Improving the efficacy of functional lung avoidance radiation therapy

By

Eric M. Wallat

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Medical Physics)

at the
University of Wisconsin – Madison
2023

Date of final oral examination: 04/12/2023

This dissertation is approved by the following members of the Final Oral Committee:

Bryan P. Bednarz, Associate Professor, Medical Physics, University of Wisconsin-Madison
Scott K. Nagle, Associate Professor, Radiology, University of Wisconsin-Madison
Joseph M. Reinhardt, Professor, Biomedical Engineering, University of Iowa
Diego Hernando, Associate Professor, Medical Physics, University of Wisconsin-Madison
Brian W. Pogue, Professor, Medical Physics, University of Wisconsin-Madison

"Why seek answers when we do not know the question?"

Siebren de Kuiper

Abstract

Department of Medical Physics

Improving the efficacy of functional lung avoidance radiation therapy

Eric M. Wallat

Functional lung avoidance radiation therapy (RT) is a technique being investigated to avoid specific regions of the lung that are more susceptible to radiation-induced damage. Reducing dose delivered to high functioning regions may reduce the occurrence of radiation-induced lung toxicities. There is a need to improve current approaches to functional avoidance RT by accounting for irradiating the airways, and developing more accurate models of post-RT ventilation change. The purpose of this work was to develop dose-response models for ventilation changes, investigate indirect ventilation decline due to airway irradiation, and incorporate functional information into the treatment planning process.

Analysis was performed in swine to investigate the effect of irradiating airways and regional ventilation change. A dose-response relationship was observed with reduction in luminal area of the airways. Regions supplied by irradiated airways experienced larger loss in ventilation post-RT compared to regions supplied by unirradiated airways. Additional analysis was performed on human subjects and quantified the relationship between airway resistance and 4DCT-based ventilation changes. Increased cumulative airway resistance was significantly associated with decreases in ventilation function for the region

those airways supplied. A normal tissue complication probability model was built to predict bronchial stenosis.

A polynomial regression model was built to predict change in ventilation based on pre-RT ventilation and dose. This model accounted for out-of-phase ventilation by using a new methodology to derive ventilation maps using multiple phases of the 4DCT, and improved a previously developed model in accuracy and gamma pass rates. Additionally, a conditional generative adversarial network model was developed, and improved upon the polynomial regression model in predicting regions of ventilation decline.

Lastly, a new methodology was developed to incorporate functional information into the treatment planning process. This approach utilized dose-painting-by-numbers to create voxelized dose objectives within the lung based on predictions from the cGAN model. The model developed for airway toxicities was used to include dose objectives to limit the occurrence of bronchial stenosis. The new workflow resulted in more spared lung function compared to standard of care as well as the currently used method for functional avoidance plans.

Acknowledgements

There are many people that I have interacted with over the past 5 years that have contributed to my growth as a researcher, a physicist, and as an individual. First I would like to thank Dr. John Bayouth for sharing his passion and expertise for medical physics, and for the invaluable advice and mentorship he has provided. His dedication to the field has inspired my future career goals, and I am grateful for his continued guidance. I have especially appreciated his attention to my well-being throughout graduate school and his light-hearted humor to get us through rough patches. He has been instrumental in my successes thus far, and I am truly thankful to have had him as my advisor. I would also like to thank my advisors across the state border at the University of Iowa; Dr. Joe Reinhardt and Dr. Gary Christensen. Their expertise and perspectives from outside of medical physics have provided me with a greater context for the work we have developed together, and the advice I received in our weekly meetings has had a large impact on my development as a researcher. I would lastly like to thank my dissertation committee: Dr. Bryan Bednarz, Dr. Brian Pogue, Dr. Diego Hernando, Dr. Joe Reinhardt, and Dr. Scott Nagle. You have all provided me with valuable guidance to help make this work a success.

There have been many additional colleagues along the way that have helped me, either technically or personally, and without them I would not be writing this dissertation. Also from the University of Iowa, Dr. Sarah Gerard and Dr. Wei Shao have both been extremely helpful and patient when I came to them with questions about machine learning. To the rest of the friends I've made in the Department of Medical Physics at Wisconsin: thank you for the countless conversations we've had to distract us from getting work done and providing an escape when I most needed it. Most importantly, thank you to Mattison Flakus and Dr. Antonia Wuschner for being the best two labmates I could've asked for. Between helping me carry all my stuff when I was on crutches after tearing my Achilles, having excellent taste in coffee, and being all around awesome people, I'm grateful we have had these past five years working together. Ant, thank you for the many occasions you've watched my cats and for listening to me talk through all the ways my code wasn't working; Mattison, thank you for all your help as we finish off the clinical trial analysis and for also being there to listen to me complain about all the things that

went wrong. Lastly, thank you both for all the time and dedication it took to collect the swine data; without it, my work would not have been possible.

Finally, I would like to thank all my family, friends, and past teachers who have supported me. Thank you to my high school science teachers: Mr. Lewis, Mr. Lui, and Mr. Voigt. They all further pushed my passion for science and fostered an environment where it was exciting to come to class and learn. Thank you to all my friends from high school and undergrad who have always been there for me (and have given me a hard time when needed). Thank you to my parents who have always pushed me to be the best version of myself and provided me with countless opportunities during my education to maximize my learning and growth. I would not be the person I am today without their endless support and love. Lastly, thank you to my wife, Brittney; her unconditional love and support as I've gone through this journey has not gone unnoticed. I cannot thank her enough for the sacrifices she has made to allow me to pursue this education and career, and for having to listen to me talk in countless meetings about "Jacobian ratios" and "the pre-RT and post-RT changes". I love you, and I am excited to see where life takes us next!

Contents

List of Figures

List of Tables

Abbreviations

3-D	**three**-**D**imensional
4-D	**four**-**D**imensional
Ai	Luminal **A**rea
AUC	**A**rea **U**nder **C**urve
CNN	**C**onvolutional **N**eural **N**etwork
CT	**C**omputed **T**omography
DLCO	**D**iffusing Capacity of **L**ungs for **C**arbon M**o**noxide
DVH	**D**ose **V**olume **H**istogram
FEV1	**F**orced **E**xpiratory **V**olume in **one** second
FVC	**F**orced **V**ital **C**apacity
LER	**L**ung **E**xpansion **R**atio
MRI	**M**agnetic **R**esonance **I**maging
MU	**M**onitor **U**nit(s)
NPT	**N**on-**P**neumonitis **T**oxicity
NTCP	**N**ormal **T**issue **C**omplication **P**robability
OAR	**O**rgan **A**t **R**isk
PBT	**P**roximal **B**ronchial **T**ree
PET	**P**ositron **E**mission **T**omography
PTV	**P**lanning **T**arget **V**olume
ROC	**R**eceiver **O**perating **C**haracteristic
ROI	**R**egion **O**f **I**nterest

RT	**R**adiation **T**herapy
SBRT	**S**tereotactic **B**ody **R**adiation **T**herapy
SOC	**S**tandard **O**f **C**are
SPECT	**S**ingle **P**hoton **E**mission **C**omputed **T**omography
TCP	**T**umor **C**ontrol **P**robability
TPS	**T**reatment **P**lanning **S**ystem
WA	**W**all **A**rea

Chapter 1

Introduction

Lung cancer is estimated to be one of the most common and deadly cancers in the United States in 2023, and radiation therapy (RT) remains one of the primary treatment techniques, especially for cases deemed inoperable [1]. However, of lung cancer patients treated with RT, approximately 20% will experience a radiation-induced lung injury (RILI), which can range from short term effects, such as pnuemonitis, to long term effects, such as fibrosis [2–4]. The severity of RILIs can also vary, and in some cases can be life-threatening. Functional avoidance RT is currently being investigated in multiple clinical trials (NCT02843568, NCT02528942, NCT02308709, NCT02002052)[5–8] as a method to alter a treatment plan in order to spare higher functioning lung tissue and reduce the occurrence of RILI in patients post-RT.

The main objective of functional avoidance RT is to maintain disease control while minimizing the normal tissue toxicities that can occur as a result of treatment. The standard of care (SOC) for lung cancer includes using volumetric measurements, such as the volume of the lung receiving 20 Gy or more (V20), to guide treatment plans in order to reduce lung toxicities. One issue with this approach is that it assumes the function of the lung is homogeneous. Previous studies have shown that functional-guided RT may be effective in preserving functional lung by preferentially avoiding higher functioning regions [9–15].

Current methods of determining ventilation via four-dimensional computed tomography (4DCT) image processing only consider one part of a complex biomechanical process. Treating patients using functional information based solely on ventilation metrics could result in irreparable damage to pulmonary vasculature and airway structures [16–19]. Incorporating additional functional information into treatment plans is essential to limit the toxicities patients experience following RT. Few studies have been published on the indirect effects of irradiating other important structures, such as airways and vasculature, and their role in ventilation post-RT. Consequently, the combination of direct and indirect ventilation damage has not been used to determine post-RT function or guide functional avoidance treatment plans.

1.1 Organization of Dissertation

The goal of this thesis is to improve the efficacy of pulmonary functional avoidance radiation therapy through more accurate dose-response modeling and consistent incorporation of relevant information into the treatment planning process. This work is organized into five parts:

1. Preliminary investigation of radiation-induced airway toxicities in a novel swine model and the impact on downstream ventilation change.

2. Retrospective analysis of radiation-induced bronchial stenosis in human subjects and subsequent indirect ventilation damage.

3. Expansion of a polynomial dose-response model to incorporate out-of-phase ventilation.

4. Development of deep learning-based dose-response models for predicting post-RT ventilation change.

5. Comparison of various treatment planning methods and development of a voxel-wise dose-painting-by-numbers approach for functional avoidance radiation therapy treatment planning.

1.2 Chapter Descriptions

Chapter 2 provides background information on pulmonary function and anatomy, the different techniques through which pulmonary function is measured and quantified, and how that information can be used in radiation therapy treatment planning. *Chapter 3* performs a preliminary investigation of radiation-induced airway toxicities and downstream ventilation changes in a swine model, and retrospectively analyzes airway toxicities in human subjects and the correlations to regional downstream ventilation decline. *Chapter 4* develops a polynomial regression model to predict post-RT ventilation change based on pre-RT ventilation and radiation dose, while accounting for effects of out-of-phase ventilation, and develops and investigates multiple machine learning-based models to predict post-RT ventilation change while maintaining spatial relationships and preserving global context. *Chapter 5* combines functional information from the airway analysis and predictive models to develop a comprehensive approach to generating functional avoidance treatment plans. Finally, *Chapter 6* summarizes the major conclusions and future direction of this work.

Chapter 2

Background and motivation

2.1 Introduction

As radiation therapy (RT) treatments have become more advanced in their ability to be more conformal to the target, additional care can be taken to reduce radiation dose to normal tissues. Additionally, advancements in computers have led to faster and more accurate estimations of dose distributions and the development of radiation treatment plans. These advancements have resulted in the ability to create highly conformal and patient-specific treatment plans that provide better tumor control and normal tissue sparing. Functional avoidance RT pushes further to include functional information into the treatment planning process to avoid specific regions of the lung based on pre-RT image analysis. This work focuses on methods to improve the modeling of ventilation response to radiation dose and incorporation of those metrics in the treatment plan. However, it is essential to understand the basic lung mechanics that are at the foundation of predictive models as well as the techniques used to derive lung function metrics. This chapter includes an overview of pulmonary physiology and anatomy, radiation-induced pulmonary

toxicities, techniques to acquire pulmonary functional information, and approaches to modeling pulmonary toxicities.

2.2 Pulmonary Physiology/Anatomy and Lung Toxicities

2.2.1 Pulmonary Function

The main role of the respiratory system is to provide a means for gas exchange, and the main organ of this system is the lung. Oxygen is brought into the body and carbon dioxide is expelled through the act of breathing, which is largely a mechanical process in which the thoracic and abdominal muscles expand or contract the lungs to create pressure gradients. During inspiration, negative pressure is created (relative to the atmosphere) around the lungs, which causes air to be forced into the airways. The air makes its way through the airways until it reaches the terminal bronchioles. Each of these terminal bronchioles ends in an acinus, a structure containing alveolar sacs branching off of the respiratory bronchioles [20]. The alveolar sacs each contain 20-30 alveoli, which are small air sacs at the end of bronchioles that allow for gas exchange through a thin blood-gas barrier [20]. There are approximately 200-500 million alveoli in the average adult human lungs, resulting in an average internal surface area of 75 m^2 despite the average lung volume being 6 liters [21]. Because of the large internal surface area, this process is sufficient to provide the body with the oxygen it needs to survive.

Figure 2.1 shows the pathway of respiration starting from the trachea and continuing through the airways until the terminal bronchioles and alveoli, where gas exchange occurs with blood during pulmonary circulation. As can be seen in Figure 2.1 and due to

the branching of the bronchi, if one region of gas exchange is damaged or nonfunctional, other surrounding regions may be relatively unaffected. This parallel nature of the lungs is partially why functional avoidance radiotherapy has been proposed, as it can potentially avoid specific regions of the lung deemed to be most at risk for damage leading to pulmonary injuries or toxicities, which are discussed in Section 2.2.2.

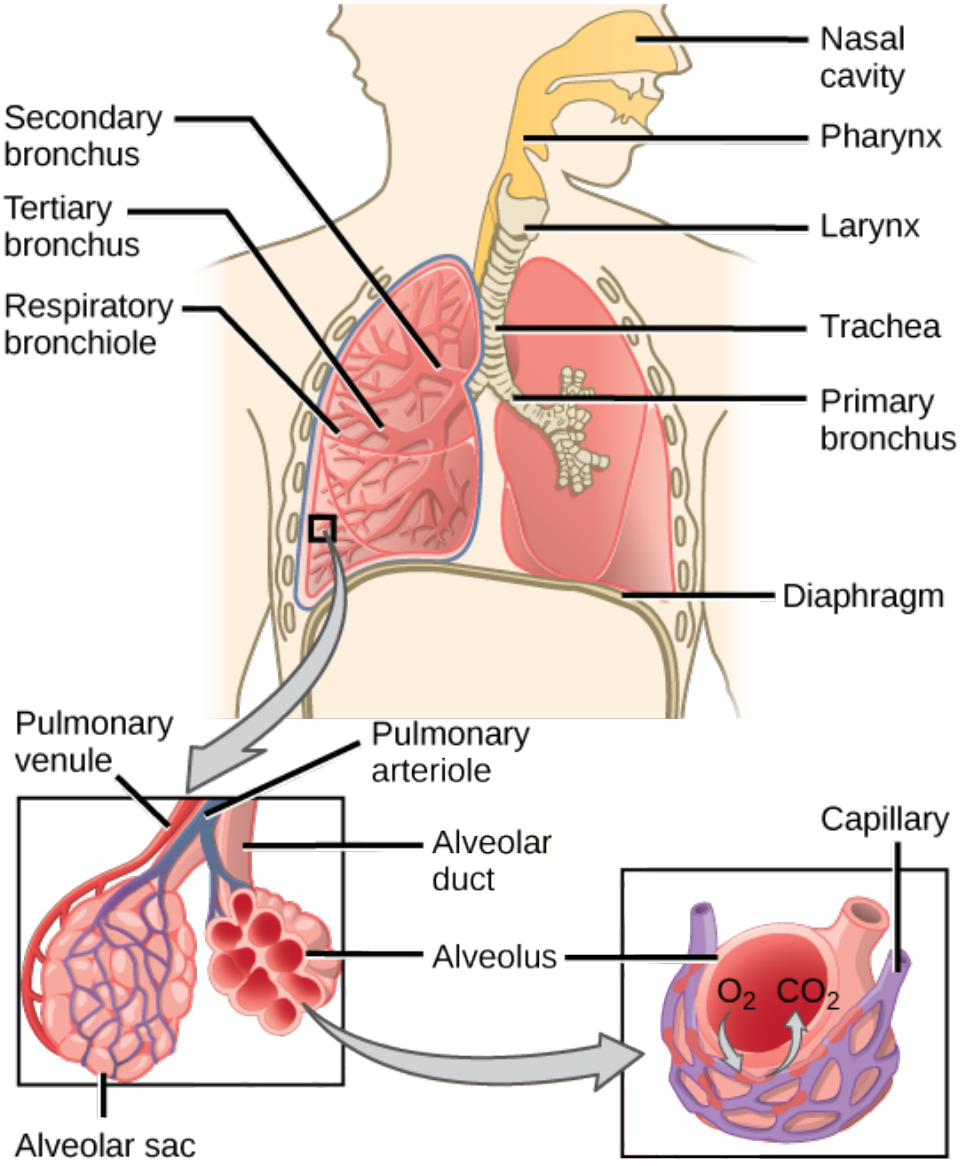


Figure 2.1: Respiratory tract pathway from OpenStax[22]

The pressure gradients created due to the difference in pressure between the atmosphere

and the alveoli are what help the gas exchange process at the blood-gas barrier. At the blood-gas barrier, the walls of the alveoli are thin enough to allow passive diffusion of oxygen and carbon dioxide due to partial pressure differences. These partial pressure differences favor transport of oxygen into the bloodstream and carbon dioxide into the alveoli [20]. The act of replenishing the blood supply in the lungs is commonly known as perfusion, and the act of replenishing the air within the lungs is commonly known as ventilation. As both processes are crucial to the pulmonary and cardiovascular systems, there exist many metrics to quantify their function, which are discussed in Section 2.3.

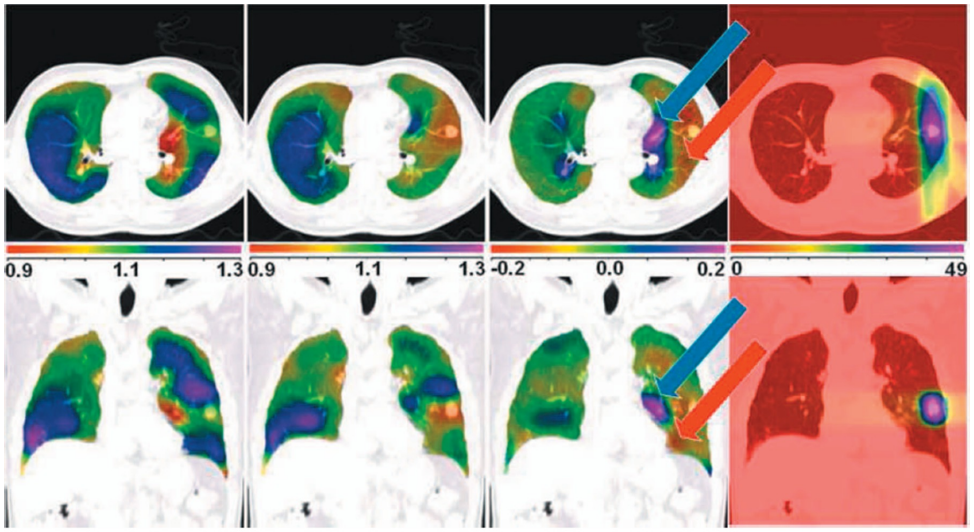
2.2.2 Radiation-Induced Lung Toxicities

It is estimated for 2023 that lung cancer will be the third most diagnosed cancer (approx. 240,000 cases) and account for the most mortalities of any cancer in the United States (approx. 130,000 deaths), in both sexes combined [1]. Approximately one-fourth of all cancer deaths are due to lung cancer, and cigarette smoking is directly linked to 82% of those cases [1]. Typically, lung cancer is asymptomatic, which can lead to later diagnosis and lower survival [1]. For disease that is diagnosed early, surgical resection offers the best chance of survival and prevention of recurrence [23]. For patients who have a later diagnosis, radiotherapy with or without concurrent chemotherapy is the most common form of treatment [24].

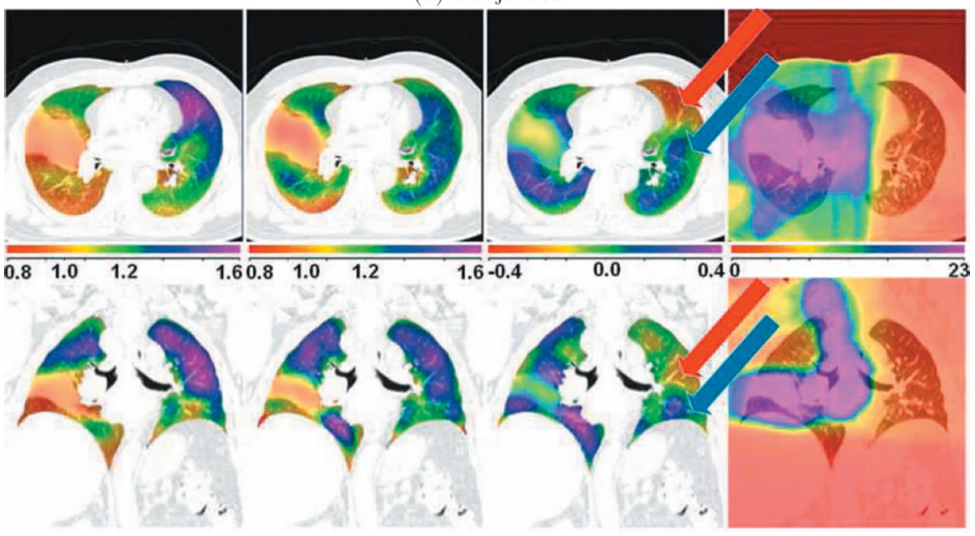
A major complication of using radiation therapy (RT) to treat lung cancer are radiation-induced lung injuries (RILIs), which includes any lung damage due to exposure of lung tissue to ionizing radiation [3]. RILIs are classified into two main groups: an acute stage (hours to days post treatment) known as radiation pneumonitis, a typically reversible inflammatory state of the lung tissue; and a late stage (months to years post treatment)

known as radiation fibrosis, an irreversible scarring and stiffening of the lung tissue [2–4]. Furthermore, this work describes two different types of radiation damage that are dependent on the spatial location of the radiation dose delivered within the lungs: direct ventilation damage and indirect ventilation damage.

Direct ventilation damage refers to the direct ventilation decline caused by irradiation of the lung parenchyma. Almost all published literature on radiation-induced pulmonary damage refers to direct ventilation damage. Ding et al. [25] reported a correlation between radiation dose and pulmonary function change as measured using 4DCT. An example of the pulmonary function change induced by the delivered radiation dose is shown in Figure 2.2 from Ding et al. [25]. Palma et al. [26] investigated post-RT lung density changes and found that increased CT density, which has been associated with histologic findings of lung tissue inflammation [27], was correlated with higher radiation dose. O'Reilly et al. [28] also found that the high ventilation volume, considered to be the regions with the highest 45% to 60% ventilation values, receiving $\geq$20 Gy, was a significant predictor of radiation pneumonitis.



(a) Subject A



(b) Subject B

Figure 2.2: Radiation-induced pulmonary function change from Ding et al. [25]. The first column is the pre-RT pulmonary function. The second column is the post-RT pulmonary function. The third column is the pulmonary function change between the images from the first two columns. The last column is the delivered radiation dose.

While functional avoidance RT has mainly focused on the direct results of irradiating healthy, normal lung tissue, few studies have been published on the indirect effects of irradiating airways and their role in ventilation change post-RT. Indirect ventilation damage is defined as the decline in ventilation in a region due to irradiated airways upstream that supply the particular region.

Characterizing the effect of radiation on the airways and resulting lung function remains an important area of development for improving functional avoidance RT. Previous studies have investigated the relationship between dose and atelectasis in patients who received stereotactic body radiation therapy (SBRT). Kazemzadeh et al. [29] investigated and created a risk model of airway collapse based on the maximum dose and diameter of a bronchial segment. Vicente et al. [30] continued this work to create a functional weighted airway sparing (FWAS) map, as seen in Figure 2.3, to avoid specific airways based on the function of the sub-volume they supplied. While atelectasis is an important clinical endpoint, it is hypothesized that there are imaging biomarkers as precursors to atelectasis that have an impact on post-RT ventilation and lung function. Prior to total collapse of an entire lobe or sub-lobe, there may be subtle changes in the airways, such as bronchial stenosis, that could contribute to losses in pulmonary function in regions "downstream" from the affected airways. No work has provided evidence of functional decline in regions supplied by irradiated airways.

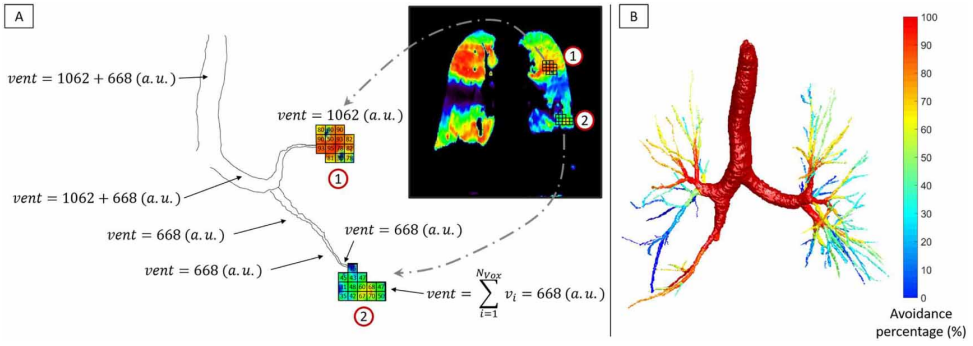


Figure 2.3: Functional weighted airway sparing (FWAS) map developed by Vicente et al. [30]. The contribution of ventilation to each sublobar volume is assigned to the terminal airways and cumulatively summed at each branch point up to the trachea. The FWAS map gives weighted importance to each bronchial segment for functional avoidance treatment planning.

There is limited work published on radiation-induced airway changes and recommended

dose parameters to be used in treatment planning. Miller et al. [16] reported a dose-response effect with external beam radiotherapy-induced stenosis, but focused only on rate of incidence and symptomatic stenosis. The results of Wang et al. [31] provided a recommended dosimetric parameter, but only for the proximal bronchial tree (PBT) and subjects treated with conventionally fractionated thoracic three-dimensional conformal radiation therapy (3DCRT). Additionally, Manyam et al. [32] investigated the results of NRG Oncology/Radiation Therapy Oncology Group (RTOG) 0813 and performed clinical validation of the PBT constraints for 5-fraction SBRT. It was found that a maximum point dose PBT constraint of $D_{0.03cc} \leq 50$ Gy had the best sensitivity and specificity for predicting grade 2 to 5 non-pneumonitis toxicity (NPT). However, little work has been done to quantify dose constraints for airway segments beyond the PBT to limit airway toxicities or develop a quantitative relationship between dose and bronchial stenosis or other radiation-induced changes. The lack of work investigating the dose-response relationship with airway toxicities and their correlation with ventilation damage is the motivation for the airway analysis and dose-response modeling discussed in Chapter 3.

2.3 Quantification of Pulmonary Function

There are many metrics and methods for quantifying pulmonary function, including pulmonary function tests (PFTs), hyperpolarized gas MRI, SPECT/PET, and 4DCT. PFTs can be considered the gold clinical standard as they are simple, widely available to administer, and effective at aiding in diagnosing or monitoring disease. However, PFTs provide a global measurement of lung function and have no ability to provide local/regional information of pulmonary function. Furthermore, PFTs can vary widely for a single patient based on patient effort during the test.

Hyperpolarized gas (HP) MRI is a relatively new method of determining pulmonary function and is able to study ventilation and gas exchange at a regional level. However, this technique is not widely available due to special equipment needed to create the hyperpolarized gas for patients to inhale before the procedure.

SPECT imaging can be considered the clinical standard of ventilation imaging and has been studied for the longest of the ventilation imaging techniques. SPECT imaging allows for imaging of both perfusion and ventilation regionally. For ventilation imaging, a radioactive agent is inhaled by the patient, and for perfusion imaging, a radioactive agent is injected into the patient's bloodstream. Typically, both ventilation and perfusion imaging are done within the same visit, allowing for comparison of the two metrics, known as V/Q matching. Very similar types of studies can also be performed using a PET scanner and positron emitters as the radioactive agent.

Lastly, 4DCTs can be used to estimate ventilation function by calculating the deformation between different phases of the breathing cycle. This technique calculates the deformations using image registration and creates the resultant image transformation. The image transformation is then used to compute the changes in Hounsfield units (HU) due to density changes in the lung from ventilation or to calculate the tissue expansion directly. This section will further describe these methods and the advantages and disadvantages of using each technique to measure pulmonary function.

2.3.1 Pulmonary Function Tests

Pulmonary function tests (PFTs) have been used to quantify pulmonary function on a global scale and include tests such as spirometry and lung volume tests. Other tests can also measure how well oxygen is exchanged into the bloodstream, known as pulse oximetry, or measure levels of fractional exhaled nitric oxide (FeNO), which is an indicator

of inflammation in the lungs [33]. Normal ranges of values for each test are gathered from a large population of healthy subjects to set standards based on age, sex, race, and height [34]. Based on an individual's test value and how it compares to the population mean, the physician can determine whether or not the patient's lung function is normal.

Spirometry is the most common form of PFT used, and it measures lung volume against time [34]. Patients are instructed to take in a maximum inspiration and exhale air as long and as quickly as possible. Metrics from the spirometry test include forced expiratory volume in one second (FEV1), forced vital capacity (FVC), and the ratio of those two volumes (FEV1/FVC). For example, patients with emphysema will have a reduced FEV1 due to a loss in lung elasticity and an inability to forcefully expire. Furthermore, the ratio between the two volumes, FEV1/FVC, can help identify if an obstructive or restrictive ventilation defect exists. These defects are commonly found in patients with asthma or chronic obstructive pulmonary disorder (COPD) [34]. Radiation fibrosis can also cause a decrease in the FEV1/FVC ratio due to the stiffening of the lungs, forcing the lungs to return to a state of exhalation. The ratio of FEV1/FVC may not always be accurate, however, because if a patient is unable to fully inhale or exhale, the measures of FEV1 and FVC may be similar, leading to a ratio near unity.

Another PFT commonly performed measures the diffusing capacity of carbon monoxide (DLCO). This test requires the patient to inhale a gas mixture containing 0.3% carbon monoxide and a low concentration of an inert gas, usually neon, methane, or helium, and then hold their breath for approximately 10 seconds [35]. While the patient is holding their breath, the carbon monoxide enters the bloodstream. The more carbon monoxide that enters the bloodstream in 10 seconds, the greater the diffusing capacity [35]. A decrease in DLCO is typically associated with emphysema or any other disease that thickens or destroys the alveoli, such as radiation fibrosis.

Lung volume tests are another effective measurement of global lung function and provide measurements of lung volumes at various lung capacities. These different volumes and capacities are shown in Figure 2.4. Lung volumes measure the amount of air for one function (i.e., inhalation or exhalation), and lung capacity is any two or more volumes. Figure 2.4 contains four lung volumes and four lung capacities. The capacities are: (1) total lung capacity (TLC), the volume of gas contained inside the lung after a maximum inspiration; (2) inspiratory capacity (IC), the maximum volume of gas inhaled from the resting expiratory level; (3) functional residual capacity (FRC), the volume of gas in the lungs at the resting expiratory level; and (4) vital capacity (VC), the maximum volume of gas that can be exhaled following maximum inspiration. The volumes are: (1) tidal volume (V_T), the volume of gas inhaled or exhaled during an unforced or normal respiratory cycle; (2) inspiratory reserve volume (IRV), the maximum volume of gas inspired from the end of normal inspiration; (3) expiratory reserve volume (ERV), the maximum volume of gas expired from the end of normal expiration; and (4) residual volume (RV), the volume of gas remaining in the lungs after maximum exhalation.

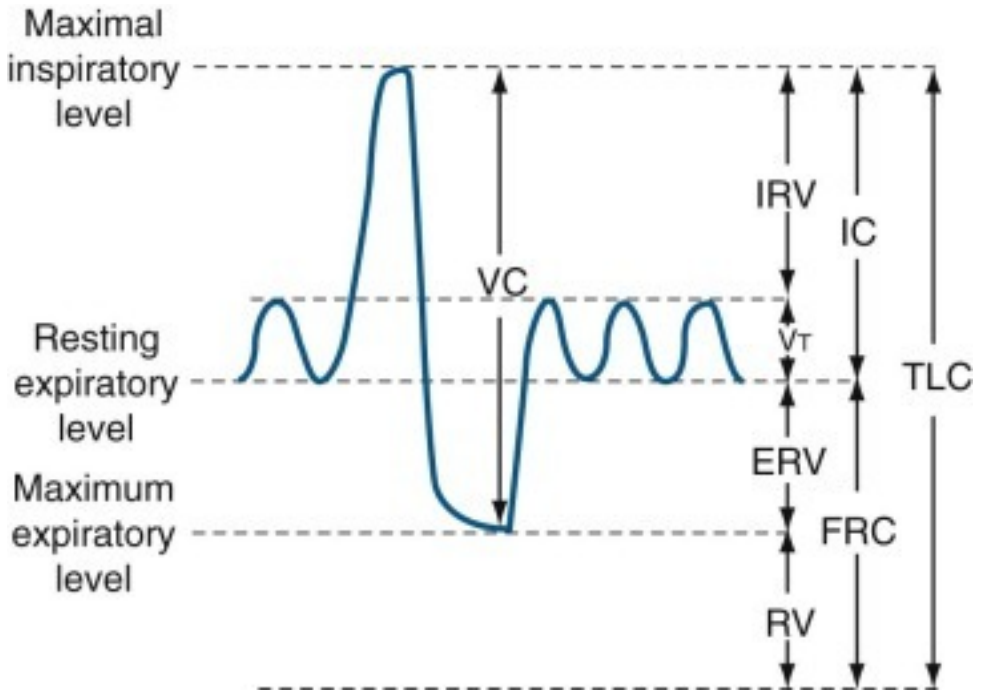


Figure 2.4: Lung volume and capacity from Gold and Koth [35].

PFTs and their measurements have been used to quantify the changes in pulmonary function as a result of radiation therapy in lung cancer patients. Henderson et al. [36] studied the effects of baseline FEV1 and DLCO on post-treatment survival and pulmonary function decrease. They found that patients with lower than baseline values for FEV1 and DLCO had no significant decrease in survival; however, patients with baseline FEV1 in the highest quartile had significantly inferior survival [36]. This result suggests that patients with higher lung function prior to receiving radiation therapy have a lower survival rate. Furthermore, they found a significant decrease in DLCO following radiation therapy. Binkley et al. [37] supported the findings of Henderson et al. [36]: radiation therapy delivered to normal and high functioning regions of the lung predicted worse pulmonary function post-RT, while RT delivered to emphysematic regions of the lung predicted improvement in pulmonary function post-RT. Another study found reduced FEV1 and DLCO 20 weeks post-RT, but these results were not found to be significantly correlated

with pulmonary toxicities [38]. Additionally, Ferrero et al. [38] found that baseline PFTs were not associated with survival, contradicting the results of Henderson et al. [36] and Binkley et al. [37]. One possible reason for this discrepancy was that survival rate was not necessarily dependent on pulmonary function but rather on other factors, including cardiac disease [38]. Stone et al. [39] found similar results supporting the conclusions of Ferrero et al. [38]: neither baseline PFT nor post-RT pulmonary function are associated with worse overall survival. They found that there were declines in FEV1 and forced vital capacity (FVC) at 12 and 24 months post-RT; however, these declines in PFTs did not predict worse overall survival [39].

One thing to note from the studies of Binkley et al. [37], Henderson et al. [36], Ferrero et al. [38], and Stone et al. [39] is that all radiation treatments delivered were stereotactic body radiation therapy (SBRT)/ stereotactic ablative radiotherapy (SABR). SBRT/SABR treatments are typically used for smaller lung lesions that are located more peripherally in the lung, as the treatments are typically delivered in five fractions. Such types of treatments would not be well tolerated near the mediastinum, where critical structures like the heart and mainstem bronchi exist. With SBRT/SABR treatments, due to the smaller volume of lung irradiated, it can be hypothesized there may be less pulmonary decline post-RT. This effect may account for the differences in results amongst the aforementioned studies and is a reason to create more regional and higher sensitivity measurements of pulmonary function. While PFTs are good for initial diagnosis or monitoring of disease, they are global measurements of lung function and have no ability to provide a local/regional analysis of lung function.

2.3.2 SPECT/PET

Lung scintigraphy, which was introduced in the 1960s, was one of the first developed imaging techniques that allowed for assessment of patients with known or suspected pulmonary embolisms (PE) [40]. Wagner et al. [40] used ^{51}Cr and ^{131}I labelled macro-aggregated albumin (MAA) to image pulmonary perfusion in humans and animals (dogs and rats) using a single detector NaI(Tl) rectilinear scanner [40, 41]. The radioactive isotopes used emit gamma rays, which can then be detected by the scintillating detector. Further development of radioisotope imaging used multiple detectors known as gamma cameras. A gamma camera consists of a scintillation crystal (typically NaI(Tl)), a light guide, and multiple photomultiplier tubes (PMTs). The emitted gamma rays from the radioisotope interact with the scintillation crystal which causes the crystal to give off visible light that is proportional to the energy of the interacting gamma ray. Light guides channel the visible light away from the gaps between individual PMTs to increase collection efficiency before finally entering the PMT. The PMTs then collect the visible light and output an electrical pulse with a current that is proportional to the energy of the original gamma ray. The signals from the PMTs are then fed into a position logic system which determines the 2D positions of the scintilation events. This use of gamma cameras, as seen in Figure 2.5, to determine the 2D distribution of a radioisotope within the body sets the basis for single photon emission computed tomography (SPECT).

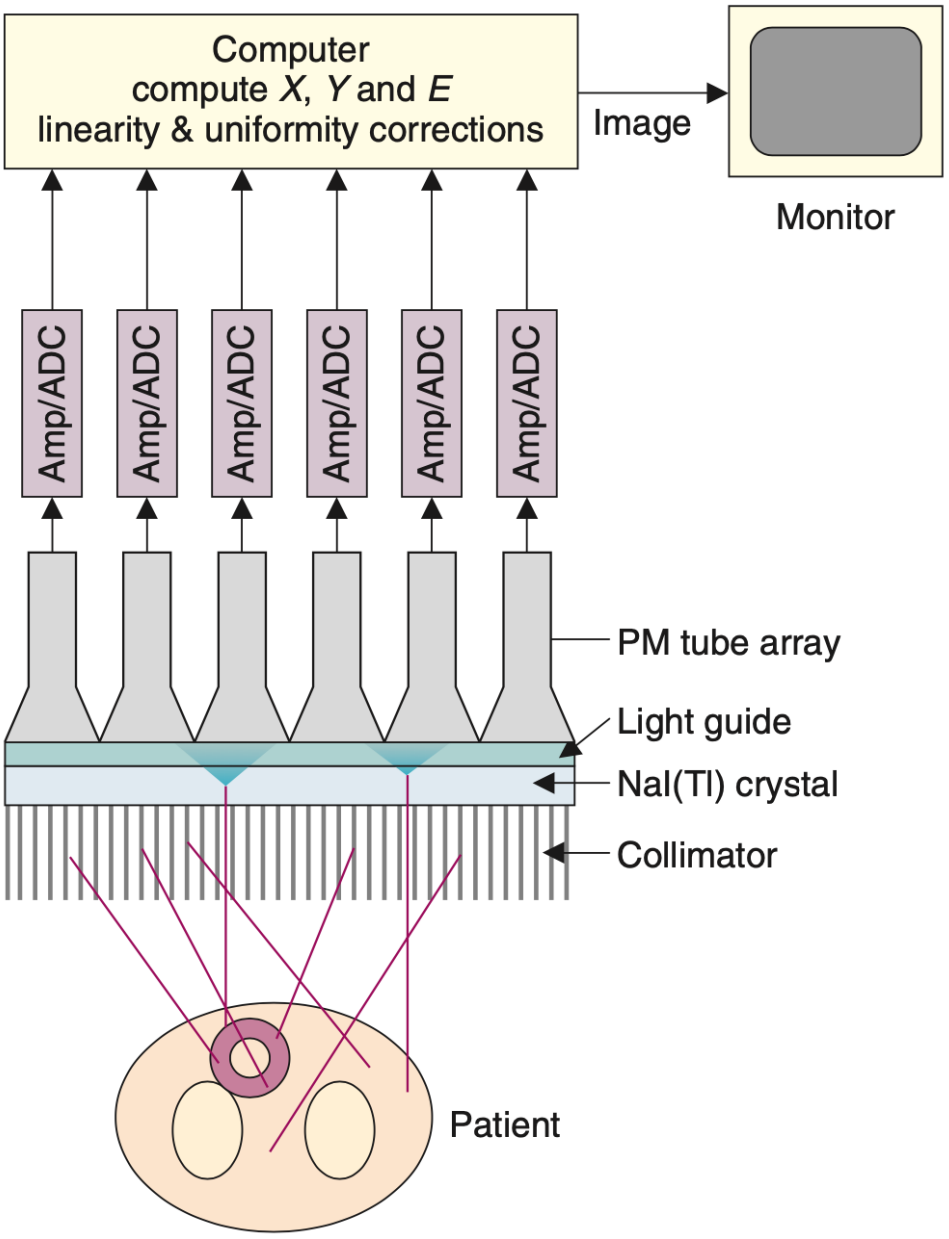


Figure 2.5: Modern gamma camera components showing collection of gamma rays emitted from a patient. Image from Cherry et al. [42]

SPECT imaging uses the same concepts of planar imaging using gamma cameras, but instead rotates the detectors around the patient to develop a 3D distribution of the acquired signal. This technique is known as computed tomography and allows for data to

be captured at all angles around the patient. The collected data can then be reconstructed into the 3D distribution using techniques like simple or filtered backprojection.

In SPECT imaging today, ^{99m}Tc is the most commonly used radionuclide, in part due to its emission of gamma rays with an energy of 140 keV [42]. ^{99m}Tc is the metastable product of ^{99}Mo, which is a by-product of nuclear reactors, and can be harvested via a generator containing ^{99}Mo by "milking" the produced ^{99m}Tc around every 24 hours when needed for a procedure. The ^{99m}Tc can then be labeled to a variety of compounds based on the area of the body to be imaged. For perfusion imaging, ^{99m}Tc can be attached to MAA, and, due to its large size, will get trapped within the smaller pulmonary capillaries causing temporary "micro-embolisms" [43, 44]. The ^{99m}Tc within the capillaries will emit gammas which can then be detected and serve as a surrogate for pulmonary perfusion function. For ventilation imaging, ^{99m}Tc can be attached to diethylenetriaminepentaacetate (DTPA) to create an aerosol which can then be inhaled. Once inhaled, the ^{99m}Tc-DTPA travels through the bronchi, condensates, and attaches to the alveoli. The distribution of ^{99m}Tc-DTPA can be imaged and ventilation defects, such as PEs, can be seen [45]. Technegas, first developed by Burch et al. [46], is another commonly used aerosol that can be inhaled and used for SPECT-based ventilation imaging. Compared to ^{99m}Tc-DTPA, Technegas particles are approximately 100 time smaller which allows for a more homogenous distribution throughout the lungs without the hotspots and clumping commonly seen when using ^{99m}Tc-DTPA [46, 47]. While SPECT imaging of perfusion and ventilation have been proven to have high sensitivity and specificity in diagnosing PEs, it is not without its limitations [48, 49]. SPECT imaging using ^{99m}Tc has a spatial resolution on the order of 1cm full width at half maximum. Additionally, in order to acquire 3D images, the gamma camera is required to rotate around the patient which takes multiple minutes to complete. The long scan times introduce errors due to breathing motion which result in a blurred image [50].

Positron emission tomography (PET) works in a similar fashion to SPECT imaging, where a radioactive tracer is injected into or inhaled by a patient and the emitted radiation is detected by the system. For PET, positron emitters are used, which emit positrons that then eventually interacts with an electron. The most likely interaction between the electron and positron is electron-positron annihilation that produces two 511 keV photons emitted at nearly 180 degrees to each other. The PET system, consisting of a ring of scintillating detectors, detects the two photons nearly simultaneously, which allows the system to localize the photons' origin along a line between the two detectors. Most PET systems are joined with a CT scanner (PET/CT) so that the activity detected by the PET imaging can be overlaid with the CT image acquired in the same setup position and at the same time. For ventilation imaging, ^{68}Ga, in the form of gallium-chloride, can be placed in already existing ^{99m}Tc generators to produce "Galligas". Galligas is the ^{68}Ga-equivalent of the previously mentioned Technegas, and an example ventilation image produced using Galligas is shown in Figure 2.6. For perfusion imaging, ^{68}Ga is attached to MAA, however this process is more labor intensive compared to the ^{99m}Tc process.

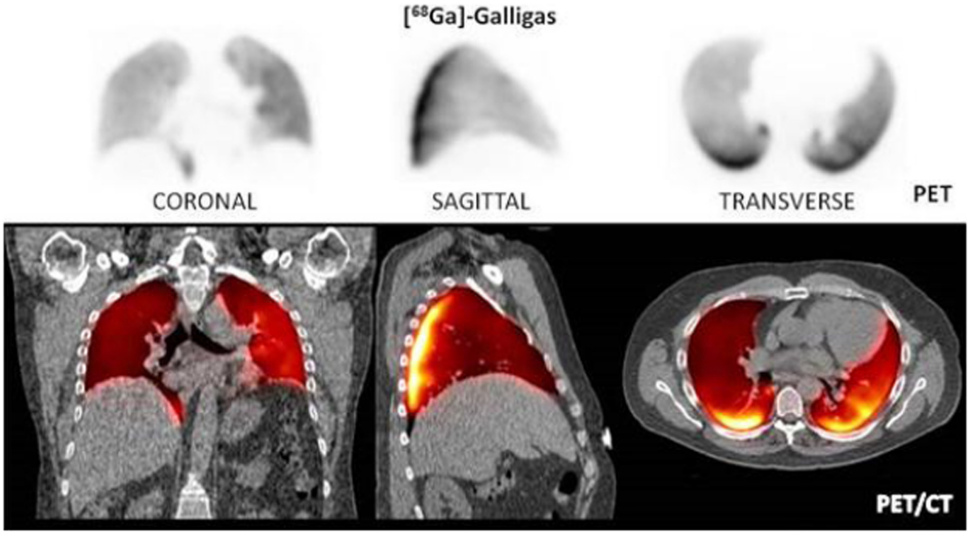


Figure 2.6: Galligas ventilation image showing the distribution of the inhaled gas. Image from Bailey et al. [50]

Le Roux et al. [51] investigated the correlation between Gallium-68 ventilation-perfusion (V/Q) PET/CT and PFTs. With V/Q imaging, ventilation and perfusion were compared to each other to find where they match and where they differ. Four distinct regions were analyzed: normal perfusion and ventilation, abnormal ventilation and perfusion (matched defects), normal ventilation and abnormal perfusion (mismatched defects), and abnormal ventilation and normal perfusion (reverse mismatched defects). They found strong correlations between the percentage of volume of lung with normal V/Q and FEV1/FVC, suggesting the future use of PET/CT for assessing local lung function. Siva et al. [15] investigated using V/Q PET/CT for functionally adapted IMRT plans. Functional avoidance plans were created using either highly-perfused lung or highly-ventilated lung volumes, and were compared to the standard-of-care treatment plans. It was found that the functional avoidance plans created using the highly-perfused volumes resulted in a dose reduction to functional lung, whereas plans created using the highly-ventilated volumes did not show improvement over standard-of-care. The authors contributed this to the image quality and noise levels of the ventilation images compared to the perfusion images due to lower activity and signal in the lungs for ventilation.

2.3.3 Hyperpolarized Gas MRI

Magnetic resonance imaging (MRI) is a form of non-ionizing diagnostic imaging, and is generally used for its excellent soft tissue contrast. MR images are formed by detecting hydrogen atoms after being perturbed within a magnetic field. The main magnetic field of the MRI machine is produced by the large, superconducting magnet that surrounds the bore and typically has a magnetic field strength of 1.5-3.0 T [52]. The main magnetic field causes protons in the body (or material of interest) to either align with or against the magnetic field, with most of the magnetic fields from the protons canceling. However, there is a slight excess of protons that align in the direction of the main magnetic field,

referred to as the net magnetization, which becomes the source of the MR signal used to produce images. The protons that are specifically imaged in typical MR imaging come from hydrogen atoms as they are the most abundant in the human body, and are positively charged, allowing for interaction with the main magnetic field. Once in the magnetic field, the protons will begin to precess at a precessional frequency that is dependent on the specific type of nuclei and the main magnetic field strength, which is about 64 MHz for a hydrogen proton in a 1.5 T magnetic field. A radiofrequency (RF) pulse is then transmitted, through a separate RF coil, at the same precessional frequency in order to transmit the RF energy to the protons. The absorbed RF energy causes the protons to tip away from the longitudinal direction and into the transverse plane. Using receiver coils, the amount of transverse magnetization can be detected and used to reconstruct images.

MR imaging can be used to image the lungs, but suffers from low signal due to the lung being a low-density material. To overcome this, other types of nuclei can be imaged, such as Xenon-129 and Helium-3, but requires specialized equipment to do so. A method known as spin exchange optical pumping, or hyperpolarization, can boost the MR signal of noble gases by up to five orders of magnitude [53], and allows the gases to be imaged within the lung. The patient will inhale the hyperpolarized gas and will be imaged by the MR scanner. Following image post-processing techniques, the distribution of gas can then be visualized within the lungs and is assumed to be the regional ventilation. From these images, ventilation defects, caused by regional airway obstruction and air trapping [54], can be observed, and quantities such as ventilation defect volume, ventilation volume, and coefficient of variation can be determined. These quantities can be used to show regional ventilation heterogeneity for various obstructive pulmonary diseases [54]. While Helium-3 was the first hyperpolarized gas to be developed and used [55], Xenon-129 has been a popular choice for researchers due to a recent global shortage of helium and its unique

properties within the pulmonary system. Xenon is soluble in blood and tissue barriers, and undergoes frequency shifts as it passes through, allowing it to be distinguished in each unique compartment [56]. This allows for ventilation and gas exchange to be imaged simultaneously as seen in Figure 2.7. Figure 2.8 shows the signal distribution for Xenon-129 in different pulmonary tissues. By thresholding the signal at different frequencies the distribution in the airspace provides information about ventilation, while the distribution in the tissue is a surrogate for how much oxygen can be exchanged through the alveoli into the bloodstream.

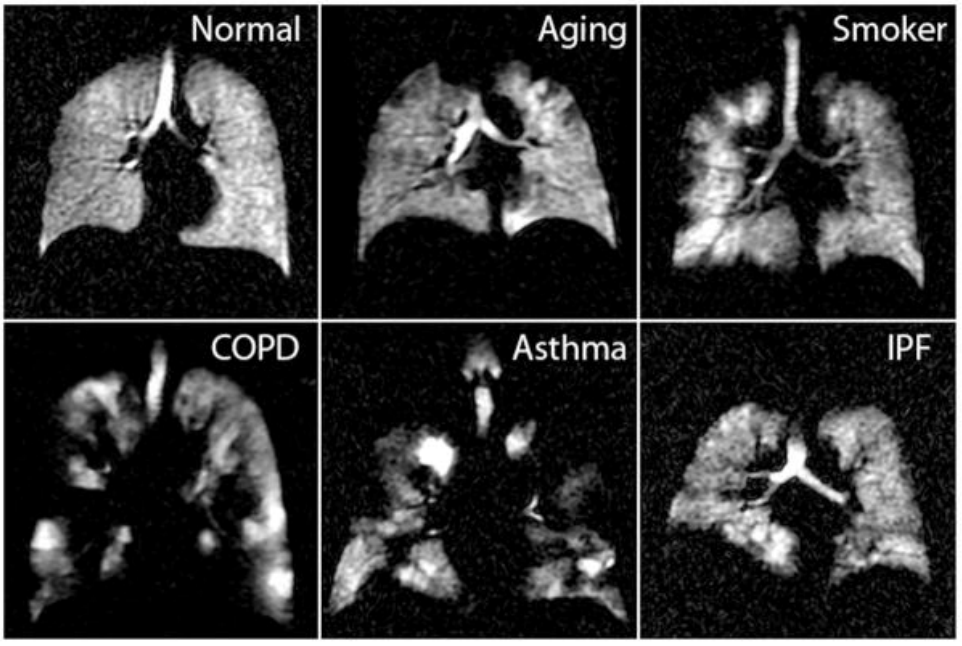


Figure 2.7: Hyperpolarized Xenon-129 ventilation images of a healthy subject, aging subject, smoker, and COPD, asthma, and idiopathic pulmonary fibrosis patients. Image from Roos et al. [57]

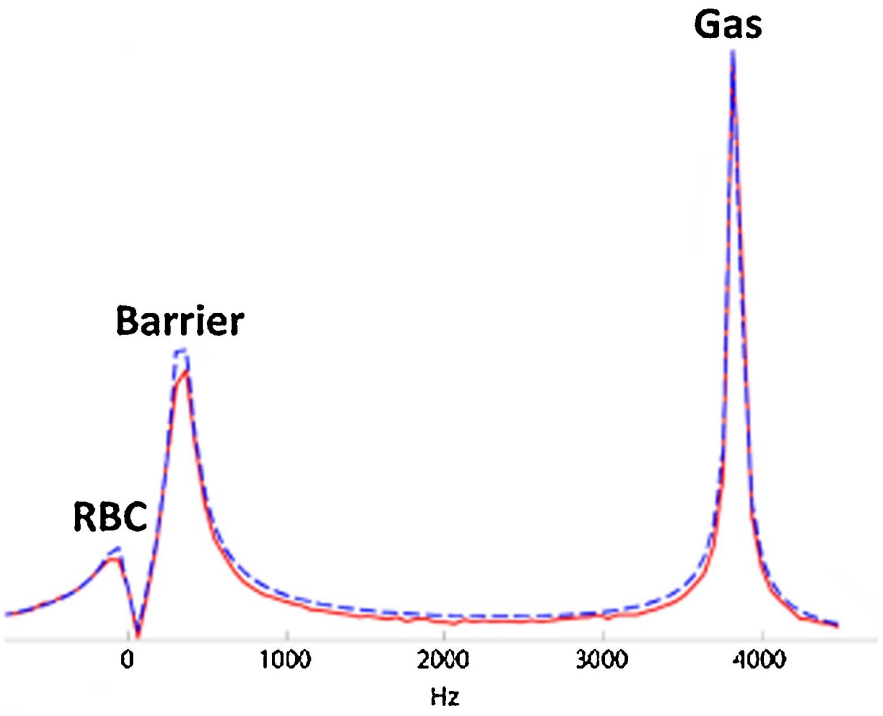


Figure 2.8: Xenon-129 MR spectrum with peaks corresponding to blood, blood-tissue interface, and gas. Figure from Ebner et al. [56].

Multiple studies have investigated using HP-MRI for guiding functional avoidance radiation therapy treatment planning. Ding et al. [58] performed a pilot study using Xenon-129 HP-MR imaging to create ventilation-based segmentation for the purposes of developing a functional avoidance treatment plan. They delineated the HP-MR images into four classifications: well-ventilated, ventilated, hypo-ventilated, and poorly ventilated. One standard-of-care plan was created and compared to a functional avoidance plan where the dose to the well-ventilated lung regions was minimized, and it was found that they were able to significantly reduce the V_{10Gy} and V_{20Gy} for the well-ventilated regions in the experimental plans. Rankine et al. [59] investigated the clinical feasibility of Xenon-129-based HP-MRI for use in functional avoidance treatment plans. Two different functional avoidance plans were created, a gas exchanged-guided plan and a ventilation-guided

plan, and were compared to the standard-of-care clinical plan. They found that the gas exchange-guided planning reduced dose to high functioning regions of the lung while still maintaining standard-of-care dose limits and plan quality. It was also found that the ventilation-guided plans reduced dose to higher gas-exchange regions, suggesting that ventilation-based plans may be clinically useful.

A major drawback to implementing HP-MRI for functional avoidance treatment planning is the need for specialized equipment. Special commercial systems are required in order to perform the necessary spin-exchange optical pumping to produce hyperpolarized gases. Not every institution has the necessary funds or trained staff to implement this type of system and workflow. An additional drawback is the need for the patient to undergo an additional procedure. Even though there is no radiation dose accumulated during the HP-MR scan, it puts an additional time burden on the patient.

2.4 4DCT Ventilation Imaging

Four-dimensional computed tomography (4DCT) has also been a modality used to estimate pulmonary function and ventilation. It is of great interest due to its minimal impact on the clinical workflow during a typical treatment for a patient undergoing radiation therapy as a 4DCT is collected during the simulation and planning process. This section will review 4DCT and CT fundamentals, as well as image registration and ventilation measurement derivation.

2.4.1 4DCT Basics

Computed tomography (CT) is a form of diagnostic imaging that uses ionizing radiation, in the form of x-rays, to create images. Images are formed by passing x-rays through the

body while the x-ray tube rotates around the patient in order to collect information from many different angles. A detector array opposite to the x-ray tube detects the x-rays that are transmitted through the body. The attenuation of the x-rays at each gantry angle is measured and produces a projection of the anatomy which can be used to create 3D images. To create 3D volumetric images, the couch is translated through the bore of the CT scanner all while the x-ray tube is rotating. The pitch of the CT scanner describes the table distance traveled per one rotation of the gantry divided by the beam width. A pitch less than 1 results in oversampling/overscanning and a pitch greater than 1 will undersample as seen in Figure 2.9. CT image pixels/voxels have intensity values that use Hounsfield Units (HU). HU can be calculated from the determined linear attenuation value of the specific voxel and applying a linear transform referenced to the linear attenuation value of water as shown in Equation 2.1. Material with a high linear attenuation, such as bone, will have larger HU values and show bright in images, while material with low linear attenuation, such as air, will have low HU values and show up dark in images.

$$HU = 1000 * \frac{\mu - \mu_{water}}{\mu_{water}} \tag{2.1}$$

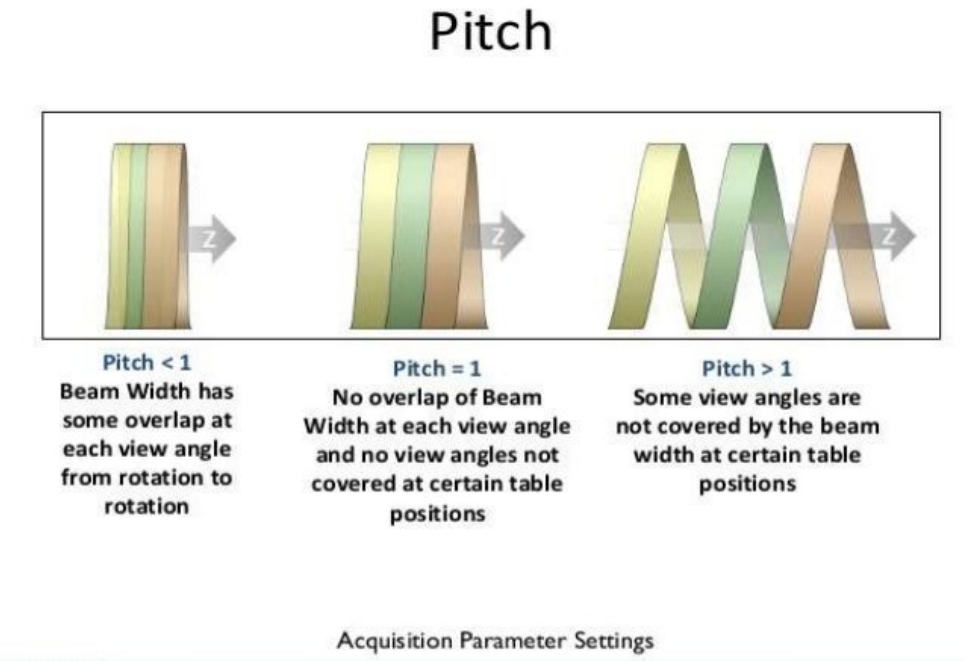


Figure 2.9: Examples of different CT pitch. Image from Elangovan [60].

A more recent advancement of CT is the ability to collect four-dimensional (4D) information during scanning. A 4DCT collects 3D image volumes during different time periods, typically different phases of the respiratory cycle, allowing for visualization of motion throughout the specified time period. This is commonly used during the simulation and treatment planning process for lung cancer patients undergoing radiation therapy. The 4DCT allows physicians to determine the extent of tumor motion during the patient's respiratory cycle. This helps define target boundaries and margins to use for treatment, and to make sure there is adequate tumor coverage as well as minimal healthy tissue overlap. In order to produce 4DCT images that align with the correct respiratory phases, breathing motion must be monitored and recorded during the collection of the 4DCT. A respiratory surrogate [61, 62], such as a strain-gauge belt or infrared reflective box, is placed on or around the patient's abdomen. As the patient breathes the respiratory surrogate will also move with the chest motion in phase with the breathing motion. Once

the scan is over, the breathing trace will be used to determine when each phase of the respiratory cycle occurred relative to the projection data collected by the CT scanner. The projection data is then sorted into each of the respective respiratory phases and the 3D volume for each phase can then be reconstructed as shown in Figure 2.10. Because each slice of the volume of interest needs to be imaged at least once during an entire respiratory cycle, the pitch is reduced by about a factor of 10 in order to capture all the necessary information for proper reconstruction.

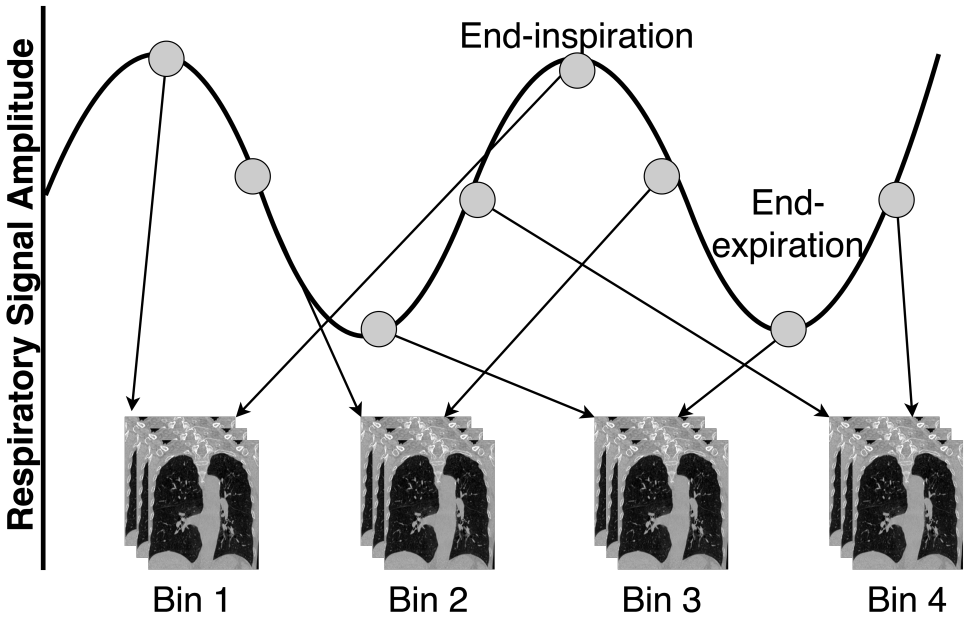


Figure 2.10: Diagram of phase binning for 4DCT reconstruction. This example shows 4 discrete phases, however it is common for 4DCT images to be reconstructed into 10 different respiratory phases.

There are a few disadvantages to working with 4DCT. The first of which is increased radiation dose. Due to the need to oversample in order to collect anatomical information from every respiratory phase, overall imaging time increases along with the radiation dose. While the increase in dose is almost negligible compared to the amount of dose received during a course of radiation therapy, it cannot be completely ignored. The effective dose from 4DCT can be 2-4 times higher than the effective dose received from a regular

helical 3DCT scan [63, 64]. However, there are efforts to reduce the amount of effective dose from 4DCT scans through iterative reconstruction techniques which can reconstruct 4DCT images of similar image quality using less imaging dose [65]. Another disadvantage of 4DCT is the occurrence of image artifacts. Yamamoto et al. [66] identified four main types of 4DCT image artifacts: blurring, duplicate structure, overlapping structure, and incomplete structure, as shown in Figure 2.11. Blurring artifacts occur due to motion that is faster than the CT gantry rotation speed and happens during a specific couch position. The other three structural artifacts occur at the interface between two adjacent couch positions, which can be caused by projection data being binned to the incorrect respiratory phase or abdominal displacement.

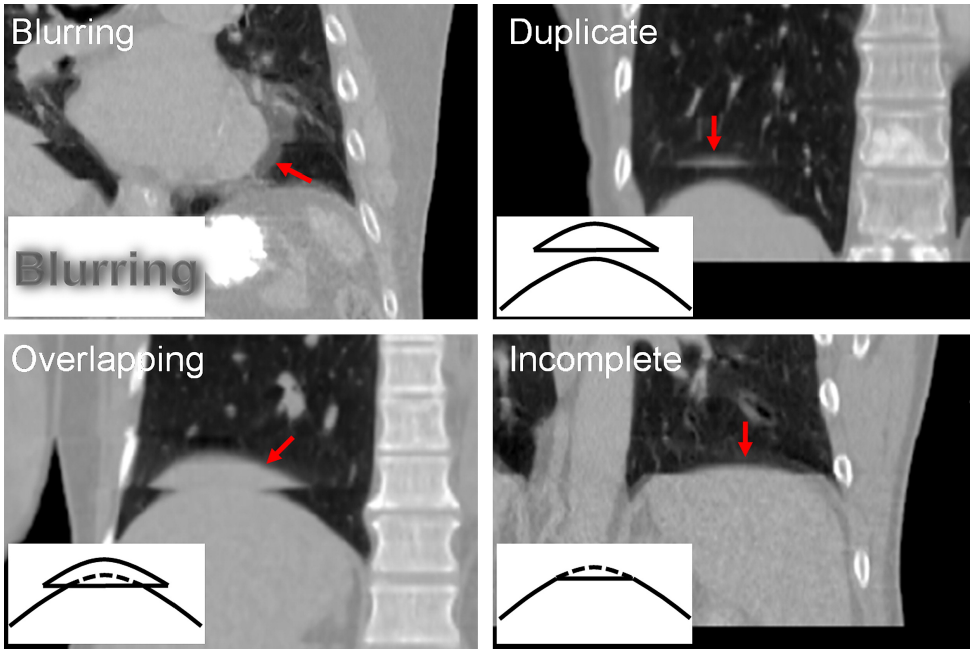


Figure 2.11: Examples of four main types of 4DCT artifacts. Image from Yamamoto et al. [66].

2.4.2 Image Registration

A critical step in deriving ventilation measurements from 4DCT scans is image registration. Image registration in medical imaging is the process of calculating a geometric transformation between a reference (stationary) scan and a moving scan. A generic image registration process can be seen in Figure 2.12. The image registration process is iterative as the algorithm improves the previous transformation iteration based on the information received from an optimizer which is used to guide the image registration algorithm to find the best result of a specified image metric. Image metrics used for image registration can be broken up into two distinct classes: voxel intensity-based and feature-based. Voxel intensity-based metrics use image voxel data directly and include metrics such as correlation coefficient (CC) and sum of the squared difference (SSD). Feature-based metrics use features (anatomical landmarks and organ boundaries) extracted from image data and include metrics such as mutual information (MI) and contour-based comparisons.

Specifically in radiotherapy, image registration has multiple uses including transfer of anatomical segmentations (either from an atlas or prior patient imaging), multi-modality treatment planning (MR, SPECT/PET tumor delineation, etc.), image-guided radiotherapy, and treatment response assessment. Treatment response assessment can include more simple evaluations such as tumor measurements, which requires a rough registration between the two scans, however for assessing pulmonary function change following radiotherapy more precise registration techniques are needed in order to evaluate function at the voxel level. Specifically for pulmonary function assessment, image registration is used to map the deformations that occur between different phases of the 4DCT for each voxel, and can be used to assess lung motion and mechanics. Ultimately, the result of image registration is a transformation matrix consisting of the mappings from the moving image to the reference image.

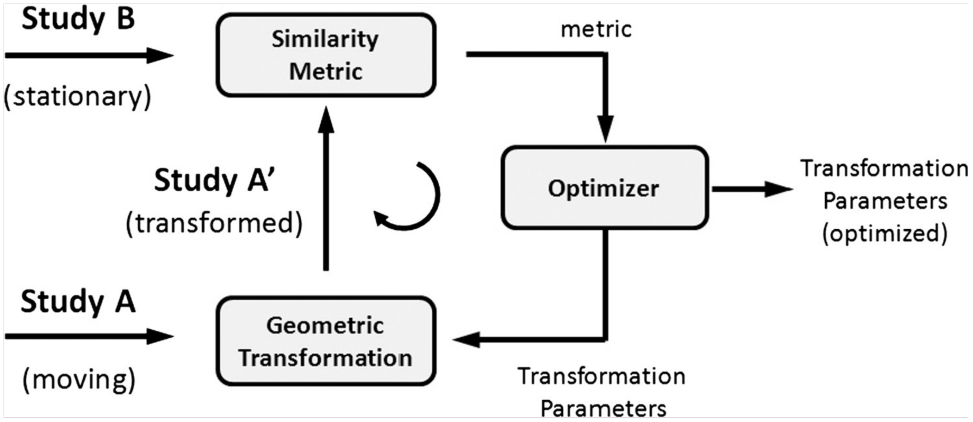


Figure 2.12: Typical workflow of registration algorithms. Image from Brock et al. [67]

This work utilizes an image registration algorithm with multi-resolution cubic B-spline parameterization and uses the sum of squared tissue volume difference (SSTVD) as the similarity metric. The spatial multi-resolution procedure from coarse to fine is used to improve speed, accuracy, and robustness [68]. The initial registration is performed at a downsampled lower resolution image, and once the correspondence is found between the two images, the algorithm uses the transformation produced as a first guess on the next higher resolution image. This process is repeated at multiple resolution levels until a transformation is produced for the original image resolution. Cubic B-spline parameterizations are used to represent the transformation as they are commonly used for shape modeling and arer efficient for nonrigid motion transformations [68]. The SSTVD similarity metric is used in this work because of its specific advantages over the regular SSD for lung image registration. SSTVD accounts for CT intensity change as the lungs expand and fill with air. When the lung tissue fills with air, the HU value of the lung tissue decreases, and during expiration/contraction, the HU value of the lung tissue increases. The volume of tissue $V(x)$ at voxel location x is determined by:

$$V(x) = v(x)\frac{HU(x) - HU_{air}}{HU_{tissue} - HU_{air}} \tag{2.2}$$

The cost function using the SSTVD metric uses the difference between the tissue volumes for each image and is defined as:

$$C_{SSTVD} = \int_{\Omega} [V_2(x) - V_1(h(x))]^2 dx \tag{2.3}$$

where $h(x)$ is the optimized correspondence mapping that maps the moving image (I_1) to the reference image (I_2).

2.4.3 Deriving Ventilation Measurements From 4DCT

Unlike ventilation measurements that can be produced through SPECT/PET or MR techniques, ventilation estimates derived from 4DCT images do not require the patient to inhale gases or undergo additional procedures. Today, there are currently two primary methods of deriving ventilation estimates from 4DCTs, both of which rely on accurate image registration. The first method estimates ventilation based on changes in HU values at the voxel-level between the maximum-inspiration and maximum-expiration CT images [69, 70]. First, the maximum inspiration and expiration images are registered to each other through deformable image registration. Then the fraction of air in each voxel, F_{air}, is calculated using the method first introduced by Simon [70] where

$$F_{air} = -\frac{HU}{1000} \tag{2.4}$$

This method assumes that HU values for voxels within the lung are composed of a linear combination of only air (HU=-1000) and tissue (HU=0). Another assumption is that the volume of tissue does not change between the expiration and inspiration images, meaning that any volumetric and/or HU changes between the two images must be due to addition or subtraction of air. This change in air volume is assumed to be directly correlated with ventilation, and is typically reported as the local volume change due to inspiration (compared to expiration), $\frac{\Delta V}{V_{ex}}$. Using this assumption, ventilation using the HU method can be given as

$$\frac{\Delta V}{V_{ex}} = \frac{F_{ex} - F_{in}}{F_{ex}(1 - F_{in})} \tag{2.5}$$

Lastly, using Equation 2.4 to make a substitution, the final form becomes

$$\frac{\Delta V}{V_{ex}} = 1000\frac{HU_{in} - HU_{ex}}{HU_{ex}(1000 + HU_{in})} \tag{2.6}$$

One of the main disadvantages of the HU-based ventilation metric is that it is susceptible to image quality and noise, which can vary based on scan acquisition parameters and image artifacts. Image noise has an inverse relationship with tube current time product (mAs) where

$$Noise \propto (mAs)^{0.5} \tag{2.7}$$

The second primary method, and is the method used throughout this work, uses the Jacobian determinant of the image transformation matrix produced from the registration between the inspiration and expiration images, as first proposed by Reinhardt et al. [71]. This method assumes that the expansion that a voxel undergoes between the inspiration

and expiration images is due to the addition of air during ventilation, and thus serves as a surrogate ventilation measurement. First, deformable image registration is performed to register the inspiration image to the expiration image, resulting in a transformation matrix, T. The ventilation metric at each voxel is then the Jacobian determinant of T for that specific voxel, where the Jacobian determinant is the determinant of the Jacobian matrix. The Jacobian matrix is a matrix of all first-order partial derivatives in each of the 3 image dimensions, and is given as

$$V_{Jac} = \det Jac(T) = \begin{vmatrix} \frac{\partial T_1}{\partial x_1} & \frac{\partial T_1}{\partial x_2} & \frac{\partial T_1}{\partial x_3} \\ \frac{\partial T_2}{\partial x_1} & \frac{\partial T_2}{\partial x_2} & \frac{\partial T_2}{\partial x_3} \\ \frac{\partial T_3}{\partial x_1} & \frac{\partial T_3}{\partial x_2} & \frac{\partial T_3}{\partial x_3} \end{vmatrix}$$

It is apparent, however, that Jacobian-based ventilation measurements are completely reliant on the accuracy of the image registration algorithm used. Previous work by Cao et al. [68] has shown the B-spline method described in Section 2.4.2 to have a mean landmark error on the order of 1mm, which is the same resolution of the 4DCT scans used throughout this work.

The two ventilation derivation techniques described calculate the local lung expansion ratio (LER) from an end-inhale phase (100IN) and end-exhale phase (0EX), which is defined as 2-phase LER (LER-2) [72]. However, this can lead to errors in the ventilation map due to out-of-phase ventilation. Out-of-phase ventilation can be defined as local lung volume change that is out-of-phase with respect to global lung contraction and expansion, and on average 21.3% of the lung is considered out-of-phase [73]. To correct for out-of-phase ventilation, LER can be calculated from the ratio of the maximum and minimum local lung volume over the entire breathing cycle, defined as LER-N, as proposed by Shao et al. [73]. The LER-N is determined for a voxel by taking the ratio of the maximum to the minimum local lung volume:

$$\text{LER-N}(x) = \max_{i\in\{1\ldots N\}} \mathbb{J}_i(x) / \min_{i\in\{1\ldots N\}} \mathbb{J}_i(x) \tag{2.8}$$

In Eq. 2.8, x is the location of the voxel, $\mathbb{J}_i$ is the Jacobian image of the ith phase, and N represents the number of phases used in the LER-N calculation and can change for a single subject longitudinally to account for effort correction. The number of phases used for each LER-N calculation is determined based on a 100 cc tidal volume criteria between all scans for a given subject. Selecting volumes that have an equivalent tidal volume (ETV) within 100 cc of each other has been shown to improve reproducibility of ventilation measurements [74]. For example, when calculating the LER-N for two scans from a given subject, one scan may have a tidal volume of 1L at the 100IN phase while the second scan may have a tidal volume of 1.2L at the 100IN phase. In this case, the second scan has a tidal volume closer to 1L at the 80IN phase. To account for this difference in tidal volumes, the first scan would use N=9 phases and the second scan would use N=8 phases (i.e. not using 100IN) to calculate their respective LER-N. This example is also illustrated in Figure 2.13.

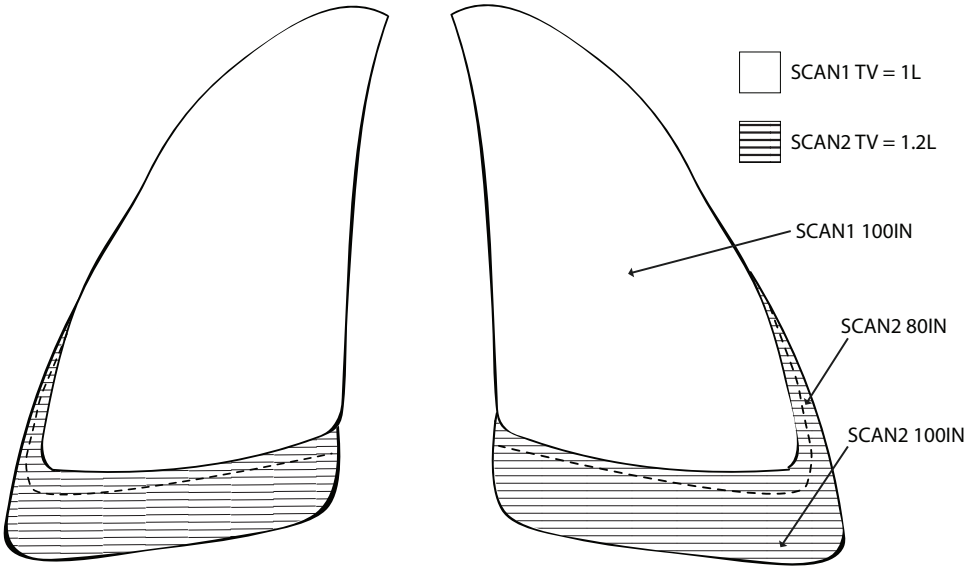


Figure 2.13: Illustration of two scans of one subject where each scan has a different tidal volume at 100IN. For the calculation of LER-N, SCAN1 would use N=10 phases and SCAN2 would use N=9 phases to account for its larger tidal volume. This technique is used as an effort correction strategy to allow for comparison of scans independent of tidal volume. Image from Wallat et al. [75].

Figure 2.14 is a diagram representing two independent voxels and their percentage of relative expansion and contraction throughout the respiration cycle. One voxel experiences out-of-phase ventilation where its maximum expansion does not occur at the maximum breathing phase (100IN) as indicated by the external surrogate. The LER-N calculation corrects for this by taking the maximum and minimum Jacobian values (i.e. amount the voxel expanded) for each voxel from the N phases of the breathing cycle.

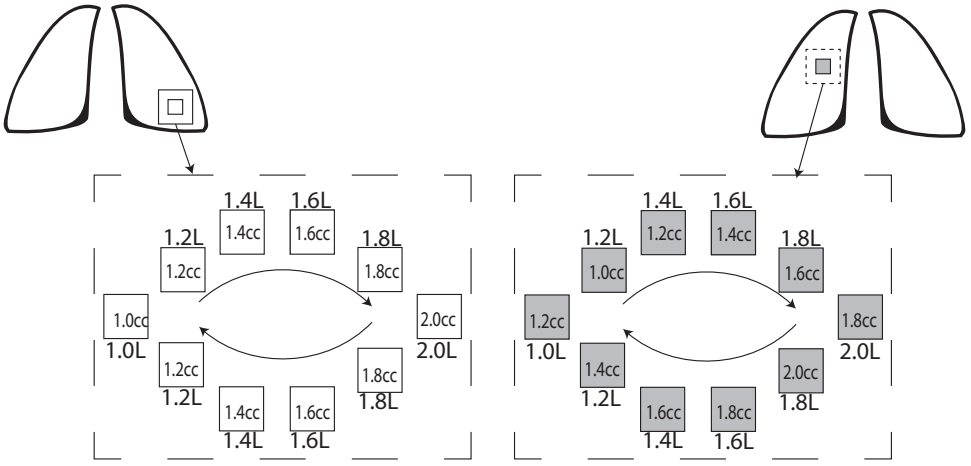


Figure 2.14: Diagram of out-of-phase ventilation. The voxel represented on the left exhibits in-phase behavior as it reaches its own local maximum expansion (2.0cc) at the global maximum expansion (2.0L) The voxel represented on the right exhibits out-of-phase behavior as it reaches its own local maximum expansion (2.0cc) at a tidal volume that is not the global maximum (1.8L). Image from Wallat et al. [75].

One disadvantage of the LER-N method is that it is computationally more expensive compared to the other previously described ventilation derivation methods. Because LER-N uses the maximum and minimum local lung volumes, N (typically N>6, with a max of 9 phases) separate registrations must be made from each of the N 3DCT images to the reference phase 3DCT image (0EX). A depiction of this pairwise registration for LER-N is shown in Figure 2.15 from Shao et al. [73].

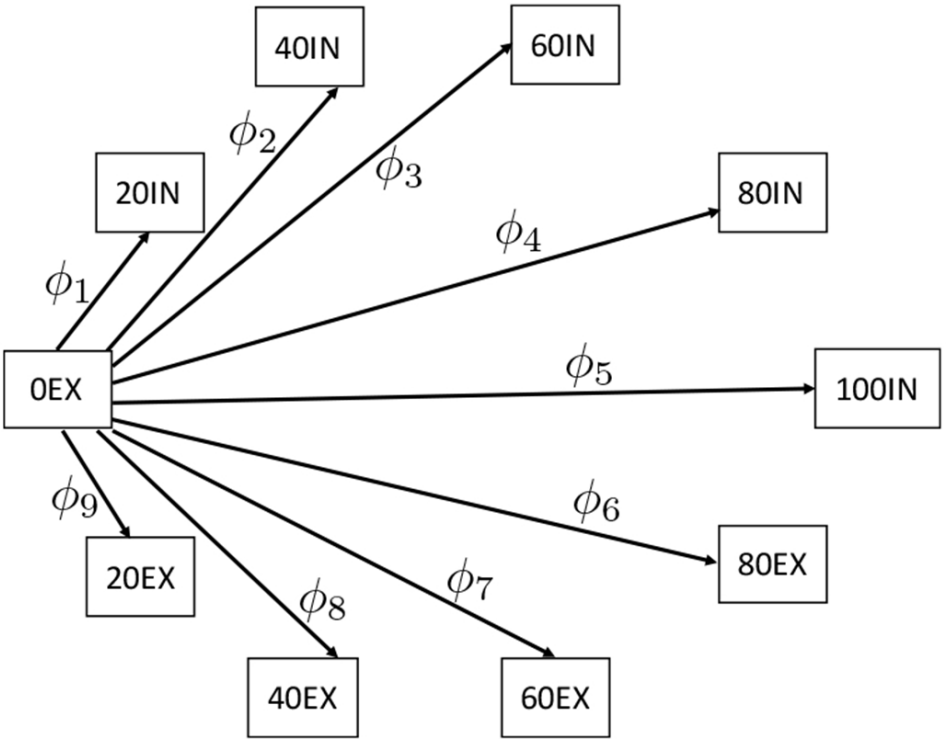


Figure 2.15: Depiction of necessary pairwise registrations for calculation of N-phase local expansion ratio (LER-N). Image from Shao et al. [73] ©2020 IEEE.

2.5 Modeling Pulmonary Functional Damage

Most approaches to functional avoidance RT use a map of a patient's lung function pre-RT to guide the treatment planning process and avoid regions of the lung deemed as high functioning. Lee et al. [76] used ^{99m}Tc-labeled macro-aggregated albumin (MAA) perfusion single photon emission computed tomography (SPECT) imaging to create high perfused lung regions of interest (ROI) that were used as avoidance structures in the treatment planning optimization process. Similarly, Vinogradskiy et al. [77] created avoidance ROIs based on four-dimensional computed tomography (4DCT) ventilation images. Vicente et al. [78] combined serial and parallel functionality by considering dose to airways in addition to parenchymal dose in order to further improve functional

sparing. Previous work by Patton et al. [79] showed a dose-response relationship with post-RT ventilation change, and is the current method used to guide functional avoidance treatment planning in NCT02843568 [5]. Instead of relying solely on pre-RT ventilation, Patton's [79] regression model predicts post-RT ventilation change based on delivered radiation dose and pre-RT ventilation at the voxel level. However, this model struggles to accurately predict ventilation decline; particularly in regions receiving low radiation doses. This is most likely due to the complex nature of pulmonary biomechanics and the resulting changes in ventilation following RT. In this section, pulmonary function modeling and prediction techniques are described.

2.5.1 NTCP Modeling

Radiation therapy is a commonly used treatment modality for cancer, but it can lead to significant toxicity to normal tissues. Normal tissue complication probability (NTCP) models have been developed to predict the likelihood of developing radiation-induced toxicities in various normal tissues. Some commonly utilized NTCP models will be briefly described in this section.

One of the most commonly used models for radiation pneumonitis is the Lyman-Kutcher-Burman (LKB) model [80], which uses a probit function to estimate the probability of a complication based on the dose and volume of irradiated lung tissue. The LKB model can be described in Equations 2.9-2.11, where D_i is the dose and v_i is the relative volume of the i-ith bin of the differential DVH, n is a parameter that describes the volume dependence of the tissue, m is the slope of the model fit, and D_{50} is the uniform dose given to the organ of interest that results in a complication probability of 50% [80]. A model was built by Kwa et al. [81] on data from 540 patients and they found that mean lung dose is a useful predictor of radiation pneumonitis. Although this model was simple,

it as been shown to perform as well as other complex models [82]. However, the LKB assumes a normal distribution of lung sensitivity and has a lack of consideration of the spatial distribution of radiation dose.

$$NTCP_{LKB} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx \tag{2.9}$$

$$t = \frac{D_{eff} - D_{50}}{mD_{50}} \tag{2.10}$$

$$D_{eff} = (\sum_{i} v_i D_i^{\frac{1}{n}})^n \tag{2.11}$$

The critical volume model developed by Niemierko and Goitein [83] is a mathematical model based on the binomial distribution of the dose-response curve, and uses the concept of "functional subunits" (FSUs). The fraction of surviving FSUs following irradiation determines the probability of normal tissue complications. Vågane and Olsen [84] modeled radiation pneumonitis in mice using a critical volume model and found spatial variation in the irradiation of partial lung volumes, supporting the hypothesis that lung function (and radiation-induced damage) may be regionally dependent.

Lastly, logistic regression is a simple but effective modeling approach to calculate probabilities of a binary outcome based on one or more predictive variables. In general, the probability of a toxicity based on a logistic regression model can be described by Equation 2.12 where β_0 and β_1 are the constant and regression coefficients, respectively. This model can be expanded to include additional regression coefficients depending on the number of significant predictor variables that are included. Logistic regression has been

used to determine significant predictors of radiation pneumonitis using irradiated volume of highly-functional lung [28], airway dose [85], and functional mean lung dose [86].

$$NTCP_{logistic} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + ...)}} \tag{2.12}$$

Despite the progress that has been made in developing NTCP models for lung toxicity, there are still challenges that need to be addressed. For example, the models need to be validated in larger and more diverse patient populations, and there is a need to develop models that can predict other types of lung toxicities, such as radiation-induced fibrosis and airway toxicities. Additionally, these statistical models require the breakdown of dose distributions into DVHs, losing spatial relationships and global context of surrounding tissues. As machine learning has grown exponentially over the past decade, radiation-induced toxicities can be modeled in a more comprehensive and precise manner.

2.5.2 Machine Learning

The beginning of modern artificial intelligence (AI) can be traced back to Alan Turing when he published his work titled "Computing Machinery and Intelligence" in 1950 [87, 88]. At the time of his publication, computer hardware and capabilities were still years behind the theory, and it wasn't until the 1960s and 1970s that AI flourished. This increased activity in the field is depicted in Figure 2.16. More recently, AI and machine learning (ML) have been utilized to develop models that are able to recognize and diagnose different diseases using diagnostic imaging [89, 90]. By leveraging the data-driven techniques of AI/ML to solve problems within healthcare, patient outcomes may be improved beyond the capabilities of human-driven analysis and computation.

Machine learning lends itself very well to functional lung imaging research as the bio-physiological processes that occur within the lung are difficult to model, and newer deep learning -based approaches may be robust enough to exceed currently existing techniques and technologies. Currently, functional lung imaging and research is based on physical properties of the image such as calculating local CT density changes [91], calculating regional volume changes using the Jacobian from deformable image registration (DIR) [71], or any of the other various techniques discussed in Section 2.3. Recently there have been works published that use machine learning to predict radiation-induced pneumonitis [92–95], ventilation maps [96–100], perfusion maps [101–104], segmentations [105–108], and DIR [109–112]. While machine learning is a broad category that describes many different types of architectures and networks, medical image analysis, including techniques used specifically for functional lung imaging, mainly rely on deep learning. This section will introduce deep learning, including popular architectures and networks used in medical image analysis, as well as an overview of model training. Note: for the purposes of describing image sizes and dimensions, a "channels last" notation will be used. For example, an image with dimensions of 128x128x128x3 refers to a 3D multi-channel image with a height, width, and depth of 128 voxels and 3 channels (h x w x d x c).

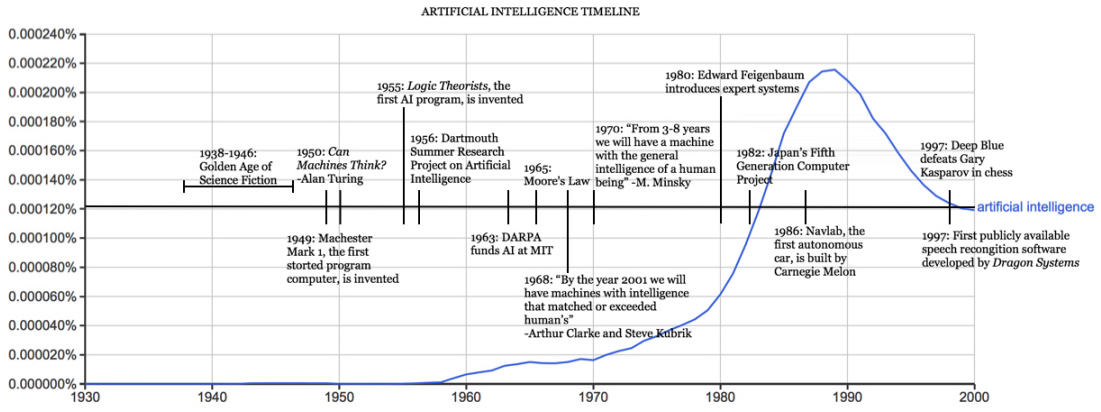


Figure 2.16: Timeline of the evolution of artificial intelligence (AI). Image from Anyoha [87].

2.5.2.1 Deep Learning

Convolutional neural networks (CNNs) are a subset of artificial neural networks (ANNs), which all fall under the category of deep learning. Deep learning is a specific type of machine learning algorithm which integrates feature extraction and output prediction into one combined architecture. Convolutional neural networks are commonly used for image and video analyzation tasks due to its ability to learn patterns in spatially correlated data. Typical input to a CNN is a multidimensional tensor or array (i.e. a 3D image stack), containing data that is spatially correlated. A CNN will have multiple layers, i, with each layer taking an input I_i and transforming it to an output I_{i+1}. Through each transformation and layer, the spatial relationships within the data is preserved. The most common types of transformation used in a layer include convolutional, transposed convolution, pooling, and nonlinearity. Each layer has a set of parameters that are learned and updated throughout the training process using a specified optimizer.

Convolutional layers perform a convolution operation using the input image and a specified kernel. Figure 2.17 depicts how basic 2D convolution works. For each convolutional layer, the amount of filters, F, is set by the user and specifies the amount of times the convolution is performed. The layer then outputs F number of convolutions which are concatenated along the channel dimension creating an activation map. The model learns and updates the kernels throughout the optimization process.

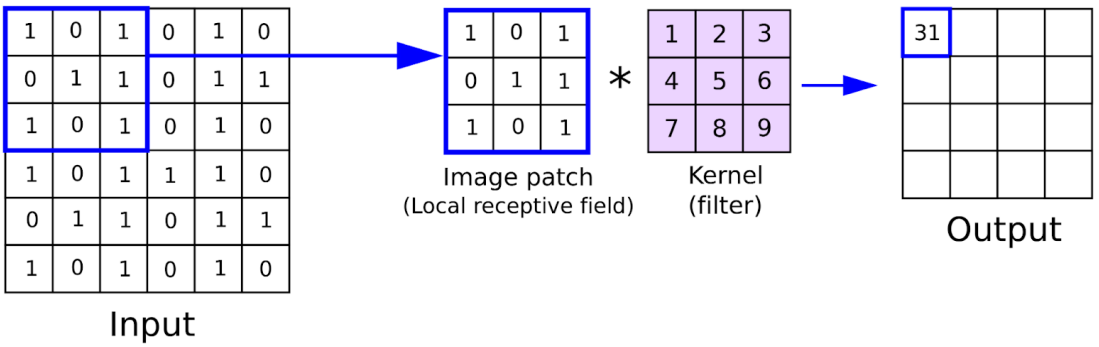


Figure 2.17: Convolutional filter applied to data. Image from LENDAVE [113].

Another common layer used is a pooling layer. A pooling layer is used to reduce the spatial size of the input. For example, a pooling layer with a pooling size of 2 and an input image with a size of 30x30x30x3 outputs an image with size 15x15x15x3. Pooling layers do not have any learnable parameters, however they are useful as they reduce the spatial dimensions of the images which reduces the amount of parameters and memory needed to train. This is especially useful when using large images and/or filters.

Lastly, an essential part of deep learning architectures is the nonlinearity layer. These layers are typically used following a convolutional layer to introduce nonlinearity into the feature maps. Similar to the pooling layer, there are no learnable parameters in the nonlinearity layers, however they allow the model to develop complex mappings between the inputs and outputs. Common nonlinearity functions include hyperbolic tangent, sigmoid, and rectified linear units (ReLU), and are shown in plots in Figure 2.18.

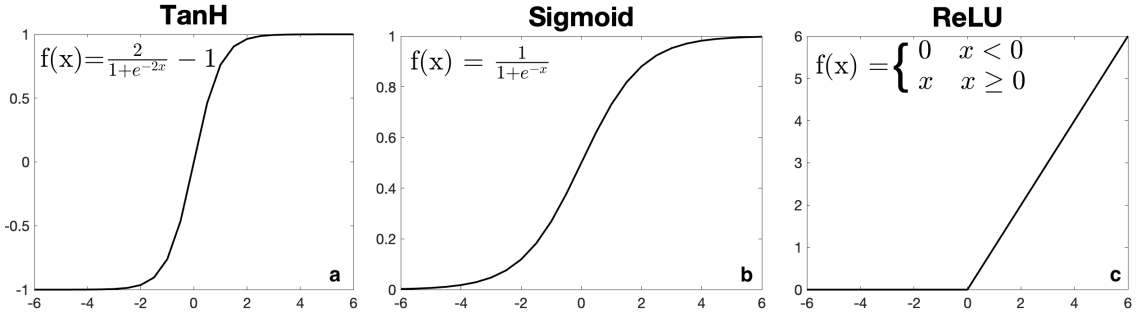


Figure 2.18: Common nonlinearity functions used in deep learning networks: (a) hyperbolic tangent, (b) sigmoid, and (c) rectified linear units (ReLU).

2.5.2.2 Common Architectures

In deep learning, there are various types of architectures that have been developed to better solve specific types of problems. As it pertains to medical imaging, these tasks that can be solved by deep learning include segmentation, classification, image regression/-translation, etc. Certain types of architectures are better suited to solve a particular problem than others, which has led to an explosion of different machine learning networks. This section will discuss a few of those networks most relevant to the content presented in this work.

As machine learning problems became more complex, additional layers were added to networks to create a deeper architecture. It was hypothesized that deeper networks have the ability to integrate information from low to high level features [114], and the levels of features can be increased by adding more depth to the network. However in a network with increasing depth, accuracy becomes saturated which then leads to degradation of the model [115]. He et al. [115] proposed solution to solve this problem was to use residual blocks which contain skip connections. Skip connections, as seen in Figure 2.19, allow for the input to skip one or more layers which is then concatenated with the output from the layers. Specifically in the first ResNet architecture, known as Resnet-34, the

skip connection performs an identity mapping which is then added to the output [115]. Using residual blocks with skip connections allows for deeper networks without increasing complexity. When compared to the VGG net [116] and other similar very "deep" neural networks that performed well on the popular ImageNet dataset [117], the Resnet-34 network won the ILSVRC 2015 classification challenge [118] while having more layers and less complexity that its competitors.

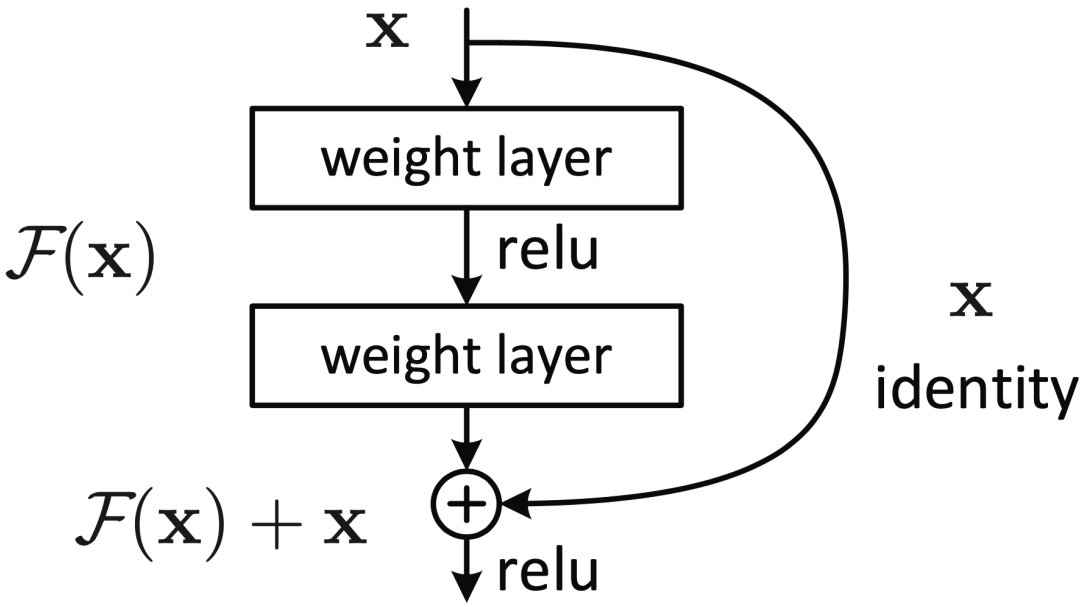


Figure 2.19: Diagram of a residual block. Image from He et al. [115] ©2016 IEEE.

Another popular network, especially in the field of image segmentation, is the U-Net, developed by Ronneberger et al. [119]. U-Net was specifically designed for semantic segmentation, which is the labeling of each voxel or pixel to its corresponding class. The structure of the U-Net architecture consists of two main parts: the encoder and the decoder. The encoder, also known as the contracting path, follows a similar pattern to typical convolutional networks. At each level of the encoder, the input will go through two convolutions, each followed by a nonlinear activation (typically ReLU) and a pooling operation with a stride of 2 for downsampling. At each of the downsampling steps the

number of features in the map is doubled. In the decoder path, also known as the expansive path, upsampling occurs through a transposed convolution with a stride of 2. The upsampled output is then concatenated with the output from the encoder path of the same level, which is followed by two convolutions with ReLU activation. This process is repeated for each level until the output is back to the original input image dimensions where a final convolution is performed to map each voxel/pixel to the desired number of classes. Overall, the architecture follows a mostly symmetrical style, which can be seen in Figure 2.20, and is the reason it is referred to as U-Net.

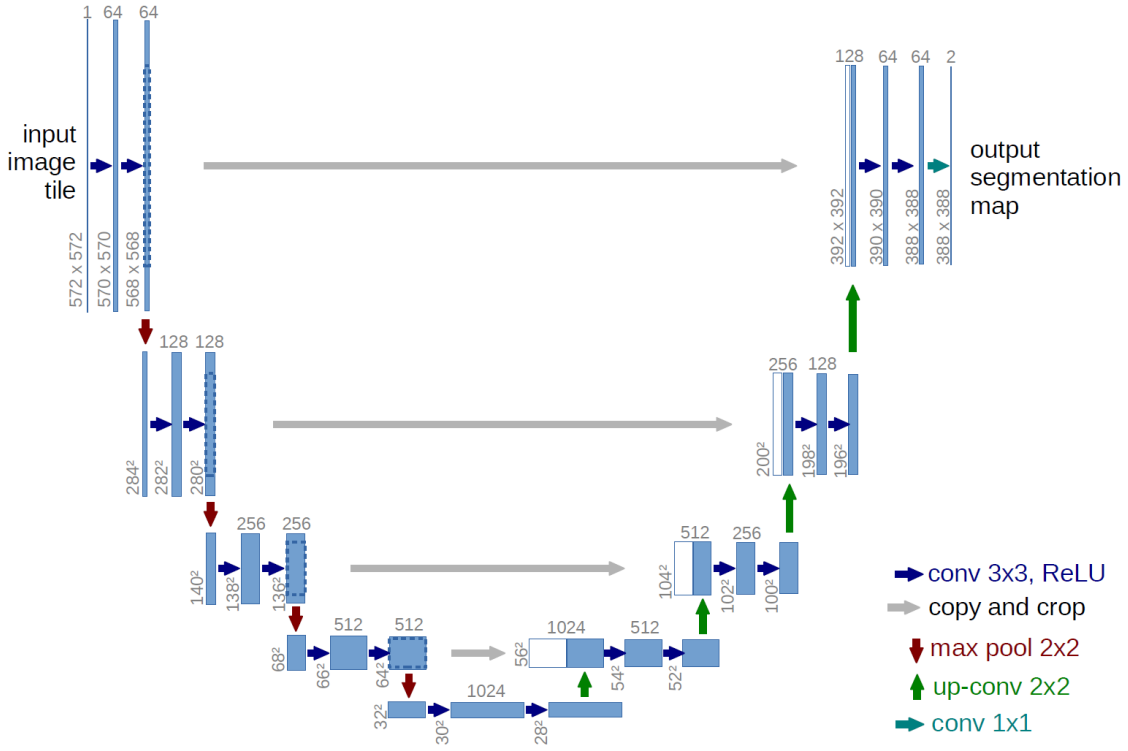


Figure 2.20: Diagram of the U-Net architecture. Image from Ronneberger et al. [119].

Lastly, generative adversarial networks (GANs) have become a popular tool for image generation and image translation tasks, and originally proposed by Goodfellow et al. [120]. A GAN is typically made up of two networks: a generator and a discriminator.

The generator used in a GAN is usually chosen based on the problem to be solved. In image segmentation or regression tasks, it is common to use the U-Net architecture as the generator network. The discriminator usually follows a typical convolutional network style used for classification. While a CNN, like U-Net, aims to optimize its solution by comparing its output to a ground truth, a GAN network has the generator and discriminator networks work towards opposite goals. The generator tries to fool the discriminator network by minimizing the error between the ground truth and generated image. The discriminator tries to identify fake data from ground truth data by minimizing the error between ground truth and fake generated data. A diagram of a general GAN architecture is depicted in Figure 2.21.

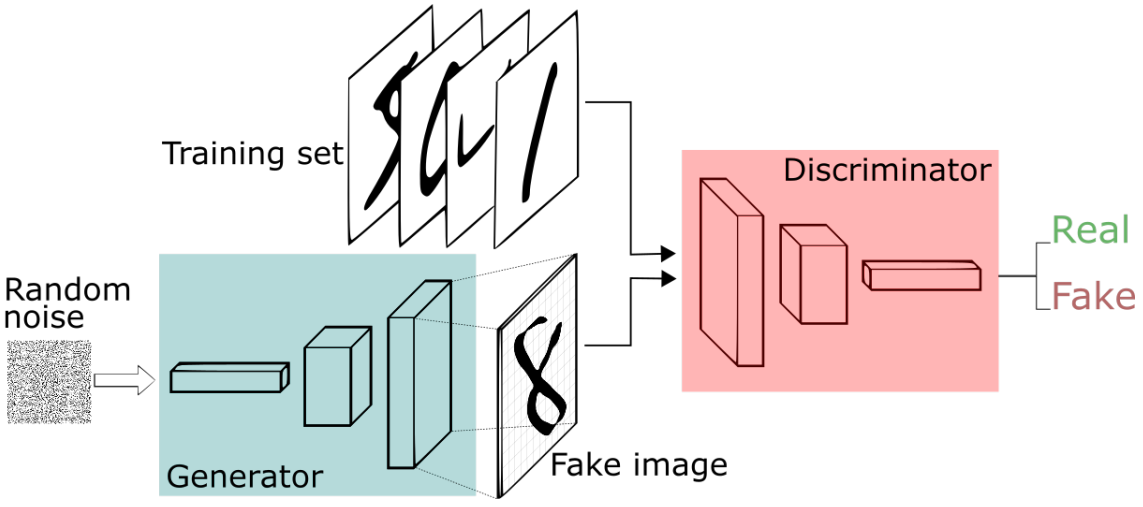


Figure 2.21: Diagram of a general GAN architecture. Image from Silva [121].

2.6 Discussion

The aims in this work were developed to address specific issues that affect the efficacy in which pulmonary functional information can be applied to lung cancer RT. First, there is a need to characterize the effects of radiation on the lower respiratory tract and resulting

lung function. Kazemzadeh et al. [29] investigated and created a risk model of airway collapse based on the maximum dose and diameter of a bronchial segment. Vicente et al. [30] continued this work to create a functional weighted airway sparing (FWAS) map to avoid specific airways based on the function of the sub-volume they supplied. While these studies looked at maximum doses to airway structures that ended in collapse, airway structural changes that occur prior to atelectasis and the subsequent effect on ventilation have not been thoroughly studied. Additionally, Manyam et al. [32] investigated the results of NRG Oncology/Radiation Therapy Oncology Group (RTOG) 0813 and performed clinical validation of the proximal bronchial tree (PBT) constraints for 5-fraction SBRT. It was found that a maximum point dose PBT constraint of $D_{0.03cc} \leq 50$ Gy had the best sensitivity and specificity for predicting grade 2 to 5 non-pneumonitis toxicity (NPT). However, no work has been done to quantify dose constraints for airway segments beyond the PBT to limit NPT. The purpose of this work's first aim is to investigate changes in imaging biomarkers as precursors to atelectasis post-RT, and develop a normal tissue complication probability (NTCP) model based on these results.

Second, there is a need for a volumetric dose-response model that more accurately predicts direct damage to the lung parenchyma and indirect effects of dose to airways that supply regions outside of the main dose distribution. Previous works have developed models to predict toxicity outcomes, such as radiation pneumonitis and fibrosis, as a result of dose to functional lung; however, they all fail to include specific regional predictions of functional decline [28, 92, 94, 122]. An additional work has created a predictive dose-response model, but only considered the relationship with dose to a voxel within the lung [123]. These studies did not include changes in ventilation that might occur due to bronchial damage. Additionally, these predictive models were limited because they lost any spatial relationship between neighboring voxels and did not account for out-of-phase ventilation. The purpose of the second aim of this work is to develop a dose-response model that more

accurately predicts direct and indirect ventilation damage, preserves spatial relationships between neighboring voxels, and accounts for out-of-phase ventilation.

Once accurate predictions of post-RT ventilation are produced, that information still must be incorporated into treatment plans in order to preserve lung function and lower the probability of toxicities occurring. Currently, discrete regions of interest (ROI) are created based on 10 Gy dose bins from 10 to 60 Gy using the information from the predictive dose-response model created by Patton [123]. There is a need to develop a method in which the functional information can be applied to the treatment plan in a continuous fashion, rather than discrete ROIs. It is hypothesized that the results of the new functional avoidance plans will predict significantly better pulmonary function preservation than the current approach. Furthermore, implementing a feedback system into the workflow will be advantageous for comparing the potential amount of functional lung spared between SOC and optimized plans. The goal of this third aim is to improve the efficiency and consistency in which functional avoidance information is applied to treatment plans through the development of a functional avoidance treatment planning workflow.

Chapter 3

Investigation of radiation-induced airway changes and downstream ventilation decline [1]

3.1 Introduction

Functional avoidance radiation therapy (RT) has mainly focused on the direct results of irradiating healthy, normal lung tissue, however there has been little work describing the impact of irradiating the airways and potential disruption to ventilation function in the regions those airways supply.

[1] Portions of this work have been published [124]

3.1.1 Radiation-induced airway toxicity

The mechanisms and impact of radiation-induced airway toxicities and the subsequent change in downstream ventilation function are still not well understood in the context of pulmonary functional avoidance radiotherapy. There have been a few previous studies that have investigated the relationship between dose and atelectasis in patients who received stereotactic body radiation therapy (SBRT). Kazemzadeh et al. [29] investigated complete airway collapse based on airway metrics collected through virtual bronchoscopy. They found through logistic regression that airway segmental diameter ($p=0.014$) and maximum dose ($p=0.007$) were significant predictors of airway segment collapse. Furthermore, they found that there was a 1.07 times higher chance of airway collapse for every 1 Gy increase in maximum dose after controlling for airway caliber [29]. Vicente et al. [30] used the toxicity model developed by Kazemzadeh et al. [29] in order to create a way in which to account for the relative functional value of each terminal airway. This led to what Vicente et al. [30] referred to as a functional weighted airway sparing (FWAS) map. Each bronchial pathway was mapped to a functional sub-lobar lung volume which determined the relative ventilation contribution of that particular sub-volume/bronchial pathway pair. Using the toxicity model from Kazemzadeh et al. [29], treatment plans were created to avoid a predicted threshold of airway collapse and it was found that treatment plans created using the FWAS resulted in superior ventilation preservation compared to a clinical standard-of-care treatment plan [30]. Atelectasis is an extreme endpoint for airway toxicity, however, and it is hypothesized there are precursors to atelectasis, such as bronchial stenosis, that may have negative effects for downstream ventilation function post-RT.

Additionally, there has been limited work focused on radiation-induced airway remodeling and toxicity models to use for radiation treatment planning. Miller et al. [16] assessed bronchial stenosis in patients treated with high dose (> 70 Gy) twice-daily radiotherapy

and found bronchial stenosis to be a significant clinical complication. They additionally acknowledged there seemed to be a dose-response relationship to the development of bronchial stenosis, but were not sufficiently powered in their analysis to make any conclusions or conduct proper analysis. Wang et al. [31] focused on central airway toxicites within the proximal bronchial tree (PBT) for subjects treated with conventionally fractionated thoracic three-dimensional conformal radiation therapy (3DCRT). They found that 17% of their patients developed PBT toxicities and that V75Gy was a significant predictor for both grade 1+ and 2+ PBT toxicities [31]. The recommended thresholds for V75 of the PBT were 6.8% and 11.9% for grade 1+ and grade 2+ PBT toxicity, respectively [31]. Lastly, Manyam et al. [32] performed validation of the results from the NRG Oncology/Radiation Therapy Oncology Group (RTOG) 0813 report. They considered multiple nonpneumonitis toxicities (NPT) which included stenosis, necrosis, fistula, hemoptysis, and clinically significant pleural effusion [32]. They validated the current RTOG PBT constraints and suggested a slightly lower maximum point dose to the PBT of $D_{0.03cc} \leq 50$ Gy [32].

The purpose of the work presented in this chapter is to retrospectively analyze radiation-induced changes in airways as precursors to atelectasis post-RT. This work presents quantitative relationships between airway dose and airway structural changes, as well as indirect local changes in ventilation in regions supplied by irradiated airways. The first section will perform a preliminary retrospective analysis in a novel swine model, and the second section investigates these mechanisms and responses within a human subject cohort.

3.2 Radiation-induced airway changes and downstream ventilation decline in a swine model

3.2.1 Materials and Methods

3.2.1.1 Animal Model and Setup

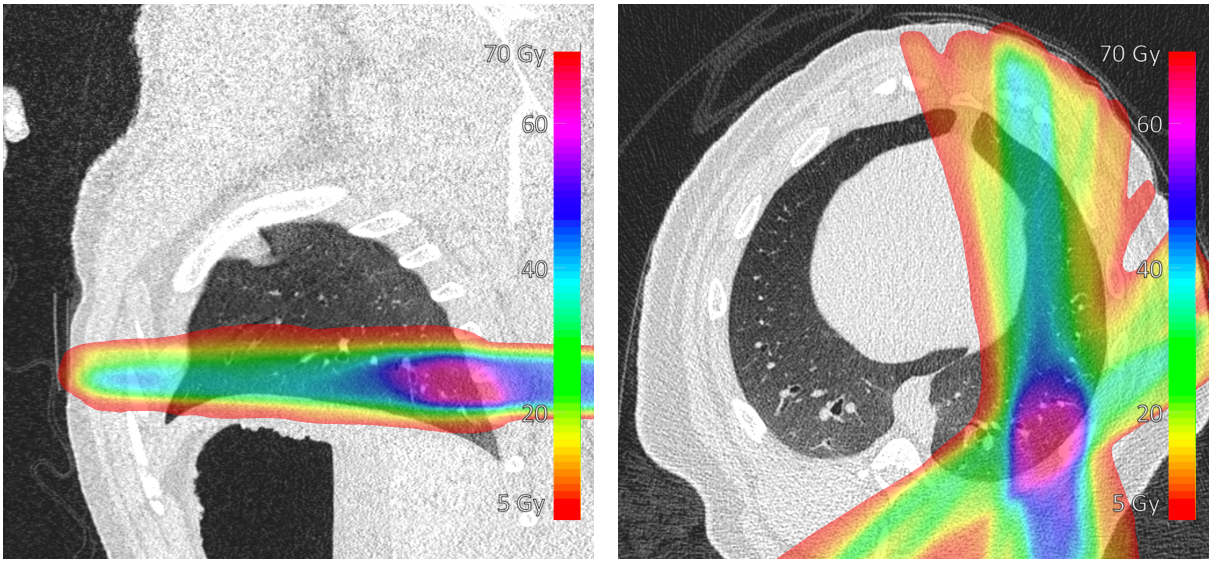
Collecting data and characterizing dose-response relationships is a current challenge in studying functional avoidance RT in humans. There is a lack of consistency between 4DCT scans in humans due to irregular breathing, which leads to unreliable function maps. Additionally, the treatment plan, dose distribution, and anatomy for each patient is vastly different, making it difficult to observe subtle changes in anatomy due to irradiation. The use of swine in radiotherapy research has been favorable due to their physiological similarity to humans [125–127]. Pitfalls of using conventional swine include their larger size and growth rates, which can prohibit longitudinal investigations of biomarker changes. This work uses the Wisconsin Miniature Swine (WMS™) [128], a breed specifically developed to have greater physiological and size similarity to humans compared to conventional swine. At full maturity mini-swine weigh 68-91 kg, which is comparable to the size of an average human male whereas conventional swine can weigh as much as 300 kg. This allows for more accurate development and validation of techniques to be eventually used with human subjects.

A total of three WMS™ from a previous experiment focused on perfusion [129] were retrospectively analyzed. The WMS™ were mechanically ventilated to have a respiratory rate of 15 breaths per minute and a consistent tidal volume of 1 L in order to match the average respiratory rate and tidal volume of human subjects. All subjects were sedated to eliminate possibilities of motion artifacts during imaging and uncertainties during

treatment delivery. Details regarding animal care and drugs administered can be found in Appendix A. All procedures as well as animal care practices were approved by an Institutional Animal Care and Use Committee (IACUC).

3.2.1.2 Treatment Scheme

Each WMS™ was prescribed a research course of 60 Gy to 95% of the planning target volume (PTV) in 5 fractions approved by IACUC. Treatments used a 6 MV flattening filter free (FFF) beam, with each fraction consisting of 10-13 beams at various gantry angles ranging from 0 to 200 degrees. Treatments were performed on the ViewRay (ViewRay Inc., Cleveland, OH) treatment system in order to maximize dose conformity and reduce the uncertainty of dose delivery due to respiratory motion. Using ViewRay's integrated 0.35T MR imaging capabilities, subjects were imaged at four frames per second in the sagittal plane to track adjacent pulmonary vasculature and target in order to restrict beam delivery to the centroid of the target. ViewRay treatments use a step-and-shoot intensity modulated radiation therapy dose delivery technique. Beam-on time ranged from 7-12 minutes per fraction, and the duty-cycle was approximately 60%. These fractions were delivered following a standard clinical SBRT schedule receiving fractions with a day in between each delivery during weekdays and 2 days over the weekend. The PTV was designated as the bifurcation of a vessel in the left inferior lung and the right lung was left unirradiated (max point dose <5 Gy). Since this experiment was not designed to study the effects of airway irradiation on downstream lung function, there was approximately 100cc of lung volume inferior to the PTV that received less than 5 Gy. Figure 3.1 shows the sagittal and axial view of a treatment plan for one WMS™ subject, and is representative of what each subject received.



(a) Sagittal View **(b)** Axial View

Figure 3.1: Representative treatment plan from one of the WMS™ subjects.

3.2.1.3 4DCT Acquisition

Each swine underwent two 4DCT scans pre-RT and two additional 4DCT scans 3 months post-RT. All scans were acquired on a Siemens SOMATOM Definition Edge 128 slice CT scanner with a 0.6 mm slice thickness, 0.5 second tube rotation time, 0.09 pitch, 120 kVp, and 100 mAs/rotation. Respiration induced movement was monitored using Varian Real-Time Position Management (RPM) (Varian Medical Systems Inc., Palo Alto, CA), which captured the respiratory traces from the swine for the 4D reconstructions. Images were reconstructed using a sharp reconstruction kernel and with a 0.6 mm x 0.6 mm x 0.6 mm voxel size. Each 4DCT was reconstructed into 10 discrete phases of the breathing cycle.

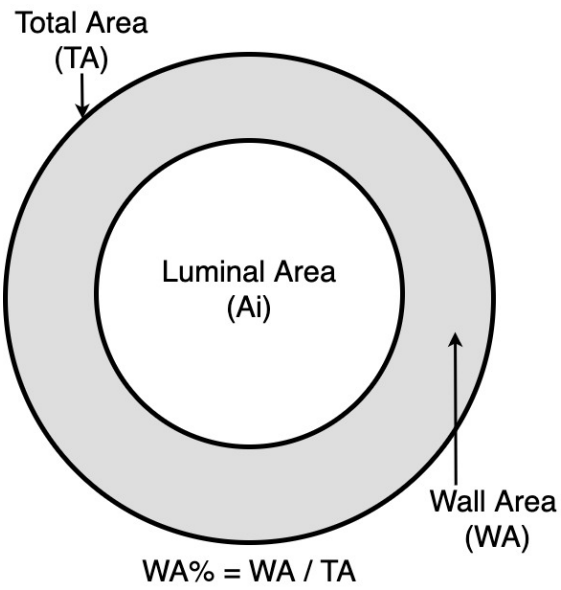
3.2.1.4 4DCT-derived Ventilation & Analysis

Ventilation maps were calculated for each of the swine and for the two timepoints (pre- and post-RT). In this work, the lung ventilation maps were derived using the Jacobian method as described by Shao et al. [73] and in Section 2.4.3. This method of calculating the local lung expansion, referred to as N-phase local expansion ratio (LER-N), uses N phases of the 4DCT to account for out-of-phase ventilation. The total change in lung volume over the breathing cycle was very consistent across all scans and subjects since the miniature swine were mechanically ventilated. Due to this consistency, N=10 phases were used for the calculation of the ventilation maps. Jacobian maps for the two scans at each time point were geometrically averaged to reduce noise within the map. Lastly, Jacobian ratios were calculated by taking the post-RT map divided by the pre-RT map, meaning values less than one describe regions that experienced a decline in ventilation/function post-RT. The regions of ventilation investigated were the inferior left lung that received less than 5 Gy (ILL), the superior left lung that received less than 5 Gy (SLL), and the contralateral right lung (RL). The 5 Gy limit for the regions was a max point dose. All regions analyzed had a mean dose near 1 Gy or less. It was assumed that the ILL region was supplied by irradiated airways, which is hypothesized to have a decrease in function due to a downstream effect of damaging the supplying airways.The SLL and RL airways were relatively unirradiated ($<$ 5 Gy max point dose), and are hypothesized to be regions that show little to no change in function.

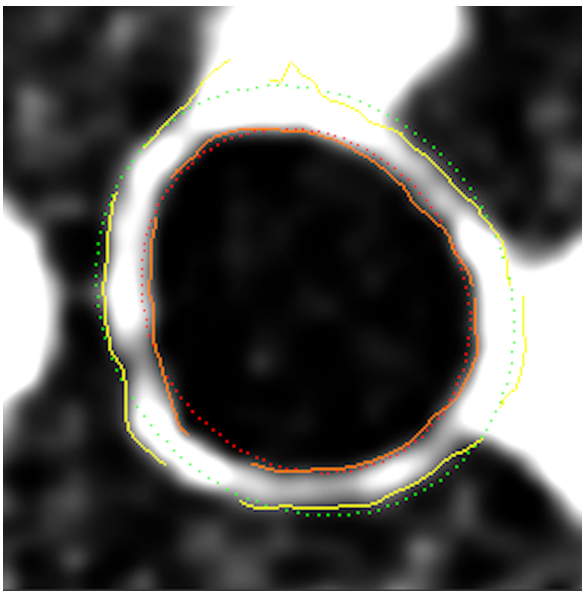
3.2.1.5 Airway Analysis

Airway segmentation was performed using 3DSlicer [130] and the Chest Imaging Platform (CIP) module (Applied Chest Imaging Laboratory, Boston, MA) on the maximum inhale phase from the pre- and post-RT 4DCTs. The single kernel phase congruency (SKPC)

method was used to segment all airways as this has been shown to give a better localization of the airway wall compared to the full width at half max (FWHM) method and is less sensitive to variations in reconstruction kernels and imaging dose [131]. Airways with inner major diameters between 4-16 mm were segmented in each lung for each swine, resulting in a total of 53 airways that were analyzed. A total of four segmentations (two scans pre-RT and two scans three months post-RT) were created for each airway to characterize the uncertainty in the measurements. Longitudinal changes from the pre- and post-RT scans were then compared by calculating the percent difference from the averaged metrics. The luminal area (Ai) and square root of wall area ($\sqrt{WA}$) for each airway were investigated as they have been shown previously to correlate with pulmonary function [132–134]. Airway segmentations were performed on 2D axial slices along the airway tree centerline, which allows for measurements of Ai and WA at that particular location. Figure 3.2b shows an example of one segmentation that was performed, where the inner and outer luminal contours are indicated in orange and yellow, respectively. The inner and outer luminal contours of the airway without reformatting along the airway axis to obtain an orthogonal slice view of the airway are indicated in red and green, respectively. Dependence of airway size was also investigated by examining the relationship between pre- and post-RT Ai sizes for a 30±5 Gy dose bin. Figure 3.2a shows a diagram of Ai and wall area in relation to a bronchial segment. WA% is calculated by taking the WA divided by the total airway area (TA). All calculations were performed automatically by the CIP module. The dose to an airway was calculated as the near-maximum dose, or the dose to the highest 2% ($D_{2\%}$), at the axial slice the airway segmentation was performed.



(a) Diagram of the airway metrics in relation to a bronchial segment



(b) Example 2D segmentation of an airway, where the inner and outer luminal contours are indicated in orange and yellow, respectively.

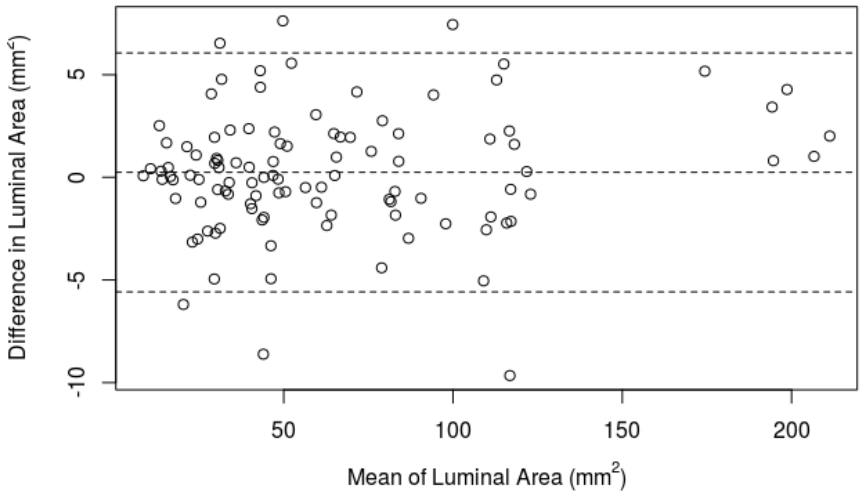
Figure 3.2: Diagram and segmentation of airway

All statistical analysis was conducted in R [135]. Repeatability of airway segmentations and metrics were assessed by Bland-Altman analysis [136] with limits of agreement set at a 95% level. The Pearson correlation coefficient was used to determine the correlation between airway metrics and dose. The Shapiro-Wilk test was used to test data for normality. Assuming a normal distribution, correlations and percent changes were tested for significance using t-tests.

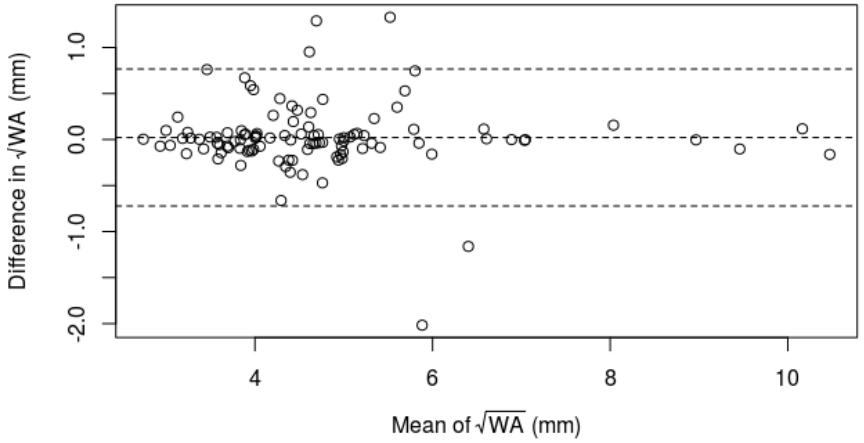
3.2.2 Results

The results of the Shapiro-Wilk test for normality for both Ai and $\sqrt{WA}$ measurements rejected the null hypothesis (p>0.1), meaning the data is normally-distributed. Bland-Altman plots for repeatability measurements of Ai and $\sqrt{WA}$ are shown in Figure 3.3. Both figures use measurements from pre- and post-RT airway measurements. Figure

3.3a shows Ai had excellent repeatability across all segmentations and metrics for both timepoints. Figure 3.3a shows no discernible trend in the data. Figure 3.3b shows larger differences in $\sqrt{WA}$ which may be due to the segmentation tool's ability to discern the edges of the airway wall, which may be especially variable in regions of fibrosis. In addition, some of the variability of the $\sqrt{WA}$ is due to measurement error of the airways which have wall thickness smaller than the image reconstruction resolution.



(a) Luminal area Bland-Altman plot



(b) $\sqrt{WA}$ Bland-Altman plot

Figure 3.3: Repeatability of luminal area and $\sqrt{WA}$ measurements using repeat scans from both pre- and post-RT. Scatter plots shows variation in measurement with Bland-Altman analysis. Top and bottom dashed lines represent the 95% limits of agreement. The middle dashed line represents the mean difference.

Figure 3.4 represents Ai percent change as a function of airway dose for the airways of the left irradiated lung for the three swine subjects, and shows Ai decreases with dose and

that percent change increases linearly with dose. The airway measurements of the right lung were not included in this plot as to not bias the fit with a majority of points located near the origin. This was due to the right lung/airways not receiving doses larger than 5 Gy or experiencing significant changes in Ai. The airways in the right (unirradiated) lung for all swine showed no significant change ($p=0.48$) in Ai post-RT compared to pre-RT, with a mean fractional change of 0.0072 ± 0.028. The airways in the inferior left (irradiated) lung of all swine were found to have a decreased Ai post-RT compared to pre-RT, and was significantly ($p<0.001$), negatively correlated (Pearson $R = -0.98$) with radiation dose.

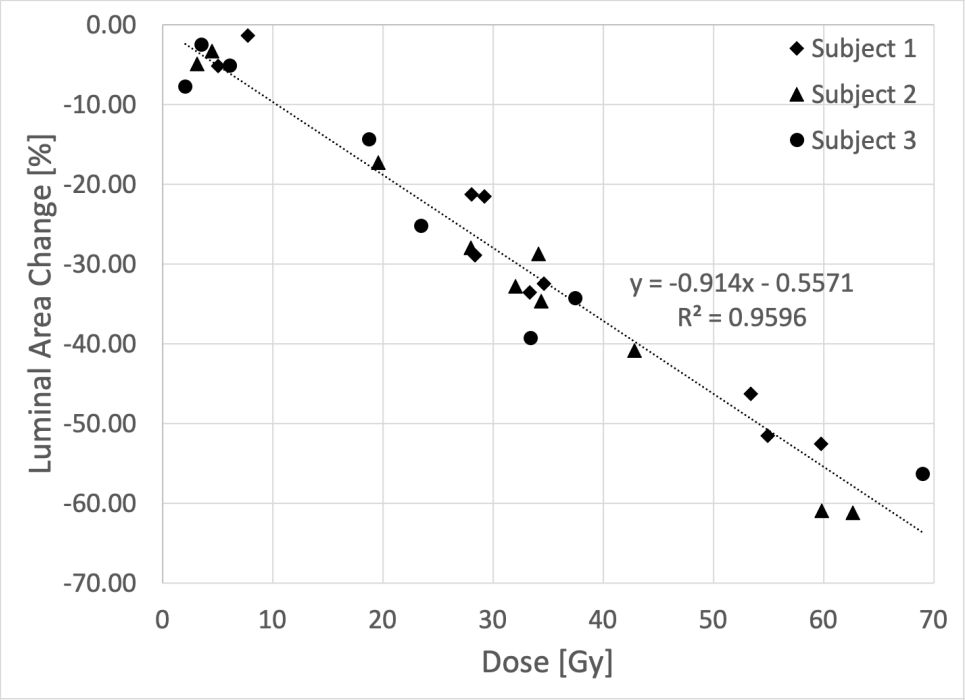


Figure 3.4: Ai percent change as a function of airway dose for airways in the inferior left (irradiated) lung across three WMS™ subjects.

Figure 3.5 shows the Jacobian ratios (post-RT/pre-RT) for three locations within the swine lungs. All regions received mean doses near 1 Gy or less, however the ILL region was downstream and supplied by airways from the highly irradiated region of the PTV.

The SLL and RL supplying airways were unirradiated (less than 1 Gy max point dose). A decrease in both the mean Jacobian ratio across all subjects and the Jacobian ratio of each individual subject of the ILL (μ=0.883) compared to the SLL (μ=0.932) and the RL (μ=0.955) was observed, however this difference was not found to be significant. This change in the ventilation of the inferior left lung for all three subjects suggests that, while on average it received less than 1 Gy, the decrease in Jacobian may be due to irradiation of the region's supplying airways. Furthermore, as shown by Patton et al. [79], a model was produced to predict Jacobian changes following irradiation of the lung. This model predicts that in order to produce a Jacobian ratio of 0.9 that is seen in the inferior left lung of the swine subjects, the regions would have had to receive at least 30 Gy, regardless of pre-RT function.

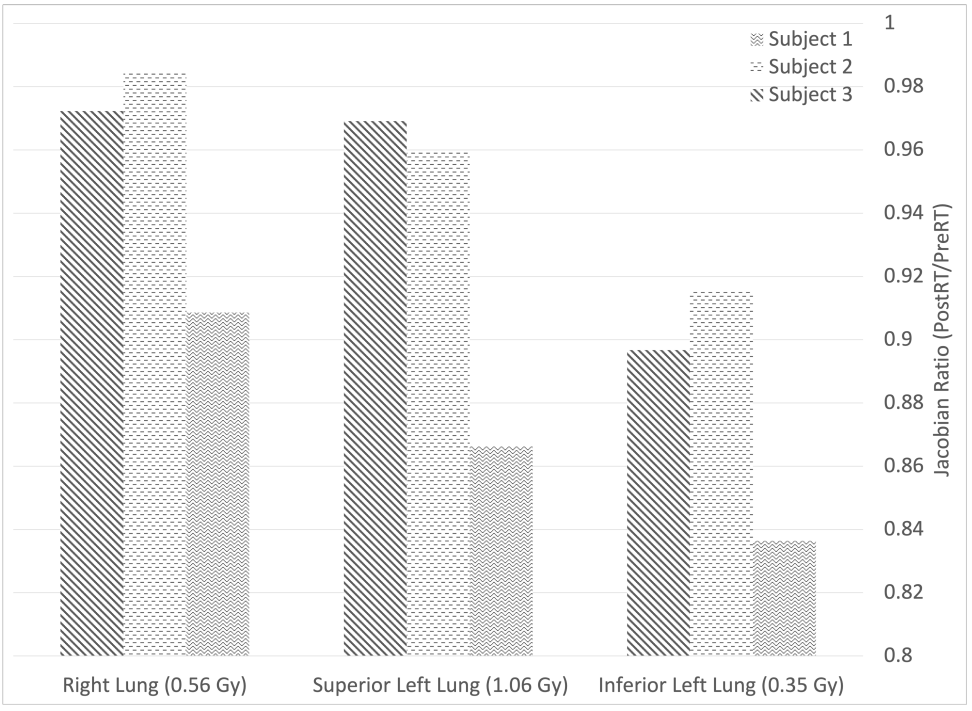


Figure 3.5: Evidence of indirect damage due to irradiation of airways. All regions received mean doses near 1 Gy or less, however the ILL region was downstream and supplied by airways from the highly irradiated region of the PTV. The SLL and RL supplying airways were unirradiated (less than 5 Gy max point dose). The average dose to the three regions across the three subjects is indicated in parentheses next to the corresponding region.

Figure 3.6 shows that post-RT Ai decreased by approximately 27% compared to pre-RT and that this decrease was constant across pre-RT Ai from 15-80 mm^2. These results analyzed airways that received a dose between 25 and 35 Gy since this dose bin contained the largest number of samples. A linear regression line was fit to this data and forced through the origin resulting in a slope of 0.729 (Pearson R = 0.998). The linear regression line was forced through the origin because this work considered only stenosis and not total collapse of the airways.

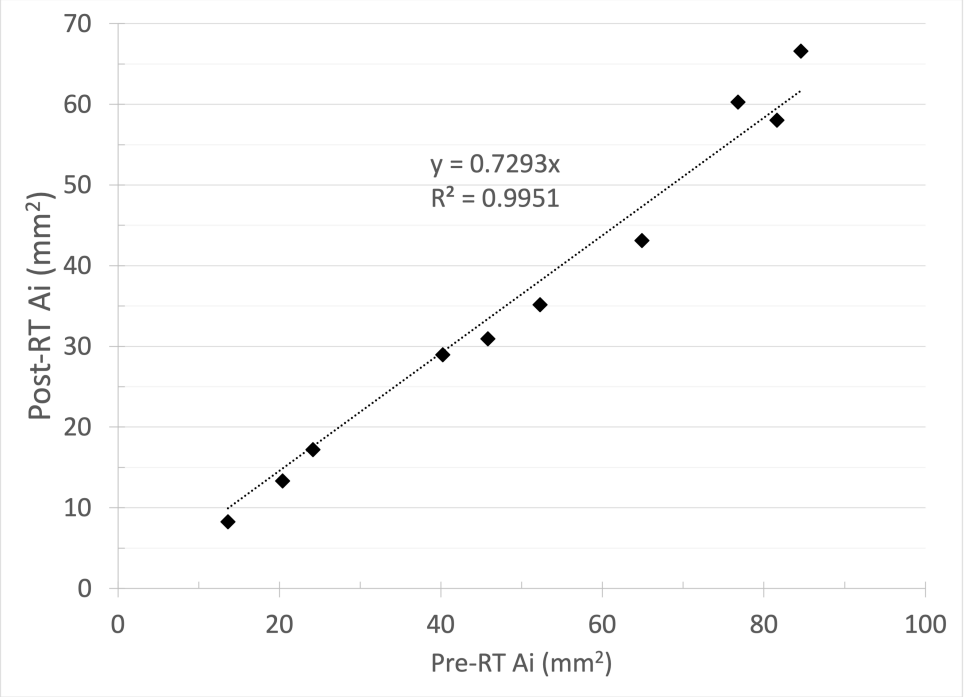


Figure 3.6: Post-RT Ai vs pre-RT Ai for a 30±5 Gy dose bin with linear regression fit.

Figure 3.7 show that there was a significant ($p<0.001$) negative correlation (Pearson R = -0.66) between the percent change in $\sqrt{WA}$ and airway dose. This correlation may be worse due to the limitation of the reconstruction resolution (0.6mm isotropic), which is on the order of the change in $\sqrt{WA}$ for smaller airways.

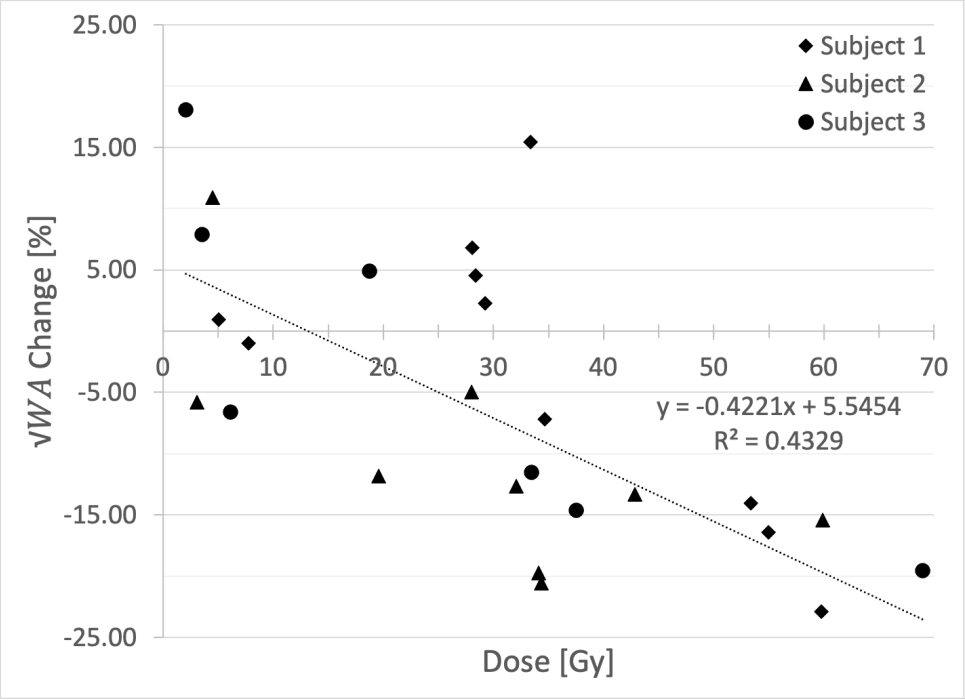
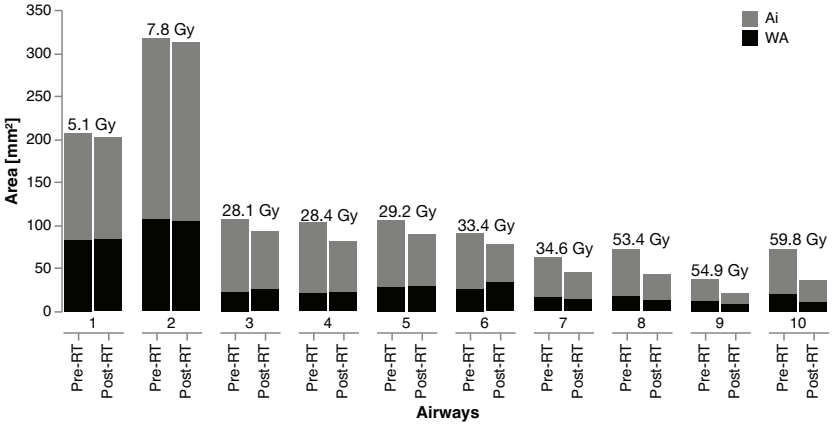
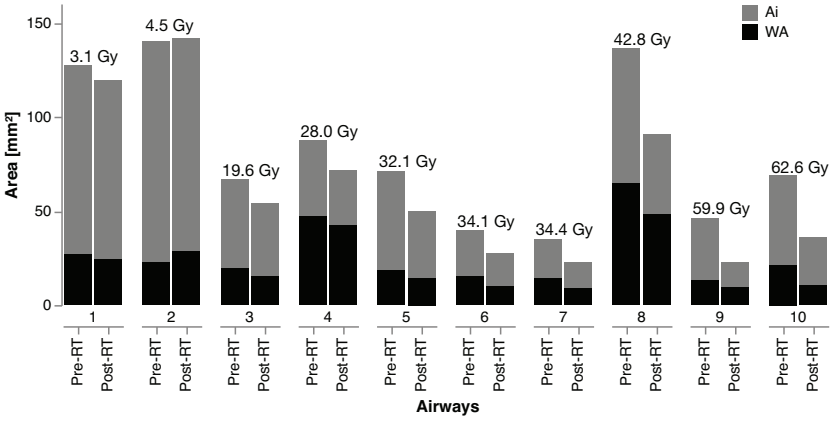


Figure 3.7: Scatter plot of $\sqrt{WA}$ change vs airway dose. Note that the correlation with airway dose is worse compared to the luminal area in Figure 3.4. The airway wall thickness is approaching the voxel size and resolution of the scan, increasing measurement noise when computing $\sqrt{WA}$.

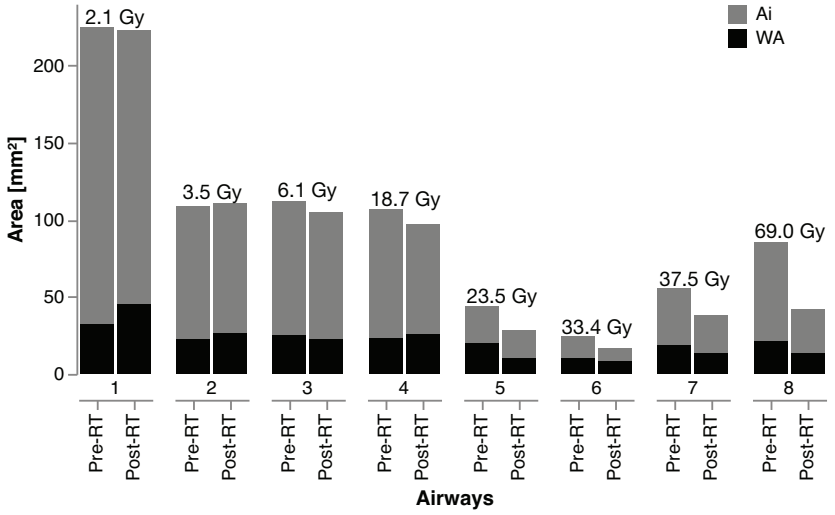
Figure 3.8 shows the WA, Ai, and $D_{2\%}$ for each airway and subject pre- and post-RT. Each figure is ordered by increasing $D_{2\%}$. Although the absolute TA decreased for almost every irradiated airway, the WA% increased for a majority of the airways since Ai (gray bars) decreased more relative to the decrease in WA (black bars).



(a) Subject 1



(b) Subject 2



(c) Subject 3

Figure 3.8: WA and Ai for each airway and subject, ordered by increasing $D_{2\%}$. $D_{2\%}$ is the dose to the highest 2% of the airway.

3.2.3 Discussion

This study demonstrated a strong linear relationship between radiation dose and change in Ai. In particular, that for every 10 Gy increase in dose to a particular airway, there was an additional 9.1% reduction in Ai of that airway. Additionally, no threshold dose level was identified for doses as low as 2.5 Gy for this effect to appear since there was a constant decrease in Ai with increasing airway dose from 2.5 to 70 Gy. Freitag et al. [137] established grading levels of bronchial stenosis based on changes in Ai, defining a decrease of greater than 25% as Grade 2 bronchial stenosis. Figure 3.4 shows that airways that were considered Grade 2 or worse post-RT received a $D_{2\%}$ of approximately 25 Gy or greater. Compared to the findings of Manyam et al. [32], these results suggest a maximum bronchial dose much lower than the suggested $D_{0.03cc} \leq 50$ Gy may be needed to minimize the occurrence of non-pneumonitis toxicities. The strong negative correlation between Ai change and radiation dose demonstrates structural changes occurring in airways beyond the PBT, and may be of clinical relevance as it pertains to pulmonary toxicities and function. Additionally, these changes are occurring at doses that are nearly half of the current dose constraint recommendations for bronchial structures.

The results shown in Figure 3.5 supports the hypothesis that regions of the lung supplied by irradiated airways, even though the region itself was relatively unirradiated, experiences a greater decline in function than unirradiated regions supplied by unirradiated airways. While this decrease was not significant, most likely due to a small sample size of n=3, there appears to be a mechanism of indirect damage that is causing the ILL to decrease in function. This decrease in function is seen despite that region receiving less than a mean dose of 1 Gy or a max dose of 5 Gy, a dose level at which little to no damage is expected. Patton et al. [79] modeled post-RT ventilation change as a function of radiation dose and found that the model did not perform well in regions downstream of irradiated regions. This model also predicts that to see a Jacobian ratio of 0.9, the region

would have to receive at least 30 Gy, regardless of pre-RT function. The hypothesis of downstream effects is further supported by the results presented in this work that the RL and SLL, whose airways were unirradiated (< 5 Gy), had higher Jacobian ratios than the ILL. It should be noted that Subject 1 appears to have a lower Jacobian ratio overall compared to the other two subjects and may be skewing the average. However, when excluding Subject 1 there is a decrease in the Jacobian ratio across the three regions evaluated. Patton et al. [79] considered a Jacobian ratio less than 0.94 as damaged lung, which is seen across all three subjects in the ILL.

When examining the 30 Gy dose bin (25-35 Gy), no dependence on Ai was found as the decrease in size from pre- to post-RT was nearly constant. However, the sample size of airways that met the criteria of this dose bin was small (n=9). Further experiments with a larger sample size, of both subjects and airways of varying doses, are needed to validate this result across a large dose range through a multi-variate analysis. This result is imperative to lung RT planning since airway dose to structures beyond the PBT are not typically considered, nor is the dose to the PBT (as long as it is within specified limits). As previously suggested, recommended dose constraints may not be accurate for PBT structures or applicable for more distal airway generations. These results suggest that there may be no threshold airway size for radiation-induced bronchial stenosis, which may potentially lead to pulmonary functional decline.

Figures 3.7 and 3.8 show that $\sqrt{WA}$ significantly decreases with airway dose. This suggests that while there is an overall shrinkage of the airway with dose, Ai decreases more in proportion to the thickness of airway wall. In other words, while the airway reduced in size, the relative thickness of the airway wall increased, resulting in increased stenosis of the airway. It has been previously shown that airway-wall thickening tends to encroach on the inner lumen through a remodeling and inflammatory process, rather than expand into surrounding lung parenchyma Hasegawa et al. [138]. Nakano et al. [132]

found that WA% was greater and Ai was smaller as the forced expiratory volume in one second (% predicted) (FEV1%P) decreased in a cohort of patients with COPD. Hasegawa et al. [138] performed a similar study, but specifically focused on third through sixth generation airways in COPD patients. Their study found WA% and Ai to be significant predictors of FEV1%P, and that correlation strength increased in further generations (i.e., as airways became smaller). Further studies have shown that more distal airways are better predictors of lung function compared to those airways located near or within the PBT [133, 134]. Kelsey et al. [17] compared the flow of airways to laminar flow through Poiseuille's law, which states that the resistance to airflow is inversely proportional to the fourth power of the radius. The hypothesis that radiation-induced bronchial stenosis is an important factor in pulmonary function and functional avoidance RT is supported by the evidence presented in the strong dose-response relationship with stenosis in Figure 3.4, coupled with the decrease in function of downstream regions in Figure 3.5.

This study was limited due to the small sample size and retrospective nature of the analysis. The radiation treatment that was delivered was not designed to study indirect effects of the radiation dose on the airways and subsequent ventilation changes downstream. As all the swine were irradiated in the inferior left lung, there was not much volume of lung downstream from the irradiated airways to study using 4DCT derived ventilation maps.

A second limitation of this work was the small (n=3) sample size of subjects. While over 50 airways were segmented as part of this analysis, it would be advantageous to perform the aforementioned improvements in a larger sample. One consideration when applying this result to human subjects, is the difference in the anatomical structure of the airways in the miniature swine compared to that of humans. Human airways exhibit a bipodial branching pattern where each airway has a bifurcation point and splits into two airways. In contrast to human airways, swine airways exhibit a monopodial branching pattern where the left and right main stem bronchi extend far inferiorly, with further generations

branching off asymmetrically. This difference in airway structure between swine and humans may result in differences in the effect of airway damage on pulmonary function. This further emphasizes the need to investigate radiation-induced airways changes in a human cohort, along with pulmonary function analysis.

3.3 A Risk Model for Radiation-Induced Bronchial Stenosis and Indirect Ventilation Damage

3.3.1 Methods

3.3.1.1 Patient Characteristics

The images used in this study were collected from 10 lung cancer patients that participated in a University of Wisconsin-Madison IRB-approved prospective study (NCT02843568) who underwent RT for non-small cell lung cancer (NSCLC). Exclusion criteria for this study included prior or future planned surgery for the existing lung cancer, prior thoracic radiotherapy, severe COPD, oxygen dependence, known underlying collagen vascular disease, under 18 years of age, and a Karnofsky score $< 60\%$. Subjects received a standard fractionation course of 2 Gy x 30 fractions. The patient information for this study cohort is summarized in Table 3.1.

Table 3.1: Study cohort characteristics

Number of Patients	10
Mean age (range)	71 y (64-89)
Gender (male/female)	6/4
Tumor location	
Right Upper	3
Right Middle	2
Right Lower	1
Left Upper	2
Left Lower	1
Mediastinum	1
Mean PTV Volume (range)	468 cc (70 - 1935)

3.3.1.2 4DCT Acquisition

The 4DCT scans used in this work were acquired pre-RT and 12 months following treatment. The post-RT timepoint of 12 months was chosen for this study as the median onset of radiation-induced bronchial stenosis has been reported to be 6-10 months [16, 19, 31, 32, 139, 140]. All scans were acquired on a Siemens SOMATOM Definition Edge 128 slice CT scanner (Siemens Healthineers AG, Erlangen, Germany) with a 0.6 mm slice thickness, 0.5 second tube rotation time, 0.09 pitch, 120 kVp, and 100 mAs/rotation. The Varian Real-Time Position Management (RPM) (Varian Medical Systems Inc., Palo Alto, CA) system monitored respiration-induced movement and captured respiratory traces used for the 4DCT reconstructions. Melodic breathing instructions were played at a rate of 15 breaths per minute during each image acquisition which has been demonstrated to increase the repeatability between scans and reduce the occurrence of

artifacts [141]. The phase of the scans used for both the 4DCT-derived ventilation map and the airway segmentation was determined by selecting the phase which minimized the difference in tidal volumes between the two timepoints. Tissue expansion using images with tidal volumes within 100cc of each other has been shown to increase the repeatability of the ventilation measurement [142].

3.3.1.3 4DCT-derived Ventilation

This work utilized an image registration algorithm with multi-resolution cubic B-spline parameterization and used the sum of squared tissue volume difference (SSTVD) as the similarity metric as described in Section 2.4.2. The Jacobian-based method described by Shao et al. [73] and Section 2.4.3 was used to generate ventilation maps for each subject at both timepoints (pre- and post-RT). Ventilation change maps (J_{ratio}) were calculated by taking the voxel-wise ratio of the post-RT and pre-RT LER-N ventilation maps.

3.3.1.4 Airway Segmentation & Analysis

Airway segmentation was performed using a commercial virtual bronchoscopy software (VIDA Diagnostics Inc., Coralville, IA). Airway measurements and segmentations up to the 11th airway generation on average were generated for the entire lung. To generate the airway segmentations, an initial segmentation was performed automatically by the software. The automatic segmentations were then visually inspected, and corrected if required. Airway measurements that were collected included luminal diameter and cross-sectional area (Ai), wall thickness (WT), and wall area percentage (WA%). Measurements were made perpendicular to the long axis of the airway and averaged along the middle third of the segment. All airway measurements were exported from VIDA into XML

datasets where they could be pulled from. The reproducibility of airway measurements produced using this software has been previously reported by Smith et al. [143].

A nearest neighbor search was performed to connect each terminal airway in the pre-RT airway tree to its nearest voxels in the LER-N ventilation map. It was assumed that each terminal airway was responsible for the ventilation to those nearest voxels. Only voxels that received less than 5 Gy were included in the analysis to ensure that any changes in ventilation were due to changes in the airways and not direct irradiation of the lung parenchyma. The post-RT airway tree was warped to the pre-RT reference frame using deformable image registration from Section 2.4.2 in order to determine the correspondence between the post- and pre-RT airway tree labels. Label pairs were used to extract airway measurements from the XML datasets as well as map to the terminal J_{ratio} regions. Airways that were within the planning target volume (PTV) were excluded from analysis due to these airways being surrounded by disease/tumor. Airways surrounded by disease may be compressed or infiltrated by the tumor making it difficult to measure the true changes in the airways post-RT due to radiation treatment.

Cumulative airway resistance for a terminal airway pathway was determined by summing all airways in series within the path. Resistance was calculated using Poiseuille's Law:

$$R = \frac{8\eta l}{\pi r^4} \tag{3.1}$$

where η is the viscosity, l is the length of the airway, and r is the radius of the airway. The viscosity was assumed to be constant at $1.7\text{x}10^{-5}$ m^2/s. Airway resistance change was defined as the ratio between the post- and pre-RT cumulative airway resistances for each terminal airway path. Lastly, the maximum point dose for each airway was defined as $D_{0.03cc}$ as this was found by Manyam et al. [32] to be a significant predictor of non-pneumonitis toxicities in SBRT patients.

3.3.1.5 Statistical Analysis

RStudio (version 2022.07.2; Posit Software, Boston, MA) was used for calculation of all statistics. Pearson correlation coefficients were first determined for correlations between changes in airway metrics, including wall area fraction ratio, luminal area ratio, and wall area ratio, and maximum dose received. The ratio was defined as the post-RT divided by the pre-RT airway metric. Wall fraction was defined as the area of the airway wall divided by the entire airway area. Logistic regression was used to calculate the normal tissue complication probability (NTCP). The toxicities of interest in this work included bronchial stenosis, which was defined as a reduction in the cross-sectional area of a bronchial segment, and regional ventilation decline, which was defined as a regional average $J_{ratio} < 0.94$. Univariate logistic regression was first used to determine the clinical and dosimetric predictors for these two endpoints. Any variables from the univariate analysis that had a p-value < 0.05 were included in the multivariate logistic regression for that endpoint. Predictors that were considered significant from univariate analysis were compared to each other to test for correlations. For variables with a strong correlation between each other the predictor with the lower p-value was selected to be included in the multivariate analysis.

Receiver operating characteristic (ROC) analysis was used to analyze the power of the selected predictors. Bootstrapping with replacement was performed with 2000 random samplings to calculate the 95% confidence intervals (CI) for the area under the ROC curves (AUC) and the coefficients of the predictors in the logistic regression models. Data was randomly split into training and testing datasets with an 80/20 train/test split. Lastly, Welch t-tests were used to compare differences in the means of airway resistances and Jacobian ratios of the fed regions.

3.3.2 Results

3.3.2.1 Bronchial Stenosis Dose-Response

The total number of pre/post airway segment pairs analyzed for dose-response with bronchial stenosis was 679. There were 79 airways excluded from analysis as they were within the PTV. Pearson correlation coefficients for maximum dose and the changes in airway metrics were calculated for wall area fraction (WAF), wall area (WA), and luminal area (Ai) ratios. All three metrics were significantly correlated with maximum dose (WAF: r=0.13, P<0.001; WA: r=-0.10, P=0.02; Ai: r=-0.15, P<0.001). Additionally, bronchial segments that received less than 10 Gy had mean and median Ai ratios of 1.21 and 1.14, respectively. Segments that received greater than 10 Gy had mean and median Ai ratios of 1.07 and 1.03, respectively. The difference between the means of the two dose groups was statistically significant (p<0.001).

Airways were considered stenosed when the ratio of post- to pre-RT Ai was less than unity. Of the 679 airway segments analyzed, 261 (38.4%) were stenosed following radiation therapy. Univariate logistic regression was performed in order to determine significant predictors of bronchial stenosis for the purposes of an NTCP model. All analyzed airway metrics and dosimetric parameters were significantly associated with bronchial stenosis, and the pre-RT wall thickness had the highest OR (OR 3.98; 95% CI 1.67-5.63; $P <$ 0.001). The maximum and average dose delivered to a bronchial segment had the same OR (OR 1.02), however the maximum dose had a lower p-value ($P = 0.004$). Table 3.2 summarizes the results of the univariate logistic regression, including the odds ratio (OR) and AUC for each variable.

Table 3.2: Univariate logistic regression analyses of dosimetric and airway parameters for airway stenosis. Confidence intervals (CI) for the odds ratio (OR) and area under the curve (AUC) were determined by bootstrapping with 2000 random samplings.

Feature	OR (95% CI)	P-value	AUC (95% CI)
Max Dose	1.02 (1.01-1.03)	0.004	0.57 (0.46-0.67)
Avg. Dose	1.02 (1.00-1.04)	0.01	0.56 (0.46-0.66)
Pre-RT D_i	1.16 (1.06-1.27)	< 0.001	0.62 (0.53-0.71)
Pre-RT A_i	1.01 (1.00-1.02)	0.01	0.62 (0.52-0.71)
Pre-RT A_o	1.01 (1.00-1.01)	0.004	0.61 (0.52-0.70)
Pre-RT WT	3.09 (1.67-5.63)	< 0.001	0.56 (0.46-0.65)
Pre-RT WA	1.02 (1.00-1.03)	0.002	0.60 (0.50-0.69)

Abbreviations: Di = inner diameter; A_i = luminal area; A_o = outer area; WT = wall thickness; WA = wall area.

From the results of the univariate logistic regression analysis, the maximum dose and pre-RT WT were selected for multivariate logistic regression analysis as they each had stronger p-values than the other variables they were found to be correlated with. Maximum dose delivered to a bronchial segment (OR 1.01; 95% CI 1.00-1.03; $P = 0.03$) and pre-RT WT (OR 2.61; 95% CI 1.42-4.73; $P = 0.002$) were found to be significantly correlated with bronchial stenosis using multivariate analysis. The results of the multivariate logistic regression analysis are summarized in Table 3.3. Figure 3.9 shows the ROC curve from the multivariate logistic regression model using the testing dataset, and has an AUC of 0.61.

Table 3.3: Multivariate logistic regression analyses using max dose and pre-RT wall thickness as the predictors for airway stenosis.

Feature	OR (95% CI)	P-value	AUC (95% CI)
Max Dose	1.01 (1.00-1.03)	0.03	0.61 (0.51-0.71)
Pre-RT WT	2.61 (1.42-4.73)	0.002	

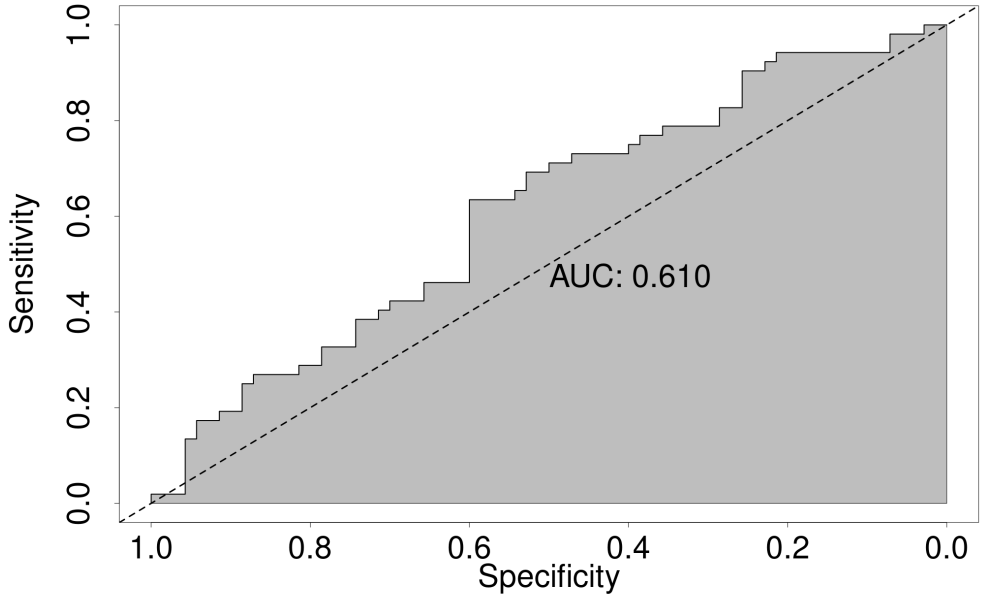


Figure 3.9: Receiver operating characteristic (ROC) curve from the multivariate logistic regression for airway stenosis. Overlaid on the plot is the area under the ROC curve (AUC) of 0.61.

3.3.2.2 Airway Resistance and Ventilation Change

A total of 328 terminal airway paths were analyzed to investigate the relationship between airway resistance change and ventilation change following radiation therapy. Of the 328 airways paths, 129 fed regions that declined in ventilation post-RT ($J_{ratio} < 0.94$) and 199 fed regions that did not decline in ventilation post-RT. Of the 129 airway paths that

fed regions of ventilation decline, 74 (57%) increased in resistance post-RT compared to the pre-RT airway path resistance. Of the 199 airway paths that did not feed regions of ventilation decline, 72 (36%) increased in resistance post-RT compared to the pre-RT airway path resistance. There was a significantly ($P < 0.001$) greater proportion of airways that increased in resistance post-RT that fed regions that experienced ventilation decline compared to those that fed regions that experienced no change or increased in ventilation. Furthermore, as shown in Figure 3.10, the average resistance ratio (post-RT/pre-RT cumulative airway path resistances) was significantly ($P = 0.004$) greater for the airways that fed regions of ventilation decline (average R-Ratio = 1.72) compared to the airways that fed regions of ventilation increase (average R-Ratio = 1.20). Lastly, as seen in Figure 3.11, the average J_{ratio} was significantly ($P < 0.001$) greater in regions that were fed by airways that decreased in resistance post-RT ($J_{ratio} = 0.98$) compared to regions that were fed by airways that increased in resistance post-RT ($J_{ratio} = 0.95$).

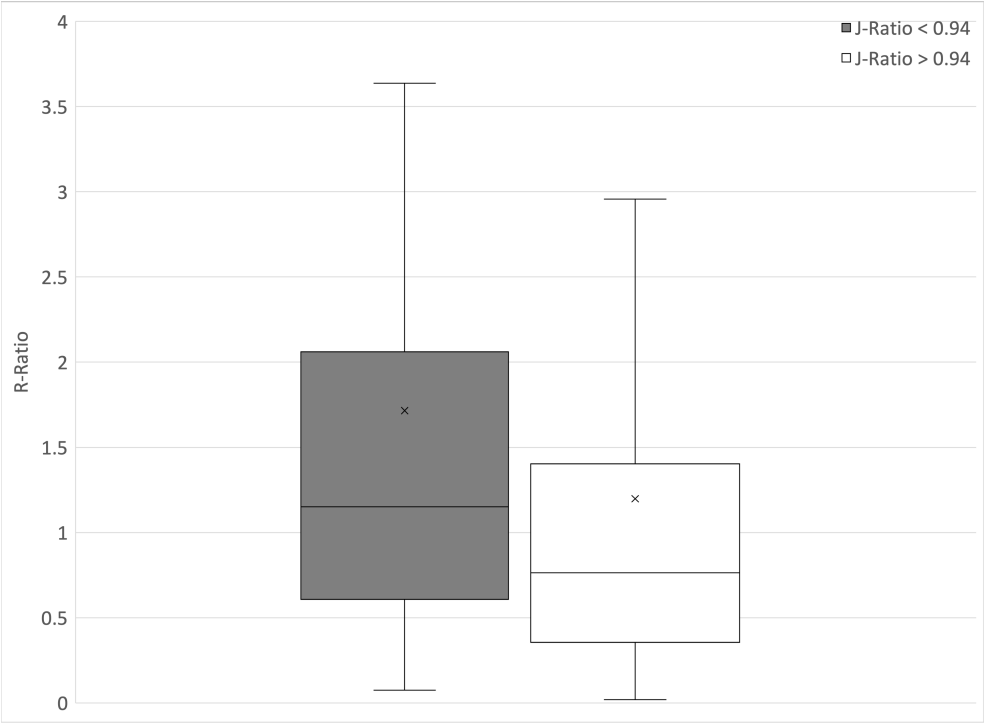


Figure 3.10: Boxplots showing the resistance ratios (post-RT/pre-RT cumulative airway resistances) of airways that fed regions that had a Jacobian ratio greater or less than 0.94. There was a significant ($P = 0.004$) difference between the average resistance ratio of the two groups.

Figure 3.11: Boxplots showing the Jacobian ratios of regions fed by airways that either decreased or increased in cumulative resistance post-RT. There was a significant ($P < 0.001$) difference between the average Jacobian ratios of the two groups.

Univariate analysis was performed and it was found that the resistance ratio (OR 1.20; 95% CI 1.01-1.43; $P = 0.03$) was significantly correlated with regional ventilation decline. Results of the univariate logistic regression analysis are summarized in Table 3.4, and the ROC curve is shown in Figure 3.12 with an AUC of 0.69.

Table 3.4: Univariate logistic regression analysis using the resistance ratio (post-RT/pre-RT resistances) as the predictor for ventilation decline, where ventilation decline was defined as a Jacobian ratio < 0.94.

Feature	OR (95% CI)	P-value	AUC (95% CI)
R-Ratio	1.20 (1.01-1.43)	0.03	0.69 (0.54-0.82)

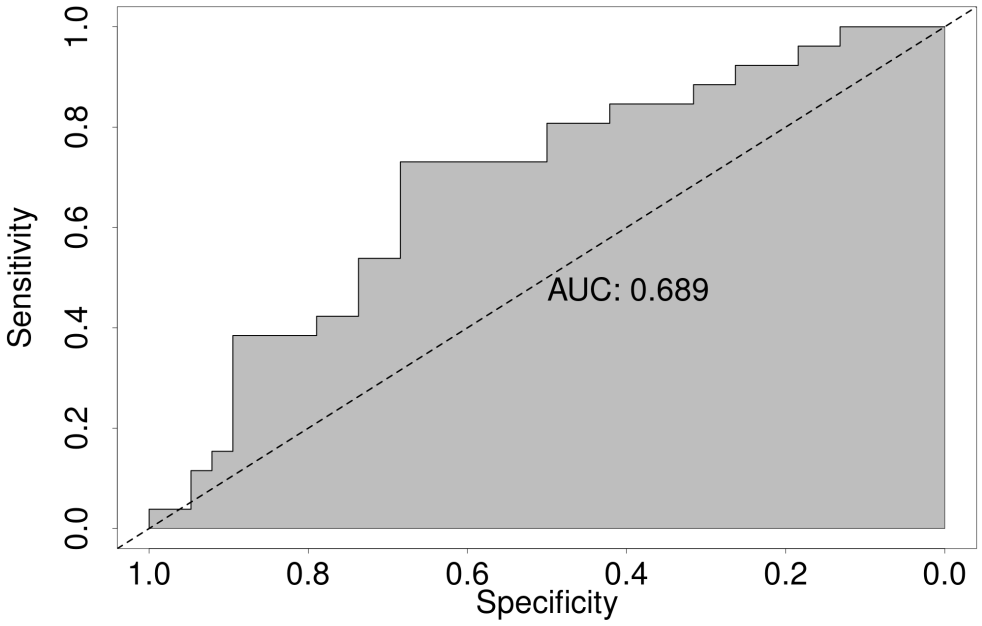


Figure 3.12: Receiver operating characteristic (ROC) curve from the univariate logistic regression for ventilation decline. Overlaid on the plot is the area under the ROC curve (AUC) of 0.69.

Figure 3.13 is an example that shows one subject with the delivered dose distribution (1A), pre-RT airway segmentation (1B), and the Jacobian ratio map thresholded to only show voxels that received less than 5 Gy and had a Jacobian Ratio less than 0.94. (1C).

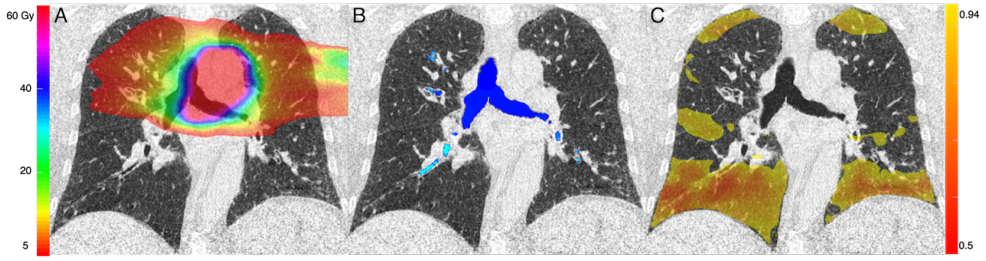


Figure 3.13: One subject showing a coronal view of the (A) delivered dose distribution, (B) segmented airways, and (C) Jacobian Ratio thresholded to only show voxels that received less than 5 Gy and had a Jacobian Ratio less than 0.94.

3.3.3 Discussion

Other studies have reported airway toxicity dose constraints for SBRT fractionated patients [29, 32, 85, 140], however this is the first study to analyze the dose-response relationship between bronchial stenosis and radiation dose for standard fractionation patients. It is also the first work to quantify the relationship between airway resistance changes and 4DCT-based ventilation changes from baseline (pre-RT) and 12-months post-RT. The results shown demonstrate that airway toxicity, specifically bronchial stenosis, is prevalent beyond the proximal bronchial tree, and that this increases airway resistance and decreases ventilation function in the regions supplied by the affected airways. It was found through multivariate analysis that the maximum dose delivered to a bronchial segment and the pre-RT wall thickness of that segment are significant predictors of bronchial stenosis following radiation therapy. Additionally, it was found that an increase in cumulative airway resistance was significantly associated with a decrease in ventilation function in the region of the lungs supplied by that airway branch.

There were significant correlations between maximum dose and the post- to pre-RT ratios of wall area fraction, wall area, and luminal area. Both the wall area ratio and luminal area ratio were weakly, negatively correlated with maximum dose, and the wall area fraction was weakly, positively correlated with maximum dose. This means that the wall area, while it decreased overall, made up a larger fraction of the total airway area post-RT compared to pre-RT. Coxson et al. [144] found that the forced expiratory volume in one second (FEV1) decreased as wall area fraction increased. The results presented in this work suggest that increased radiation dose to a bronchial structure increases the wall area fraction, which can lead to decreased global lung function. These results are particularly important to pulmonary functional avoidance radiation therapy as most clinical trials in this space largely ignore the effects of irradiating the airways, especially those beyond the proximal bronchial tree. Based on the results shown, it is imperative that bronchial dose

is considered and minimized when developing a treatment plan to reduce the amount of indirect ventilation damage that occurs.

From the example subject shown in Figure 3.13, it can be seen that lower right lobe experiences a decline in ventilation post-RT while receiving radiation dose less than 5 Gy. As seen in the work by Patton et al. [79], doses below 5 Gy are not predicted to cause direct ventilation damage. The cumulative airway resistances of the airway branches feeding this region were found to have increased in resistance post-RT compared to pre-RT, and it is hypothesized that the increased resistance is responsible for the decrease in ventilation seen. This hypothesis is further supported by the significant association between regional ventilation decline and cumulative airway resistance from the univariate logistic regression analysis. The univariate analysis showed that the resistance ratio had an OR of 1.20 (95% CI 1.01-1.43) and an AUC of 0.69. By combining this result with the results of the multivariate analysis for bronchial stenosis, it is a reasonable conclusion that radiation dose to a bronchial segment can cause stenosis, leading to increased airway resistance and a decline in ventilation in the region fed by those airways.

While this work was successful in showing a relationship between bronchial stenosis, radiation dose, and ventilation change, there were a few limitations. First, this study was performed on only 10 subjects. By performing the methodology on a larger cohort of subjects there may be additional information that can be determined from the dose-response of the airway metrics. Another limiting factor was that all subjects analyzed were treated with a standard fractionation of 2 Gy x 30 fractions. This type of fractionation schedule was selected for this study as it maximized the number of airways receiving radiation dose and airways that fed regions distal to the main volume of irradiation. SBRT fractionation treatments (10 Gy x 5 fractions) are typically prescribed for tumors not near the mediastinum due to previous reports of central airway toxicities as well as concerns due to high biological radiation dose to the heart. Due to the regions of irradiation

for SBRT treatments, there will be less airways and lung parenchyma "downstream" from the PTV, however the response due to irradiation and the effect on ventilation may differ from the results provided here. Additionally, due to only analyzing standard fractionation subjects, there is a significant correlation between airway caliber and dose received. All the PTVs analyzed were centrally located meaning that the larger airways were generally closer to the highest regions of the dose distribution. The results of the multivariate analysis showed that the pre-RT wall thickness was a significant predictor of bronchial stenosis, however there is a significant ($p<0.001$) correlation with maximum dose. It may be necessary to only use the results of the univariate analysis as a way to predict bronchial stenosis until further analysis can be performed on a wider variety of dose distributions. A third limitation is the resolution of the CT images as well as the resolution of the airway segmentation software and its ability to segment small airways. The image resolution used for this work was 0.6mm isotropic voxels, however, the smallest airways segmented experienced changes in metrics, such as wall thickness and inner diameter, that were on the order of the imaging resolution. Additionally, the software is only able to automatically segment airways down to a certain size before it labels non-airway tissues as airway. At this point, manual intervention is required which may lead to missed or incorrectly segmented airways. Lastly, the use of Poiseuille's Law assumes laminar flow, which occurs throughout most of the airway tree. However, this is a simplification as both laminar and turbulent flow occur within the airways and the calculation of resistance where turbulent flow exists depends on the density of the gas instead of the viscosity.

3.4 Chapter Summary

While there have been previous studies that have shown a dose-response relationship with bronchial stenosis, no work has provided a quantitative relationship between airway dose and stenosis, nor considered these effects in conjunction with ventilation change. The results of this work support the hypothesis that airway dose beyond the PBT should be considered during functional avoidance treatment planning to avoid radiation-induced bronchial stenosis and potentially pulmonary function decline. Additionally, significant correlations were found between the maximum dose delivered to a bronchial segment, the pre-RT wall thickness, and the occurrence of bronchial stenosis. Furthermore, significant associations were found between increased cumulative airway resistance and regional ventilation decline, supporting the hypothesis that radiation-induced bronchial stenosis has an indirect effect on ventilation function. The results from this work should be considered for future clinical trials involving functional lung sparing as dose to bronchial structures may impact regional ventilation function in addition to the direct effects already seen from irradiation of lung parenchyma.

Chapter 4

Modeling normal lung tissue response to radiation dose [1]

4.1 Introduction

4DCT-derived ventilation maps are assumed to be a surrogate for lung ventilation and can be calculated using various techniques. Kipritidis et al. [146] estimated lung function directly using Hounsfield unit values of the time-averaged 4DCT. Reinhardt et al. [147] used the Jacobian determinant of the transformation computed from image registration and is the method used in this work. The Jacobian determinant method works on the assumption that the expansion of a voxel is caused by the addition of air from ventilation.

Previous works have calculated the local lung expansion ratio (LER) from the end inhale phase (100IN) and the end exhale phase (0EX), which is defined as 2-phase LER (LER-2) [72]. However, this can lead to errors in the ventilation map due to out-of-phase ventilation. Out-of-phase ventilation can be defined as local lung volume change that is

[1] Portions of this work have been published [75, 145]

out-of-phase with respect to global lung contraction and expansion [72]. To correct for out-of-phase ventilation, the LER can be calculated from the ratio of the maximum and minimum local lung volume over the entire breathing cycle, as proposed by Shao et al. [72]. This calculation of LER is defined as N-phase LER (LER-N).

Vinogradskiy et al. [86] suggested that incorporating ventilation-based functional imaging could help improve prediction of radiation pneumonitis. Through logistic regression they found that dose-function metrics had on average greater AUC values than their traditional dose-volume equivalents [86]. Faught et al. [13] found that including a variety of functional information resulted in a higher predictive power than using strictly total lung dose. Additionally, by using CT ventilation-based functional lung metrics to guide treatment planning, they observed predicted reductions in grades 2+ and 3+ pneumonitis of 7.1% and 4.7%, respectively [13]. A previous regression-based model has been developed by Patton [123] to predict changes in post-RT ventilation based on pre-RT ventilation measurements and the dose delivered. One of the issues with this approach was that ventilation measurements were calculated using only the end-inhale and end-exhale phases of the 4DCT, which does not account for out-of-phase ventilation.

More recently, machine learning-based models have grown in popularity due to their higher predictive accuracy. Katsuta et al. [95] trained a kernel-based support vector machine (SVM) to predict radiation pneumonitis and found that their predictive power, via the AUC, significantly improved when Jacobian-based dose-function features were added. Huang et al. [93] used a fuzzy clustering neural network to predict radiation-induced pneumonitis based on ventilation imaging-based dose-function metrics. However, to date there has been no work attempting to predict regional ventilation change following radiation therapy.

The first section of work in this chapter develops and quantifies the accuracy of a dose-response model using ventilation measurements calculated from all phases of the 4DCT

(LER-N). Through using a "simpler" modeling approach using polynomial regression, it is possible to obtain physical values associated with the predicted outcomes, compared to the more black box nature of deep learning solutions. The second section of this chapter trains, validates, and compares multiple deep learning-based models to predict regional pulmonary ventilation changes. It is hypothesized that training a deep learning model on volumetric images will result in a higher model sensitivity in predicting regions of ventilation decline.

4.2 Modeling the impact of out-of-phase ventilation on normal lung tissue response to radiation dose

4.2.1 Methods

4.2.1.1 4DCT Acquisition and Datasets

34 human subjects from a University of Wisconsin-Madison (UW) IRB-approved prospective study (NCT02843568) and an additional 8 human subjects from a University of Iowa (UI) IRB-approved trial (NCT01039649) who underwent RT were used in this study. Exclusion criteria for this study included prior or future planned surgery for the existing lung cancer, prior thoracic radiotherapy, severe COPD, oxygen dependence, known underlying collagen vascular disease, under 18 years of age, and a Karnofsky score $< 60\%$. Each subject had two 4DCT scans separated by five minutes acquired pre-RT and three months post-RT for a total of four scans. The patient information for this study cohort is summarized in Table 4.1.

Table 4.1: UW and UI cohort characteristics. Note that one subject did not have a tumor stage recorded during their study enrollment.

Number of Patients	42
Mean age (range)	68 y (30-88)
Gender (male/female)	26/16
Tumor type	
Adenocarcinoma	20
Squamous Cell Carcinoma	10
NSCLC	7
Others	5
Tumor location	
Right Upper	13
Right Middle	4
Right Lower	9
Left Upper	11
Left Lower	5
Mean PTV Volume (range)	354.2 cc (13.7-1447.1)
Mean Prescription Dose (range)	56 Gy (36-70)
Karnofsky (range)	90 (70-100)
SBRT/Standard Fractionation	19/23
Tumor Stage	
I	20
II	0
III	17
IV	4

All scans for the UI dataset were acquired on a Siemens Biograph 30-slice CT scanner

(Siemens AG, Munich, Germany) with a 2 mm slice thickness, 0.5 second tube rotation time, 0.1 pitch, 120 kVp, and 700 mAs. The scanner used an Anzai AZ-773V (Anzai Medical Co., Tokyo, Japan) with a strain gauge belt and audible timing cues were played throughout the scan to increase the repeatability of respiration (RESP@RATE, Intercure Ltd., Lod, Israel).

All scans for the UW datasets were acquired on a Siemens SOMATOM Definition Edge 128 slice CT scanner with a 0.6 mm slice thickness, 0.5 second tube rotation time, 0.09 pitch, 120 kVp, and 100 mAs/rotation. Respiration induced movement was monitored using the Varian Real-Time Position Management (RPM) (Varian Medical Systems Inc., Palo Alto, CA), which captured the respiratory traces from the patients for the 4D reconstructions. Additionally during each 4DCT acquisition, breathing instructions were played to increase the repeatability between scans and reduce the occurrence of artifacts. Breathing instructions were played at a rate of 15 breaths per minute.

4.2.1.2 Image Registration and Data Preprocessing

Each 4DCT was reconstructed into 10 discrete phases of the breathing cycle, with 20% (20IN), 40% (40IN), 60% (60IN), 80% (80IN) and 100% (100IN) inspiration phases and 80% (80EX), 60% (60EX), 40% (40EX), 20% (20EX) and 0% (0EX) expiration phases. Han et al. [148] describes more in-depth the respiration cycle as it pertains to 4DCT imaging.

Deformable image registration was used to register each of the 9 phases (20-100) to the 0EX phase. The registration algorithm uses a multi-resolution B-spline parameterization as described in Section 2.4.2 [149]. Additionally, to account for HU variation with lung density change, a sum of squared tissue volume difference (SSTVD) similarity metric is used. By using the method proposed by Reinhardt et al.[147], the LER of each voxel is

estimated by calculating the Jacobian determinant of the transformation. This method of calculating ventilation has also been shown to result in more reproducible results versus CT intensity-based methods [150].

4.2.1.3 Local Expansion Ratio From Multiple 4DCT Phases

According to Shao et al.[72][151], only calculating the LER from two phases of the breathing cycle will underestimate LER when there is out-of-phase ventilation. To account for this, the LER is calculated using multiple phases of the breathing cycle (LER-N) as describe in Section 2.4.3.

4.2.2 Model Training & Cross-Validation

The method for building the predictive model was similar to the method described by Patton [123]. The predictive model was built using polynomial regression based on the delivered dose distribution and the change in ventilation. From the pre-RT Jacobian value and the dose of a voxel, the model was designed to predict the change in ventilation of the voxel. To train the model, 27 of the 34 subjects from the UW dataset were used. For subjects who had multiple scans at a single timepoint that were deemed acceptable (i.e. no excessive artifacts), the Jacobian maps for the scans were averaged using the geometric mean. The training data consisted of the average post-RT Jacobian to pre-RT Jacobian ratio, $J_{postRT/preRT}$, and were binned based on dose and pre-RT Jacobian values. Dose bins ranged from 0 Gy to 60 Gy in 5 Gy increments and pre-RT Jacobian bins ranged from 1.0 to 1.6 in increments of 0.05. The average $J_{postRT/preRT}$ for all voxels were then calculated for each bin. 3D polynomial regression was performed using the dose distribution and pre-RT Jacobian as the independent variables and the average $J_{postRT/preRT}$ as the dependent variable. Twenty-five regression models ranging from first

to fifth order polynomials in both dose and pre-RT Jacobian domains were created for each iteration of the cross validation. This means that there were 27 models created for each of the 25 different combinations of polynomial orders. The polynomial models are referred to as polyXY, where X is the order of the polynomial for dose and Y is the order of the polynomial for pre-RT Jacobian. The fits of each model were calculated using the method of least-squares and the adjusted coefficient of determination (R^2_{adj}) was calculated for each fit. R^2_{adj} is a metric that informs how well the model fits the data, but also penalizes for overfitting when higher-order terms are used. This allows for fair comparison between the fits of models that have different ordered polynomials[152]. For each polynomial regression model a predicted post-RT ventilation map was created. Normalization was achieved by multiplying the $J_{predict}$ map by the ratio of J_{preRT} to $J_{predict}$, resulting in the predicted ventilation map having the same tidal volume as the pre-RT map.

4.2.2.1 Validation

Validation of the model was initially performed using a leave-one-out cross-validation as described by Patton [123]. Cross-validation was performed in order to determine the degree of polynomial fit for dose and J_{preRT}. The various models were assessed using the gamma pass rate with a distance-to-agreement term of 2mm and an intensity difference term of 6%, which is equal to the standard deviation of the Jacobian values between repeat scans[153]. The gamma analysis was performed by comparing J_{postRT} and $J_{predict}$ to test the accuracy of each model's predictive ability. The models were also compared using R^2_{adj}. From the results of the cross-validation, the best model of the 25 various polyXY models was determined by choosing the model that had the highest quadratic sum of the gamma pass rates and R^2_{adj}.

Additional validation was performed on the model determined from cross-validation using the remaining 7 subjects from the UW dataset and the 8 subjects from the UI dataset. The inclusion of the UI dataset provided data that was acquired on a different scanner, at a different institution, with different treatment planning parameters, and was collected almost 15 years prior to the UW dataset. Using the UI dataset tested the robustness and flexibility of the model. Gamma pass rate was used again to determine the similarity between the predicted and actual post-RT ventilation maps. A Student's t-test compared the similarity of the N-phase model and 2-phase model gamma pass rates, positive predictive value (PPV), true positive rate (TPR), true negative rate (TNR), and accuracy (ACC). PPV, TPR, TNR, and ACC are calculated from true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). A TP is when the model correctly predicts a decrease in function of 6% or greater, a FP is when the model incorrectly predicts a decrease in function of 6% or greater, a TN is when the model correctly predicts no decline in function, and a FN is when the model incorrectly predicts no decline in function. The threshold value of 6% was used due to the Jacobian ratio of repeat scans having a standard deviation near 6%. PPV, TPR, TNR, and ACC were calculated at the voxel level and then averaged for each subject. The final values given for each of the metrics is the average of all the additional validation subjects for the polynomial model that was determined from cross-validation. P-values reported were calculated on the patient level (i.e. n=15). The equations for PPV, TPR, TNR, and ACC are given in Equation 4.1, Equation 4.2, Equation 4.3, and Equation 4.4, respectively.

$$PPV = \frac{TP}{TP + FP} \tag{4.1}$$

$$TPR = \frac{TP}{TP + FN} \tag{4.2}$$

$$TNR = \frac{TN}{TN + FP} \tag{4.3}$$

$$ACC = \frac{TP + TN}{P + N} \tag{4.4}$$

P and N in the equation for ACC represent the total number of positives and negatives observed in the ground truth J_{postRT} data, respectively.

All metrics used to validate the predictive model were calculated using voxels in the lung that received doses of 20 Gy or higher. A mean lung dose of 20 Gy and percent volume of the lung receiving 20 Gy (V_{20}) have been shown to be predictors of when lung toxicities begin to occur and are used commonly in treatment planning [154, 155]. Additionally, it has been shown that regions of the lung that receive greater than 20 Gy show a greater decline in ventilation post-RT than regions that receive less than 20 Gy [123].

To investigate the N-phase model's ability to predict post-RT ventilation in regions of out-of-phase ventilation in the pre-RT ventilation map, the Jacobian ratio of post-RT to predicted Jacobian values ($J_{postRT/predict}$) were calculated for both models and compared. Ideally, this ratio should be unity which would indicate that the predicted post-RT Jacobian value is equal to the actual post-RT Jacobian value for a particular voxel. For this analysis, only voxels that were determined as out-of-phase pre-RT were analyzed and included voxels receiving all dose values.

As previously shown by Patton et al. [156], regions of the lung labeled as high function prior to RT (J_{preRT} >1.1) have a greater reduction in ventilation compared to those regions considered low function pre-RT. Voxels that received 20 Gy or more, were predicted to decline in function by more than 6%, and were identified as high function pre-RT were

analyzed using the LER-N model to determine the model's ability to predict regions of the lung that have the largest impact on post-RT ventilation.

4.2.3 Results

Figure 4.1 shows the results for the subject with the greatest amount of lung considered out-of-phase, with approximately 50% of the lung, mostly in the inferior regions, out-of-phase. Consequently, the 2-phase model provides a poor prediction of the ventilation map post-RT. By incorporating information from multiple phases of the breathing cycle, the out-of-phase ventilation was accounted for and a more accurate prediction of the ventilation map was produced. This is an extreme case for out-of-phase ventilation as it was determined for this cohort that an average of 14% of the lung is out-of-phase, but this shows the usefulness of including N-phase data in the model.

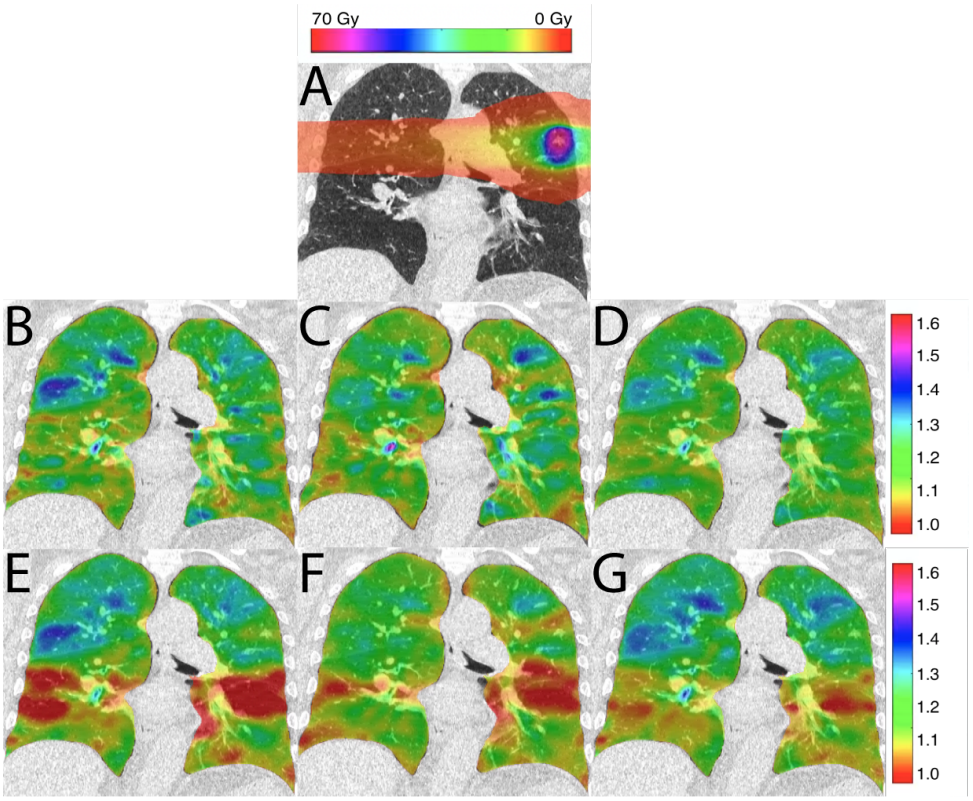


Figure 4.1: Images from one subject with severe out-of-phase ventilation. A) Dose distribution overlaid on pre-RT CT B) LER-N pre-RT Jacobian map C) LER-N actual post-RT Jacobian map D) LER-N predicted post-RT Jacobian map E) LER-2 pre-RT Jacobian map F) LER-2 actual post-RT Jacobian map G) LER-2 predicted post-RT Jacobian map. The inferior portions of the lungs in the LER-2 maps are largely out-of-phase, which is corrected by LER-N.

Figure 4.2 shows the cumulative 2D histogram of LER-N vs LER-2 Jacobian values for all voxels for all subjects used in validation. A logarithmic scale is used for better visualization. The area labeled as A lays between the line $y = x$ and $y = 1.06x$ and represents voxels considered in-phase ventilation, which constitute 86.1% of all the voxels. Region B is considered out-of-phase ventilation where LER-N and LER-2 calculated the voxel as low function (J < 1.1). Region C is out-of-phase ventilation where LER-N calculated the voxel as high function (J > 1.1) and LER-2 calculated the voxel as low function. Lastly, region D is out-of-phase ventilation where LER-N and LER-2 calculated the voxel as high

function. P(A), P(B), P(C), and P(D) are the percent of voxels in each region as noted in Figure 4.2.

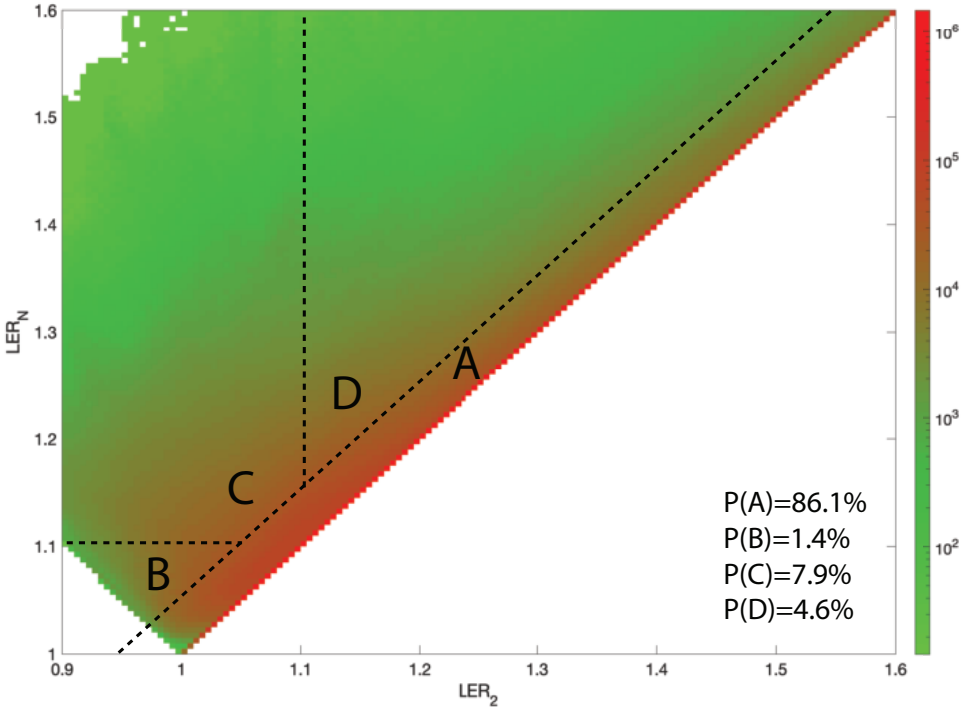


Figure 4.2: 2D cumulative histogram of LER-N vs LER-2 Jacobian values. Region A lies between the lines y=x and y=1.06x and is considered in-phase ventilation. Regions B, C, and D are considered out-of-phase.

The calculation of out-of-phase ventilation determined by using LER-N matches well with the findings of Shao et al. [72]. In this study it was determined that 13.9% of all voxels were considered out-of-phase compared to 19.3% from Shao et al. [72], which included a different cohort of subjects. Note that Shao et al. [72] considered voxels out-of-phase if the LER-N>1.05*LER-2, while this work considers voxels out-of-phase when the LER-N>1.06*LER-2. The value 1.06 was selected as it is the maximum intensity difference term in the gamma pass rate.

4.2.3.1 Cross-validation

Cross-validation showed that the gamma pass rate of the 25 models ranged between 57.45% and 58.29%, with the average gamma pass rate being 57.95%. The gamma pass rates and R^2_{adj} for the 25 different polynomial models are summarized in Tables 4.2 and 4.3.

Table 4.2: Average gamma pass rates for each polynomial model.

	Dose Polynomial Order				
J_{preRT} Polynomial Order	**1**	**2**	**3**	**4**	**5**
1	58.29%	58.23%	57.64%	58.22%	58.25%
2	58.19%	58.17%	57.48%	57.99%	57.94%
3	58.01%	58.03%	57.45%	58.00%	57.95%
4	57.98%	57.81%	58.09%	58.00%	57.95%
5	58.14%	57.97%	58.02%	57.89%	57.95%

Table 4.3: Average R^2_{adj} for each polynomial model.

	Dose Polynomial Order				
J_{preRT} Polynomial Order	**1**	**2**	**3**	**4**	**5**
1	84.93%	92.56%	95.94%	97.50%	97.49%
2	92.25%	92.57%	96.05%	97.75%	97.75%
3	92.27%	92.63%	96.04%	97.78%	97.76%
4	92.24%	92.69%	97.81%	97.82%	97.97%
5	92.31%	92.72%	97.96%	97.97%	97.97%

The gamma pass rates and R^2_{adj} values for each model were summed in quadrature, choosing the highest value as the model to continue analysis with. It was determined

that a polynomial fit using third order for dose and fifth order for pre-RT Jacobian value (poly35) best modeled the input data.

The polynomial surface fits for the LER-N and the LER-2 for poly35 are represented in Figures 4.3 and 4.4, respectively.

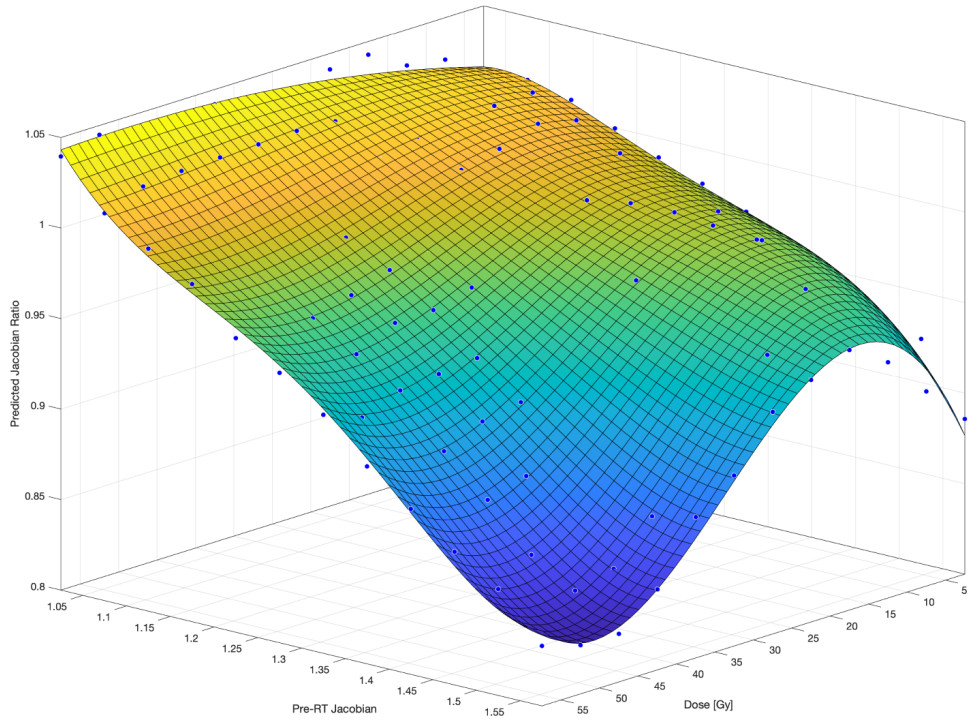


Figure 4.3: Predicted Jacobian ratio polynomial fit (shaded surface) using the LER-N training dataset (points). Polynomial with third order for dose and fifth order for pre-RT ventilation.

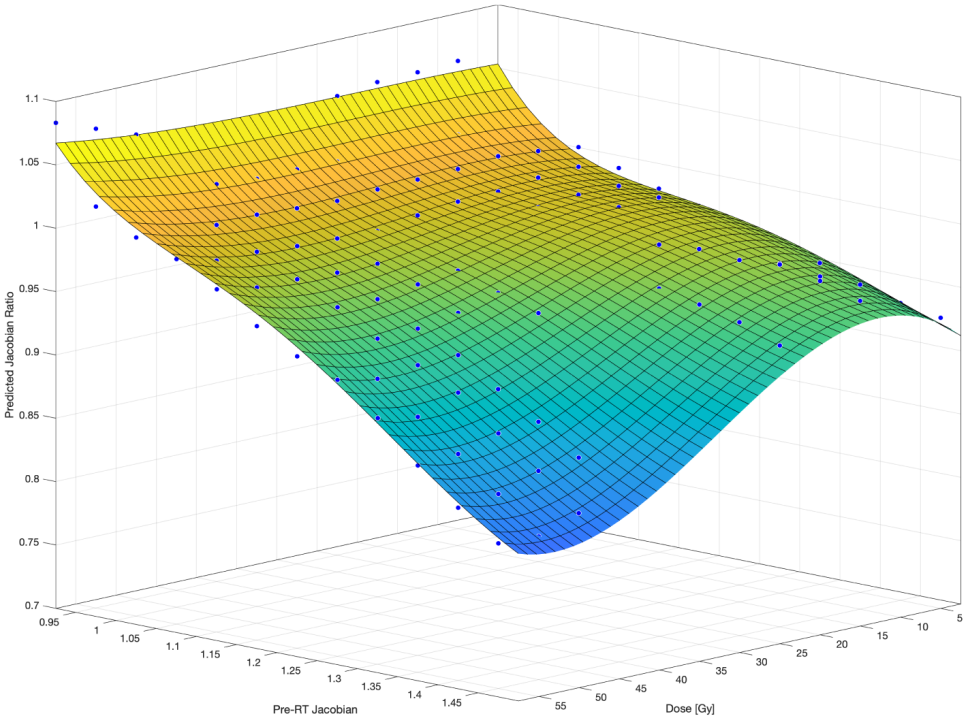


Figure 4.4: Predicted Jacobian ratio polynomial fit (shaded surface) using the LER-2 training dataset (points). Polynomial with third order for dose and fifth order for pre-RT ventilation.

The equation corresponding to the poly35 model fit for the LER-N model is given as

$$\begin{aligned} z = & -33.2 + 0.6x + 123.5y + 0.002x^2 - 2.1xy \\ & - 174.6y^2 - 2.4 \times 10^{-6}x^3 - 0.004x^2y + 2.5xy^2 \\ & + 120.5y^3 - 2.6e \times 10^{-6}x^3y + 0.003x^2y^2 \\ & - 1.4xy^3 - 40.3y^4 + 4.9 \times 10^{-6}x^3y^2 - 0.001x^2y^3 + 0.3xy^4 + 5.2y^5 \end{aligned} \tag{4.5}$$

where x is the dose in Gray, y is the J_{PreRT}, and z is the predicted Jacobian ratio.

The equation corresponding to the poly35 model fit for the LER-2 model is given as

$$
\begin{aligned}
z = {} & 36.8 - 0.03x - 134.4y + 1.7 \times 10^{-5}x^2 + 0.07xy \\
& + 200.2y^2 - 5.8 \times 10^{-6}x^3 + 6.4 \times 10^{-4}x^2y - 0.06xy^2 \\
& - 147.9y^3 + 6.6 \times 10^{-6}x^3y - 9.3 \times 10^{-4}x^2y^2 \\
& + 0.03xy^3 + 54.1y^4 - 1.2 \times 10^{-7}x^3y^2 + 2.2 \times 10^{-4}x^2y^3 \\
& - 0.003xy^4 - 7.9y^5
\end{aligned} \tag{4.6}
$$

where x is the dose in Gray, y is the J_{PreRT}, and z is the predicted Jacobian ratio.

The poly35 model was then used for additional validation using the remaining UW subjects and UI subjects.

4.2.3.2 Additional Validation

Using the additional UW subjects and the UI subjects for further validation, it was found that for voxels that received 20 Gy or greater, there was a significant increase from 52% to 59% (p=0.03) in the gamma pass rates of the LER-N model predicted post-RT Jacobian maps to the actual post-RT Jacobian maps, relative to the LER-2 model. Additionally, the accuracy significantly increased from 68% to 75% (p=0.03) using the LER-N model, relative to the LER-2 model. Table 4.4 summarizes the mean, variance, Pearson Correlation, and p-value for PPV, TPR, TNR, ACC, and gamma pass rate for both models.

The voxels labeled as out-of-phase pre-RT were analyzed for both LER-N and LER-2 models to observe the effect of out-of-phase ventilation on the model's ability to predict post-RT ventilation. Figure 4.5 is a histogram of $J_{postRT/predict}$ for the LER-2 and LER-N models. Using the LER-N model to predict post-RT Jacobian values for each voxel

Table 4.4: Mean, variance, Pearson Correlation, and p-value for PPV, TPR, TNR, ACC, and gamma pass rate for both models.

	Mean	Variance	Pearson Correlation	P(T<=t) two-tail
PPV (LER-2)	64.05%	10.67%	89.37%	0.51
PPV (LER-N)	67.26%	11.72%		
TPR (LER-2)	23.33%	4.60%	66.10%	0.069
TPR (LER-N)	15.13%	1.72%		
TNR (LER-2)	91.79%	2.15%	31.44%	0.24
TNR (LER-N)	96.08%	0.35%		
ACC (LER-2)	67.69%	1.42%	47.36%	**0.026**
ACC (LER-N)	74.62%	1.20%		
Gamma (LER-2)	51.90%	2.78%	82.87%	**0.026**
Gamma (LER-N)	58.63%	3.94%		

resulted in a mean $J_{postRT/predict}$ of 1.002 (SD=0.094), which is closer to the desired value of unity and significantly different ($p << 0.001$) from the LER-2 mean of 1.061 (SD=0.12).

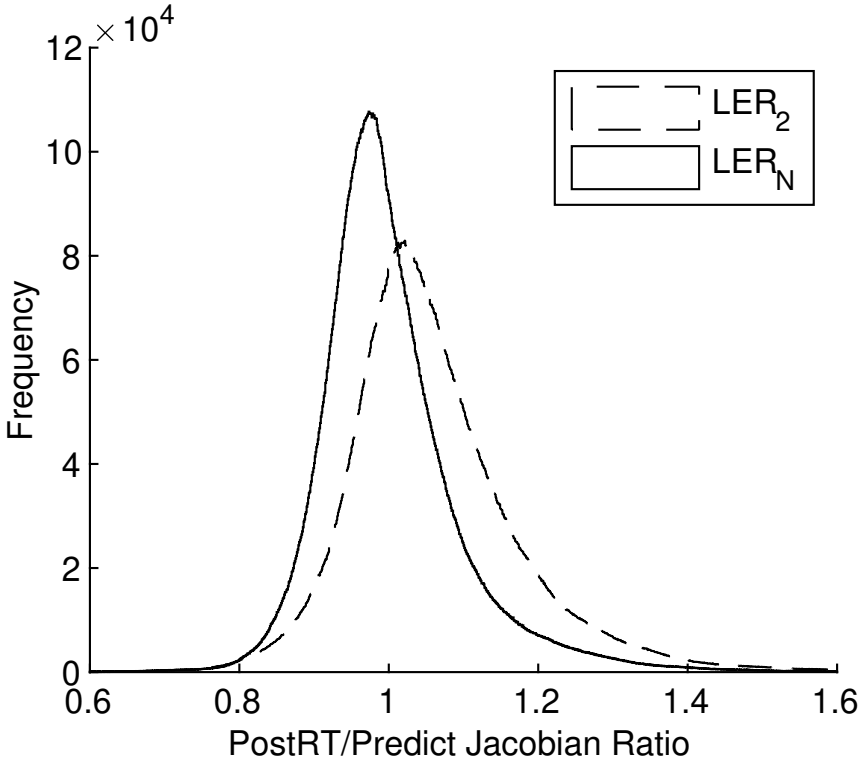


Figure 4.5: Histogram of $J_{postRT/predict}$ for voxels determined as out-of-phase. The dashed outline represents LER values calculated using LER-2 and the solid outline represents LER values calculated using LER-N.

Figure 4.6 is a histogram of $J_{postRT/predict}$ for all voxels that received 20 Gy or more, were predicted to decline in function by more than 6%, and were identified as high-function pre-RT ($J_{preRT} > 1.1$), using LER-N. 66% of the voxels had a J_{postRT} smaller than $J_{predict}$ indicating lung damage occurred in regions not predicted by the LER-N model.

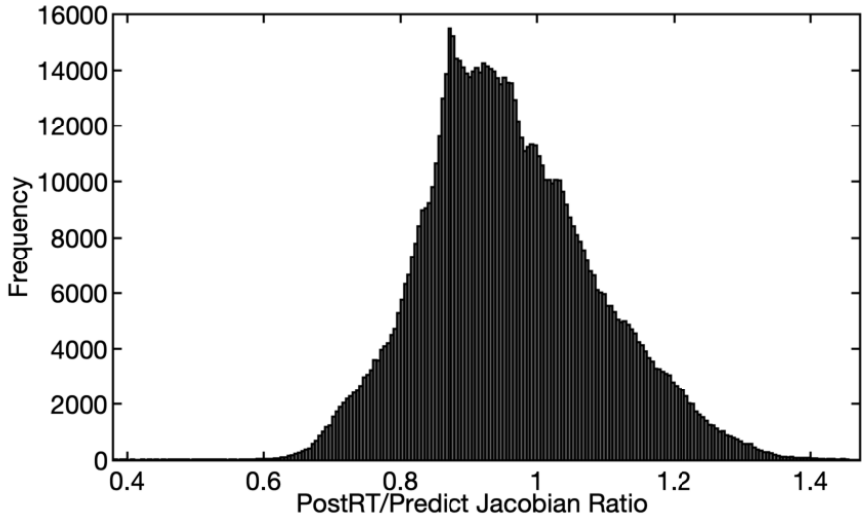


Figure 4.6: Histogram of $J_{postRT/predict}$ for all voxels that received 20 Gy or more, were predicted to decline in function by more than 6%, and were identified as high-function pre-RT ($J_{preRT} > 1.1$). Jacobian values were calculated using LER-N.

4.2.4 Discussion

It was found that when calculating LER using multiple phases, the dose response polynomial regression model is able to significantly improve the accuracy of the predicted post-RT ventilation map compared to the dose response model based on LER-2. This is due to the LER-N maps being less noisy and less dependent on anomalies that may occur at any given phase, and accounting for out-of-phase ventilation. This is the first study to use LER-N to build a dose response model to predict post-RT ventilation change on the voxel level.

The subject shown in Figure 4.1 has extreme out-of-phase breathing. It is hypothesized that this is due to the subject modifying their method of breathing during the scan, switching between abdominal breathing and thoracic breathing. Abdominal breathing can cause the external surrogate to be out-of-phase in relation to the internal respiratory cycle. This is mostly mitigated by providing the audio breathing instructions, but

additional studies to improve repeatability in ventilation measurements are needed to more accurately model response to radiation therapy [157]. Additional coaching may be required to ensure patients do not switch between abdominal and thoracic breathing.

From Table 4.4, it can be seen that the variance for the metrics calculated are relatively low. This is due to the regions analyzed being restricted to the 20 Gy isodose line and the amount of voxels inside that region that exhibit a change in ventilation larger than the 6% threshold is relatively low. Within this region there are approximately 10 times more voxels that experience no change compared to those that do change post-RT.

When only the voxels that were labeled as out-of-phase pre-RT were examined, the LER-N model was significantly better at predicting post-RT ventilation change compared to the LER-2 model. However, while the LER-N model resulted in a more accurate prediction of the post-RT ventilation maps than the LER-2, there was a reduction in TPR and room to further improve the predictive power of the model. From Figure 4.6, J_{postRT} tends to be smaller than $J_{predict}$. 66% of the voxels had J_{postRT} smaller than $J_{predict}$, meaning that the model underestimates the decline in function a voxel will undergo due to dose, possibly due to lung damage in regions "upstream" of those directly receiving the radiation dose.

There are likely additional parameters that need to be incorporated into the model in order to more accurately predict post-RT ventilation change. The current model fails to identify all damage such as pulmonary vasculature damage and airway damage [29], and changes in pressure within the lung due to inflammation and edema. This lack of variables may be the reason for the model underestimating the reduction in ventilation observed post-RT. Additionally, there may not be large improvements in the accuracy of the N-phase post-RT predictions for some subjects depending on the amount of out-of-phase ventilation for the scan. The amount of out-of-phase voxels compared to in-phase voxels are relatively low, making it less likely that the out-of-phase voxels lie within the dose distribution or 20 Gy cutoff. The difference in the predicted post-RT ventilation

maps produced by the two models would be small for regions with small amounts of out-of-phase voxels. However, using LER-N to calculate ventilation is most useful for patients who exhibit out-of-phase breathing as this can drastically affect the ventilation maps, and thus affect the functional avoidance plan.

4.3 Predicting Pulmonary Ventilation Damage After Radiation Therapy for Non-Small Cell Lung Cancer Using Deep Learning

The polynomial regression model was developed to predict ventilation changes following radiation therapy, however the predictive accuracy of this model suggested that there were additional damage mechanisms that were not properly incorporated. The previous model had only considered dose to regions of the lung and pre-RT ventilation, and were fit to fifth-order polynomials. Additionally, this model was created at a voxel level and lost any spatial relationship between adjacent voxels. In this section, various machine learning models were developed to predict pulmonary ventilation change maps as a result of pre-RT ventilation and the delivered dose distribution. Predictions from the ML models were compared using 8-fold cross-validation to predictions from the polynomial regression model described in the previous section, with a focus on improvement of properly identifying regions or voxels that experienced a decline in ventilation post-RT. Accurate predictions of post-RT pulmonary ventilation change as a result of RT can improve the clinical utility of functional avoidance RT and can be delivered with more confidence that the intended function sparing will occur.

4.3.1 Methods

4.3.2 4DCT Acquisition & Datasets

The images used in this study were collected from 82 lung cancer patients that participated in a University of Wisconsin-Madison IRB-approved prospective study (NCT02843568) who underwent RT for non-small cell lung cancer (NSCLC). Exclusion criteria for this study included prior or future planned surgery for the existing lung cancer, prior thoracic radiotherapy, severe COPD, oxygen dependence, known underlying collagen vascular disease, under 18 years of age, and a Karnofsky score $< 60\%$. Each subject had two 4DCT scans separated by five minutes acquired pre-RT and three months post-RT for a total of four scans. Subjects either received a standard fractionation course of 2-2.2 Gy x 30 fractions or a stereotactic body radiation therapy (SBRT) course of 10-12 Gy x 5 fractions. One subject received an SBRT course of treatment with 8 Gy x 5 fractions. The patient information for this study cohort is summarized in Table 4.5.

Table 4.5: Study cohort characteristics

Number of Patients	82
Mean age (range)	74 y (50-94)
Gender (male/female)	43/39
Tumor location	
Right Upper	33
Right Middle	9
Right Lower	12
Left Upper	18
Left Lower	10
Mean PTV Volume (range)	220.2 cc (8.5-1935.1)
Mean Prescription Dose (range)	54.6 Gy (40-66)
SBRT/Standard Fractionation	48/34

All scans were acquired on a Siemens SOMATOM Definition Edge 128 slice CT scanner (Siemens Healthineers AG, Erlangen, Germany) with a 0.6 mm slice thickness, 0.5 second tube rotation time, 0.09 pitch, 120 kVp, and 100 mAs/rotation. The Varian Real-Time Position Management (RPM) (Varian Medical Systems Inc., Palo Alto, CA) system monitored respiration-induced movement and captured respiratory traces used for the 4DCT reconstructions. Images were reconstructed using a 'sharp' reconstruction kernel (Siemens B50f). Melodic breathing instructions were played at a rate of 15 breaths per minute during each image acquisition which has been demonstrated to increase the repeatability between scans and reduce the occurrence of artifacts [141].

4.3.2.1 4DCT-derived Ventilation

This work utilized an image registration algorithm with multi-resolution cubic B-spline parameterization and used the sum of squared tissue volume difference (SSTVD) as the similarity metric as described in Section 2.4.2. The Jacobian-based method described by Shao et al. [73] and Section 2.4.3 was used to generate ventilation maps for each subject at both timepoints (pre- and post-RT). LER-N maps for the two scans at each time point were geometrically averaged to reduce noise within the map introduced by random changes in patient respiratory patterns between scans.

4.3.2.2 Data Preprocessing

As the total number of subjects available in the training dataset was relatively small (n=82), data augmentation was performed to increase the dataset size and training performance. The augmentation was performed to include various LER-N ventilation maps created at different tidal volumes (i.e., changing the number of phases used, N). For each subject, the tidal volume at each discrete breathing phase was calculated, and up to 4 (40% - 100% inspiration, in increments of 20%) separate LER-N maps were created at each timepoint and the breathing phases that had tidal volumes within 100cc had their LER-N maps paired together. Tissue expansion using images with tidal volumes within 100cc of each other has been shown to increase the repeatability of the ventilation measurement [142]. This increased the size of the dataset from 82 to 242.

All images were downsampled to 3 mm isotropic voxels for training due to memory constraints. For each subject, a bounding box was calculated using a lung mask [158, 159] and was used to crop all images for the subject to both reduce image size and remove irrelevant input voxels (i.e. anatomy outside of the lungs). Ventilation maps were clipped to a range of 1 to 2 and normalized by the max. Lastly, dose images were clipped to

a range of 0 to 75 and normalized by the max. The maximum threshold values used for clipping were determined by calculating the 99.7 percentile (3 standard deviations). Values greater than the thresholds were considered extreme outliers and were clipped.

4.3.3 Deep Learning Models

Several different types of network architectures were investigated to determine the optimal model for predicting ventilation decline. Specifically, three types of machine learning networks were investigated: Seg3DNet, a variation on the classical U-Net, ResNet3D, a residual neural network, and a conditional generative adversarial network (cGAN).

4.3.3.1 Network Architectures

The Seg3DNet network architecture has been successfully used for directly estimating the local tissue expansion between lungs imaged at two lung volumes as well as pulmonary fissure detection [96, 106–108]. The Seg3DNet architecture is well suited for this work as it is designed to be less memory-intensive compared to SegNet and U-Net [119, 160]. This allows the network to learn information from entire stacks of images instead of cropped ROIs which might lose important spatial information from regions surrounding the ROIs. This method sacrifices image resolution in order to preserve global context, which is crucial for properly modeling local expansion. Figure 4.7 shows the Seg3DNet architecture as described by Gerard et al. [106]. He variance scaling initialization [161] was used for convolutional filter parameters and the final activation layer was a sigmoid function. The network used the Nadam optimizer with an initial learning rate of 10E-5. For this work, the input to the network was a 4D tensor with two 3D volumetric images stacked along the channel dimension corresponding to the dose distribution and pre-RT ventilation map. The output of the model was the voxel-wise ratio of the post-RT

ventilation map divided by the pre-RT ventilation map, referred to as the Jacobian ratio (J_{ratio}). The J_{ratio} map represents local ventilation change where voxels with a J_{ratio} less than 0.94 constitutes ventilation decline, as defined in previous studies [79, 142].

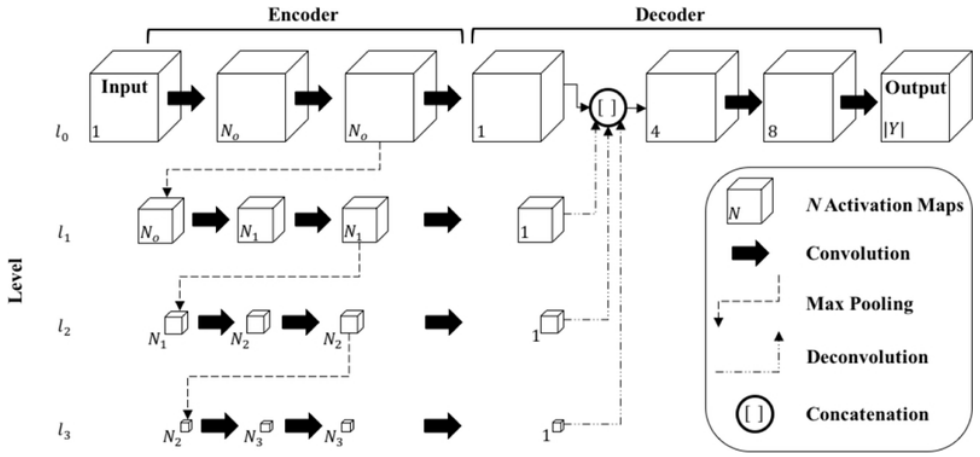


Figure 4.7: Seg3DNet architecture for the dose-response model. Four resolution levels (L=4) are proposed, with the number of activation maps designated as $N_i = 2^{i+5}$ for i = 0,..,L. At each resolution level the spatial resolution of the image representation is downsampled by a factor of 2. Reprinted from "FissureNet: A Deep Learning Approach For Pulmonary Fissure Detection in CT Images", Gerard, S.E., Patton, T.J., Christensen, G.E., et al. Figure 2 (IEEE Trans Med Imaging 2019;38(1):156-166)[106]. © 2019 IEEE.

The ResNet3D network was based on the High-Res3DNet [162] and the 3D ResNet [103]. The model architecture, as seen in Figure 4.8, consisted of three groups of residual blocks (ResBlocks). Each ResBlock contained a pair of atrous convolutions where the output was added to the skip connection, and the size of the dilation (one, two, and four) and number of filters, F, (16, 32, and 64) increased in each of the three groups. He variance scaling initialization [161] was used for convolutional filter parameters and dropout layers were implemented in the second and third groups with a rate of 0.25. The final activation layer used a sigmoid function. The network used the Nadam optimizer with an initial learning rate of 10E-5. The input was a concatenation along the channel dimension of

the source image tensor (pre-RT ventilation map and dose distribution), and the output image was the J_{ratio}.

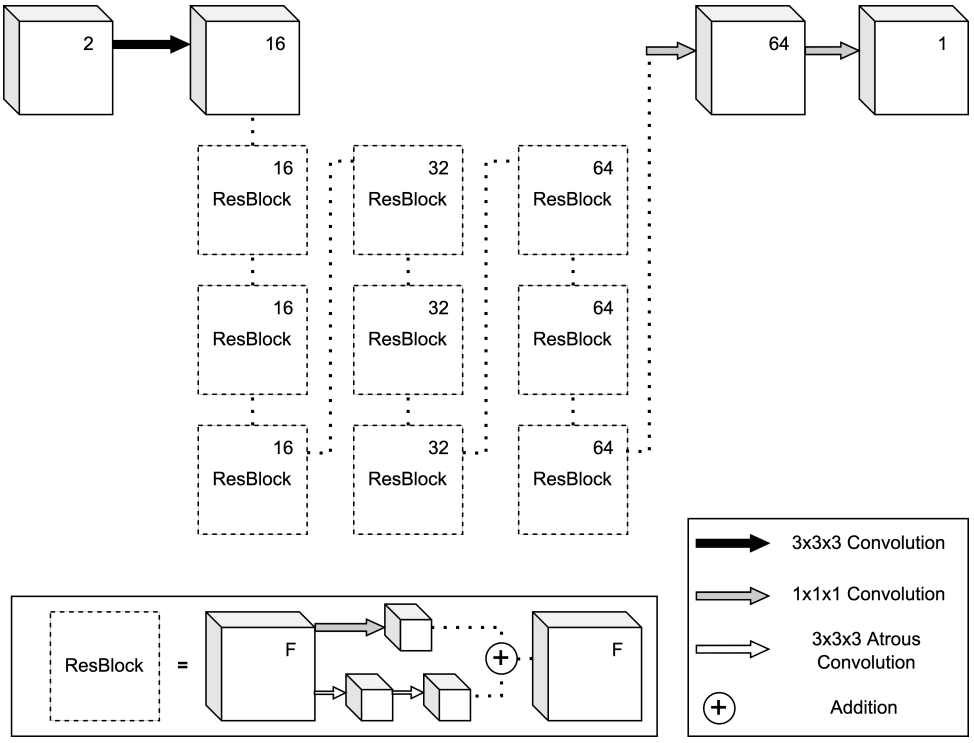


Figure 4.8: Diagram of residual neural network architecture used in this work. Input to the model was a 2 channel, 4D tensor, and was trained to output the voxel-wise ratio of post-RT to pre-RT ventilation maps (J_{ratio}). The number of filters, F, used in each convolution is labeled in the top right corner of each block.

A generative adversarial network (GAN) is a specific type of machine learning architecture that simultaneously trains a generator network and a discriminator network to produce plausible synthetic images [120]. This work utilized a conditional GAN (cGAN) similar to the Pix2Pix architecture, a network specifically designed for image-to-image translation through which generated output images are conditional on the input source image [163]. The network consisted of one generator network, whose purpose was to generate plausible images in the target domain, and one discriminator network, whose purpose was to distinguish between real images in the target domain and images produced by the generator network. The generator takes in the image input, which in this work was a 4D

tensor consisting of two 3D volumetric images stacked along the channel dimension, and translates it to the output domain via voxel-to-voxel transformation. The target image in this work is the J_{ratio}.

The discriminator network was a deep convolutional neural network (DCNN) designed for image classification. Specifically, this work implemented the PatchGAN discriminator which classifies patches as either fake or real, and runs convolutionally across the entire input image before averaging the result of each patch to give an overall classification score. The network consisted of four layers of strided convolutions (stride = 2) with increasing number of filters (64, 128, 256, 512) and kernel size of 4 voxels. Each strided convolution was followed by batch normalization and LeakyRelu activation with a slope of 0.2. The last 2 layers had a stride of one, with the final convolutional layer mapping to a single channel output followed by a sigmoid activation layer. The discriminator network used the Adam optimizer with Nesterov momentum (Nadam) with an initial learning rate of 1E-5. The loss was weighted by 0.5 for every discriminator update in order to slow down the discriminator training process. The input to the discriminator was a concatenation along the channel dimension of the source image tensor (pre-RT ventilation map and dose distribution) and the target image (J_{ratio} = post-RT/pre-RT ratio map), resulting in a 4D tensor with 3 channels.

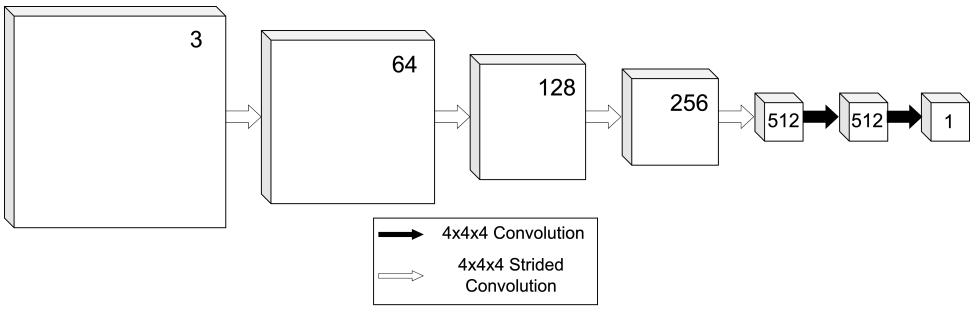


Figure 4.9: Diagram of the discriminator network architecture used in this work. Input to the discriminator was a 3 channel, 4D tensor, resulting in an image patch of size 8x8x8. The number of filters, F, used in each convolutions is labeled in the top right corner of each block.

The generator network used the same ResNet3D architecture as previously described. The generator network used the Nadam optimizer with an initial learning rate of 0.01. The input to the generator was a concatenation along the channel dimension of the source image tensor (pre-RT ventilation map and dose distribution), and the output image was the J_{ratio}.

Eight separate models were trained in order to perform 8-fold cross-validation for model evaluation, with 6 folds leaving out 10 subjects each and two folds leaving out 11 subjects each. All work was implemented in Tensorflow v2.6.0. The Seg3DNet and ResNet3D models were trained for 1000 epochs and employed early stopping which stopped training when the validation loss did not improve over 50 epochs to prevent the models from overfitting to the data. The cGAN models were trained for 1000 epochs, and models were saved every 10 epochs resulting in a total of 100 different models for each fold. All models for all architectures were trained on a NVIDIA A100 80GB GPU and implemented in Tensorflow v2.6.0. The Seg3DNet, ResNet3D, and cGAN training time per fold took 25 hours, 12 hours, and 56 hours, respectively.

4.3.3.2 Loss Function

The loss function used in this work for the cGAN models followed a similar method to the loss function that was used in the original Pix2Pix GAN [163], where the discriminator was trained in a standalone manner minimizing the negative log likelihood of discriminating between real and fake images (adversarial loss). The generator was trained using a composite loss function of both the adversarial loss for the discriminator model and an image quality loss to encourage the generator to output images that are plausible translations of the input. The adversarial loss can be defined as

$$\mathcal{L}_{adversarial} = \mathbb{E}_{x,y}[logD(x,y)] + \mathbb{E}_{x,z}[log(1 - D(x, G(x,z)))] \tag{4.7}$$

where x is the observed image, y is output, and z is a random noise vector. In this work, random noise was introduced both through the variability in the input images as well as the dropout layers within the generator as proposed by Isola et al. [163].

An asymmetrical structural similarity index measure (SSIM) was used as the secondary loss function for the generator, which was based on the work of Porter et al. [103], who used an asymmetrical mean absolute error (AMAE) for predicting perfusion defects. Gerard et al. [96] used SSIM as the loss function to train a CNN to estimate local tissue expansion as it has been shown to improve perceptual image quality. The purpose of using an asymmetrical loss is to increase penalization for predictions that under-estimate ventilation decline. In general, SSIM is defined as

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_2)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{4.8}$$

where x and y are windows with size 11×11×11 voxels from the predicted image (J_{cGAN}) and ground truth image (J_{LER-N}), respectively; μ_x and μ_y are corresponding averages, σ_x and σ_y are the variances of x and y, σ_{xy} is the covariance of x and y, and c1 and c2 are constants. A sliding window was used to generate a SSIM map for the entire image, which was implemented as a convolutional layer with fixed filters. To determine the asymmetrical loss (aSSIM), the SSIM was multiplied by a factor α, defined as

$$\alpha = \frac{ln(2)}{ln(2 + \delta + \epsilon)} \tag{4.9}$$

where δ is the voxel-wise difference between the ground truth and predicted image, and ϵ is a pre-determined parameter that determines the weight of the penalization for underestimating damage. As epsilon approaches zero, the higher the penalty, and was set at $\epsilon = 0.001$. The loss for each image was taken as the mean aSSIM within the lung mask.

The total composite loss can then be defined as:

$$\mathcal{L}_{total} = 0.5 * \mathcal{L}_{adversarial} + 50 * aSSIM \tag{4.10}$$

The weights for each of the loss terms (0.5 for the adversarial loss and 50 for the aSSIM loss) were set with a 1:100 ratio as was used by Isola et al. [163].

The aSSIM loss function was used for both the ResNet3D and Seg3DNet models.

4.3.3.3 Performance and Statistics

Cross-validation was performed on the 8 different models for each network architecture for evaluation. The group of subjects in each fold that was left out only had one LER-N ventilation map as to not skew the validation results by including datasets with similar data. Additionally, the data used for validation was not downsampled or cropped, with the only preprocessing step being min-max normalization. The cGAN model used for evaluation from each fold was determined by selecting the model which had the minimum generator loss, while the best models saved from early stopping were used for the Seg3Dnet and ResNet3D models. The full-sized image was used for testing as this is what would be used in a clinical application of the model in order to provide avoidance ROIs for treatment planning. Additional testing was performed using data that underwent the same preprocessing steps as the training set (i.e. cropping and downsampling) and found no statistically significant differences between results of the full-sized images

and cropped/downsampled images. The results presented are thus from testing using the full-sized images.

Model performance was assessed by receiver operating characteristic (ROC) metrics such as true positive rate (TPR), true negative rate (TNR), accuracy (ACC), and Dice similarity coefficient (DSC). In this work, a true positive is defined as a voxel that was correctly predicted to have ventilation damage, corresponding to $J_{ratio} \leq 0.94$ (i.e., a decline in ventilation of greater than 6%), whereas a false positive was a voxel that was incorrectly predicted to have a $J_{ratio} \leq 0.94$ when the ground truth voxel had a J_{ratio} that was > 0.94. Ground truth and predicted images were masked such that the metrics only considered voxels inside the lung mask and ignored voxels in other anatomical regions. Youden's index (J) [164] was an additional metric used to evaluate the models' balance between sensitivity and specificity, and is defined as

$$J = sensitivity + specificity - 1 \tag{4.11}$$

J corresponds to the height above the chance line on a ROC curve and serves as the probability of an informed decision.

Furthermore, model performance and metrics were compared to the results of a previously developed polynomial regression model described in Chapter 5, subsequently referred to as POLY. The paired samples t-test was used to compare the metrics between the Seg3DNet, ResNet3D, cGAN, and POLY models. Results were considered statistically significant when $p<0.05$.

4.3.4 Results

Results of the 8-fold cross validation across each fold, as well as the average of all folds, are summarized in Table 4.6 and 4.7.

Table 4.6: 8-fold cross validation results for the cGAN and POLY models in the form of mean (SD).

	cGAN					POLY				
Fold no.	TPR	TNR	ACC	DSC	J	TPR	TNR	ACC	DSC	J
1	0.77 (0.10)	0.61 (0.18)	0.65 (0.13)	0.48 (0.15)	0.38 (0.11)	0.05 (0.03)	1.00 (0.01)	0.80 (0.11)	0.08 (0.06)	0.04 (0.03)
2	0.84 (0.08)	0.63 (0.13)	0.67 (0.10)	0.50 (0.10)	0.45 (0.08)	0.08 (0.08)	0.99 (0.02)	0.80 (0.07)	0.13 (0.11)	0.07 (0.07)
3	0.70 (0.24)	0.59 (0.25)	0.66 (0.13)	0.45 (0.16)	0.29 (0.09)	0.14 (0.19)	0.97 (0.05)	0.80 (0.11)	0.19 (0.22)	0.11 (0.15)
4	0.89 (0.08)	0.36 (0.20)	0.49 (0.16)	0.46 (0.14)	0.28 (0.14)	0.28 (0.20)	0.92 (0.10)	0.76 (0.07)	0.32 (0.18)	0.20 (0.13)
5	0.77 (0.21)	0.58 (0.23)	0.64 (0.16)	0.48 (0.16)	0.37 (0.20)	0.17 (0.14)	0.97 (0.05)	0.79 (0.06)	0.23 (0.19)	0.13 (0.13)
6	0.48 (0.23)	0.82 (0.13)	0.76 (0.10)	0.41 (0.13)	0.31 (0.12)	0.09 (0.08)	0.98 (0.02)	0.80 (0.11)	0.14 (0.11)	0.07 (0.06)
7	0.66 (0.22)	0.68 (0.19)	0.67 (0.11)	0.41 (0.17)	0.34 (0.15)	0.17 (0.18)	0.96 (0.06)	0.81 (0.08)	0.20 (0.17)	0.12 (0.13)
8	0.70 (0.19)	0.65 (0.16)	0.66 (0.11)	0.47 (0.13)	0.35 (0.10)	0.15 (0.13)	0.97 (0.05)	0.78 (0.06)	0.22 (0.13)	0.10 (0.09)
Average	0.72 (0.21)	0.62 (0.21)	0.65 (0.14)	0.46 (0.14)	0.34 (0.14)	0.14 (0.15)	0.97 (0.05)	0.79 (0.08)	0.19 (0.16)	0.11 (0.11)

Box plots of TPR, TNR, ACC, and DSC for the cGAN, ResNet3D, Seg3DNet, and POLY models, averaged across all cross-validation folds, are shown in Figure 4.10. There was a statistically significant difference in the means of all four metrics. The average true positive volume increased from 104±119 cc in the POLY model to 565±332 cc in the

Table 4.7: 8-fold cross validation results for the ResNet3D and Seg3DNet models in the form of mean (SD).

	ResNed3D					Seg3DNet				
Fold no.	TPR	TNR	ACC	DSC	J	TPR	TNR	ACC	DSC	J
1	0.21 (0.07)	0.91 (0.02)	0.75 (0.09)	0.26 (0.06)	0.12 (0.06)	0.00 (0.00)	1.00 (0.00)	0.78 (0.12)	0.00 (0.00)	0.00 (0.00)
2	0.25 (0.08)	0.90 (0.02)	0.77 (0.07)	0.30 (0.07)	0.16 (0.08)	0.00 (0.00)	1.00 (0.00)	0.79 (0.08)	0.00 (0.00)	0.00 (0.00)
3	0.29 (0.15)	0.83 (0.06)	0.72 (0.08)	0.30 (0.15)	0.12 (0.11)	0.25 (0.23)	0.79 (0.16)	0.70 (0.10)	0.23 (0.18)	0.04 (0.09)
4	0.45 (0.25)	0.78 (0.12)	0.71 (0.05)	0.39 (0.16)	0.23 (0.15)	0.14 (0.08)	0.86 (0.03)	0.68 (0.09)	0.17 (0.05)	0.00 (0.07)
5	0.32 (0.17)	0.83 (0.07)	0.72 (0.06)	0.32 (0.16)	0.15 (0.13)	0.48 (0.13)	0.57 (0.09)	0.56 (0.08)	0.32 (0.13)	0.06 (0.16)
6	0.33 (0.11)	0.84 (0.04)	0.74 (0.09)	0.31 (0.10)	0.17 (0.08)	0.09 (0.04)	0.94 (0.03)	0.77 (0.11)	0.13 (0.05)	0.03 (0.04)
7	0.32 (0.21)	0.85 (0.06)	0.74 (0.06)	0.28 (0.12)	0.17 (0.16)	0.00 (0.00)	1.00 (0.00)	0.81 (0.08)	0.00 (0.00)	0.00 (0.00)
8	0.24 (0.12)	0.89 (0.03)	0.75 (0.06)	0.28 (0.10)	0.13 (0.11)	0.85 (0.12)	0.12 (0.11)	0.28 (0.09)	0.34 (0.09)	0.03 (0.07)
Average	0.30 (0.16)	0.86 (0.07)	0.74 (0.07)	0.30 (12)	0.16 (0.12)	0.23 (0.31)	0.78 (0.31)	0.67 (0.20)	0.15 (0.16)	0.01 (0.08)

cGAN model, and the average false negative volume decreased from 654±361 cc in the POLY model to 193±163 cc in the cGAN model. Additionally, the average false positive volume increased from 66±103 cc in the POLY model to 995±534 cc in the cGAN model, and the average true negative volume decreased from 2758±1109 cc in the POLY model to 1829±1119 cc in the cGAN model. All volume changes were statistically significant and are summarized in Table 4.8.

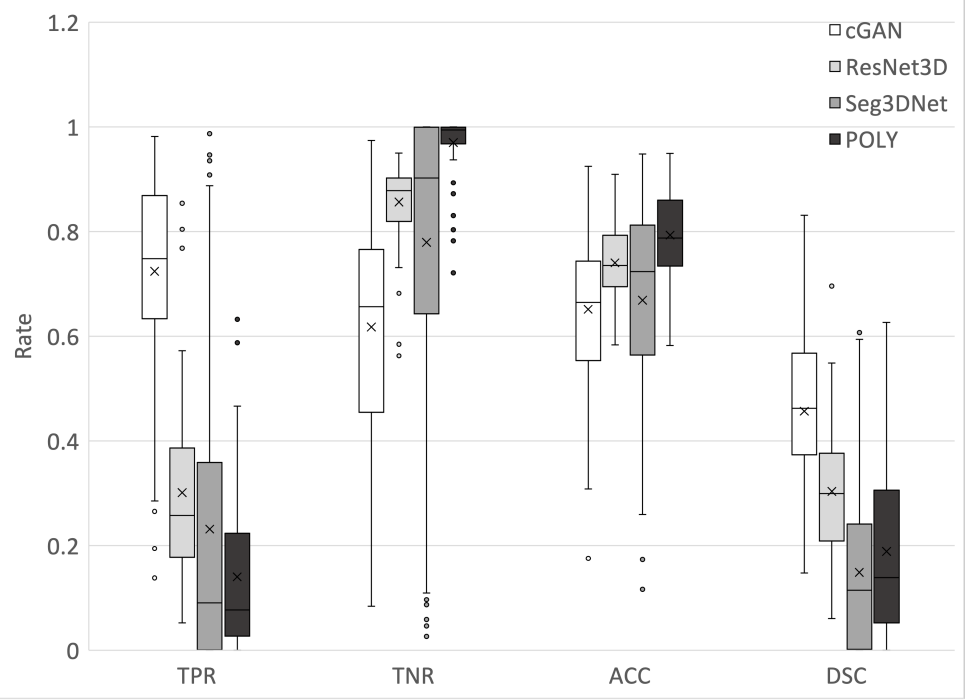


Figure 4.10: Box plots of true positive rate (TPR), true negative rate (TNR), Dice similarity coefficient (DSC), and accuracy (ACC) for the cGAN, ResNet3D, Seg3DNet, and POLY models, averaged across all subjects from the cross-validation folds. There was a significant difference between the means for all metrics between the two models.

Table 4.8: Average volumes of true positives, false negatives, false positives, and true negatives for the cGAN and POLY models in the form of mean±SD, averaged across all subjects from the cross-validation folds.

Volumes (cc)	cGAN	POLY
True Positives	565±332	104±119
False Negatives	193±163	654±361
False Positives	995±534	66±103
True Negatives	1829±1119	2758±1109

Figures 4.11-4.13 show the coronal views of subjects from the testing dataset that represent the mean model performance, the best performing standard fractionation dataset, and best performing SBRT dataset, respectively, with the delivered dose distribution,

ground truth J_{ratio}, POLY model predicted J_{ratio}, and cGAN model predicted J_{ratio}. The first row of images (B-D) in Figures 4.11-4.13 shows the entire range of J_{ratio} values, and the second row of images (E-G) shows J_{ratio} maps thresholded to only show voxels with a $J_{ratio} \leq 0.94$.

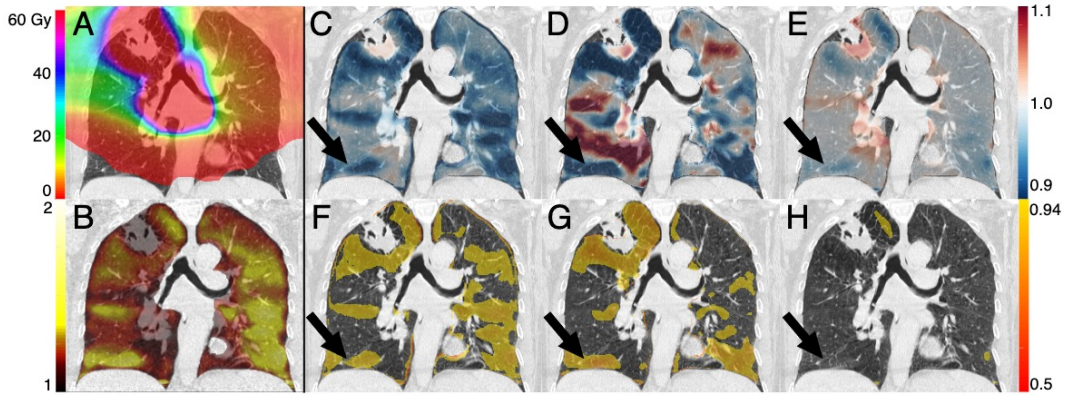


Figure 4.11: Coronal view of one subject from the testing dataset that represents the mean model performance with a true positive rate of 0.70 and true negative rate of 0.68. (A) the delivered dose distribution, (B) pre-RT LER-N, (C-E) cGAN, ground truth, and POLY J_{ratio}, respectively, (F-H) cGAN, ground truth, and POLY J_{ratio}, respectively, thresholded to only show damaged ($J_{ratio} \leq 0.94$) voxels. The arrows in B-G highlight a region of ventilation decline, downstream from the main dose distribution, correctly predicted by the cGAN model and missed by the POLY model.

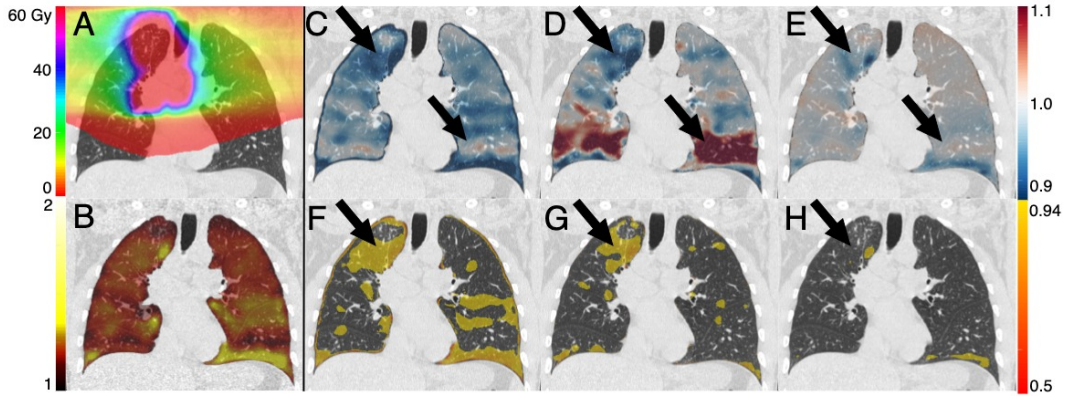


Figure 4.12: Coronal view of the best performing conventional fractionation subject from the testing dataset with a true positive rate of 0.82 and true negative rate of 0.64. (A) the delivered dose distribution, (B) pre-RT LER-N, (C-E) cGAN, ground truth, and POLY J_{ratio}, respectively, (F-H) cGAN, ground truth, and POLY J_{ratio}, respectively, thresholded to only show damaged (J_{ratio} $\leq$0.94) voxels. The arrows pointing to the region in the superior right lung are where the cGAN model was able to properly predict ventilation damage. The arrows pointing to the inferior region of the left lung shows the cGAN model failed to predict increases in ventilation and instead predicted ventilation decline.

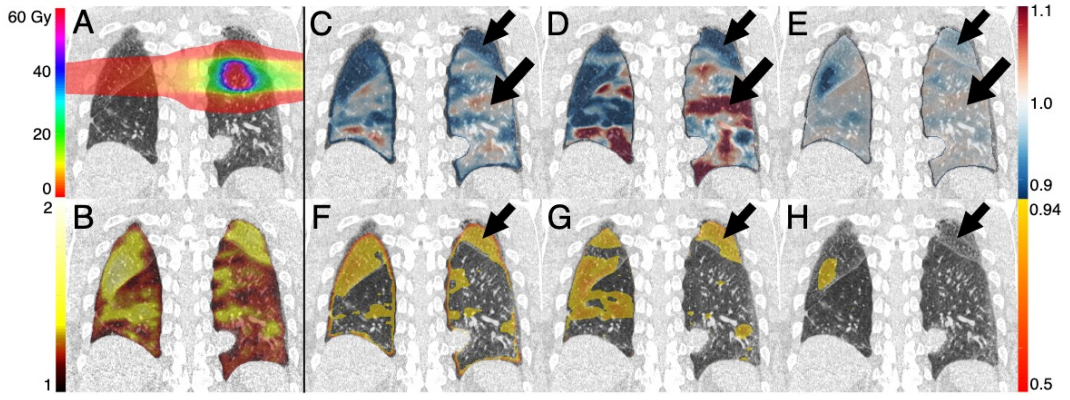


Figure 4.13: Coronal view of the best performing SBRT subject from the testing dataset with a true positive rate of 0.64 and true negative rate of 0.79. (A) the delivered dose distribution, (B) pre-RT LER-N, (C-E) cGAN, ground truth, and POLY J_{ratio}, respectively, (F-H) cGAN, ground truth, and POLY J_{ratio}, respectively, thresholded to only show damaged (J_{ratio} $\leq$0.94) voxels. The cGAN model correctly predicted ventilation decline, indicated by arrows pointing to the superior left lung, and ventilation increase/preservation, indicated by the arrows pointing to the middle left lung.

4.3.5 Discussion

In this study, a DL-based framework that predicts ventilation change following radiotherapy at the voxel level was evaluated. This DL-based model was able to successfully predict ventilation decline, and outperform the previously developed polynomial regression model from Section 4.2. This is the first study to develop a DL-based approach for predicting local ventilation change resulting from radiotherapy treatment for NSCLC.

In the context of functional avoidance lung radiotherapy, it is desirable to identify more regions of the lung that may experience a decline in ventilation due to irradiation rather than correctly predicting regions that will not change or have an increase in ventilation post-RT. By increasing the sensitivity of the model, larger volumes of functional lung can be spared and decrease the probability of the patient developing RILIs. Therefore, while the cGAN model had a significantly lower TNR and ACC, the results show that the cGAN model had a significantly higher TPR and DSC, which is more relevant for the model's clinical utility of predicting radiation-induced ventilation decline in functional lung avoidance treatment planning. Using the output of the model, the predicted J_{ratio}, ROIs can be produced and included in the treatment planning system where the treatment plan can be further optimized to reduce dose to regions predicted to decline in ventilation. It should be noted that the optimization may result in a dose distribution that would push dose into regions that were predicted to not change and/or increase in ventilation, making it necessary to perform a secondary check with the model and the newly optimized dose distribution to verify functional sparing. Additionally, the cGAN model had an average Youden's index (TPR+TNR-1) of 0.34, over three times the Youden's index of 0.11 from the POLY model, indicating a more balanced model and further from the random chance line. Finally, the cGAN model significantly increased the average true positive volume and reduced the average false negative volume which will help create more effective functional avoidance RT plans.

From Figures 4.11-4.13, it can be seen that the POLY model predicts little to no change (i.e., voxels with values near 1.0 and shaded white) while the cGAN model overestimates ventilation decline (i.e., voxels with values < 0.94 and shaded blue). Across the three example subjects shown, the cGAN model generally underestimates both the intensity and volume of voxels that increase or are preserved in ventilation post-RT ($J_{ratio} > 1$). In the first row (C-E) of Figures 4.11-4.13, the cGAN model does not predict many voxels to have $J_{ratio} > 1$ (i.e., regions of ventilation increase/preservation and shaded red). However, the cGAN model, relative to the POLY model, correctly predicts more ventilation decline. It is hypothesized that the decrease in ventilation, indicated by the arrow in the right inferior lobe, in Figure 4.11 is due to the high dose delivered to the right mainstem bronchus and the subsequent generation of bronchi. The cGAN model did well in this subject in correctly identifying this region of ventilation decrease while the POLY model predicted no change. Figure 4.12 shows the best performing standard fractionation subject where the cGAN model correctly predicted ventilation damage, as indicated by the arrow pointing to the region in the superior right lung. Figure 4.12 also shows the cGAN model failed to predict increases in ventilation, indicated by arrows pointing to the inferior region of the left lung, and instead predicted ventilation decline. Figure 4.13 shows the best performing SBRT subject where the cGAN model correctly predicted ventilation decline, indicated by arrows pointing to the superior left lung, and ventilation increase/preservation, indicated by the arrows pointing to the middle left lung.

While almost all folds of the cross-validation resulted in a TPR greater than 0.65, there was one fold (fold 6) that had a lower TPR and higher TNR than the other folds. For two subjects within fold 6 that performed poorly in TPR (TPR < 0.20), it was found that the volume of positives (i.e. declined in function post-RT; $J_{ratio} < 0.94$) was less than half of the average volume of positives across all folds. Additionally, it appears the model was unable to correctly predict voxels that declined in function that were in the contralateral

lung, received less than 1 Gy, and were not downstream from any irradiated airways. One possible explanation could be that these two subjects did not breath consistently during one of the 4DCT acquisitions, however the exact source of this discrepancy was unable to be determined and should be a focus of future research.

In patients with mediastinal or bronchial tumors, ventilation may increase following RT in regions that are ventilated by airways that were previously blocked by the mass. This phenomena may explain the increases (dark red regions) in ventilation seen in Figure 4.11C, where in the middle right lung it would be expected to observe a decrease in ventilation due to the region receiving 20-30 Gy. Irradiating airways has also been shown to affect ventilation in regions downstream from the irradiated region. Vicente et al. [78] found an increase in ventilation preservation when airways were included in functional lung avoidance radiotherapy planning, and Chapter 3 found a decrease in lung ventilation in the regions of the lung supplied by irradiated airways. Future work to improve this model and functional lung avoidance treatment planning should further investigate the impact of 'indirect' ventilation damage due to irradiation of the airways.

This work used both SBRT (5 fraction) and standard fractionation (30 fraction) subjects within the training and testing datasets. When the results were stratified by fractionation cohort, there was a statistically significant difference in the average TPR (5 fraction: 0.65, 30 fraction: 0.82) and TNR (5 fraction: 0.66, 30 fraction: 0.55) between the two cohorts. To investigate this difference between cohorts new models were trained following the same methodology as previously stated. First, a model was trained completely from scratch where the only difference was that the dose distributions were first converted into biologically effective doses (BED) with an $\alpha/\beta = 3$. The model trained with BED dose distributions performed significantly worse than the original cGAN model. An additional four models were created through transfer learning from the worst and best performing cross-validation folds (folds 4 and 6) from the original model. The worst and

best performing models were fine tuned and tested using either the SBRT or standard fractionation subjects corresponding to that fold. It was found that for all fine-tuned models there was no statistically significant improvement on the original cGAN model. This shows that the original model with both fractionation cohorts combined results in the best performing model.

While a DL-based model was successfully built to predict ventilation decline from radiotherapy, there are a few limitations to this study. First, the training of the model may have been limited by the size of the training dataset. Many other medical image-based DL studies use 2D images for the purposes of training as splitting up volumetric images into separate 2D slices greatly increases the size of the dataset. In this study the network was trained using 3D volumes, allowing the network to model local expansion changes as well as ventilation changes occurring distally from the target, which would not be possible using 2D slices or extracted patches. Due to this constraint, image resolution and training size were sacrificed. Ventilation maps are inherently noisy, making downsampling of the training images less important; however, the small training size makes it difficult for the model to learn complex or uncommon relationships in the data. Second, the asymmetrical loss function appeared to have over-predicted damage in most cases. This is partly due to using SSIM as the main loss function, which optimizes local structure similarity and not intensity of the values. Further tuning the asymmetrical factor as well as adding an additional term to account for intensity differences could potentially alleviate this issue. Another limitation of this study is that all data used in this work originated from a single clinic. Patients that participated in this clinical trial received audio guidance during 4DCT acquisition to minimize the potential for artifacts. Deriving ventilation maps from 4DCTs without audio guidance may lead to artifacts that would impact the training or prediction of the model. Further, different institutions may have variation in

the treatment plans created and use different algorithms/processing techniques to create 4DCT-derived ventilation maps. Lastly, the dose distributions and ventilation maps used were deformed to the end-exhale (0EX) breathing phase, however, patients were free breathing during actual treatment delivery. Depending on the location and motion of the target during treatment, the actual dose delivered to the surrounding lung tissue may differ from the dose distribution used to train and evaluate the model. A future iteration of the model should account for breathing motion by incorporating ventilation maps deformed to various phases throughout the breathing cycle.

While it was not in the scope of this study to develop new functional lung avoidance treatment plans using the predictions from the cGAN model, future work should analyze these new treatment plans to determine whether it results in better lung function sparing. Ultimately, the purpose of developing a model that can predict post-RT pulmonary ventilation change is to utilize it in the creation of functional lung avoidance treatment plans that will spare healthy lung tissue and decrease the risk of the patient developing RILIs following treatment. The increase in the true positive volume in the cGAN model was quantified, however this does not directly translate to the volume of functional lung that will be spared during treatment. Factors such as target coverage and other organs at risk dose constraints heavily influence the amount in which the treatment plan can be modified to account for functional lung sparing.

4.4 Chapter Summary

In the first section of this chapter, a polynomial regression model was created by using ventilation images derived using the LER-N method. The LER-N model produced more repeatable ventilation maps allowing for potentially better normal tissue sparing, and improved upon a previously developed polynomial regression model. This new model

also gives insight into the importance of both pre-RT ventilation function and the delivered dose to that particular voxel in predicting post-RT ventilation change. The second section of this chapter developed multiple DL-based models designed to predict post-RT pulmonary ventilation changes resulting from radiotherapy was developed and evaluated. The results show that compared to the LER-N polynomial regression model, a cGAN was able to increase the sensitivity of detecting ventilation damage following radiotherapy. The new DL-based model can be used to create more sensitive functional lung avoidance treatment plans that potentially result in superior lung sparing.

Chapter 5

Integrating Functional Avoidance Information Into Treatment Planning

5.1 Introduction

Currently, functional avoidance treatment plans are created by using discrete avoidance structures that are based on the functional information derived from pre-RT 4DCTs or various other functional lung imaging modalities. Some potential failures of this approach are the discretization of the avoidance structures and the differing approaches to manipulating dose out of the avoidance structures by dosimetrists. Additionally, there is no secondary check with the dose-response model to verify that the optimized plan achieves the desired functional lung sparing.

Some previous studies have investigated the incorporation of pulmonary functional information into the treatment planning process. St-Hilaire et al. [165] implemented SPECT-based optimization for functional avoidance of the lung and found that the perfusion-weighted volume receiving 10 Gy was reduced by 2.2%. They also found that this technique would occasionally cause overdosage regions in the target volume [165]. However, follow-up predictions were not conducted to determine the predicted functional sparing that would occur between the various plans they created. Ding et al. [58] investigated using hyperpolarized xenon-129 MR images to guide treatment planning in a small (n=10 patients) pilot study. For well-ventilated lung areas it was found that the function-based treatment plan significantly reduced V5Gy, V10Gy, and V20Gy compared to the clinical standard-of-care plan [58]. Again, this study did not include follow-up data nor predicted pulmonary function sparing. Functional regions were also discretely segmented from the HP MR images by designating four separate ROIs: well-ventilated, ventilated, hypo-ventilated, and poorly ventilated.

In this chapter, functional avoidance information and models developed in Chapter 4 will be integrated into a treatment planning system for efficient creation and optimization of functional avoidance radiotherapy treatment plans. Overall, four different functional avoidance treatment plans will be compared to the clinical SOC treatment plan to evaluate differences in OAR doses, PTV coverage, and functional lung and airway sparing. Lastly, the cGAN model developed in Chapter 4 will be utilized to compare predicted ventilation damage across all five treatment plans.

5.2 Methods

5.2.1 Patient Characteristics

The data used in this study was collected from 6 lung cancer patients that participated in a University of Wisconsin-Madison IRB-approved prospective study (NCT02843568) who underwent RT for non-small cell lung cancer (NSCLC). Exclusion criteria for this study included prior or future planned surgery for the existing lung cancer, prior thoracic radiotherapy, severe COPD, oxygen dependence, known underlying collagen vascular disease, under 18 years of age, and a Karnofsky score $< 60\%$. Each subject had two 4DCT scans separated by five minutes acquired pre-RT and three months post-RT for a total of four scans. Subjects received a standard fractionation course of 2 Gy x 30 fractions. The patient information for this study cohort is summarized in Table 5.1.

Table 5.1: Study cohort characteristics

Number of Patients	6
Mean age (range)	70 y (64-76)
Gender (male/female)	3/3
Tumor location	
Right Upper	1
Right Middle	1
Right Lower	1
Left Upper	1
Left Lower	1
Mediastinum	1
Mean PTV Volume (range)	327.5 cc (118.7-780.3)
Prescription Dose	60 Gy

All scans were acquired on a Siemens SOMATOM Definition Edge 128 slice CT scanner (Siemens Healthineers AG, Erlangen, Germany) with a 0.6 mm slice thickness, 0.5 second tube rotation time, 0.09 pitch, 120 kVp, and 100 mAs/rotation. The Varian Real-Time Position Management (RPM) (Varian Medical Systems Inc., Palo Alto, CA) system monitored respiration-induced movement and captured respiratory traces used for the 4DCT reconstructions. Melodic breathing instructions were played at a rate of 15 breaths per minute during each image acquisition which has been demonstrated to increase the repeatability between scans and reduce the occurrence of artifacts [141]. The averaged pre-RT 4DCT was used for the purposes of treatment planning.

5.2.2 4DCT-derived Ventilation

This work utilized an image registration algorithm with multi-resolution cubic B-spline parameterization and used the sum of squared tissue volume difference (SSTVD) as the similarity metric as described in Section 2.4.2. The Jacobian-based method described by Shao et al. [73] and Section 2.4.3 was used to generate ventilation maps for each subject at both timepoints (pre- and post-RT). LER-N maps for the two scans at each time point were geometrically averaged to reduce noise within the map introduced by random changes in patient respiratory patterns between scans.

5.2.3 Dose-painting-by-numbers

Dose-painting-by-numbers (DPBN) is a technique for radiotherapy treatment planning where radiation dose is prescribed at the voxel level. It is then the job of the optimizer to recreate the idealized dose distribution as closely as possible while still producing a physically deliverable plan. Almost all previous uses of dose-painting-by-numbers in treatment planning have been for prescribing non-uniform radiation dose distributions to the target

volume. Multiple clinical trials (NCT01168479 [166], NCT04910308[90], NCT01024829 [167]) have investigated the use of dose-painting-by-numbers to escalate dose to tumors based on imaging biomarkers. Prior to creating the treatment plan, imaging of the tumor site is performed and imaging biomarkers are extracted. van Diessen et al. [167] used FDG PET scans to locate regions of high pre-treatment FDG-uptake in non-small cell lung cancer tumors, and redistributed more dose to those specific regions of the tumor. Petit et al. [168] used dose-painting-by-numbers to redistribute tumor dose based on differences in oxygen levels determined from oxygen diffusion-consumption modeling. However, no studies have applied dose-painting-by-numbers to achieve normal/functional lung tissue sparing.

In this chapter, DPBN is utilized in three different ways. First, minimum dose values for voxels within the PTV from the standard-of-care (SOC) plan are utilized as a reference for the functional avoidance plans to encourage similar PTV coverage. Second, maximum dose values for voxels in regions outside of the lungs from the SOC plan are utilized as a reference for the functional avoidance plans to limit doses to OARs. Lastly, maximum dose values for voxels within the lung are derived from one of the predictive models from Chapter 4 in order to reduce dose to regions of the lung that are predicted to decline in ventilation. In summary, all voxels within the lung ROI were assigned individual maximum dose objectives to limit each voxel to a specific dose value based on the predicted ventilation decline due to parenchymal dose, the PTV voxels were assigned individual minimum dose objectives to mimic the same coverage and dose as was achieved in the SOC plan, and voxels outside both the PTV and lungs were assigned dose values to minimize dose to OARs.

5.2.4 Treatment Plan Creation

Five treatment plans were created for each subject, comprising of the clinical SOC treatment plan and four functional avoidance treatment plans. The clinical SOC treatment plans were developed to meet institutional clinical goals and follow standards as set by OAR recommendations from the Radiation Therapy Oncology Group (RTOG) report 0617, and were previously created by a certified medical dosimetrist. Functional avoidance optimized plans were to follow and meet the same clinical goals as the SOC plans. The following sections will explain the methods used to create each of the four functional avoidance treatment plans.

5.2.4.1 Functional Avoidance Treatment Plan 1 (Opt1)

The Opt1 plans used 6 discrete ROIs generated from the polynomial regression model developed in Section 4.2 to avoid functional regions of the lung. For each subject, the pre-RT ventilation map was flattened into a one-dimensional array and an array of equivalent length was created representing a uniform dose (i.e., every value assigned 10 Gy). The two arrays were used as inputs for the model which then produced a voxelized ventilation change (i.e. Jacobian ratio (J_{ratio})) array. This was performed 6 times, where a different uniform dose was used (10-60 Gy, in 10 Gy intervals) to create 6 different arrays corresponding to the J_{ratio} that would occur to each voxel at 6 discrete dose levels. The arrays were reshaped to the original image size and masked to only consider voxels within the lungs. The J_{ratio} maps for each dose-level were thresholded to consider only the voxels that were predicted to have a J_{ratio} less than 0.90. The 6 thresholded J_{ratio} maps were then imported into RayStation (RaySearch Laboratories, Stockholm, Sweden) and the threshold ROI tool was used to create ROIs from the 6 different thresholded J_{ratio} maps. The SOC plan, beamsets, and objective functions for optimization were copied to create

the starting point for the Opt1 plan. From the 6 functional ROIs, 6 objective functions were added to the optimization to influence the optimizer to reduce dose to the functional regions defined. Optimization is not reset for the creation of the Opt1 plan, but rather continued in order to maintain a similar level of PTV coverage and OAR doses. Figure 5.1 shows an example subject with the six ROIs generated that are used for the optimization process.

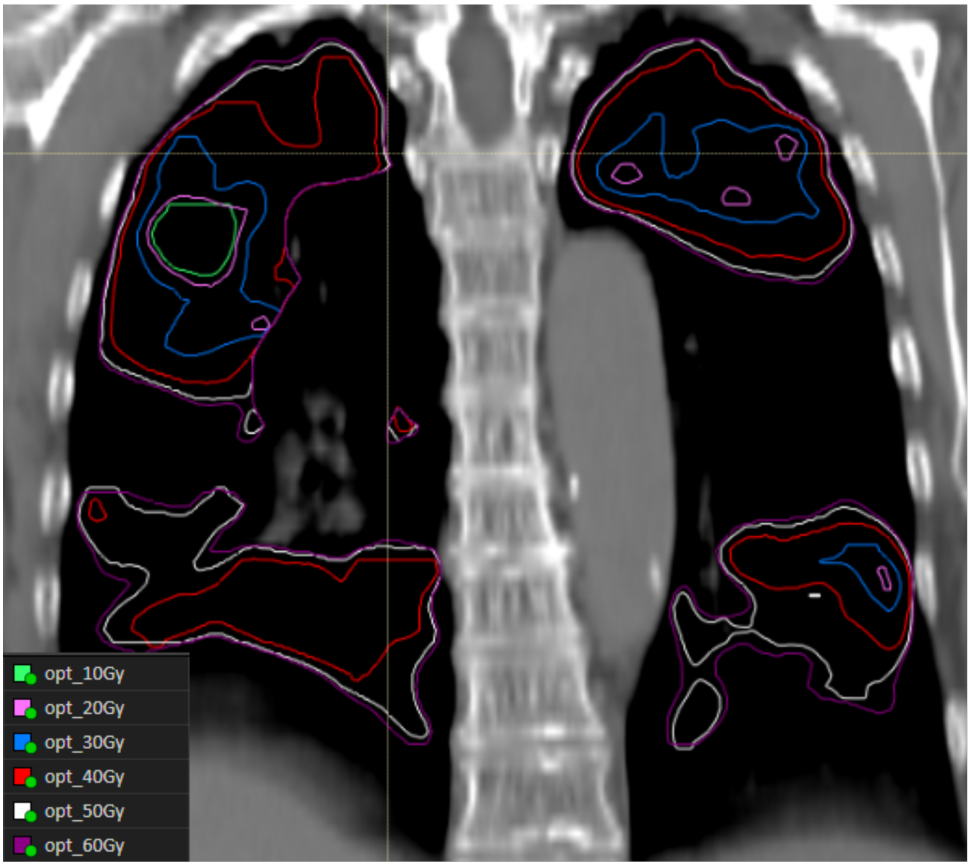


Figure 5.1: Example of the six ROIs created from the polynomial regression model predictions for creating regions of avoidance. Optimization objectives are generated for each ROI to create the Opt1 treatment plan.

5.2.4.2 Functional Avoidance Treatment Plan 2 (Opt2)

The Opt2 plans were generated using a method similar to simulated annealing to determine the maximum dose to each voxel. First, 80% of the maximum dose value from the SOC plan was assigned to each voxel within an ROI defined as the lungs minus a 5 mm expansion of the PTV (PTV+5mm), while the dose within the PTV+5mm ROI was set to the same values as the SOC plan. The 5 mm expansion was to ensure that the treatment planning system optimizer was not given a steep dose gradient off of the target that it was unable to achieve. Second, this initial 'dose distribution' and the pre-RT ventilation map were input to the polynomial regression model from Chapter 4 in the same manner as was performed for the Opt1 plan. The output from the model was the J_{ratio}, and for voxels within the lung where $J_{ratio} < 0.90$, the dose value (in the case of the 1st iteration, 80% of the SOC maximum dose) was assigned to that voxel location in the dose array. In each subsequent iteration, the dose values to voxels with a predicted $J_{ratio} < 0.90$ were decremented by 1 Gy and re-input into the polynomial regression model with the same pre-RT ventilation array. Voxels with a $J_{ratio} < 0.90$ had the new dose value updated to their voxel location in the dose map. This process repeated itself until dose had been decremented down to 1 Gy. The end result was a dose map masked to the lungs, where each voxel value represents the maximum dose that voxel can receive in order to prevent a decline in ventilation of greater than 10% post-RT. For some subjects, there were voxels that still had a predicted $J_{ratio} < 0.90$ with a maximum dose of 1 Gy due to this technique. The procedure to create the reference dose distribution is depicted in Figure 5.2.

This reference dose distribution was imported back into RayStation as a new treatment plan referred to as 'Reference'. The SOC beamsets were copied, and three DPBN optimization objectives were created based on the imported 'Reference' dose map as described previously. Although the SOC beamsets were copied to the Opt2 plan, optimization was

reset in order to develop a dose distribution that relied on the new optimization objectives. Additional optimization objectives, such as maximum dose constraints to spinal cord or esophagus, were added when necessary in order to meet OAR clinical goals. Lastly, the DPBN method as previously described was used to create the voxelized dose objectives for the PTV, lungs, and other OARs referenced from the imported 'Reference' plan.

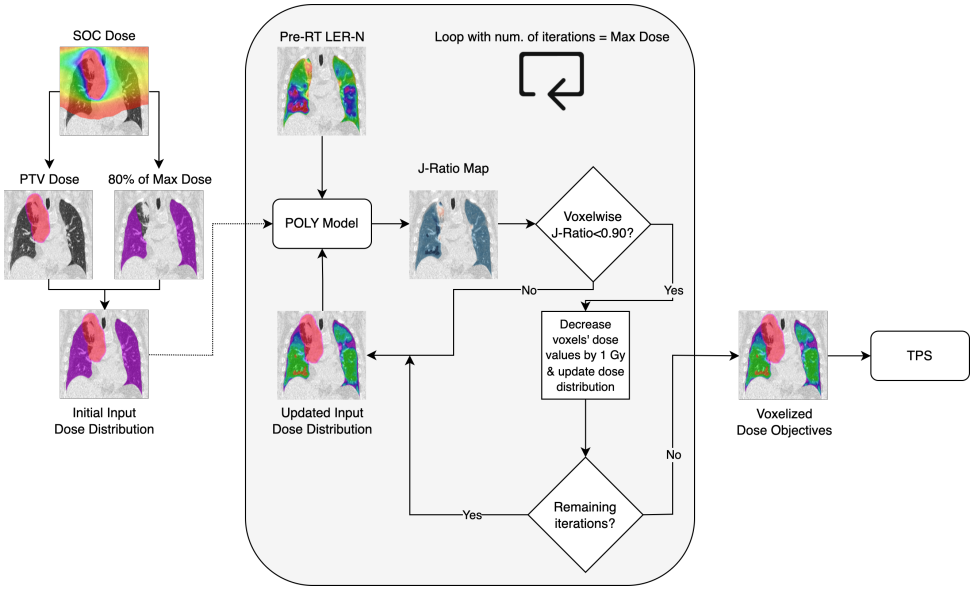


Figure 5.2: Workflow for creating the reference dose distribution for the Opt2 treatment plan. The resultant reference dose distribution can then be imported into the treatment planning system where voxelized dose objectives can be derived from.

5.2.4.3 Functional Avoidance Treatment Plan 3 (Opt3)

The Opt3 plans were created using the cGAN model developed in Chapter 6 to predict the ventilation change post-RT as a result of the delivered dose distribution. Due to the nature of the machine learning model, there was not a direct correspondence between decreasing dose to a voxel and the minimization of the J_{ratio}. Therefore, decreasing the dose and iterating through various dose levels as was done for the Opt2 plan was not effective for developing an optimized plan using the cGAN model. Alternatively, an optimizer, PyGAD[169], was used to generate a voxelized dose map which utilized a loss

function designed to minimize the volume of voxels with a $J_{ratio} < 0.90$. The optimizer was initialized with the SOC dose distribution as a starting point and was run for 500 iterations. Additional optimizer parameters can be found in Table 5.2. The output of the optimizer was a dose map corresponding to the maximum dose values for each voxel within the lungs that minimized the predicted volume of voxels with a $J_{ratio} < 0.90$. Dose values within the PTV+5mm ROI and outside the lungs remained the same from the SOC plan. Similarly to the Opt2 plan, this dose map was imported into the treatment planning system as a new 'Reference' dose distribution, and the original SOC beamsets were again copied. The voxelized dose objectives were generated using the new reference dose map and a new dose distribution was created after restarting optimization using the new optimization objectives. Figure 5.3 shows the workflow for generating the reference dose map using the cGAN model and PyGAD optimizer.

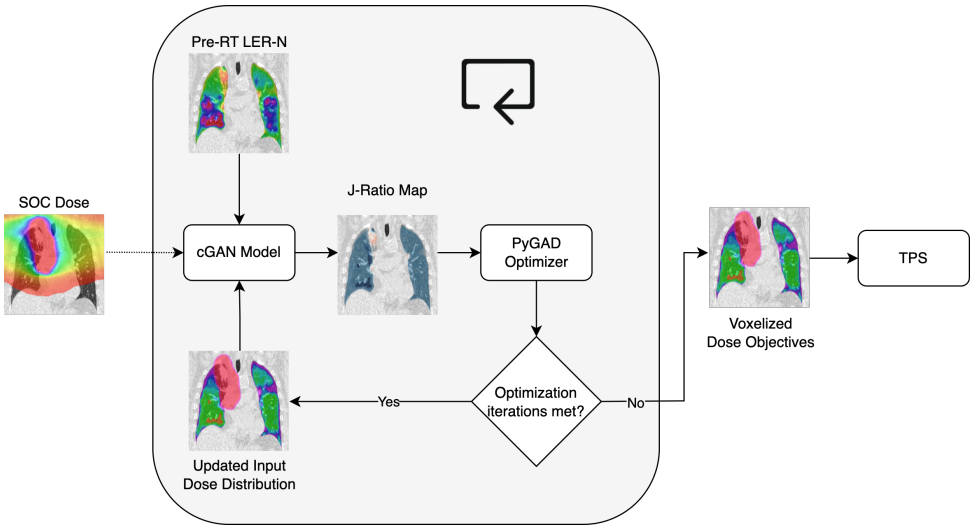


Figure 5.3: Workflow for creating the reference dose distribution for the Opt3 treatment plan. The PyGAD optimizer alters the dose map every iteration in order to minimize the volume of voxels with a $J_{ratio} < 0.90$.

Table 5.2: PyGAD optimizer parameters

Number of Generations	500
Number of Parents Mating	10
Solutions per Population	20
Mutation Type	Random
Mutation Percent	5
Mutation by Replacement	True

5.2.4.4 Functional Avoidance Treatment Plan 4 (Opt4)

The last functional avoidance treatment plan generated, Opt4, used the Opt3 as an initial dose distribution. Airway dose objectives were created using airway segmentations and the results of the toxicity model for bronchial stenosis developed in Chapter 3. Airway segmentations were subtracted from the lung segmentations to ensure that dose objectives relating to those structures did not overlap and cause conflicts in optimization. Five dose objectives were created for 10 Gy intervals between 10-50 Gy inclusive, specifying the specific airway segments and the maximum dose they each should receive to keep the probability of bronchial stenosis under 37%. This cutoff threshold of 37% was selected as it was the threshold at which the logistic regression model had the largest J-index, maximizing both sensitivity and specificity of the model prediction. The Opt3 plan was copied, the airway dose objectives were added, and optimization was continued to spare dose delivered to the airway structures.

5.2.5 Analysis

The functional avoidance and SOC plans were compared using several metrics. First, target coverage for each plan was normalized to the prescription coverage to allow for comparison across the various ROIs. Additional PTV coverage was evaluated by the following metrics: D98% $\geq$ 54 Gy, D99.5% $\geq$ 42 Gy, and D0.03% $\leq$ 66 Gy. To compare functional lung sparing between plans the dose to highly functional lung, defined as voxels with a pre-RT LER-N value greater than 1.20, and maximum airway segment doses were evaluated. The airways were evaluated for the five dose bins (10-50 Gy) as specified in the methods for the development of treatment plan Opt4. Additionally, the dose distributions created for each plan were used as input with the pre-RT ventilation map to generate predicted J_{ratio} maps using the cGAN model developed in Chapter 4. The predicted volume of damaged lung was compared across all plans and subjects. Organs at risk (OAR) dose was assessed for heart, esophagus, and spinal cord, and the dose limits are summarized in Table 5.3.

Table 5.3: Organs at risk dose limits for conventional fractionation radiotherapy.

Organs At Risk	Dose Limits
Heart	V60Gy < 30%
	Mean Dose < 30 Gy
	V45Gy < 60%
	V40Gy < 80%
Esophagus	Mean Dose < 34 Gy
	D0.1cc < 63 Gy
Spinal Cord	D0.1cc < 50 Gy

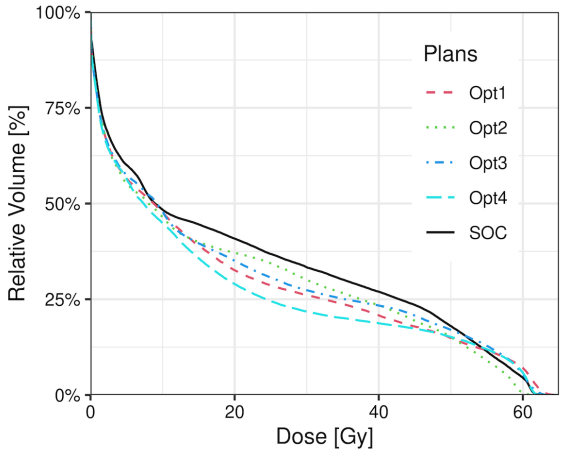
Lastly, plan modulation was assessed by comparing planned monitor units (MU) between each plan. Plan deliverability was not explicitly assessed through phantom measurements.

5.2.6 Statistical Analysis

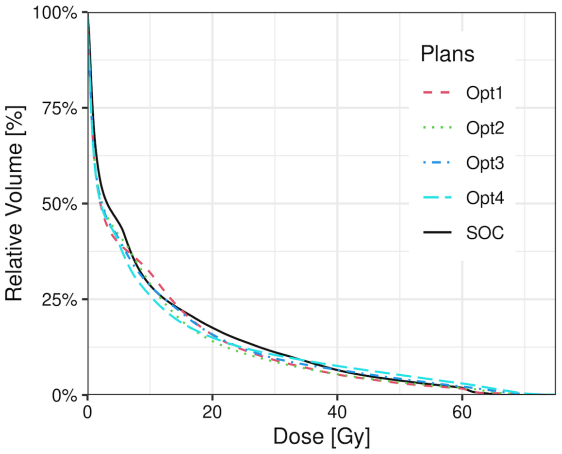
RStudio (version 2022.07.2; Posit Software, Boston, MA) was used for calculation of all statistics. Paired-sample t-tests were conducted to evaluate whether there was a significant difference of the volume of predicted damaged lung between each of the functional avoidance plans and the SOC plan. T-tests were also used to compare the planned MUs between plans. Tukey's range test was used to test for significant differences between the various maximum dose levels delivered to the bronchial segments across all the treatment plans.

5.3 Results

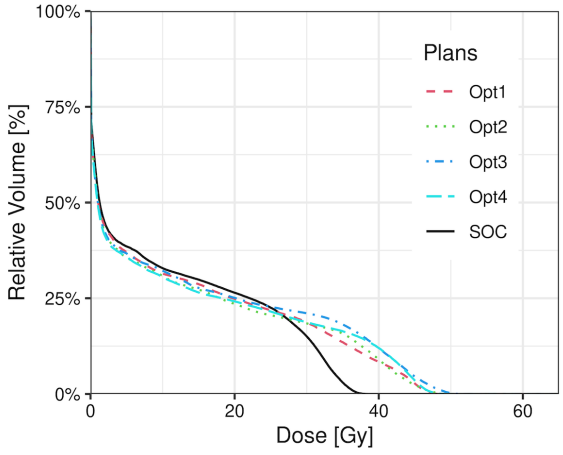
All functional avoidance plans met the dose constraints for OARs as listed in Table 5.3, but failed to meet the PTV goal of D0.03% $\leq$ 66 Gy. The remaining PTV coverage goals were met. The dose-volume histograms (DVH) for each of the 5 treatment plans were averaged over the six subjects and are shown in Figure 5.4. In general, the PTVs for all functional avoidance plans showed greater hotspots compared to the SOC plans, however all functional avoidance plans achieved a reduction in dose to highly functional lung.



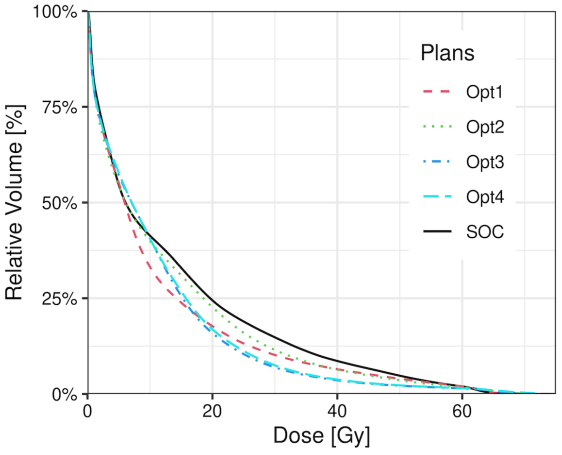
(a) DVH for the esophagus



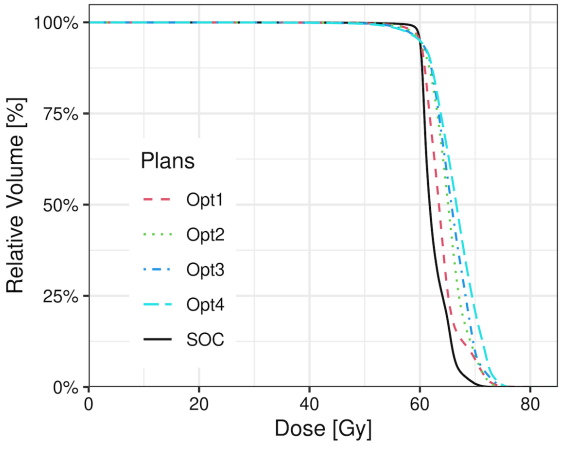
(b) DVH for the heart



(c) DVH for the spinal cord



(d) DVH for high functioning lung



(e) DVH for the PTV

Figure 5.4: Dose-volume histograms (DVH) for the esophagus, heart, spinal cord, high functioning lung, and the planning target volume (PTV). DVHs for each of the five plans (standard-of-care (SOC), Opt1, Opt2, Opt3, Opt4) are shown. Each DVH was averaged across the 6 subjects.

Table 5.4: Treatment planning dose-volume metrics for the standard-of-care (SOC), Opt1, Opt2, Opt3, and Opt4 plans.

		SOC	Opt1		Opt2		Opt3		Opt4	
Structure	Planning metric	Absolute values	Absolute values	Relative Δ from SOC	Absolute values	Relative Δ from SOC	Absolute values	Relative Δ from SOC	Absolute values	Relative Δ from SOC
Cord	D0.01cc,Gy	34.3±5.1	38.8±9.4	4.5±4.9	40.7±10.4	6.4±6.8	43.6±8.1	9.4±3.3	43.4±7.4	9.1±4.9
Esophagus	D0.01cc,Gy	59.5±4.0	54.6±11.2	-4.9±8.1	57.4±4.2	-2.0±1.3	54.2±11.8	-5.3±7.9	53.5±13.2	-6.0±9.3
	Davg,Gy	20.9±8.8	18.2±13.0	-2.7±5.4	18.8±10.2	-2.1±3.7	19.0±12.2	-1.9±3.9	16.9±12.2	-4.0±4.2
Heart	V45Gy,%	4.9±6.4	4.1±6.0	-0.8±2.2	4.3±6.8	-0.7±1.9	5.4±7.1	0.5±2.1	6.4±9.2	1.5±3.9
	Davg,Gy	10.1±8.4	9.3±8.4	-0.8±2.1	9.1±9.0	-1.0±1.3	9.6±9.3	-0.5±1.3	9.6±10.5	-0.5±2.9
Lungs	V20Gy,%	24.0±9.5	17.0±7.7	-6.9±4.7	22.1±11.4	-1.8±5.1	15.5±7.7	-8.5±3.9	16.3±7.6	-7.7±3.8
	V5Gy,%	53.5±21.2	53.9±24.7	0.4±6.0	53.5±27.2	-0.1±8.8	55.7±29.2	2.3±11.3	55.8±27.3	2.4±11.0
	MLD,Gy	10.4±4.9	8.6±4.2	-1.8±1.1	9.5±5.6	-0.9±0.9	8.2±4.6	-2.2±1.0	8.5±4.4	-1.9±1.5
Functional Lung	fMLD,Gy	13.3±4.2	11.2±3.6	-2.2±1.5	12.2±4.9	-1.1±1.1	10.8±4.0	-2.6±1.1	11.0±3.8	-2.3±1.5
	fV10Gy,%	41.9±16.0	33.8±11.7	-8.0±7.0	40.4±20.1	-1.5±5.8	41.1±17.7	-0.7±6.7	41.2±15.9	-0.7±8.6
	fV20Gy,%	25.4±9.1	18.6±7.1	-6.8±5.0	23.4±11.7	-1.9±4.9	16.8±8.3	-8.5±3.7	17.8±8.0	-7.5±4.0
	fV30Gy,%	15.8±6.1	11.2±4.9	-4.7±3.0	12.5±6.9	-3.3±3.9	8.1±4.6	-7.8±2.6	8.6±4.7	-7.2±2.7
	fV40Gy,%	9.7±5.3	7.5±3.6	-2.2±2.1	7.5±±4.4	-2.2±2.4	4.7±±2.9	-5.0±2.6	4.9±3.0	-4.7±2.4
	fV50Gy,%	5.9±±3.5	4.9±2.8	-1.0±1.0	4.7±3.0	-1.3±0.9	3.2±2.0	-2.7±1.6	3.3±2.1	-2.6±1.6
	fV60Gy,%	3.1±1.9	2.7±2.0	-0.3±0.4	2.7±2.0	-0.4±0.4	2.2±1.5	-0.9±0.8	2.3±1.5	-0.8±0.8
Airways	10GyMax,Gy	58.2±2.6	52.5±8.3	-5.7±7.6	58.4±5.9	0.2±4.7	48.2±10.7	-10.0±8.6	**30.7±7.4**	-27.4±7.4
	20GyMax,Gy	49.8±5.7	50.1±11.4	0.3±11.4	49.6±6.9	-0.3±3.9	36.6±8.3	-13.2±6.2	**30.2±6.1**	-19.7±4.3
	30GyMax,Gy	39.8±14.5	37.3±19.4	-2.5±10.1	36.0±14.4	-3.8±10.5	28.8±14.3	-10.9±6.3	26.2±12.1	-13.6±3.8
	40GyMax,Gy	32.1±16.9	27.4±15.4	-4.8±7.6	32.2±15.4	0.1±5.0	25.8±8.3	-6.4±9.8	23.5±6.7	-8.6±13.5
	50GyMax,Gy	3.3±3.0	4.0±4.4	0.5±1.2	3.5±3.8	0.1±2.2	3.6±3.7	0.2±0.6	3.9±4.1	0.4±0.9

Abbreviations: fV20Gy = volume of functional lung (pre-RT LER-N≥1.2) ≥ 20 Gy; MLD = mean lung dose; fMLD = functional mean lung dose; 20GyMax = maximum dose to 20 Gy max dose airway segments.
All results are presented as the mean ± standard deviation. Absolute values in **bold** font were significantly ($p<0.05$) different compared to the SOC plan.

Table 5.5 shows the predicted volumes of damaged lung (J_{ratio} <0.90) for each subject and plan. On average across all 6 subjects, the Opt4 plan significantly (p=0.04) reduced the predicted volume of damaged lung by 1.5%.

Table 5.5: Predicted volumes of damaged lung for each subject and treatment plan. P-values are from paired-sample t-tests between each of the functional avoidance plans and the standard-of-care (SOC) plan.

	Predicted Damaged Lung Volume (cc)				
Subject No.	SOC	Opt1	Opt2	Opt3	Opt4
1	2537	2537	2529	2528	2540
2	3669	3642	3659	3614	3615
3	2828	2788	2802	2784	2782
4	1709	1706	1712	1708	1707
5	2479	2444	2497	2441	2437
6	2722	2671	2678	2638	2633
Average	2658	2632	2647	2619	2619
		***p*=0.02**	p=0.27	***p*=0.03**	***p*=0.04**

Figure 5.5 shows the boxplots for each treatment plan, separated by the maximum airway dose bins. There was a significant (p<<0.001) difference in the average dose for the 10 Gy max dose bin between the Opt4 treatment plan and all other plans. Additionally, there was a significant (p<<0.001) difference in the average dose for the 20 Gy max dose bin between the Opt4 treatment plan and the SOC, Opt1, and Opt2 plans. For the 30 Gy and 40 Gy dose bins, the Opt4 plan showed dose reductions compared to the other treatment plans, but these differences were insignificant. Lastly, the 50 Gy dose bin showed no differences across all plans, however the airways within this dose bin were well below the maximum dose of 50 Gy.

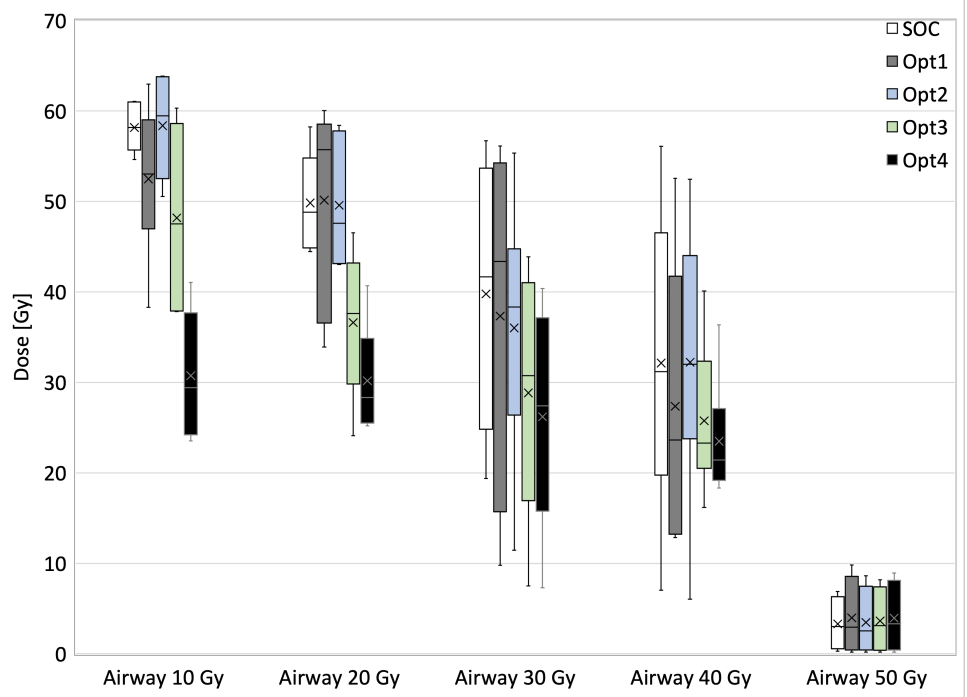


Figure 5.5: Boxplots for each of the five treatment plans and for each of the five maximum airway dose bins (10-50 Gy). There was a significant difference in the average dose for the 10 Gy max dose bin between the Opt4 treatment plan and all other plans. Additionally, there was a significant difference in the average dose for the 20 Gy max dose bin between the Opt4 treatment plan and the SOC, Opt1, and Opt2 plans.

Figure 5.6 summarizes the planned MUs for each plan and subject. The Opt1 (p=0.04) and Opt4 (p=0.008) plans showed significantly greater planned MUs compared to the SOC plan.

Table 5.6: Planned monitor units (MUs) for each plan and subject. There was a significant increase in MUs between the standard-of-care (SOC) plan and the Opt1 (p=0.04) and Opt4 (p=0.008) plans.

	Planned MU				
Subject No.	SOC	Opt1	Opt2	Opt3	Opt4
1	457	984	505	661	993
2	373	699	624	634	705
3	352	584	454	532	648
4	538	748	754	769	854
5	765	765	759	844	905
6	674	868	752	840	864
Average	527	775	641	713	828
		$\boldsymbol{p=0.04}$	$p=0.62$	$p=0.17$	$\boldsymbol{p=0.008}$

Figure 5.6 shows one subject's various treatment plans and the pre-RT LER-N map. The SOC plan for this subject resulted in greater dose deposited to the anterior right lung, near regions of high function as designated in the pre-RT LER-N map. The Opt3 and Opt4 plans reduced dose to the anterior right lung as well as portions of the posterior right lung. The left main bronchus, in addition to other airway segments not seen in the shown axial slice, also showed a dose reduction in the Opt3 and Opt4 plans, compared to the SOC plan.

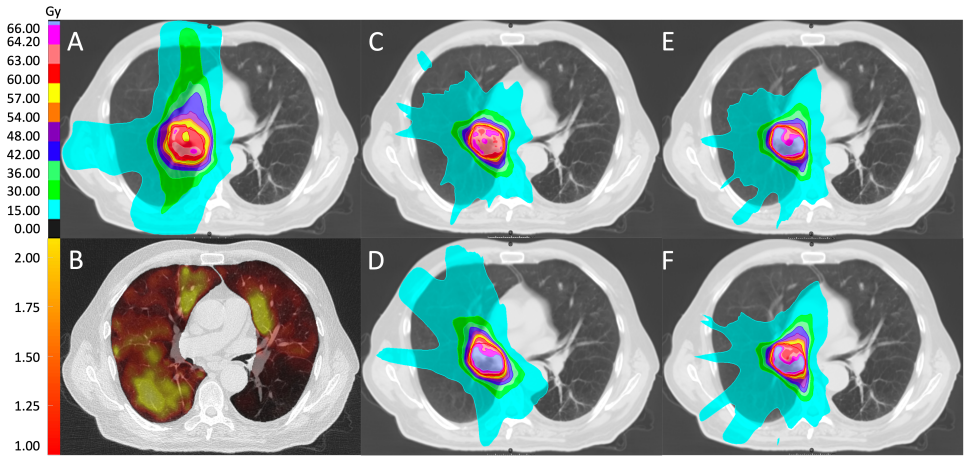


Figure 5.6: One subject showing the five different treatment plans created and the pre-RT LER-N map. A) Standard-of-care dose distribution, B) pre-RT LER-N map, C) Opt1 dose distribution, D) Opt2 dose distribution, E) Opt3 dose distribution, and F) Opt4 dose distribution.

5.4 Discussion

This work demonstrates that the use of 4DCT-derived ventilation information across different treatment planning techniques resulted in clinically-acceptable plans that achieved superior functional lung sparing compared to the clinical SOC plans. While other studies [8, 10, 11, 165, 170–172] have investigated the efficacy and feasibility of functional avoidance treatment plans, no studies to date have used predictive modeling or the DPBN approach used in this work. Additionally, this is the first work to consider airway dose reduction for bronchial stenosis. Vicente et al. [78] investigated functionally weighted airway sparing (FWAS) in combination with regional functional avoidance objectives, however the main endpoint for airway toxicity was complete airway collapse and it was assumed that an airway was considered collapsed if it had a predicted probability of collapse $\geq$5%.

Overall, each of the functional avoidance plans generated were able to meet clinical OAR goals and maintain PTV coverage while reducing dose to highly functional lung. It can be seen in the DVHs in Figure 5.4 that the doses to the esophagus, heart, and spinal cord are similar or better across all functional avoidance plans compared to the SOC plan. The maximum spinal cord dose increased for all functional avoidance plans compared to the SOC treatment plan, however all doses remained within recommended dose constraints. Esophagus maximum and average dose was slightly reduced across all functional avoidance treatment plans compared to the SOC plan. Depending on risk tolerance determined by the physician for the patient, it is possible additional dose could be given to the esophagus in the functional avoidance plans in order to spare further functional lung. The standard dose constraints for lung (V20Gy and MLD) were reduced for all functional avoidance treatment plans, while V5Gy slightly increased. Lastly, while all functional plans reduced the volume of functional lung (pre-RT LER-N $\geq$1.2) that received 10-60 Gy, these differences were not significant compared to the SOC treatment plan. This is further seen in Table 5.5 where there is a significant reduction in the predicted volume of damaged lung in the Opt4 plan compared to the SOC plan, however this reduction was only an average of 39 cc. These results suggest that further optimization must be done on the functional avoidance plans in order to have a larger impact on functional lung sparing.

Airway doses were significantly reduced between the SOC and Opt4 treatment plans. For airway segments that were in the 10 Gy max dose bin, the average maximum dose reduced significantly ($p<<0.001$) from 58.2$\pm$2.6 Gy in the SOC treatment plan to 30.7$\pm$7.4 Gy in the Opt4 treatment plan. Additionally, for airway segments in the 20 Gy max dose bin, the average maximum dose reduced significantly ($p<<0.001$) from 49.8$\pm$5.7 Gy to 30.2$\pm$6.1 Gy in the SOC and Opt4 treatment plans, respectively. There was little change across all five treatment plans in the 50 Gy max dose bin, however all plans were well

below 50 Gy resulting in little to no weight given to these structures during treatment plan optimization. Similarly, the 40 Gy max dose bin airways received doses less than 40 Gy on average across all five treatment plans. Lastly, while the difference in airway doses for the 30 Gy max dose bin was not significant between the SOC and Opt4 treatment plans, it should be noted that the Opt4 plan was able to successfully reduce dose to below 30 Gy on average to those airway structures. Overall, there was a significant ($p<<0.001$) reduction in the predicted probability of bronchial stenosis across all airways of 1.2% between the SOC and Opt4 treatment plans.

There were multiple limitations of the work presented in this chapter. First, the number of subjects (6) analyzed was small and may have contributed to finding few significant changes between the functional avoidance plans and the SOC treatment plan. For example, conducting power analysis for the fV20Gy metric with a significance level of 0.05 and power of 80%, at least 22 subjects would be required to show significant differences between the SOC and Opt4 plans. A second limitation of this work was the inability to incorporate the machine learning model into the treatment planning system in order to use the predictions from the model to further optimize the functional avoidance treatment plan. Currently, dose must be exported to another machine in order to generate the predicted lung ventilation change, create a new reference dose distribution, re-import this data back into the treatment planning system, and re-optimize. This is a time consuming process that makes this technique incompatible in a clinical workflow. Future work should consider a more seamless integration between the treatment planning system, the predictive models, and the optimization of the reference dose. Another limitation was treatment planning expertise as the SOC and Opt1 plans were created by professional dosimetrists, however the Opt2-Opt4 plans were not. Although the Opt2-Opt4 plans were clinically acceptable plans, there was most likely improvements that could be made to them to increase functional lung sparing, reduce OAR doses, and improve PTV coverage.

Lastly, while this work showed some improvements in the Opt4 functional lung sparing compared to the other functional avoidance treatment plans there is still not a consistent methodology for pushing dose out of regions susceptible to ventilation damage. As with other OARs, further dose sparing can be achieved through additional re-optimization and manipulation of dose objectives. This process can be time consuming and is not necessarily practical in a clinical environment. Improvements should be made to minimize the amount of time needed to maximize functional lung sparing. Approaches such as multi criteria optimization (MCO)[173] and machine learning [174] are currently being investigated as tools to dynamically create dose distributions that can be adjusted in real time, and should be considered for the development of functional avoidance treatment plans.

5.5 Chapter Summary

In this chapter multiple methods for creating pulmonary functional avoidance treatment plans for ventilation sparing were generated and compared. On average it was found that the DPBN method combined with airway sparing (Opt4) resulted in the greatest reduction in predicted damaged lung volume. Additionally it was found that the airway sparing plan (Opt4) was able to significantly reduce the average maximum dose to airway structures in the 10 and 20 Gy max dose bins, showing potential for this method to be used clinically to reduce potential for airway toxicities. In general all functional avoidance plans increased the plan complexity as seen from the increase in the planned MUs per fraction. Overall this work showed the efficacy of new voxelwise dose objectives for achieving clinically-acceptable ventilation sparing plans. Future work should focus on integration of the predictive models into the treatment planning system to reduce the need for data transfer between systems, and minimize the potential for errors that may arise as a result of transferring data. Furthermore, by integrating the predictive models

into the treatment planning systems it will be easier for adjustments to be made to functional avoidance plans in real-time.

Chapter 6

Conclusions & Future Work

6.1 Summary and Limitations

In this work, techniques and methods to improve the efficacy of pulmonary functional avoidance radiotherapy were investigated. Changes in airway structures were analyzed before and after radiotherapy to quantify the relationship between radiation-induced bronchial stenosis and ventilation changes downstream. Airway remodeling was observed in both a novel swine model and in a human subject cohort. Additionally, multiple predictive models were developed, trained, and validated to predict post-RT ventilation changes. The first model was a polynomial regression model that used the LER-N method to derive ventilation maps instead of the previously utilized LER-2 method. Multiple deep learning models were created to predict post-RT ventilation change using volumetric images as input to preserve global context across the entire lung volume. Finally, the results of the airway toxicity analysis and ventilation change predictive models were used to develop a new method in creating functional avoidance treatment plans. This new approach

was compared with previously used methods and the standard-of-care treatment plans to investigate the potential for lung ventilation sparing.

6.1.1 Investigation of radiation-induced airway remodeling and regional ventilation changes

6.1.1.1 Radiation-induced airway changes and downstream ventilation decline in a swine model

Section 3.2 performed a retrospective analysis of radiation-induced changes in airways in a novel miniature swine model to quantify the relationships between airway dose and airway structural changes, as well as regional changes in ventilation function. A strong linear relationship was found between radiation dose and the change in the luminal area of a bronchial segment. Furthermore, it was found that the dose level causing Grade 2 or worse bronchial stenosis was approximately 25 Gy, which is half the recommended maximum dose to prevent non-pneumonitis toxicities recommended by RTOG 0813 and validated by Manyam et al. [32]. It was also observed that there was a greater decline in ventilation function in unirradiated regions that were supplied by irradiated airways, compared to unirradiated regions supplied by unirradiated airways. This supports the hypothesis that there is "downstream" ventilation damage caused by irradiating the supplying airways.

Part of the limitations in this study was due to the small sample size of swine studied (n=3). Significant changes were unable to be observed for the downstream ventilation changes. A further limitation was that while the swine were mechanically ventilated and were imaged and treated within a relatively controlled environment, there were still inter-subject variabilities that existed that prevent strong conclusions to be drawn from the results. Additionally for this cohort of swine, the radiation treatment plan that

was delivered was not specifically planned prospectively to study downstream effects of irradiating the airways. The delivered dose distribution was placed relatively inferiorly in the left lung, reducing the volume of lung to analyze that was supplied by irradiated airways that also did not receive substantial levels of dose. Lastly, while swine and human lungs have some anatomical similarities, the airways are quite different. Swine airways exhibit a monopodial branching pattern, meaning the left and right main stem bronchi extend the length of the lungs, with smaller branches and generations branching off asymmetrically throughout the lung. Human airways have a bipodial branching pattern with each airway segment containing a bifurcation point, splitting into two separate airways of the same generation. This key difference in airway structures between swine and humans may cause dissimilar responses to radiation dose and regional ventilation changes as a result of bronchial stenosis. These limitations described in the swine study were a key motivating factor for extending the airway analysis to a human subject cohort.

6.1.1.2 A Risk Model for Radiation-Induced Bronchial Stenosis and Indirect Ventilation Damage

In Section 3.3 additional analysis was performed retrospectively on human subject imaging data to further examine the relationship between airway structure changes and downstream ventilation change. A similar but weaker dose-response relationship was observed between radiation dose and bronchial stenosis. A commercial virtual bronchoscopy software was used to calculate metrics for each airway segment out to the 11th generation of airways, allowing for comparison of resistance changes between the pre-RT and post-RT images. Through this analysis it was found there was a significant relationship between the change in resistance of an airway path and the ventilation change in the region those airways supplied. From these results, a toxicity model was developed to predict the probability of an individual bronchial segment to experience stenosis based on the delivered

maximum dose and the initial wall thickness. This toxicity model can be utilized in treatment planning by creating ROIs for various maximum dose thresholds, which can then be used as dose objectives when planning to avoid exceeding those dose tolerances to each airway segment. By creating a treatment plan which meets these dose objectives, the probability of bronchial stenosis and the resultant loss in ventilation will be minimized.

Multiple limitation existed in the analysis of human airways. First, due to the time intensive nature of segmenting airway trees out to the resolution of the image, only 10 human subjects were included for analysis. While the commercial bronchoscopy software used was able to automatically segment most airway trees, it failed to provide accurate segmentations in regions of advanced disease and dense, fibrotic tissue. As the human subjects in this study are patients receiving radiotherapy for lung cancer, both of these issues were highly prevalent. There is a need for improved tools and segmentation techniques that are able to overcome the issue of accurate automatic segmentation in these particular regions. Another limitation was that all the subjects analyzed received standard fractionation radiotherapy across 30 fractions. Because these treatments are typically reserved for diseases near the mediastinum that can not tolerate the higher biological doses associated with SBRT fractionated treatments, there may be a slight bias in the model. The bias that exists shows a weak correlation between maximum dose and airway caliber; this is due to the fact that larger airways exist closer to the mediastinum and therefore they tend to receive higher doses during standard fractionation treatments. This factor should be kept in mind when applying this model for functional avoidance purposes as well as future studies to validate the findings. Lastly, simplifying assumptions were made for airflow within the airways in order to calculate resistance. These assumptions should again be considered when making clinical decisions using these toxicities models, and should be investigated further using more sophisticated techniques.

6.1.2 Modeling normal lung tissue response to radiation dose

6.1.2.1 Modeling the impact of out-of-phase ventilation on normal lung tissue response to radiation dose

The polynomial regression model presented in Section 4.2 improved in gamma pass rates and accuracy compared to a previously developed model by Patton [123]. Specifically in regions considered out-of-phase, the LER-N model was significantly more accurate at predicting post-RT ventilation compared to the LER-2 model. One weakness of the LER-N model was that it decreased in sensitivity to regions of ventilation decline. In general, the predicted post-RT ventilation values tended to under predict ventilation damage. It was found that for highly functional voxels that received 20 Gy or more and were predicted to decline in ventilation post-RT, 66% had a ground truth post-RT ventilation value smaller than the predicted value. This indicates that greater lung damage was occurring than the model was able to predict.

One of the largest limitations of the LER-N model was that the values the model was trained on were flattened into one-dimensional arrays and lost all global context and volumetric information. It should be apparent that lung ventilation changes are dependent on more than the dose delivered to one particular voxel and may be the reason for the model performing poorly in identifying regions of ventilation decline. This also means that pulmonary vasculature and airway damage is unaccounted for within this model. An additional limitation of the model compared to the LER-2 model is that for patients who do not experience long periods of out-of-phase breathing, the added benefit of the LER-N method may not substantially impact predicted ventilation decline. These two majors issues with the polynomial regression model provide motivation to develop more complex models that are able to overcome these limitations.

6.1.2.2 Predicting Pulmonary Ventilation Damage After Radiation Therapy for Non-Small Cell Lung Cancer Using Deep Learning

Section 4.3 presented multiple deep learning-based models for predicting radiation-induced pulmonary ventilation changes. Three different models were investigated including the Seg3DNet, ResNet3D, and a conditional generative adversarial network (cGAN). Each model was validated using 8-fold cross-validation and compared to the previously developed polynomial regression model. All three deep-learning -based models were able to improve in sensitivity compared to the POLY model, however the cGAN significantly outperformed all other models. Compared to the POLY model, the cGAN had a three-fold improvement in the average Youden's index (TPR+TNR-1) and over a five-fold improvement in average true positive rate. One of the consequences in the asymmetrical loss function designed to heavily penalize under-prediction of ventilation damage was that the cGAN model then over-predicted regions of ventilation decline. In general the cGAN model had a higher rate of false positives, predicting regions of ventilation decline that did not exist within the ground truth images.

A limitation of this work was due to the limited availability of data to use for training and validation. This was partly addressed by pre-processing techniques that were employed as well as using cross-validation, however the size of the dataset used was relatively small compared to typical dataset sizes used for deep learning tasks. One potential way to address this issue is to utilize open-source datasets to pre-train models to be used for transfer learning. Most open-source datasets containing 4DCTs rarely contain radiotherapy data such as dose distributions, and none to date outside of the data collected in this clinical trial provide 4DCTs that have been collected post-RT. Additionally, utilizing datasets from other institutions will increase the variety of the training data and make for a more generalizable model. Another limitation of this work is lack of explainability in the models' ability to predict post-RT ventilation change. There has been recent work

on developing explainable GAN models in an attempt to understand the decision making process and reasoning for predictions made by deep learning models. Wu et al. [175] has investigated creating an explainable GAN network for CT angiography images, which is important for clinical translation of these models.

6.1.3 Integrating Functional Avoidance Information Into Treatment Planning

Lastly, Section 5.1 compared multiple methods of creating functional avoidance treatment plans to a clinical standard-of-care plan. Four different functional avoidance methods were tested including an ROI-based approach using the polynomial regression model, a DBPN method using the polynomial regression model, a DBPN method using the cGAN model, and a method that also included airway avoidance structures. The method using the DPBN approach with the cGAN model and airway avoidance structures resulted in the greatest reduction in predicted damaged lung volume. This method also significantly reduced dose to airway structures to minimize the probability of stenosis. However, a trend seen across all functional avoidance plans was a significant increase in the planned MUs compared to the standard-of-care plan, indicating greater plan complexity and modulation.

A limitation of this work was the lack of integration of the predictive models into the treatment planning system. While the workflow is possible without this integration, it makes the entire process more cumbersome and time consuming, making it less feasible in a clinical environment. Additionally, a more seamless integration with the treatment planning system would allow for better "on the fly" adaptation to treatment plans based on the predicted ventilation change for a given dose distribution during the optimization process. Another limitation of this work was again a small sample size of subjects

analyzed. Power analysis suggested at least 22 subjects would be required to show significant differences between plans. However, the preliminary work shown here suggests the methodologies proposed may be feasible in producing clinically acceptable plans that better spare lung ventilation compared to standard-of-care and currently utilized functional avoidance treatment planning methods.

6.2 Future Work

While the methods, models, and analysis presented in this work improve the efficacy in which functional avoidance information can be utilized for lung cancer radiotherapy, there is additional work to be done to validate results and further improve clinical integration.

First, additional swine studies should be performed in order to further validate the dose-response relationship as well as the pathology of radiation-induced bronchial stenosis. While the results presented suggested a strong, linear dose-response relationship with bronchial stenosis, the mechanisms are not well understood. Hasegawa et al. [138] suggested that airway-wall thickening encroaches on the inner lumen of the airway, however it was seen that both airway wall thickness and airway caliber decreased as a function of radiation dose. Additionally, the pathology of downstream ventilation decline should be studied to understand the inflammatory response that occurs as a result of irradiating the airways. Performing these additional experiments within the swine models allows for the pathology studies and provides a controlled environment for analyzing the airway responses further. Within the human subject cohort, additional segmentations should be made across a larger group of subjects to further validate the toxicities models developed. Additionally, these toxicities models should be trained and tested using a greater variety of dose distributions; both in location and fractionation.

Future work in regards to post-RT ventilation modeling should focus on the creation of explainable models. Explainable models will provide a better basis for decision making when using the deep learning models in the treatment planning process for functional avoidance plans. An additional helpful change to the cGAN model would be to develop a cycle GAN model, which is a specific type of GAN that learns to map images from domain A to domain B and vice versa, while also ensuring that the mapping is bijective and invertible. To achieve this, a cycle GAN consists of two GANs, one for each domain, and the cycle consistency loss is calculated by comparing the original image and the image reconstructed after going through both generators. This loss encourages the generators to learn a more meaningful mapping between the two domains, and also helps to prevent mode collapse. Applying the cycle GAN to the ventilation model could potentially allow for the mapping and prediction of an "ideal" dose distribution based on a desired ventilation map.

Lastly, future work to improve the integration of functional information into the treatment planning process should focus on the use of the predictive models directly within the treatment planning system to avoid multiple time consuming data transfers. As stated before, this would also aid in real time decision making by being able to view the dose distribution and the predicted ventilation change. Additionally, expanding this work to include perfusion-based metrics will further enhance the ability to create treatment plans that consider all facets of lung function.

6.3 Conclusion

This work presents multiple improvements to pulmonary functional avoidance radiotherapy through the development of an airway toxicity model and subsequent analysis of downstream ventilation changes, a deep-learning based predictive dose-response model

for regional ventilation change, and new methods for the integration of functional information into the treatment planning process. Future work is needed to further validate the results presented, however this work improved upon previous methods and models that are currently used in a clinical trial and should be considered for use in future trials.

Appendix A

Details Regarding the use of Wisconsin Miniature Swine in Airway Results

A.1 Methods Regarding Care and Use of Wisconsin Miniature Swine

A.1.1 Indwelling Catheter Placement

Dantrolene was administered prophylactically (5.9 mg/kg) in a small amount of feed to prevent against malignant hyperthermia. Anesthesia was administered using an injectable Telazol/Xylazine cocktail (3 mg/kg of Telazol and 1.5 mg/kg of Xylazine). Gas anesthesia was then administered via a nose cone (isoflurane at 2%). The subjects' necks were shaved and scrubbed bilaterally. Surgical depth of anesthesia was confirmed by testing toe pinch response and palpebral reflex. An indwelling central venous catheter was placed

percutaneously in the vena cava (placement confirmed by fluoroscopy) and secured in place with suture. The catheter was sutured to the skin twice using finger-trap suture technique as well as using 2 butterfly clamps. The subject was outfitted with a pocketed catheter jacked, placed over a spandex shirt.

Prior to catheter placement 20 mL of blood was collected for serum and plasma samples. A 30 mL 0.9% saline flush was administered and anesthesia was discontinued. The subject was moved to the recovery pen for 2.5 hours before being returned to housing and given a feed ration.

A.1.2 Imaging and fraction delivery sessions

Dantrolene was administered prophylactically (5.9 mg/kg) in a small amount of feed to prevent against malignant hyperthermia. Anesthesia was administered using an injectable Telazol/Xylazine cocktail (3 mg/kg of Telazol and 1.5 mg/kg of Xylazine). Gas anesthesia was then administered via a nose cone (isoflurane at 2%). The subject was intubated and placed in a transport cart to be transported to the CT suite.

Once in the CT suite, Propofol was administered again (2.4 mg/kg) and Telazol (1.2 mg/kg) was administered in the left hind leg. The subject was mechanically ventilated at 15 breaths per minute and 1000 mL/breath. Mechanical ventilation of subject was periodically adjusted to appropriately maintain SPO2 and ETCO2 while allowing for desired image acquisition and irradiation. For all 4DCTs, ventilation was reduced to 8 breaths per minute and for all contrast enhanced CTs the subject was placed in an inspiration breath hold with a 1000 mL tidal volume. Two contrast-enhanced scans were performed. Each contrast scan involved injection of 80 mL of Omnipaque 300 followed by 50 mL of saline administered intravenously.

Once imaging was complete, ventilation, gas anesthesia, and saline administration were discontinued while the subject was transported from the CT suite to the MRI-Guided LINAC vault. Once there, ventilation was resumed at 750 mL/breath and 8 breaths per minute with an inspiratory: expiratory ratio of 2:1. Irradiation treatment of the lung was performed to deliver 12 Gy to the left lobe targeted at a vessel bifurcation.

Once irradiation was complete, anesthesia was discontinued and the subjects were weaned of mechanical ventilation. The subjects were transported to housing where they were extubated.

A.1.3 Animal Care During Study

All Wisconsin Miniature Swine were housed onsite in facilities managed by the Biomedical Research Model Services group at the University of Wisconsin. Facilities undergo frequent inspection by the university Animal Care and Use Committee (ACUC) to ensure ethical treatment of animal subjects.

Bibliography

[1] Rebecca L. Siegel, Kimberly D. Miller, Nikita Sandeep Wagle, and Ahmedin Jemal. Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, 73(1):17–48, 2023. doi: https://doi.org/10.3322/caac.21763. URL https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.3322/caac.21763.

[2] Kavitha M Prezzano, Sung Jun Ma, Gregory M Hermann, Charlotte I Rivers, Jorge A Gomez-Suescun, and Anurag K Singh. Stereotactic body radiation therapy for non-small cell lung cancer: A review. *World Journal of Clinical Oncology*, 10 (1):14–27, jan 2019. ISSN 2218-4333. doi: 10.5306/wjco.v10.i1.14. URL https://www.wjgnet.com/2218-4333/full/v10/i1/14.htm.

[3] Marisol Arroyo-Hernández, Federico Maldonado, Francisco Lozano-Ruiz, Wendy Muñoz-Montaño, Mónica Nuñez-Baez, and Oscar Arrieta. Radiation-induced lung injury: current evidence. *BMC Pulmonary Medicine*, 21(1):1–12, 2021. ISSN 14712466. doi: 10.1186/s12890-020-01376-4. URL https://doi.org/10.1186/s12890-020-01376-4.

[4] L. B. Marks, X. Yu, Z. Vujaskovic, W. Small Jr, R. Folz, and M. S. Anscher. Radiation-induced lung injury. *Seminars in Radiation Oncology*, 13(3):333 – 345, 2003. ISSN 1053-4296. doi: https://doi.org/10.1016/S1053-4296(03)00034-1. URL http://www.sciencedirect.com/science/article/pii/S1053429603000341.

[5] John Bayouth. Partial Distributions of Radiation Induced Pulmonary Changes Can Be Modeled and Utilized for Image-Guided Planning and Therapy to Improve Pulmonary Function Preservation and Hence the Therapeutic Ratio for Lung Cancer. *Medical Physics*, 43(6):3371, 2016.

[6] Y Vinogradskiy, T Waxweiler, Q Diot, R Castillo, T Guerrero, E Castillo, B Kavanagh, L Schubert, and M Miften. Su-c-bra-06: Developing clinical and quantitative guidelines for a 4dct-ventilation functional avoidance clinical trial. *Medical Physics*, 42(6Part2):3196–3197, 2015. doi: https://doi.org/10.1118/1.4923816. URL https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4923816.

[7] Tokihiro Yamamoto, Sven Kabus, Matthieu Bal, Paul Keall, Stanley Benedict, and Megan Daly. The first patient treatment of computed tomography ventilation functional image-guided radiotherapy for lung cancer. *Radiotherapy and Oncology*, 118(2):227–231, 2016. ISSN 18790887. doi: 10.1016/j.radonc.2015.11.006. URL http://dx.doi.org/10.1016/j.radonc.2015.11.006.

[8] Brian P. Yaremko, Dante P.I. Capaldi, Khadija Sheikh, David A. Palma, Andrew Warner, A. Rashid Dar, Edward Yu, George B. Rodrigues, Alexander V. Louie, Mark Landis, Michael Sanatani, Mark D. Vincent, Jawaid Younus, Sara Kuruvilla, Jeff Z. Chen, Abigail Erickson, Stewart Gaede, Grace Parraga, and Douglas A. Hoover. Functional Lung Avoidance for Individualized Radiotherapy (FLAIR): Results of a Double-Blind, Randomized Controlled Trial. *International Journal of Radiation Oncology*Biology*Physics*, 2022. ISSN 03603016. doi: 10.1016/j.ijrobp.2022.04.047. URL https://doi.org/10.1016/j.ijrobp.2022.04.047.

[9] Yoshiyuki Shioyama, Si Young Jang, H. Helen Liu, Thomas Guerrero, Xuanmin

Wang, Isis W. Gayed, William D. Erwin, Zhongxing Liao, Joe Y. Chang, Melenda Jeter, Brian P. Yaremko, Yerko O. Borghero, James D. Cox, Ritsuko Komaki, and Radhe Mohan. Preserving Functional Lung Using Perfusion Imaging and Intensity-Modulated Radiation Therapy for Advanced-Stage Non-Small Cell Lung Cancer. *International Journal of Radiation Oncology Biology Physics*, 68(5): 1349–1358, 2007. ISSN 03603016. doi: 10.1016/j.ijrobp.2007.02.015.

[10] Emma L. Bates, Christopher M. Bragg, Jim M. Wild, Matthew Q F Hatton, and Rob H. Ireland. Functional image-based radiotherapy planning for non-small cell lung cancer: A simulation study. *Radiotherapy and Oncology*, 93(1):32–36, 2009. ISSN 01678140. doi: 10.1016/j.radonc.2009.05.018. URL `http://dx.doi.org/10.1016/j.radonc.2009.05.018`.

[11] Tokihiro Yamamoto, Sven Kabus, Tobias Klinder, Cristian Lorenz, Jens Von Berg, Thomas Blaffert, Billy W. Loo, and Paul J. Keall. Investigation of four-dimensional computed tomography-based pulmonary ventilation imaging in patients with emphysematous lung regions. *Physics in Medicine and Biology*, 56(7):2279–2298, 2011. ISSN 00319155. doi: 10.1088/0031-9155/56/7/023.

[12] Yevegeniy Vinogradskiy, Quentin Diot, Brian Kavanagh, Tracey Schefter, Laurie Gaspar, Moyed Miften, Yevegeniy Vinogradskiy, Quentin Diot, Brian Kavanagh, Tracey Schefter, and Laurie Gaspar. Spatial and dose – response analysis of fibrotic lung changes after stereotactic body radiation therapy Spatial and dose – response analysis of fibrotic lung changes after stereotactic body radiation therapy. *Medical Physics*, 081712(8):1–9, 2013. doi: 10.1118/1.4813916.

[13] Austin M. Faught, Yuya Miyasaka, Noriyuki Kadoya, Richard Castillo, Edward Castillo, Yevgeniy Vinogradskiy, and Tokihiro Yamamoto. Evaluating the Toxicity Reduction With Computed Tomographic Ventilation Functional Avoidance

Radiation Therapy. *International Journal of Radiation Oncology Biology Physics*, 99(2):325–333, 2017. ISSN 1879355X. doi: 10.1016/j.ijrobp.2017.04.024. URL http://dx.doi.org/10.1016/j.ijrobp.2017.04.024.

[14] Katherina P. Farr, Ditte S. Møller, Azza A. Khalil, Stine Kramer, Anni Morsing, and Cai Grau. Loss of lung function after chemo-radiotherapy for NSCLC measured by perfusion SPECT/CT: Correlation with radiation dose and clinical morbidity. *Acta Oncologica*, 54(9):1350–1354, 2015. ISSN 1651226X. doi: 10.3109/0284186X.2015.1061695. URL http://www.tandfonline.com/doi/abs/10.3109/0284186X.2015.1061695{#}.VqvQscv2apo.

[15] Shankar Siva, Roshini Thomas, Jason Callahan, Nicholas Hardcastle, Daniel Pham, Tomas Kron, Rodney J. Hicks, Michael P. MacManus, David L. Ball, and Michael S. Hofman. High-resolution pulmonary ventilation and perfusion PET/CT allows for functionally adapted intensity modulated radiotherapy in lung cancer. *Radiotherapy and Oncology*, 115(2):157–162, 2015. ISSN 18790887. doi: 10.1016/j.radonc.2015.04.013. URL http://dx.doi.org/10.1016/j.radonc.2015.04.013.

[16] Keith L. Miller, Timothy D. Shafman, Mitchell S. Anscher, Su Min Zhou, Robert W. Clough, Jennifer L. Garst, Jeffrey Crawford, Julian Rosenman, Mark A. Socinski, William Blackstock, Gregory S. Sibley, and Lawrence B. Marks. Bronchial stenosis: An underreported complication of high-dose external beam radiotherapy for lung cancer? *International Journal of Radiation Oncology Biology Physics*, 61(1):64–69, 2005. ISSN 03603016. doi: 10.1016/j.ijrobp.2004.02.066.

[17] Chris R. Kelsey, Daniel Kahn, Donna R. Hollis, Keith L. Miller, Su Min Zhou, Robert W. Clough, and Lawrence B. Marks. Radiation-induced narrowing of the tracheobronchial tree: An in-depth analysis. *Lung Cancer*, 52(1):111–116, 2006. ISSN 01695002. doi: 10.1016/j.lungcan.2005.11.007.

[18] Robert Timmerman, Ronald McGarry, Constantin Yiannoutsos, Lech Papiez, Kathy Tudor, Jill DeLuca, Marvene Ewing, Ramzi Abdulrahman, Colleen DesRosiers, Mark Williams, and James Fletcher. Excessive toxicity when treating central tumors in a phase II study of stereotactic body radiation therapy for medically inoperable early-stage lung cancer. *Journal of Clinical Oncology*, 24(30): 4833–4839, 2006. ISSN 0732183X. doi: 10.1200/JCO.2006.07.5937.

[19] Kristin Karlsson, Jan Nyman, Pia Baumann, Peter Wersäll, Ninni Drugge, Giovanna Gagliardi, Karl Axel Johansson, Jan Olov Persson, Eva Rutkowska, Owe Tullgren, and Ingmar Lax. Retrospective cohort study of bronchial doses and radiation-induced atelectasis after stereotactic body radiation therapy of lung tumors located close to the bronchial tree. *International Journal of Radiation Oncology Biology Physics*, 87(3):590–595, 2013. ISSN 03603016. doi: 10.1016/j.ijrobp.2013.06.2055.

[20] Jason Bates. *Lung Mechanics*. Cambridge University Press, 2009. ISBN 9780521509602.

[21] Yoshiharu Ohno, Hiroto Hatabu, and Hans-Ulrich Kauczor. *Pulmonary Functional Imaging*. Springer Nature Switzerland, 2021. ISBN 978-3-030-43538-7. URL `http://link.springer.com/10.1007/978-3-030-43539-4`.

[22] OpenStax. *Concepts of Biology by OpenStax*. XanEdu Publishing Inc, paperback edition, 5 2016. ISBN 978-1506696539.

[23] Julian R. Molina, Ping Yang, Stephen D. Cassivi, Steven E. Schild, and Alex A. Adjei. Non-Small Cell Lung Cancer: Epidemiology, Risk Factors, Treatment, and Survivorship. *Mayo Clinic Proceedings*, 83(5):584–594, May 2008. ISSN 00256196. doi: 10.4065/83.5.584. URL `https://linkinghub.elsevier.com/retrieve/pii/S0025619611607350`.

[24] Kimberly D. Miller, Rebecca L. Siegel, Chun Chieh Lin, Angela B. Mariotto, Joan L. Kramer, Julia H. Rowland, Kevin D. Stein, Rick Alteri, and Ahmedin Jemal. Cancer treatment and survivorship statistics, 2016. *CA: A Cancer Journal for Clinicians*, 66(4):271–289, July 2016. ISSN 00079235. doi: 10.3322/caac.21349. URL `http://doi.wiley.com/10.3322/caac.21349`.

[25] Kai Ding, John E. Bayouth, John M. Buatti, Gary E. Christensen, and Joseph M. Reinhardt. 4DCT-based measurement of changes in pulmonary function following a course of radiation therapy. *Medical Physics*, 37(3):1261–1272, 2010. ISSN 00942405. doi: 10.1118/1.3312210.

[26] David A. Palma, John van Sörnsen de Koste, Wilko F.A.R. Verbakel, Andrew Vincent, and Suresh Senan. Lung Density Changes After Stereotactic Radiotherapy: A Quantitative Analysis in 50 Patients. *International Journal of Radiation Oncology*Biology*Physics*, 81(4):974–978, November 2011. ISSN 03603016. doi: 10.1016/j.ijrobp.2010.07.025. URL `https://linkinghub.elsevier.com/retrieve/pii/S0360301610009569`.

[27] Ghazaleh Ghobadi, Laurens E. Hogeweg, Hette Faber, Wim G. J. Tukker, Jacobus M. Schippers, Sytze Brandenburg, Johannes A. Langendijk, Robert P. Coppes, and Peter van Luijk. Quantifying Local Radiation-Induced Lung Damage From Computed Tomography. *International Journal of Radiation Oncology, Biology, Physics*, 76(2):548–556, February 2010. ISSN 0360-3016. doi: 10.1016/j.ijrobp.2009.08.058. URL `https://www.redjournal.org/article/S0360-3016(09)03038-7/fulltext`. Publisher: Elsevier.

[28] Shannon O'Reilly, Varsha Jain, Qijie Huang, Chingyun Cheng, Boon Keng Kevin Teo, Lingshu Yin, Miao Zhang, Eric Diffenderfer, Taoran Li, William Levin, Ying Xiao, Lei Dong, Steven Feigenberg, Abigail T Berman, and Wei Zou. Dose to

Highly Functional Ventilation Zones Improves Prediction of Radiation Pneumonitis for Proton and Photon Lung Cancer Radiation Therapy. In *International Journal of Radiation Oncology Biology Physics*, volume 107, pages 79–87. Elsevier Inc., 2020. doi: 10.1016/j.ijrobp.2020.01.014. URL https://doi.org/10.1016/j.ijrobp.2020.01.014.

[29] Narges Kazemzadeh, Arezoo Modiri, Santanu Samanta, Yulong Yan, Ross Bland, Timothy Rozario, Henky Wibowo, Puneeth Iyengar, Chul Ahn, Robert Timmerman, and Amit Sawant. Virtual Bronchoscopy-Guided Treatment Planning to Map and Mitigate Radiation-Induced Airway Injury in Lung SAbR. *International Journal of Radiation Oncology • Biology • Physics*, 102(1):210–218, sep 2018. ISSN 0360-3016. doi: 10.1016/j.ijrobp.2018.04.060. URL https://doi.org/10.1016/j.ijrobp.2018.04.060.

[30] E Vicente, A Modiri, J Kipritidis, A Hagan, K Yu, H Wibowo, Y Yan, D R Owen, M M Matuszak, P Mohindra, R Timmerman, and A Sawant. Functionally weighted airway sparing (FWAS): a functional avoidance method for preserving post-treatment ventilation in lung radiotherapy. *Physics in Medicine & Biology*, 65(16), aug 2020. ISSN 1361-6560. doi: 10.1088/1361-6560/ab9f5d. URL https://iopscience.iop.org/article/10.1088/1361-6560/ab9f5d.

[31] Weili Wang, Martha M. Matuszak, Chen Hu, Ke Colin Huang, Eileen Chen, Douglas Arenberg, Jeffrey L. Curtis, Shruti Jolly, Jian Yue Jin, Mitchell Machtay, Randall K. Ten Haken, and Feng Ming (Spring) Kong. Central Airway Toxicity After High Dose Radiation: A Combined Analysis of Prospective Clinical Trials for Non-Small Cell Lung Cancer. *International Journal of Radiation Oncology Biology Physics*, 108(3):587–596, 2020. ISSN 1879355X. doi: 10.1016/j.ijrobp.2020.05.026. URL https://doi.org/10.1016/j.ijrobp.2020.05.026.

[32] Bindu V. Manyam, Kyle Verdecchia, Gregory M.M. Videtic, Tingliang Zhuang, Neil M. Woody, Wei Wei, Zi Ouyang, and Kevin L. Stephans. Validation of RTOG 0813 Proximal Bronchial Tree Constraints for Pulmonary Toxicity With Stereotactic Body Radiation Therapy for Central Non-small Cell Lung Cancer. *International Journal of Radiation Oncology Biology Physics*, 107(1):72–78, 2020. ISSN 1879355X. doi: 10.1016/j.ijrobp.2020.01.009. URL https://doi.org/10.1016/j.ijrobp.2020.01.009.

[33] Pulmonary function tests. URL https://www.nhlbi.nih.gov/health-topics/pulmonary-function-tests.

[34] Harpreet Ranu, Michael Wilde, and Brendan Madden. Pulmonary function tests. *The Ulster medical journal*, 80(2):84–90, may 2011. ISSN 2046-4207. doi: 10.1007/978-3-030-43539-4_2. URL http://link.springer.com/10.1007/978-3-030-43539-4_2http://www.ncbi.nlm.nih.gov/pubmed/22347750http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3229853.

[35] Warren M. Gold and Laura L. Koth. Pulmonary Function Testing. In *Murray and Nadel's Textbook of Respiratory Medicine*, volume 2507, pages 407–435.e18. Elsevier, 2016. ISBN 9789896540821. doi: 10.1016/B978-1-4557-3383-5.00025-7. URL https://linkinghub.elsevier.com/retrieve/pii/B9781455733835000257.

[36] Mark Henderson, Ronald McGarry, Constantin Yiannoutsos, Achilles Fakiris, David Hoopes, Mark Williams, and Robert Timmerman. Baseline Pulmonary Function as a Predictor for Survival and Decline in Pulmonary Function Over Time in Patients Undergoing Stereotactic Body Radiotherapy for the Treatment of Stage I Non-Small-Cell Lung Cancer. *International Journal of Radiation Oncology Biology Physics*, 72(2):404–409, 2008. ISSN 03603016. doi: 10.1016/j.ijrobp.2007.12.051.

[37] Michael S. Binkley, Martin T. King, Joseph B. Shrager, Karl Bush, Aadel A. Chaudhuri, Rita Popat, Michael F. Gensheimer, Peter G. Maxim, H. Henry Guo, Maximilian Diehn, Viswam S. Nair, and Billy W. Loo. Pulmonary function after lung tumor stereotactic ablative radiotherapy depends on regional ventilation within irradiated lung. *Radiotherapy and Oncology*, 123(2):270–275, 2017. ISSN 18790887. doi: 10.1016/j.radonc.2017.03.021. URL http://dx.doi.org/10.1016/j.radonc.2017.03.021.

[38] Cinzia Ferrero, Serena Badellino, Andrea Riccardo Filippi, Luana Focaraccio, Matteo Giaj Levra, Mario Levis, Francesco Moretto, Roberto Torchio, Umberto Ricardi, and Silvia Novello. Pulmonary function and quality of life after VMAT-based stereotactic ablative radiotherapy for early stage inoperable NSCLC: A prospective study. *Lung Cancer*, 89(3):350–356, 2015. ISSN 18728332. doi: 10.1016/j.lungcan.2015.06.019.

[39] Brandon Stone, Victor S. Mangona, Matthew D. Johnson, Hong Ye, and Inga S. Grills. Changes in Pulmonary Function Following Image-Guided Stereotactic Lung Radiotherapy: Neither Lower Baseline Nor Post-SBRT Pulmonary Function Are Associated with Worse Overall Survival. *Journal of Thoracic Oncology*, 10(12): 1762–1769, 2015. ISSN 15561380. doi: 10.1097/JTO.0000000000000670. URL http://dx.doi.org/10.1097/JTO.0000000000000670.

[40] Henry N. Wagner, David C. Sabiston, John G. McAfee, Donald Tow, and Howard S. Stern. Diagnosis of massive pulmonary embolism in man by radioisotope scanning. *New England Journal of Medicine*, 271(8):377–384, August 1964. doi: 10.1056/nejm196408202710801. URL https://doi.org/10.1056/nejm196408202710801.

[41] Henry N. Wagner, David C. Sabiston, Masahiro Iio, John G. McAfee, Jon K. Meyer, and James K. Langan. Regional pulmonary blood flow in man by radioisotope

scanning. *JAMA*, 187(8), February 1964. doi: 10.1001/jama.1964.03060210051012. URL https://doi.org/10.1001/jama.1964.03060210051012.

[42] Simon R. Cherry, James A. Sorenson, and Michael E. Phelps. *Physics in nuclear medicine*. Elsevier/Saunders, Philadelphia, 4th ed edition, 2012. ISBN 9781416051985.

[43] George V. Taplin and Norman S. MacDonald. Radiochemistry of macroaggregated albumin and newer lung scanning agents. *Seminars in Nuclear Medicine*, 1(2):132–152, April 1971. doi: 10.1016/s0001-2998(71)81013-9. URL https://doi.org/10.1016/s0001-2998(71)81013-9.

[44] Alan M. Kumar and J. Anthony Parker. VENTILATION/PERFUSION SCINTIGRAPHY. *Emergency Medicine Clinics of North America*, 19(4):957–974, November 2001. doi: 10.1016/s0733-8627(05)70229-7. URL https://doi.org/10.1016/s0733-8627(05)70229-7.

[45] Nelson P Trujillo, Jonathan P Pratt, Sachin Talusani, Robert A Quaife, David Kumpe, and James L Lear. Dtpa aerosol in ventilation/perfusion scintigraphy for diagnosing pulmonary embolism. *Journal of Nuclear Medicine*, 38(11):1781–1783, 1997.

[46] WILLIAM M. BURCH, PAUL J. SULLIVAN, and CHRISTOPHER J. MCLAREN. Technegas - a new ventilation agent for lung scanning. *Nuclear Medicine Communications*, 7(12):865–872, December 1986. doi: 10.1097/00006231-198612000-00003. URL https://doi.org/10.1097/00006231-198612000-00003.

[47] Jonas Jögi, Björn Jonson, Marie Ekberg, and Marika Bajc. Ventilation-perfusion SPECT with 99mTc-DTPA versus Technegas: a head-to-head study in obstructive and nonobstructive disease. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 51(5):735–741, 2010. ISSN 15355667. doi: 10.2967/jnumed.109.073957.

[48] Paul J. Roach, Denis J. Gradinscak, Geoffrey P. Schembri, Elizabeth A. Bailey, Kathy P. Willowson, and Dale L. Bailey. SPECT/CT in v/q scanning. *Seminars in Nuclear Medicine*, 40(6):455–466, November 2010. doi: 10.1053/j.semnuclmed.2010.07.005. URL `https://doi.org/10.1053/j.semnuclmed.2010.07.005`.

[49] Paul J. Roach, Geoffrey P. Schembri, and Dale L. Bailey. V/q scanning using SPECT and SPECT/CT. *Journal of Nuclear Medicine*, 54(9):1588–1596, August 2013. doi: 10.2967/jnumed.113.124602. URL `https://doi.org/10.2967/jnumed.113.124602`.

[50] Dale L. Bailey, Enid M. Eslick, Geoffrey P. Schembri, and Paul J. Roach. 68Ga PET Ventilation and Perfusion Lung Imaging—Current Status and Future Challenges. *Seminars in Nuclear Medicine*, 46(5):428–435, 2016. ISSN 15584623. doi: 10.1053/j.semnuclmed.2016.04.007. URL `http://dx.doi.org/10.1053/j.semnuclmed.2016.04.007`.

[51] Pierre Yves Le Roux, Shankar Siva, Daniel P. Steinfort, Jason Callahan, Peter Eu, Lou B. Irving, Rodney J. Hicks, and Michael S. Hofman. Correlation of 68Ga ventilation-perfusion PET/CT with pulmonary function test indices for assessing lung function. *Journal of Nuclear Medicine*, 56(11):1718–1723, 2015. ISSN 01615505. doi: 10.2967/jnumed.115.162586.

[52] Robert A. Pooley. Fundamental physics of MR imaging. *RadioGraphics*, 25(4): 1087–1099, July 2005. doi: 10.1148/rg.254055027. URL https://doi.org/10.1148/rg.254055027.

[53] Giles Santyr, Matthew Fox, Kundan Thind, Elaine Hegarty, Alexei Ouriadov, Michael Jensen, Timothy J. Scholl, Jacob Van Dyk, and Eugene Wong. Anatomical, functional and metabolic imaging of radiation-induced lung injury using hyperpolarized MRI. *NMR in Biomedicine*, 27(12):1515–1524, 2014. ISSN 10991492. doi: 10.1002/nbm.3180.

[54] Stanley J. Kruger, Scott K. Nagle, Marcus J. Couch, Yoshiharu Ohno, Mitchell Albert, and Sean B. Fain. Functional imaging of the lungs with gas agents. *Journal of Magnetic Resonance Imaging*, 43(2):295–315, 2016. ISSN 15222586. doi: 10.1002/jmri.25002.

[55] Sean Fain, Mark L. Schiebler, David G. McCormack, and Grace Parraga. Imaging of lung function using hyperpolarized helium-3 magnetic resonance imaging: Review of current and emerging translational methods and applications. *Journal of Magnetic Resonance Imaging*, 32(6):1398–1408, November 2010. doi: 10.1002/jmri.22375. URL https://doi.org/10.1002/jmri.22375.

[56] Lukas Ebner, Jeff Kammerman, Bastiaan Driehuys, Mark L. Schiebler, Robert V. Cadman, and Sean B. Fain. The role of hyperpolarized 129xenon in MR imaging of pulmonary function. *European Journal of Radiology*, 86(5):343–352, jan 2017. ISSN 0720048X. doi: 10.1016/j.ejrad.2016.09.015. URL https://linkinghub.elsevier.com/retrieve/pii/S0720048X16302844.

[57] Justus E. Roos, Holman P. McAdams, S. Sivaram Kaushik, and Bastiaan Driehuys. Hyperpolarized Gas MRI: Technique and Applications. *Magnetic resonance imaging clinics of North America*, 23(2):217–229, May 2015. ISSN 1064-9689. doi: 10.

1016/j.mric.2015.01.003. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4428591/.

[58] Yi Ding, Lu Yang, Qian Zhou, Jianping Bi, Ying Li, Guoliang Pi, Wei Wei, Desheng Hu, Qiuchen Rao, Haidong Li, Li Zhao, An Liu, Dongsu Du, Xiao Wang, Xin Zhou, Guang Han, and Kun Qing. A pilot study of function-based radiation therapy planning for lung cancer using hyperpolarized xenon-129 ventilation MRI. *Journal of Applied Clinical Medical Physics*, 23(3), March 2022. ISSN 1526-9914, 1526-9914. doi: 10.1002/acm2.13502. URL https://onlinelibrary.wiley.com/doi/10.1002/acm2.13502.

[59] Leith J. Rankine, Ziyi Wang, Chris R. Kelsey, Elianna Bier, Bastiaan Driehuys, Lawrence B. Marks, and Shiva K. Das. Hyperpolarized 129Xe Magnetic Resonance Imaging for Functional Avoidance Treatment Planning in Thoracic Radiation Therapy: A Comparison of Ventilation- and Gas Exchange-Guided Treatment Plans. *International Journal of Radiation Oncology, Biology, Physics*, 111(4):1044–1057, November 2021. ISSN 0360-3016. doi: 10.1016/j.ijrobp.2021.07.002. URL https://www.redjournal.org/article/S0360-3016(21)00871-3/fulltext. Publisher: Elsevier.

[60] Vivek Elangovan. Computed tomography dose index, 2014. URL https://www.slideshare.net/VivekElangovan1/ctdi-42682868.

[61] Joanne Mcnamara. *Investigation of two respiratory monitoring systems used for 4D CT and respiratory gating.* PhD thesis, Wollongong, 2008.

[62] Christian Heinz, Michael Reiner, Claus Belka, Franziska Walter, and Matthias Söhn. Technical evaluation of different respiratory monitoring systems used for 4D CT acquisition under free breathing. *Journal of Applied Clinical Medical Physics*, 16(2):334–349, 2015. ISSN 15269914. doi: 10.1120/jacmp.v16i2.4917.

[63] Shinichiro Mori, Susumu Ko, Takayoshi Ishii, and Kanae Nishizawa. Effective doses in four-dimensional computed tomography for lung radiotherapy planning. *Medical Dosimetry*, 34(1):87–90, March 2009. doi: 10.1016/j.meddos.2008.08.002. URL https://doi.org/10.1016/j.meddos.2008.08.002.

[64] Patricia Hubbard, Jason Callahan, Jim Cramb, Ray Budd, and Tomas Kron. Audit of radiation dose delivered in time-resolved four-dimensional computed tomography in a radiotherapy department. *Journal of Medical Imaging and Radiation Oncology*, 59(3):346–352, March 2015. doi: 10.1111/1754-9485.12284. URL https://doi.org/10.1111/1754-9485.12284.

[65] George Noid, An Tai, Guang-Pei Chen, Jared Robbins, and X. Allen Li. Reducing radiation dose and enhancing imaging quality of 4dct for radiation therapy using iterative reconstruction algorithms. *Advances in Radiation Oncology*, 2(3):515–521, July 2017. doi: 10.1016/j.adro.2017.04.003. URL https://doi.org/10.1016/j.adro.2017.04.003.

[66] Tokihiro Yamamoto, Ulrich Langner, Billy W. Loo, John Shen, and Paul J. Keall. Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. *International Journal of Radiation Oncology*Biology*Physics*, 72(4):1250–1258, November 2008. doi: 10.1016/j.ijrobp.2008.06.1937. URL https://doi.org/10.1016/j.ijrobp.2008.06.1937.

[67] Kristy K. Brock, Sasa Mutic, Todd R. McNutt, Hua Li, and Marc L. Kessler. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132: Report. *Medical Physics*, 44(7):e43–e76, 2017. ISSN 00942405. doi: 10.1002/mp.12256.

[68] Kunlin Cao, Kai Ding, Gary E. Christensen, and Joseph M. Reinhardt. Tissue volume and vesselness measure preserving nonrigid registration of lung CT images.

Medical Imaging 2010: Image Processing, 7623(March 2010):762309, 2010. ISSN 16057422. doi: 10.1117/12.844541.

[69] Thomas Guerrero, Kevin Sanders, Josue Noyola-Martinez, Edward Castillo, Yin Zhang, Richard Tapia, Rudy Guerra, Yerko Borghero, and Ritsuko Komaki. Quantification of regional ventilation from treatment planning ct. *International Journal of Radiation Oncology Biology Physics*, 62:630–634, 2005. ISSN 03603016. doi: 10.1016/j.ijrobp.2005.03.023.

[70] Brett A Simon. Non-invasive imaging of regional lung function using x-ray computed tomography. *Journal of Clinical Monitoring and Computing*, 16:433–442, 2000.

[71] Joseph M. Reinhardt, Kai Ding, Kunlin Cao, Gary E. Christensen, Eric A. Hoffman, and Shalmali V. Bodas. Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation. *Medical Image Analysis*, 12(6):752–763, December 2008. doi: 10.1016/j.media.2008.03.007. URL `https://doi.org/10.1016/j.media.2008.03.007`.

[72] W. Shao, T.J. Patton, S.E. Gerard, Y. Pan, J.M. Reinhardt, J.E. Bayouth, O.C. Durumeric, and G.E. Christensen. Detecting out-of-phase ventilation using 4dct to improve radiation therapy for lung cancer. pages 251–259, 2018.

[73] Wei Shao, Taylor J. Patton, Sarah E. Gerard, Yue Pan, Joseph M. Reinhardt, Oguz C. Durumeric, John E. Bayouth, and Gary E. Christensen. N-Phase Local Expansion Ratio for Characterizing Out-of-Phase Lung Ventilation. *IEEE Transactions on Medical Imaging*, 39(6):2025–2034, 2020. ISSN 1558254X. doi: 10.1109/TMI.2019.2963083.

[74] K. Du, J.M. Reinhardt, G.E. Christensen, K. Ding, and J.E. Bayouth. Respiratory effort correction strategies to improve the reproducibility of lung expansion measurements. *Medical Physics*, 40(12):123504, 2013. doi: 10.1118/1.4829519. URL https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4829519.

[75] Eric M. Wallat, Mattison J. Flakus, Antonia E. Wuschner, Wei Shao, Gary E. Christensen, Joseph M. Reinhardt, Andrew M. Baschnagel, and John E. Bayouth. Modeling the impact of out-of-phase ventilation on normal lung tissue response to radiation dose. *Medical Physics*, 47(7):3233–3242, jul 2020. ISSN 00942405. doi: 10.1002/mp.14146.

[76] Eunsin Lee, Jing Zeng, Robert S. Miyaoka, Jatinder Saini, Paul E. Kinahan, George A. Sandison, Tony Wong, Hubert J. Vesselle, Ramesh Rengan, and Stephen R. Bowen. Functional lung avoidance and response-adaptive escalation (FLARE) RT: Multimodality plan dosimetry of a precision radiation oncology strategy: Multimodality. *Medical Physics*, 44(7):3418–3429, 2017. ISSN 00942405. doi: 10.1002/mp.12308.

[77] Yevgeniy Vinogradskiy, Richard Castillo, Edward Castillo, Leah Schubert, Bernard L. Jones, Austin Faught, Laurie E. Gaspar, Jennifer Kwak, Daniel W. Bowles, Timothy Waxweiler, Jingjing M. Dougherty, Dexiang Gao, Craig Stevens, Moyed Miften, Brian Kavanagh, Inga Grills, Chad G. Rusthoven, and Thomas Guerrero. Results of a Multi-Institutional Phase 2 Clinical Trial for 4DCT-Ventilation Functional Avoidance Thoracic Radiation Therapy. *International Journal of Radiation Oncology Biology Physics*, 112(4):986–995, 2022. ISSN 1879355X. doi: 10.1016/j.ijrobp.2021.10.147. URL https://doi.org/10.1016/j.ijrobp.2021.10.147.

[78] Esther M. Vicente, Arezoo Modiri, John Kipritidis, Kun Chang Yu, Kai Sun, Jochen Cammin, Arun Gopal, Jingzhu Xu, Sina Mossahebi, Aaron Hagan, Yulong Yan, Daniel Rockwell Owen, Pranshu Mohindra, Martha M. Matuszak, Robert D. Timmerman, and Amit Sawant. Combining Serial and Parallel Functionality in Functional Lung Avoidance Radiation Therapy. *International Journal of Radiation Oncology Biology Physics*, 113(2):456–468, 2022. ISSN 1879355X. doi: 10.1016/j.ijrobp.2022.01.046. URL https://doi.org/10.1016/j.ijrobp.2022.01.046.

[79] Taylor J. Patton, Sarah E. Gerard, Wei Shao, Gary E. Christensen, Joseph M. Reinhardt, and John E. Bayouth. Quantifying ventilation change due to radiation therapy using 4DCT Jacobian calculations. *Medical Physics*, 45(10):4483–4492, 2018. ISSN 00942405. doi: 10.1002/mp.13105.

[80] Gerald J. Kutcher and C. Burman. Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method gerald. *International Journal of Radiation Oncology*Biology*Physics*, 16(6):1623–1630, 1989. ISSN 0360-3016. doi: https://doi.org/10.1016/0360-3016(89)90972-3. URL https://www.sciencedirect.com/science/article/pii/0360301689909723.

[81] Stefan L. S. Kwa, Joos V. Lebesque, Jacqueline C. M. Theuws, Lawrence B. Marks, Mike T. Munley, Gunilla Bentel, Dieter Oetzel, Uwe Spahn, Mary V. Graham, Robert E. Drzymala, James A. Purdy, Allen S. Lichter, Mary K. Martel, and Randall K. Ten Haken. Radiation pneumonitis as a function of mean lung dose: an analysis of pooled data of 540 patients. *International Journal of Radiation Oncology*Biology*Physics*, 42(1):1–9, 1998. ISSN 0360-3016. doi: https://doi.org/10.1016/S0360-3016(98)00196-5. URL https://www.sciencedirect.com/science/article/pii/S0360301698001965.

[82] Lawrence B. Marks, Soren M. Bentzen, Joseph O. Deasy, Feng-Ming (Spring) Kong, Jeffrey D. Bradley, Ivan S. Vogelius, Issam El Naqa, Jessica L. Hubbs, Joos V. Lebesque, Robert D. Timmerman, Mary K. Martel, and Andrew Jackson. Radiation Dose–Volume Effects in the Lung. *International Journal of Radiation Oncology*Biology*Physics*, 76(3):S70–S76, March 2010. ISSN 03603016. doi: 10.1016/j.ijrobp.2009.06.091. URL https://linkinghub.elsevier.com/retrieve/pii/S0360301609032933.

[83] Andrzej Niemierko and Michael Goitein. Modeling of normal tissue response to radiation: The critical volume model. *International Journal of Radiation Oncology*Biology*Physics*, 25(1):135–145, 1993. ISSN 0360-3016. doi: https://doi.org/10.1016/0360-3016(93)90156-P. URL https://www.sciencedirect.com/science/article/pii/036030169390156P.

[84] Randi Vågane and Dag R. Olsen. Analysis of normal tissue complication probability of the lung using a reliability model. *Acta Oncologica*, 45(5):610–617, January 2006. ISSN 0284-186X. doi: 10.1080/02841860600658245. URL https://doi.org/10.1080/02841860600658245. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02841860600658245.

[85] H. Tekatli, M. Duijm, E. Oomen-de Hoop, W. Verbakel, W. Schillemans, B.J. Slotman, J.J. Nuyttens, and S. Senan. Normal Tissue Complication Probability Modeling of Pulmonary Toxicity After Stereotactic and Hypofractionated Radiation Therapy for Central Lung Tumors. *International Journal of Radiation Oncology*Biology*Physics*, 100(3):738–747, March 2018. ISSN 03603016. doi: 10.1016/j.ijrobp.2017.11.022. URL https://linkinghub.elsevier.com/retrieve/pii/S036030161734141X.

[86] Y. Vinogradskiy, R. Castillo, E. Castillo, S.L. Tucker, Z. Liao, T. Guerrero, and M.K. Martel. Use of 4-dimensional computed tomography-based ventilation imaging to correlate lung dose and function with clinical outcomes. *International Journal of Radiation Oncology*Biology*Physics*, 86(2):366 – 371, 2013. ISSN 0360-3016. doi: https://doi.org/10.1016/j.ijrobp.2013.01.004. URL `http://www.sciencedirect.com/science/article/pii/S0360301613000230`.

[87] Rockwell Anyoha. The history of artificial intelligence, Aug 2017. URL `https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/`.

[88] Alan Turing. Computing machinery and intelligence. *Mind*, 49:433–460, 1950.

[89] Varun Gulshan, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tom Madams, Jorge Cuadros, Ramasamy Kim, Rajiv Raman, Philip C. Nelson, Jessica L. Mega, and Dale R. Webster. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 316(22): 2402, December 2016. doi: 10.1001/jama.2016.17216. URL `https://doi.org/10.1001/jama.2016.17216`.

[90] Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, January 2017. doi: 10.1038/nature21056. URL `https://doi.org/10.1038/nature21056`.

[91] Richard Castillo, Edward Castillo, Josue Martinez, and Thomas Guerrero. Ventilation from four-dimensional computed tomography: Density versus Jacobian methods. *Physics in Medicine and Biology*, 55(16):4661–4685, 2010. ISSN 00319155. doi: 10.1088/0031-9155/55/16/004.

[92] Shifeng Chen, Sumin Zhou, Junan Zhang, Fang Fang Yin, Lawrence B. Marks, and Shiva K. Das. A neural network model to predict lung radiation-induced pneumonitis. *Medical Physics*, 34(9):3420–3427, 2007. ISSN 00942405. doi: 10.1118/1.2759601.

[93] Peng Huang, Hui Yan, Zhihui Hu, Zhiqiang Liu, Yuan Tian, and Jianrong Dai. Predicting radiation pneumonitis with fuzzy clustering neural network using 4DCT ventilation image based dosimetric parameters. *Quantitative Imaging in Medicine and Surgery*, 0(0):0–0, 2021. ISSN 22234292. doi: 10.21037/qims-20-1095.

[94] Gilmer Valdes, Timothy D Solberg, Marina Heskel, Lyle Ungar, and Charles B Simone. Using machine learning to predict radiation pneumonitis in patients with stage I non-small cell lung cancer treated with stereotactic body radiation therapy. *Physics in Medicine and Biology*, 61(16):6105–6120, 2016. ISSN 13616560. doi: 10.1088/0031-9155/61/16/6105.

[95] Yoshiyuki Katsuta, Noriyuki Kadoya, Shina Mouri, Shohei Tanaka, Takayuki Kanai, Kazuya Takeda, Takaya Yamamoto, Kengo Ito, Tomohiro Kajikawa, Yujiro Nakajima, and Keiichi Jingu. Prediction of radiation pneumonitis with machine learning using 4D-CT based dose-function features. *Journal of Radiation Research*, 63(1):71–79, 2022. ISSN 0449-3060. doi: 10.1093/jrr/rrab097.

[96] Sarah E. Gerard, Joseph M. Reinhardt, Gary E. Christensen, and Raul San Jose Estepar. Estimating Local Tissue Expansion in Thoracic Computed Tomography Images Using Convolutional Neural Networks. *Proceedings - International Symposium on Biomedical Imaging*, 2020-April:1856–1860, 2020. ISSN 19458452. doi: 10.1109/ISBI45749.2020.9098413.

[97] Yuncheng Zhong, Yevgeniy Vinogradskiy, Liyuan Chen, Nick Myziuk, Richard Castillo, Edward Castillo, Thomas Guerrero, Steve Jiang, and Jing Wang. Technical Note: Deriving ventilation imaging from 4DCT by deep convolutional neural network. *Medical Physics*, 46(5):2323–2329, 2019. ISSN 00942405. doi: 10.1002/mp.13421.

[98] Andrew Westcott, Dante P.I. Capaldi, David G. McCormack, Aaron D. Ward, Aaron Fenster, and Grace Parraga. Chronic obstructive pulmonary disease: Thoracic CT texture analysis and machine learning to predict pulmonary ventilation. *Radiology*, 293(3):676–684, 2019. ISSN 15271315. doi: 10.1148/radiol.2019190450.

[99] Dante P.I. Capaldi, Fumin Guo, Lei Xing, and Grace Parraga. Pulmonary Ventilation Maps Generated with Free-breathing Proton MRI and a Deep Convolutional Neural Network. *Radiology*, 298(2):427–438, 2021. ISSN 15271315. doi: 10.1148/RADIOL.2020202861.

[100] Nicholas J. Tustison, Brian B. Avants, Zixuan Lin, Xue Feng, Nicholas Cullen, Jaime F. Mata, Lucia Flors, James C. Gee, Talissa A. Altes, John P. Mugler, and Kun Qing. Convolutional Neural Networks with Template-Based Data Augmentation for Functional Lung Image Quantification. *Academic Radiology*, 26(3): 412–423, 2019. ISSN 18784046. doi: 10.1016/j.acra.2018.08.003. URL `https://doi.org/10.1016/j.acra.2018.08.003`.

[101] Ge Ren, Jiang Zhang, Tian Li, Haonan Xiao, Lai Yin Cheung, Wai Yin Ho, Jing Qin, and Jing Cai. Deep Learning-Based Computed Tomography Perfusion Mapping (DL-CTPM) for Pulmonary CT-to-Perfusion Translation. *International Journal of Radiation Oncology Biology Physics*, 110(5):1508–1518, 2021. ISSN 1879355X. doi: 10.1016/j.ijrobp.2021.02.032. URL `https://doi.org/10.1016/j.ijrobp.2021.02.032`.

[102] Ge Ren, Sai Kit Lam, Jiang Zhang, Haonan Xiao, Andy Lai yin Cheung, Wai Yin Ho, Jing Qin, and Jing Cai. Investigation of a Novel Deep Learning-Based Computed Tomography Perfusion Mapping Framework for Functional Lung Avoidance Radiotherapy. *Frontiers in Oncology*, 11(March):1–12, 2021. ISSN 2234943X. doi: 10.3389/fonc.2021.644703.

[103] Evan M. Porter, Nicholas K. Myziuk, Thomas J. Quinn, Daniela Lozano, Avery B. Peterson, Duyen M. Quach, Zaid A. Siddiqui, and Thomas M. Guerrero. Synthetic pulmonary perfusion images from 4DCT for functional avoidance using deep learning. *Physics in Medicine and Biology*, 66(17), 2021. ISSN 13616560. doi: 10.1088/1361-6560/ac16ec.

[104] Pietro Nardelli, Rubén San José Estépar, Farbod N. Rahaghi, and Raúl San José Estépar. Functional-Consistent CycleGAN for CT to Iodine Perfusion Map Translation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12502 LNCS:109–117, 2020. ISSN 16113349. doi: 10.1007/978-3-030-62469-9_10.

[105] Jiantao Pu, Joseph K Leader, Jacob Sechrist, Cameron A Beeche, Jatin P Singh, Iclal Ocak, and Michael Risbano. Automated identification of pulmonary arteries and veins depicted in non-contrast chest CT scans. *Medical Image Analysis*, page 102367, 2022. ISSN 13618415. doi: 10.1016/j.media.2022.102367. URL `https://doi.org/10.1016/j.media.2022.102367`.

[106] Sarah E. Gerard, Taylor J Patton, Gary E Christensen, John E Bayouth, and Joseph M Reinhardt. FissureNet: A Deep Learning Approach For Pulmonary Fissure Detection in CT Images. *IEEE Trans Med Imaging.*, 38(1):156–166, 2019. doi: doi:10.1109/TMI.2018.2858202.

[107] Sarah E. Gerard and Joseph M. Reinhardt. Pulmonary lobe segmentation using a sequence of convolutional neural networks for marginal learning. *Proceedings - International Symposium on Biomedical Imaging*, 2019-April(Isbi):1207–1211, 2019. ISSN 19458452. doi: 10.1109/ISBI.2019.8759212.

[108] Sarah E. Gerard, Jacob Herrmann, David W. Kaczka, Guido Musch, Ana Fernandez-Bustamante, and Joseph M. Reinhardt. Multi-resolution convolutional neural networks for fully automated segmentation of acutely injured lungs in multiple species. *Medical Image Analysis*, 60:101592, February 2020. doi: 10.1016/j.media.2019.101592. URL `https://doi.org/10.1016/j.media.2019.101592`.

[109] Jiayi Lu, Renchao Jin, Enmin Song, Guangzhi Ma, and Manyang Wang. Lung-CRNet: A convolutional recurrent neural network for lung 4DCT image registration. *Medical Physics*, 48(12):7900–7912, 2021. ISSN 00942405. doi: 10.1002/mp.15324.

[110] Tobias Fechter and Dimos Baltas. One-Shot Learning for Deformable Medical Image Registration and Periodic Motion Tracking. *IEEE Transactions on Medical Imaging*, 39(7):2506–2517, 2020. ISSN 1558254X. doi: 10.1109/TMI.2020.2972616.

[111] Koen A. J. Eppenhof and Josien P. W. Pluim. Pulmonary CT registration through supervised learning with convolutional neural networks. *IEEE Transactions on Medical Imaging*, 38(5):1097–1105, May 2019. doi: 10.1109/tmi.2018.2878316. URL `https://doi.org/10.1109/tmi.2018.2878316`.

[112] Zhuoran Jiang, Fang-Fang Yin, Yun Ge, and Lei Ren. A multi-scale framework with unsupervised joint training of convolutional neural networks for pulmonary deformable image registration. *Physics in Medicine & Biology*, 65(1):015011, January 2020. doi: 10.1088/1361-6560/ab5da0. URL `https://doi.org/10.1088/1361-6560/ab5da0`.

[113] VIJAYSINH LENDAVE. What is a convolutional layer?, Jun 2021. URL `https://analyticsindiamag.com/what-is-a-convolutional-layer/`.

[114] Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8689 LNCS(PART 1): 818–833, 2014. ISSN 16113349. doi: 10.1007/978-3-319-10590-1_53.

[115] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-December:770–778, 2016. ISSN 10636919. doi: 10.1109/CVPR.2016.90.

[116] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pages 1–14, 2015.

[117] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. doi: 10.1109/CVPR.2009.5206848.

[118] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015. doi: 10.1007/s11263-015-0816-y.

[119] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *IEEE Access*, 9:16591–16603, may 2015.

ISSN 2169-3536. doi: 10.1109/ACCESS.2021.3053408. URL https://ieeexplore.ieee.org/document/9330594/http://arxiv.org/abs/1505.04597.

[120] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. *Communications of the ACM*, 63(11):139–144, jun 2014. ISSN 15577317. doi: 10.1145/3422622. URL http://arxiv.org/abs/1406.2661.

[121] Thalles Silva. An intuitive introduction to generative adversarial networks (gans), Jan 2018. URL https://www.freecodecamp.org/news/an-intuitive-introduction-to-generative-adversarial-networks-gans-7a2264a81394.

[122] George Rodrigues, Michael Lock, David D'Souza, Edward Yu, and Jake Van Dyk. Prediction of radiation pneumonitis by dose-volume histogram parameters in lung cancer - A systematic review. *Radiotherapy and Oncology*, 71(2):127–138, 2004. ISSN 01678140. doi: 10.1016/j.radonc.2004.02.015.

[123] T.J. Patton. *Quantifying and Modeling Radiation Therapy-induced Ventilation Changes and Investigating the Robustness of 4DCT-based Functional Avoidance*. PhD thesis, University of Wisconsin-Madison, 2018.

[124] Eric M. Wallat, Antonia E. Wuschner, Mattison J. Flakus, Gary E. Christensen, Joseph M. Reinhardt, Dhanansayan Shanmuganayagam, and John E. Bayouth. Radiation-induced airway changes and downstream ventilation decline in a swine model. *Biomedical Physics and Engineering Express*, 7(6), 2021. ISSN 20571976. doi: 10.1088/2057-1976/ac3197.

[125] J. W. Hopewell, M. Rezvani, and H. F. Moustafa. The pig as a model for the study of radiation effects on the lung. *International Journal of Radiation Biology*, 76(4): 447–452, 2000. doi: 10.1080/095530000138439. URL https://doi.org/10.1080/095530000138439.

[126] Jong Geol Lee, Sunhoo Park, Chang Hwan Bae, Won Suk Jang, Sun Joo Lee, Dal Nim Lee, Jae Kyung Myung, Cheol Hyeon Kim, Young Woo Jin, Seung Sook Lee, and Sehwan Shim. Development of a minipig model for lung injury induced by a single high-dose radiation exposure and evaluation with thoracic computed tomography. *Journal of Radiation Research*, 57(3):201–209, 2016. ISSN 13499157. doi: 10.1093/jrr/rrv088.

[127] Jacqueline P. Williams, Stephen L. Brown, George E. Georges, Martin Hauer-Jensen, Richard P. Hill, Amy K. Huser, David G. Kirsch, Thomas J. MacVittie, Kathy A. Mason, Meetha M. Medhora, John E. Moulder, Paul Okunieff, Mary F. Otterson, Michael E. Robbins, James B. Smathers, and William H. McBride. Animal models for medical countermeasures to radiation exposure. *Radiation Research*, 173(4):557–578, 2010. ISSN 00337587. doi: 10.1667/RR1880.1.

[128] Dominic T. Schomberg, Armando Tellez, Jennifer J. Meudt, Dane A. Brady, Krista N. Dillon, Folagbayi K. Arowolo, Joan Wicks, Serge D. Rousselle, and Dhanansayan Shanmuganayagam. Miniature swine for preclinical modeling of complexities of human disease for translational scientific discovery and accelerated development of therapies and medical devices. *Toxicologic Pathology*, 44(3):299–314, 2016. doi: 10.1177/0192623315618292. URL `https://doi.org/10.1177/0192623315618292`. PMID: 26839324.

[129] Antonia E. Wuschner, Eric M. Wallat, Mattison J. Flakus, Dhanansayan Shanmuganayagam, Jennifer Meudt, Gary E. Christensen, Joseph M. Reinhardt, Jessica R. Miller, Michael J. Lawless, Andrew M. Baschnagel, and John E. Bayouth. Radiation-induced Hounsfield unit change correlates with dynamic CT perfusion better than 4DCT-based ventilation measures in a novel-swine model. *Scientific Reports*, 11(1):1–14, 2021. ISSN 2045-2322.

[130] Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka, John Buatti, Stephen Aylward, James V. Miller, Steve Pieper, and Ron Kikinis. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging*, 30 (9):1323–1341, nov 2012. ISSN 0730725X. doi: 10.1016/j.mri.2012.05.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0730725X12001816.

[131] Raúl San José Estépar, George G. Washko, Edwin K. Silverman, John J. Reilly, Ron Kikinis, and Carl-Fredrik Westin. Accurate airway wall estimation using phase congruency. In Rasmus Larsen, Mads Nielsen, and Jon Sporring, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, pages 125–134, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. ISBN 978-3-540-44728-3.

[132] Yasutaka Nakano, Shigeo Muro, Hiroaki Sakai, Toyohiro Hirai, Kazuo Chin, Mitsuhiro Tsukino, Koichi Nishimura, Harumi Itoh, Peter D. Pare, James C. Hogg, and Michiaki Mishima. Computed tomographic measurements of airway dimensions and emphysema in smokers. *American Journal of Respiratory and Critical Care Medicine*, 162(3):1102–1108, 2000. doi: 10.1164/ajrccm.162.3.9907120. URL https://doi.org/10.1164/ajrccm.162.3.9907120. PMID: 10988137.

[133] Shin Matsuoka, Yasuyuki Kurihara, Kunihiro Yagihashi, Makoto Hoshino, and Yasuo Nakajima. Airway dimensions at inspiratory and expiratory multisection ct in chronic obstructive pulmonary disease: Correlation with airflow limitation. *Radiology*, 248(3):1042–1049, 2008. doi: 10.1148/radiol.2491071650. URL https://doi.org/10.1148/radiol.2491071650. PMID: 18710993.

[134] M Hasegawa, H Makita, Y Nasuhara, N Odajima, K Nagai, Y Ito, T Betsuyaku, and M Nishimura. Relationship between improved airflow limitation and changes

in airway calibre induced by inhaled anticholinergic agents in copd. *Thorax*, 64 (4):332–338, 2009. ISSN 0040-6376. doi: 10.1136/thx.2008.103671. URL https://thorax.bmj.com/content/64/4/332.

[135] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL https://www.R-project.org/.

[136] J. Martin Bland and Douglas G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476):307–310, 1986. ISSN 0140-6736. doi: https://doi.org/10.1016/S0140-6736(86)90837-8. URL https://www.sciencedirect.com/science/article/pii/S0140673686908378. Originally published as Volume 1, Issue 8476.

[137] L. Freitag, A. Ernst, M. Unger, K. Kovitz, and C. H. Marquette. A proposed classification system of central airway stenosis. *European Respiratory Journal*, 30 (1):7–12, 2007. ISSN 09031936. doi: 10.1183/09031936.00132804.

[138] Masaru Hasegawa, Yasuyuki Nasuhara, Yuya Onodera, Hironi Makita, Katsura Nagai, Satoshi Fuke, Yoko Ito, Tomoko Betsuyaku, and Masaharu Nishimura. Airflow limitation and airway dimensions in chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine*, 173(12):1309–1315, 2006. doi: 10.1164/rccm.200601-037OC. URL https://doi.org/10.1164/rccm.200601-037OC. PMID: 16556695.

[139] A. Jackson, C. Wang, E. Yorke, D. Geldblum, A. Apte, J. Yang, A. Rimner, and A. Wu. MA02.06 Dose-Volume Factors Predicting Airway Stenosis After SBRT for Ultra-Central Lung Tumors. *Journal of Thoracic Oncology*, 14(10):S253–S254, October 2019. ISSN 15560864. doi: 10.1016/j.jtho.2019.08.506. URL https://linkinghub.elsevier.com/retrieve/pii/S155608641931189X.

[140] Juliët E. van Hoorn, Max Dahele, and Johannes M. A. Daniels. Late Central Airway Toxicity after High-Dose Radiotherapy: Clinical Outcomes and a Proposed Bronchoscopic Classification. *Cancers*, 13(6):1313, March 2021. ISSN 2072-6694. doi: 10.3390/cancers13061313. URL https://www.mdpi.com/2072-6694/13/6/1313.

[141] J Bayouth, T Patton, S Gerard, G Christensen, A Baschnagel, and J Reinhardt. Reducing patient respiration induced 4dct image artifacts. *MEDICAL PHYSICS*, 44(6):3005 – 3006, 2017. ISSN 2473-4209.

[142] Kaifang Du, Joseph M. Reinhardt, Gary E. Christensen, Kai Ding, and John E. Bayouth. Respiratory effort correction strategies to improve the reproducibility of lung expansion measurements. *Medical Physics*, 40(12):1–15, 2013. ISSN 00942405. doi: 10.1118/1.4829519.

[143] Benjamin M. Smith, Eric A. Hoffman, Dan Rabinowitz, Eugene Bleecker, Stephanie Christenson, David Couper, Kathleen M. Donohue, Meilan K. Han, Nadia N. Hansel, Richard E. Kanner, Eric Kleerup, Stephen Rennard, and R. Graham Barr. Comparison of spatially matched airways reveals thinner airway walls in COPD. The Multi-Ethnic Study of Atherosclerosis (MESA) COPD Study and the Subpopulations and Intermediate Outcomes in COPD Study (SPIROMICS). *Thorax*, 69(11):987–996, November 2014. ISSN 0040-6376, 1468-3296. doi: 10.1136/thoraxjnl-2014-205160. URL https://thorax.bmj.com/content/69/11/987. Publisher: BMJ Publishing Group Ltd Section: Chronic obstructive pulmonary disease.

[144] Harvey O. Coxson, Brendan Quiney, Don D. Sin, Li Xing, Annette M. McWilliams, John R. Mayo, and Stephen Lam. Airway Wall Thickness Assessed Using Computed

Tomography and Optical Coherence Tomography. *American Journal of Respiratory and Critical Care Medicine*, 177(11):1201–1206, June 2008. ISSN 1073-449X. doi: 10.1164/rccm.200712-1776OC. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2408438/.

[145] Eric M. Wallat, Antonia E. Wuschner, Mattison J. Flakus, Sarah E. Gerard, Gary E. Christensen, Joseph M. Reinhardt, and John E. Bayouth. Predicting pulmonary ventilation damage after radiation therapy for nonsmall cell lung cancer using a resnet generative adversarial network. *Medical Physics*, n/a(n/a), 2023. doi: https://doi.org/10.1002/mp.16311. URL https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.16311.

[146] J. Kipritidis, M.S. Hofman, S. Siva, J. Callahan, P. Le Roux, H.C. Woodruff, W.B. Counter, and P.J. Keall. Estimating lung ventilation directly from 4d ct hounsfield unit values. *Medical Physics*, 43(1):33–43, 2016. doi: 10.1118/1.4937599. URL https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4937599.

[147] J.M. Reinhardt, K. Ding, K. Cao, G.E. Christensen, E.A. Hoffman, and S.V. Bodas. Registration-based estimates of local lung tissue expansion compared to xenon ct measures of specific ventilation. *Medical Image Analysis*, 12(6):752 – 763, 2008. ISSN 1361-8415. doi: https://doi.org/10.1016/j.media.2008.03.007. URL http://www.sciencedirect.com/science/article/pii/S1361841508000303. Special issue on information processing in medical imaging 2007.

[148] D. Han, J.E. Bayouth, S. Bhatia, M. Sonka, and X. Wu. Characterization and identification of spatial artifacts during 4d-ct imaging. *Medical Physics*, 38(4): 2074–2087, 2011. doi: 10.1118/1.3553556. URL https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.3553556.

[149] K. Cao, K. Ding, J.M. Reinhardt, and G.E. Christensen. Improving intensity-based lung ct registration accuracy utilizing vascular information. *International Journal of Biomedical Imaging*, 2012. doi: doi:10.1155/2012/285136.

[150] K. Du, J.E. Bayouth, K. Ding, G.E. Christensen, K. Cao, and J.M. Reinhardt. Reproducibility of intensity-based estimates of lung ventilation. *Medical Physics*, 40(6Part1):063504, 2013. doi: 10.1118/1.4805106. URL `https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4805106`.

[151] W. Shao, T.J. Patton, S.E. Gerard, Y. Pan, J.M. Reinhardt, O.C. Durumeric, J.E. Bayouth, and G.E. Christensen. N-phase local expansion ratio for characterizing out-of-phase lung ventilation. *IEEE TRANS MED IMAGING*, 2019.

[152] S. Yang and G. Berdine. Model selection and model over-fitting. *The Southwest Respiratory and Critical Care Chronicles*, 3(12):52–55, 2015.

[153] M.J. Flakus, A.E. Wuschner, E.M. Wallat, W. Shao, S.E. Gerard, G.E. Christensen, J.M. Reinhardt, and J.E. Bayouth. Improving the accuracy of 4dct-based ventilation measurements using multiple phases. *Medical Physics*, 46:E378, 2019.

[154] S. Ramella, L. Trodella, T.C. Mineo, E. Pompeo, G. Stimato, D. Gaudino, V. Valentini, F. Cellini, M. Ciresa, M. Fiore, A. Piermattei, P. Russo, A. Cesario, and R.M. D'Angelillo. Adding ipsilateral v20 and v30 to conventional dosimetric constraints predicts radiation pneumonitis in stage iiia–b nsclc treated with combined-modality therapy. *International Journal of Radiation Oncology*Biology*Physics*, 76(1):110 – 115, 2010. ISSN 0360-3016. doi: https://doi.org/10.1016/j.ijrobp.2009.01.036. URL `http://www.sciencedirect.com/science/article/pii/S0360301609001321`.

[155] M.V. Graham, J.A. Purdy, B. Emami, W. Harms, W. Bosch, M.A. Lockett, and C.A. Perez. Clinical dose–volume histogram analysis for pneumonitis after 3d treatment for non-small cell lung cancer (nsclc). *International Journal of Radiation Oncology*Biology*Physics*, 45(2):323 – 329, 1999. ISSN 0360-3016. doi: https://doi.org/10.1016/S0360-3016(99)00183-2. URL `http://www.sciencedirect.com/science/article/pii/S0360301699001832`.

[156] T. J. Patton, S. E. Gerard, W. Shao, G. E. Christensen, J. M. Reinhardt, and J. E. Bayouth. Quantifying ventilation change due to radiation therapy using 4dct jacobian calculations. *Medical Physics*, 45(10):4483–4492, 2019. doi: 10.1002/mp.13105.

[157] J.E. Bayouth, T.J. Patton, S.E. Gerard, W. Shao, G.E. Christensen, J.M. Reinhardt, and A.M. Baschnagel. Signal processing analysis of breathing rates show improved spectral coherence when human subjects receive musical melody and voice instruction guidance. *Medical Physics*, 2018.

[158] Sarah E. Gerard, Jacob Herrmann, Yi Xin, Kevin T. Martin, Emanuele Rezoagli, Davide Ippolito, Giacomo Bellani, Maurizio Cereda, Junfeng Guo, Eric A. Hoffman, David W. Kaczka, and Joseph M. Reinhardt. CT image segmentation for inflamed and fibrotic lungs using a multi-resolution convolutional neural network. *Scientific Reports*, 11(1):1455, January 2021. ISSN 2045-2322. doi: 10.1038/s41598-020-80936-4. URL `https://www.nature.com/articles/s41598-020-80936-4`. Number: 1 Publisher: Nature Publishing Group.

[159] Sarah E. Gerard, Jacob Herrmann, David W. Kaczka, Guido Musch, Ana Fernandez-Bustamante, and Joseph M. Reinhardt. Multi-resolution convolutional

neural networks for fully automated segmentation of acutely injured lungs in multiple species. *Medical Image Analysis*, 60:101592, February 2020. ISSN 1361-8415. doi: 10.1016/j.media.2019.101592. URL https://www.sciencedirect.com/science/article/pii/S136184151930132X.

[160] V Badrinarayanan, A Kendall, and R Cipolla. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12):2481–2495, 2017. ISSN 1939-3539 VO - 39. doi: 10.1109/TPAMI.2016.2644615.

[161] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, 2015. URL https://arxiv.org/abs/1502.01852.

[162] Wenqi Li, Guotai Wang, Lucas Fidon, Sebastien Ourselin, M. Jorge Cardoso, and Tom Vercauteren. On the compactness, efficiency, and representation of 3D convolutional networks: Brain parcellation as a pretext task. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10265 LNCS:348–360, 2017. ISSN 16113349. doi: 10.1007/978-3-319-59050-9_28.

[163] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks, 2016. URL https://arxiv.org/abs/1611.07004.

[164] W. J. Youden. Index for rating diagnostic tests. *Cancer*, 3(1):32–35, 1950. doi: https://doi.org/10.1002/1097-0142(1950)3:1⟨32::AID-CNCR2820030106⟩3.0.CO;2-3.

[165] Jason St-Hilaire, Caroline Lavoie, Anne Dagnault, Frédéric Beaulieu, Francis Morin, Luc Beaulieu, and Daniel Tremblay. Functional avoidance of lung in plan optimization with an aperture-based inverse planning system. *Radiotherapy and Oncology*, 100(3):390–395, September 2011. ISSN 0167-8140. doi: 10.1016/j.radonc.2011.09.003. URL https://www.sciencedirect.com/science/article/pii/S0167814011005287.

[166] Linda G. W. Kerkmeijer, Veerle H. Groen, Floris J. Pos, Karin Haustermans, Evelyn M. Monninkhof, Robert Jan Smeenk, Martina Kunze-Busch, Johannes C. J. de Boer, Jochem van der Voort van Zijp, Marco van Vulpen, Cédric Draulans, Laura van den Bergh, Sofie Isebaert, and Uulke A. van der Heide. Focal Boost to the Intraprostatic Tumor in External Beam Radiotherapy for Patients With Localized Prostate Cancer: Results From the FLAME Randomized Phase III Trial. *Journal of Clinical Oncology*, 39(7):787–796, March 2021. ISSN 0732-183X. doi: 10.1200/JCO.20.02873. URL https://ascopubs.org/doi/10.1200/JCO.20.02873. Publisher: Wolters Kluwer.

[167] Judi van Diessen, Dirk De Ruysscher, Jan-Jakob Sonke, Eugène Damen, Karolina Sikorska, Bart Reymen, Wouter van Elmpt, Gunnar Westman, Gitte Fredberg Persson, Edith Dieleman, Hedvig Bjorkestrand, Corinne Faivre-Finn, and José Belderbos. The acute and late toxicity results of a randomized phase II dose-escalation trial in non-small cell lung cancer (PET-boost trial). *Radiotherapy and Oncology*, 131:166–173, February 2019. ISSN 0167-8140. doi: 10.1016/j.radonc.2018.09.019. URL https://www.sciencedirect.com/science/article/pii/S0167814018335023.

[168] Steven F. Petit, André L. A. J. Dekker, Renaud Seigneuric, Lars Murrer, Natal A. W. van Riel, Marianne Nordsmark, Jens Overgaard, Philippe Lambin, and Bradly G. Wouters. Intra-voxel heterogeneity influences the dose prescription for

dose-painting with radiotherapy: a modelling study. *Physics in Medicine and Biology*, 54(7):2179–2196, March 2009. ISSN 0031-9155. doi: 10.1088/0031-9155/54/7/022. URL https://doi.org/10.1088/0031-9155/54/7/022. Publisher: IOP Publishing.

[169] Ahmed Fawzy Gad. Pygad: An intuitive genetic algorithm python library, 2021.

[170] Yvette Seppenwoolde, Martijn Engelsman, Katrien De Jaeger, Sara H. Muller, Paul Baas, Daniel L. McShan, Benedick A. Fraass, Marc L. Kessler, José S. A. Belderbos, Liesbeth J. Boersma, and Joos V. Lebesque. Optimizing radiation treatment plans for lung cancer using lung perfusion information. *Radiotherapy and Oncology*, 63(2): 165–177, May 2002. ISSN 0167-8140. doi: 10.1016/S0167-8140(02)00075-0. URL https://www.sciencedirect.com/science/article/pii/S0167814002000750.

[171] Gazi Md Daud Iqbal, Hao Zhang, Wareen D'Souza, Lidan Ha, and Jay M. Rosenberger. Four-dimensional computed tomography-based ventilation imaging in intensity-modulated radiation therapy treatment planning for pulmonary functional avoidance. *Journal of Applied Clinical Medical Physics*, page e13920, February 2023. ISSN 1526-9914. doi: 10.1002/acm2.13920. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/acm2.13920. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/acm2.13920.

[172] L. J. Rankine, Z. Wang, E. Bier, C. R. Kelsey, L. B. Marks, B. Driehuys, and S. K. Das. Pulmonary Gas Exchange-guided Functional Avoidance Treatment Planning for Thoracic Radiation Therapy Using Hyperpolarized 129Xe Magnetic Resonance Imaging. *International Journal of Radiation Oncology, Biology, Physics*, 108 (3):S102–S103, November 2020. ISSN 0360-3016. doi: 10.1016/j.ijrobp.2020.07.2281. URL https://www.redjournal.org/article/S0360-3016(20)33700-7/fulltext. Publisher: Elsevier.

[173] Sophia C. Kamran, Birgit S. Mueller, Peter Paetzold, Joseph Dunlap, Andrzej Niemierko, Thomas Bortfeld, Henning Willers, and David Craft. Multi-criteria optimization achieves superior normal tissue sparing in a planning study of intensity-modulated radiation therapy for RTOG 1308-eligible non-small cell lung cancer patients. *Radiotherapy and Oncology*, 118(3):515–520, March 2016. ISSN 0167-8140. doi: 10.1016/j.radonc.2015.12.028. URL https://www.sciencedirect.com/science/article/pii/S016781401600030X.

[174] Davood Hajinezhad, Afshin Oroojlooy, Mohammadreza Nazari, Xin Hunt, Jorge Silva, Colette Shen, Bhisham Chera, and Shiva K. Das. Machine learning to generate adjustable dose distributions in head-and-neck cancer radiation therapy, 2022. URL https://arxiv.org/abs/2207.07168.

[175] Chulin Wu, Heye Zhang, Jiaqi Chen, Zhifan Gao, Pengfei Zhang, Khan Muhammad, and Javier Del Ser. Vessel-gan: Angiographic reconstructions from myocardial ct perfusion with explainable generative adversarial networks. *Future Generation Computer Systems*, 130:128–139, 2022. ISSN 0167-739X. doi: https://doi.org/10.1016/j.future.2021.12.007. URL https://www.sciencedirect.com/science/article/pii/S0167739X2100488X.