treatment to reach a constant value.

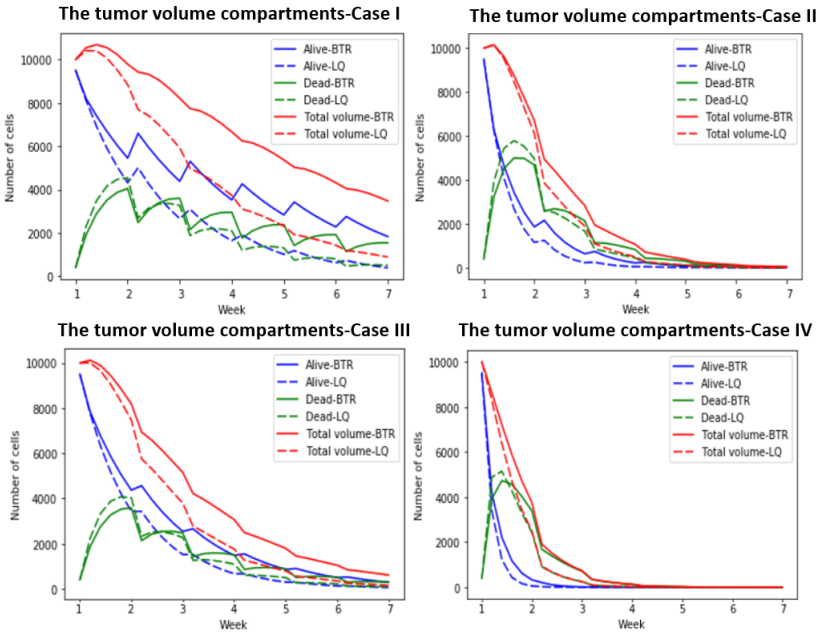
3.3.2.2 The Effect of Biological Parameters

We study the effect of radiosensitivity (α) and the tumor growth factor (τ_g) on tumor volume response considering four tumor cases (see Table 3.3). The following parameters were selected based on the ranges in Table 3.3: $\tau_d = 4\ (days^{-1}),\ \rho = 0.8,\ \alpha/\beta =$

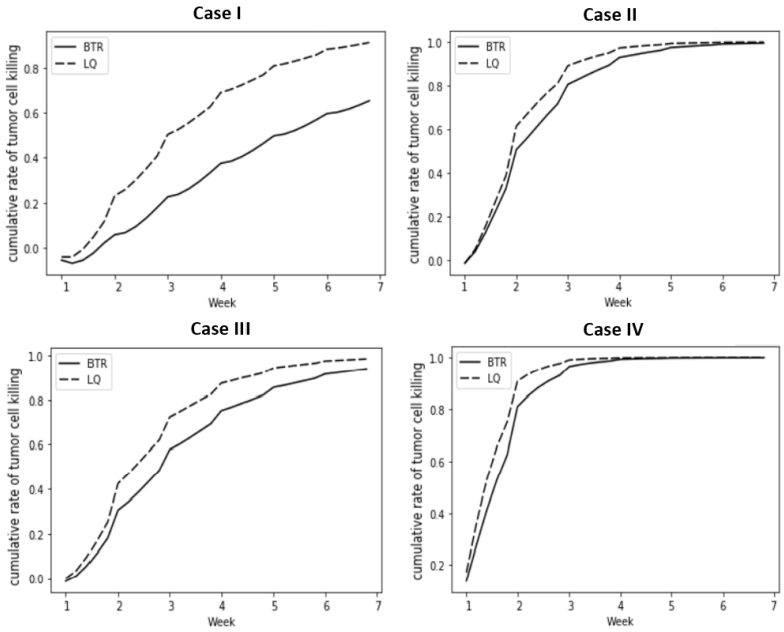
$10\ (Gy)$, $OER = 1.8$ for all cases, $\alpha = 0.135\ (Gy^{-1})$ and $\tau_g = 5\ (days^{-1})$ for Case I, $\alpha = 0.282\ (Gy^{-1})$ and $\tau_g = 5\ (days^{-1})$ for Case II, $\alpha = 0.135\ (Gy^{-1})$ and $\tau_g = 15\ (days^{-1})$ for Case III, $\alpha = 0.282\ (Gy^{-1})$ and $\tau_g = 15\ (days^{-1})$ for Case IV. Experimental results of the proposed tumor response model will be compared with those of the conventional LQ response model. As it is common in practice, we assumed that the treatment plan is to deliver five equal fractional doses of 2 Gy in 6 weeks and no treatment will be given on the weekends. For each case, we simulated the weekly tumor response based on the LQ model and the proposed response model. Figure 3.6 shows the tumor response curves and cumulative tumor cell-kill rate compared with the conventional LQ response model for each tumor case. Figure 3.7 presents the tumor volume changes for each tumor case based on LQ tumor response model (Figure 3.7(a)) and biological tumor response (BTR) model (Figure 3.7(b)).

As shown in Figure 3.6, both models behaved similarly for Case IV which is assumed to have the highest α and τ_g values. The difference between the two models is more noticeable when the tumor is less radio-sensitive (e.g., Case II and Case IV), and the impact of tumor growth and reoxygenation (which is not considered in the LQ model) became more apparent. The results indicate that our tumor response model is more sensitive to tumor reoccurrence risk and shows more realistic results than the LQ model for less radio-sensitive tumors (e.g., Case I and Case III).

The best treatment outcome (i.e., no more remaining tumor cells after the fourth week of treatment) was observed in Case IV, which is the most radio-sensitive case with slow proliferation. Increasing the proliferation rate (Case II) resulted in a slightly worse treatment outcome by taking six weeks to eradicate the whole tumor. However, we observed that a lower tumor radio-sensitivity lead to a less desirable treatment outcome for both slow proliferating (Case III) and fast proliferating (Case I) tumors. As expected, the worst treatment outcome was observed in Case I, which has the lowest radio-sensitivity with fast



(a)

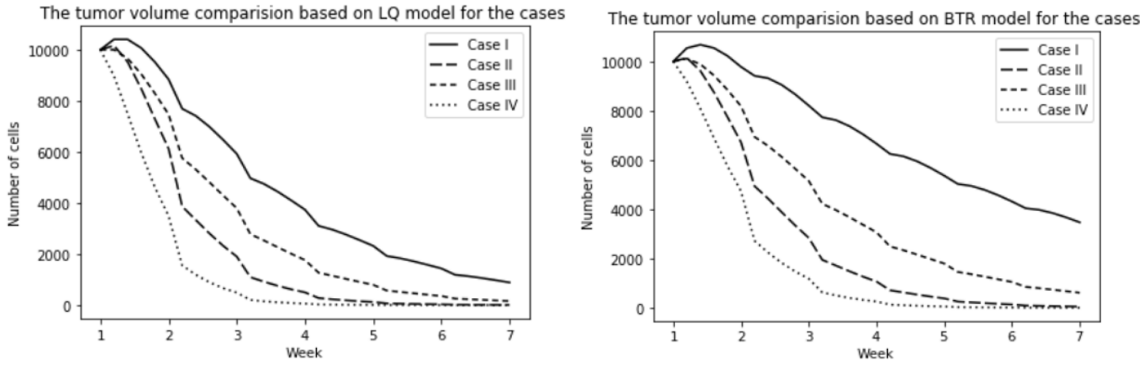


(b)

Figure 3.6: Radiosensitivity effect based on tumor response model versus LQ model; (a) tumor response curves, (b) cumulative tumor cell-kill rate curves

tumor cell proliferation.

We consider exponential tumor growth with constant rate of $1/\tau_g$ which can be found



(a) (b)

Figure 3.7: Illustration of tumor volume changes for each tumor case based on (a) LQ tumor response model and (b) biological tumor response (BTR) model

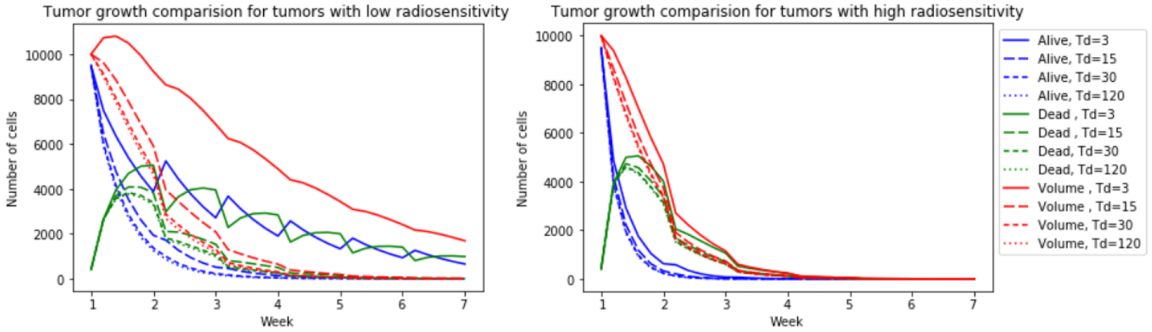


Figure 3.8: Tumor growth comparison

based on the tumor doubling time ($\tau_g = T_d/Ln(2)$). Higher value of the tumor growth rate (τ_g) would lead to higher tumor doubling time and slower tumor growth. We considered one tumor case with high radiosensitivity ($\alpha = 0.282$) and another one with low radiosensitivity ($\alpha = 0.135$) and simulate the tumor response using four different T_d of 3, 15, 30, and 120 days with the same treatment plan (5 equal fractional dose of 2 Gy during a period of six weeks). Figure 3.8 represents the tumor respond curves for four assumed tumor cases with different values of T_d. The tumor response curves of both cases show that tumor volume and number of viable tumor cells are higher for lower T_d values (fast growing tumor) during the course of treatment. Moreover, the effect of changing T_d is more noticeable for less radiosensitive case as expected based on previous observations. The radioresistance case

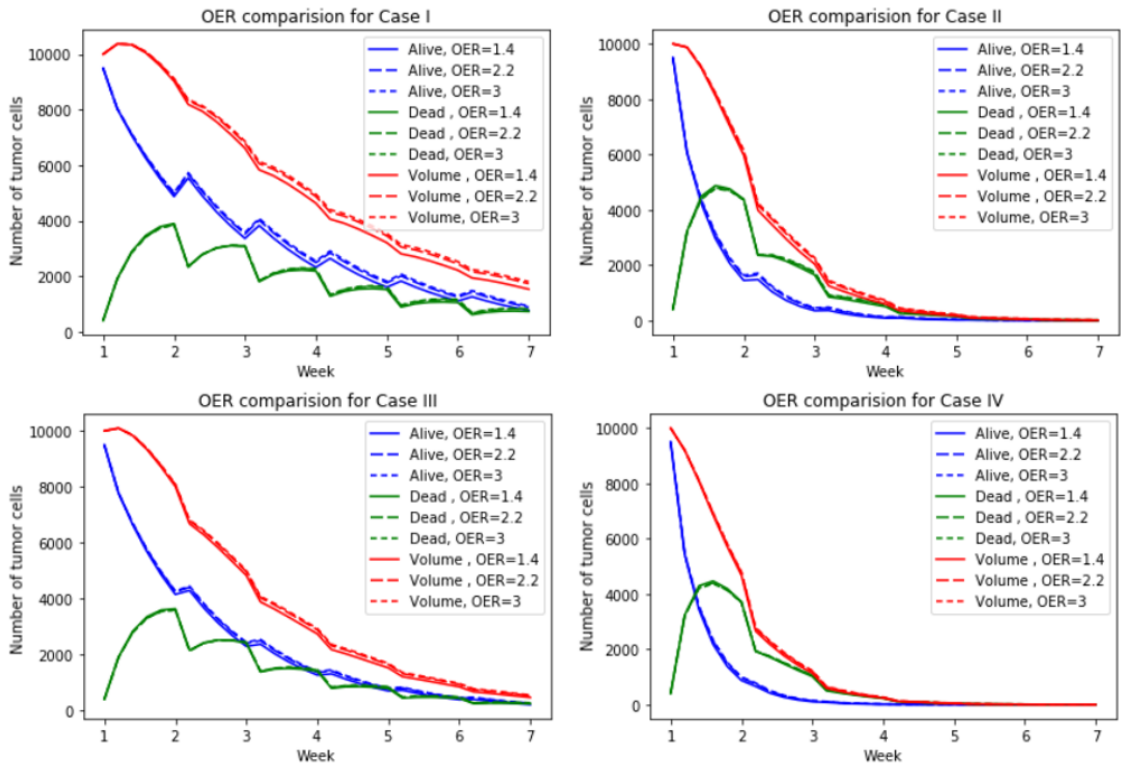


Figure 3.9: OER parameter effect on the tumor response curves from the proposed tumor response model

with the lowest doubling time or fastest proliferation ($T_d = 3$ days) has the highest final number of remaining tumor cells which means that the number of repopulated tumor cells were higher than the number of cells killed by RT.

The effect of changing OER parameter on tumor volume regressions based on the proposed tumor response model is explored for the four assumed cases and it is shown in Figure 3.9. Increasing the value of OER means that the hypoxia effect is more severe for the tumor; therefore, the tumor would be less radiosensitive. As shown in figure 3.9, changing OER does not affect the tumor response curves for radio sensitive tumors (Case II and Case IV) because the value of radiosensitivty parameter (α) is still high even in hypoxic phase. However, higher value of OER can lead to larger volume of tumor at the end of treatment. Also, faster tumor growth can enhance the effect of increasing OER value.

To find the impact of changing tumor decay parameter (τ_d) in tumor radiation response,

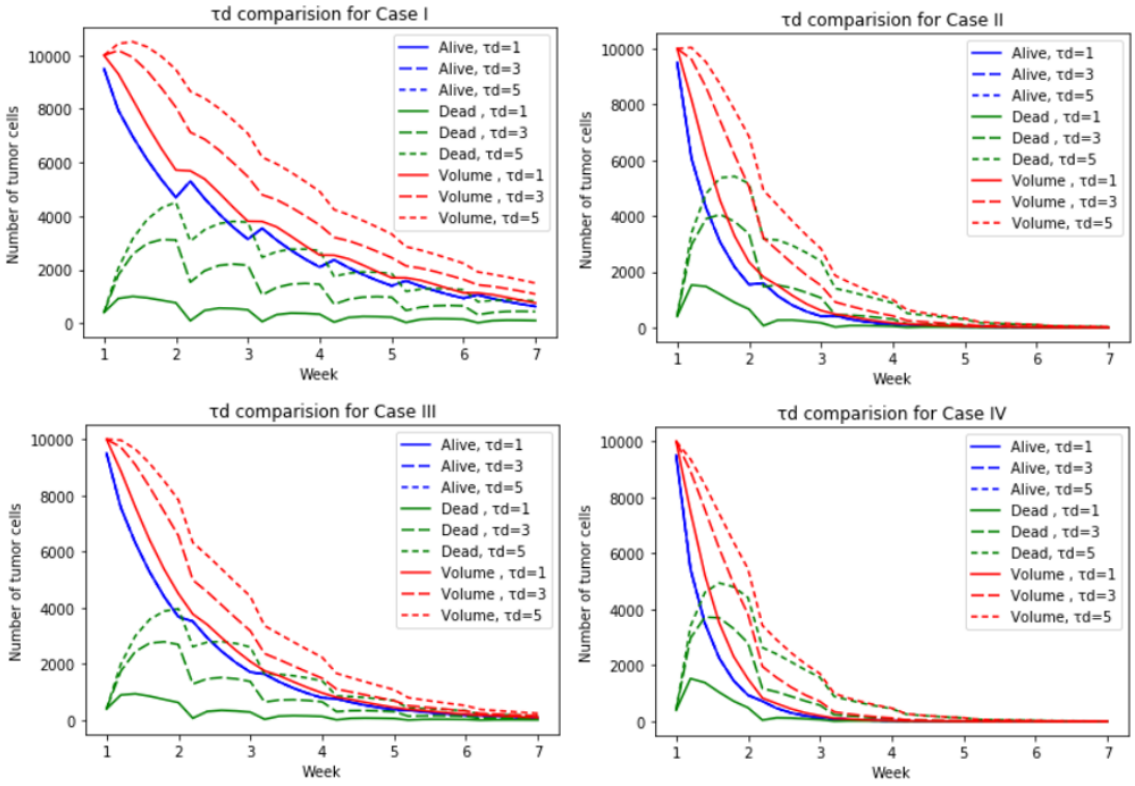


Figure 3.10: Decay parameter effect on the tumor response curves from the proposed tumor response model

we considered three common values for this parameter for each case and the results is presented in Figure 3.10. This figure shows that changing the value of τ_d only affect the number of dead cells which are in the process of decay. Having longer tumor decay time leads to have less decay of dead cells; as a result, the total number of tumor cells will increase. Also, we can see that changing τ_d affects the tumor response for all cases in a same way.

We assumed the same α/β ratio for all previous analyses and we investigate the sensitivity of the tumor response model to α/β values for tumors with high and low radiosensitivity. Figure 3.11 shows the simulated tumor radiation response using four α/β values of 2, 4, 8, 10 Gy for high and low radiosensitive tumors. The results shows that higher α/β decrease tumor volume regression rate during the course of treatment. This depletion is more noticeable for low radiosensetive tumors because tumors with high radiosensetivity

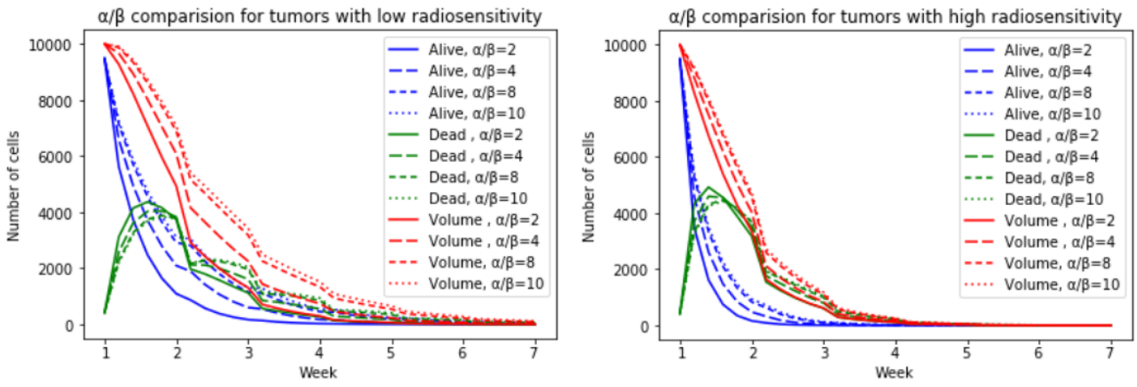


Figure 3.11: α/β parameter effect on the tumor response curves from the proposed tumor response model

can make up the changes in α/β.

3.3.2.3 The Effect of Variable Fractionation Scheme

The variable dose fractionation is another important factor affecting the tumor biological response during radiation therapy treatment. We investigated the impact of having different fractionation plans on the tumor BED, OAR BED, and tumor cell killing rate for the four cancer cases. We assumed that the treatment protocol is to deliver 60 Gy prescription dose five fractions per week for 6 weeks. We considered one plan with an equal fraction dose of $d = 2$, and two other plans with the variable fraction dose $d \in \{1.8, 2, 2.2\}$ with approximately same OAR toxicity ($BED^{\phi} \pm 1\%$) as equal dose plan. Table 3.4 summarizes the weekly dose per fraction and the values of BED^{T}, BED^{Φ}, and tumor cell killing rate $(1 - SF)$ for the four tumor cases. As shown in the table, the BED^{Φ} for both variable plans were the same, and it was increased by 0.3% compared to the equal dose plan. The value of BED^{T} was also increased for Plan (a) by 0.50% and decreased by 0.44% for Plan (b), on average.

Figure 3.12 shows the Tumor BED comparison based on these three plan for four tumor cases. As we can see from the figure, Plan (a) has higher BED than the Equal dose plan while Plan (b) has lower BED than the two other plans. Also, Case I shows the most

sensitivity to fractionation scheme. Furthermore, the tumor cell killing rate (1-Surviving fraction) was the highest for Plan (a) and the lowest for Plan (b) for all cases. This difference was more noticeable for cases with a worse treatment outcome (e.g., Case I). The results suggest that the variable fractionation can change the treatment outcome in terms of the tumor BED. The treatment outcome can be improved by changing the fractionation scheme specifically for the cases with a low BED.

Table 3.4: Weekly dose per fraction, tumor BED, OAR BED and tumor cell killing rate ($1 - SF$) based on each plan for the four tumor cases

	Weekly dose per fraction (Gy)								
	W1	W2	W3	W4	W5	W6	BED^T	BED^Φ	$1 - SF(\%)$
Case I									
Equal dose plan	2	2	2	2	2	2	55.61	23.02	97.70
Plan (a)	2	2.2	2.2	1.8	1.8	2	55.94	23.20	97.92
Plan (b)	2	2	1.8	1.8	2.2	2.2	55.25	23.20	97.28
Case II									
Equal dose plan	2	2	2	2	2	2	57.15	23.02	99.42
Plan (a)	2	2.2	2.2	1.8	1.8	2	57.40	23.20	99.50
Plan (b)	2	2	1.8	1.8	2.2	2.2	56.93	23.20	99.28
Case III									
Equal dose plan	2	2	2	2	2	2	72.77	23.02	99.42
Plan (a)	2	2.2	2.2	1.8	1.8	2	73.17	23.20	99.48
Plan (b)	2	2	1.8	1.8	2.2	2.2	72.32	23.20	99.33
Case IV									
Equal dose plan	2	2	2	2	2	2	78.44	23.02	99.90
Plan (a)	2	2.2	2.2	1.8	1.8	2	78.72	23.20	99.91
Plan (b)	2	2	1.8	1.8	2.2	2.2	78.31	23.20	99.90

3.3.3 RL Environment Generation and Variability Analysis

The environment of RL algorithm is ART treatment planning environment which includes all possible ART policies based on all possible tumor volume changes scenarios. At each time stage, a number of tumor volume cases and their associated BED and surviving fraction based on the state and action in the previous time stage are calculated and used in

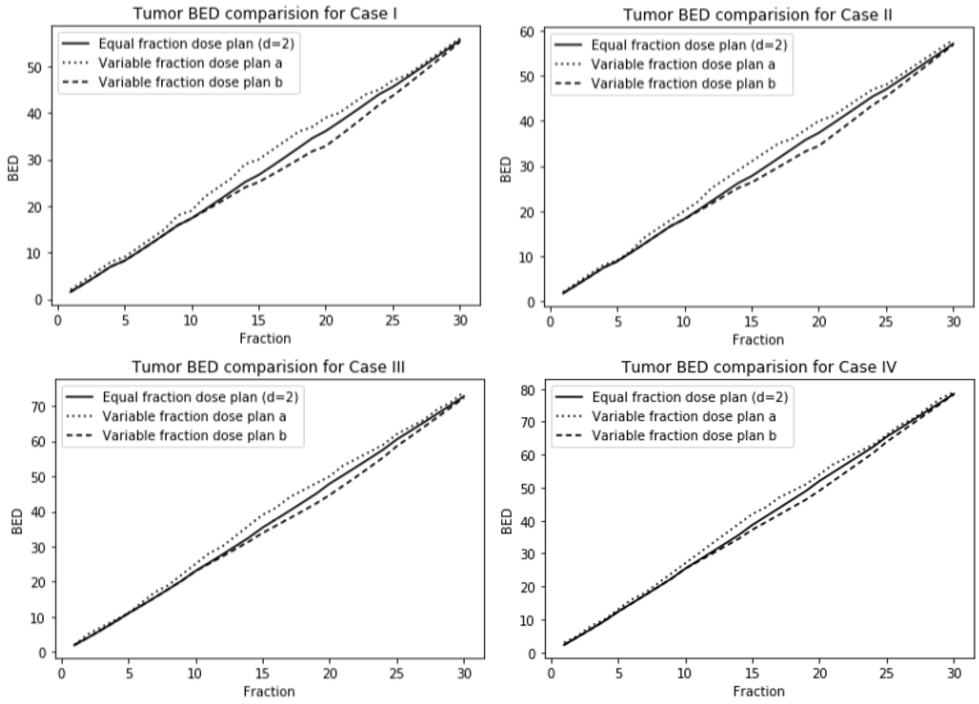


Figure 3.12: The effect of variable fractionation scheme on tumor BED for all four cases

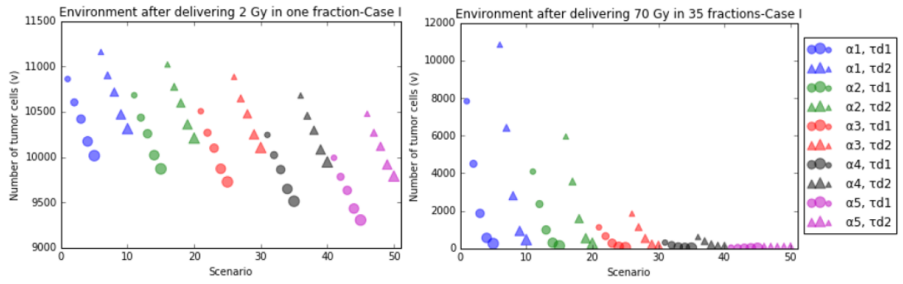
determining state and immediate reward. To incorporate the variability inherent in biological parameters within the RL environment, we choose a set of values for each parameter based on the possible range of it at each episode for a specific cancer site. As a result, the RL agent can see all possible values of parameters and take robust actions.

To better understand the existing biological uncertainty in tumor radiation response, the RL environment was generated based on the assumed ranges for the four general tumor cases. A set of values for each parameter was chosen randomly using uniform distribution. Since the tumor response model is very sensitive to α and τ_g values, we chose five random values for each one. two values for T_d are considered because the variability within the possible range of T_d was lower than α and τ_g. Since the observed effect of OER on tumor radiation response was negligible, We assumed $\alpha/\beta = 10$ and $OER = 1.8$ which are the most common reported values for them.

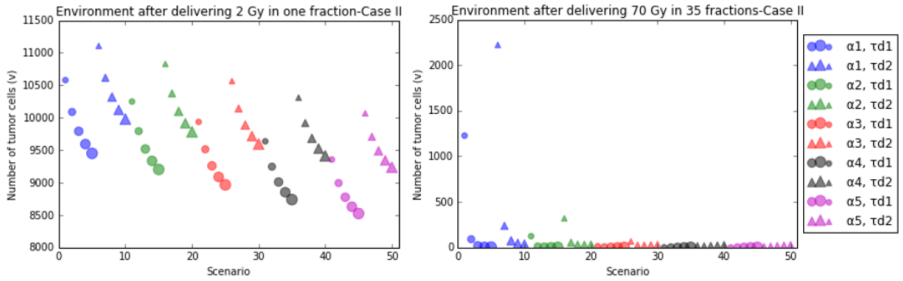
Figure 3.13 shows the number of remaining tumor cells (tumor volume) in RL environment for each assumed tumor case based on one set of values for uncertain parameters after delivering one fraction and 35 fractions of 2 Gy. Based on this figure, We can see that the uncertainty inherent in biological parameters can result in having variable treatment outcome in terms of tumor cell survival. This figure includes the variability caused by each uncertain biological parameter in RL environment for Case I (3.13(a)), Case II (3.13(b)), Case III (3.13(c)), and Case IV (3.13(d)). The number of remaining tumor cells for each value of α is shown by different colors in which the blue color and purple color represent the lowest and highest value of α, respectively. Circle is marker of fast tumor decay ($\tau_d = 2\ days$) and triangle is marker of slow tumor decay ($\tau_d = 6\ days$). Also, different marker size represents τ_g values such that larger values of τ_g are shown with larger marker size.

As shown in the figure, the variability in treatment outcome (residual tumor volume) is higher for less radiosensitive cases (Case I and Case III) than high radiosensitive cases (Case II and Case IV) which confirms the previous observations about high sensitivity of tumor response model to radiosensitivity parameter (α). Therefore, we can say that for each tumor case there is a threshold for α parameter that any α value larger than this threshold will result in total removal of tumor cells. Also, lower values of τ_g leads to have higher number of remaining tumor cells and higher variability among scenarios. Moreover, slower tumor decay (τ_d) lower tumor volume at each treatment stage and lower variability in final treatment outcome. In summary, we can conclude that variability among scenarios is almost the same for all four cases after the first fraction but the tumor volume range is different. Furthermore, the final treatment outcome in terms of residual tumor volume is the worst for case I which has low radiosensitivity and high rate of proliferation and the best outcome is for case IV with high radiosensitivity and slow proliferation as expected based on the parameters. Also, Case III shows more variability and worse final outcome

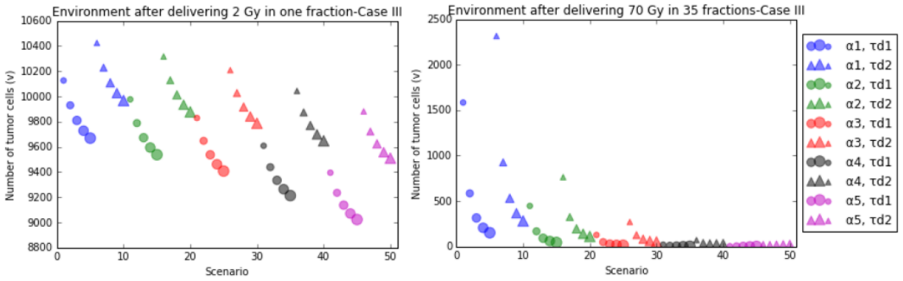
than case II which shows the higher sensitivity of response model to radiosensitivity than growth rate.



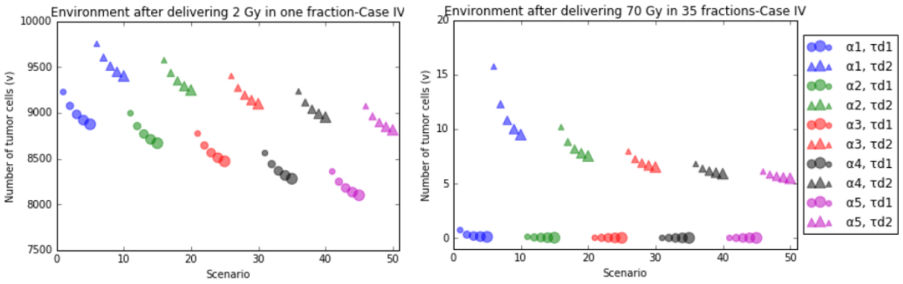
(a) Case I



(b) Case II



(c) Case III



(d) Case IV

Figure 3.13: Tumor volume variability in the RL environment for four tumor cases

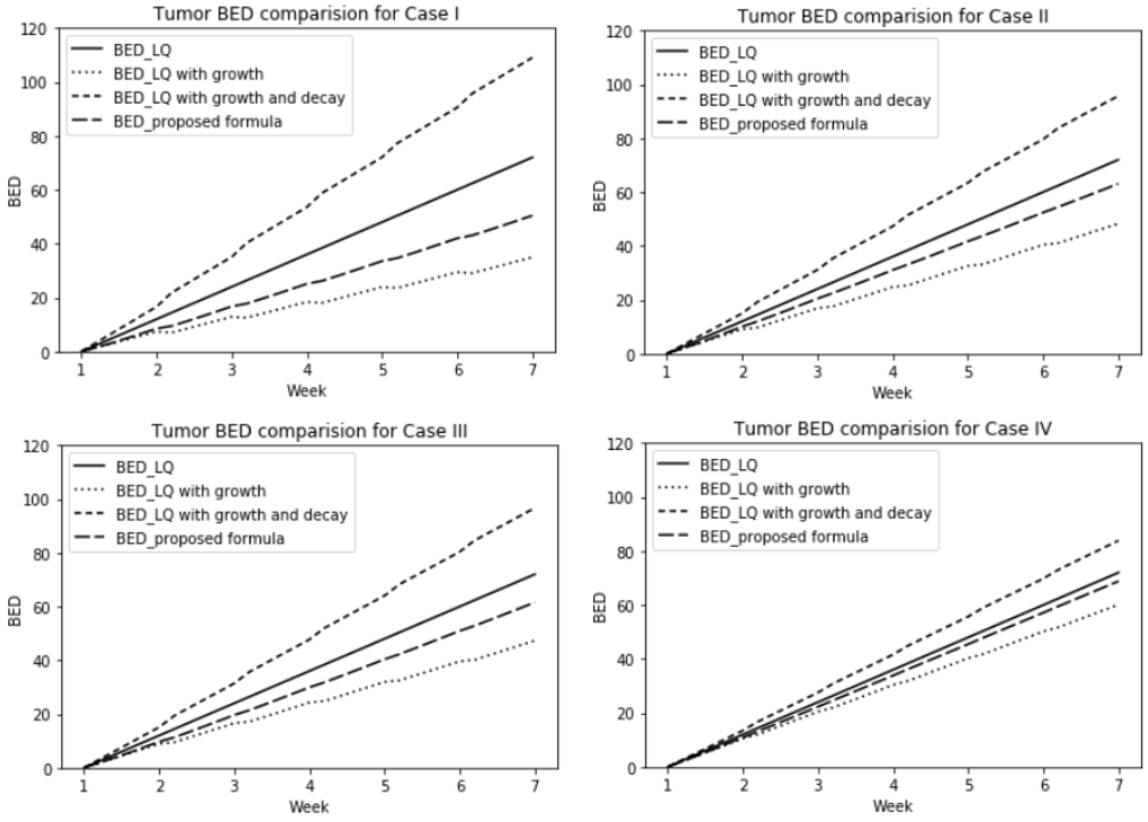


Figure 3.14: Tumor BED comparison based on different models for all four tumor cases

3.3.4 Sensitivity Analysis of BED Function

As discussed in section (2.3.1), we proposed equation (3.18) to estimate BED based on biological parameters, tumor response model, and tumor volume sub-compartments. In this section, we compare the proposed BED function with the BED formulations in the literature including BED based on LQ model[1] [6, 7], LQ model with repopulation [2] [10, 29], and LQ model with repopulation and decay [3] [29]. Figure 3.14 shows the tumor BED during the course of treatment (35 fractions of 2 Gy) based on four BED formulations for each assumed tumor case.

As shown in the figure, LQ model estimates an equal BED for all cases because of

[1] $BED_i = d_i(1 + \frac{d_i}{\alpha/\beta})$
[2] $BED_i = d_i(1 + \frac{d_i}{\alpha/\beta}) - \frac{\Delta t_i}{\alpha\ \tau_g}$
[3] $BED_i = d_i(1 + \frac{d_i}{\alpha/\beta}) - \frac{\Delta t_i}{\alpha\ \tau_g} + -\frac{\Delta t_i}{\alpha\ \tau_d}$

the assumed equal α/β ratio. So, BED based on LQ model cannot reflect the different treatment outcome of these cases. Moreover, incorporating tumor growth in LQ model results in having lower BED values which is the same for cases with the same tumor growth rate and cannot reflect the variability in treatment outcome between these cases (e.g., Case I and Case II). Also, considering both tumor decay and growth can neglect the effect of tumor growth and result in misleading information about the treatment outcome due to the inaccurate value of tumor decay parameter. Therefore, we proposed a weighted formulation for BED that considers tumor growth and decay and hypoxia effect in each volume sub-compartment based on the tumor response to radiation.

Figure 3.14 represents that the estimated BED based on the proposed model is the lower than the BED from LQ model and greater than the BED from LQ model with repopulation. Case IV and Case I show the highest estimated BED (closest to LQ model) and lowest estimated BED (closest to LQ model with repopulation), respectively, which confirms the previous observation for treatment outcome in terms of remaining tumor volume cells after 6 weeks of treatment. Also, Case II shows the slightly higher BED than Case III as expected. As a result, we can conclude that higher radiosensitivity and slower tumor repopulation increase the proposed estimated BED and eradication of tumor cells.

3.3.5 A Case Study on a Clinical Lung Cancer Cancer Case

We evaluated the performance of our proposed RT treatment planning framework on an actual clinical non-small cell lung cancer (NSCLC) case obtained from the MD Anderson Cancer Center (MDACC), Houston, TX. The patient went through a four-dimensional CT imaging as a part of a routine treatment simulation before starting the radiation therapy. The target volume and normal structures were manually contoured on axial slices of the planning CT scan by a physician. The anatomy was discretized into voxels of 2.5 mm (L) $\times$ 2.5 mm (W) $\times$ 2.5 mm (H). Table 3.5 lists the organs of interest, voxel counts of each

organ, and the prescribed treatment protocol and requirements.

Table 3.5: Organs of interest, voxel counts of each organ, and dose-volume requirements for the volumes of interest

Structure	Structure Type	Number of Voxels	Dose Requirements
Planning target volume (PTV)	Target	59,030	Volume receiving at least the prescribed dose: $\geq 95\%$
Heart	OAR	43,180	Volume receiving doses higher than 45 Gy: $\leq 65\%$
Total lung	OAR	287,616	Volume receiving doses higher than 20 Gy: $\leq 45\%$

We made the following assumptions to construct the RL environment. As it was done in the clinic, five fractions of radiation dose are to be delivered to the target in each week, skipping the treatment during the weekends to allow healthy tissues to recover. A total of 35 fractions, $N = 35$, will be delivered to complete the treatment. The optimal plan allows the total delivered dose to be in the range between $d_l = 68$ Gy and $d_u = 72$ Gy [146] based on the tumor biological characteristics. The set of possible actions (i.e., fraction dose) of $A = \{1.8, 2.0, 2.2\}$ was assumed to include a 0.2 Gy deviation from the conventional fraction dose, which is reasonable in ART fractionation scheme [146, 147].

Increasing the tumor's BED while keeping the OAR BED at a safe level can help improve tumor control without elevating the OAR toxicity. Therefore, our goal is to increase the BED of the tumor compared to the one from the reference plan, while keeping the OAR BED to be, at most, $+5\%$ from the reference plan. Also, the surviving fraction was capped at 0.01% to ensure the elimination of all tumor cells. The corresponding penalty coefficients in the reward function were determined by manual adjustments to achieve the desired goal of the treatment plan according to the treatment planner's preference.

Ranges of biological parameters for lung cancer were chosen based on the literature to set up the RL environment for tumor response during the treatment period. We assumed

an exponential tumor growth rate of $\tau_g \in [10, 60]$ days [144, 148, 149] and and a tumor decay factor of $\tau_d \in [2, 6]$ days [139, 150], the uncertain ranges of $\alpha_p \in [0.20, 0.365]$ Gy^{-1} [139, 144] and $\alpha/\beta \in [4, 10]$ Gy [139, 145, 151, 152]. The standard value of oxygen enhancement ratio for lung cancer was considered as a constant value of $OER = 1.7$.

The total lung radiation toxicity is one of the most important metrics in determining the quality of RT plan for lung cancer. In this study, we assumed the total lung as an OAR with $\alpha = 0.3\ Gy^{-1}$ and $\alpha^{\phi}/\beta^{\phi} = 3$ [10, 153, 154], OAR repopulation rate $\tau_g^{\phi} = 15$ days, and repair rates $\tau_r^{\phi} = 3.5$ days [9, 21] in the RL environment. We also considered a constant OAR dose sparing factor of $\gamma = 0.7$ [10].

At each episode of the training process, three values for α and τ_g, and two values for τ_d are randomly chosen from the defined ranges which resulted in generating 50 tumor volume cases. At each time stage i within an episode (i.e., each fraction), 50 tumor volume cases and their associated $BED(s_i, a_i)$ and surviving fraction based on the current state s_i and selected action a_i were calculated and used in determining the state and immediate reward.

To find the optimal adaptation points, we need to determine the optimal action at each time epoch (i.e., the beginning of each week). Hence, we trained the RL model using Algorithm 1 to find the optimal Q value function $Q^*(s, a)$, as a prediction of the reward function, with a minimum loss function value $L(w)$ in Equation (3.32). The model with the minimum loss value was achieved at 44000^{th} episode of training, which is the last episode before over-fitting happens. The network loss within an episode in every 1000 iterations is depicted in Figure 3.15. As shown in the figure, the network training loss has fluctuations at the beginning of training and it decreased as the episode increased. This behavior is reasonable as the RL model tries to learn over time to predict the Q value function with the highest precision.

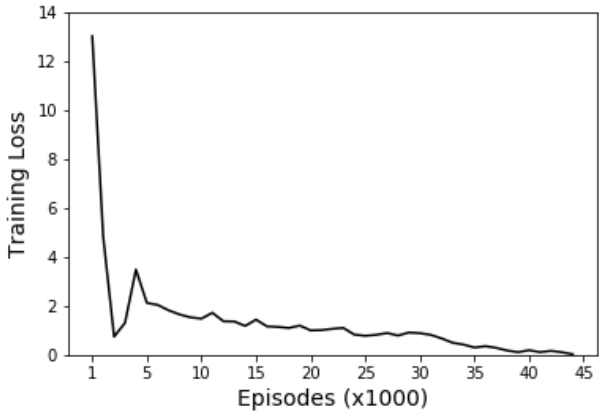
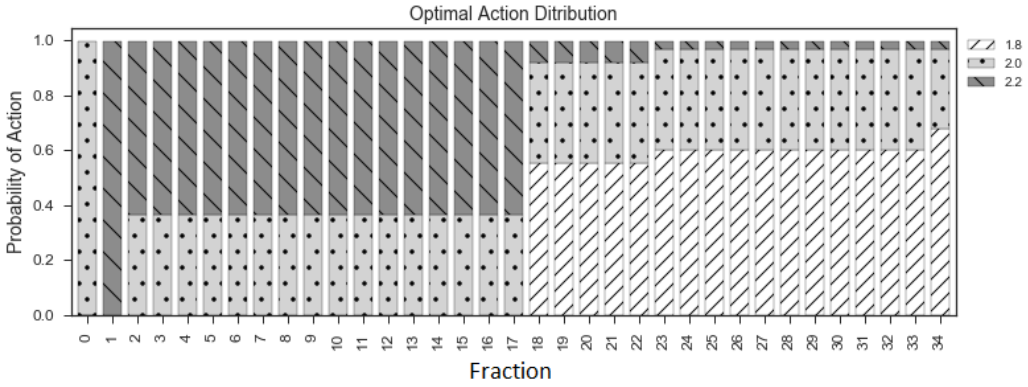


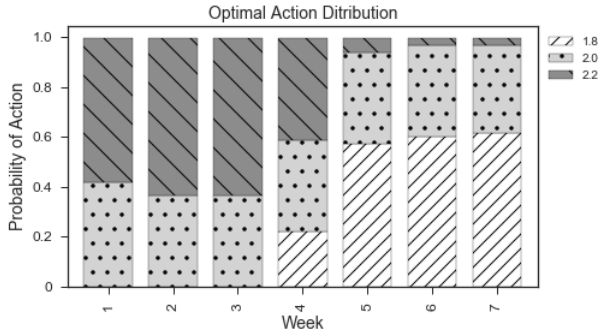
Figure 3.15: Training loss value in every 1000 episodes

The best trained RL model is used to find the optimal fractionated plans for 500 scenarios, which are generated based on different sets of selected biological parameters (i.e., α, τ_g, τ_d) of a lung cancer patient. For each scenario, an optimal action (i.e., dose) was selected to obtain the optimal plan per fraction. Then, the expected dose amount over 500 scenarios was calculated to generate the optimal fractionation scheme. Figure 3.16(a) shows the distribution of optimal actions at each fraction. Note that a weekly fractionation is typically used for ART in a clinical practice. Hence, we determined the distribution of optimal actions for each week by concatenating distributions of five fractions within a week (see Figure 3.16(b)). As shown in Figure 3.16, the most probable policy is to increase the fraction dose at the beginning of the treatment then decrease the dose at some point.

Using the weekly distribution of optimal actions, the expected dose amount over 500 scenarios for each week, sum of the probability of each dose in action set multiplied by the dose, was calculated to determine the optimal fractionation scheme (i.e., $d_i = P(d_i = 2.2)\times 2.2 + P(d_i = 2)\times 2 + P(d_i = 1.8)\times 1.8$). Since changing the weekly fractionated dose by a small amount is not feasible in clinical practice, we considered week i as an adaptation point if the difference between the current dose regime and the projected dose for week i is greater than a threshold (i.e., $|d_i - d_{i-1}| \geq 0.03\ Gy$). Once the adaptation points are



(a) Distribution of optimal actions for 35 fractions



(b) Distribution of optimal actions for 7 weeks

Figure 3.16: Distribution of optimal actions; (a) distribution of optimal actions for 35 fractions, (b) distribution of optimal actions for 7 weeks of treatment.

identified, the ART weekly dose fractionation scheme is determined in such a way that the new weekly dose amount will be the average dose over all previous weekly doses since the last adaptation point. We used the conventional fractionation schedule with an equal weekly fractionated dose as a reference plan to compare with our proposed plan. Table 3.6 presents the weekly fractionation for the lung cancer patient based on the expected value of actions over the week, corrected values for ART, and the reference plan.

The optimal ART schedule (*ART(I)*) is to have an initial radiation dose of $2.13\ Gy$ for the first three weeks, and decrease it to $2.04\ Gy$ at week four, and finally drop the dose to $1.86\ Gy$ for the rest of the treatment. This implies that delivering a higher radiation dose at the beginning of the treatment will cause more damage to the tumor cells and this will

Table 3.6: Weekly dose per fraction and total dose for the generated ART plan and the reference plan

	Weekly dose per fraction (Gy)							
	W1	W2	W3	W4	W5	W6	W7	Total Dose
Optimal fractionated plan	2.12	2.13	2.13	2.04	1.89	1.85	1.85	70.05
Optimal ART plan	2.13	2.13	2.13	2.04	1.86	1.86	1.86	70.05
Reference plan	2.00	2.00	2.00	2.00	2.00	2.00	2.00	70.00

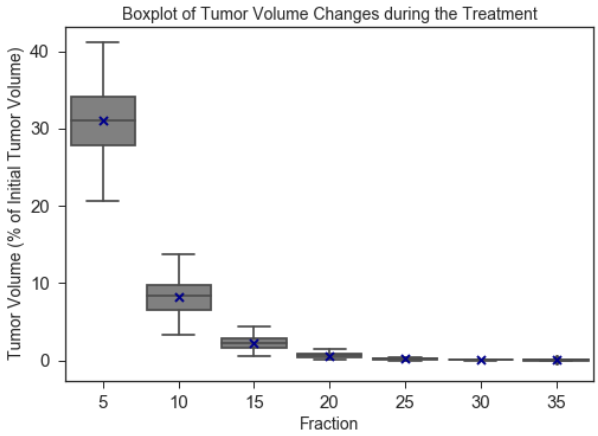


Figure 3.17: Box plot of the weekly tumor volume as a percentage of initial tumor volume among the assumed scenarios

change the dose requirement for the rest of the treatment to be lower.

The total dose of the proposed plan is slightly higher than the reference plan. This is because delivering a higher amount of radiation dose in a treatment will likely increase the total biological effective dose of the tumor and OARs cells. However, we can only increase the total dose by a certain amount (i.e., at most $+5\%$ of the reference plan) to maintain the desired range of OAR BED based on the assumed preferences. Figure 3.17 shows the box plot of the weekly tumor volume as a percentage of initial tumor volume among the assumed scenarios which has a decreasing trend during the course of treatment.

3.3.5.1 Plan Evaluation: Biological Comparisons

To quantify the extent of potential biological benefits of the proposed approach, final surviving fraction (SF) of the tumor and biological effective dose (BED) of tumor and OAR for reference plan and the optimal ART plan based on the assumed decision policy were compared. BED is one of the most important biological measures to evaluate the quality of a RT treatment plan because BED of a structure can represent the biological damage in the structure very well. So, increasing the tumor BED while keeping the OAR BED in a acceptable level can help to achieve more damage to the tumor without increasing the harm to the healthy organs. Tumor's surviving fraction is also important to ensure the effectiveness of the treatment plan in term of removing all tumor cells which was set to be lower than a desired ϵ value (0.01 %) in the current experiment.

Table 3.7 summarizes these biological metrics from the experiments based on the generated scenarios for the reference plan and the developed ART plan. In terms of BED, the generated ART plan performed better than the reference plan for the tumor, increasing the mean value by 2.01%, while affecting the OAR BED by 0.49% (compared to the reference plan). Furthermore, The tumor's surviving fraction was also well controlled in the generated plan and the probability of having lower values of SF ($P(SF < 0.01\%)$) among the scenarios was improved by 42.31%.

Table 3.7: Biological measures for the optimal and the reference plan

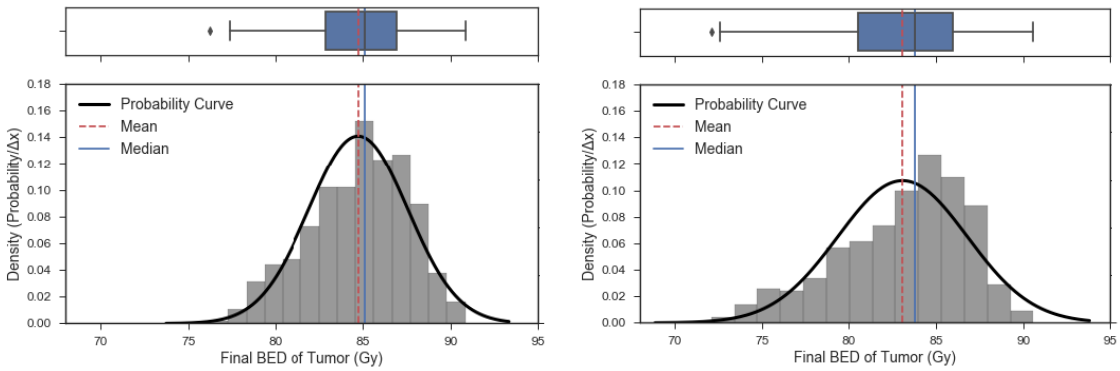
	BED^T	BED^Φ	$1 - SF(\%)$	$P(SF < 0.01\%)$
Optimal ART plan	84.74 (95% CI [84.48,85.01])	26.63	99.98	0.37
Reference plan	83.07 (95% CI [82.74, 83.41])	26.50	99.97	0.26

Measuring the outcome solely based on the mean BED value may hide some individual worst-case scenarios whose values are much lower than the desired value for the tumor,

which can lead to undesirable effects on the treatment outcome. Therefore, it is important to reduce the variability of tumor BED values under different scenarios, where the RL framework can help address the issue.

Figure 3.18 shows the histogram, the estimated probability density function, and the box plot of the final BED of the tumor for each of the plans corresponding to 500 scenarios. The sum of the probability densities is equal to 1. The optimal ART plan resulted in the final BED^T distribution, which appears to follow a normal distribution with a small variance and higher values around the average BED. As a comparison, the reference plan's distribution has a wider variance and is skewed left. Also, the ART plan has a shorter left tail distribution compared to the reference plan, which means that the ART plan will result in a smaller number of undesirable worst-cases for BED^T. This suggests that the ART plan is more likely to produce a treatment plan with a greater final tumor BED than the reference plan.

We further evaluate the variability in the solution for each plan using commonly used variability metrics in statistics, including median, range, and interquartile range (IQR). In Figure 3.18, mean and median values are marked with a red dash line and a blue line, respectively. We observed that the ART plan resulted in a smaller variability compared to the reference plan. First, the BED distribution of the ART plan has a narrower spread. Second, both the mean and median values of the ART plan are higher than those of the reference plan. Furthermore, the ART plan resulted in a smaller difference between the mean and median values compared to the reference plan. Third, the IQR can be visualized using the box width in the box plot. We can see that the ART plan exhibited a 25% narrower IQR than the reference plan. Also, the range of tumor BED values is reduced by 21%. Overall, all variability measures of the optimal ART plan were lower than those of the reference plan. This can be interpreted as meaning that the ART plan performs better than the reference plan in terms of improving the final BED^T and reducing its variability in



(a) ART plan (b) Reference plan

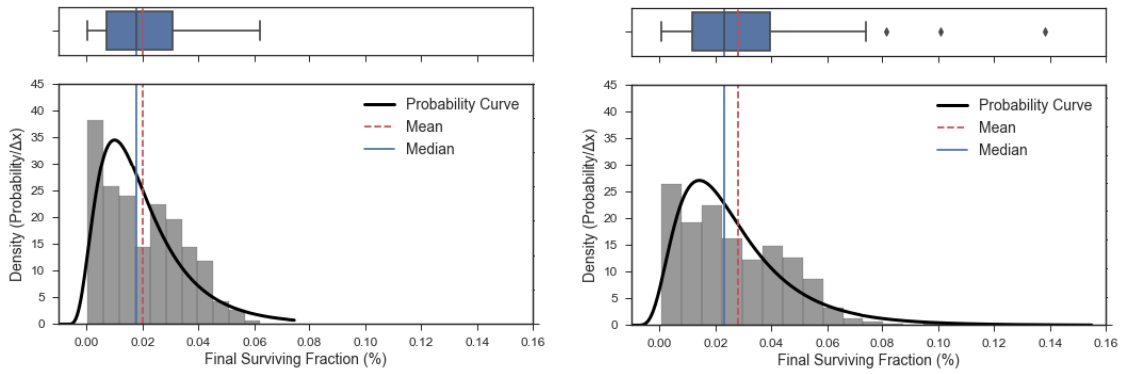
Figure 3.18: Histogram, estimated probability density function (PDF), and box plot of the final tumor BED for the 500 generated scenarios based on (a) the ART plan, and (b) the Reference plan.

uncertain biological parameters.

We used the normal distribution for estimating the PDF for the BED values because it has the best fitting results compared with other well-known distributions (i.e., Lognormal, Gamma, and Beta), which can be explained based on the central limit theorem [155] and randomness of biological parameters. However, a lognormal distribution was used to estimate PDF of SF values because the relationship between the SF and the BED of the tumor is logarithmic. Figure 3.19 shows the histogram, the estimated probability density function, and the box plot of the final tumor's SF from each plan.

The ultimate goal of treatment planning is to produce a plan whose final SF is minimized. As shown in the figure, the final SF distribution is skewed right. The long tail on the right side corresponds to less undesired cases having a large SF (i.e., worst-cases). Comparing the two treatment plans in Figure 3.19, the ART plan reduced the right tail distribution of the final tumor SF as well as the upper quartile of the box plot. Hence, the ART plan outperformed the reference plan in terms of controlling the worst-case tumor SF. Moreover, the final SF values of the ART plan showed a tighter distribution and higher density around lower values compared to the reference plan. The ART plan also resulted in reducing the IQR and value ranges by 11.95% and 46.79%, respectively. Therefore, we

claim that the ART plan has an advantage over the reference plan in developing treatment plans under the biological parameter uncertainty.



(a) ART plan (b) Reference plan

Figure 3.19: Histogram, estimated probability density function (PDF), and box plot of the final tumor SF for the 500 generated scenarios based on (a) the ART plan, and (b) the Reference plan.

3.3.5.2 Effect of Different Decision-Making Preferences

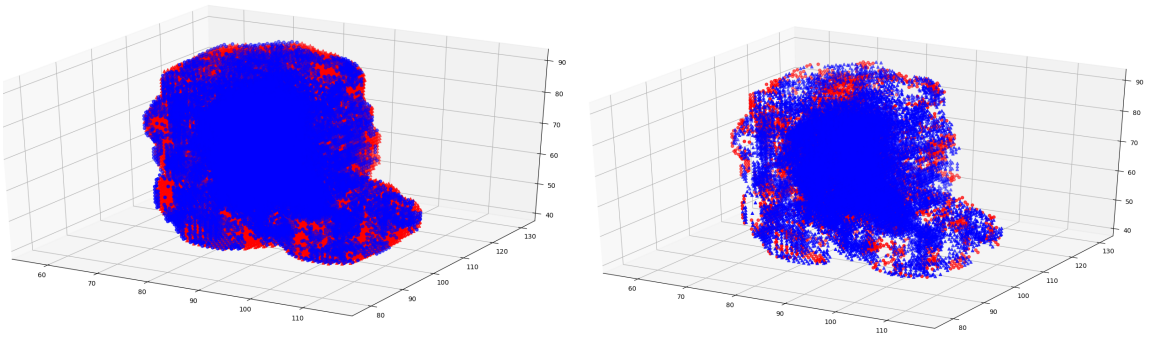
We further investigate the effect of changing a planner's preference and assumptions on the final optimal fractionation schedule as well as the quality of treatment in terms of biological and dosimetric measures. Based on the patient's characteristics and the type of cancer, the treatment planner can set different priorities and goals in the RT treatment planning. For instance, one may want to develop a plan to ensure that the target is receiving the required dose by controlling the surviving fraction ($SF < \epsilon$), while reducing the radiation damage to OARs in terms of the BED. This can be done by assigning a higher weight on the OAR BED part of the reward function than the tumor BED, while satisfying the surviving fraction. We expect that the OAR toxicity will improve as a result, and this leads to a better patient recovery from radiotherapy. In this regard, we also explored the effect of changing OARs sparing factor (γ) in the BED formulation by finding the optimal plan based on two values of $\gamma = 0.7$ and $\gamma = 0.4$. Table 3.8 shows the optimal fractionation schedule, total

dose, and OAR BED reduction percentage (ΔBED^{Φ}) based on this plan along with different values of γ. Compared to the conventional plan with a prescription dose of 70 Gy, the ART plan reduced the total dose by 2.14% and 0.71% based on the new preference (or priority) with $\gamma = 0.7$ and $\gamma = 0.4$, respectively. This reduction in total radiation dose may not seem to be significant, but this will result in better OAR sparing with a lower amount of radiation. In both plans, the total treatment dose is decreased because the priority was made to reduce the OAR BED, while having the same or better surviving fraction of the tumor cells and limiting the BED in target volume to a clinically desired level. As we can see from Table 3.8, the OAR BED is improved by 9.26% and 16.12% using $\gamma = 0.7$ and $\gamma = 0.4$, respectively, compared to the reference plan on tumor.

Furthermore, the ART with $\gamma = 0.7$ resulted in a more aggressive plan than the one with $\gamma = 0.4$. This is because the OAR will likely reduce the radiation exposure by lowering the value of γ. Thus, a lower penalty value was assigned to the OAR BED in the reward function and we can see a higher BED^{Φ} depletion even with a smaller amount of dose reduction. The results from this experiment show that the proposed treatment planning framework is effective to develop a plan that preserves more healthy cells. Therefore, the planner can develop the best plan according to the patient characteristics and the physician's preference.

Table 3.8: Weekly dose per fraction and OAR BED based on the new policy and different values of γ

	Weekly dose per fraction (Gy)								
	W1	W2	W3	W4	W5	W6	W7	Total Dose	ΔBED^{Φ}
$\gamma = 0.7$	2.00	2.00	1.90	1.90	1.90	2.00	2.00	68.50	-9.26%
$\gamma = 0.4$	2.00	2.00	2.00	2.00	1.95	1.95	2.00	69.50	-16.12 %



(a) First adaptation point (b) Second adaptation point

Figure 3.20: The simulated residual tumor volume (blue points) and the removed voxels (red points) at (a) the first adaptation point (beginning of week 4), and (b) the second adaptation point (beginning of week 5).

3.3.5.3 Plan Evaluation: Dose-Volume Results

Dose-volume metrics are commonly used to evaluate the treatment plan quality. For the purpose of simulating the ART procedure in this section, treatment plans were adapted to the patient volumetric changes at each adaptation point. Planning target volumes (PTVs) were generated analogously on the basis of the average estimated residual tumor volume over all scenarios using the proposed tumor response model. We removed the tumor cells receiving a dose higher than the tolerance threshold from k outer layers (i.e., *2.5mm* per layer) of the tumor volume until achieving the estimated residual tumor volume at each adaptation point. Figure 3.20 shows the residual tumor volume and the removed voxels at each adaptation point after the first iteration.

To evaluate the OAR toxicity of the generated plans, the dose-volume metrics were calculated for the heart and total lung for three plans: the reference plan, ART(I) with $\gamma = 0.7$ (see Section 3.3.5), and ART(II) with $\gamma = 0.7$ (see Section 3.3.5.2). For the heart, V_{45} and D_2 were measured to examine the level of a high dose of radiation, which is critical for a serial organ. Also, V_{20} and mean dose were measured to account for the average spread of radiation dose in the total lung, which is a parallel organ. Table 3.9 summarizes these dose-volume metrics for the proposed plan and the reference plan. Note

that all dose-volume values were normalized to have 95% of PTV receiving at least 70 Gy for the comparison purpose. We make the following observations regarding the ART plans in comparison to the reference plan based on the table. First, all metrics of the ART plans for the heart and total lung were lower compared to the reference plan. Second, ART(I) reduced V_{45} and D_2 of the heart by 3.78% and 0.10%, respectively. Third, ART(II) decreased V_{45} and D_2 of the heart by 6.01% and 2.48%, respectively. Finally, both ART(I) and ART(II) reduced V_{20} of total lung by 3.56% and 4.28%, and mean lung dose by 2.14% and 5.63%, respectively.

Overall, the ART plans outperformed the reference plan by reducing OAR toxicity. ART(II) showed slightly lower values on all measured dose-volume metrics, which is a direct result of the planner's preference to keep the OAR toxicity at the desired level.

Table 3.9: Comparison of dose-volume metrics for optimal plans under two policies and the reference plan

	Optimal ART plan		
	Policy I	Policy II	Reference plan
Heart			
$V_{45}(\%)$	11.69	11.42	12.15
$D_2(Gy)$	71.84	70.13	71.91
Total Lung			
$V_{20}(\%)$	30.85	30.62	31.99
$Mean(Gy)$	18.77	18.10	19.18

3.4 Conclusion

Multiple studies demonstrate the benefits of ART in terms of healthy tissue sparing and tumor cell reduction. Considering the biological features of tumor and healthy organs in treatment planning and adapting the plan to biological changes during the course of treatment is the key motivation for ART. In this paper, we developed a novel biological response

model that incorporates important biological factors for tumor and healthy organs to predict the tumor volume regressions during the treatment. Then, we proposed an automated framework using Reinforcement Learning and an optimization method to find the optimal adaptation points for ART and dynamically adapt the plan considering the tumor's uncertain biological response over time. We aimed to achieve a plan with a maximum final tumor control, while minimizing or maintaining the OARs toxicity levels by finding the actions to maximize the RL reward function. After finding the adaptation points, an optimization model was solved to find the optimal beamlet intensities satisfying clinical dose-volume requirements for the patient based on the predicted tumor volume and the proposed fractionation dose determined by the RL approach at each adaptation point.

We evaluated the performance of our proposed RT treatment planning framework using a clinical non-small cell lung cancer (NSCLC) case. We also analyzed the proposed approach under various assumptions and decision priorities to see the trade off in terms of tumor coverage and OARs toxicity. The proposed ART plans were assessed and compared with the reference plan (i.e, equal dose fractionation) based on biological and dose-volume metrics. The results showed that the proposed approach can help the treatment planner to achieve a robust solution under high levels of uncertainty in the biological parameters. Using the proposed method, it is not only possible to control the biological aspect of the treatment and tumor biological response uncertainty, but it also helps satisfy dose-volume requirements and clinical limits of the treatment. Furthermore, the proposed reinforcement learning framework can help achieve a robust solution under uncertainty in the biological parameters, while reducing the variability in the solution and improving the control on the worst-cases. The proposed approach enables the physicians to find an appropriate personalized ART plan in terms of fractionation dose and the timing of the adaptations. Two major benefits of this approach are to reduce the time and effort to collect large-scale datasets and avoid the need for taking expensive CT images at each visit. The proposed RL approach

can be easily applied to various types of cancer, ART methods, and different treatment planning preferences.

For the future work, the predicted tumor response to the radiation should be validated or corrected by obtaining actual imaging data for every visit during the course of RT treatment (e.g., at the determined adaptation points). This will help further enhance the proposed approach. Moreover, robust optimization techniques can be used in a beamlet optimization model to handle physical uncertainties while controlling the biological uncertainties within the reinforcement learning framework.

Chapter 4

Assessment of Radiation-Induced Lymphopenia Risk for Cancer Patients Treated with Photon Versus Proton Therapy

4.1 Introduction

Significant lymphocyte count depletion (i.e., lymphopenia) is a common toxicity of radiation therapy and is associated with worse disease control in a number of solid tumors, including esophageal cancer [65–68]. Because lymphocytes have a substantial role in the body's anticancer immune response, severe lymphopenia can reduce a patients' survival even in the early stages of tumor progression [68–72].

Multiple recent studies have shown that severe lymphopenia is strongly associated with poor treatment outcomes in a number of solid tumors such as cervical [72, 73], pancreatic [74, 75], rectal [76], lung [69, 77], and esophageal [66, 78] cancers. Thus, preservation of the lymphocytes from radiation damage is crucial for the effectiveness of radiation therapy, and it is critical to understand the clinical and dosimetric factors affecting the severity and incidence of RIL and develop strategies for its mitigation.

Radiation-induced lymphopenia (RIL) commonly occurs in conventional photon radiation therapy, presumably due to the high radiosensitivity of lymphocytes and the large low and medium dose bath of photon therapy. Dose distribution patterns from protons and photons can differ greatly, and the dosimetric advantages of state-of-the-art proton therapy over photon therapy in terms of sparing organs at risk and normal tissue have been demonstrated in several clinical studies [80, 81]. Also, intensity-modulated proton therapy (IMPT) performs better than intensity-modulated radiation therapy (IMRT) in terms of dose sparing and robustness towards common anatomical changes in esophageal cancer patients [82].

Moreover, RIL risk likely varies by treatment modality. Recent studies have reported greater lymphocyte depletion in patients treated with photon therapy than with proton therapy [66, 68, 77–79] For example, Shiraishi et al. [66] reported that proton beam therapy was associated with a lower risk of grade 4 lymphopenia compared with IMRT in esophageal cancer patients receiving neoadjuvant chemoradiotherapy.

In this study, we aimed to evaluate the relationship between severe lymphopenia and delivered radiation dose to the patient using both measured and estimated absolute lymphocyte count (ALC). We also modeled expected ALC depletion kinetics in esophageal cancer patients treated with three different modalities: IMRT, passive-scattering proton therapy (PSPT), and IMPT.

4.2 Methodology

In this section, we will describe two prediction models for lymphocyte depletion based on radiation doses and then report the patient selection for model validation as well as treatment planning of IMRT and IMPT for patients treated with PSPT.

4.2.1 Patient Selection and Treatment Planning

Ten esophageal cancer patients treated with PSPT (delivered in 28 fractions) at MD Anderson Cancer Center were included in this study. Baseline ALC and weekly ALC measurements during radiation treatment were recorded. Five patients had six weekly measurements; four patients had five measurements; and one patient had only four measurements. ALC nadir was defined as the lowest among the weekly measurements for each patient. The average ALC nadir for the 10 patients was $0.34\ K/\mu L$, ranging from $0.07\ K/\mu L$ to $0.68\ K/\mu L$. Three of the patients had grade 2 lymphopenia, four had grade 3, and Three had grade 4. All patients received concurrent chemoradiation therapy, during which chemo regimens were doublets of a taxane, fluorouracil, or platinum-based compound. Examples of the PSPT dose distributions for all patients can be found in Figure 4.1.

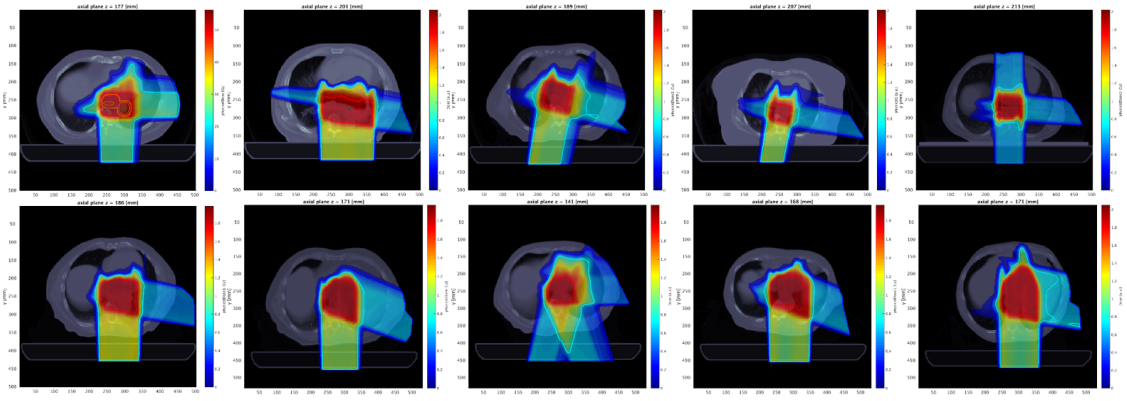


Figure 4.1: Dose distributions on an axial plane of PSPT plans for 10 esophageal cancer patients.

We used MatRad [156], a research-oriented treatment planning system, to create IMRT and IMPT plans for each patient. The prescription dose to the clinical target volume (CTV) was 50.4 Gy in 28 fractions for all patients. In optimization of the IMRT and IMPT plans, the same dosimetric criteria were used for each patient, but objective weights and constraints were adjusted, when necessary, to achieve the best possible target coverage and normal tissue sparing. All three plans were normalized to have 95% of the planning target

volume (PTV) receive the prescription dose per patient.

4.2.2 ALC Depletion Prediction Using a Piecewise-linear Lymphocyte Survival Function

On the basis of the dose distributions in each patient, we estimated the ALC during treatment by using a piecewise-linear relationship between lymphocyte survival and dose per fraction for each modality. The piecewise-linear function was modeled by interpolating previous findings about radiation-induced lymphocyte death (Yovino et al 2013, Nakamura et al 1990). Nakamura et al. (1990) reported the percentages of surviving lymphocytes as 90%, 50%, 10%, and 0% for radiation doses of 0.5 Gy, 2 Gy, 3 Gy, and 6 Gy or higher for each fraction, respectively. This estimation assumes that all circulating blood cells may receive doses by the end of treatment (after 28 fractions in this study). The lymphocyte survival probability for each voxel i, S_i, after receiving a fractional dose d (in Gy), can be calculated using the following piecewise-linear function

$$S_i(d) = \begin{cases} -0.2\,d + 1 & 0 \le d < 0.5 \\ -0.26\,d + 1.03 & 0.5 \le d < 2 \\ -0.4\,d + 1.3 & 2 \le d < 3 \\ -0.03\,d + 0.2 & 3 \le d < 6 \\ 0 & 6 \le d. \end{cases} \quad . \tag{4.1}$$

We assumed that the initial number of lymphocytes in 1 μL of body volume can be estimated by multiplying the pretreatment ALC value for each patient by the percentage of blood in 1 unit of body volume. We also assumed that all lymphocytes are in the blood and lymphocytes are distributed uniformly throughout the irradiated volume. The average

percentage of blood in the human body is 7% of body weight/volume [157], so the number of lymphocytes in 1 μL of body volume before treatment can be calculated as $L_0 = ALC_0 \times 0.07\ (cells \times 1000/\mu L)$. Therefore, the total number of lymphocytes in the whole-body volume before radiation was

$$N_b = (ALC_0 \times 0.07) \times (body\ volume) = L_0 \times (\nu \times N_\nu)\,, \tag{4.2}$$

where ν is the volume of each voxel and N_ν is the total number of voxels in the body for each patient.

The total number of remaining lymphocytes in the body volume after 1 fraction was calculated by summing the number of surviving lymphocytes in all voxels as follows

$$N_1 = \sum_{i=1}^{N_\nu} S_i(d)\ \nu\ L_0. \tag{4.3}$$

To find the ALC value after treatment, the total probability of lymphocyte survival after delivering k fractions (TP_k) was calculated using the ratio of the remaining lymphocytes to the initial number of lymphocytes in the body as follows

$$TP_k = \frac{\sum_{i=1}^{N_\nu} S_i^k\ \nu\ L_0}{\nu\ L_0\ N_\nu} = \frac{\sum_{i=1}^{N_\nu} S_i^k}{N_\nu}. \tag{4.4}$$

Thus, the expected final value of ALC after k fractions ($k \geq 1$) can be estimated using the following equation

$$ALC_k = TP_k \times ALC_0\ (cells/\mu L). \tag{4.5}$$

In this planning study, the total number of fractions was assumed to be 28 for all patients, i.e., $k = 28$. Note that the weekly ALC measurements available were for 6, 5, or 4

weeks for the 10 patients and we assume the measurements were taken in the end of each week of 5 fractions. Meanwhile, the predictions of ALC use 28 fractions for all patients.

4.2.3 ALC Prediction Using Exponential Curve Fitting

For all 10 patients undergoing PSPT for esophageal cancer, ALC counts during PSPT exhibited exponential decay as treatment proceeded. Thus, alternatively, we modeled patient ALC using an exponential function of accumulated delivered dose. We used the weekly ALC data points to fit an exponential function based on the total delivered dose to the body (i.e., the sum of doses in all voxels in treatment field) after k fractions ($D_k > 0$) as

$$ALC(D_k) = a.\ exp(-b.\ D_k) + c, \tag{4.6}$$

where a is a fixed parameter indicating the initial ALC before starting the treatment ALC_0, b is an index of an individual patient's lymphocytes' sensitivity to dose, and c is added to the exponential function to account for the replenishment of lymphocytes after irradiation. Note that previous studies have shown that ALC loss follows exponential decay in setting of total body radiation in primates [158] and in humans (accidental exposure) [159]. An exponential fitting method was also used to study ALC loss in partial body radiation therapy [84].

Using the weekly ALC data points and delivered dose in each week, we found the fitted values of the b and c parameters. Then, we used the same fitted function to predict ALC during treatment for PSPT, IMRT, and IMPT plans. The goodness of fit was tested using PSPT data.

In addition, we used the first three weeks' data to fit the exponential ALC function of dose to predict the final ALC after the entire treatment course. The rationale for this

approach was to determine whether patient-specific factors, including lymphocyte radiation sensitivity, derived from initial treatment fractions, is predictive of loss of lymphocytes by the end of treatment.

4.3 Numerical Experiments and Results

4.3.1 Model Validation

Fifteen esophageal cancer patients, treated with IMRT, PSPT, and IMPT (5 per modality) at MD Anderson Cancer Center with the same treatment prescription delivering 50.4 Gy in 28 fractions and identical normal tissue constraints, were selected to validate the models. Important metrics for measured values and predictions are summarized in 4.1. The average mean body doses were 14.44, 7.37, and 6.12 Gy for IMRT, PSPT, and IMPT treatments, respectively. The average ALC nadir for the 15 patients treated with IMRT, PSPT, and IMPT was 0.17, 0.33, and 0.39 $K/\mu L$, respectively. The average predicted ALC nadirs after treatment were 0.15, 0.32, and 0.37 $K/\mu L$ after IMRT, PSPT, and IMPT treatments using the piecewise-linear model, respectively, and 0.12, 0.30, and 0.36 $K/\mu L$ using the exponential model. The mean squared error was 0.005, 0.023, and 0.003 for IMRT, PSPT, and IMPT treatments based on the piecewise-linear model, and 0.005, 0.005, and 0.004 based on the exponential model.

Table 4.1: Mean body dose, ALC baseline, real and predicted ALC nadirs, and associated errors for patients treated with IMRT, PSPT, and IMPT. Unit for ALC values is $K/\mu L$. Values for ALC are presented as mean±SD deviation.

RT modality	Mean Body Dose	ALC_0	Real ALC	Piecewise-linear model			Exponential model		
				Predicted ALC	MSE	MAE	Predicted ALC	MSE	MAE
IMRT	14.44	1.42 ± 0.46	0.17 ± 0.10	0.15 ± 0.09	0.005	0.064	0.12 ± 0.06	0.005	0.053
PSPT	7.37	1.41 ± 0.57	0.33 ± 0.18	0.32 ± 0.15	0.023	0.104	0.30 ± 0.16	0.005	0.057
IMPT	6.12	1.55 ± 0.57	0.39 ± 0.26	0.37 ± 0.24	0.003	0.040	0.36 ± 0.24	0.004	0.058

For each group of patients, ΔALC (baseline $-$ nadir) using the piecewise-linear and

the exponential models were also calculated. The real values and predictions are summarized in Table 4.2. The average predicted ΔALC for IMRT patients were 1.27 $K/\mu L$ and 1.30 $K/\mu L$ using piecewise-linear and exponential model, respectively, which were in agreement with the measured ΔALC of 1.25 $K/\mu L$ with a mean squared error of 0.005 for both models and mean absolute error of 0.064 and 0.053 for piecewise-linear and exponential model, respectively. Similarly, for PSPT patients, the average ΔALC were 1.08 $K/\mu L$, 1.09 $K/\mu L$, and 1.11 $K/\mu L$ for the measured, piecewise-linear model predictions and exponential model predictions, respectively. The estimated ΔALC and ALC nadirs had a mean squared error of 0.023 and 0.005, and a mean absolute error of 0.104 and 0.057 for piecewise-linear and exponential model, respectively. Finally, the average predicted ΔALC for IMPT patients were 0.99 $K/\mu L$ and 0.99 $K/\mu L$ using piecewise-linear and exponential model, respectively, compared with the measured value of 0.97 $K/\mu L$. The mean squared error and mean absolute error of estimated values were 0.003 and 0.04 for the piecewise-linear model, and 0.004 and 0.058 for the exponential model. These results reassure that these two simple models of lymphocyte survival can provide reasonably good predictions for different radiation modalities.

Table 4.2: Real ΔALC, predicted values and associated errors for patients treated with IMRT, PSPT, and IMPT. Unit for ALC values is $cells \times 1000/\mu L$. Values for ALC are presented as mean±SD deviation.

RT modality	**Real ΔALC**	**Piecewise-linear model**			**Exponential model**		
		Predicted ΔALC	**MSE**	**MAE**	**Predicted ΔALC**	**MSE**	**MAE**
IMRT	1.25±0.44	1.27±0.43	0.005	0.064	1.30 ± 0.46	0.005	0.053
PSPT	1.08±0.52	1.09±0.52	0.023	0.104	1.11 ± 0.54	0.005	0.057
IMPT	0.97±0.58	0.98±0.60	0.003	0.040	0.99±0.59	0.004	0.058

4.3.2 Dosimetric Characteristics of IMRT, PSPT, and IMPT Plans

We first evaluated the dose distributions of the PSPT plans employed to treat the 10 patients and the IMRT and IMPT plans generated for this comparison study in terms of dose-volume metrics (e.g., mean body dose, V_5, V_{10}, etc.). An example of the dose distributions of IMRT, PSPT, and IMPT plans for a patient (Patient 5) can be found in Figure 4.2. As expected, the radiation dose using the IMPT plan conformed more closely to the PTV than did the PSPT and IMRT plans, and the IMRT plan delivered the highest doses (and largest dose baths) to the body.

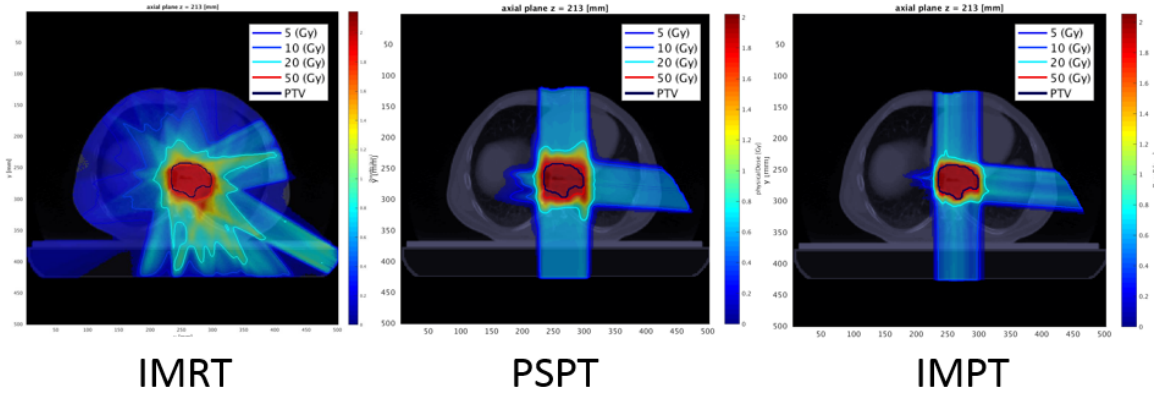


Figure 4.2: Dose distributions on an axial plane of IMRT, PSPT, and IMPT plans on an axial plane for Patient 5

The mean body doses (MBD), averaged among the 10 patients, were 7.46 Gy, 4.84 Gy, and 3.85 Gy for IMRT, PSPT, and IMPT plans, respectively. The fractions of the body volume that received different doses were consistently higher for photon therapy (IMRT) plans than for proton therapy (IMPT and PSPT) plans, especially at low doses such as 5 Gy or 10 Gy. For each dose-volume index, IMPT plans outperformed PSPT plans. More detailed comparison of body dose-volume metrics for the 10 patients among the 3 modalities can be found in Figure 4.3.

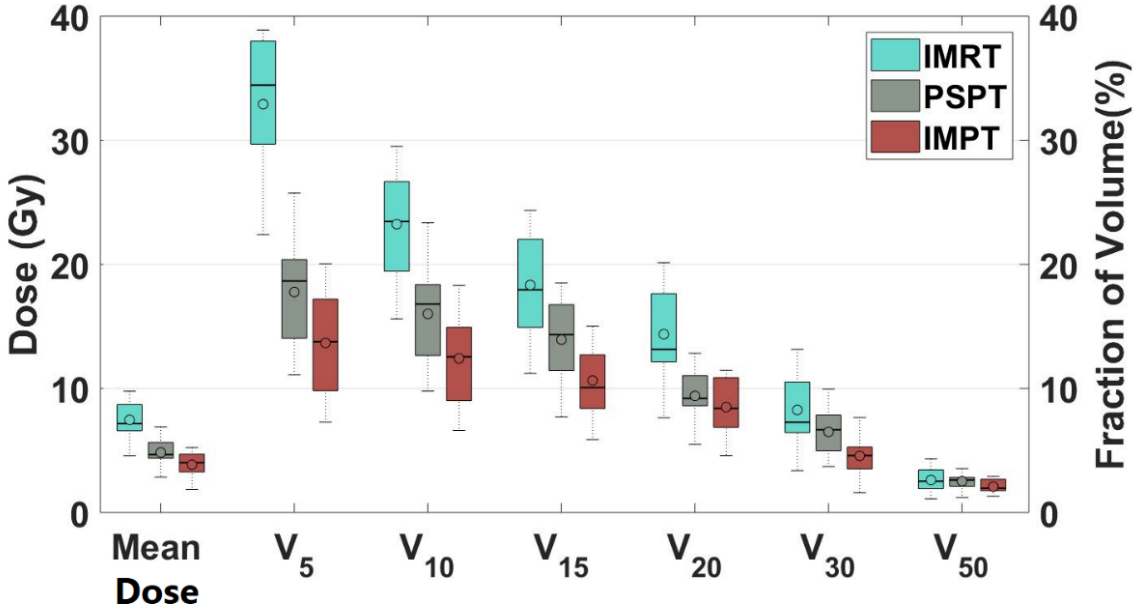


Figure 4.3: Box plots illustrating different dose-volume indices for the total irradiated volume for three treatment plans (IMRT, PSPT, and IMPT) in 10 patients. V_x is the fraction of volume receiving more than x Gy dose.

4.3.3 Lymphocyte Survival Based on Piecewise-linear Function of Dose

Using the piecewise-linear lymphocyte survival function, we estimated ALCs as a function of delivered radiation dose to each voxel in the patient body for IMRT, PSPT, and IMPT plans for each patient. Note that several studies have shown that patients who receive concurrent chemotherapy exhibit greater treatment-related lymphocyte depletion than patients who receive radiotherapy alone [66, 77, 78]. Also, chemotherapy type is a significant factor in determining the level of ALC nadir (the minimum ALC value over the course of or after radiation therapy), along with radiation modality and mean body dose [66]. As all of our patients had undergone chemotherapy during the course of radiation therapy, we added a factor for "chemo effect" to correct the predicted ALC in this linear dose model. The chemo effect factor was determined by the average of the differences between measured and predicted ALCs for PSPT data and was applied to all final predictions. It was

calculated as 26% (SD=0.08) in this study and then applied to the final prediction of ALC.

The average predicted ALCs after treatment were 0.27 $K/\mu L$ (95% CI [0.21, 0.33]), 0.35 $K/\mu L$ (95% CI [0.27, 0.42]), and 0.37 $K/\mu L$ (95% CI [0.29, 0.44]) for IMRT, PSPT, and IMPT plans, respectively. Figure 4.4(a) shows box plots of predicted ALCs at the end of treatment courses for the 3 treatment modalities. Figure 4.4(b) shows box plots of predicted ALC changes before and after treatment ($\Delta ALC = baseline - nadir$). Proton plans showed smaller ALC reductions than did photon plans, and the differences between PSPT and IMPT were relatively small.

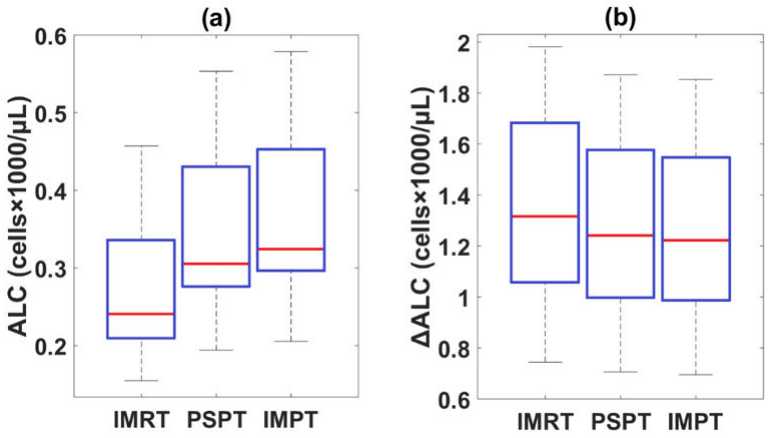


Figure 4.4: Box plots of the piecewise-linear model prediction results for (a) final ALCs, (b) ALC changes after the three treatment modalities for 10 esophageal cancer patients

The predicted ALC changes for the 10 patients based on the 3 plans are shown in Figure 4.5(a). ΔALC values for IMRT plans were higher than those of IMPT and PSPT plans for all 10 patients. This supports the hypothesis that proton modalities cause less ALC depletion than do photons. Actual ALC changes from measured ALC data for PSPT treatments are indicated by black diamonds. We observed that the ALC estimates for PSPT plans were relatively close to the real measured data, with a mean absolute error of 0.075 and a mean squared error of 0.010.

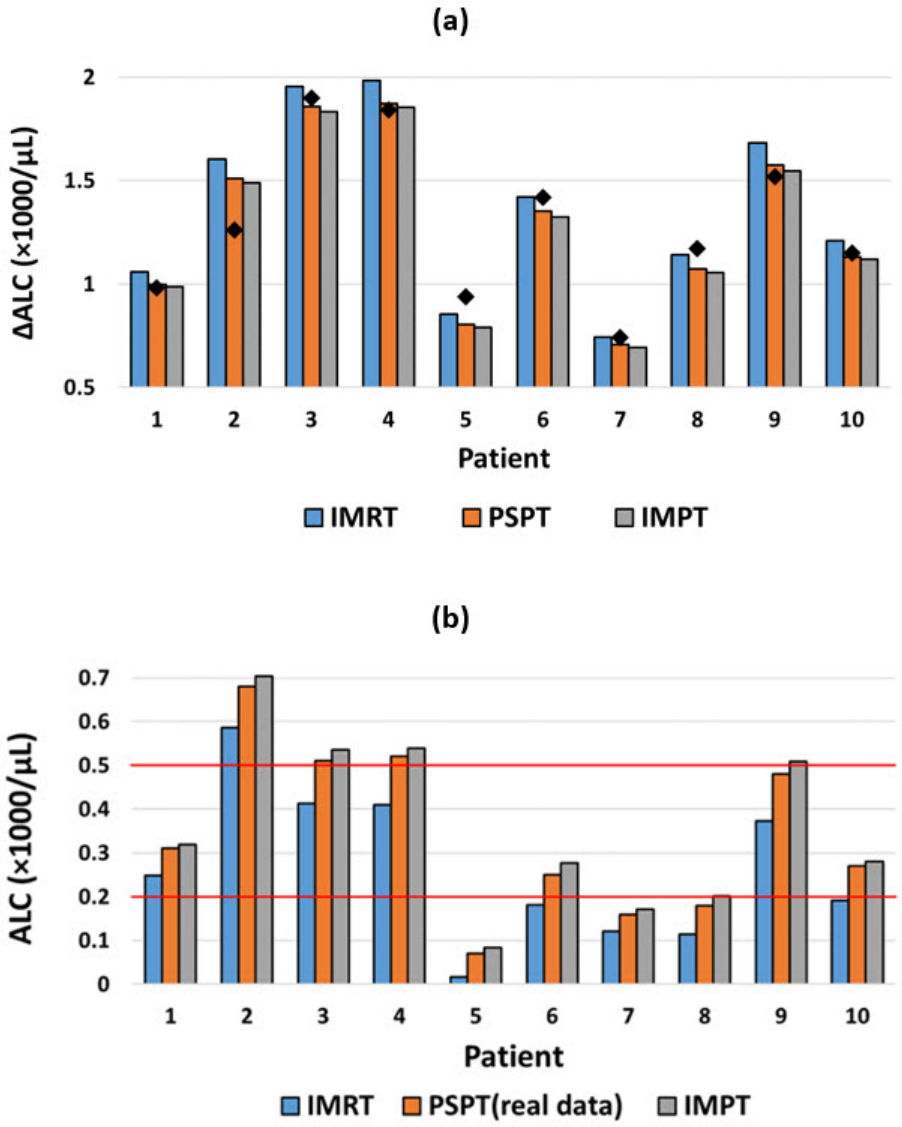


Figure 4.5: (a) Comparison of predicted ΔALC using the piecewise-linear method for IMRT, PSPT, and IMPT plans for 10 esophageal cancer patients. (b) Predicted ALC nadirs for IMRT and IMPT treatments using the piecewise-linear method versus the measured ALC nadir for PSPT plans.

Grade 4 lymphopenia (G4L) and grade 3 lymphopenia (G3L), according to the National Cancer Institute's Common Toxicity Criteria for Adverse Events v5.0, are defined as $ALC < 200\ cells/\mu L$ and $ALC < 500\ cells/\mu L$, respectively. Recent studies reported a lower risk of grade 4 lymphopenia for patients treated with proton therapy compared with patients treated with photon therapy [66, 78, 79]. Figure 4.5(b) shows the predicted ALC

nadirs for IMRT and IMPT treatments, which were calculated by subtracting the predicted ALC change from the measured ALC baseline, for each of the 10 patients. In other words, this figure demonstrates the predicted ALC nadir if each patient had been treated by IMRT or IMPT instead of PSPT. Grade 4 lymphopenia occurred in Patients 5, 7, and 8 after PSPT treatment. Based on the predicted ALC nadirs, grade 4 lymphopenia might have been avoided for Patient 8 if she or he had been treated with IMPT instead of PSPT. Patients 6 and 10 had grade 3 lymphopenia after PSPT; however, they may have developed grade 4 lymphopenia if they had been treated with IMRT. Patients 3 and 4 were predicted to have developed grade 3 lymphopenia if IMRT had been used instead of PSPT.

4.3.4 Lymphocyte Survival Based on Exponential Fitting

Weekly measurements of ALC and the total delivered dose for PSPT were used to fit an exponential curve for each patient. Figure 4.6 (a) shows box plots of predicted posttreatment ALC values for all patients using this approach for all 3 modalities, and Figure 4.6 (b) shows box plots of predicted ALC changes before and after treatment (ΔALC). The predicted final ALC was the lowest for IMRT; IMPT was estimated to result in a higher ALC than PSPT. These results agree with those of the piecewise-linear function approach. The fitted exponential curve and data points used for fitting for all patients are shown in Figure 4.7. The average estimated ALCs after treatment, calculated using exponential fitting, were 0.14 $K/\mu L$ (95% CI [0.08, 0.19]), 0.22 $K/\mu L$ (95% CI [0.14, 0.30]), and 0.33 $K/\mu L$ (95% CI [0.19, 0.45]) for IMRT, PSPT, and IMPT plans, respectively, which shows the same trend as the previous approach.

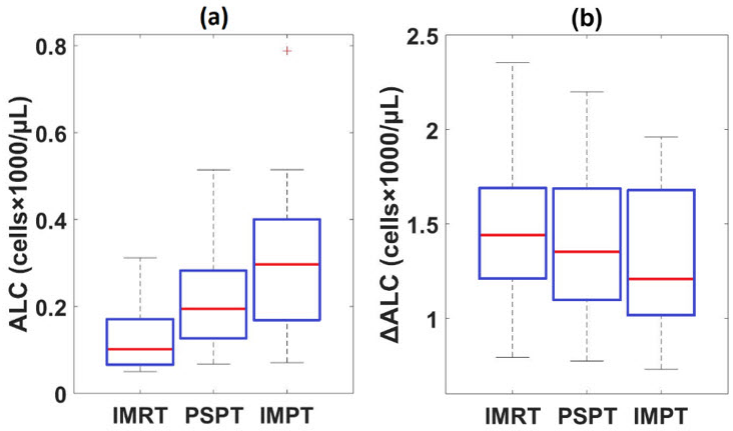


Figure 4.6: Box plots of the exponential fitting method prediction results for (a) final ALCs, (b) ALC changes after the three treatment modalities for 10 esophageal cancer patients

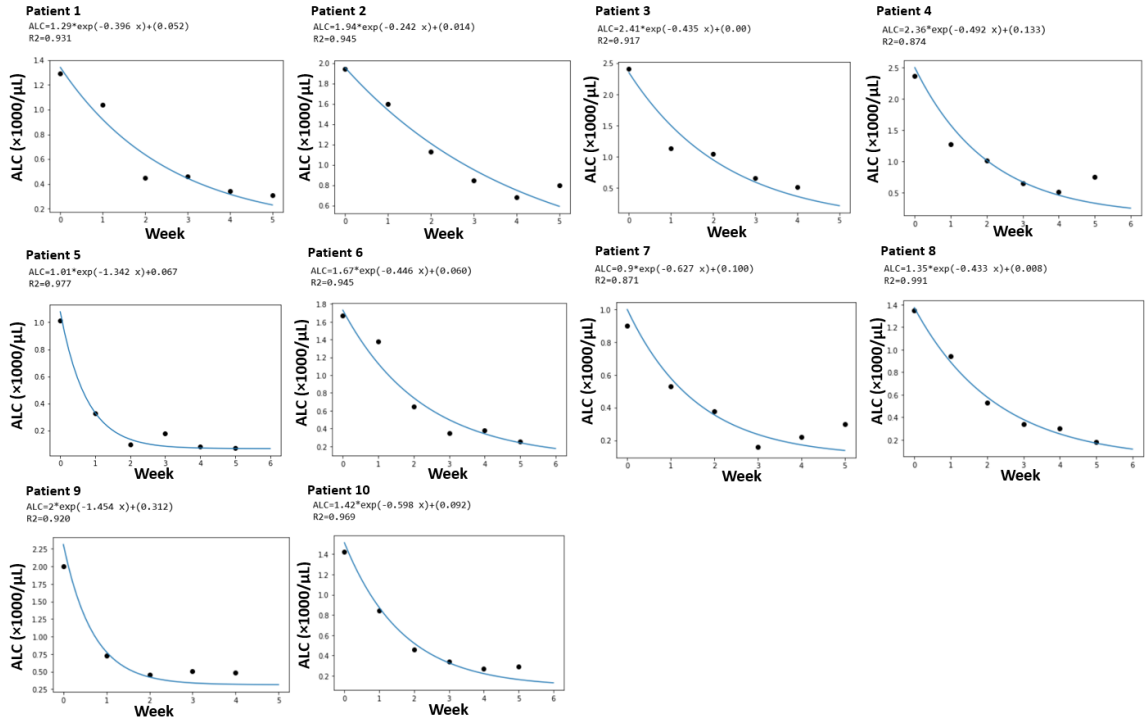


Figure 4.7: The exponential curves fitted with measured weekly ALC data for 10 esophageal cancer patients treated with PSPT.

A comparison of the predicted ALC change using the exponential model is shown in Figure 4.8(a). The ΔALC values for IMRT were higher than those for IMPT and PSPT for all 10 patients. Meanwhile, the estimated ALCs had a mean absolute error of 0.125 and a

mean squared error of 0.023, higher than for the piece-wise linear method. Figure 4.8(b) shows the predicted ALC nadirs for IMRT and IMPT treatments, similar to Figure 4.6(b). For example, Patients 7 and 8 may have avoided grade 4 lymphopenia if treated with IMPT, and Patients 1 and 6 had grade 3 lymphopenia with PSPT, but might have had grade 4 if treated with IMRT.

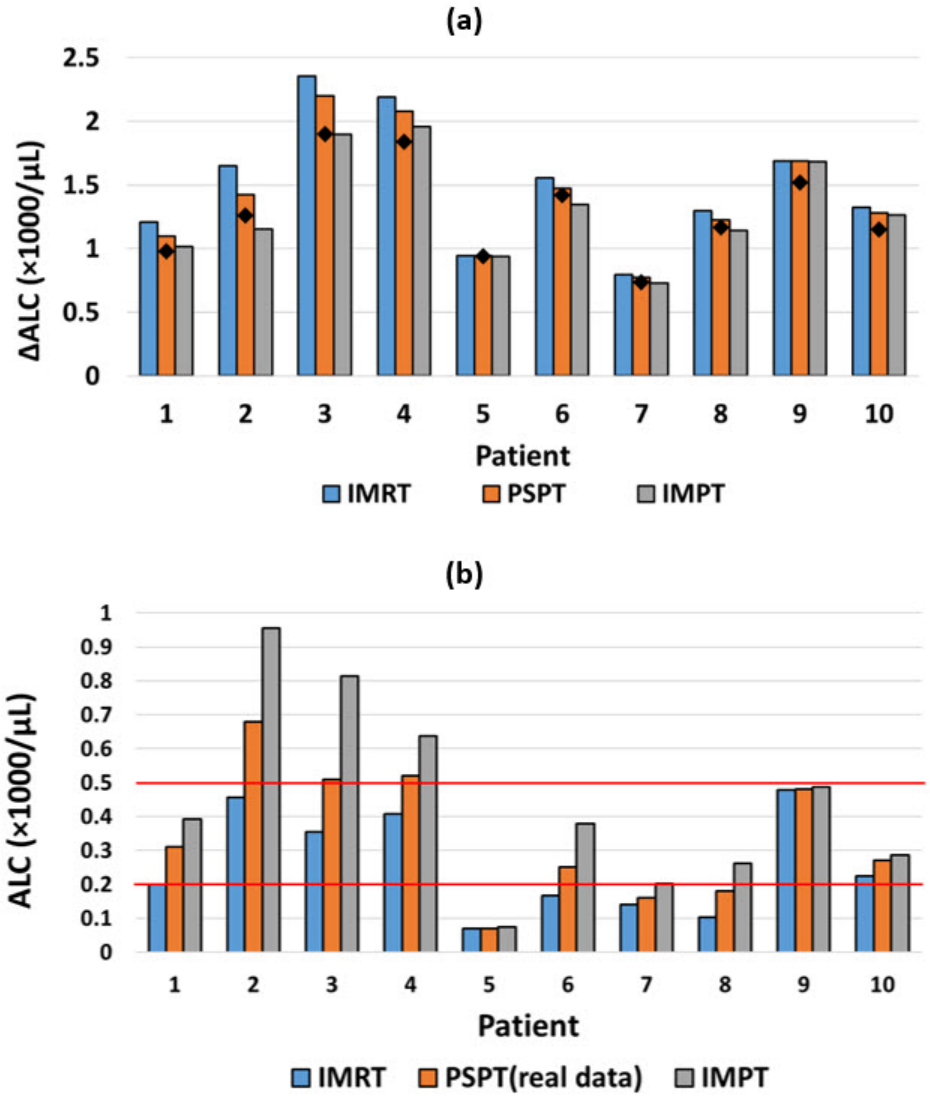


Figure 4.8: (a) Comparison of predicted ΔALC using the exponential fitting method for IMRT, PSPT, and IMPT plans for 10 esophageal cancer patients. (b) Predicted ALC nadirs for IMRT and IMPT treatments using the exponential fitting method versus the measured ALC nadir for PSPT plans.

To predict lymphopenia in the early stages of treatment, we used the ALC-dose data for the first three weeks to estimate the exponential ALC function. The fitted exponential curves using 3-week data for all patients are shown in Figure 4.9. Figure 4.10(a) illustrates the measured ALC nadirs and the estimated final ALC based on 3-week and all-week data after PSPT treatment. ALC predictions based on 3-week data were lower than ALC predictions using all weekly data. Figure 4.10(b) shows an R-squared comparison between 2 exponential fittings, which was higher for the first exponential fitting for all 10 patients. The average R-squared values for exponential fitting using all weekly data and 3-week data were 93.5% and 90.0%, respectively.

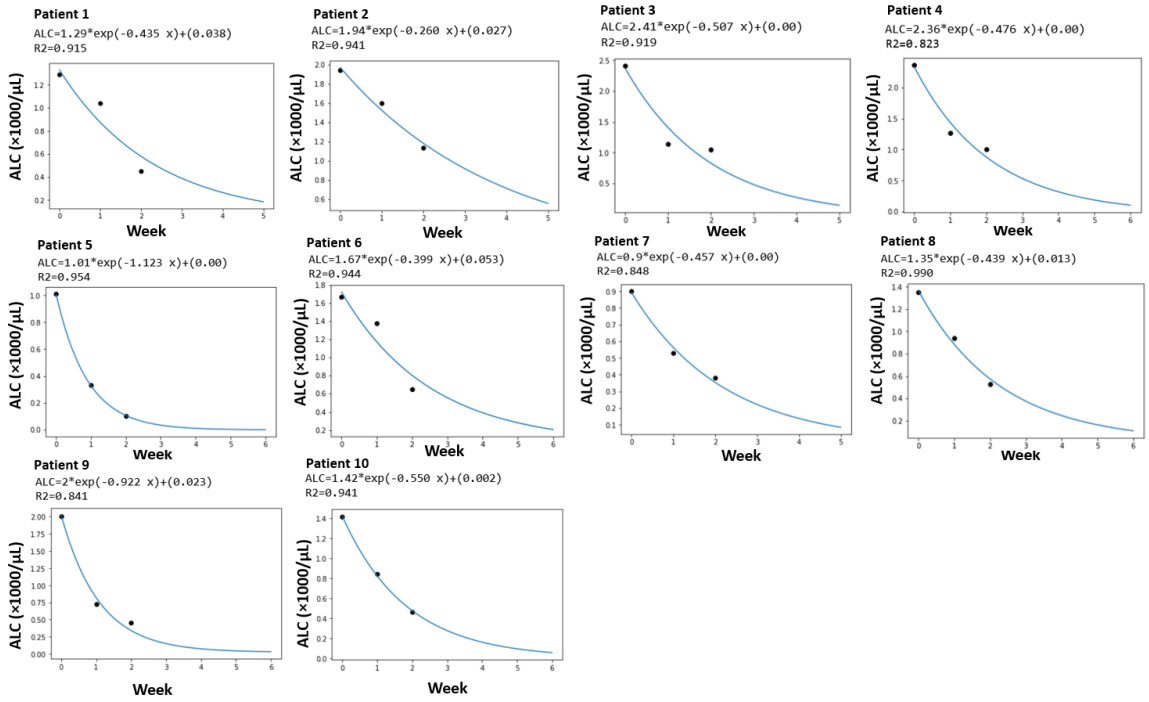


Figure 4.9: The fitted exponential curve based on the measured ALC of the first 3 weeks for 10 esophageal cancer patients treated with PSPT

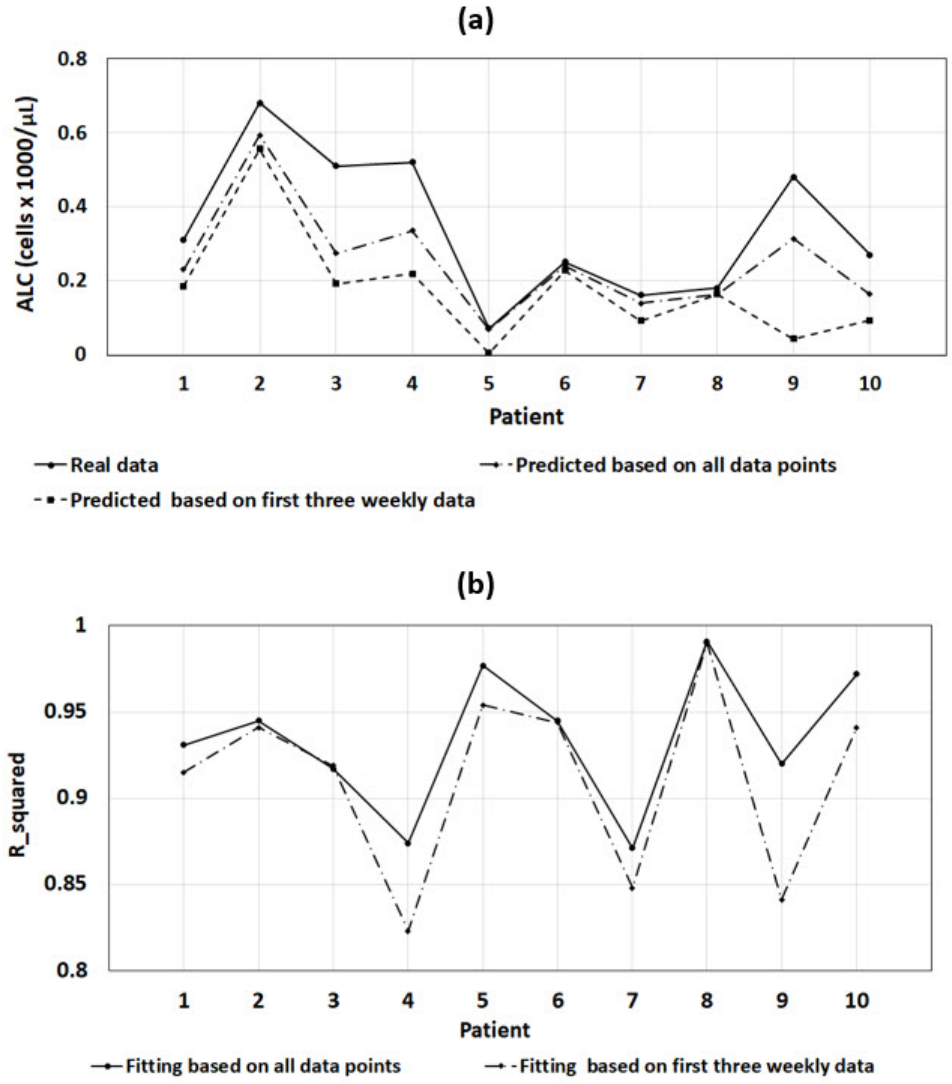


Figure 4.10: (a) Measured ALC nadirs and estimated posttreatment ALC based on the fitted exponential model using the first three weeks data and all weekly data from PSPT treatments for 10 patients. (b) R-squared comparison for two exponential fittings.

4.4 Discussion

The present planning study indicates that IMPT treatment might lead to less lymphocyte depletion than PSPT and that IMRT may produce the most lymphocyte depletion. It is worth noting that this study focuses on comparing RIL risks from different treatments for

the same individual patient, rather than associating RIL with treatments for prospective patients. Studies like this one can generate hypotheses for clinical trials for investigating the RIL risks entailed by different radiation modalities. Further studies based on clinical data of photon therapy (e.g., IMRT, VMAT) are required to support the hypothesis that proton therapy outperforms Photon therapy in terms of reducing RIL risk.

The piecewise-linear model of lymphocyte survival based on fractional voxel dose was able to predict the trend of ALC changes during treatment according to the PSPT data. While this model is simplistic, it could be a straightforward method to compute a population-based estimate of ALCs before the start of treatment once the dose distribution is known. However, one limitation of the linear methods is the assumption that lymphocyte distribution throughout the treatment field is homogeneous. This is not be the case where organs, for example, spleen in the upper abdomen, may have concentration of lymphocytes, and may have a greater influence on lymphocyte depletion than the body as a whole.

The exponential lymphocyte survival model of total body dose, built from fitting measured ALC data, was also able to predict ALC depletion during treatment. However, it was less accurate than the linear method for predicting the ALC nadir. One possible cause may be uncertainty in ALC measurement. For example, erroneous increases in ALC were seen for Patients 2, 4 and 7 (see Figure 4.7). In order to mitigate the sensitivity of this model to uncertain fitting data (i.e., ALC measurements) and, more importantly, improve the predictive power, we will investigate approaches to incorporate pre-clinical predictors into the model in our future work, such as patient age and BMI [85]. Figure 4.11 shows the comparison of measured and predicted ALC nadirs based on the linear and exponential methods for all patients. The predicted ALC nadirs are mostly lower than the measured ones because the predictions were calculated for 28 fractions but the measured nadirs only were from 4 to 6 weeks among patients. In addition, the impact of increased dose (or more fractions) on ALC depletion appears to be higher for the exponential method than the linear

method, for example, Patient 3 as an extreme case where ALC nadir was measured in week 4 (i.e., after 20 fractions).

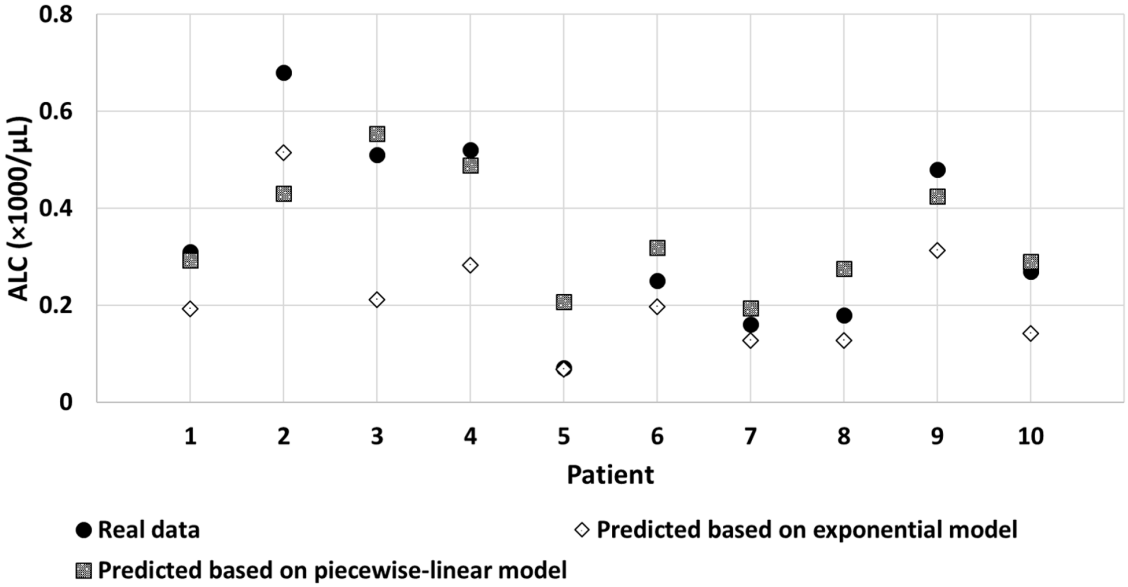


Figure 4.11: Measured ALC nadirs and the estimated posttreatment ALC based on fitted exponential model and piecewise-linear model.

Although it is not clear that the exponential fitting approach will be of benefit in predicting ALC for new patients, as it requires large data sets, this approach is useful in comparative studies of different dose patterns for individual patients, such as in the present work. In addition, by only fitting the ALC data in initial weeks of treatment (of the first few fractions), one could estimate of ALC nadir early in treatment course based on the dose distribution and the consideration of individual patients' lymphocyte sensitivity and make mid-course correction with adaptive plans if and when needed.

This study also motivates further studies to investigate the clinical factors that affect RIL risk of different radiation modalities. With help of research on continuing better understanding of lymphocyte distribution throughout the treatment field, radiation dose could be optimized accordingly to avoid lymphocyte killing. For example, IMPT and IMRT plans in this study were optimized using the same conventional dosimetric criteria as the PSPT

plans. Additional immune sparing could be possible by optimizing plans with constraints on dose received by volumes of the body (and immune organs at risk such as the spleen, heart, etc.), which is most promising for IMPT due to its high complexity and flexibility in modulation. Such methods to enhance the ability of IMPT to minimize lymphopenia risk but without compromising tumor coverage and other normal tissues at risk will be studied in our future work.

4.5 Conclusion

This treatment planning study assessed RIL risk and the impact of different dose distributions of IMRT, PSPT, and IMPT on ALCs for 10 esophageal cancer patients. Two methods are proposed to estimate posttreatment ALC. Results from both approaches showed significant lymphocyte reduction associated with treatment. Proton plans showed a lower risk of lymphopenia after the treatment course than did photon plans, and IMPT plans outperformed PSPT plans in terms of lymphocyte preservation.

Chapter 5

A Hybrid Deep Learning Model for Forecasting Lymphocyte Depletion During Radiation Therapy

5.1 Introduction

Radiation therapy (RT) is an effective treatment option for many cancer patients. An RT patient undergoes a series of treatment sessions over several weeks to deliver a prescribed dose of radiation to the tumor. The clinical goal of RT is to maximize the radiation-induced damage to the tumor, killing all cancerous cells, while minimizing toxic effects on surrounding healthy tissues [160].

Recent studies have shown that the absolute lymphocyte count (ALC) is very sensitive to radiation exposure; by killing the circulating lymphocytes in the radiation field, RT suppresses the immune system [68], [70]. The resulting reduction in ALC causes radiation-induced lymphopenia (RIL), a common toxic effect of RT [69], [71]. Clinical studies have shown that severe lymphopenia can reduce the survival of patients with a number of solid tumors, including esophageal cancer [78], [66]. Severe RIL can be diagnosed by measuring the ALC in the patient's blood. Grade 4 RIL, according to the Common Toxicity Criteria for

Adverse Events, version 5.0, is defined as an ALC less than 0.2 $cells \times 1000/\mu L$ during and immediately following the course of RT. The ability to reliably predict radiation-induced ALC depletion on the basis of pretreatment factors (i.e., dosimetric factors, treatment factors, and patient-specific factors) would improve RT planning. Specifically, predicting the risk of lymphocyte depletion during early RT fractions could identify patients who are at high risk of severe lymphopenia (i.e., grade 4 RIL) and who may stand to benefit from RIL mitigation strategies and modified treatments that may ultimately improve their survival [84, 85].

Several studies have shown strong associations between pretreatment factors and the risk of severe RIL in various cancers [66, 71, 78, 84–86, 161]. Some of these studies have attempted to predict RIL based on different set of pretreatment parameters which are summarized in Table 5.1. van Rossum et al. [85] showed the significance of age, planning target volume, body mass index, radiation modality, and baseline ALC in relation to grade 4 RIL for esophageal cancer patients and developed a pretreatment clinical nomogram based on these factors to determine the risk of grade 4 RIL for new patients. Zhu et al. [86] developed a hybrid deep learning model to classify patients with grade 4 RIL based on patient characteristics and dosimetric features but they did not investigate the ALC kinetics during RT treatment for individual patients. Ebrahimi et al. [161] performed a posttreatment analysis based on weekly ALC measurements for ten esophageal cancer patients and showed that the ALC depletion during the course of RT can be fitted to a piecewise-linear or an exponential model as a function of radiation dose for ten esophageal cancer patients. However, they did not consider other significant patient-specific clinical factors in their models. Therefore, comprehensive models that can forecast the kinetics of lymphocyte loss after fractionated radiation exposures in order to identify high-risk patients are lacking. It is critical to fill this gap and provide a comprehensive prediction model that can forecast the

ALC regressions during RT to develop RIL mitigation strategies at the right time and improve the effectiveness of RT for cancer patients. In this study, we aimed to predict weekly radiation-induced lymphocyte depletion in esophageal cancer patients during the course of RT on the basis of significant pretreatment or early-treatment information.

Table 5.1: Summary of recent studies on RIL prediction

		van Rossum et al. [85]	Zhu et al. [86]	This Paper
Considered Variables	Dosimetric Features		✓	✓
	Patient Clinical Characteristics	✓[a]	✓	✓
	ALC Measurements			✓
Method	RIL Patient Classification		✓	
	RIL Risk Prediction	✓		
	Forecasting weekly ALC values			✓

[a] Among all clinical features, only age, planning target volume, body mass index, radiation modality, and baseline ALC were considered.

The application of artificial intelligence and machine learning methods to extract insights from data is becoming increasingly attractive in many fields, including healthcare. Although many healthcare applications have been developed, those that can predict disease progression [87], [88], treatment outcomes [89], or potential side effects [90], [91] play an important role in improving patients' care. Deep learning models have been developed to extract information from various kinds of data and for many tasks [92], [93]. Recurrent neural networks (RNNs) achieved significant results in extracting temporal information from sequential data such as text, audio, video, and time series [92]. The main advantage of RNNs is that they can maintain memory of recent events and update their current state based on both past states and current input data [94], [95]. Hochreiter and Schmidhuber [96] proposed the long short-term memory (LSTM) network as an improved variant of the RNN to handle the long-term dependency and vanishing gradient issues of RNNs. LSTM networks have been widely used for various kinds of tasks, including speech recognition [97], [98], image captioning [99], [100], trajectory prediction [101], [102], and text embedding [103], [104]. However, an LSTM network cannot be used alone for the

current problem because the significant features that may predict RIL do not have uniform characteristics. A potential solution for this issue is to develop a stacked structure.

In a stacked structure, the algorithm nonlinearly integrates predictors in order to achieve higher prediction accuracy and reduce generalization error. Deep-stacked models can outperform state-of-the-art deep learning and machine learning models such as tree-based ensemble models and extreme gradient boosting algorithms [105], [106]. Therefore, we propose a hybrid deep-stacked model that combines a deep neural network with an LSTM network for different groups of features in a stacked structure. The proposed structure consists of 4 channels to process 4 categories of features with different characteristics; LSTM is used to process the sequential features. We established 3 models to predict an ALC depletion trend on the basis of pretreatment, first-week, and second-week treatment information. To evaluate the performance of our proposed hybrid deep-stacked model, we calculated well-known prediction metrics and compared the results with other common prediction methods.

In summary, the contribution of this paper is:

- A hybrid deep-stacked structure based on pretreatment information is proposed to predict RIL for new esophageal cancer patients during the course of RT.

- The proposed hybrid deep-stacked structure can use information from different groups of features with different characteristics to predict weekly ALC without requiring a large amount of data to process too many features at the same time, while reducing bias and the adverse effects of any noise in the data.

- The developed model is flexible, interpretable, and can be extended easily to account for early-treatment predictions (i.e., at the end of week 1 or 2), and a discriminative kernel layer was developed to distinguish the importance of each value in the input sequence at different times by assigning different weights to each one.

- The ability to predict ALC depletion trend during the course of RT based on pretreatment clinical information would enable physician to evaluate individual RT treatment plans for lymphopenia risk and identify patients at high risk who would benefit from modified treatment approaches.

The rest of this paper is organized as follows: section 2 covers the data description, data preprocessing, and our proposed hybrid deep-stacked model; section 3 presents the experimental results and discusses the key findings of our study; and section 4 concludes and provides directions for further research.

5.2 Materials and Methods

5.2.1 Data description

This study was approved by the University of Texas MD Anderson Cancer Center institutional review board. All methods performed here were in accordance with the Health Insurance Portability and Accountability Act. Data from 860 patients who received concurrent chemoradiotherapy (with or without surgery) for biopsy-proven esophageal cancer between January 2004 and November 2017 at MD Anderson Cancer Center were used for this study. All patients were treated with proton or photon radiation modalities with a total radiation dose of 50.4 Gy over five weeks. All included patients also had available baseline ALC values and 3 or more documented weekly ALC values during treatment.

5.2.2 Variable Selection

The variables of interest for the prediction were ALC after each week of treatment. Predictor variables were selected on the basis of their clinical relevance, their low level of

missingness (<20%), and the results of the correlation analyses presented in previous studies on this data set by Zhu et al. [86] and van Rossum et al. [85]. As a result, 53 features were selected as predictors and categorized into 4 main groups on the basis of their clinical and analytical characteristics: (1) dosimetric features contained 30 dose-volume metrics such as $V_5,\ V_{10},\ \ldots\ , V_{45}$ (V_x refers to the percentage of the organ volume that received at least x Gy radiation dose) and mean dose for 3 organs at risk: the lung, heart, and spleen; (2) other numerical treatment-related and patient-specific parameters were nondosimetric numerical features including age, body mass index (BMI), total blood volume, planning target volume, and blood component profiles at baseline (red blood cells, white blood cells, and others); (3) nondosimetric categorical features such as RT modality (proton or photon), race, sex, tumor location, tumor histologic characteristics, and use of induction chemotherapy; (4) sequential features included the sequence of 5 weekly ALC values. The features groups and variables description are also summarized in Table 5.2.

Zhu et al. [86] reported high collinearity between dosimetric features based on the high variance inflation factor of each dosimetric predictor. This is because of the sequential and highly intercorrelated nature of dose-volume histograms. Dimensionality reduction techniques can be used to address the collinearity problem by combining highly correlated variables into a set of uncorrelated variables [162, 163]. The literature showed that t-distributed stochastic neighbor embedding (t-SNE) is highly capable of non-linear embedding and retaining the main information while preserving the non-linear variance and local structure of data [164–166]. Moreover, we tested the performance of t-SNE over other common linear and non-linear dimensionality reduction methods such as principal component analysis (PCA), linear discriminative analysis (LDA), Isomap, locally linear embedding (LLE), uniform manifold approximation and Projection (UMAP) in a prediction model based on the dosimetric features. t-SNE outperformed rest of these methods in terms of features

Table 5.2: Features groups and variables description

Features Groups	Variables
Dosimetric Features	V_x^{Heart}, V_x^{Lung}, V_x^{Spleen} (%); $V_x \in \{5, 10, 15, 20, 25, 30, 35, 40, 45\}$ Mean heart dose, mean lung dose, mean spleen dose (Gy) [b]
Nondosimetric Numerical Features	Age Body mass index ($kg\ m^{-2}$) Total blood volume (l) Planning target volume (cm^3) Blood component profiles at baseline: RBC ($10^7 \mu l^{-1}$), HB (g/dl) , HT (%), WBC ($10^9\ l^{-1}$), ANC ($10^3\ \mu l^{-1}$), PLC ($10^3\ \mu l^{-1}$), MC ($10^3\ \mu l^{-1}$) [c]
Nondosimetric Categorical Features	RT modality −Proton beam −Photon beam Race −white −Other Sex −Female −Male Tumor location −Upper and middle of esophagus −Distal Tumor histologic characteristics −Adenocarcinoma −Squamous cell carcinoma Induction of chemotherapy −Yes −No
Sequential Features	ALC baseline ($10^3\ \mu l^{-1}$) ALC values after each week of RT treatment: $ALC_x(10^3\ \mu l^{-1});\ \ \forall x \in \{1,\ 2,\ 3,\ 4,\ 5\}$

[a] V_x Refers to the percentage of the volume that received at least x Gy radiation dose

[b] Average radiation dose delivered to heart, lung, and spleen structures

[c] Abbreviations: RBC, red blood cell count; HB, hemoglobin level; HT, hematocrit level; WBC, white blood cell count; ANC, absolute neutrophil count; PLC, platelet count; MC, monocyte count; ALC, absolute lymphocyte count.

variance in the reduced dimension and weekly ALC prediction error based on the dimensionality reduced dosimetric features. Therefore, we used t-SNE dimensionality reduction

method to reduce the effect of this severe collinearity among dosimetric features without losing main information. t-SNE is a nonconvex and nonlinear method to reduce the dimensionality of data by considering the similarity between features that follow a conditional exponential probability distribution as

$$P(j|i) = \frac{\exp \frac{\|x_i - x_j\|^2}{2\sigma_i^2}}{\sum_{i \neq j} \exp \frac{\|x_i - x_j\|^2}{2\sigma_i^2}} \ , \tag{5.1}$$

where $p(j|i)$ is the similarity between features x_i and x_j considering the original data features as $x_1,\ x_i,\ \ldots\ , x_N$. The total similarity between these 2 variables is the mean value of 2 conditional probabilities divided by the number of features ($p_{ij} = \frac{p(j|i)+p(i|j)}{2N}$). The new d dimensional data $y_1,\ y_i,\ \ldots\ , y_d$ must reflect the p_{ij} as much as possible. So, q_{ij} can be estimated as

$$q_{ij} = \frac{\left(1 + \|y_i - y_j\|^2\right)^{-1}}{\sum_{i \neq j} \left(1 + \|y_i - y_j\|^2\right)^{-1}} \ . \tag{5.2}$$

The values of new features can be calculated by minimizing the Kullback–Leibler (KL) divergence between the distributions of data before (P) and after (Q) dimensionality reduction as follows

$$\min\ KL(P|\,(Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}}. \tag{5.3}$$

In our proposed model, t-SNE was implemented with 3 components, an optimal perplexity value of 60, a learning rate of 10, and principal component analysis (PCA) initialization in a maximum of 5000 iterations.

5.2.3 Data Preprocessing

There were some missing values for some features in the data set; these were considered missing at random. To avoid losing information by removing the missing values, they were imputed. Missing values in nondosimetric numerical features and dosimetric features were imputed by median value, and missing values in sequential features were handled with multiple imputation. Multiple imputation with Bayesian ridge regression was used to impute missing ALC values for weeks 4 and 5 in 2 steps: first, we imputed 9 missing ALC values for week 4; then, we imputed 124 missing ALC values for week 5. The imputation did not change the data distribution and variability.

Due to the uncertainty caused by ALC measurement error, there were some patients with odd weekly ALC trends, which we considered as outliers. Since removing these outliers from the data set was not feasible because of the size of the data set, we used Holt's double exponential smoothing (DES) method to remove noise from the data. DES is a popular smoothing method for time series with trends. It assigns exponentially decreasing weights to observations as the observations get older [167]. The smoothing method helps to remove or reduce volatility or other types of noise and allows important patterns to stand out. The equations used to determine the smoothed values with this method are

$$F_{t+m} = S_t + mT_t \ , \tag{5.4}$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \ , \tag{5.5}$$

$$\text{and} \quad S_t = \alpha y_t + (1-\alpha)(S_{t-1} + T_{t-1}), \tag{5.6}$$

where y_t refers to the actual value at time t, F_{t+m}, T_t, and S_t are the forecast for the period $t+m$, the trend estimate, and the exponentially smoothed series, respectively. α denotes the process smoothing constant, and β refers to the trend smoothing constant

($-1 \leq \alpha,\ \beta \leq 1$). DES with a damped trend was done only for patients with unreasonable trends, which was defined by RT experts as having an increased ALC after weeks 2 or 3, or an increase in ALC value greater than 0.1 after weeks 4 or 5. Min-max normalization (i.e., $x = \frac{x - x_{min}}{x_{max} - x_{min}}$) was applied to all features, and the data were split into training and test sets in a ratio of 7:3 (602:258), respectively, using a stratified random sampling scheme.

5.2.4 Hybrid Deep-Stacked Model Structure

The general idea of developing our hybrid stacked model was to combine the knowledge from 4 channels of features and train a meta-model. This was expected to reduce bias and achieve a robust model that reduced the effect of any possible noise caused by the imputation or the randomness of the data. The hybrid deep-stacked model with 4 channels of input (based on the 4 categories of features) was developed to predict the weekly ALC depletion trend during RT using pretreatment information. Each channel of input was separately processed to predict weekly ALC values for five weeks of RT treatment (output layer size of 5 nodes). Then, a fully connected neural network integrated the outputs of 4 channels to do the final prediction. The first 3 branches of the structure consisted of dense layers that parallelly predicted a sequence of 5 weekly ALC values from dosimetric, nondosimetric numerical, and nondosimetric categorical features. Since the nondosimetric categorical features were sparse, the least absolute shrinkage and selection operator (Lasso) regularization and dropout methods were added for nondosimetric categorical features to avoid the adverse effect of sparsity on the prediction and overfitting. The last branch in the structure considered the sequential features, for which we developed an encoder-decoder LSTM network structure to encode the sequential input (i.e., the encoder) and to predict a sequence of weekly ALC values (i.e., the decoder). Since we aimed to make our pretreatment predictions on the basis of pretreatment information only, we only included the baseline ALC value at the beginning of the treatment (i.e., week 0) as an input; thus, we had

a one-to-many LSTM structure. Then, all predictions from each branch were concatenated and fed into combined dense layers to predict the weekly ALC values. Fig 1. represents the model structure schema and data preprocessing flow. We refer to this model as "HDS-t0" throughout the paper.

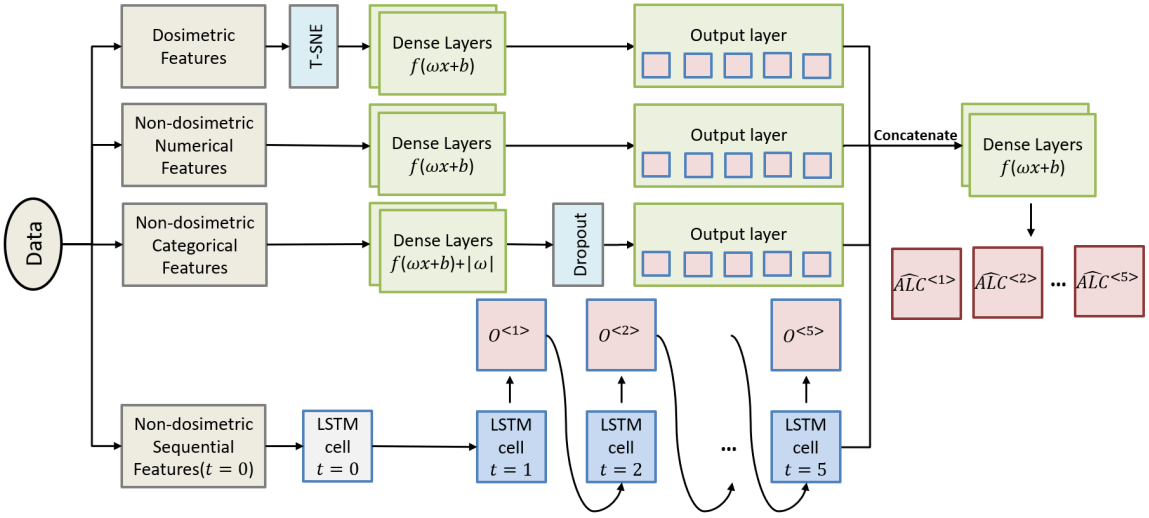


Figure 5.1: The model structure for predictions based on pretreatment data (HDS-t0)

The inner connections of the LSTM cells for prediction based on sequential features were based on the mathematical expressions as follows

$$f_t = \sigma\left(W_{fh}\, h_{t-1} + W_{fx}\, x_t + b_f\right), \tag{5.7}$$

$$i_t = \sigma\left(W_{ih}\, h_{t-1} + W_{ix}\, x_t + b_i\right), \tag{5.8}$$

$$\widetilde{c}_t = \tanh\left(W_{\widetilde{c}h}\, h_{t-1} + W_{\widetilde{c}x}\, x_t + b_{\widetilde{c}}\right), \tag{5.9}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \widetilde{c}_t, \tag{5.10}$$

$$o_t = \sigma\left(W_{oh}\, h_{t-1} + W_{ox}\, x_t + b_o\right), \tag{5.11}$$

$$\text{and} \quad h_t = o_t \tanh\left(c_t\right). \tag{5.12}$$

Where h_{t-1}, x_t, c_{t-1} denote the current hidden state, the input of the cell, and the current cell state vectors of the LSTM, respectively. W and b are the weight matrices and bias vector parameters of each layer in the LSTM cells. f_t is the forget gate's activation vector, which decides what information will be discarded from the cell state. This decision was made by a sigmoid function that returned 1 when it completely kept the information or 0 when it completely discarded all the information. Moreover, $\widetilde{c}_t$, i_t, o_t are the cell input, the input/update gate, and output gate activation vectors, respectively. The input gate decides what information is to be updated (sigmoid function in i_t) and what new information (hyperbolic tangent function in $\widetilde{c}_t$) is to be added and stored in the cell state, and the output gate uses updated cell state and hidden state information to decide what information can be output.

It is feasible to change the treatment plan for RT patients to avoid or mitigate grade 4 RIL. Therefore, the ability to predict lymphopenia at the early stages of RT could help to validate pretreatment predictions or prompt modification of the treatment plans for high-risk patients without losing much time. Thus, we extended our model to account for early-treatment predictions at the first and second weeks of RT to be used for validation or correction of pretreatment predictions. The model was extended to make predictions at the end of the first week of RT by adding the measured ALC value data for the first week to the input sequence in LSTM encoder. Also, the second week data was added to make predictions at the end of the second week of treatment. Therefore, a sequence-to-sequence (many-to-many) LSTM structure combined with a deep discriminative kernel was proposed to make predictions at the end of the first or second week of the RT treatment. The discriminative kernel was developed to reflect the correlation between input and output sequences by generating importance weights for each item in the input sequence to predict each item in the output sequence. The discriminative kernel automatically assigned higher weights to the most relevant part of the input sequence for each output and enabled the model to focus on

the most important information. The output of this layer were discriminative weights and weighted values of LSTM encoder output (i.e., context vector), which were calculated as

$$\alpha_{ts} = W_1 \tanh\left(W_2\, h_t + W_3\, \bar{h}_s\right), \tag{5.13}$$

$$w_{ts} = \frac{\exp\left(\alpha_{ts}\right)}{\sum_{s'=1}^{S} \exp\left(\alpha_{ts'}\right)}, \tag{5.14}$$

$$\text{and} \quad a_{context} = \sum_{s=1}^{S} w_{ts}\, \bar{h}_s\,, \tag{5.15}$$

where α_{ts} was used to evaluate the encoder hidden state at time t, h_t, and the decoder hidden state at time s, $\bar{h}_s$, which were normalized using a softmax function as shown in equation 5.14. Then, the context vector was calculated as a weighted average over all decoder hidden states as shown in equation 5.15.

Throughout this paper, we refer to the modified structure for predictions after 1 and 2 weeks of treatment as "HDS-t1" and "HDS-t2", respectively. Fig. 2 and Fig. 3 show the HDS-t1 and HDS-t2 model structures, consecutively.

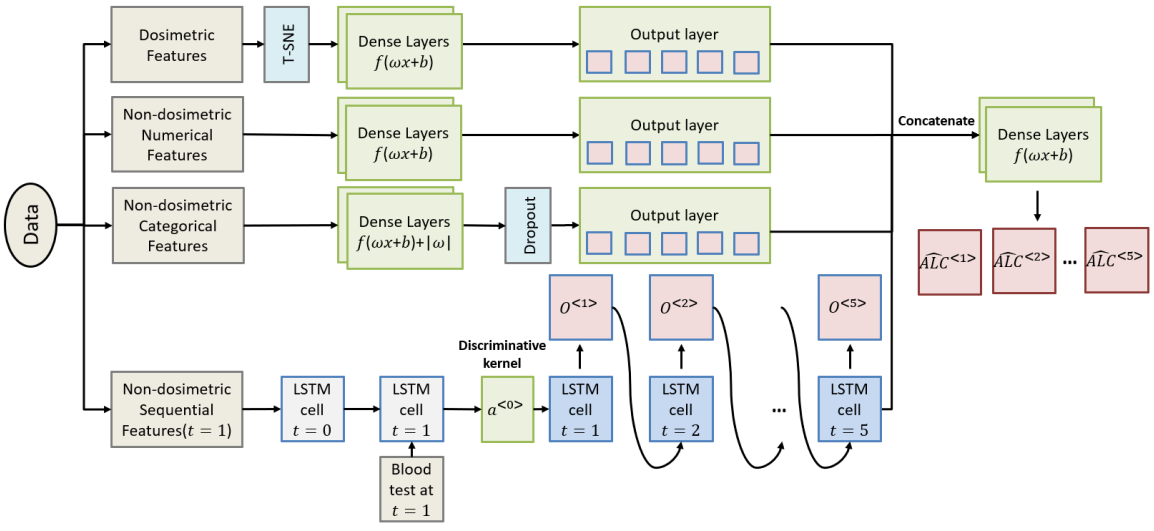


Figure 5.2: The model structure for predictions after 1 week of treatment (HDS-t1)

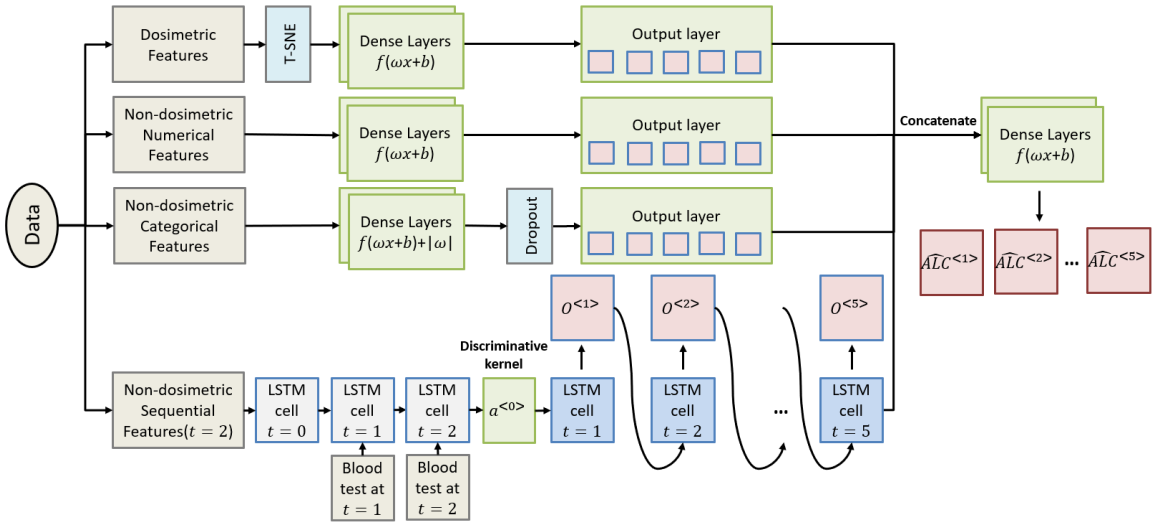


Figure 5.3: The model structure for predictions after 2 weeks of treatment (HDS-t2)

5.2.5 Training Algorithm and Model Configuration

Algorithm 5.1 shows the training flow of the proposed models. Each deep learning model was trained with 75 epochs using a batch size of 16 and was implemented using Adam optimizer [168] with a learning rate of 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$. The L1 regularizer weight of 0.2 and dropout rate of 0.2 were considered for the dense layers of nondosimetric categorical features. The regularization of the hyperparameters of the training set was adjusted empirically until the loss functions of both the training and testing sets declined according to similar trends and without significant gaps between them.

5.2.6 Evaluation Metrics

In order to evaluate the performance of the proposed models, several important prediction metrics were calculated using predictions for the test data, including mean square error (MSE), normalizeroot mean square error (NRMSE), mean absolute error (MAE), and

Algorithm 5.1 Training process

Initialization: Initial state, initial network weights $\theta_0 \rightarrow \theta$, size of output sequence N_{out}, number of epochs N_e, number of batches N_b, $e = 1$, $b = 1$;
for $e \leq N_e$ **do**
  for $b \leq N_b$ **do**
    Feed each feature group to the corresponding branch and obtain the output:
    $Y_{Dose} = f_{Dose}(X_{Dose})$
    $Y_{NonDose} = f_{NonDose}(X_{NonDose})$
    $Y_{NonDoseCat} = f_{NonDoseCat}(X_{NonDoseCat})$
    $X_{encode}, h_0, c_0 = f_{encode}(X_{Seq}, initial\ state)$
    Obtain each LSTM output sequence starting from $i = 1$:
    if $|X_{Seq}| > 1$ **then**
      $a_{context}, w_0 = f_{discr}(X_{encode}, h_0)$
      for $i \leq N_{out}$ **do**
        $X^i_{encode}, h_i, c_i = f_{encode}(a_{context}, h_{i-1}, c_{i-1})$
        $a_{context}, w_i = f_{discr}(X_{encode}, h_i)$
    else
      for $i \leq N_{out}$ **do**
        $X^i_{encode}, h_i = f_{encode}(X_{encode}, h_{i-1})$
  $X_{decode} = \left[X^1_{encode}, X^2_{encode}, \cdots, X^{N_{out}}_{encode}\right]$
  $Y_{out} = f_{final}(Y_{Dose}, Y_{NonDose}, Y_{NonDoseCat}, X_{decode})$
  Optimize the loss using Adam optimizer
  Calculate the batch total loss using $MSE = \left\|Y - \hat{Y}\right\|^2$
  Update network weights $\theta + \Delta\theta \rightarrow \theta$

explained variance (EV). These evaluation metrics were defined as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2, \tag{5.16}$$

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}{n}}}{(Y_{max} - Y_{min})}, \tag{5.17}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \widehat{Y}_i\right|, \tag{5.18}$$

$$\text{and} \quad EV = 1 - \left(Var(Y_i - \widehat{Y}_i)/Var(Y)\right), \tag{5.19}$$

where Y_i and $\widehat{Y}_i$ are the true values and predicted values, respectively.

5.2.7 Comparison Models

To evaluate the performance of our proposed hybrid deep-stacked model, we compared the results with other off-the-shelf prediction methods. Support vector machine (SVM), linear regression (LR), LR with lasso regularization (LR-Lasso), LR with ElasticNet (LR-ElasticNet), regression with stochastic gradient descent (SGD), decision tree, extra tree, and random forest models were developed for comparison with the proposed model. The same training and testing sets used for the hybrid deep-stacked model were used for all of these models. The hyperparameters for each model were selected based on a grid search to make the best possible predictions.

5.3 Results

5.3.1 Prediction Based on Pretreatment Data

The data were split into training and test sets in a ratio of 7:3 (602:258), respectively, using a stratified random sampling scheme. So, the proposed model was trained using data from 602 patients in the training set and another 258 patients in the test set. The MSE value of predictions using the baseline hybrid deep-stacked (HDS-t0) model for the 258 patients in the test set was 0.046. Fig. 4 shows the true ALC depletion trends versus the predicted curves for 15 randomly selected patients in the test set. As shown in the figure, the model provided accurate predictions with only small errors; these predictions are therefore suitable for use in pretreatment analysis to opt out patients at high risk of significant ALC reduction after RT.

As explained in section 5.2.7, eight off-the-shelf prediction methods were employed for comparison with the proposed model using the same training and test set. Table 5.3

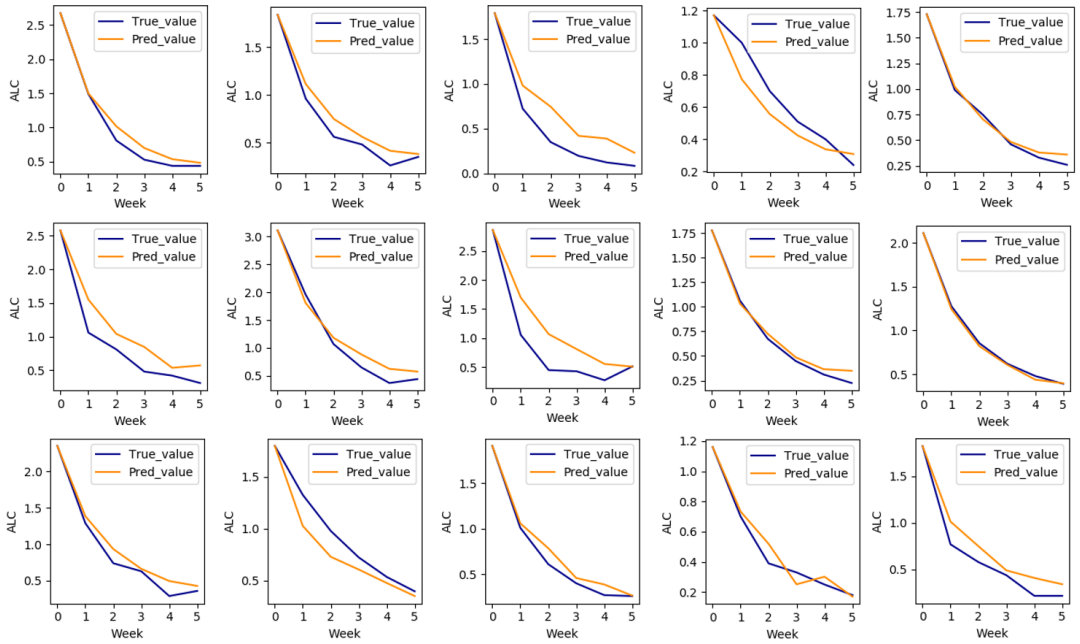


Figure 5.4: Predicted ALC trends using the HDS-t0 model (orange) versus the real values (blue) for 15 randomly selected patients in the test set. Unit for ALC values is cells $\times$ $1000/\mu L$ and ALC at week 0 refers to the baseline ALC.

compares the performance metrics of eight prediction models: SVM, LR, LR-Lasso, LR-ElasticNet, SGD, decision tree, extra tree, random forest, and the proposed hybrid model for pretreatment predictions (HDS-t0). As we can see from the table, LR achieved the best results among the eight prediction methods, with the lowest MSE, 0.0657. The proposed HDS-t0 model outperformed the LR model, with a reduction of 30.6% in the MSE of the predicted values. Our model also improved upon several other metrics, including the normalized RMSE (NRMSE) (-16.8%), MAE (-5.82%), and EV (+21.2%), compared to the best off-the-shelf model (i.e., LR).

5.3.2 Predictions After 1 and 2 Weeks of Treatment

As discussed in section II.D, the proposed hybrid deep-stacked model was extended to use the ALC value obtained after 1 week of RT to forecast an ALC trend for the rest of the treatment. Fig. 5 shows the true ALC depletion trends versus the predicted curves for

Table 5.3: Comparison of prediction performance metrics (MSE, NRMSE, MAE, and EV) of eight common prediction models and the proposed HDS-t0 model for predictions based on pretreatment information

	MSE	NRMSE	MAE	EV
HDS-t0	0.0456	0.0332	0.1457	0.7260
SVM	0.0736	0.0422	0.1635	0.5552
Random Forest	0.0698	0.0411	0.1596	0.5745
LR-Elastic Net	0.0859	0.0456	0.1771	0.4764
Decision Tree	0.0776	0.0433	0.1716	0.5269
Extra Tree	0.0691	0.0409	0.1593	0.5787
SGD	0.0839	0.0450	0.1722	0.4891
LR-Lasso	0.0859	0.0456	0.1771	0.4764
Linear Regression	0.0657	0.0399	0.1547	0.5992

the same 15 patients in the test set. As shown in the figure, the extended model, HDS-t1, provided more accurate predictions than did the HDS-t0 model. The HDS-t1 model achieved an MSE value of 0.014 for the test set predictions, a reduction of 69.6% compared to the HDS-t0 model. These results suggest that data from early stages of RT can be used to estimate the final patient response to treatment with more confidence than the pretreatment data. Therefore, our model's predictions after the first week can be used to validate the pretreatment predictions or modify the treatment plan for patients at high risk of grade 4 RIL during the early stages of treatment. To evaluate the effect of collecting more data during treatment on our model's ALC predictions, we also predicted the future ALC values after 2 weeks of treatment using the same test set with the HDS-t2 model. Fig. 6 shows the true ALC depletion trends and the predicted curves using the HDS-t2 model for the same 15 patients in the test set. Prediction metrics were calculated for the predictions based on pretreatment data using the HDS-t0 model and data from after the first and second weeks of treatment using the HDS-t1 and HDS-t2 model structures, respectively. Table 5.4 summarizes the MSE, NRMSE, MAE, and EV of each model for the predicted weekly ALC values of patients in the test set. As shown in the table, using the first-week

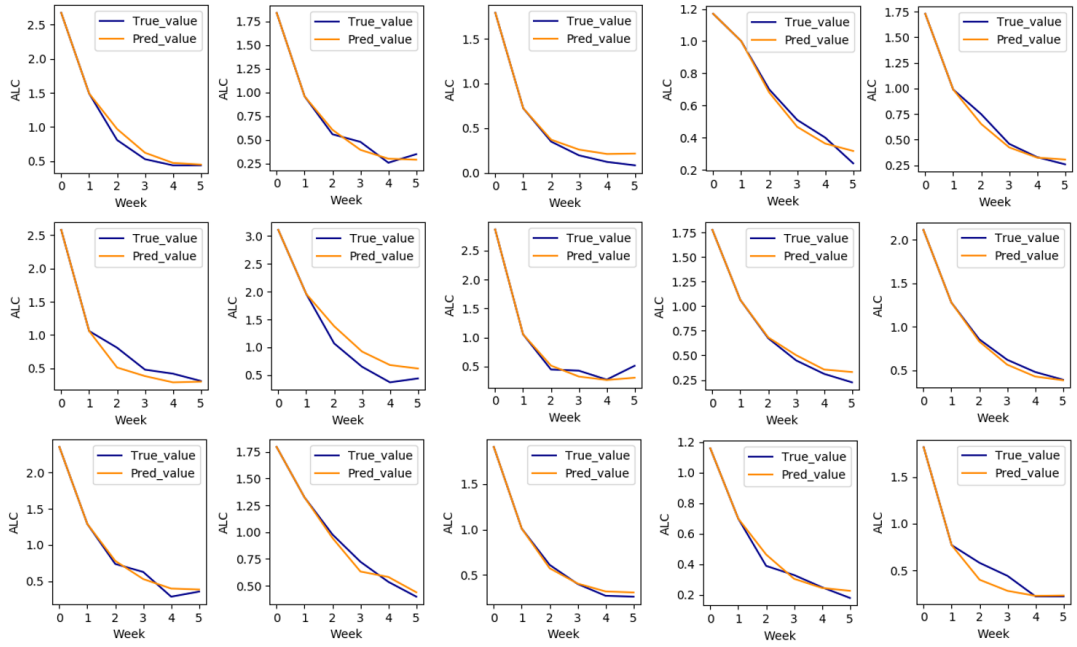


Figure 5.5: Predicted ALC trends based on the HDS-t1 model (orange) versus real values (blue) for 15 randomly selected patients in the test set. Unit for ALC values is cells $\times$ $1000/\mu L$ and ALC at week 0 refers to the baseline ALC.

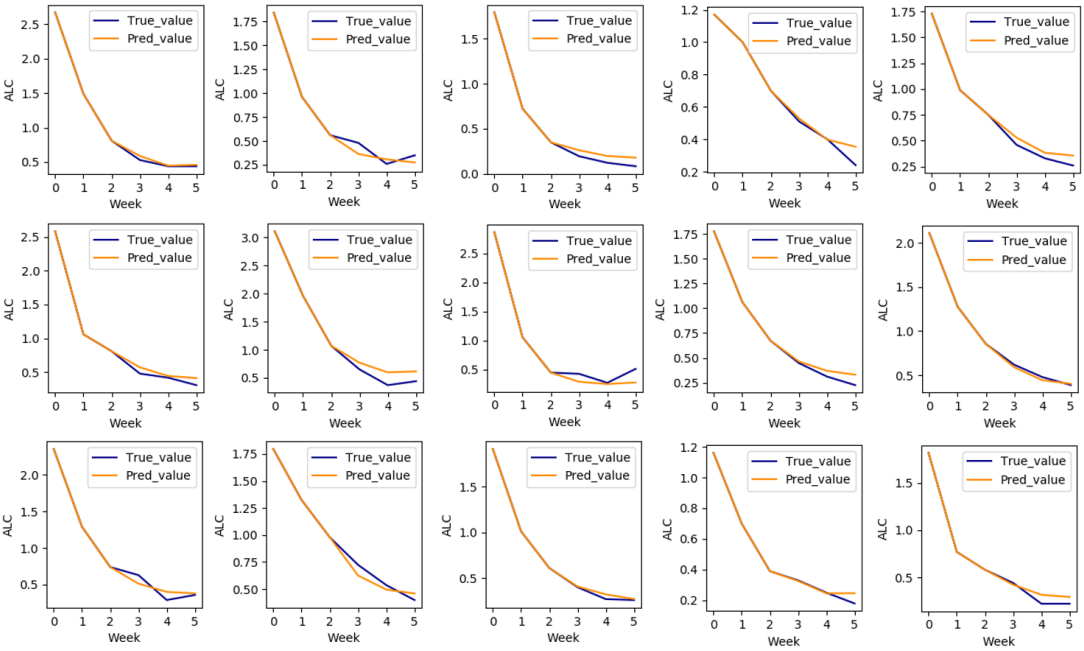


Figure 5.6: Predicted ALC trends based on the HDS-t2 model (orange) versus real values (blue) for 15 randomly selected patients in the test set. Unit for ALC values is cells $\times$ $1000/\mu L$ and ALC at week 0 refers to the baseline ALC.

data reduced the MSE of predictions by 69.6% compared to the model based on pretreatment data. Moreover, adding second-week data improved the MSE value by 42.9% over the HDS-t1 model and 82.6% over the HDS-t0 model. Therefore, we can conclude that augmenting the model with the first-week treatment data can significantly improve the pretreatment predictions, although the magnitude of improvement is smaller when additional weekly data (i.e., week 2 data) are included. This suggests that the difference between the baseline ALC value and that measured after the first week can provide very useful information to predict the ALC trend during the rest of treatment. By updating the predictions with the measured ALC after the first week of treatment, we can reduce the risk of false-negative pretreatment risk predictions. Also, physicians could use this method to validate the pretreatment predictions, update treatment plans, or develop mitigation strategies. This model could be also used after delivery of 1 fraction of radiation instead of after 1 week to reduce time and cost. Fig. 7 and Fig. 8 show the scatter plots and distributions of weekly

Table 5.4: Comparison of prediction performance metrics (MSE, NRMSE, MAE, and EV) for predictions based on HDS-t0, HDS-t1, and HDS-t2 models

	MSE	**NRMSE**	**MAE**	**EV**
HDS-t0	0.046	0.033	0.146	0.726
HDS-t1	0.014	0.018	0.069	0.917
HDS-t2	0.008	0.014	0.046	0.954

ALC values for real and predicted values based on each model. As shown in these figures, the distribution of predicted values using the HDS-t1 model was closer to the distribution of real values than were the pretreatment predictions using the HDS-t0 model. Similarly, the HDS-t2 model achieved more accurate predictions than the HDS-t1 model. Fig. 9 represents the box plot of the residual values (i.e., $ALC_i - \widehat{ALC_i}$) normalized by the mean ALC value within each week. This figure shows that the median of the error (i.e., the normalized residual value) within each week was almost zero for all 3 models, which suggests that the models performed well to predict weekly ALC for more than 50% of the patients

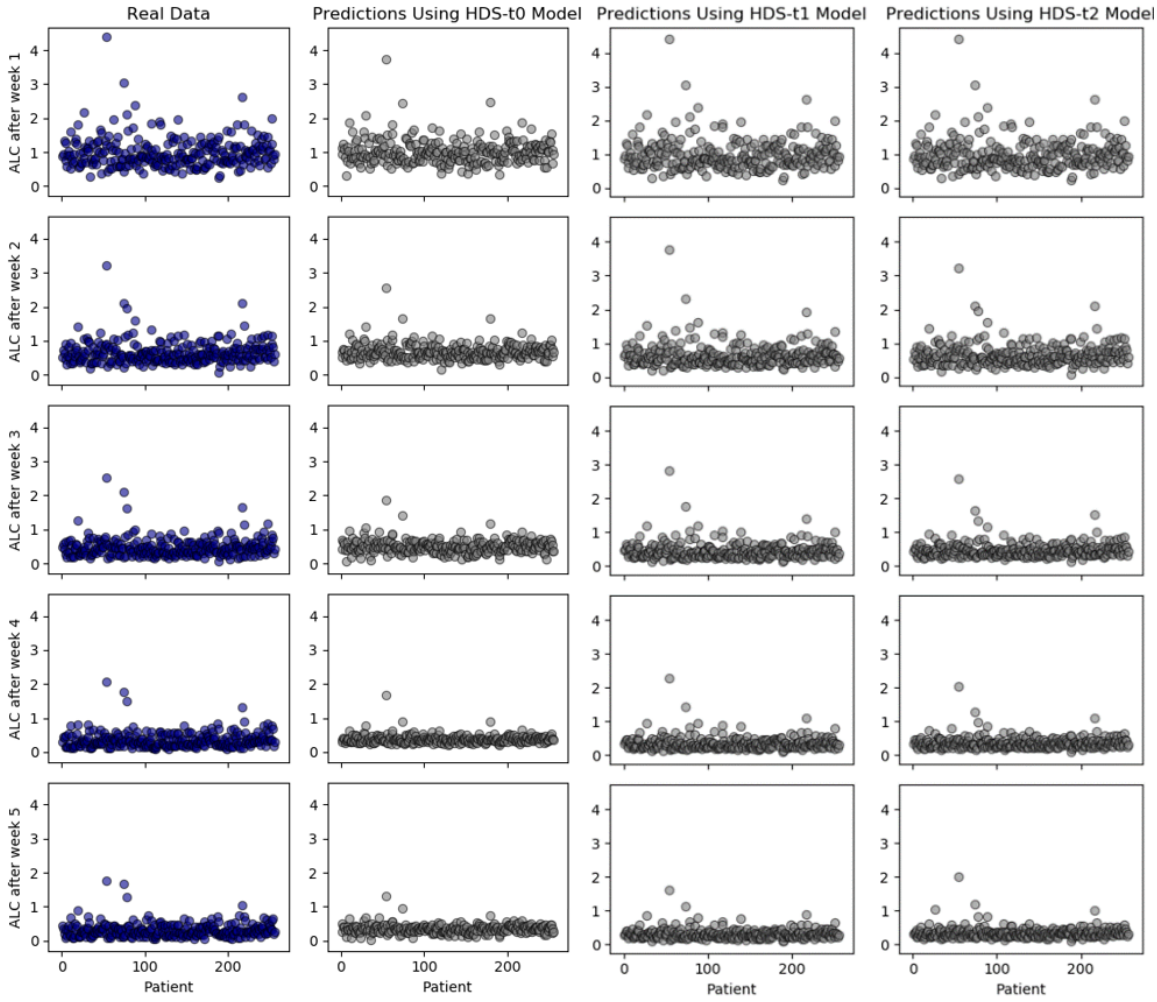


Figure 5.7: Scatter plots of weekly ALC values for real and predicted values from each model. Unit for ALC values is cells $\times\ 1000/\mu L$.

in the test set. Moreover, the range and interquartile range of the error in weeks 3 to 5 was the lowest for the HDS-t2 model, and the HDS-t1 model showed lower error values than the HDS-t0 model. This result is in agreement with our previous results comparing the models' performance.

The minimum ALC value during treatment, known as the ALC nadir, is an important factor in determining the occurrence of grade 4 RIL for RT patients. Thus, we determined the minimum ALC value during the five weeks of treatment for the real data and 3 predictions. Fig. 10 shows the histogram, box plot, and kernel density estimation of the ALC

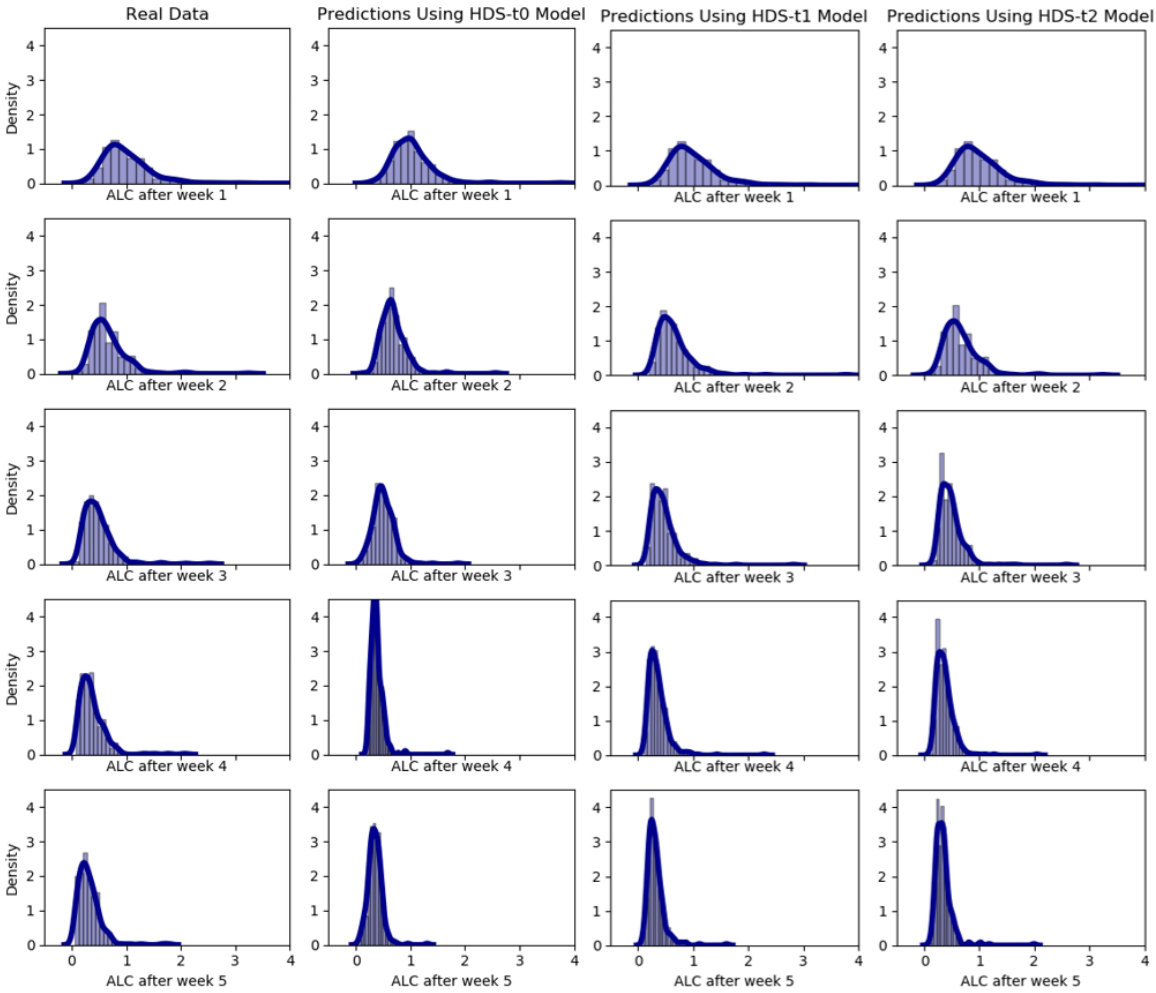


Figure 5.8: The distribution of weekly ALC values for real and predicted values based on each model. Unit for ALC values is cells $\times\ 1000/\mu L$.

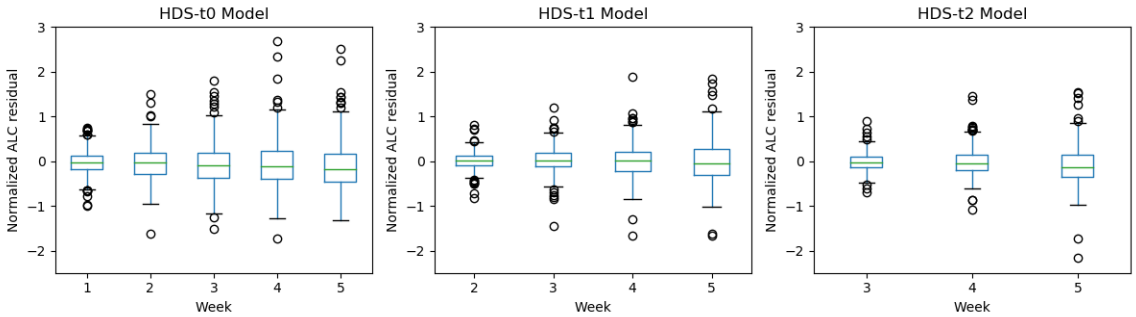


Figure 5.9: Box plots of predicted ALC values normalized to the mean ALC value for each week based on each model.

nadir during five weeks of treatment based on the real values and the predicted values from the HDS-t0, HDS-t1, and HDS-t2 models. As shown in the figure, the HDS-t2 model achieved the best predictions and the most similar distribution to the real data in terms of the range (real: 1.722, HDS-t2: 1.915), mean (real: 0.298, HDS-t2: 0.315), median (real: 0.25, HDS-t2: 0.318), interquartile range (real: 0.189, HDS-t2: 0.142), and kernel density estimation.

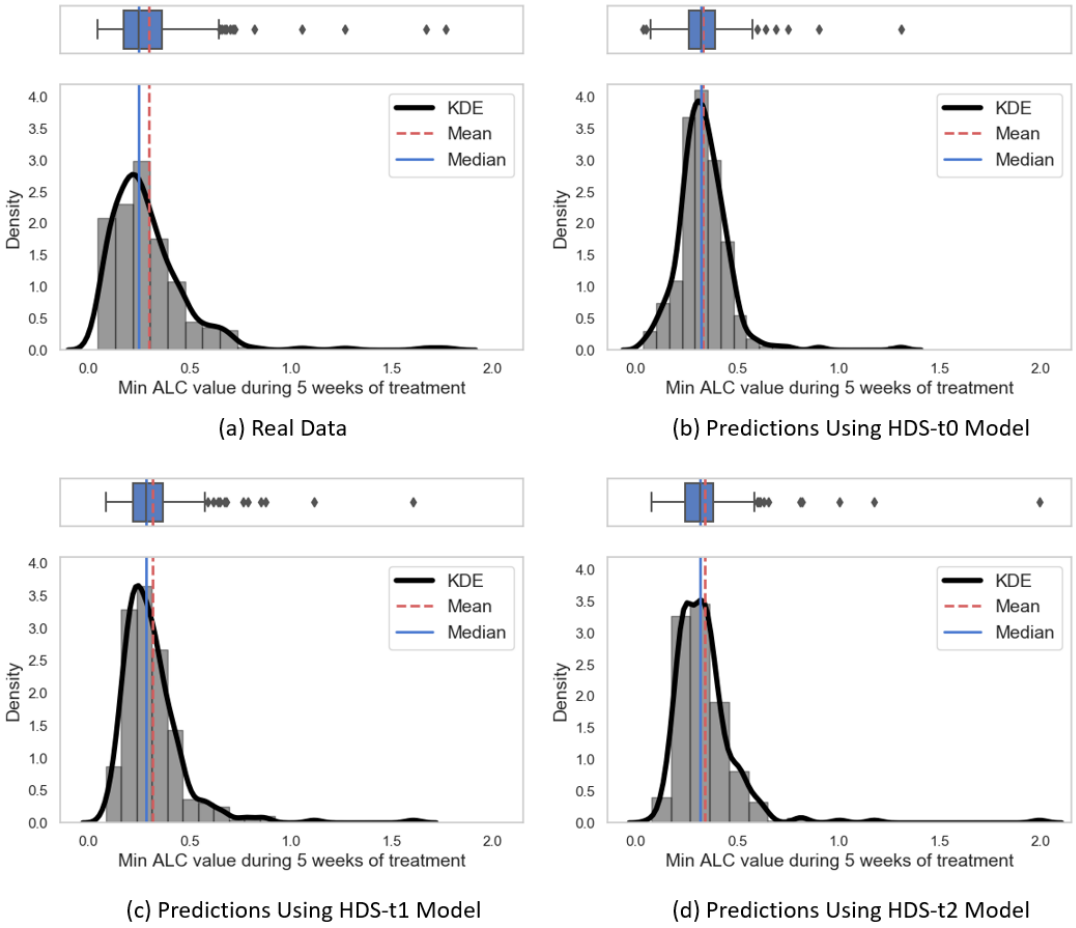
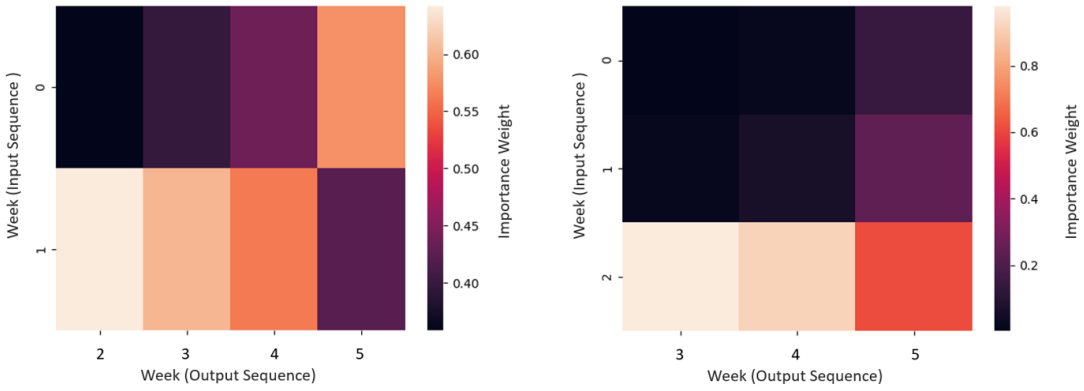


Figure 5.10: Histograms, box plots, and kernel density estimations (KDE) of minimum ALC values during five weeks of treatment for (a) real data; (b) predicted values using HDS-t0; (c) predicted values using HDS-t1; (d) predicted values using HDS-t2 models.

For the two extended models, HDS-t1 and HDS-t2, a discriminative layer was added to evaluate the importance weight of each value in the input sequence. The importance weights of each data value in the input sequence for each predicted value in the output

sequence were obtained from this discriminative layer. Using the HDS-t1 model, the input sequence contained the ALC value at the baseline (i.e., week 0) and after the first week of treatment (i.e., week 1), and the output sequence was the predicted ALC values at the ends of weeks 2 to 5. Fig. 11 (a) shows the importance weights for each input-output pair of the HDS-t1 model. As shown in the figure, the importance of baseline ALC was higher for the predicted ALC values of the last weeks of treatment (i.e., weeks 4 and 5) than the early ones (i.e., weeks 2 and 3). Likewise, Fig. 11 (b) presents the importance weights for each input-output pair based on the obtained results from the HDS-t2 model, which suggest the same conclusion as HDS-t1. Therefore, we can conclude that the models were able to capture long-term as well as short-term dependencies.



(a) Importance weights from the HDS-t1 model

(b) Importance weights from the HDS-t2 model

Figure 5.11: Heatmaps of importance weights from output of the discriminative kernel for each input-output pair based on the obtained results from HDS-t1 (a) and HDS-t2 (b) models

5.4 Discussion

In this paper, a hybrid deep-stacked model is proposed to predict RT-induced lymphocyte depletion for esophageal cancer patients during the course of RT. The proposed stacked structure processed 4 categories of features in 4 channels, which reduced the bias and adverse effects of any possible noise in the data, followed by a fully connected neural network

to integrate the outputs from 4 channels and make final predictions. First, the model was used to predict weekly ALC values during RT treatment course based on pretreatment information. Then, the model was extended to account for predictions made after the initial part of treatment (i.e., at the end of weeks 1 or 2), and a discriminative kernel layer was developed to evaluate the importance weight of each value in the input sequence. To our best knowledge, this is one of the first works to develop a comprehensive deep learning prediction model that can effectively use different categories of clinical features to forecast the weekly ALC regression during RT in order to identify high-risk patients.

In order to evaluate the performance of the proposed models, important prediction metrics were compared with those from eight off-the-shelf prediction methods. The results showed that the proposed model outperformed these off-the-shelf prediction methods in predicting weekly ALC values. Moreover, using the extended model based on the first-week data reduced the MSE of predictions compared to the model based on the pretreatment data. We conclude that augmenting the model with data from early stages of treatment (i.e., weeks 1 or 2) can significantly improve ALC predictions. Therefore, the HDS-t0 model using pretreatment data can be used in RT treatment planning to predict lymphocyte depletion during the course of RT. This prediction will help to select patients for RT and develop lymphopenia-mitigating strategies to ultimately improve patients' survival. After treatment is started, further predictions in the early stages of treatment can be used to validate the pretreatment predictions and, if necessary, modify the treatment plan for high-risk patients in order to preserve lymphocytes.

Although our proposed model has achieved good prediction performance compared to other state-of-art prediction methods, there is still some room for future work. First, the model evaluation only on the basis of model performance measures (e.g., mean squared error) cannot guarantee whether an individual prediction on a given patient should be trusted in clinical practice. Therefore, it is important to quantify the risk associated with each

prediction for a given individual patient to evaluate the reliability predictions to be used in clinical practice. In future work, we will investigate the benefits of using probabilistic deep learning approaches to predict RIL risk for RT cancer patients. Secondly, although the size of our data set might be sufficient to validate the rationale of the proposed model, it would be ideal to use data from a different institution and cancer type to further evaluate the robustness of our approach. In the next step, further analysis and discussions will be made based on different patient profiles to investigate the impact of different clinical factors on RIL and to estimate the risk associated with each predicted ALC value. This study also motivates further studies to investigate the effect of 1 fraction of radiation instead of 1 week by using fraction-based clinical data to reduce time and cost of identifying high-risk patients.

5.5 Conclusion

In this study, we proposed a new deep learning model based on deep neural network and long short-term memory (LSTM) network in a stacked structure to predict ALC depletion trend using pretreatment clinical information. The proposed model performed well in predicting radiation-induced lymphocyte depletion. Also, this model was implemented to account for predictions after 1 and 2 weeks of RT treatment which significantly improved ALC predictions. The ability to predict weekly ALC values during the course of RT treatment will enable physicians to identify patients who are at high risk of severe RIL and who would benefit from treatment replanning, use of different modality, or developing mitigations strategies. Moreover, our proposed deep learning method is interpretable and it is capable of providing the weights of each feature group in making the final prediction. It is flexible and can be transferred to predict other related toxic effects of RT.

Chapter 6

A Decision-Making Framework for Radiation Therapy Treatment Selection Including Lymphopenia Risk and Its Predictive Uncertainty

6.1 Introduction

Radiation therapy (RT) is one of the most common treatment options for cancer patients. The effectiveness of radiation therapy treatment relies on body's immune system response [169]. However, RT also suppresses the immune system by killing circulating lymphocytes in the radiation field since lymphocytes are highly radiosensitive [68, 70, 170]. As a result, RT causes a reduction in absolute lymphocytes count (ALC) which leads to radiation-induced lymphopenia (RIL), a common toxicity of RT [67, 69, 72]. Several clinical studies showed that high-grade RIL (grade 3 ALC $\leq$ 500 cells/μL or grade 4 ALC $\leq$ 200 cells/μL) correlates with poorer overall survival of patients with solid tumors, including cervical, pancreatic, rectal, lung, and esophageal cancer [71, 75, 77, 78, 83, 171].

Several studies have reported the strong associations between pretreatment factors including treatment-related characteristics (e.g., treatment modality, dose distribution patterns, fractionation regimens, etc.) [72, 77, 78, 84] and patient-specific factors (e.g., age, BMI, total blood volume, ALC baseline, etc.) [66, 71, 84–86]. Recent studies have reported greater lymphocyte depletion and higher risk of grade 4 lymphopenia in patients treated with photon therapy than with proton therapy [66, 68, 85, 161, 172]. For example, Ebrahimi et al. [161] assessed the RIL risks for ten esophageal cancer patients comparing proton and photon treatment modalities based on two models (i.e., a piecewise-linear and an exponential model) as a function of radiation dose distribution. Their results showed that proton plans carried a lower risk of lymphopenia after the RT treatment course than did photon plans.

Preservation of the lymphocytes from radiation damage is crucial for the effectiveness of RT [77, 78, 83]. Therefore, the ability to reliably predict RIL based on pretreatment factors (i.e., dosimetric factors, clinical, and patient-specific characteristics) would improve RT planning. van Rossum et al. [85] developed a pretreatment clinical nomogram based on age, planning target volume, body mass index, radiation modality, and baseline ALC to determine the risk of grade 4 RIL for new patients. Zhu et al. [86] introduced a hybrid deep learning model to classify patients with grade 4 RIL using dosimetric and clinical information. Recently, Ebrahimi et al. [173] proposed a hybrid deep learning model to forecast radiation-induced ALC depletion trend before or at the early stages of RT treatment using dosimetric, treatment- and patient-related clinical information. Although most of the prediction models achieved a good performance on large data sets using the standard statistical measures (e.g., accuracy, mean squared error), these metrics are based on averages over patients who may have a different characteristics and we cannot evaluate the prediction risks for a given individual patient using these measures alone [109].

The ability to identify patients who are at high risk of grade 4 RIL (G4RIL) based

on prediction models is very helpful to improve patient survival. However, poor performance for prediction models in clinical practice can have adverse consequences for patients. Therefore, it is important for clinicians to have some sense of when they can trust the prediction model results. Also, improving the management of risk and uncertainties in clinical decisions is reported as a potential opportunity to enhance the treatment outcome in medical practice [108]. Nevertheless, the model evaluation only on the basis of model performance measures (e.g., accuracy, mean squared error) cannot guarantee whether an individual prediction on a given patient should be trusted in clinical practice [109].

There are two main uncertainties in prediction results: (1) aleatoric uncertainty which is inherent noise and randomness in the real data due to the measurements and data collection errors. For example, in our data the weekly ALC values are subject to measurement errors since the measured ALC values are rounded to the nearest 100. (2) epistemic uncertainty or model uncertainty due to inductive assumptions or inadequate model, knowledge, and data [110, 111]. These two uncertainties need to be estimated sufficiently in uncertainty quantification models to account for predictions uncertainty. Consequently, in this study, we aimed to extend the deep learning model in Ebrahimi et al. [173] and develop a hybrid deep learning structure that can estimate the uncertainty associated with each predicted ALC value for a given individual patient.

Several methods have been developed to quantify the uncertainty of linear/nonlinear regression models and more complex prediction models such as neural networks in different applications by calculating prediction intervals [112–115] or determining a trust score [116]. A prediction interval can be computed for a neural network model with the assumption of normally distributed error for the neural network to accounts for aleatoric uncertainties [112, 114]. Neural network-based prediction intervals have been widely used in predicting health conditions and detecting diseases [116–119]. However, most of these methods require a large data sets and long trainings. Moreover, Bayesian inference has

shown a good performance in quantifying uncertainty of predictions for traditional machine learning models such as random forest [120], SVM [121], and variety of deep learning models, including neural networks [122–124], LSTM [125], CNN [126–128], and RNN [129]. However, these methods are computationally expensive and require significant modifications in training process. Furthermore, Deep Gaussian Processes are another effective alternative way to model the uncertainty of predictions as a non-parametric Bayesian approach by considering a Gaussian distribution over latent variables with respect to the input samples [130–132]. Unlike the Bayesian methods, this approach is easy to implement and archives high quality predictive uncertainty estimates. In this study, we take an advantage of this method in an ensembled neural network structure to estimate the uncertainty of ALC prediction.

Ebrahimi et al [173] showed that the ALC depletion prediction could be highly improved by incorporating the ALC measurement after one week of the treatment. Because it is still feasible to change the treatment plan at early stages of treatment (e.g., week 1 or 2) we can validate/correct the decision based on the pretreatment data using the new predictions. As a result, we can improve identification of high-risk patients and minimize false negative risk which is very important in current problem (i.e., determining grade 4 RIL for RT patients). Therefore, it is important to make the best use of prediction models before or at the early stages of the treatment and their estimated uncertainties to improve patient survival in clinical practice. A general decision-making framework based on the prediction models that can help physicians to use the results of complex deep learning models easily in their decisions for a given individual patient in the clinical practice is lacking. Therefore, to fill this gap, we propose a hybrid decision-making framework for selecting patients for RT treatment using the predicted values of ALC and their associated risk to be used in clinical practice with the goal of avoiding grade 4 RIL for cancer patients. This decision-making framework enables physicians to identify patients who are at high risk of grade 4

RIL and who may stand to benefit from treatment replanning, use of a different modality, or a pharmacological intervention and ultimately improve survival outcomes.

In summary, the contribution of this study is:

- We developed a deep learning model to predict the weekly ALC values and uncertainty associated with each predicted ALC value for individual patients to be used in estimation of prediction intervals. Estimating the prediction intervals for individual patients enables practical implications of predictive models in clinical decision-making by considering individual prediction risks.

- Different groups of patients with different pretreatment characteristics were assessed in terms of ALC predictions uncertainties based on results from the proposed deep learning model.

- We also proposed a comprehensive hybrid decision-making framework for selecting patients for RT treatment using the predicted values of ALC and their associated risk to be used in clinical practice with the goal of avoiding grade 4 RIL for cancer patients.

- This decision-making framework is flexible, straightforward, easy to interpret by clinicians, and can be modified to account for different levels of risk and enables physicians to easily take the advantage of complex deep learning models in their decisions for an individual patient.

- The effect of any treatment modifications (e.g., changing treatment modality) on the individual risk of G4RIL for each patient can be assessed using the proposed decision-making framework.

The rest of this chapter is organized as follows: section 2 covers the data description, data preprocessing, and our proposed hybrid deep learning model and decision-making

framework; section 3 presents the experimental results of the models; section 4 discusses the key findings of our study and directions for further research; finally, section 5 concludes this study.

6.2 Materials and Methods

6.2.1 Data Description and Preprocessing

Data of 860 esophageal cancer patients who received concurrent chemoradiotherapy were used for this study. All patients were treated with protons or photons radiation modalities with a dose of 50.4 Gy between January 2004 and November 2017 at MD Anderson Cancer Center. Pre- and weekly during RT ALCs, clinical characteristics and dosimetric parameters were extracted from data bases. All included patients also had available baseline ALC values and 3 or more documented weekly ALC values during treatment.

The variables of interest for the prediction were ALC after each week of treatment and estimated uncertainty of each prediction. We used the same predictor variables as the previous study on this data set by Ebrahimi et al. [173] which were selected on the basis of their clinical relevance, their low level of missingness (<20%), and the results of the correlation analyses as presented by Zhu et al. [86] and van Rossum et al. [85]. Therefore, 52 features were categorized into 4 main groups on the basis of their clinical and analytical characteristics: (1) dosimetric features contained 30 dose-volume metrics such as $V_5,\ V_{10},\ \ldots\ ,V_{45}$ (V_x refers to the percentage of the volume that received at least x Gy radiation dose) and mean dose for 3 organs at risk: the lung, heart, and spleen; (2) other numerical treatment-related and patient-specific parameters were nondosimetric numerical features including age, body mass index, total blood volume, planning target volume, and blood component profiles at baseline (red blood cells, white blood cells, and others); (3) nondosimetric categorical features such as RT modality (proton or photon), race, sex,

tumor location, tumor histologic characteristics, and use of induction chemotherapy; (4) sequential features included the sequence of 5 weekly ALC values.

Data preprocessing steps has been done based on the instructions in Ebrahimi et al [173] including handling the high collinearity among dosimetric features using t-distributed stochastic neighbor embedding (t-SNE) dimensionality reduction [166], imputing the missing values in in sequential features using multiple imputation approach, and removing the noise from the sequential ALC measurements by implementing Holt's double exponential smoothing (DES) method [167].

6.2.2 Deep Learning Model for Uncertainty Quantification of Predicted ALC Values

We aim to develop a deep learning model that can predict its own uncertainty. Considering y as all possible values of target variable and θ as the distribution parameters (e.g., μ and σ for a Gaussian distribution), $f(y|\theta)$ represents the likelihood of y given distribution parameters of θ. We assumed that the target variables (i.e., ALC values at each week) for each patient follows a Gaussian distribution with shape and scale parameters of μ and σ , respectively. So, a deep learning model must be developed to predict the distribution parameters (θ) instead of y. In current study, we used the hybrid deep learning model proposed by Ebrahimi et al. [173] to forecast ALC depletion during the course of RT treatment as a reference model and modified it to account for uncertainty predictions. This model consists of 3 channels of dense neural networks and one channel of LSTM network to process 4 categories of features including dosimetric, nondosimetric numerical, nondosimetric categorical, and ALC sequential values to predict weekly ALC values parallelly. Then, the information from each channel concatenated in a fully connected neural

network to make the final prediction. This structure showed a good performance in predicting weekly ALC depletion during the RT treatment course. The modified deep learning structure for predicting weekly ALC values and their predictive uncertainties is shown in Figure 6.1.

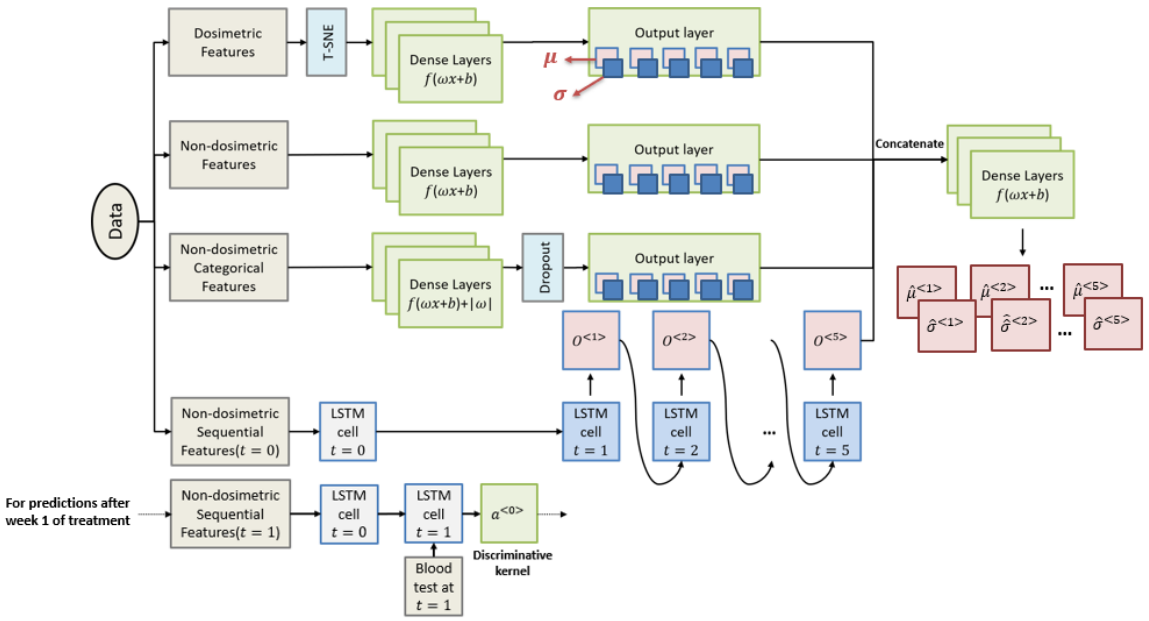


Figure 6.1: The proposed network structure for uncertainty quantification of ALC predictions

Firstly, we increased the size of output layer in each channel to match the number of distribution parameters in θ. So, μ and σ of ALC will be predicted for each week. Secondly, we used the nonnegative activation functions (i.e., ReLU) in each dense layer so that the output will be consistent with the logical upper bound and lower bound of the parameters in the chosen distribution. Finally, we need to adjust the loss function accordingly. Since the mean square error loss function cannot capture prediction uncertainty, we used maximum likelihood estimation (MLE) method to calculate the loss function. To find the correct values of θ which best describe the target distribution, we maximized the likelihood $L(\theta|Y)$ given the value of the target variable $y = Y$. As minimizing the negative log likelihood is equivalent to maximum likelihood estimation [174], we used a negative log likelihood

as the loss function to find the values of θ that maximize the likelihood $L(\theta|Y)$ having the value of the target variable $y = Y$ [131]. The loss function can be calculated as

$$Loss = -\log L(\theta \mid Y) = -\log\left(\frac{1}{\sqrt{2\pi\sigma}}\ \exp\left(-\frac{(Y-\mu)^2}{2\sigma^2}\right)\right). \tag{6.1}$$

By minimizing the negative log likelihood loss function during the training, the model will converge to the correct θ. For multiple samples, we can take average of the distributions and calculate mean negative log likelihood (MNLL) as

$$L_{MNLL} = \frac{\sum_{i=1}^{N_s} L(\theta_i,\ y_i)}{N_s}\ . \tag{6.2}$$

The training process is summarized in Algorithm 6.1. The model was trained with 100 epochs using a batch size of 16. Adam optimizer with a learning rate of 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$ was used to minimize the loss during the training. The L1 regularizer weight of 0.2 and dropout rate of 0.2 were considered for the dense layers of nondosimetric categorical features. The regularization of the hyperparameters of the training set was adjusted empirically.

After the training, we will have a model that takes a set of input features, X, and return a set of parameters, θ, which represent the probability distribution of the target variable, Y. Therefore, we can calculate prediction intervals to evaluate the reliability of the prediction and insure whether the model is confident about the predicted target value or not.

Algorithm 6.1 Training process

Initialization: Initial state, initial network weights $\theta_0 \rightarrow \theta$, size of output sequence N_{out}, number of epochs N_e, number of batches N_b, sample size N_s, $\mu^{(0)=0}$, $\sigma^{(0)=1}$, $i = 1$, $j = 1$;
for $i \leq N_e$ **do**
  for $j \leq N_b$ **do**
    Feed each feature group to the corresponding branch and obtain the output:
    $Y_{Dose}^{(i,j)} = f_{Dose}(X_{Dose}^{(i,j)})$
    $Y_{NonDose}^{(i,j)} = f_{NonDose}(X_{NonDose}^{(i,j)})$
    $Y_{NonDoseCat}^{(i,j)} = f_{NonDoseCat}(X_{NonDoseCat}^{(i,j)})$
    $X_{encode}^{(i,j)}, h_0, c_0 = f_{encode}(X_{Seq}^{(i,j)}, initial\ state)$
    Obtain each LSTM output sequence starting from $k = 1$:
    if $\left|X_{Seq}^{(i,j)}\right| > 1$ **then**
      $a_{context}, w_0 = f_{discr}(X_{encode}^{(i,j)}, h_0)$
      for $k \leq N_{out}$ **do**
        $X_{encode}^{k}, h_k, c_k = f_{encode}(a_{context}, h_{k-1}, c_{k-1})$
        $a_{context}, w_k = f_{discr}(X_{encode}^{(i,j)}, h_k)$
    else
      for $k \leq N_{out}$ **do**
        $X_{encode}^{k}, h_k = f_{encode}(X_{encode}^{(i,j)}, h_{k-1})$
  $X_{decode}^{(i,j)} = \left[X_{encode}^{1}, X_{encode}^{2}, \cdots, X_{encode}^{N_{out}}\right]$
  $Y_{out}^{(i,j)} = \theta^{(i,j)} = [\mu^{(i,j)}, \sigma^{(i,j)}] = f_{final}(Y_{Dose}^{(i,j)}, Y_{NonDose}^{(i,j)}, Y_{NonDoseCat}^{(i,j)}, X_{decode}^{(i,j)})$
  Calculate $loss = -\log L_{MNLL}(\mu^{(i,j)}, \sigma^{(i,j)}, Y_{target})$
  Optimize the loss using Adam optimizer
  Update network weights $\theta^{(i,j)} + \Delta\theta^{(i,j)} \rightarrow \theta^{(i,j+1)}$
  Update μ and σ as $\mu^{(i,j)} + \Delta\mu^{(i,j)} \rightarrow \mu^{(i,j+1)}$ and $\sigma^{(i,j)} + \Delta\sigma^{(i,j)} \rightarrow \sigma^{(i,j+1)}$
  if $\sigma \leq 0$ **then**
    $\sigma = 0$

6.2.3 Hybrid Decision-Making Framework for RT Patient Selection

We proposed a decision-making framework to decide if a given RT treatment plan is safe for a patient considering the G4RIL risk. The general idea of developing this decision-making framework was to create a use case for the ALC prediction models (e.g., pretreatment and early treatment prediction models) in clinical practice with the goal of improving identification of patients at high risk of G4RIL and minimizing false negative risk (i.e., missing the patients at high-risk of G4RIL). Therefore, in this decision-making framework, we consider the predicted weekly ALC values and their associated uncertainties to decide

whether we can trust the pretreatment predictions or we need to be cautious and make another prediction after finishing a part of the treatment (e.g., one week of treatment) to validate/correct the pretreatment predictions. Figure 6.2 shows the flowchart of our proposed decision-making framework. The possible ranges for ALC prediction intervals are also marked by letters in the figure.

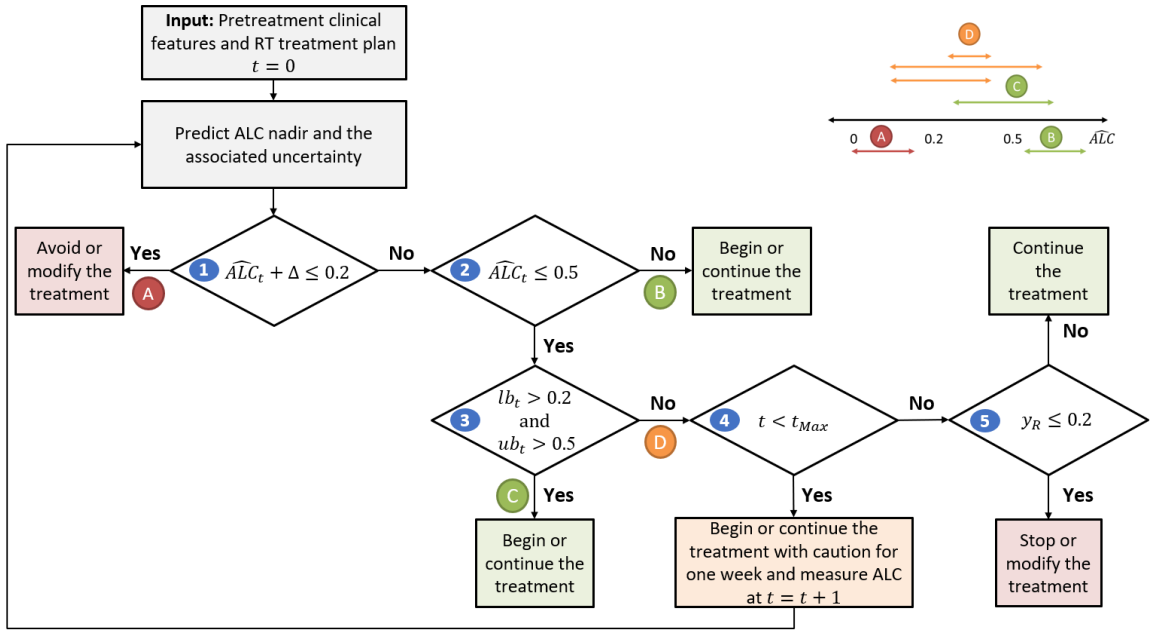


Figure 6.2: Flowchart of the proposed decision-making framework for selecting patients for RT treatment

As shown in the flowchart, following steps will be taken based on the proposed decision-making framework. In the first step, we use the pretreatment clinical information to predict weekly ALC values and their estimated uncertainty in terms of normal distribution parameters (e.g., mean and standard deviation). Then, we determine the ALC nadir ($\widehat{ALC}$) and its uncertainty based on the predicted values. Since we want to be risk averse as much as possible, we avoid the treatment whenever the predicted ALC nadir plus an uncertainty term Δ, which is adjustable based on physicians preferences, is lower than ALC threshold for the grade 4 RIL (i.e., $\widehat{ALC}$<0.2). Otherwise, to be more confident about the risk of G4RIL, we also compare the predicted ALC nadir with grade 3 RIL threshold (i.e., $\widehat{ALC}$<0.5) in

the second step. If the predicted ALC nadir is greater than the threshold of 0.5, we can conclude that the patient will not experience neither of grade 3 nor grade 4 RIL so that the RT treatment is safe for the patient. If not, there is a possibility that the patient might develop grade 3 RIL and we need to be more cautious; so, we consider the uncertainty of prediction by looking at the prediction intervals based on the selected confidence level α (i.e., $\widehat{ALC} \pm z_{\alpha/2}$) to decide if we can trust the pretreatment predictions or not. In the third step, if the lower prediction interval lb and the upper prediction interval ub were greater than the G4RIL and G3RIL thresholds, respectively, which means we are $(1-\alpha)\,\%$ confident that the patient will not experience grade 4 RIL considering the uncertainties. Also, the risk of grade 3 RIL is low because the upper prediction interval is not less than 0.5. As a result, we can conclude that the RT treatment is safe for the patient in terms of risk of grade 4 RIL. Otherwise, although $\widehat{ALC}$ is greater than G4RIL threshold, there is a risk that the patient experience either grade 3 RIL (i.e., $ub < 0.5$) or grade 4 RIL (i.e., $lb < 0.2$) due the uncertainty in the predictions. Therefore, we should begin the first part of treatment (e.g., one week) with caution then measure the ALC and make predictions using the early treatment prediction model. Next, since it is feasible to change the treatment plan only in the early stages of the treatment, we repeat the steps 1-3 until a predefined t_{Max} which can be selected by physician based the treatment modality, cancer type, and patient characteristics. When we could not make a decision for a given patient in early time steps (i.e., $t < t_{Max}$) and the t_{Max} is reached, we need to finalize the decision; so, in step 4, we check if we reached to t_{Max} point or not. Then, we use a linear regression based on the predicted ALC nadirs and their associated uncertainty, in terms of standard deviation, to calculate the expected ALC nadir and compare it with grade 4 RIL threshold (step 5). This regression line can be fitted using the same training set as we used for the prediction models. Then, the fitted parameters determine the weight of predicted ALC nadir at each time step and the

weighted sum of them can be calculated for final decision as

$$y_R = \beta_0 \ \widehat{ALC}_0 + \beta_1 \ \widehat{ALC}_1 + \cdots + \beta_{t_{Max}} \widehat{ALC}_{t_{Max}} + \gamma_0 \ \widehat{\sigma}_{\widehat{ALC}_0} + \gamma_1 \ \widehat{\sigma}_{\widehat{ALC}_1} + \cdots + \gamma_{t_{Max}} \ \widehat{\sigma}_{\widehat{ALC}_{t_{Max}}} \ , \tag{6.3}$$

where β_t and γ_t refer to the weights for predicted ALC nadir $\widehat{ALC}_t$ and the associated standard deviation $\widehat{\sigma}_{\widehat{ALC}_t}$ at time step $t, \ \forall t \in \{0, 1, \ldots, \ t_{Max}\}$.

In summary, the output of this decision-making framework will be a decision about a proposed RT treatment plan for a patient which minimized the risk of G4RIL. Therefore, different treatment plans and modalities can be evaluated for individual G4RIL risk and the safest treatment can be chosen for the patient.

6.2.4 Evaluation Metrics

In order to evaluate the performance of the proposed deep learning model in predicting weekly ALC values, two important prediction metrics were calculated using predictions for the test data, including mean square error (MSE) and mean absolute error (MAE), these evaluation metrics were defined as

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \widehat{Y}_i \right)^2, \tag{6.4}$$

$$\text{and} \quad MAE = \frac{1}{n} \sum_{i=1}^{n} \left| Y_i - \widehat{Y}_i \right|. \tag{6.5}$$

For the decision-making framework, since the output is whether a patient could be selected for a specific RT treatment based on G4RIL risk or not, this problem can be inferred as a classification problem in which we classify the patients into two group: (1) patients at high risk of G4RIL as the positive class, (2) patients at low risk of G4RIL as the negative class. Therefore, the ability of the decision-making framework to classify high-risk

patients can be assessed using common classification evaluation metrics such as accuracy, recall, precision, and F1 score. These metrics can be calculated as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} , \tag{6.6}$$

$$Precision = \frac{TP}{TP + FP} , \tag{6.7}$$

$$Recall = \frac{TP}{TP + FN} , \tag{6.8}$$

$$\text{and} \quad F1\ Score = \frac{2 \times Recall \times Precision}{Recall + Precision} , \tag{6.9}$$

Where TP indicates the number of true positive cases (i.e., high-risk patients who are correctly identified); FP, false positive cases (i.e., low-risk patients who are incorrectly classified as high-risk patients); FN, false negative cases (i.e., high-risk patients who are incorrectly classified as low-risk patients); and TN, true negative cases (i.e., low-risk patients who are correctly identified).

6.3 Results

6.3.1 Evaluation of Predictive Uncertainty Quantification Method

The proposed deep learning model was trained using data from 602 patients in the training set and tested for another 258 patients in the test set. For each patient, parameters of Normal distribution (μ and σ) were predicted for ALC value in 5 weeks of treatment. The predicted μ and σ for each week were considered as the estimated ALC and its uncertainty at that week. The model achieved MSE value of 0.054 for weekly ALC predictions for the 258 patients in the test while using the pretreatment information. Moreover, using the early treatment prediction model and augmenting the measured ALC value after the first week of treatment resulted in a reduced MSE of 0.017 as we expected based on the results of the

previous study by Ebrahimi et al. [173]. Figure 6.3 shows the true weekly ALC values versus the predicted curves based on pretreatment information along with the 90% and 95% prediction intervals for 10 randomly selected patients in the test set. As shown in the figure, although the model provided accurate predictions with only small errors, but the width of prediction interval varied for different patients. This observation confirms that we cannot trust the predictions for all patients and shows the importance of estimating the uncertainty of the predictions for each individual patient. Having the prediction intervals helps physicians to evaluate the predictions and use them for pretreatment analysis in clinical practice with more confidence.

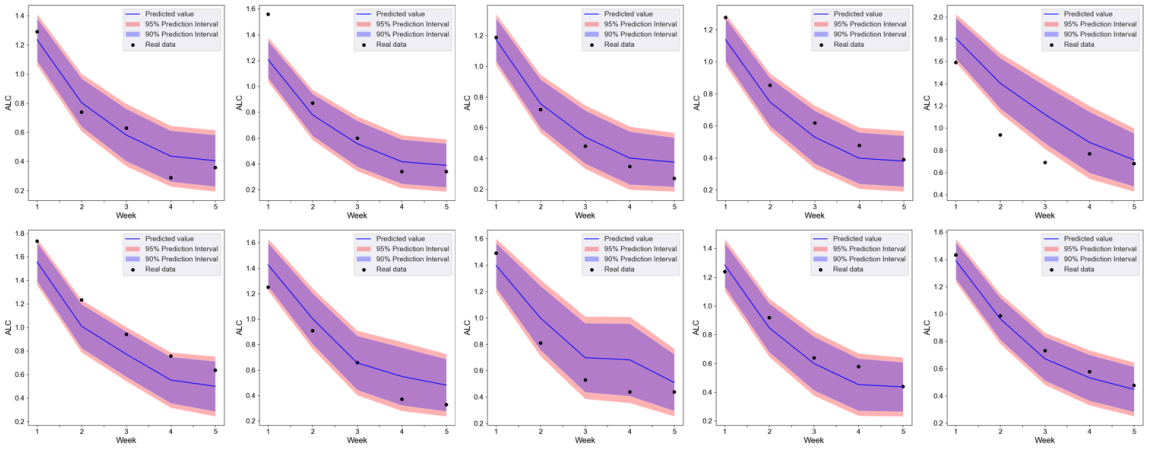


Figure 6.3: Predicted ALC trend and its 90% and 95% prediction intervals using the model based on pretreatment information (blue) versus the real values (blackdots) for 10 randomly selected patients in the test set

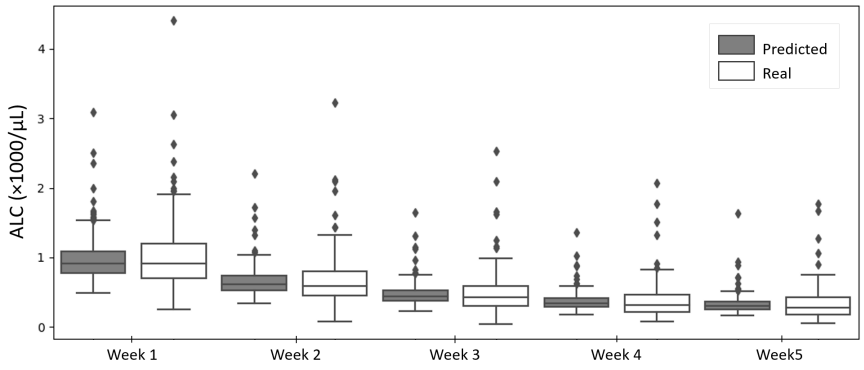
To further evaluate the performance of the model, we calculated the average of predicted ALCs, prediction intervals, and real ALCs over all patients in the test set for each week. Table 6.1 summarizes the comparison of important statistical metrics including mean, standard deviation (SD), median, interquartile range (IQR), and prediction intervals (PI) for predicted versus real weekly ALC values for all patients in the test set. We can see from the table that the variability of real ALC measurements after each week of treatment in terms of IQR and SD decreased as we go from week 1 to week 5 of the treatment. The

same trend for SD and IQR as the real data was observed in the predicted ALCs. Moreover, the mean and median ALC had a decreasing trend from week 1 to week 5 of the treatment as a result of RT based on the predicted values which is also in agreement with the real ALC measurements. Furthermore, the model performed very well in prediction of weekly ALC values with small MSE and MAE which improved for last weeks because of the improvement of the information in hidden states of LSTM network.

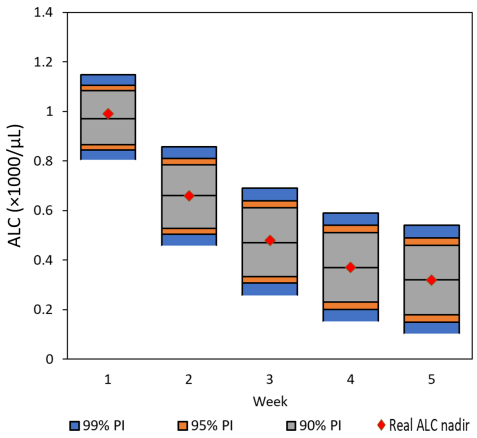
Table 6.1: Comparison of important statistical metrics including mean, standard deviation (SD), median, interquartile range (IQR), and prediction intervals (PI) for predicted versus real weekly ALC values for all patients in the test set

	Real ALC nadir				**Predicted ALC nadir**				**MSE**	**MAE**
	Mean	**SD**	**Median**	**IQR**	**Mean, (95%PI)**	**SD**	**Median**	**IQR**		
Week 1	0.99	0.46	0.91	0.49	0.97, (0.84, 1.11)	0.31	0.91	0.31	0.096	0.228
Week 2	0.66	0.34	0.59	0.36	0.66, (0.50, 0.81)	0.21	0.61	0.22	0.064	0.180
Week 3	0.48	0.28	0.43	0.28	0.47, (0.31, 0.64)	0.16	0.44	0.15	0.052	0.166
Week 4	0.37	0.23	0.32	0.25	0.37, (0.20, 0.54)	0.13	0.34	0.13	0.041	0.150
Week 5	0.32	0.21	0.27	0.24	0.32, (0.15, 0.49)	0.13	0.30	0.11	0.032	0.133

Figure 6.4(a) represents the box plot of weekly ALC values based on the real data versus the predictions using pretreatment information. This figure shows that the prediction model could capture the distribution and variability of the real data very well. Also, Figure 6.4(b) depicted the average 99%, 95%, 90% prediction intervals of predicted ALC for each week. In this figure, the mean value of predicted ALC is in the middle of the intervals and the mean of real ALC values is shown by red diamonds. As we can see from the figure, real ALC weekly means are within the prediction intervals based on different confidence levels. Also, the mean of predicted ALCs for each week and mean of real ALCs for each corresponding week were very close to each other which means that the model achieved very good performance in predicting weekly ALC values.



(a)



(b)

Figure 6.4: Analysis of weekly ALC values; (a) box plot of weekly ALC values based on the real data versus the predictions using pretreatment information; (b) the average 99%, 95%, 90% prediction intervals of predicted ALC for each week.

6.3.2 Decision-making Framework Results

6.3.2.1 Comparison of Decision-making Framework with Classification Methods

To evaluate the performance of our proposed decision-making framework, we compared the results with a recently published classification deep learning model on the current data set developed by [86]. They showed that their proposed classification model achieved superior performance in classification of patients at high risk of G4RIL compared with the popular classification models such as logistic regression, support vector machines, and random forest.

To have a fair comparison, we used the same data inclusion exclusion criteria as [86] and trained the pretreatment and early treatment prediction models using the same training set with 505 patients, then, we tested the decision-making framework using the same test set with 216 patients based on the predictions made by the models.

As we mentioned earlier, our proposed decision-making framework is flexible and risk parameters (i.e., Δ and confidence level α for prediction intervals), time threshold for modifying the treatment (t_{Max}) can be adjusted based on the physician's preference for considering the predictions uncertainties in their decisions. In this experiment, we assumed Δ=0 since we wanted to be risk averse about low ALC nadirs as much as possible and avoid the treatment for any patient with predicted ALC nadir lower that 0.2. We also assumed that $t_{Max} = 1$ (week) and the treatment can be modified up to the beginning of the second week. Moreover, the linear regression model in Equation (1) was fitted using the same training set as we used for the prediction models. The standard deviation of the predicted ALC nadir based on pretreatment information ($\widehat{\sigma}_{\widehat{ALC}_0}$) and predicted ALC nadir after the first week of the treatment ($\widehat{ALC}_1$) were significant variables with the p-value of 0.034 and 0.001, respectively. The fitted regression line was as

$$y_R = \beta_1 \, \widehat{ALC}_1 + \gamma_0 \, \widehat{\sigma}_{\widehat{ALC}_0} = 1.1 \, \widehat{ALC}_1 - 0.93 \, \widehat{\sigma}_{\widehat{ALC}_0} \, . \tag{6.10}$$

First, we tested the model for different values of α to select the best one based on our another preference which is minimizing the false negative rate (i.e., high-risk patients who are incorrectly classified as low-risk patients) to improve RT patients survival. Therefore, we calculated recall, precision, and F1 score using $(1-\alpha)\%\epsilon\{68,\ 85,\ 90,\ 93,\ 95,\ 97,\ 99,\ 99.7\}$ in the decision-making framework for patients in the test set. Figure 6.5 shows the comparison of these $1-\alpha$ values. As we can see from the figure, increasing the $1-\alpha$ resulted

in higher precision and lower recall as we were stricter about classifying patients as positive class, this is similar to a classification problem in which we set a high discriminative threshold for positive class. As a result, lower number of patients classified as positive class and number of false positive cases decreased, so, the precision increased. At the same time, we might misclassify some positive class cases as negative class, therefore, the number of false negative cases increased which resulted in lower recall value. In order to have a reasonable balance between false negative and false positive errors, we considered F1 score as our metric to select the best α value. As shown in Figure 6.5, $(1-\alpha)\% = 95\%$ achieved the highest F1 score value. So, we used α =0.05 for the rest of the numerical experiments to calculated prediction intervals in decision-making framework.

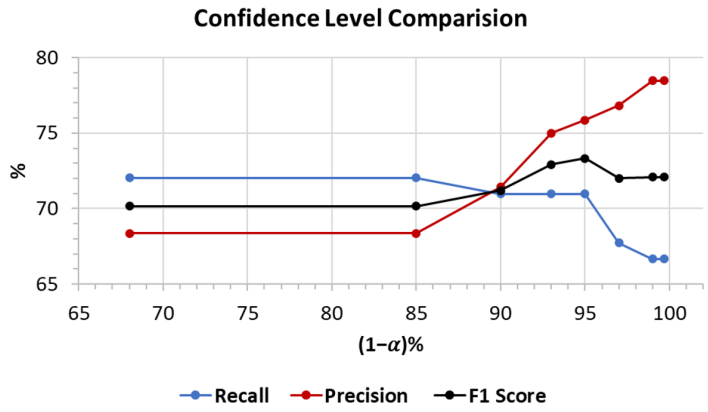


Figure 6.5: Comparison of recall, precision, and F1 score of decision-making framework results for different levels of α used in prediction intervals

The decision-making framework results for the 216 patients in the test set were evaluated using important classification metrics including accuracy, recall, precision, and F1 score, which were also reported in [86]. The results showed that using the decision-making framework increased recall and precision by 8.2% (0.656→0.710) and 2.6% (0.759→0.740), respectively, compared to the results from [86] model. So, we can conclude that the number of both false negative and false positive cases decreased and the amount of improvement in the number of false negative cases were about 3 times more than the improvement

in the number false positive cases as we wanted to be more cautious about false negative cases. Also, the F1 score and accuracy increased by 5.5% (0.695→0.733) and 1.2% (0.769→0.778), respectively. Since the classification model by [86] achieved the best results among common classification models, we can conclude that our proposed decision-making framework also outperformed logistic regression, support vector machines, and random forest models in classifying G4RIL patients. Table 6.2 also represent the summary of classification metrics for the classification results based on the decision-making framework and models in [86].

Table 6.2: Summary of classification metrics including accuracy, recall, precision, and F1 score for the classification results based on the decision-making framework and classification methods

	Accuracy	**Recall**	**Precision**	**F1 Score**
Decision-making framework	0.778	0.710	0.759	0.733
Classification model by Zhu et al.[86]	0.769	0.656	0.740	0.695
Logistic regression	0.717	0.563	0.681	0.616
Logistic regression with elastic-net regularization	0.722	0.575	0.621	0.589
Random forest	0.718	0.644	0.651	0.621
Support vector machines	0.699	0.736	0.604	0.575

6.3.2.2 Examples of Decision-making Framework Results for Individual Patient Cases

To further assess the performance of the proposed decision-making framework, we took one individual patient example for five possible actions in the decision-making framework. Having $t_{Max} = 1\ (week)$, there are five possible actions based on the decision-making framework: (I) Avoid or modify the treatment, (II) Begin the treatment when $\widehat{ALC} > 0.5$, (III) Begin the treatment when $\widehat{ALC} \leq 0.5$, $lb > 0.2$ and $ub > 0.5$, (IV) Begin the treatment with caution for one week then continue the treatment, (V) Begin the treatment with caution for one week then stop the treatment. Table 6.3 presents some pretreatment characteristics of the selected patients, real treatment outcome, and the selected action based on the decision-making framework.

As shown in the table, the predicted ALC nadir for Patient I was lower than 0.2; so,

Table 6.3: Pretreatment characteristics, real treatment outcome, and selected action based on the decision-making framework for five randomly selected patient cases

	Patient I	**Patient II**	**Patient III**	**Patient IV**	**Patient V**
Age	76	56	65	55	79
BMI	27.74	22.66	22.92	33.05	21.52
Baseline ALC	1.37	3.65	2.92	3.06	1.77
Treatment Modality	IMRT	IMRT	IMRT	IMRT	IMPT
ALC nadir	0.10	0.55	0.46	0.47	0.14
G4RIL	Yes	No	No	No	Yes
Predicted ALC nadir at t=0, (95%PI)	0.18, (0.07, 0.29)	0.54, (0.16, 0.93)	0.46, (0.22, 0.71)	0.30, (0.03, 0.56)	0.28, (0.15, 0.39)
Predicted ALC nadir at t=1, (95%PI)				0.41 (0.26, 0.55)	0.28 (0.19, 0.36)
Decision-making framework decision	(I) Avoid the treatment	(II) Begin the treatment	(III) Begin the treatment	(IV) Begin the treatment with caution for one week then continue the treatment	(V) Begin the treatment with caution for one week then stop the treatment

the decision-making framework recommended to avoid the proposed RT treatment for this patient which is in agreement with the real data that this patient developed G4RIL during the RT treatment course. For Patient II, the decision-making framework suggested to begin the treatment as the predicted ALC nadir was greater than 0.5 and the risk of both G4RIL and G3RIL were low for this patient. The real data also showed that this patient did not experience neither of G4RIL nor G3RIL and the decision was correct. The decision-making framework decided to begin the RT treatment for Patient III as $\widehat{ALC} > 0.2$, $lb > 0.2$ and $ub > 0.5$ and we are 95% confident that the real ALC nadir is within this prediction interval and higher than 0.2, so, the risk of G4RIL is low for the patient. This action is also consistent with the real data in which the patient did not developed G4RIL. Although the predicted ALC nadir for Patient IV was greater than 0.2 and lower than 0.5 but the lower prediction interval lb was lower than the critical threshold of 0.2; so, the decision-making framework considered this prediction as an uncertain prediction and recommended to Begin the treatment with caution for one week then measure the ALC and make another prediction based on early treatment prediction model after one week. The predicted ALC nadir

based on the new prediction was also greater than 0.2 but it showed less uncertainty than the pretreatment prediction as the width of prediction interval decreased and $lb > 0.2$ and $ub > 0.5$. Therefore, the decision-making framework advised that completing the RT treatment is safe for the patient which agrees with the real data as this patient did not developed G4RIL. For Patient V, the decision-making framework recommended to begin the treatment with caution for one week then measure the ALC and make another prediction afterwards because $\widehat{ALC} < 0.5$ and both G4RIL and G3RIL are probable based on the 95% prediction intervals. Although the predicted ALC nadir based on the predictions after week 1 was greater than 0.2 and lower than 0.5 but the prediction intervals were $lb < 0.2$ and $ub < 0.5$ and there was still a high chance of G4RIL, we needed to be more cautious about the decision. However, we reached to the $t_{Max} = 1$ and we cannot make another prediction, so, we should use the linear regression model to make a decision. Since the $y_R \leq 0.2$ the decision-making framework identified this patient at high risk of G4RIL and suggested to stop the treatment which is consistent with the real data as the patient developed G4RIL in real practice. Figure 6.6 shows the real ALC depletion, predicted weekly ALCs and the associated 95% and 90% prediction intervals for these patient examples.

6.4 Discussions

6.4.1 Analysis of ALC Prediction Risk for Different Treatment Modalities and ALC Baseline Values

Deep learning has shown great potential in many medical applications such as disease progression [88, 175], treatment outcomes [89], or potential side effects [90, 91]. Despite the success of standard DL methods in solving various healthcare problems, they cannot provide information about the reliability of their predictions [107]. The decision-making in medical applications are mostly life-and-death decisions; so, quantifying reliability of

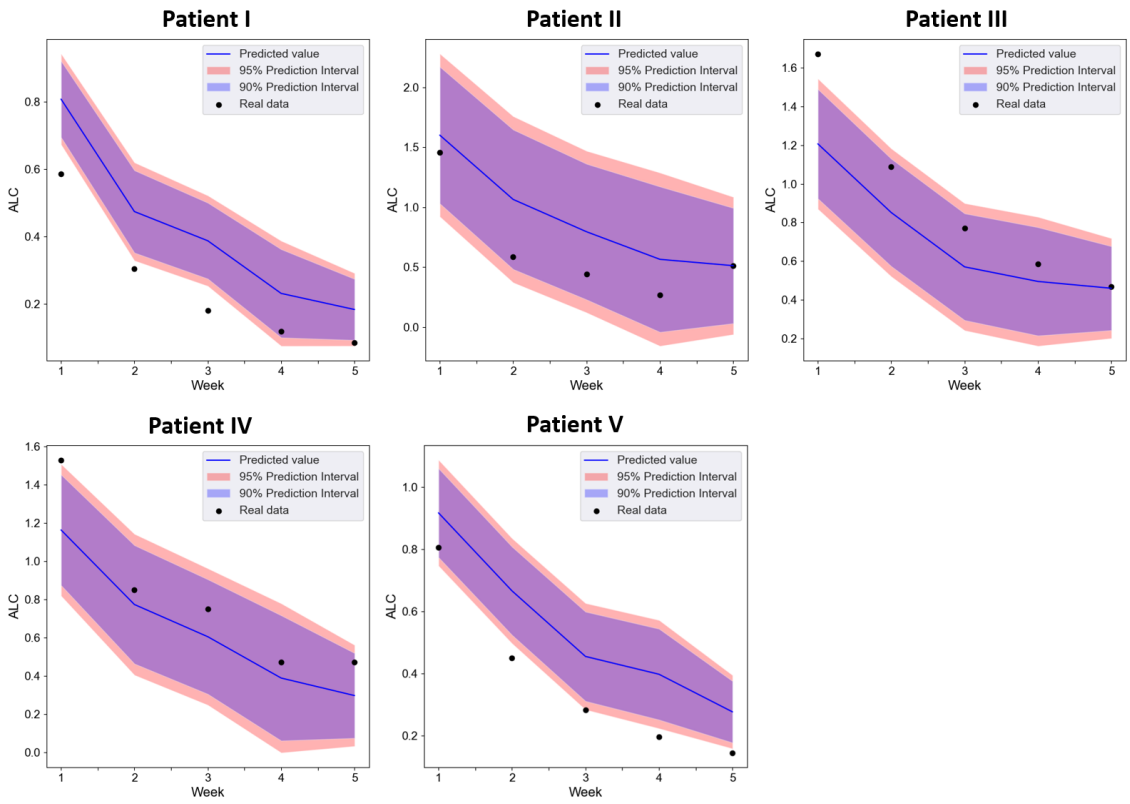


Figure 6.6: The real and predicted weekly ALCs and the associated 95% and 90% prediction intervals for five randomly selected patient examples.

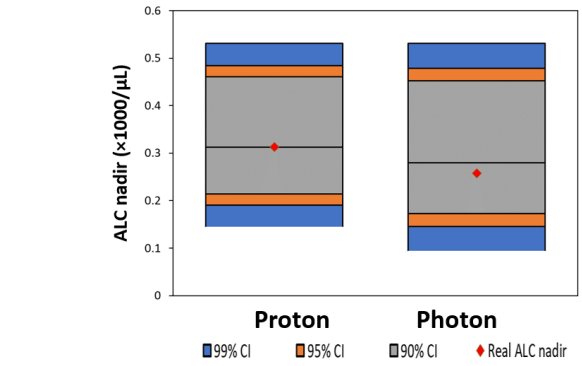
predictions is crucial. However, most efforts in development of prediction models in this field have focused on improving average accuracy of the algorithm, with little consideration of risk management

In this study, we developed a deep learning model to predict the uncertainty of predicted ALC values for each individual patient to be considered in clinics for identifying patients who may experience a large lymphocyte depletion after RT treatment course. The model achieved a good performance in prediction weekly ALC and their associated prediction intervals for patients in the test set. This model can also be used to assess different groups of patients with different pretreatment characteristics in terms of ALC predictions uncertainties. The significance of treatment modality and baseline ALC in predicting G4RIL have been shown in several studies [66, 85, 86, 173]. Therefore, we analyzed the effect of these

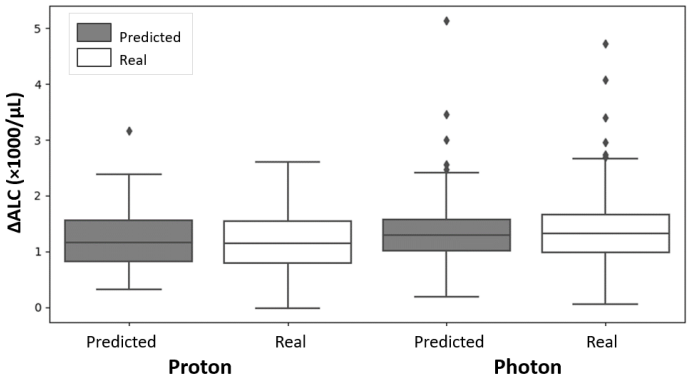
significant parameters on the final predicted ALC nadir and its uncertainty for the patients in the test set.

First, we separated the patients in the test set based on their treatment modality to two groups of proton and photon therapy patients and calculated the mean of predicted ALC nadir, associated prediction intervals, and real ALC nadir for each group. Figure 6.7(a) shows the average 99%, 95%, 90% prediction intervals of predicted ALC for proton therapy and photon therapy patient groups. As we can see from the figure, the mean ALC nadir based on predicted and real values were lower for the photon therapy patients than the proton therapy patients which is consistent with the results of previous studies that photon therapy caused greater lymphocyte depletion compared to the proton therapy treatment. Moreover, the prediction intervals were slightly wider for photon therapy patients as the variability in the real ALC nadir values for photon therapy patients in terms of standard deviation was higher (proton: 0.11, photon: 0.12). For each group of patients, ALC depletion, ΔALC (i.e., baseline ALC – ALC nadir), based on the real ALC values and predictions were calculated and the box plot of the predicted and real ΔALC for patients treated with proton therapy and photon therapy is shown in Figure 6.7(b).

We also explored the effect of baseline ALC value on the final predicted ALC nadir and its uncertainty. We splinted the patients based on the median of baseline ALC into two groups of bottom 50% ($0 < ALC_0 < 1.54$) and top 50% ($1.54 \leq ALC_0$). Table 6.8 summarizes the predicted and real ALC nadir values and their uncertainties in terms of statistical metrics including mean, prediction interval width, median, and standard deviation for these two groups of patients. The average predicted ALC nadirs for patients in the first and second group were 0.25 K/μL and 0.39 K/μL, respectively, which were close to the real measured ALC nadirs of 0.22 K/μL and 0.35 K/μL in each group with relatively small errors. Figure 6.8 shows the average 99%, 95%, 90% prediction intervals of predicted ALC for patients in each group of patients. As we can see from the table and figure, the



(a)



(b)

Figure 6.7: Analysis of RT modality; (a) the average 99%, 95%, 90% prediction intervals of predicted ALC for proton therapy and photon therapy patient groups; (b) the box plot of the predicted and real ΔALC for patients treated with proton therapy and photon therapy.

patients in the top 50% ($1.54 \leq ALC_0$) group, which were the half with the largest ALC baseline values, showed higher mean ALC nadir based on both real data and pretreatment predictions. Also, the prediction interval was the widest for patients in this half compared to another one which can be explained by the higher variance in real ALC nadirs for these patients. Therefore, we can conclude that the model was able to successfully reflect the uncertainty of real data in the uncertainty quantification of predictions. Moreover, we can observe that larger baseline ALC led to larger final ALC nadir predictions which is consistent with the trend in the real ALC nadir values.

Table 6.4: Comparison of important statistical metrics including mean, standard deviation (SD), median, and prediction intervals (PI) for predicted versus real weekly ALC values for the patients within the top 50% and bottom 50% based on ALC baseline

	$ALC_0 \pm SD$	Real ALC nadir		Predicted ALC nadir	
		Mean±SD	Median	Mean, (95% PI)	Median
$0 < ALC_0 < 1.54$	1.14±0.27	0.25±0.14	0.22	0.25, (0.13, 0.37)	0.25
$1.54 \leq ALC_0$	2.10 ±0.62	0.35±0.26	0.28	0.39, (0.19, 0.58)	0.35

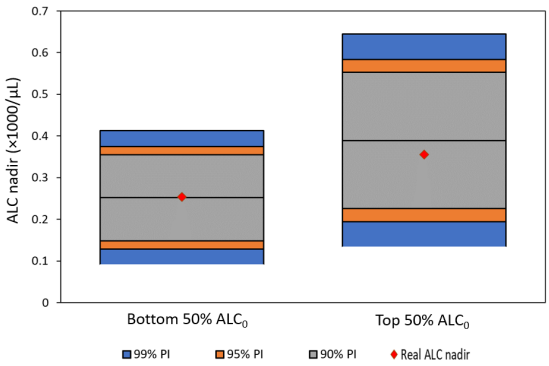


Figure 6.8: The average 99%, 95%, 90% prediction intervals of predicted ALC for patients in the bottom 50% ($0 < ALC_0 < 1.54$) and top 50% ($1.54 \leq ALC_0$) groups based on baseline ALC.

6.4.2 Discussions on Decision-Making Framework

Although deep learning methods showed a good performance in predicting severe RIL but most of these approaches are like a black box and too complicated to be used in clinical practice. Therefore, it is important to make the best use of prediction models in a general decision-making framework that is easy to interpret and use in clinics. Moreover, Ebrahimi et al. [173] reported 70% improvement in MSE of weekly ALC predictions after augmenting the measured ALC at the end of the first week of RT treatment. However, it is not feasible to begin the treatment for all patients and measure the ALC in clinical practice as it is time and cost consuming and may have adverse consequences for high-risk patients. Therefore, it is important to identify highly uncertain predictions, which need to be validated or corrected after delivering initial part of the treatment, for some patients. In this study, A decision-making framework was proposed with the goal of providing the best use

of prediction models in clinical practice by making reliable decisions. In this regard, this framework detects the pretreatment predictions with high uncertainty and improves them by delivering a portion of treatment to a patient and using the measured data at early stages of the treatment.

As mentioned in the results section, the MSE of weekly ALC predictions were 0.054 and 0.017 based on the predictions using the pretreatment data for all patients and early treatment prediction after the first week of treatment for all patients. Using the decision-making framework resulted in MSE of 0.024 which was 55% lower than the pretreatment predictions and 41% higher than the predictions after the week 1. Although the MSE is still higher than the predictions made after the first week but this approach is more applicable in clinical practice as it can reduce time, cost, and effort by using early treatment predictions only for the patients with high prediction uncertainties.

Our results showed that the proposed decision-making framework showed a good performance in identifying patients at high risk of G4RIL. Therefore, it can be used to evaluate different RT treatment plans for an individual patient with the goal of minimizing the G4RIL risk and improving the patient survival. Using this decision-making framework enables the physicians to predict the effect of any possible modification in the RT treatment plan which can benefit the patient. For example, several studies reported that proton therapy treatments showed a lower risk of G4RIL than photon therapy patients [66, 161]. Therefore, changing the treatment modality from photon to proton could be a potential treatment modification that may help reduce G4RIL risk for a patient.

In this study, we evaluated the impact of changing the treatment modality to proton therapy, for photon therapy patients who were identified at high risk of G4RIL based on the decision-making framework. The equivalent proton therapy treatment plans were created based on the same dosimetric criteria and prescription dose to the clinical target volume

(CTV) of 50.4 Gy in 28 fractions using MatRad [156], a research-oriented treatment planning system. For optimization of proton therapy (i.e., IMPT) plans, we used a constant relative biological effectiveness (RBE) of 1.1 [176] and all plans were normalized to have 95% of the planning target volume (PTV) receive the prescription dose. Table 6.5 presents the summary of treatment and patient characteristics as well as the prediction results for each patient based on proton and photon treatment modalities.

Table 6.5: Important treatment- and patient-related features and prediction results for three patients that selected by decision-making framework for treatment modifications

		Patient 1		**Patient 2**		**Patient 3**	
Patient characteristics	**Age**	68		53		56	
	BMI	26.30		34.25		28.57	
	ALC0	1.57		2.42		2.10	
	Real ALC nadir	0.03		0.07		0.06	
	G4RIL	Yes		Yes		Yes	
Treatment characteristics	**Treatment Modality**	IMRT	IMPT	IMRT	IMPT	IMRT	IMPT
	Mean lung dose	18.17	12.61	13.43	7.04	14.13	9.19
	Mean heart dose	38.36	10.95	34.41	20.20	40.42	23.27
	Mean spleen dose	13.92	11.59	38.27	16.21	22.33	13.98
Prediction results	**Predicted ALC nadir**	0.18	0.31	0.17	0.48	0.22	0.40
	95% Prediction Intervals	[0.06, 0.30]	[0.22, 0.41]	[0.01, 0.34]	[0.26, 0.70]	[0.04, 0.4]	[0.22, 0.59]
	Decision	Avoid treatment	Begin treatment	Avoid treatment	Begin treatment	Avoid treatment	Begin treatment

As shown in the table, changing the treatment modality from IMRT to IMPT for patient 1 reduced the mean dose to three considered organs at risk (OARs) of lung, heart, and spleen. Therefore, we changed the treatment modality and all dosimetric features in the data to make new predictions for weekly ALCs during the course of RT treatment. The modified treatment plan achieved higher predicted weekly ALCs and the decision-making framework changed the decision for this patient based on the predicted ALC nadir and its prediction intervals. The results suggested that this patient would benefit from changing the treatment modality from IMRT to IMPT. We can also see the same observation for patient 2 and 3. Therefore, we can conclude that the proposed decision-making framework

performed well in discriminating different treatment modalities and showing the effect of treatment modifications on treatment outcome. This will help the physicians to evaluate several treatment plans (e.g., different treatment modalities) for an individual patient and select the one with the lowest risk of G4RIL. Figure 6.9 also shows the real weekly ALC, predicted weekly ALCs and the associated 95% and 90% prediction intervals for both IMRT and IMPT modalities for the three patient cases.

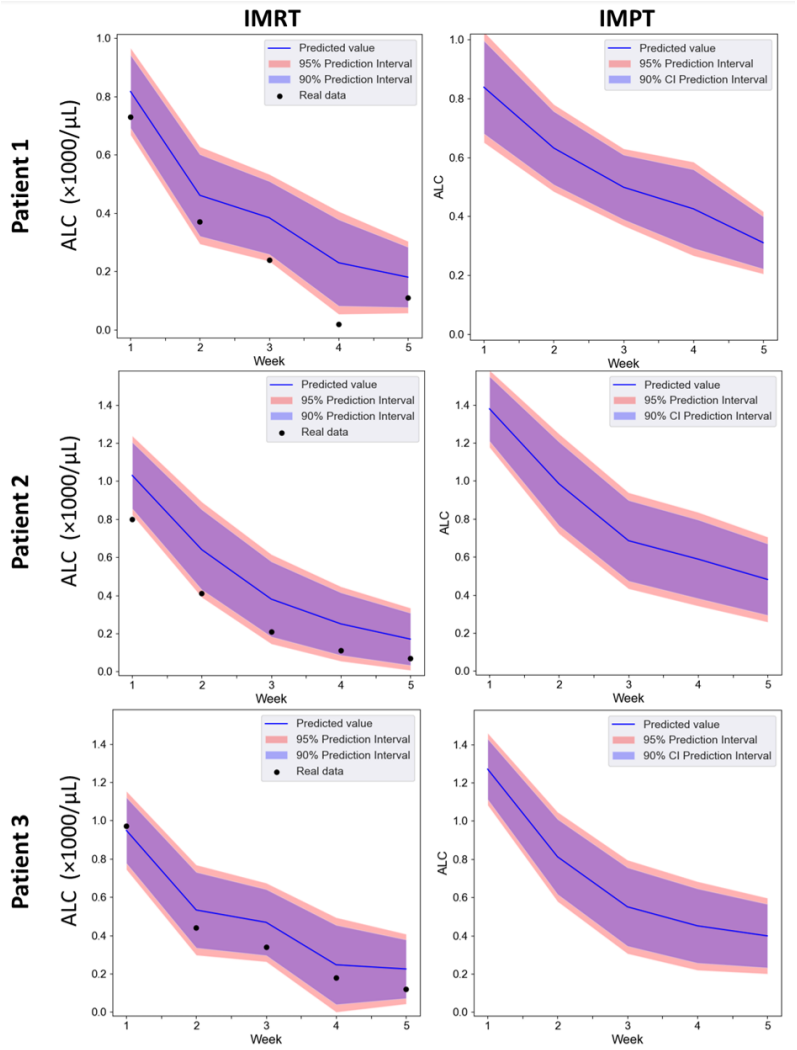


Figure 6.9: The real weekly ALC, predicted weekly ALCs and the associated 95% and 90% prediction intervals for both IMRT and IMPT modalities for the three randomly selected patient cases

6.4.3 Directions for Further Research

Although our proposed decision-making approach has achieved a good performance in identifying patients at high risk of G4RIL compared to other classification methods, there is still some room for future work. First, although the size of our data set might be sufficient to validate the rationale of the proposed model, it would be ideal to use data from a different institution and cancer type to further evaluate the robustness of our approach. Secondly, this study motivates further studies to investigate the effect of treatment modifications in different aspects such as changing treatment modality, beam angles, using variable fractionation and adaptive RT, using the decision-making framework for a large number of patients. Third, we assumed that the ALC predictions have a normal distribution and we estimated the parameters of a normal distribution for each predicted ALC to account for prediction uncertainties. So, alternative distributions such as Weibull and Log normal can be tested to investigate the effect different distribution on the prediction intervals. Moreover, as Bayesian deep learning approaches showed a good performance in uncertainty quantification of prediction models; In future work, we will also investigate the benefits of using Bayesian deep learning approaches to improve prediction RIL risk for RT cancer patients.

6.5 Conclusion

In this study, we proposed a deep learning approach to predict the weekly ALC values and uncertainty associated with each predicted ALC value for individual patients to be used in estimation of prediction intervals. The proposed model performed well in predicting radiation-induced lymphocyte depletion and predictive uncertainty quantification. The

ability to estimate the uncertainty of predictions will enable physicians to assess the individual predictions risk in clinical practice. Moreover, we proposed a novel hybrid decision-making framework for selecting patients for RT treatment using the predicted values of ALC and their associated risk. Our results showed that the decision-making framework yielded a good performance in classifying patients at-high risk of G4RIL. This decision-making framework can help physicians to easily take the advantage of complex deep learning models in their decisions for each individual patient and identify patients who are at high risk of severe RIL and who would benefit from treatment replanning, use of different modality, or developing mitigations strategies. Furthermore, this approach is flexible, adjustable based on clinical preferences, and can be transferred to predict other related toxic effects of RT.

Chapter 7

Summary and Future Work

In summary, the primary goal of this dissertation was to address the current challenges of radiation therapy treatment planning and provide practical and reliable solutions to improve treatment quality and patient survival.

In Chapter 3, a biological-based treatment planning framework was proposed such that it not only controls the biological aspect of the treatment and incorporates the uncertainty of the tumor's biological response, but also ensures the dose-volume requirements and clinical limits of the treatment without needing to deal with complex optimization models. The proposed reinforcement learning framework for ART planning can help the decision-maker to achieve a robust solution under high levels of uncertainty in the biological parameters while reducing the variability in the solution and improving the control on the worst cases. Furthermore, using the proposed comprehensive biological response model to estimate tumor volume regressions can reduce the time and effort from collecting large-scale data sets and avoid the need for taking expensive CT images at each visit. The performance of the proposed RT treatment planning framework was tested using a clinical non-small cell lung cancer (NSCLC) case. The results were compared with the conventional fractionation schedule (i.e, equal dose fractionation) as a reference plan. The results showed that the proposed approach performed well in achieving a robust optimal ART treatment

plan under high uncertainty in the biological parameters. This approach enables the physicians to find an appropriate personalized ART plan in terms of dose fractionation and the timing of the adaptations. It also can be used to identify patients who would benefit from ART as an alternative to the conventional equal-dose plan. The proposed approach is flexible enough to support a wide range of treatment objectives and preferences for different decision-makers and various cancer types. Moreover, the contribution of this chapter can be extended by adjusting the reinforcement learning environment to account for another radiation therapy treatment planning problem, such as radiation beam angle optimization. In this case, the environment should reflect variable tumor response to radiation based on different sets of beam angles and the agent aims to find the optimal set of beam angles which maximize treatment effectiveness and patient survival considering the biological uncertainties. Also, robust optimization techniques can be used in a beamlet optimization model to handle physical uncertainties while controlling the biological uncertainties within the reinforcement learning framework.

In Chapter 4, we addressed the role of RT in immunosuppression caused by significant lymphocyte loss during the course of radiation therapy treatment. Two mathematical models were proposed to approximate lymphocyte depletion based on radiation dose distributions for cancer patients. In the first model, we use a piecewise-linear relationship between lymphocyte survival and radiation dose. The second model assumes an exponential function for ALC depletion, and it uses a non-linear regression to estimate post-treatment ALC. Moreover, the impact of different radiation modalities and dose distributions on developing severe lymphopenia after the treatment were assessed for ten esophageal cancer patients and the potential post-treatment lymphocyte survival outcomes based on the proposed models were compared for photon and proton-based modalities. Results showed significant lymphocyte reduction is associated with treatment modalities and proton plans

outperformed photon plans in terms of lymphocyte preservation. This study motivates further research to investigate the clinical factors that affect RIL risk of different radiation modalities. With help of research on continuing better understanding of lymphocyte distribution throughout the treatment field, the radiation dose could be optimized accordingly to avoid lymphocyte killing. Moreover, additional immune protection could be possible by optimizing plans with constraints on dosages received by volumes of the body (and immune organs at risk such as the spleen, heart, etc.). Such methods to enhance the ability of treatment plans to minimize lymphopenia risk without compromising tumor coverage and other normal tissues at risk should be studied in future research into this area of study.

In Chapter 5, a novel hybrid deep learning model in a stacked structure was proposed to predict RT-induced lymphocyte depletion during the course of RT treatment for cancer patients based on the pretreatment clinical information. The proposed model structure consisted of four channels, one channel based on a long short-term memory (LSTM) network and three channels based on deep neural networks, to process four categories of features followed by a fully connected neural network to integrate the outputs of four channels and predict the weekly ALC values. Using this structure, we can efficiently use information from different groups of features with different characteristics to predict weekly ALC without requiring a large amount of data to process too many features while reducing bias and the adverse effects of any noise in the data. As a result, this approach can help the physicians to identify high-risk patients and select them for modified treatment approaches or mitigation strategies. The proposed model was trained and tested on a data set of 860 esophageal cancer patients who received RT treatment. First, the model was used to predict weekly ALC values during RT treatment course based on pretreatment information for a cohort of esophageal cancer patients. Then, the model was extended to account for predictions made after the initial part of treatment (i.e., at the end of weeks 1 or 2). So, a discriminative kernel was developed to extract temporal features and assign different

weights to each part of the input sequence which enables the model to focus on the most relevant parts. In order to evaluate the performance of the proposed models, we compared the model with 8 commonly used prediction models by calculating important prediction metrics. The results showed that the proposed model outperformed these off-the-shelf prediction methods in predicting weekly ALC values. Moreover, using the extended model based on the first-week data reduced the MSE of predictions compared to the model based on the pretreatment data. Therefore, we can conclude that augmenting the model with data from early stages of treatment (i.e., weeks 1 or 2) can significantly improve ALC predictions. Unlike most deep learning models, the proposed stacked structure is interpretable, and it can provide the weights of each feature group in making the final prediction. Although our proposed model achieved a good prediction performance compared to other state-of-art prediction methods, there is still some room for future work. As the next step, we investigated the benefits of using probabilistic deep learning approaches to estimate the risk of prediction for individual patients in Chapter 6. Another way to extend the contribution of this chapter is to take advantage of clustering methods to define clusters of patients based on the similarity in clinical characteristics and consider them as a feature to be fed into the deep learning model. Moreover, further analysis and risk assessments should be made based on different patient profiles to investigate the impact of different clinical factors on RIL and to estimate the risk associated with each predicted ALC value.

In Chapter 6, we extended our proposed deep learning model and provided an approach to predict the weekly ALC values and their associated uncertainties in form of prediction intervals for individual patients. The performance of the model was evaluated using a data set of esophageal cancer patients. Furthermore, the effect of different patient- and treatment-specific factors on RIL risk were assessed. The results showed that the proposed model performed well in predicting radiation-induced lymphocyte depletion and predictive uncertainty quantification. Estimation of prediction intervals for a given individual patient

enables practical implications of predictive models in clinical decision-making by considering individualized prediction risks. Next, we proposed a comprehensive hybrid decision-making framework to select patients for RT treatment based on the ALC prediction and its predictive uncertainty for a given patient with the goal of minimizing severe RIL risk for the patient. The decision-making framework was tested to identify high-risk patients using real data under different scenarios and compared with commonly used classification methods. Our results showed that the decision-making framework yielded a good performance in classifying patients at high-risk of severe RIL. The proposed decision-making framework can help physicians to take the advantage of complex deep learning models in their decisions and identify high-risk patients who may benefit from treatment modifications. Also, the effect of any treatment modifications on RIL risk for a given patient can be evaluated and the safest treatment plan can be chosen for the patient. Furthermore, this approach is flexible based on clinical preferences, and can be transferred to predict other related toxic effects of RT. For future work, one can combine the beamlet intensity optimization with this decision-making framework to find the optimal beamlet intensities for a given treatment plan. Therefore, we can generate a personalized treatment plan that meets dose-volume requirements while minimizing the RIL risk for the patient. Moreover, further statistical analysis should be made to investigate the impact of uncertainty in each pretreatment clinical feature on the estimated uncertainty of predicted ALC values. As a result, physicians can understand the effect of noise and uncertain data on performance of prediction models.

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: A Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7–30, 2020.

[2] H. R. Withers, "The four R's of radiotherapy," in *Advances in radiation biology*. Elsevier, 1975, vol. 5, pp. 241–271.

[3] G. G. Steel, T. J. McMillan, and J. Peacock, "The 5Rs of radiobiology," *International journal of radiation biology*, vol. 56, no. 6, pp. 1045–1048, 1989.

[4] S. Khaled and K. D. Held, "Radiation biology: a handbook for teachers and students," 2012.

[5] J.-E. Bibault, I. Fumagalli, C. Ferté, C. Chargari, J.-C. Soria, and E. Deutsch, "Personalized radiation therapy and biomarker-driven treatment strategies: a systematic review," *Cancer Metastasis Reviews*, vol. 32, no. 3-4, pp. 479–492, Dec. 2013.

[6] B. Douglas and J. Fowler, "The effect of multiple small doses of x rays on skin reactions in the mouse and a basic interpretation," *Radiation research*, vol. 178, no. 2, pp. AV125–AV138, 2012.

[7] J. F. Fowler, "The linear-quadratic formula and progress in fractionated radiotherapy," *The British Journal of Radiology*, vol. 62, no. 740, pp. 679–694, 1989.

[8] J. Fowler, "Biological factors influencing optimum fractionation in radiation therapy," *Acta oncologica*, vol. 40, no. 6, pp. 712–717, 2001.

[9] F. Saberian, A. Ghate, and M. Kim, "Optimal fractionation in radiotherapy with multiple normal tissues," *Mathematical medicine and biology: a journal of the IMA*, vol. 33, no. 2, pp. 211–252, 2016.

[10] T. Bortfeld, J. Ramakrishnan, J. N. Tsitsiklis, and J. Unkelbach, "Optimization of radiation therapy fractionation schedules in the presence of tumor repopulation," *INFORMS Journal on Computing*, vol. 27, no. 4, pp. 788–803, 2015.

[11] T. Wheldon and J. Kirk, "Mathematical derivation of optimal treatment schedules for the radiotherapy of human tumours. fractionated irradiation of exponentially growing tumours," *The British journal of radiology*, vol. 49, no. 581, pp. 441–449, 1976.

[12] T. Wheldon, J. Kirk, and J. Orr, "Optimal radiotherapy of tumour cells following exponential-quadratic survival curves and exponential repopulation kinetics," *The British journal of radiology*, vol. 50, no. 597, pp. 681–682, 1977.

[13] C. Armpilia, R. Dale, and B. Jones, "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*, vol. 77, no. 921, pp. 765–767, 2004.

[14] G. Brunton and T. Wheldon, "Characteristic species dependent growth patterns of mammalian neoplasms," *Cell Proliferation*, vol. 11, no. 2, pp. 161–175, 1978.

[15] H. McAneney and S. O'Rourke, "Investigation of various growth mechanisms of solid tumour growth within the linear-quadratic model for radiotherapy," *Physics in Medicine & Biology*, vol. 52, no. 4, p. 1039, 2007.

[16] D. Corwin, C. Holdsworth, R. C. Rockne, A. D. Trister, M. M. Mrugala, J. K. Rockhill, R. D. Stewart, M. Phillips, and K. R. Swanson, "Toward patient-specific, biologically optimized radiation therapy plans for the treatment of glioblastoma," *PLoS ONE*, vol. 8, no. 11, 2013.

[17] C. H. Holdsworth, D. Corwin, R. D. Stewart, R. Rockne, A. D. Trister, K. R. Swanson, and M. Phillips, "Adaptive IMRT using a multiobjective evolutionary algorithm integrated with a diffusion-invasion model of glioblastoma," *Physics in medicine and biology*, vol. 57, no. 24, pp. 8271–8283, 2012.

[18] M. Le, H. Delingette, J. Kalpathy-Cramer, E. R. Gerstner, T. Batchelor, J. Unkelbach, and N. Ayache, "Personalized radiotherapy planning based on a computational tumor growth model," *IEEE Transactions on Medical Imaging*, vol. 36, no. 3, pp. 815–825, 2017.

[19] H. Lee, Y. C. Ahn, D. Oh, H. Nam, Y. I. Kim, and S. Y. Park, "Tumor volume reduction rate measured during adaptive definitive radiation therapy as a potential prognosticator of locoregional control in patients with oropharyngeal cancer," *Head & Neck*, vol. 36, no. 4, pp. 499–504, 2014.

[20] D. J. Brenner, L. R. Hlatky, P. J. Hahnfeldt, E. J. Hall, and R. K. Sachs, "A convenient extension of the linear-quadratic model to include redistribution and reoxygenation," *International Journal of Radiation Oncology, Biology, Physics*, vol. 32, no. 2, pp. 379–390, May 1995.

[21] Y. Yang and L. Xing, "Optimization of radiotherapy dose-time fractionation with consideration of tumor specific biology," *Medical physics*, vol. 32, no. 12, pp. 3666–3677, 2005.

[22] J. Jeong, K. Shoghi, and J. Deasy, "Modelling the interplay between hypoxia and proliferation in radiotherapy tumour response," *Physics in Medicine & Biology*, vol. 58, no. 14, p. 4897, 2013.

[23] Y. R. Lawrence, M. Werner-Wasik, and A. P. Dicker, "Biologically conformal treatment: biomarkers and functional imaging in radiation oncology," *Future Oncology*, 2008.

[24] C. South, M. Partridge, and P. Evans, "A theoretical framework for prescribing radiotherapy dose distributions using patient-specific biological information," *Medical physics*, vol. 35, no. 10, pp. 4599–4611, 2008.

[25] M. Kim, A. Ghate, and M. H. Phillips, "A stochastic control formalism for dynamic biologically conformal radiation therapy," *European Journal of Operational Research*, vol. 219, no. 3, pp. 541–556, 2012.

[26] L. M. Wein, J. E. Cohen, and J. T. Wu, "Dynamic optimization of a linear–quadratic model with incomplete repair and volume-dependent sensitivity and repopulation," *International Journal of Radiation Oncology • Biology • Physics*, vol. 47, no. 4, pp. 1073–1083, Jul. 2000.

[27] M. Kim, A. Ghate, and M. H. Phillips, "A Markov decision process approach to temporal modulation of dose fractions in radiation therapy planning," *Physics in Medicine and Biology*, vol. 54, no. 14, pp. 4455–4476, Jun. 2009.

[28] A. Ghate, "Dynamic Optimization in Radiotherapy," in *Transforming Research into Action*, ser. INFORMS TutORials in Operations Research. INFORMS, Sep. 2011, pp. 60–74.

[29] J. Unkelbach, D. Craft, T. Hong, D. Papp, J. Ramakrishnan, E. Salari, J. Wolfgang,

and T. Bortfeld, "Exploiting tumor shrinkage through temporal optimization of radiotherapy," *Physics in medicine and biology*, vol. 59, no. 12, pp. 3059–3079, 2014.

[30] O. Nohadani and A. Roy, "Robust optimization with time-dependent uncertainty in radiation therapy," *IISE Transactions on Healthcare Systems Engineering*, vol. 7, no. 2, pp. 81–92, 2017.

[31] J. G. Scott, A. Berglund, M. J. Schell, I. Mihaylov, W. J. Fulp, B. Yue, E. Welsh, J. J. Caudell, K. Ahmed, T. S. Strom, E. Mellon, P. Venkat, P. Johnstone, J. Foekens, J. Lee, E. Moros, W. S. Dalton, S. A. Eschrich, H. McLeod, L. Harrison, and J. F. Torres-Roca, "A genome-based model for adjusting radiotherapy dose (gard): a retrospective, cohort-based study," *The Lancet Oncology*, vol. 18, no. 2, pp. 202–211, 2017.

[32] Z. Huang, N. A. Mayr, W. T. Yuh, S. S. Lo, J. F. Montebello, J. C. Grecula, L. Lu, K. Li, H. Zhang, N. Gupta, and J. Z. Wang, "Predicting outcomes in cervical cancer: a kinetic model of tumor regression during radiation therapy," *Cancer research*, vol. 70, no. 2, pp. 463–470, 2010.

[33] A. J. van de Schoot, P. de Boer, J. Visser, L. J. A. Stalpers, C. R. N. Rasch, and A. Bel, "Dosimetric advantages of a clinical daily adaptive plan selection strategy compared with a non-adaptive strategy in cervical cancer radiation therapy." *Acta oncologica*, vol. 56 5, pp. 667–674, 2017.

[34] S. Ramella, M. Fiore, S. Silipigni, M. C. Zappa, M. Jaus, A. M. Alberti, P. Matteucci, E. Molfese, P. Cornacchione, C. Greco, L. Trodella, E. Ippolito, and R. M. D'Angelillo, "Local control and toxicity of adaptive radiotherapy using weekly CT imaging: Results from the LARTIA trial in stage III NSCLC," *Journal of Thoracic Oncology*, vol. 12, no. 7, pp. 1122–1130, 2017.

[35] A. Belfatto, M. Riboldi, D. Ciardo, A. Cecconi, R. Lazzari, B. A. Jereczek, R. Orecchia, G. Baroni, and P. Cerveri, "Adaptive mathematical model of tumor response to radiotherapy based on CBCT data," *IEEE journal of biomedical and health informatics*, vol. 20, no. 3, pp. 802–809, 2016.

[36] X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.-F. Yin, and Q. J. Wu, "A planning quality evaluation tool for prostate adaptive IMRT based on machine learning," *Medical Physics*, vol. 38, no. 2, pp. 719–726, 2011.

[37] C. Dial, E. Weiss, J. V. Siebers, and G. D. Hugo, "Benefits of adaptive radiation therapy in lung cancer as a function of replanning frequency," *Medical Physics*, vol. 43, no. 4, p. 1787, 2016.

[38] O. Veresezan, I. Troussier, A. Lacout, S. Kreps, S. Maillard, A. Toulemonde, P.-Y. Marcy, F. Huguet, and J. Thariat, "Adaptive radiation therapy in head and neck cancer for clinical practice: state of the art and practical challenges," *Japanese Journal of Radiology*, vol. 35, no. 2, pp. 43–52, 2017.

[39] B. Saka, R. L. Rardin, M. P. Langer, and D. Dink, "Adaptive intensity modulated radiation therapy planning optimization with changing tumor geometry and fraction size limits," *IIE Transactions on Healthcare Systems Engineering*, vol. 1, no. 4, pp. 247–263, 2011.

[40] M. Guckenberger, A. Richter, J. Wilbert, M. Flentje, and M. Partridge, "Adaptive radiotherapy for locally advanced non–small-cell lung cancer does not underdose the microscopic disease and has the potential to increase tumor control," *International Journal of Radiation Oncology* Biology* Physics*, vol. 81, no. 4, pp. e275–e282, 2011.

[41] Y. Zheng, H. Singh, L. Zhao, E. V. Ramirez, S. Rana, K. Prabhu, L. S. Doh, and G. L. Larson, "Adaptive radiation therapy for lung cancer using uniform scanning proton beams: Adaptation strategies, practical considerations, and clinical outcomes," *International Journal of Radiation Oncology • Biology • Physics*, vol. 93, no. 3, p. S29, 2015.

[42] P. Berkovic, L. Paelinck, Y. Lievens, A. Gulyban, B. Goddeeris, C. Derie, V. Surmont, W. D. Neve, and K. Vandecasteele, "Adaptive radiotherapy for locally advanced non-small cell lung cancer, can we predict when and for whom?" *Acta Oncologica*, vol. 54, no. 9, pp. 1438–1444, 2015.

[43] M. Zarepisheh, T. Long, N. Li, Z. Tian, H. E. Romeijn, X. Jia, and S. B. Jiang, "A dvh-guided imrt optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning," *Medical physics*, vol. 41, no. 6Part1, p. 061711, 2014.

[44] Y. Kawata, H. Arimura, K. Ikushima, Z. Jin, K. Morita, C. Tokunaga, H. Yabu-uchi, Y. Shioyama, T. Sasaki, H. Honda, and M. Sasaki, "Impact of pixel-based machine-learning techniques on automated frameworks for delineation of gross tumor volume regions for stereotactic body radiation therapy," *Physica Medica*, vol. 42, pp. 141–149, 2017.

[45] P. Zhang, A. Rimner, E. Yorke, J. Hu, B. Ravindranath, G. Mageras, and J. Deasy, "Predicting spatial distribution of residual tumor post radiation therapy based on pretreatment pet/ct for locally advanced non-small cell lung cancer," *International Journal of Radiation Oncology• Biology• Physics*, vol. 93, no. 3, p. E557, 2015.

[46] M. Surucu, K. K. Shah, I. Mescioglu, J. C. Roeske, W. Small Jr, M. Choi, and B. Emami, "Decision trees predicting tumor shrinkage for head and neck cancer:

Implications for adaptive radiotherapy," *Technology in cancer research & treatment*, vol. 15, no. 1, pp. 139–145, 2016.

[47] G. Guidi, N. Maffei, B. Meduri, E. D'Angelo, G. Mistretta, P. C. A. Ciarmatori, A. Bernabei, S. Maggi, M. Cardinali, V. Morabito, F. Rosica, S. Malara, A. Savini, G. Orlandi, C. D'Ugo, F. Bunkheila, M. Bono, S. Lappi, C. Blasi, F. Lohr, and T. Costi, "A machine learning tool for re-planning and adaptive RT: a multicenter cohort investigation," *Physica Medica*, vol. 32, no. 12, pp. 1659–1666, 2016.

[48] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1238–1274, 2013.

[49] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.

[50] M. Glavic, R. Fonteneau, and D. Ernst, "Reinforcement learning for electric power system decision and control: Past considerations and perspectives," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 6918–6927, 2017.

[51] Z. Wen, D. O'Neill, and H. Maei, "Optimal demand response using device-based reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2312–2324, 2015.

[52] Y. Ling, S. A. Hasan, V. Datla, A. Qadir, K. Lee, J. Liu, and O. Farri, "Diagnostic inferencing via improving clinical concept extraction with deep reinforcement learning: A preliminary study," in *Machine Learning for Healthcare Conference*, 2017, pp. 271–285.

[53] H.-H. Tseng, Y. Luo, S. Cui, J.-T. Chien, R. K. Ten Haken, and I. El Naqa, "Deep reinforcement learning for automated radiation adaptation in lung cancer," *Medical physics*, vol. 44, no. 12, pp. 6690–6705, 2017.

[54] I. Fox and J. Wiens, "Reinforcement learning for blood glucose control: Challenges and opportunities," 2019.

[55] M. Tejedor, A. Z. Woldaregay, and F. Godtliebsen, "Reinforcement learning application in diabetes blood glucose control: A systematic review," *Artificial Intelligence in Medicine*, vol. 104, p. 101836, 2020.

[56] C. Yu, Y. Dong, J. Liu, and G. Ren, "Incorporating causal factors into reinforcement learning for dynamic treatment regimes in hiv," *BMC medical informatics and decision making*, vol. 19, no. 2, pp. 19–29, 2019.

[57] Z. Liu, C. Yao, H. Yu, and T. Wu, "Deep reinforcement learning with its application for lung cancer detection in medical internet of things," *Future Generation Computer Systems*, vol. 97, pp. 1–9, 2019.

[58] J. Stember and H. Shalu, "Deep reinforcement learning to detect brain lesions on mri: a proof-of-concept application of reinforcement learning to medical images," *arXiv preprint arXiv:2008.02708*, 2020.

[59] C. Yu, J. Liu, and S. Nemati, "Reinforcement learning in healthcare: A survey," *arXiv preprint arXiv:1908.08796*, 2019.

[60] A. Coronato, M. Naeem, G. De Pietro, and G. Paragliola, "Reinforcement learning for intelligent healthcare applications: A survey," *Artificial Intelligence in Medicine*, vol. 109, p. 101964, 2020.

[61] M. Naeem, S. T. H. Rizvi, and A. Coronato, "A gentle introduction to reinforcement learning and its application in different fields," *IEEE Access*, 2020.

[62] J. Futoma, A. Lin, M. Sendak, A. Bedoya, M. Clement, C. O'Brien, and K. Heller, "Learning to treat sepsis with multi-output gaussian process deep recurrent q-networks," 2018.

[63] B. K. Petersen, J. Yang, W. S. Grathwohl, C. Cockrell, C. Santiago, G. An, and D. M. Faissol, "Precision medicine as a control problem: Using simulation and deep reinforcement learning to discover adaptive, personalized multi-cytokine therapy for sepsis," *arXiv preprint arXiv:1802.10440*, 2018.

[64] Y. Li, W. Zhang, C.-X. Wang, J. Sun, and Y. Liu, "Deep reinforcement learning for dynamic spectrum sensing and aggregation in multi-channel wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, pp. 464–475, 2020.

[65] K. S. Sellins and J. J. Cohen, "Gene induction by gamma-irradiation leads to dna fragmentation in lymphocytes." *The Journal of Immunology*, vol. 139, no. 10, pp. 3199–3206, 1987.

[66] Y. Shiraishi, P. Fang, C. Xu, J. Song, S. Krishnan, E. J. Koay, R. J. Mehran, W. L. Hofstetter, M. Blum-Murphy, J. A. Ajani, R. Komaki, B. Minsky, R. Mohan, B. P. Hobbs, and S. H. Lin, "Severe lymphopenia during neoadjuvant chemoradiation for esophageal cancer: A propensity matched analysis of the relative risk of proton versus photon-based radiation therapy," *Radiotherapy and Oncology*, vol. 128, no. 1, pp. 154–160, 2018.

[67] K. Holub, A. Vargas, and A. Biete, "Radiation-induced lymphopenia: the main aspects to consider in immunotherapy trials for endometrial and cervical cancer patients," *Clinical and Translational Oncology*, pp. 1–9, 2020.

[68] S. G. Ellsworth, "Field size effects on the risk and severity of treatment-induced lymphopenia in patients undergoing radiation therapy for solid tumors," *Advances in radiation oncology*, vol. 3, no. 4, pp. 512–519, 2018.

[69] J. L. Campian, X. Ye, M. Brock, and S. A. Grossman, "Treatment-related lymphopenia in patients with stage iii non-small-cell lung cancer," *Cancer investigation*, vol. 31, no. 3, pp. 183–188, 2013.

[70] S. A. Grossman, X. Ye, G. Lesser, A. Sloan, H. Carraway, S. Desideri, and S. Piantadosi, "Immunosuppression in patients with high-grade gliomas treated with radiation and temozolomide," *Clinical Cancer Research*, vol. 17, no. 16, pp. 5473–5480, 2011.

[71] B. P. Venkatesulu, S. Mallick, S. H. Lin, and S. Krishnan, "A systematic review of the influence of radiation-induced lymphopenia on survival outcomes in solid tumors," *Critical reviews in oncology/hematology*, vol. 123, pp. 42–51, 2018.

[72] O. Cho, M. Chun, S.-J. Chang, Y.-T. Oh, and O. K. Noh, "Prognostic value of severe lymphopenia during pelvic concurrent chemoradiotherapy in cervical cancer," *Anticancer research*, vol. 36, no. 7, pp. 3541–3547, 2016.

[73] E. S. Wu, T. Oduyebo, L. P. Cobb, D. Cholakian, X. Kong, A. N. Fader, K. L. Levinson, E. J. Tanner III, R. L. Stone, A. Piotrowski, S. Grossman, and K. L. Roche, "Lymphopenia and its association with survival in patients with locally advanced cervical cancer," *Gynecologic oncology*, vol. 140, no. 1, pp. 76–82, 2016.

[74] A. Balmanoukian, X. Ye, J. Herman, D. Laheru, and S. A. Grossman, "The association between treatment-related lymphopenia and survival in newly diagnosed patients with resected adenocarcinoma of the pancreas," *Cancer investigation*, vol. 30, no. 8, pp. 571–576, 2012.

[75] A. T. Wild, X. Ye, S. G. Ellsworth, J. A. Smith, A. K. Narang, T. Garg, J. Campian, D. A. Laheru, L. Zheng, C. L. Wolfgang, P. T. Tran, S. A. Grossman, and J. M. Herman, "The association between chemoradiation-related lymphopenia and clinical outcomes in patients with locally advanced pancreatic adenocarcinoma," *American journal of clinical oncology*, vol. 38, no. 3, p. 259, 2015.

[76] J. L. Campian, X. Ye, G. Sarai, J. Herman, and S. A. Grossman, "Severe treatment-related lymphopenia in patients with newly diagnosed rectal cancer," *Cancer investigation*, vol. 36, no. 6, pp. 356–361, 2018.

[77] C. Tang, Z. Liao, D. Gomez, L. Levy, Y. Zhuang, R. A. Gebremichael, D. S. Hong, R. Komaki, and J. W. Welsh, "Lymphopenia association with gross tumor volume and lung v5 and its effects on non-small cell lung cancer patient outcomes," *International Journal of Radiation Oncology* Biology* Physics*, vol. 89, no. 5, pp. 1084–1091, 2014.

[78] R. Davuluri, W. Jiang, P. Fang, C. Xu, R. Komaki, D. R. Gomez, J. Welsh, J. D. Cox, C. H. Crane, C. C. Hsu, and S. H. Lin, "Lymphocyte nadir and esophageal cancer survival outcomes after chemoradiation therapy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 99, no. 1, pp. 128–135, 2017.

[79] D. M. Routman, A. Garant, S. C. Lester, C. N. Day, W. S. Harmsen, C. T. Sanheuza, H. H. Yoon, M. A. Neben-Wittich, J. A. Martenson, M. G. Haddock, C. L.

Hallemeier, and K. W. Merrell, "A comparison of grade 4 lymphopenia with proton versus photon radiation therapy for esophageal cancer," *Advances in radiation oncology*, vol. 4, no. 1, pp. 63–69, 2019.

[80] P. Blanchard, A. S. Garden, G. B. Gunn, D. I. Rosenthal, W. H. Morrison, M. Hernandez, J. Crutison, J. J. Lee, R. Ye, C. D. Fuller, A. S. R. Mohamed, K. A. Hutcheson, E. B. Holliday, N. G. Thaker, E. M. Sturgis, M. S. Kies, X. R. Zhu, R. Mohan, and S. J. Frank, "Intensity-modulated proton beam therapy (impt) versus intensity-modulated photon therapy (imrt) for patients with oropharynx cancer–a case matched analysis," *Radiotherapy and Oncology*, vol. 120, no. 1, pp. 48–55, 2016.

[81] A. C. Moreno, S. J. Frank, A. S. Garden, D. I. Rosenthal, C. D. Fuller, G. B. Gunn, J. P. Reddy, W. H. Morrison, T. D. Williamson, E. B. Holliday, J. Phan, and P. Blanchard, "Intensity modulated proton therapy (impt)–the future of imrt for head and neck cancer," *Oral oncology*, vol. 88, pp. 66–74, 2019.

[82] D. Moeller, M. Nordsmark, T. Nyeng, M. Alber, and L. Hoffmann, "Anatomical changes in oesophageal cancer patients: Posterior beam impt is more robust than imrt," in *Radiotherapy and Oncology*, vol. 127. Elsevier, Ireland LTD Elsevier House, Brookvale Plaza, East Park Shannon, 2018, pp. S525–S526.

[83] C. Grassberger, T. S. Hong, T. Hato, B. Y. Yeap, J. Y. Wo, M. Tracy, T. Bortfeld, J. A. Wolfgang, C. E. Eyler, L. Goyal, J. W. Clark, C. H. Crane, E. J. Koay, M. Cobbold, T. F. DeLaney, R. K. Jain, A. X. Zhu, and D. G. Duda, "Differential Association Between Circulating Lymphocyte Populations With Outcome After Radiation Therapy in Subtypes of Liver Cancer," *International Journal of Radiation Oncology*Biology*Physics*, vol. 101, no. 5, pp. 1222–1225, 2018.

[84] S. G. Ellsworth, A. Yalamanchali, H. Zhang, S. A. Grossman, R. Hobbs, and J.-Y. Jin, "Comprehensive analysis of the kinetics of radiation-induced lymphocyte loss in patients treated with external beam radiation therapy," *Radiation research*, vol. 193, no. 1, pp. 73–81, 2020.

[85] P. S. van Rossum, W. Deng, D. M. Routman, A. Y. Liu, C. Xu, Y. Shiraishi, M. Peters, K. W. Merrell, C. L. Hallemeier, R. Mohan, and S. H. Lin, "Prediction of severe lymphopenia during chemoradiation therapy for esophageal cancer: Development and validation of a pretreatment nomogram," *Practical radiation oncology*, vol. 10, no. 1, pp. e16–e26, 2020.

[86] C. Zhu, S. H. Lin, X. Jiang, Y. Xiang, Z. Belal, G. Jun, and R. Mohan, "A novel deep learning model using dosimetric and clinical information for grade 4 radiotherapy-induced lymphopenia prediction," *Physics in Medicine & Biology*, vol. 65, no. 3, p. 035014, 2020.

[87] I. S. Stafford, M. Kellermann, E. Mossotto, R. M. Beattie, B. D. MacArthur, and S. Ennis, "A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases," *npj Digital Medicine*, vol. 3, no. 1, pp. 30–30, 2020.

[88] C. K. Fisher, A. M. Smith, and J. R. Walsh, "Machine learning for comprehensive forecasting of Alzheimer's Disease progression," *Scientific Reports*, vol. 9, no. 1, p. 13622, 2019.

[89] S. Yousefi, F. Amrollahi, M. Amgad, C. Dong, J. E. Lewis, C. Song, D. A. Gutman, S. H. Halani, J. E. V. Vega, D. J. Brat, and L. A. D. Cooper, "Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models," *Scientific Reports*, vol. 7, no. 1, pp. 1–11, 2017.

[90] B. C. Munsell, C.-Y. Wee, S. S. Keller, B. Weber, C. Elger, L. A. T. da Silva, T. Nesland, M. Styner, D. Shen, and L. Bonilha, "Evaluation of machine learning algorithms for treatment outcome prediction in patients with epilepsy based on structural connectome data," *Neuroimage*, vol. 118, pp. 219–230, 2015.

[91] J. Weiss, F. Kuusisto, K. Boyd, J. Liu, and D. Page, "Machine Learning for Treatment Assignment: Improving Individualized Risk Attribution," *AMIA Annu. Symp. Proc. AMIA Symp*, vol. 2015, pp. 1306–1315, 2015.

[92] Y. Yu, X. Si, C. Hu, and J. Zhang, "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures," *Neural Comput*, vol. 31, no. 7, pp. 1235–1270, 2019.

[93] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical Systems and Signal Processing*, vol. 107, pp. 241–265, 2018.

[94] J. L. Elman, "Finding Structure in Time," *Cognitive Science*, vol. 14, no. 2, pp. 179–211, 1990.

[95] B. Tung, V.-W. Chen, and Soo, "A comparative study of recurrent neural network architectures on learning temporal sequences," *Proceedings of International Conference on Neural Networks (ICNN'96)*, vol. 4, pp. 1945–1950, 1996.

[96] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[97] A. Shewalkar, D. Nyavanandi, and S. A. Ludwig, "Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM and GRU," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 4, pp. 235–245, 2019.

[98] A. Graves, N. Jaitly, and A. Mohamed, "Hybrid speech recognition with Deep Bidirectional LSTM," in *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, 2013, pp. 273–278.

[99] C. Wang, H. Yang, C. Bartz, and C. Meinel, "Image Captioning with Deep Bidirectional LSTMs," in *Proceedings of the 24th ACM international conference on Multimedia*, 2016, pp. 988–997.

[100] X. Zhu, L. Li, J. Liu, Z. Li, H. Peng, and X. Niu, "Image captioning with triple-attention and stack parallel LSTM," *Neurocomputing*, vol. 319, pp. 55–65, 2018.

[101] S. Dai, L. Li, and Z. Li, "Modeling Vehicle Interactions via Modified LSTM Models for Trajectory Prediction," *IEEE Access*, vol. 7, pp. 38 287–38 296, 2019.

[102] F. Altché and A. D. L. Fortelle, "An LSTM network for highway trajectory prediction," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 353–359.

[103] Y. Chen, J. Yuan, Q. You, and J. Luo, "Twitter Sentiment Analysis via Bi-sense Emoji Embedding and Attention-based LSTM," in *Proceedings of the 26th ACM international conference on Multimedia*, 2018, pp. 117–125.

[104] H. Palangi, L. Deng, Y. Shen, J. Gao, X. He, J. Chen, X. Song, and R. Ward, "Deep Sentence Embedding Using Long Short-Term Memory Networks: Analysis and Application to Information Retrieval," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 4, pp. 694–707, 2016.

[105] G. Zhong, K. Zhang, H. Wei, Y. Zheng, and J. Dong, "Marginal Deep Architecture: Stacking Feature Learning Modules to Build Deep Learning Models," *IEEE Access*, vol. 7, pp. 30 220–30 233, 2019.

[106] M. Jiang, J. Liu, L. Zhang, and C. Liu, "An improved Stacking framework for stock index prediction by leveraging tree-based ensemble models and deep learning algorithms," *Phys. Stat. Mech. Its Appl*, vol. 541, pp. 122 272–122 272, 2020.

[107] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi, "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," *Information Fusion*, 2021.

[108] S. Makridakis, R. Kirkham, A. Wakefield, M. Papadaki, J. Kirkham, and L. Long, "Forecasting, uncertainty and risk; perspectives on clinical decision-making in preventive and curative medicine," *International Journal of Forecasting*, vol. 35, no. 2, pp. 659–666, 2019.

[109] P. D. Myers, K. Ng, K. Severson, U. Kartoun, W. Dai, W. Huang, F. A. Anderson, and C. M. Stultz, "Identifying unreliable predictions in clinical risk models," *npj Digital Medicine*, vol. 3, no. 1, pp. 1–8, 2020.

[110] E. Hüllermeier and W. Waegeman, "Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods," *Machine Learning*, vol. 110, no. 3, pp. 457–506, 2021.

[111] H. Jiang, B. Kim, M. Y. Guan, and M. Gupta, "To trust or not to trust a classifier," *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pp. 5546–5557, 2018.

[112] G. Chryssolouris, M. Lee, and A. Ramsey, "Confidence interval prediction for neural network models," *IEEE Transactions on Neural Networks*, vol. 7, no. 1, pp. 229–232, 1996.

[113] H. M. D. Kabir, A. Khosravi, M. A. Hosen, and S. Nahavandi, "Neural Network-Based Uncertainty Quantification: A Survey of Methodologies and Applications," *IEEE Access*, vol. 6, pp. 36 218–36 234, 2018.

[114] I. Rivals and L. Personnaz, "Construction of confidence intervals for neural networks based on least squares estimation," *Neural Networks*, vol. 13, no. 4-5, pp. 463–484, 2000.

[115] C. H. Zhang and S. S. Zhang, "Confidence intervals for low dimensional parameters in high dimensional linear models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 76, no. 1, pp. 217–242, 2014.

[116] A. Mayr, T. Hothorn, and N. Fenske, "Prediction intervals for future BMI values of individual children - a non-parametric approach by quantile boosting," *BMC Medical Research Methodology*, vol. 12, no. 1, 2012.

[117] R. Haskins, P. G. Osmotherly, F. Tuyl, and D. A. Rivett, "Uncertainty in clinical prediction rules: the value of credible intervals," *journal of orthopaedic & sports physical therapy*, vol. 44, no. 2, pp. 85–91, 2014.

[118] K. McGeechan, G. Liew, P. Macaskill, L. Irwig, R. Klein, B. E. K. Klein, J. J. Wang, P. Mitchell, J. R. Vingerling, P. T. V. M. de Jong, J. C. M. Witteman, M. M. B. Breteler, J. Shaw, P. Zimmet, and T. Y. Wong, "Prediction of Incident Stroke Events Based on Retinal Vessel Caliber: A Systematic Review and Individual-Participant Meta-Analysis," *American Journal of Epidemiology*, vol. 170, no. 11, pp. 1323–1332, 2009.

[119] H. Nishiura, "Early Detection of Nosocomial Outbreaks Caused by Rare Pathogens: A Case Study Employing Score Prediction Interval," *Osong Public Health and Research Perspectives*, vol. 3, no. 3, pp. 121–127, 2012.

[120] R. Tanno, A. Ghosh, F. Grussu, E. Kaden, A. Criminisi, and D. C. Alexander, "Bayesian Image Quality Transfer," *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*, pp. 265–273, 2016.

[121] R. Trinchero, M. Larbi, H. M. Torun, F. G. Canavero, and M. Swaminathan, "Machine Learning and Uncertainty Quantification for Surrogate Models of Integrated Devices With a Large Number of Parameters," *IEEE Access*, vol. 7, pp. 4056–4066, 2019.

[122] A. Y. K. Foong, D. R. Burt, Y. Li, and R. E. Turner, "On the Expressiveness of Approximate Inference in Bayesian Neural Networks," *Cs*, 2020.

[123] P. Izmailov, W. J. Maddox, P. Kirichenko, T. Garipov, D. Vetrov, and A. G. Wilson, "Subspace Inference for Bayesian Deep Learning," *Uncertainty in Artificial Intelligence*, pp. 1169–1179, 2020.

[124] J. Maroñas, R. Paredes, and D. Ramos, "Calibration of deep probabilistic models with decoupled bayesian neural networks," *Neurocomputing*, vol. 407, pp. 194–205, 2020.

[125] T. Vandal, M. Livingston, C. Piho, and S. Zimmerman, "Prediction and Uncertainty Quantification of Daily Airport Flight Delays," *International Conference on Predictive Applications and APIs*, pp. 45–51, 2018.

[126] A. Filos, S. Farquhar, A. N. Gomez, T. G. J. Rudner, Z. Kenton, L. Smith, M. Alizadeh, A. D. Kroon, and Y. Gal, "A Systematic Comparison of Bayesian Deep Learning Robustness in Diabetic Retinopathy Tasks," *Eess*, 2019.

[127] C. Leibig, V. Allken, M. S. Ayhan, P. Berens, and S. Wahl, "Leveraging uncertainty information from deep neural networks for disease detection," *Scientific Reports*, vol. 7, no. 1, pp. 17 816–17 816, 2017.

[128] A. Tousignant, P. Lemaître, D. Precup, D. L. Arnold, and T. Arbel, "Prediction of Disease Progression in Multiple Sclerosis Patients using Deep Learning Analysis of MRI Data," *International Conference on Medical Imaging with Deep Learning*, pp. 483–492, 2019.

[129] R. Harper and J. Southern, "A Bayesian Deep Learning Framework for End-To-End Prediction of Emotion from Heartbeat," *IEEE Transactions on Affective Computing*, pp. 1–1, 2020.

[130] O. Bousquet, U. von Luxburg, and G. Rätsch, *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures*. Springer, 2011, vol. 3176.

[131] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," *arXiv: 1612.01474*, 2017.

[132] D. J. C. Mackay and D. J. C. M. Kay, *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.

[133] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA: A Cancer Journal for Clinicians*, vol. 69, no. 1, pp. 7–34, 2019.

[134] M. Zaghian, W. Cao, W. Liu, L. Kardar, S. Randeniya, R. Mohan, and G. Lim, "Comparison of linear and nonlinear programming approaches for "worst case dose" and "minmax" robust optimization of intensity-modulated proton therapy dose distributions," *Journal of Applied Clinical Medical Physics*, vol. 18, no. 2, pp. 15–25, Mar. 2017.

[135] E. J. Hall and A. J. Giaccia, *Radiobiology for the Radiologist*. Lippincott Williams & Wilkins, 2006, vol. 6.

[136] B. G. Wouters, "Cell death after irradiation: how, when and why cells die," *Basic Clinical Radiobiology*, p. 27, 2009.

[137] D. J. Carlson, R. D. Stewart, and V. A. Semenenko, "Effects of oxygen on intrinsic radiation sensitivity: A test of the relationship between aerobic and hypoxic linear-quadratic (lq) model parameters a," *Medical physics*, vol. 33, no. 9, pp. 3105–3115, 2006.

[138] P. Paul-Gilloteaux, V. Potiron, G. Delpon, S. Supiot, S. Chiavassa, F. Paris, and S. V. Costes, "Optimizing radiotherapy protocols using computer automata to model tumour cell death as a function of oxygen diffusion processes," *Scientific reports*, vol. 7, no. 1, pp. 1–14, 2017.

[139] J. Jeong, J. H. Oh, J.-J. Sonke, J. Belderbos, J. D. Bradley, A. N. Fontanella, S. S. Rao, and J. O. Deasy, "Modeling the cellular response of lung cancer to radiation therapy for a broad range of fractionation schedules," *Clinical Cancer Research*, vol. 23, no. 18, pp. 5469–5479, 2017.

[140] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear programming: theory and algorithms*. John Wiley & Sons, 2013.

[141] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[142] X. Liu, W. Yu, F. Liang, D. Griffith, and N. Golmie, "On deep reinforcement learning security for industrial internet of things," *Computer Communications*, vol. 168, pp. 20–32, 2021.

[143] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Thirtieth AAAI conference on artificial intelligence*, 2016.

[144] J. Uzan and A. Nahum, "Radiobiologically guided optimisation of the prescription dose and fractionation scheme in radiotherapy using biosuite," *The British journal of radiology*, vol. 85, no. 1017, pp. 1279–1286, 2012.

[145] C. Van Leeuwen, A. Oei, J. Crezee, A. Bel, N. Franken, L. Stalpers, and H. Kok, "The alfa and beta of tumours: a review of parameters of the linear-quadratic model, derived from clinical radiotherapy studies," *Radiation oncology*, vol. 13, no. 1, pp. 1–11, 2018.

[146] M. C. Roach, J. D. Bradley, and C. G. Robinson, "Optimizing radiation dose and fractionation for the definitive treatment of locally advanced non-small cell lung cancer," *Journal of thoracic disease*, vol. 10, no. Suppl 21, p. S2465, 2018.

[147] J.-J. Sonke, M. Aznar, and C. Rasch, "Adaptive radiotherapy for anatomical changes," *Seminars in radiation oncology*, vol. 29, no. 3, pp. 245–257, 2019.

[148] S. Y. El Sharouni, H. Kal, and J. Battermann, "Accelerated regrowth of non-small-cell lung tumours after induction chemotherapy," *British journal of cancer*, vol. 89, no. 12, pp. 2184–2189, 2003.

[149] A. E. Nahum and J. Uzan, "(radio) biological optimization of external-beam radiotherapy," *Computational and mathematical methods in medicine*, vol. 2012, 2012.

[150] Y. Watanabe, E. L. Dahlman, K. Z. Leder, and S. K. Hui, "A mathematical model of tumor growth and its response to single irradiation," *Theoretical Biology and Medical Modelling*, vol. 13, no. 1, p. 6, 2016.

[151] A. Santiago, S. Barczyk, U. Jelen, R. Engenhart-Cabillic, and A. Wittig, "Challenges in radiobiological modeling: can we decide between lq and lq-l models based on reviewed clinical nsclc treatment outcome data?" *Radiation Oncology*, vol. 11, no. 1, p. 67, 2016.

[152] M. Stuschke and C. Pöttgen, "Altered fractionation schemes in radiotherapy," in *Controversies in the Treatment of Lung Cancer*. Karger Publishers, 2010, vol. 42, pp. 150–156.

[153] Y. Seppenwoolde, J. V. Lebesque, K. De Jaeger, J. S. Belderbos, L. J. Boersma, C. Schilstra, G. T. Henning, J. A. Hayman, M. K. Martel, and R. K. Ten Haken, "Comparing different ntcp models that predict the incidence of radiation pneumonitis," *International Journal of Radiation Oncology* Biology* Physics*, vol. 55, no. 3, pp. 724–735, 2003.

[154] A. Nahum and G. Kutcher, "Biological evaluation of treatment plans," in *Handbook of radiotherapy physics theory and practice*, A. N. P. Mayles and J. C. Rosenwald, Eds. London, UK: Taylor & Francis, 2007, ch. 36, pp. 731–771.

[155] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*. Wiley, 2014.

[156] H.-P. Wieser, E. Cisternas, N. Wahl, S. Ulrich, A. Stadler, H. Mescher, L.-R. Müller, T. Klinge, H. Gabrys, L. Burigo, A. Mairani, S. Ecker, B. Ackermann, M. Ellerbrock, K. Parodi, O. Jäkel, and M. Bangert, "Development of the open-source dose calculation and optimization toolkit matrad," *Medical physics*, vol. 44, no. 6, pp. 2556–2568, 2017.

[157] J. R. Cameron, J. G. Skofronick, R. M. Grant, and R. L. Morin, "Physics of the body," *Medical Physics*, vol. 27, no. 2, pp. 425–425, 2000.

[158] A. M. Farese, K. G. Hankey, M. V. Cohen, and T. J. MacVittie, "Lymphoid and myeloid recovery in rhesus macaques following total body x-irradiation," *Health physics*, vol. 109, no. 5, p. 414, 2015.

[159] R. E. Goans, E. C. Holloway, M. E. Berger, and R. C. Ricks, "Early dose assessment in criticality accidents," *Health physics*, vol. 81, no. 4, pp. 446–449, 2001.

[160] G. J. Lim, L. Kardar, S. Ebrahimi, and W. Cao, "A risk-based modeling approach for radiation therapy treatment planning under tumor shrinkage uncertainty," *European Journal of Operational Research*, vol. 280, no. 1, pp. 266–278, 2020.

[161] S. Ebrahimi, G. Lim, A. Liu, S. H. Lin, S. G. Ellsworth, C. Grassberger, R. Mohan, and W. Cao, "Radiation-Induced Lymphopenia Risks of Photon Versus Proton Therapy for Esophageal Cancer Patients," *Int J Part Ther*, 2021.

[162] C. F. Dormann, J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, and S. Lautenbach, "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance," *Ecography*, vol. 36, no. 1, pp. 27–46, 2013.

[163] L. V. D. Maaten, E. Postma, V. Den, and J. Herik, "Dimensionality reduction: a comparative," *J Mach Learn Res*, vol. 10, pp. 13–13, 2009.

[164] T. P. Zhou H, Wang F, "t-Distributed Stochastic Neighbor Embedding Method with the Least Information Loss for Macromolecular Simulations," *Journal of chemical theory and computation*, vol. 14, pp. 5499–5510, 2018.

[165] L. V. D. Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of machine learning research*, no. 11, pp. 9–9, 2008.

[166] L. V. D. Maaten, "Learning a parametric embedding by preserving local structure," in *Artificial Intelligence and Statistics*. PMLR, 2009, pp. 384–391.

[167] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *International Journal of Forecasting*, vol. 20, no. 1, pp. 5–10, 2004.

[168] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv:14126980*, 2017.

[169] H. B. Stone, L. J. Peters, and L. Milas, "Effect of host immune capability on radiocurability and subsequent transplantability of a murine fibrosarcoma," *Journal of the National Cancer Institute*, vol. 63, no. 5, pp. 1229–1235, 1979.

[170] S. Yovino, L. Kleinberg, S. A. Grossman, M. Narayanan, and E. Ford, "The etiology of treatment-related lymphopenia in patients with malignant gliomas: modeling radiation dose to circulating lymphocytes explains clinical observations and suggests methods of modifying the impact of radiation on immune cells," *Cancer investigation*, vol. 31, no. 2, pp. 140–144, 2013.

[171] J. Liu, Q. Zhao, W. Deng, J. Lu, X. Xu, R. Wang, X. Li, and J. Yue, "Radiation-related lymphopenia is associated with spleen irradiation dose during radiotherapy in patients with hepatocellular carcinoma," *Radiation Oncology*, vol. 12, no. 1, 2017.

[172] P. Fang, W. Jiang, R. Davuluri, C. Xu, S. Krishnan, R. Mohan, A. C. Koong, C. C. Hsu, and S. H. Lin, "High lymphocyte count during neoadjuvant chemoradiotherapy is associated with improved pathologic complete response in esophageal cancer," *Radiotherapy and Oncology*, vol. 128, no. 3, pp. 584–590, 2018.

[173] S. Ebrahimi, G. Lim, B. Hobbs, S. H. Lin, R. Mohan, and W. Cao, "A hybrid deep learning model for forecasting lymphocyte depletion during radiation therapy," *Medical Physics*, 2021.

[174] D. A. Nix and A. S. Weigend, "Estimating the mean and variance of the target probability distribution," *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)*, vol. 1, pp. 55–60, 1994.

[175] N. Brancati, G. D. Pietro, M. Frucci, and D. Riccio, "A Deep Learning Approach for Breast Invasive Ductal Carcinoma Detection and Lymphoma Multi-Classification in Histological Images," *IEEE Access*, vol. 7, pp. 44 709–44 720, 2019.

[176] H. Paganetti, A. Niemierko, M. Ancukiewicz, L. E. Gerweck, M. Goitein, J. S. Loeffler, and H. D. Suit, "Relative biological effectiveness (RBE) values for proton beam therapy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 53, no. 2, pp. 407–421, 2002.