UNIVERSITY OF COPENHAGEN

DEPARTMENT OF COMPUTER SCIENCE
IMAGE SECTION

Cardiac Dose Trends & Toxicity in Radiotherapy:

Evaluating Registry, Dosimetry, and Imaging Data with AI and Statistical Methods

PhD Thesis

Nora Jarrett Forbes

Chairperson **Erik Bjørnager Dam**
Department of Computer Science
University of Copenhagen (KU)

Reviewers **Marianne C. Aznar**
Division of Cancer Sciences
The University of Manchester

Anders Nymark Christensen
Department of Applied Mathematics and Computer Science
Technical University of Denmark (DTU)

Supervisors **Ivan Vogelius** Department of Clinical Medicine/Oncology
Sune Darkner Department of Computer Science
Jens Petersen Department of Computer Science
Joanne Reekie CHIP/PERSIMUNE

31st of July, 2024

This thesis has been submitted to the PhD School of The Faculty of Science, University of Copenhagen.

Nora Jarrett Forbes
Cardiac Dose Trends & Toxicity in Radiotherapy:

Assessment Committee: Erik B. Dam, Marianne C. Aznar, and Anders N. Christensen
Supervisors: Ivan Vogelius, Sune Darkner, Jens Petersen, and Joanne Reekie
31st of July, 2024

University of Copenhagen
Department of Computer Science
Image Section
Universitetsparken 1
2100 Copenhagen

Abstract

Objectives

The research focuses on the impact of radiation therapy (RT) on cardiac health, particularly by examining how advancements in RT technology and increased awareness of cardiotoxic risks have affected cardiac dose levels over time. The research examines the relationship between cardiac radiation doses, patient-specific factors, and long-term outcomes, including cardiovascular disease (CVD) and overall survival.

Methods

The studies analyzed data from a cohort of over 10,000 patients treated for various cancers (breast, lung, lymphoma, and esophageal) between 2009 and 2020. Artificial intelligence (AI) segmentation tools were used to delineate heart structures from CT scans and calculate dose metrics such as mean heart dose (MHD) and volumetric measures of these structures. The research also used statistical methods, including the Aalen-Johansen estimator and Cox proportional hazards, to assess the relationship between radiation doses, patient characteristics, and adverse outcomes collected from registry-based data.

Results

Study I found a significant decrease in cardiac doses over the study period, particularly in high-dose exposures. However, a recent increase in high-dose volumes for breast cancer treatments was noted. **Study II** revealed that patient-specific factors, such as pre-existing cardiac conditions and age, were stronger predictors of cardiovascular issues than radiation dose itself, with no clear dose-response relationship established with CVD.

Conclusions Results highlight the potential of AI tools to generate individually defined structures and calculate precise dose metrics, as well as the value of linking registry databases to obtain reliable results for a substantial retrospective cohort. We can conclude that advances in RT techniques and increased clinical awareness have generally led to improved heart sparing. It is still important to spare heart tissue, but tumor control should be prioritized. Additional analysis, such as substructure segmentation or so-called image-based "data mining techniques", seems to be a promising direction for further understanding of cardiotoxicity and individualization of radiation therapy.

Dansk Resumé

Mål

Forskningen fokuserer på virkningen af strålebehandling (RT) på hjertesundhed, især ved at undersøge hvordan fremskridt inden for RT-teknologi og øget bevidsthed om kardiotoksiske risici har påvirket hjertedosisniveauer over tid. Studierne undersøger sammenhængen mellem hjertestråledoser, patientspecifikke faktorer og langsigtede resultater, herunder kardiotoksicitet og overlevelse.

Metoder

Undersøgelserne analyserede data fra en kohorte på over 10.000 patienter behandlet for forskellige kræftformer (bryst, lunge, lymfom og kræft i spiserør) mellem 2009 og 2020. AI-segmenteringsværktøjer blev brugt til at indtegne hjertestrukturer fra CT-scanninger og finde mål for stråledosis såsom middelhjerte dosis (MHD) og volumetriske mål for disse strukturer. Forskningen brugte også statistiske metoder, herunder Aalen-Johansen estimator og Cox proportional hazards modeller, til at vurdere sammenhængen mellem stråledoser, patientkarakteristika og bivirkninger opsamlet fra registerdata.

Resultater

Undersøgelse I fandt et signifikant fald i hjertedoser i løbet af undersøgelses- perioden, især for højdosis eksponeringer. Der blev dog noteret en nylig stigning i højdosisvolumener til brystkræftbehandlinger. **Undersøgelse II** afslørede, at patientspecifikke faktorer, såsom allerede eksisterende hjertesygdomme og alder, var stærkere forudsigere for kardiovaskulære problemer end selve stråledosen, uden at der er etableret en klar dosis-respons-relation.

Konklusioner

Undersøgelserne fremhæver potentialet i AI-værktøjer til at generere individuelt definerede strukturer og beregne præcise dosismålinger, såvel som værdien af at sammenkæde registerdatabaser for at opnå pålidelige resultater for en væsentlig ret- rospektiv kohorte. Vi kan konkludere, at fremskridt inden for RT-teknikker og øget klinisk bevidsthed generelt har ført til forbedret hjertesparing. Det er fortsat vigtigt at skåne hjertevæv, men tumorkontrol bør prioriteres. Yderligere analyser fx substruktursegmentering eller såkaldte billedbaserede “data mining-teknikker” synes at være en lovende retning for yderligere forståelse af kardiotoksicitet og individualisering af stråleterapi.

Acknowledgement

> *Alone we can do so little;*
> *together we can do so much.*
>
> **— Helen Keller**

My name is on the cover of this thesis, but this work could not have been completed without the help and support of so many lovely humans. I am deeply grateful to everyone who made this thesis possible. I will attempt to acknowledge all those who have enriched the last few years of my life and work. But please know that my appreciation is indescribable, and these words will always fall far short.

First and foremost, my heartfelt thanks go to the patients whose data contributed to this research. Although I did not interact directly with the patients, I am indebted to the clinicians who shared their invaluable insights into patient experiences, helping to bridge the gap between the data and the human stories they represent. I hope this work can contribute to the ongoing efforts to improve patient care.

Tusund tak for the patience, guidance, and inspiration from my many advisors and mentors - Ivan, Sune, Jens, Lena, Jo, and Soren. Thank you Ivan for your reliable support, especially when we encountered unexpected obstacles and moments of doubt. Thank you Sune for seeing promise in my research goals from the beginning and enabling this incredible opportunity. Thank you Jens for graciously sharing your skills and time to help generate the medical data that are the foundation of this work. I am so grateful to the team at PERSIMUNE - namely Jo, Riia, Daniel, Sadaf, Jamshed, and Jacob - for their dedication and hard work in making registry data accessible to researchers like myself. Special thanks to my coauthors Cynthia and Abraham for paving the way and involving me in their PhD work from the start.

To the RH basement and especially the "Head and Neck and Nora" office, for always making me feel included. Julie you showed me the ropes from day one and our 'mandatory meetings' with Anne always provided insights and comic relief. To be surrounded by so many brilliant and impassioned physician researchers inspires me to continue sharing my statistical proclivities in the hospital setting.

To my DIKU colleagues who have become akin to family - particularly Roberta, Renfei, Libby, Thomas, Lennard, and Marloes. Your friendship and support have been a constant source of joy and encouragement. Roberta, you've been my rock through thick and thin; Renfei, your delicious home-cooked meals kept my belly and heart warm; Libby, you've been the best and silliest little buddy ever.

To my wider community - including Kamilla, Lotte, Henriette, Tessie, Oriel, and Charlie - your companionship has made Denmark feel like home. Copenhagen, with its wonderful coffee and pastries, has been a nurturing environment for growth, even during the difficulties of the global pandemic and enduring intercontinental distance from loved ones. Cycling the Danish countryside, playing fodbold with the Pan Idræt team, invigorating winter swims at the Nordhavn Vinterbadeclub, and practicing movement and mindfulness at Rört (even becoming a saunagus meister) has kept me sane and brought me life amidst the long gray winters and through the bright and sunny summers.

Lastly, I must thank my family, whose support has been my foundation. To my parents and sisters, whose love was always felt from across the globe. And to my wife Emily - your unwavering faith in me, even when it seemed unreasonable, has stoked my belief in myself and been an invaluable source of strength... May we never let an ocean keep us apart again.

Funding and Support

Grant No. 801199
Marie Skłodowska-Curie
European Union's Horizon 2020 research and innovation programme

Danish Cancer Society

Grant No. R231-A13976
DCCC Radiotherapy

Danmarks
Grundforskningsfond
Danish National
Research Foundation

Grant No. 126
CHIP & PERSIMUNE

UNIVERSITY OF COPENHAGEN

REGION H Rigshospitalet

DCCC Radiotherapy

Danish Comprehensive Cancer Center
Radiotherapy Research Center.
Supported by The Danish Cancer Society

PERSIMUNE

CENTRE OF EXCELLENCE FOR PERSONALISED MEDICINE
OF INFECTIOUS COMPLICATIONS IN IMMUNE DEFICIENCY

chip

Centre of Excellence for
Health, Immunity and Infections

Scientific Contributions

The following are completed **first-author articles** that form the core content of this thesis. Manuscripts are included in ***Appendix A***, and medRxiv online *(see QR codes)*:

Forbes, N. J., ..., Peterson, J.*, and Vogelius, I.*; (2024-02). *Time trends in cardiac doses in 10,000 patients receiving curative thoracic radiation therapy between 2009 and 2020.* medRxiv. DOI: 10.1101/2024.02.18.24303007. [1]
****Submitted to the Journal of Clinical Oncology*

Forbes, N. J., ..., and Vogelius, I.; (2024-08). *Cardiac dose-volume analysis of 9,411 patients with registry-based outcome data for cardiotoxicity and overall survival.* medRxiv. DOI: 10.1101/2024.08.12.24. [2]
****Submitted to the Journal of Radiotherapy and Oncology*

The following are supporting **research documentation** included in this thesis and representing further contributions to this work. These can be found in ***Appendix B***:

Forbes, N., and Vogelius, I.; (2021). *Radiotherapy exposure and association with observed cardiovascular toxicity in patients treated for cancer at Rigshospitalet* [Scientific Ethics Committee Report]. National Videnskabsetisk Komité, 82427.

Forbes, N.; (2023). *Clinical and dosimetric risk factors for ischemic heart disease* [Statistical Analysis Plan]. DOI: 10.6084/m9.figshare.24247372.v1. [3]

Forbes, N., ... and Vogelius, I.; (2022). MO-0716 Radiotherapy exposure and association with observed cardiovascular toxicity in over 5000 patients. Radiotherapy and Oncology, 170, S627-S628. [4]

The following are peer-reviewed **co-authored publications** that represent collaborative contributions. These have been presented in past PhD theses and are therefore not included in this thesis, but can be found on the respective online journals:

Smith, A. G., Petersen, J., Terrones-Campos, C., Berthelsen, A. K., Forbes, N. J.,... and Vogelius, I. R.; (2022). RootPainter3D: Interactive-machine-learning enables rapid and accurate contouring for radiotherapy. Medical Physics, 49(1), 461-473. DOI: 10.1002/mp.15353. [5]

Terrones-Campos, C., Ledergerber, B., Forbes, N., ... and Vogelius, I. R.; (2023). Prediction of radiation-induced lymphopenia following exposure of the thoracic region and associated risk of infections and mortality. Clinical Oncology, 35(7), e434-e444. DOI: 10.1016/j.clon.2023.04.003. [6]

Dissemination: Below is a selection of accepted scientific presentations showcasing this work:

- Rigshospitalet Oncology Research Day *(Apr 2023 and Mar 2022)*
- Öresund Workshop on Radiotherapy *(Mar 2023 and Oct 2022)*
- DCCC Radiotherapy AI Workshop *(Dec 2022)*
- DCCC Radiotherapy Annual Scientific Meeting (Nov 2022 and 2021)
- Responsible Machine Learning in Healthcare Workshop *(Oct 2022)*
- ESTRO Annual Congress *(May 2022)*
- Center of Excellence for Personalized Medicine of Infections Complications in Immune Deficiency *(Mar 2022)*
- DCCC Cardiotoxicity Workshop *(Nov 2021)*
- DTU, DIKU & AAU Summer School on Geometric Deep Learning *(Aug 2021)*

Roles and Environments:

- ***Computer Science Department - IMAGE Section at KU:*** Here, I gained valuable insights from computer science professors and fellow PhD students. I organized monthly workshops to facilitate knowledge sharing and contributed statistical and clinical insights to the team.
- ***Oncology Department - Radiation Research Group at Rigshospitalet:*** As one of the few non-clinicians in the group, I expanded my clinical knowledge while serving as a computational resource and providing statistical expertise to my colleagues.
- ***PERSIMUNE at CHIP:*** In this interdisciplinary environment, I contributed my expertise in medical registry and EHR data to help clean and structure the data warehouse.

Mentoring: I co-advised two MS computer science students focused on automated assessment of tumor regression during radiation treatment using cone beam computed tomography.

Unpublished Works:

- I collaborated with another CS PhD fellow pursing methods in neural network saliency and explainability. Unfortunately, this work did not provide publishable results.
- I collaborated with faculty in the department of Neuroscience to assess impacts of meditation on mental health during the COVID-19 pandemic. The scope of this work fell outside my PhD aims and was therefore handed off prior to publication. (But thank you Ron for providing such convincing data that I simply had to start meditating.)

Contents

Reading Guide

This thesis consists of eight chapters, each contributing to an in-depth exploration of the impact of radiation therapy (RT) on cardiac health using modern computational resources on a large cohort. The chapters are organized for either sequential or selective reading to fit the reader's interest and background. The first three chapters introduce the relevant background information, objectives, and data materials of this research. The next three chapters summarize each of the two completed studies and provide a discussion of how this work contributes to existing research. The next chapter describes the ongoing and prospective research that may follow. And finally, I tie this work together in a brief conclusion.

Chapter 1 - Background is made up of six sections. The first three sections detail clinical aspects including cancer prevalence and treatment modalities, RT objectives and techniques, and potential risks to cardiac health. The last three sections delve into the technical methods relevant to this data, including segmentation techniques, observational study design, and statistical methods in survival analysis.

Chapter 2 - Objectives describes the motivation for this research, outlining relevant literature and emphasizing the importance of studying the cardiotoxic effects of RT in cancer patients. It presents the research aims and touches on the potential clinical relevance that such data-driven insights can have on patient care.

Chapter 3 - Data Materials covers the important and arduous process of data collection and cleaning, including:

- **Identifying** the patient population through treatment codes
- **Extracting** associated image and 3D dosimetry data from the record and verify system
- **Segmenting** whole heart and cardiac substructures on 14000+ CT scans with open-source artificial intelligence tools
- **Validating** these models through a visual multi-step review process
- **Calculating** and converting dose metrics including mean heart dose (MHD) and volumetric measures.
- **Navigating** national restrictions to patient protected data and proceeding through the process of acquiring data approvals
- **Linking** dose data to demographic characteristics, cardiotoxicity events (via ICD codes from the LPR database), and mortality events (from the CPR registry)
- **Cleaning** the real-world data, merging multi-modal datasets, and identifying inconsistencies and outliers

Chapters 4 & 5 gives a brief overview of each study within the thesis. The full manuscripts are available in Appendix A and supporting documents can be found in Appendix B:

- **Study I:** *Time trends in cardiac doses in 10,000 patients receiving curative thoracic radiation therapy between 2009 and 2020*
- **Study II:** *Cardiac dose-volume analysis of 9,411 patients with registry-based outcome data for cardiotoxicity and overall survival*

Chapter 6 - Discussion provides a broader perspective on this work. It summarizes the key findings from the two studies and breaks down their implications. Additionally, I situate the research within the broader literature and explore its clinical and technical impact.

Chapter 7 - Prospective Work outlines potential directions for future research, highlighting promising avenues for further studies. It includes a preliminary analysis of ongoing cardiac substructure dose trends and explores opportunities for continued monitoring, cohort expansion, more detailed analyses, and the incorporation of additional data.

Finally, **Chapter 8 - Conclusion** provides a brief final word on the work and the direction of the field as a whole.

The Audience. *This PhD thesis is a multidisciplinary effort that integrates computer science tools for image analysis with statistical methods applied to large registry databases in RT. Consequently, I anticipate a diverse audience with varied expertise and have tailored the thesis accordingly.*

Nomenclature

AI Artificial Intelligence

ART Adaptive Radiotherapy

CAD Coronary Artery Disease

CAC Coronary Artery Calcium

$CBCT$ Cone-Beam Computed Tomography

CCI Charlson Comorbidity Index

CIF Cumulative Incidence Function

CI Confidence Interval

$CNNs$ Convolutional Neural Networks

CPR Central Person Register

CT Computed Tomography

CTV Clinical Target Volume

CVD Cardiovascular Disease

$DCCC$ Danish Comprehensive Cancer Center

$DBCG - IMN$ Danish Breast Cancer Group Internal Mammary Node

$DIBH$ Deep Inspiration Breath Hold

$DICOM$ Digital Imaging and Communications in Medicine

DL Deep Learning

$DMCG$ Danish Multidisciplinary Cancer Groups

DVH Dose-Volume Histogram

EHR Electronic Health Record

$EQD2$ Equivalent Dose in 2 Gy Fractions

$ESTRO$ European Society for Radiation Oncology

$GDPR$ General Data Protection Regulation

GTV Gross Tumor Volume

Gy Gray (unit of radiation dose)

$HER2$ Human Epidermal Growth Factor Receptor 2

HF Heart Failure

$IBDM$ Image-Based Data Mining

$ICDCodes$ International Classification of Diseases Codes

IHD Ischemic Heart Disease

$IGRT$ Image-Guided Radiotherapy

$IMRT$ Intensity-Modulated Radiation Therapy

KM Kaplan-Meier

$LASSO$ Least Absolute Shrinkage and Selection Operator

LPR Landspatientregisteret (National Patient Register)

MHD Mean Heart Dose

$MR - Linac$ Magnetic Resonance Linear Accelerator

ML Machine Learning

$NIfTI$ Neuroimaging Informatics Technology Initiative

$NSCLC$ Non-Small Cell Lung Cancer

NVK Danish National Committee on Health Research Ethics

OAR Organs at Risk

OS Overall Survival

PCA Principal Component Analysis

PCI Prophylactic Cranial Irradiation

PHM Proportional-Hazards Model

PTV Planning Target Volume

$QUANTEC$ Quantitative Analysis of Normal Tissue Effects in the Clinic

RT Radiotherapy

$RTTs$ Radiation Therapy Technologists

SAP Statistical Analysis Plan

$SCLC$ Small Cell Lung Cancer

TNM Tumor size (T), nodal involvement (N), and metastasis (M) Staging

VD Valvular Disease

$V5$ Heart volume receiving at least 5 Gy

$V30$ Heart volume receiving at least 30 Gy

$VMAT$ Volumetric Modulated Arc Therapy

VMK Danish National Medical Research Ethics Committee

$3D - CRT$ Three-Dimensional Conformal Radiotherapy

1 Background

> *The best thing about being a statistician is that you get to play in everyone's backyard.*
>
> — **John Tukey**

The following chapter provides essential background information to establish some context for the presented work. The research in this thesis spans multiple scientific fields, necessitating a multidisciplinary approach. For computer scientists, understanding the clinical elements (cancer, radiation therapy (RT), and cardiac disease) is beneficial to grasp the overall impact of the work. Likewise, a description of the more technical analysis methods (segmentation, study design, and statistical methods) is presented for the medical audience. A thorough reading of these topics may not be necessary to understand the research content, but this framing may help clarify the motivation behind the work and emphasize its significance.

1.1 Cancer

1.1.1 Demographics

In 2022, approximately 20 million new cancer cases were reported worldwide [7]. This volume is projected to rise to 35 million by 2050, representing a 75% increase [8]. This substantial increase is not solely attributable to population growth but also reflects rising age-standardized incidence rates. Furthermore, improvements in cancer screening programs and diagnostic methods contribute to more frequent diagnoses of cancer and improved prognoses.

Globally cancer is responsible for 9.7 million deaths per year, accounting for about one in six deaths and ranking second in leading causes of death. However, cancer survival has improved substantially due to advancements in prevention strategies, early detection programs, and treatment options. Currently, an estimated 53.5 million cancer patients survive five years after diagnosis [7].

The growing number of cancer patients and improvements in survival rates indicate an increased need to address the chronic side effects of cancer treatment [9]. To provide the best patient care and achieve the best long-term outcomes for patients, it is essential to thoroughly understand the associated risks.

Figure 1.1: **Continuum of Cancer Care.** Source: National Cancer Strategy, 2017-2026 [12]

1.1.2 Cancer Care Continuum

A comprehensive cancer care approach includes prevention, early detection, diagnosis, treatment, and transitions to either survivorship or end-of-life care [10, 11] (Fig. 1.1). Such an approach depends on a multidisciplinary team of healthcare providers working together to deliver timely interventions, manage symptoms, and enhance the patient well being. By tailoring patient care to the specific needs arising throughout their cancer journey, this model emphasizes the importance of continuous follow-up and monitoring to address late-onset complications and limit the risk of recurrence.

1.1.3 Cancer Onset

Tumors arise from abnormal cell proliferation, forming irregular masses of tissue. Benign tumors are typically not harmful and remain localized, whereas malignant tumors (cancers) can invade nearby tissues and have the capacity to spread (metastasize) into distant parts of the body.

Cancer develops through genetic mutations, either inherited or caused by random errors or carcinogen exposure. These mutations disrupt essential cellular processes like growth, division, and programmed cell death by interfering with normal cell cycle regulation through a complex set of mechanisms (Fig. 1.2) [13]. When mutations impair DNA repair mechanisms, genetic instability increases, leading to further mutations and a higher likelihood of cancer progression.

Common risk factors for cancer include behavioral factors (excesses weight, sedentary lifestyle, poor nutrition, heavy drinking, smoking, etc.), environmental exposures (radiation, carcinogenic chemicals, infections such as HPV, etc.), and individual characteristics (age, family history, etc.) [14, 15]. Understanding these causes aids in developing prevention strategies and improving early detection methods for high-risk populations.

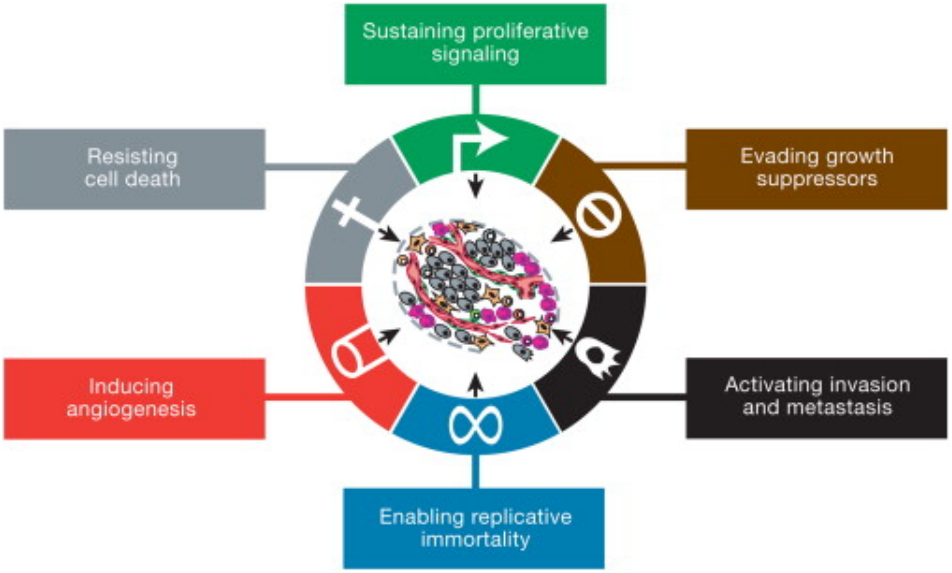


Figure 1.2: "**The Hallmarks of Cancer** - This illustration encompasses the six hallmark capabilities originally proposed in our 2000 perspective. The past decade has witnessed remarkable progress toward understanding the mechanistic underpinnings of each hallmark." Source: Hanahan et. al., 2011 [13]

1.1.4 Diagnosis and Classification

Systematic screening methods, such as mammograms, Pap smears, and colonoscopies, facilitate early detection [16]. Timely and accurate cancer diagnosis can improve the prognosis as earlier stage cancers are often less aggressive and more receptive to treatments [17]. Precise classification of cancer types, which vary in characteristics and behavior, is essential for delivering personalized care and developing effective treatment strategies [18].

Cancer diagnosis involves a set of examinations - physical assessment, imaging studies, laboratory tests, and biopsies. These tools help in identifying the type and extent of disease. Tumor classification is based on various factors, each influencing the overall treatment approach:

- **TNM Staging:** A method to describe **T**umor size, **N**odal involvement, and **M**etastasis, with disease ranging from localized (Stage I) to advanced (Stage IV) [19, 20].
- **Histology:** Microscopic characteristics of cancerous tissue vary depending on the originating cell type. Tumors are graded on the aggressiveness and extent of differentiation from normal tissue structure [21].
- **Biomarkers:** Genetic mutations and molecular phenotypes (e.g., BRCA1 / 2 and HER2 status for breast cancer) affect cancer behavior and response to treatment [22].
- **Site:** Treatment response and prognosis vary significantly based on the primary tumor's location. Each organ's unique structure and function shape how cancer develops and progresses, requiring tailored diagnostic and therapeutic approaches. Primary site data also aids in tracking epidemiological trends and implementing targeted public health measures [23].

1.1.5 Treatment Options

Available treatment options depend heavily on the cancer classification and its progression. ***Surgical removal*** of the tumor and adjacent margins remains a key technique for controlling many cancers, particularly effective for localized tumors. ***Radiation therapy*** destroys cancerous cells by damaging their DNA with high-energy radiation; this is often employed when surgery is not feasible or as an adjunct to other treatments. ***Chemotherapy*** drugs target rapidly dividing cells, making them particularly effective against widespread and metastatic cancers [24]. Cancer care strategies are continually evolving, with the field of oncology advancing rapidly. ***Modern treatment modalities*** - such as targeted therapy, stem cell therapy, hormone therapy, and immunotherapy - specifically address the biological characteristics and behaviors of tumors.

The timing of cancer therapies is crucial for optimizing outcomes. ***Neoadjuvant*** therapy (given before primary treatment like surgery) shrinks tumors and targets micrometastases. ***Adjuvant*** therapy (administered afterward) eradicates residual cancer cells to prevent recurrence. ***Concomitant*** therapy (simultaneous delivery such as chemoradiation) enhances effectiveness by combining treatments. These strategies are tailored to specific cancer types and patient needs.

1.1.6 Outcome Measures

Patient outcomes are measured to evaluate the quality of care and guide improvements in support services. The foremost concern in cancer treatment is ensuring the patient's ***survival***. Though a cancer patient may survive the initial cancer, they may experience risks of morbidity and later mortality. Nausea, vomiting, fatigue, pain, and cognitive changes are ***immediate effects*** impact quality of life, tolerance to treatment, and adherence to the prescribed plan. ***Late effects*** of the disease and the treatment can affect the long-term well-being, functioning, and quality of life of cancer survivors. Physical symptoms can include chronic pain, fatigue, neuropathy, endocrine disorders, pulmonary complications, immunosuppression, and cardiovascular issues.

1.2 Radiation Therapy

1.2.1 Objectives of Radiation Therapy

Radiation therapy has one of two primary objectives: 1) ***Palliative RT*** is used in advanced cancer cases where a cure is not feasible due to the extensive spread or aggressive nature of the disease. The primary goal of palliative RT is to alleviate symptoms (e.g. pain, bleeding, and pressure on organs) thereby improving the patient's quality of life. While treatment side effects are considered, the focus is on managing acute toxicity. These treatments typically involve lower doses and shorter treatment durations. 2) ***Curative RT*** aims to eradicate cancer cells and achieve long-term remission. This approach involves administering higher radiation doses over a more extended period, specifically targeting localized cancers to maximize cancer cell destruction while minimizing damage to surrounding healthy tissues. The success of curative RT is measured by long-term cancer remission and overall survival rates.

1.2.2 Effects and Side Effects

Curative intent RT aims to minimize damage to surrounding healthy tissues while effectively controlling tumors. For localized tumors, RT is employed to either destroy cancer cells or inhibit their ability to proliferate, essentially eradicating the cancer. Inadequate initial dosing can lead to tumor recurrence, potentially with more aggressive behavior. While increasing the radiation dose can improve tumor control, it also raises the risk of damage to organs at risk (OAR) and subsequent likelihood of long-term complications. Precision in delivery, achieved through advanced techniques, allows for higher dose delivery to the cancerous tissue while preserving normal tissues.

1.2.3 Classic Techniques

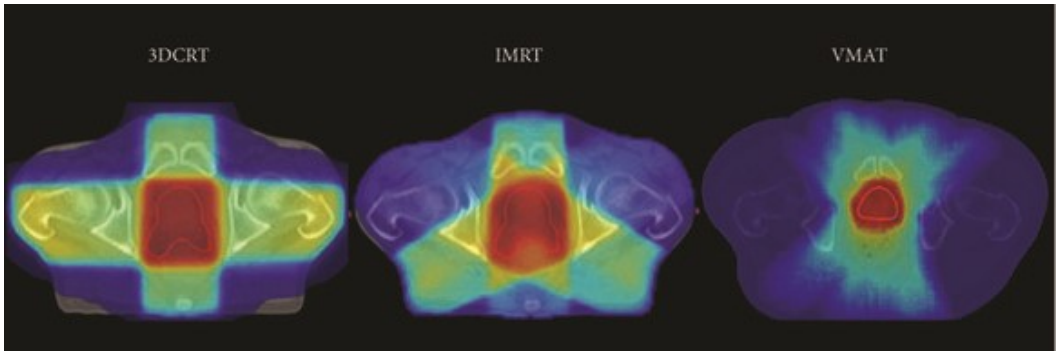


Figure 1.3: "Examples of dose distribution of a 3DCRT, IMRT-5, and VMAT treatment plan calculated on the same patient. The red surface represents the high-dose regions, the yellow surface the intermediate-high-dose regions, the dark blue surface the low-dose regions, and the azure blue surface the intermediate-dose regions." Source: Vanneste et. al., 2016 [25]

Radiation therapy for cancer treatment began in 1895 after the discovery of X-rays and became a standard practice by the mid-20th century [26, 27]. ***Three-dimensional conformal RT (3D-CRT)*** incorporated 3D imaging techniques to shape the radiation beams to better target tumors and avoid adjacent tissue. Control over radiation delivered by the beams became possible in the late 1990s with the development of ***intensity-modulated RT (IMRT)***, revolutionizing treatment by offering more precise targeting of tumors [28]. ***Volumetric modulated arc therapy (VMAT)*** further enhanced the accuracy of tumor targeting through the rotation of the linear accelerator, changing not only the intensity but also the shape of the beam during delivery. These innovations have iteratively improved the capacity to maximize tumor dose while minimizing exposure to healthy tissues (Fig. 1.3) [25].

1.2.4 Treatment Planning

Before treatment can commence, simulation and treatment planning are conducted. The treatment area is imaged with a computed tomography (CT) scan for anatomical contouring and dose planning. The radiation oncologist delineates the target areas (Fig. 1.4) and identifies critical OARs to minimize their exposure. In the thorax, these typically include the heart, lungs, esophagus, and spinal cord. [29, 30, 31].

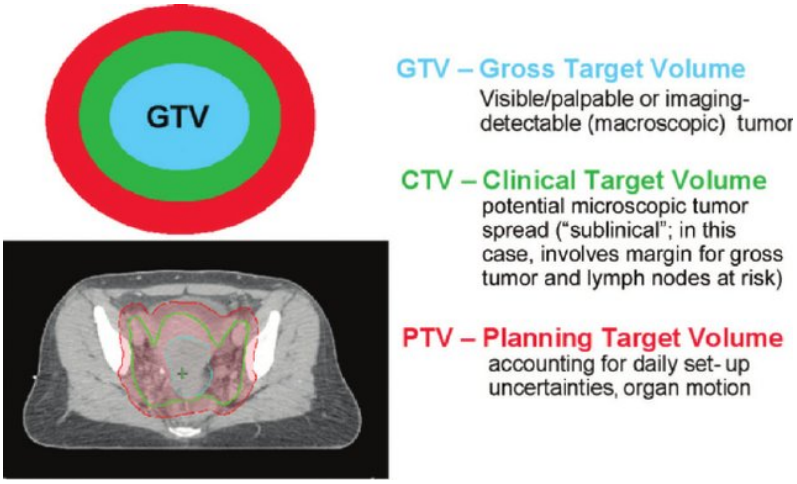


Figure 1.4: "Principles of target definition in radiation therapy planning. Diagram (top) and axial CT image (bottom) illustrate the definitions of the GTV (blue), CTV (green), and PTV (red)". Source: Xu-Welliver et. al., 2014 [32]

Radiation Therapy Technologists (RTTs) and Medical Physicists use a 3D treatment planning system, such as Varian Eclipse, to create a treatment plan. This system considers the contours, calculates the radiation dose distribution, and simulates how the radiation beams will interact with the proximal anatomy. The plan involves configuring the number, angles, and shapes of the radiation beams to optimize target dose accuracy while protecting surrounding tissues. Dose calculations are commonly performed using algorithms like the Anisotropic Analytical Algorithm (AAA) for speed and accuracy, and Acuros XB for precision in heterogeneous tissues [33]. These are improvements from Pencil Beam Convolution and Collapsed Cone Convolution calculations, but newer computationally heavy Monte Carlo Simulations offer the most accurate dose calculations.

1.2.5 Treatment Evaluation

A dose volume histogram (DVH), or cumulative dose volume frequency distribution, graphically represents the percentage of an organ at risk (OAR) or a target by dose (Fig. 1.5). DVHs are critical in the evaluation and optimization of RT plans, as they allow clinicians to assess whether the intended dose is delivered effectively to the tumor and if healthy tissue sparing is sufficient. Achieving this balance is essential for minimizing healthy tissue complication probability (NTCP) and maximizing target control probability (TCP) [35].

Within DVHs, common metrics such as Vx values (e.g., V5, V30) represent the percent of a volume of tissue receiving at least X a specified radiation dose. These values are crucial for understanding how much of an organ is exposed to potentially harmful doses. Additional key metrics derived from DVHs include Dmax, Dmean, and Dmin, which denote the maximum, mean, and minimum doses within a tissue, respectively. Equivalent Uniform Dose (EUD) provides a theoretical dose that would result in the same biological effect as the actual non-uniform dose distribution. Conformity Index (CI) and Homogeneity Index (HI) are used to measure conformity to a target volume or uniformity of dose distribution across a target volume, respectively, balancing tumor coverage and healthy tissue sparing [36].

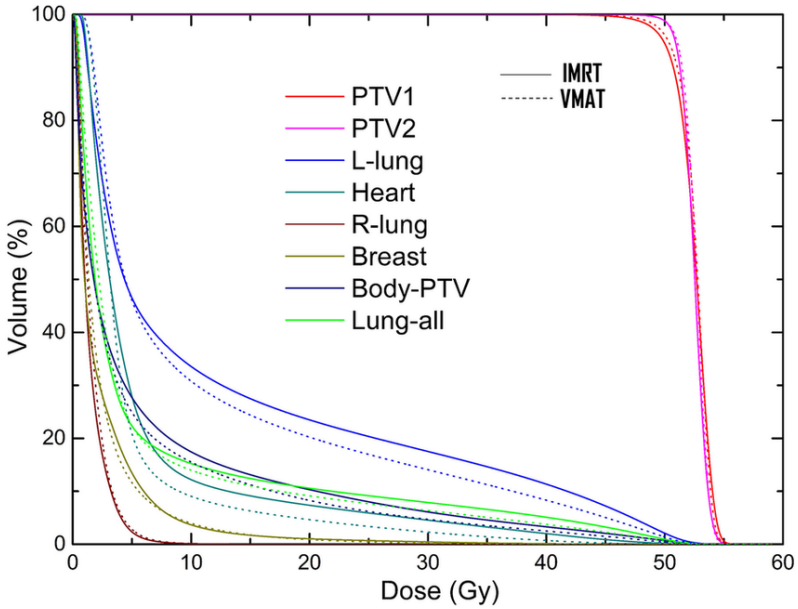


Figure 1.5: "Average dose-volume histogram (DVH) comparison for PTVs and OARs with IMRT and VMAT plans. (Solid line is IMRT and dot line is VMAT)." Source: Zhang et. al., 2021 [34]

Beyond DVHs, other evaluation techniques include direct analysis of dose distribution on CT or MRI scans, allowing for visual inspection of dose delivery across the anatomy. Isodose curves, representing lines of equal dose within the tissue, are used alongside DVHs to ensure adequate tumor coverage and appropriate dose fall-off around OARs. Metrics such as TCP and NTCP further refine the assessment of treatment success and potential side effects [36].

NTCP modeling serves as a predictive tool for estimating the likelihood of adverse effects in healthy tissues due to radiation exposure. These models quantify the association between the delivered dose and the probability of complications, enabling clinicians to optimize treatment by balancing effective tumor control with minimizing harm to surrounding tissues. By incorporating patient-specific characteristics and biological parameters, NTCP modeling contributes to safer and more personalized radiotherapy, particularly when OARs are close to the tumor [37].

1.2.6 Quality Assurance and Delivery

Once the plan is finalized, it undergoes a quality assurance process to ensure it meets clinical and safety standards. The plan is reviewed against standardized guidelines, such as those developed by multidisciplinary cancer groups (DMCGs) in Denmark [38]. Here thresholds for normal tissue exposure and the prescribed doses for the targets are defined. Standardized guidelines enhance quality, safety, and effectiveness by ensuring consistency with the latest scientific evidence and best practices. Additionally, RT guidelines support personalized treatment by providing protocols tailored to specific cancer types and stages, facilitating coordination among multidisciplinary teams involved in cancer care.

The full prescribed dose of radiation is split into smaller doses called fractions, typically delivered in daily sessions over a few weeks. Conventional RT typically delivers 1.8 to 2.2

Gy per fraction accumulating to 50 to 70 Gy in total prescribed dose over 5 to 7 weeks. This fractionation schedule maximizes treatment effectiveness while allowing normal tissue the time to repair in between sessions, reducing risks of side effects. Each session is brief, with patient positioning verified using imaging guidance, such as cone-beam CT, to ensure precise alignment before radiation delivery.

1.2.7 Advanced Techniques

Fractionation Changes

Hyperfractionation involves delivering smaller, more frequent doses, which can reduce late side effects while maintaining treatment effectiveness. This approach is particularly beneficial for cancers like those in the head and neck, where preserving surrounding tissue function is crucial. ***Hypofractionation*** administers higher doses per session (2.2 to 3.5 Gy) over a shorter period and is increasingly used for breast and prostate cancers. It offers greater convenience and emerging evidence suggests it may not increase damage to surrounding tissues as once feared. ***Stereotactic Body Radiotherapy (SBRT)***, or ***Stereotactic Radiosurgery (SRS)*** when applied cranially, delivers very high doses (6-20 Gy per fraction) in just 1 to 5 sessions. This enables precise targeting of small, well-defined tumors, such as in early-stage lung cancer, where surgery may not be an option. A novel technique, ***FLASH therapy***, is also being explored, which involves delivering a single ultra-high dose (>40 Gy) very rapidly to minimize damage to normal tissues.

Delivery Methods

While classic RT uses high-energy photons (a.k.a. X-rays) to target tumors, it is possible to use the other particles. ***Electron therapy*** has a shorter penetration than photons and can be used to treat superficial lessons [39]. ***Particle therapy*** is the use of subatomic particles and may yield potentially greater biological effectiveness against certain types of cancer [40]. ***Proton therapy*** is particularly exciting for normal tissue sparing in pediatric patients and for tumors proximal to critical structures because it now allows for precise dose distribution with minimal exit dose.

Leveraging Imaging

Image-guided RT (IGRT) was developed to incorporate imaging techniques on the accelerator during treatment, ensuring accurate tumor localization and patient positioning during treatment delivery. ***Magnetic Resonance Linear Accelerator (MR-Linac)*** is an emerging technique building upon IGRT integrating continuous, real-time MR imaging, to improve soft tissue visualization and enhanced capabilities for adaptive strategies. ***Adaptive radiotherapy (ART)*** leverages imaging to not only ensure delivery accuracy, but adjustment the RT plan according to changes in the tumor position, shape, and size over the course of treatment. The integration of ***Machine Learning (ML)*** driven ***Artificial Intelligence (AI)*** tools in RT planning and delivery is improving treatment accuracy and efficiency, optimization of treatment plans, and prediction of outcomes. The goal of these improved techniques is always to minimize harm to normal tissue while maximizing the therapeutic effect.

1.2.8 Common Thoracically Irradiated Cancers

The table below contains a brief diagnosis-specific description of demographics, risk factors, treatment options, RT approaches, and prognosis:

Cancer Type	Demographics	Risk Factors	Treatment Options	RT Approach	Prognosis
Breast Cancer	~12.9% lifetime risk for women; Predominantly affects women (~99%); Commonly occurs in patients over 50 years of age (median age ~62 years)	BRCA1/BRCA2 mutations, family history, smoking, excess weight, sedentary lifestyle, radiation exposure, prolonged estrogen exposure	**Early-Stage**: Surgery is the primary treatment, often followed by RT to eliminate residual disease; Systemic therapy, including chemotherapy, HER2-targeted therapy, and hormone therapy, is common for higher-risk patients **Advanced-Stage**: Treatment is more individual tumor characteristics, with combinations of surgery, RT, and systemic therapies	**Target**: Breast or chest wall, with possible boost to the tumor bed; lymph nodes (if involved) **Conventional**: 50 Gy in 25 fractions (2 Gy per fraction) **Hypofractionation**: 40-42.5 Gy in 15 fractions (2.67-2.85 Gy per fraction), increasingly common	Early detection and targeted therapy significantly improve outcomes; 5-year survival rate is high (~90%), particularly for localized disease; Local recurrence around 5-10% within 10 years for early-stage disease; Metastasis for patients with more aggressive or advanced disease is ~20-30%
Lymphoma (Hodgkin's)	Accounts for about 0.5% of all cancers; Slightly more common in males (~55%); Has a bimodal age distribution affecting young adults (15-40) and older adults (55+)	Infections from Epstein-Barr virus (EBV), immune deficiencies, genetic syndromes	**Early-Stage**: Chemotherapy regimen combined with RT targeting involved lymph nodes **Advanced-Stage**: More intensive chemotherapy regimens, with RT for bulky disease or specific sites	**Target**: Lymph nodes involved in the disease and surrounding tissues (if necessary) **Conventional**: 20-30 Gy in 10-15 fractions (2 Gy per fraction) over 2-3 weeks **Hypofractionation**: 20-30 Gy in 5-10 fractions (3-5 Gy per fraction)	High 5-year survival rate (~87%), especially for early stages; Advances in treatment have significantly improved survival rates; Recurrence(~10-30%) depends on the stage at diagnosis and the response to initial treatment; more common in advanced stages
Lymphoma (Non-Hodgkin's)	Represents about 4% of all cancers; More common in men (~60%); Primarily affects older adults over 60 years of age (median age ~67 years)	Immunosuppression from Human Immunodeficiency Virus (HIV), genetic predispositions	**Early-Stage**: RT alone or combined with chemotherapy regimen **Advanced-Stage**: Chemotherapy with or without immunotherapy is standard, with RT for localized disease or specific indications	**Target**: Involved lymph nodes and any affected extranodal sites **Conventional**: 30-40 Gy in 15-20 fractions (2 Gy per fraction) over 3-4 weeks	Advances in treatment have improved survival rates (~72% 5-year survival rate); Recurrence and prognosis vary widely by subtype; Indolent types have a better prognosis but may recur frequently, while aggressive lymphomas have a recurrence rate of about 30-40% after initial treatment

Cancer Type	Demographics	Risk Factors	Treatment Options	RT Approach	Prognosis
Lung Cancer (NSCLC)	Accounts for approximately 13% of all cancers; More common in men (~55-60%); Typically affects those over 65 years of age (median age ~70 years)	Smoking, asbestos exposure, family history, Chronic Obstructive Pulmonary Disease (COPD), poor air quality	**Early-Stage**: Surgery is the first option; RT is used when surgery is not feasible, or as adjuvant therapy postoperatively **Locally Advanced-Stage**: Combination of chemotherapy drugs, targeted therapy, and RT are often used together **Advanced-Stage**: Systemic therapies, including immunotherapy are common, with palliative RT to control symptoms	**Target**: Primary tumor and regional lymph nodes **Conventional**: 60-66 Gy in 30-33 fractions (2 Gy per fraction) over 6-7 weeks **SBRT**: 48-60 Gy in 3-5 fractions (10-20 Gy per fraction) for early-stage, inoperable tumors, now common for stage I	Poorer prognosis for late-stage; Early-stage detection is crucial but often diagnosed late, leading to lower survival rates; ~25% 5-year survival rate for localized disease; Local and distant recurrence is common, with rates as high as 30-55% within the first two years after treatment for early-stage disease; Advanced-stage disease has even higher recurrence rates
Lung Cancer (SCLC)	Represents 13-15% of all lung cancers; More common in men (~55-60%); Typically affects older adults (median age ~70 years)	Smoking, secondhand smoke, genetic predispositions, occupational hazards	**Early-Stage**: Rare, typically chemotherapy is used due to the aggressive nature; RT is added to improve local control **Advanced-Stage**: Chemotherapy remains the primary treatment, often with thoracic RT; Prophylactic cranial irradiation is common to prevent brain metastasis	**Target**: Tumor and mediastinal lymph nodes **Conventional**: 45-50.4 Gy in 25-28 fractions (1.8 Gy per fraction) over 5-6 weeks **Hyperfractionation**: 45 Gy in 30 fractions (1.5 Gy per fraction, twice daily) over 3 weeks	Very poor prognosis (5-year survival rate <10%) due to rapid progression and early metastasis; ~70-80% of patients experience recurrence within the first year after treatment; SCLC is known for its aggressive nature and high recurrence rates
Esophageal Cancer	Incidence is about 1% of all cancers; More common in men (~70-80%); Primarily affects adults over 55 years of age (median age ~68 years)	Smoking, heavy alcohol use, chronic inflammation from Gastroesophageal Reflux Disease (GERD), Barrett's esophagus	**Early-Stage**: Surgery is the primary approach **Locally Advanced-Stage**: Chemoradiation is commonly used to shrink the tumor before surgery; RT alone used when inoperable **Advanced-Stage**: Palliative chemotherapy, RT, and sometimes immunotherapy used to manage symptoms and slow progression	**Target**: Tumor, regional lymph nodes, and surrounding tissues (if necessary) **Conventional**: 50 Gy in 25-28 fractions (1.8-2 Gy per fraction) over 5-6 weeks	5-year survival rate ~20%; Poor prognosis due to late diagnosis; Early-stage detection and aggressive treatment are crucial for improving outcomes; ~40-60% recurrence within the first 2-3 years after treatment; Majority are metastatic rather than local

1.3 Cardiac Disease

1.3.1 Demographics

The number one cause of death globally, cardiovascular disease (CVD) claims approximately 17.9 million lives annually, representing almost a third of all deaths worldwide. This statistic underscores the significant burden CVD places on the population's health globally. Many individuals survive cardiac events including strokes and heart attacks, which often leads to a range of additional impacts. Beyond financial and emotional impacts, survivors can experience physical effects, including reduced mobility, chronic pain, and other long-term health complications that require ongoing medical attention.

1.3.2 Disease Types

Cardiovascular diseases encompass a wide array of heart and blood vessel conditions. Among the most prevalent are:

- ***Ischemic Heart Disease (IHD)***: IHD occurs when the heart muscle is damaged due to reduced blood flow and oxygen supply, typically as a result of atherosclerotic plaque buildup in the coronary arteries. This can cause chest pain and, in severe cases, myocardial infarction (MI) or heart attack.
- ***Valvular Disease (VD)***: VD involves the dysfunction of one or more heart valves, which can become either narrowed (stenosis) or leaky (regurgitation), disrupting the normal blood flow and increasing heart stress.
- ***Heart Failure (HF)***: HF arises when the heart isn't able to pump blood effectively, leading the body's needs to be insufficiently met by this poor blood flow. This condition may be caused by previous myocardial infarctions, hypertension, or cardiomyopathies, and is characterized by symptoms such as shortness of breath, fluid retention, and fatigue.

1.3.3 General Risk Factors

Cardiac risk factors encompass lifestyle habits, medical conditions, and genetic predispositions. Key modifiable risk factors include poor diet, smoking habits, sedentary lifestyle, and heavy alcohol consumption, all of which increase chances of diabetes, obesity, and high blood pressure and cholesterol. Such physical conditions increase the likelihood of cardiac issues including heart attacks and strokes. Environmental factors, such as air pollution, also play a significant role in heart disease. Additionally, chronic stress and poor mental health can exacerbate these risks.

Non-modifiable factors, including family history, age, and sex, also contribute significantly. Older age and genetic predispositions increase the risk of cardiovascular events, with men facing a higher risk at a younger age, while women's risk increases post-menopause. Managing these risks through lifestyle changes, medical treatment, and regular screenings is critical for preventing cardiovascular diseases [41].

1.3.4 Treatment-Related Risk Factors

Cardio-oncology is a developing field at the intersection of cardiovascular health and oncology. Cancer patients face heightened cardiovascular risk due to shared risk factors, a direct impact of cancer, and cancer treatments. Cancer treatments, while effective against cancer, carry significant cardiotoxic risks that vary depending on treatment type and patient factors. Chemotherapy agents, especially anthracyclines, can cause cardiac problems through oxidative stress and DNA damage. Hormone therapies may increase the risk of cardiovascular events by altering hormone levels that impact cardiovascular health. Targeted therapies may affect blood vessel growth and repair. Immunotherapies may enhance immune responses that also affect the heart. Radiation therapy to the thorax can cause direct damage to the heart and blood vessels. Combining multiple treatments can amplify cardiotoxic effects, particularly in patients with pre-existing cardiovascular conditions [42, 43, 44, 45, 46, 47].

1.3.5 Radiation-Induced Cardiotoxicity

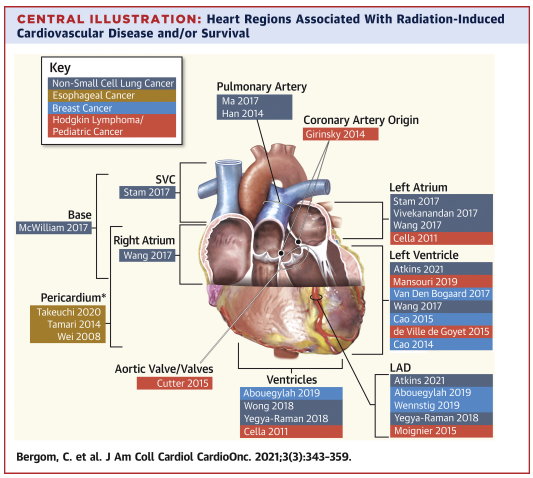


Figure 1.6: "Heart Regions Associated With Radiation-Induced Cardiovascular Disease and or Survival First author and year of publication are listed. Highlighted colors indicate cancer type (see Key). Studies demonstrating associations between total heart doses and outcomes are not included. *Pericardium, not including the heart. LAD = left anterior descending artery; SVC = superior vena cava." Source: Bergom et. al., 2021 [48]

History

The concern over radiation-induced cardiotoxicity emerged as RT became a prevalent treatment for thoracic cancers like breast cancer and lymphoma. Initial cases of radiation-induced heart damage highlighted serious complications such as pericarditis and myocardial fibrosis. These early observations laid the groundwork for subsequent research into the long-term cardiovascular risks associated with RT [49]. The understanding of RT's impact on the heart deepened by the 1970s and 1980s, with clinicians systematically studying late cardiac complications in long-term cancer survivors [50]. Seminal studies in the 1990s and 2000s, such as those by Gagliardi and Darby, quantified the dose-response relationship of cardiac risk and radiation dose, further solidifying the understanding of radiation-induced heart disease (Fig. 1.7) [51, 52, 53, 54, 55, 56, 57]. More recent work has focused on

exploration of detailed dose-response relationships and a shift toward refining radiation techniques to minimize heart exposure (Fig.1.6) [58, 59, 48].

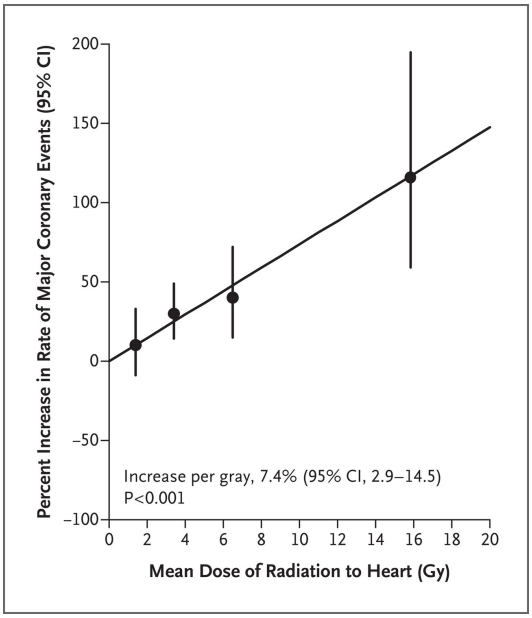


Figure 1.7: "Rate of Major Coronary Events According to Mean Radiation Dose to the Heart, as Compared with the Estimated Rate with No Radiation Exposure to the Heart." Source: Darby et.al., 2013 [54]

Mechanisms and Risk Factors

Radiation can cause acute and chronic damage to the heart, with risks ranging from pericarditis to more severe conditions like CAD, cardiomyopathy, and HF. Acute risks include pericarditis, characterized by inflammation of the pericardium, leading to chest pain, fever, and abnormal heart rhythms shortly after RT. Chronic risks are more severe and diverse, involving the acceleration of atherosclerosis, plaque buildup in coronary arteries, and subsequent IHD (Fig. 1.8)[60, 61].

The contributions of Albert van der Kogel and Peter van Luijk have been instrumental in elucidating the biological mechanisms underlying radiation-induced heart damage. Van der Kogel's research has provided insights into the vascular and fibrotic changes post-RT, which contribute to long-term cardiac disease. Van Luijk extended this research to explore how radiation damage to the heart can exacerbate lung damage in thoracic cancer treatments, highlighting the need for precise treatment planning to minimize combined cardiopulmonary complications [62].

Prevention and Monitoring

Mitigating the cardiotoxic effects of radiation therapy (RT) involves both technological advancements and clinical strategies aimed at minimizing cardiac exposure and enhancing patient safety. Techniques like proton therapy or deep inspiration breath-holding (DIBH) have been pivotal in decreasing cardiac exposure during RT, thereby lowering the risk of radiation-induced heart damage. Regular monitoring of cardiac function using echocardiograms and biomarkers is also crucial for the early detection and timely intervention of cardiotoxic effects [63, 64, 65, 66, 67, 68, 69, 70].

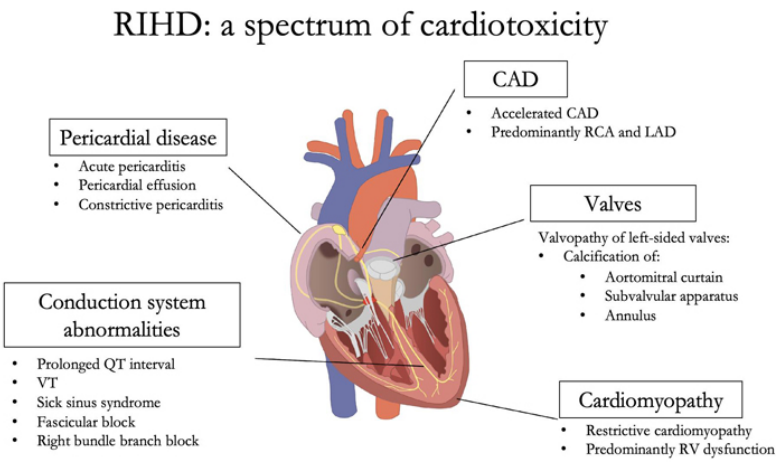


Figure 1.8: "Cardiotoxicity spectrum secondary to RIHD. RCA: right coronary artery; LAD: left anterior descending; VT: ventricular tachycardia; RV: right ventricle" Soiurce: Mehta et. al., 2023 [61]

The American Society of Clinical Oncology (ASCO) provides comprehensive a guideline on managing and preventing cardiotoxicity in cancer patients undergoing treatments including chemotherapy and radiotherapy. These guidelines emphasize the importance of identifying patients at risk for CVD at the time of diagnosis and recommend routine screening for traditional cardiovascular risk factors like hypertension and diabetes [71, 72, 73, 74].

Cardioprotective strategies, including lifestyle modifications and medications (e.g. ACE inhibitors or beta-blockers), are beneficial for improving long-term outcomes in cancer patients. For those at high risk, continuous monitoring and potential modifications to treatment plans are key to balancing the benefits of cancer therapy against the risks of cardiotoxicity [75, 76].

1.4 Segmentation, Delineation, and Contouring

1.4.1 Terminology

Segmentation, delineation, and contouring are closely related but are often used in specific contexts. They are sometimes used interchangeably, but they can reflect different stages or perspectives in the medical imaging and treatment planning process. While they can overlap in meaning, each term has a nuanced use:

- **Segmentation:** Broadly categorizing or dividing an image into regions, often used by computer scientists and in automated processes.
- **Delineation:** The precise identification and outlining of specific structures, typically used by medical practitioners.
- **Contouring:** The drawing of detailed boundaries around areas of interest is commonly used by physicists in treatment planning.

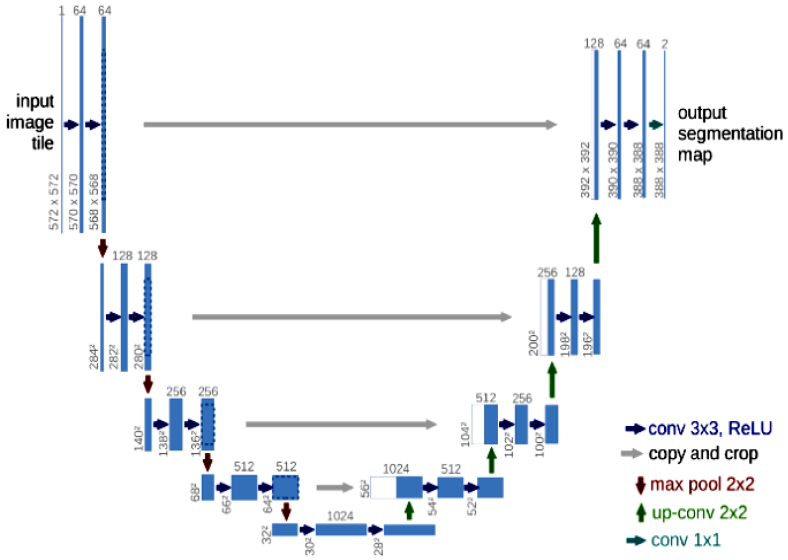


Figure 1.9: UNET model. Source: Ronneberger et. al., 2015 [79].

1.4.2 Techniques

Contouring of OAR in CT scans for RT is a critical step to ensure accurate targeting of tumors while minimizing damage to healthy tissues. Techniques have continuously developed and evolved overtime, being employed with advances in technology and computational methods. Here are the primary techniques used historically and currently:

Manual Segmentation

The boundaries of organs and tumors on each CT slice can be manually delineated by the structures of interest. This can be completed using a simple tool, such as mouse-driven contouring in RT planning software. This can offer high accuracy and control by experienced clinicians, but requires significant time and labor and can be subjective and vulnerable to human error [77].

Atlas-Based Segmentation

This approach uses pre-segmented anatomical atlases (reference images) that are registered to the patient's CT images. The atlas is deformed to fit the patient's anatomy using image registration techniques, providing a semi-automated way to delineate structures. This is faster than manual segmentation and leverages prior anatomical knowledge, but accuracy depends on the quality of the atlas and the registration process and may still require manual adjustments [78].

Automated Segmentation

U-Net's capability for pixel-wise segmentation represents a significant advancement over conventional methods by automatically generating adaptable segmentation masks across various imaging modalities including CT. It is less time-consuming and less prone to variability than manual delineation, and unlike atlas-based registration, it does not rely on pre-defined templates that may not adapt well to individual patient anatomy. This automation has the potential to enhance efficiency in clinical workflows while reducing the need for extensive manual input and associated labor [80].

Well-trained U-Net models can achieve high accuracy and consistency, significantly reducing time and labor. However, they require large annotated datasets for training, and their performance may vary depending on imaging modalities and patient anatomies. In contrast, semi-automated segmentation combines manual input with automated tools like thresholding, edge detection, and region growing. While these tools reduce segmentation time and maintain clinician control, they still require considerable manual intervention and fine-tuning.

1.4.3 In Practice

The evolution of OAR delineation techniques on CT scans has progressed from labor-intensive manual methods to sophisticated automated and AI-driven approaches. Hybrid approaches combine elements of manual, semi-automated, and fully automated techniques to leverage the strengths of each. Automated algorithms provide initial contours, which are then reviewed and adjusted by clinicians. This balances efficiency and accuracy, reducing workload while ensuring clinical oversight, but can require integration of multiple tools and processes, and may still involve some manual effort. Each technique has advantages and limitations, with current trends focusing on leveraging advanced computational methods and multimodal imaging to enhance accuracy, efficiency, and clinical outcomes in RT. While we are not, and may never be, at a place to implement fully automated segmentation methods without human oversight in the clinic, it does pose significant research opportunities [81].

1.5 Study Design

1.5.1 Registry Databases

The Danish National Registry databases are some of the world's most comprehensive and well-maintained health data systems. For decades, Denmark has been committed to systematic data collection with the goal of better informing public health, research, and policy.

The Cancer Registry started in 1943, is one of the oldest cancer registries in the world that tracks incidences, diagnoses, and treatments (Fig. 1.10). Implementing the Central person register (CPR) in 1968 assigned a unique identification to each Danish resident and allowed for the linkage of individual data across various databases and registries [82]. In 1977, the National Patient Registry (LPR) was created as one of the first health-specific registries, to collect data on hospital admissions, treatments, and procedures. The National Health Service Register, established in 1990, expanded this data collection to primary care services and enabled the tracking of outpatient services. The Danish National Prescription Registry, established in 1994, enabled tracking of pharmacy prescriptions. The Danish National Biobank was established in 2012 and provided biological samples to support health research.

The digitization of patient health data, through electronic health records (EHR), has improved drastically in recent decades. This development has enabled improved access and linkage of medical data to aid research and public health efforts. Advancements in technology and application of data science will allow for further utility and impact of this data.

Figure 1.10: "Examples of Danish health Registries, serving as valuable research tools" Source: Sørensen et. al., 2023 [83]

1.5.2 Sensitive Patient Data

Data protection is crucial when accessing and using sensitive health data, and the 2018 implementation of General Data Protection Regulation (GDPR) has reinforced these regulations. Sensitive patient data encompasses any information that can identify a patient and relate to their health, including personal identifiers, medical records, genetic and biometric data, health insurance details, and lifestyle factors. GDPR mandates that data must be processed lawfully, for specified purposes, minimized, accurate, stored only as long as necessary, and protected against unauthorized access and breaches.

In Denmark, ethical approval for research involving human subjects is overseen by the National Scientific Ethics Committee (NVK) and the Scientific Medical Ethics Committees (VMK). The NVK focuses on complex research areas, such as extensive genome mapping and health data science without biological material, while the VMK reviews clinical trials and other medical research to ensure compliance with ethical standards and data protection laws. To comply with GDPR, organizations are required to maintain data protection strategies (e.g. encryption, anonymization, access limitations, audits) protect sensitive patient data.

1.5.3 Observational versus Clinical Trials

There is much conversation surrounding the strengths and limitations of observational studies, especially when compared to clinical trials. The reality is that statistically, the clinical trial remains the gold standard for reaching scientific conclusions. However, numerous challenges make them unrealistic in some settings. Firstly, they are exorbitantly expensive, and research funding is limited. In addition, some scientific questions cannot be ethically explored through clinical trials, where randomization of patients is not appropriate.

Due to the resource-heavy nature of clinical trials, the size of data sets is often limited. So another major benefit of an observation study is the sheer size of data sets made possible. However, often the data available may lack critical information that would not be collected in a standard process flow (Fig. 1.11) [84].

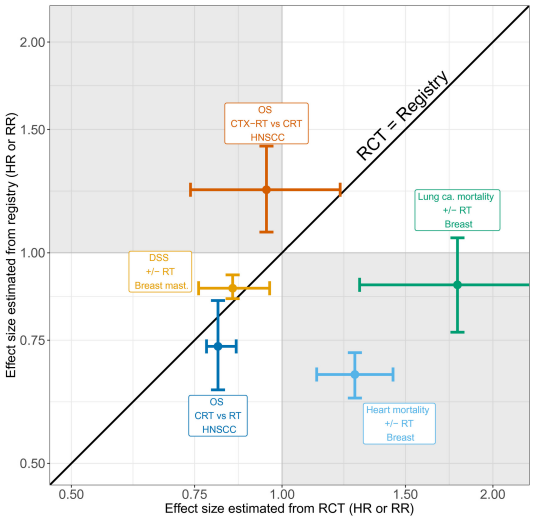


Figure 1.11: "A comparison of effect-size estimates from randomized controlled trials and registry-based analyses. The schematic shows published effect-size estimates from randomized controlled trials (x-axis) and registry-based analyses (y-axis). Concordant effect sizes are indicated by the black identity line. We see examples of registry-based studies over- and underestimating effects, as well as being relatively in agreement." Source: Vogelius and Bentzen, 2020 [84]

1.6 Statistical Methods

1.6.1 Statistical Analysis Plans (SAP)

A SAP is a comprehensive document that outlines the detailed statistical methods and procedures to be used in the analysis of study data. It specifies the primary and secondary endpoints, analysis populations, data handling procedures, and statistical tests to be applied. In clinical trials, SAPs are particularly crucial, outlining the statistical methodologies for assessing the efficacy and safety of interventions. This helps ensure that results are robust and credible, which is vital for regulatory approval and the scientific community's trust in the findings. While less common observational studies, SAPs provide a predefined framework for data analysis, enhancing transparency, reproducibility, and consistency. They are a good data practice and should become more commonplace, as they help reduce biases by specifying analysis methods and endpoints in advance, ensuring rigorous and unbiased data interpretation. [85]

1.6.2 Survival Analysis and Competing Risk

Survival analysis encompasses methods to quantify the waiting time between events, which is fundamental for medical research analyzing time-to-event data such as the onset of adverse events [86]. Kaplan-Meier (KM) estimates and competing risk analyses are complementary methods for approaching survival analysis, each offering unique insights. The KM method offers a non-parametric survival function estimate for this time-to-event analysis. It provides a straightforward step function plot to visualize the probability of survival at different times, making it widely used in clinical settings to evaluate the effectiveness of treatment and event

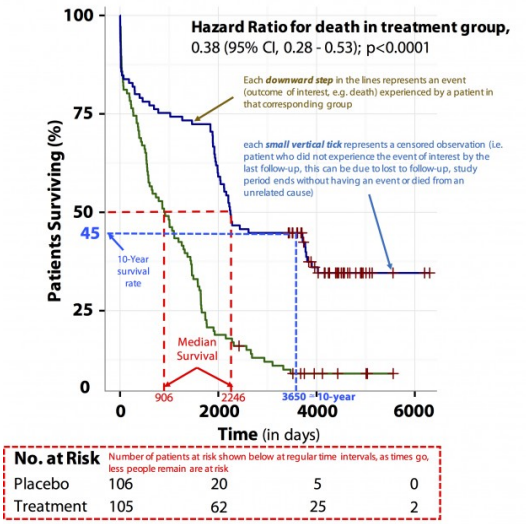


Figure 1.12: "Kaplan-Meier curve (or Survival curve) *(Blue line represents treatment group and green line represents control group)*" Source: Albarqouni 2018 [87]

timing such as disease onset or recurrence (Fig. 1.12) [88]. The KM method accounts for censored data, such as patients that are "lost to follow-up" or don't experience the event [89].

*The **Kaplan-Meier (KM) estimator** of the survival function is defined as:*

$$\hat{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right)$$

where: t_i represents the time of each event, d_i is the number of events at time t_i, and n_i is the number of individuals at risk just before time t_i.

Competing risk analysis, on the other hand, accounts for the presence of other potential events that could impact the probability of the event in question from occurring. The cumulative incidence function (CIF) gives realistic probabilities of events when competing risks are significant. It provides correct absolute risks alongside competing events, measuring the risk of the event occurring conditional on surviving. This can provide a more nuanced understanding each specific event's probability of occurring in the presence of others [90].

*The unbiased **Aalen-Johansen (AJ) estimator** of CIF for a specific type of event k at time t is defined as:*

$$\hat{F}_k(t) = \int_0^t \hat{S}(u^-) d\hat{\Lambda}_k(u)$$

where: $\hat{S}(u^-)$ is the KM estimator of the survival function just before time u and $\hat{\Lambda}_k(u)$ is the cumulative hazard function for event k.

For medical providers and researchers, the trade-offs between KM and competing risk analysis are significant. While neither method is necessarily better or more correct, they offer distinct advantages based on whether conditional or absolute risk is more appropriate to the

context. Researchers and clinicians should follow editorial guidelines that recommend specific methods for defined endpoints, and should always report KM estimates with confidence intervals or standard error. Understanding when to apply KM versus competing risk analysis is crucial for accurate survival analysis, as both methodologies offer valuable but different insights into patient outcomes. Both methods require careful interpretation and selection based on the goals of the study and the nature of the events being analyzed.

Kaplan-Meier estimates provide conditional risk—the probability of an event occurring over time, assuming that no competing events occur. This method is most relevant in clinical and biological studies, where the focus is on understanding the likelihood of specific outcomes such as survival or disease progression, independent of competing risks.KM estimates are particularly useful in scenarios like dose-fractionation effects or evaluating new technologies in normal tissue, as they consider only those individuals still at risk of experiencing the event at each time point.

On the other hand, competing risk analysis offers absolute risk, incorporating the possibility that alternative events (such as death) may prevent the event of interest from occurring. This makes it particularly useful in guiding policy and resource allocation, as it reflects the realistic incidence of various outcomes within a population. The CIF provides a more accurate depiction of the absolute likelihood of an event when competing risks are significant, making it valuable for informing public health decisions by giving a true sense of the overall burden of different outcomes [91].

1.6.3 Multivariate Methods

Multivariate survival analysis methods allow researchers to examine how multiple variables affect time-to-event data. One common approach is the Cox proportional hazards model, which estimates a hazard function, allowing for control of various covariates and providing a hazard ratio for each, measuring the relative hazard or risk [92].

*The **Cox Proportional Hazards Model** is given by:*

$$h(t|X) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)$$

where: $h(t|X)$ is the hazard function at time t given covariates $X_1, X_2, \ldots, X_p$, $h_0(t)$ is the baseline hazard function, and $\beta_1, \beta_2, \ldots, \beta_p$ are the coefficients of the covariates.

Additionally, the **Fine-Gray model** is used when considering competing risks, modeling the **subdistribution hazard** to provide insights into how covariates affect the CIF. This method is particularly important for analyzing absolute risks when competing events prevent the occurrence of the given event, and offers similar trade offs the Cox model as the Aalen-Johansen estimator offered to the KM estimator [93].

*The **Fine-Gray Model** for the subdistribution hazard of event type k is given by:*

$$h_k(t|X) = h_{k0}(t) \exp(\gamma_1 X_1 + \gamma_2 X_2 + \ldots + \gamma_p X_p)$$

where: $h_k(t|X)$ is the subdistribution hazard for event type k, $h_{k0}(t)$ is the baseline subdistribution hazard, and $\gamma_1, \gamma_2, \ldots, \gamma_p$ are the coefficients of the covariates.

1.6.4 Hazard versus Risk

In survival analysis, **hazard** and **risk** describe related but distinct concepts. The hazard rate refers to the risk of an event happening at a given moment in time, conditional on surviving until that point. The **hazard ratio (HR)** compares the hazard rates between two groups, such as exposed versus unexposed individuals, and reflects the relative risk over time.

In contrast, ***Absolute risk*** refers to the overall probability that an event occurs during a specified period of time. This is often the most intuitive measure for communicating risk to patients and is particularly relevant when discussing late adverse effects of cancer treatment, such as those seen in childhood cancer survivors.

- **Relative risk (RR)** compares the absolute risk in two different groups. For example, if a cohort of patients receiving radiation therapy has twice the risk of developing heart disease compared to a non-exposed group, the relative risk would be 2.0.
- **Excess relative risk (ERR)** quantifies the additional risk attributable to exposure. It is calculated as the relative risk minus 1. For instance, an ERR of 0.5 would indicate that exposure increases the risk by 50% compared to the baseline.

In clinical contexts, researchers often focus on excess absolute risk (EAR) and ERR to provide clearer estimates of the impact of treatments, particularly when considering long-term outcomes, as seen in studies of childhood cancer survivors who face increased risks of adverse health events after radiation therapy [91].

For example, in radiation oncology, understanding both the **absolute risk** of an adverse event (such as cardiac toxicity) and the **excess relative risk** associated with radiation exposure helps clinicians balance the risks and benefits of treatment. The **PENTEC Review** highlights the importance of these metrics, emphasizing that absolute risk is often more helpful in clinical decision-making, while relative risk and hazard ratios are valuable for assessing the strength of associations in research settings.

Measure	Abb	Definition	*AR1* and *AR2* known	*AR1* and effect measure known
Absolute risk	AR	The probability or chance of an event. This probability is often specified at a given time after exposure or at a given attained age.		
Relative risk or relative risk ratio	RR	The ratio of absolute risks of a specific event in 2 groups, such as exposed relative to unexposed individuals.	RR= AR2/AR1	AR2= RR·AR1
Excess absolute risk	EAR	The difference between the absolute risks in 2 groups.	EAR= AR2−AR1	AR2= AR1+EAR
Excess relative risk	ERR	Defined as relative risk minus 1. Often specified as a percentage.	ERR= AR2−AR1AR1	AR2= (1+ERR)·AR1
Hazard ratio	HR	The ratio of hazard rates in 2 groups. The hazard rate at a given time T is the probability of an event per unit of time conditional on not having experienced the event before time T.	HR= ln(1−AR2)ln(1−AR1)	AR2= 1−[1−AR1]HR
Odds ratio	OR	Ratio of odds of an event in 2 groups, such as exposed and unexposed children. Odds of an event is equal to the probability q that it occurs divided by the probability it does not occur, that is, q/(1 − q).	OR= AR2·(1−AR1)AR1·(1−AR2)	AR2= AR1·OR1+AR1·(OR−1)
Standardized incidence ratio	SIR	Ratio of the observed number of events in a specific, often small, group relative to what would be expected if the group had the same event rate as a larger comparator population.	SIR ≈ RR if condition in small sample is rare in comparator population	SIR ≈ RR if condition in small sample is rare in comparator population

Figure 1.13: "Risk measures frequently used in pediatric cancer survivorship studies" Source: Bentzen et. al., 2024 [94]

Objectives

2

2.1 Motivation

The potential cardiotoxic side effects of RT have been documented for decades. These risks garnered considerable attention following the publication of Darby et al.'s 2013 case-control study investigating major cardiac events in patients with breast cancer [54]. This was a pivotal study that quantified a dose-response relationship between increased cardiac dose and subsequent onset of IHD that was linear. This illustrated the importance of cardiac sparing and spurred further research into radiation-induced cardiotoxicity, advancing normal tissue-sparing techniques for thoracically irradiated patients [59]. Multifactorial analysis of radiation-induced heart disease has been further explored in studies such as Van Nimwegen et al.'s 2015 retrospective cohort analysis of Hodgkin's lymphoma survivors [55]. More recent research has focused on understanding long-term cardiovascular side effects of RT, extending beyond these major coronary events to further investigate arrhythmias, heart failure, valvular disease, and more [95].

While these studies have paved the way for a better understanding of cardiac health risks, RT is an ever-evolving field resulting in much of this foundational work is no longer representative of modern RT techniques. For example, Darby's study includes patients treated up until 2001, preceding the broad adoption of VMAT. Further, it excluded newer techniques like DIBH, which is now standard practice in treating left-sided breast cancer [65]. To accurately determine current risks and assess the need for improvements in patient care, dose-response relationships must be evaluated using data that reflects modern RT techniques. By leveraging a large, contemporary dataset, our research offers a more representative analysis of current patients than previous case-control studies, allowing us to validate and extend earlier findings within the context of today's clinical practices.

In addition, the analytic techniques applied in these foundational studies do not take full advantage of the capabilities of modern computational improvements. Estimating doses to normal tissues in retrospective data can be challenging and individual reconstruction is often impractical [56]. Conventional methods for contouring OARs rely on computationally heavy atlas-based approaches that apply deformable registration to a reference patient. Modern machine learning models offer greater flexibility and efficiency, providing an opportunity to analyze large and varied research cohorts. Well-trained U-net segmentation algorithms can provide high-quality fully-automatic delineation of individual contours that can outperform atlas-based registration techniques [96, 97]. By fostering collaboration between radiation oncology and computer science, we can harness the strengths of medical expertise and computational resources; by implementing these AI tools we achieve organ-at-risk contouring and assess complex dose estimation in a large retrospective cohort.

The acquisition of reliable data on risk factors and outcomes remains a challenge in observational longitudinal research. In an open healthcare system, loss to follow-up is rampant,

which limits power, poses the risk of biased results, and restricts the ability to appropriately assess treatment outcomes. The transition to EHR systems enables researchers to access much larger datasets. The development of national registry databases in Denmark allows for the linkage of a wide array of detailed patient medical data. We can leverage these powerful resources to identify preexisting conditions, late effects, and comprehensive mortality information.

2.2 Research Aims

The overarching research aim is to elucidate the association between cardiac dose and toxicity. We leverage clinical and public health datasets paired with advanced computational tools to achieve this goal. The specific research aims are described as follows:

Quantify the change in radiation dose to the heart over time. Awareness, technical improvements, and shifting clinical priorities may result in differences in radiation exposure to the heart. We aim to quantify these changes by assessing a cohort of thoracically irradiated cancer patients from 2009 to 2020. Specifically, we use 3D-dosimetry data and AI segmentation tools on CT scans extracted from the record and verify system. This aim is addressed in Study I.

Characterize the associations between mean heart dose (MHD) and volumetric dose measures. The MHD is a metric typically calculated in older epidemiological studies; this metric often relies on estimations from atlas patients. We compare MHD to a high and low volumetric measure - volume receiving at least five or thirty Gy (V5 or V30). This aim spans Studies I and II.

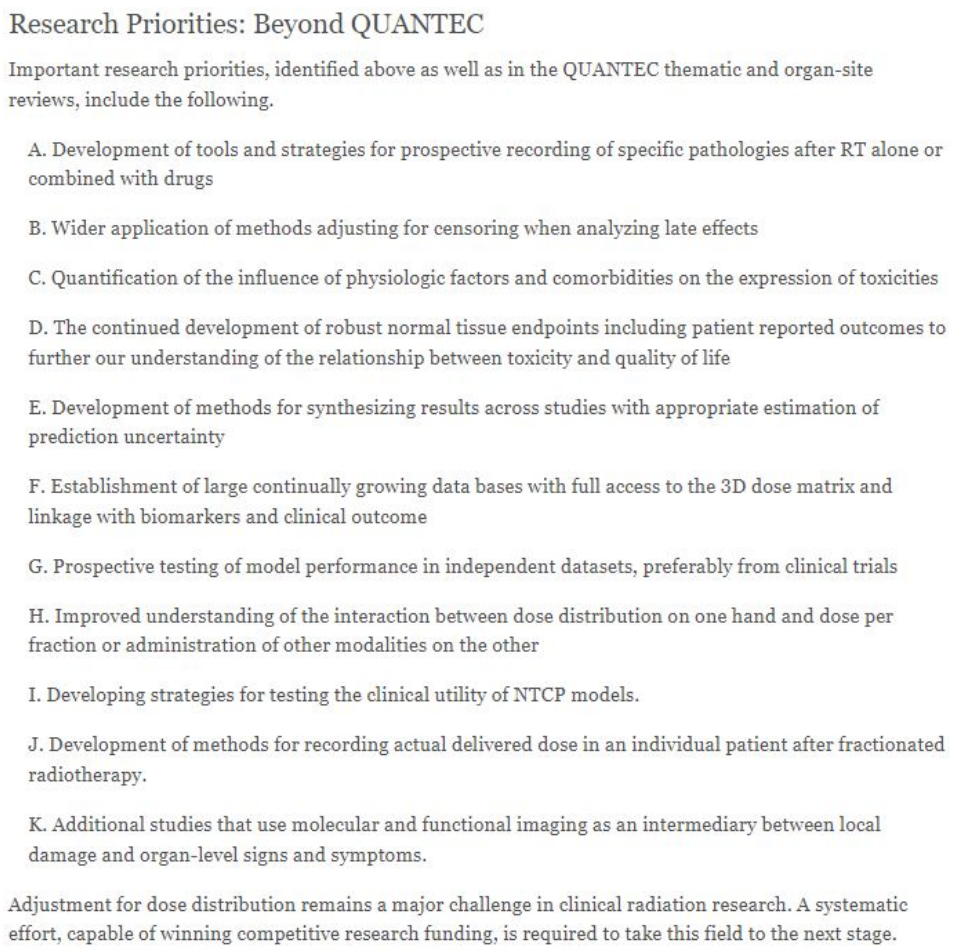
Quantify the risk of cardiac exposure on patient death and late cardiotoxic effects. This research investigates cumulative incidences of death and three cardiac diseases; We primarily focus on IHD, but also explore VD and HF. Outcomes are obtained via the Danish national patient registries - CPR and LPR. This aim is addressed in Study II.

Measure the comparable effect of risks that are treatment-related versus individual patient-specific factors on outcomes. The risk of cardiotoxic late effects and death are assessed using univariate KM models and multivariate Cox proportional-hazards models (PHM). Factors considered included cardiac dose and patient demographics (age, sex, and prior cardiac disease). This aim is addressed in Study II.

This research presents an opportunity to enhance data quality and methodological rigor, thereby bolstering the robustness of our findings. Leveraging detailed treatment records, advanced analytical techniques, and access to comprehensive patient data, we develop a methodology tailored to this multimodal dataset. Our approach aims to elucidate dose-response relationships across a significant cohort of patients receiving thoracic radiation for four cancer types over a 12-year period. Ultimately, this study directly addresses a critical clinical gap by providing insights that can refine risk assessments for cardiotoxicity and improve patient care outcomes.

2.3 Clinical Relevance

The aims of this research directly address an unmet clinical need by providing insights that can refine cardiotoxicity risk assessments and enhance patient care. Moreover, these research aims align with the priorities outlined in the 2010 IJROBP special issue focused on Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) (Fig. 2.1) [98]. By addressing these aims, our work will contribute to a deeper clinical understanding of the cardiotoxic risks faced by patients receiving thoracic radiation. The findings will support improvements in individualized patient care, paving the way for data-driven, tailored treatment approaches that ultimately strive for better patient outcomes.



Research Priorities: Beyond QUANTEC

Important research priorities, identified above as well as in the QUANTEC thematic and organ-site reviews, include the following.

A. Development of tools and strategies for prospective recording of specific pathologies after RT alone or combined with drugs

B. Wider application of methods adjusting for censoring when analyzing late effects

C. Quantification of the influence of physiologic factors and comorbidities on the expression of toxicities

D. The continued development of robust normal tissue endpoints including patient reported outcomes to further our understanding of the relationship between toxicity and quality of life

E. Development of methods for synthesizing results across studies with appropriate estimation of prediction uncertainty

F. Establishment of large continually growing data bases with full access to the 3D dose matrix and linkage with biomarkers and clinical outcome

G. Prospective testing of model performance in independent datasets, preferably from clinical trials

H. Improved understanding of the interaction between dose distribution on one hand and dose per fraction or administration of other modalities on the other

I. Developing strategies for testing the clinical utility of NTCP models.

J. Development of methods for recording actual delivered dose in an individual patient after fractionated radiotherapy.

K. Additional studies that use molecular and functional imaging as an intermediary between local damage and organ-level signs and symptoms.

Adjustment for dose distribution remains a major challenge in clinical radiation research. A systematic effort, capable of winning competitive research funding, is required to take this field to the next stage.

Figure 2.1

Data Materials

3

3.1 Overview

This chapter contains a detailed description of materials and methods used for generating data for this PhD research. For further study-specific details please refer to the manuscripts in the appendix. Below is a consort diagram detailing the cohort selection process:

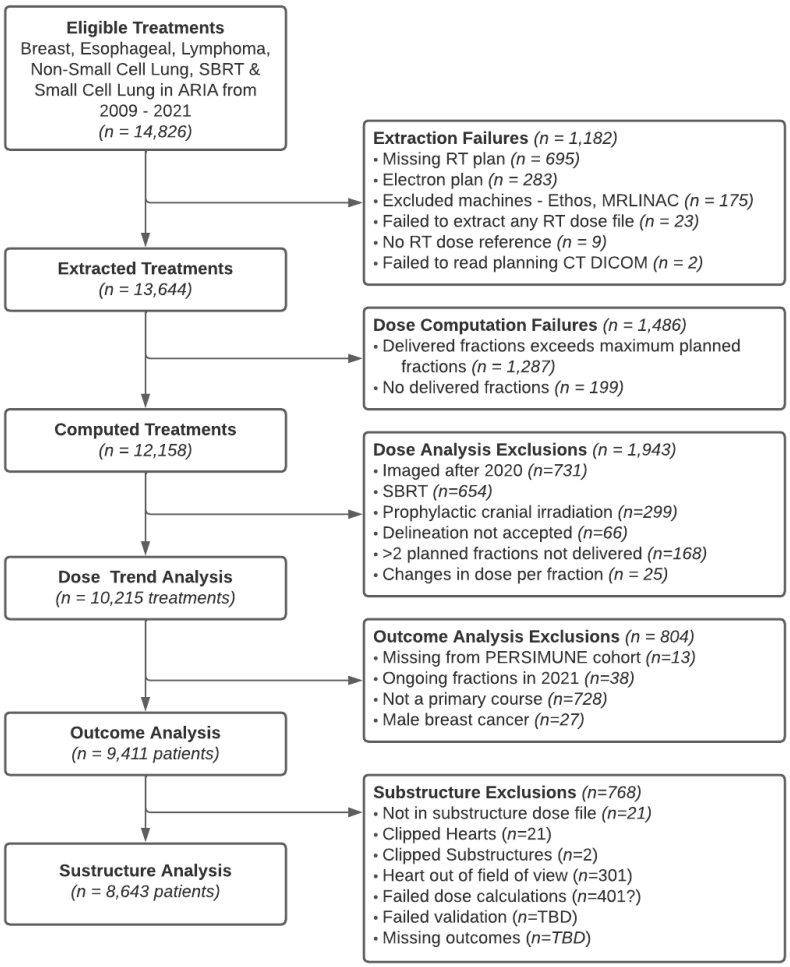


Figure 3.1: Consort Diagram

The data is presented in the following order, preceeded by a section on data formating and storage structure:

- **Image Data:** Accessing CT scans, segmenting hearts, and validating AI algorithms
- **Dose Data:** Calculating and converting dose metrics
- **Registry Data:** Generating usable patient demographics and outcomes data

3.2 Data Formats and Structure

3.2.1 Digital Imaging and Communications in Medicine

Digital Imaging and Communications in Medicine (DICOM) is an internationally recognized standardized form for data storing and transmission of medical images. At Rigshospitalet, RT data is managed through a user interface system and subsequently stored in a DICOM-formatted database. Although working with DICOM data requires more upstart skills, its standardization and high data quality make it superior to other formats, particularly given the large volume of data involved in this project.

3.2.2 Neuroimaging Informatics Technology Initiative

Nearly Raw Raster Data (NRRD) files, a raster-based format for representing spatial data, posed significant challenges outside the DICOM environment, therefore the decision was made to standardize on Neuroimaging Informatics Technology Initiative (NIfTI) for this project. NIfTI, originally developed for neuroimaging, has now become common in broader radiological applications. The Python package NiBabel is used for reading and writing NIfTI files. This format can support up to seven dimensions: the first three for spatial orientation, the fourth for time, and the remaining for additional uses. A 348-byte header contains crucial information such as voxel size and orientation, which are necessary for spatial alignment and accurate volumetric calculations. NiLearn's image processing and resampling tools are employed to ensure that all images align correctly in space.

3.2.3 Data Storage

All data is securely stored on a hospital computer network drive, accessible through a secure shell connection for running Python scripts. Each patient's treatment is associated with a course serial number (prefix 'cs'), resulting in a dedicated folder for each course. Within each course folder, there are subfolders for each CT scan (prefix 'ct'), containing the planning CT scan file (planningct.nii) and the associated heart segmentation file (heart.nii). Additionally, there is a subfolder for each radiation plan (prefix 'rp'), containing the dose matrix file (RD.'dose_id'.nii). Below is an example of the data structure for one radiotherapy course in this cohort, which includes two separate plans. The identification numbers are notably long and complex due to the use of DICOM data sources (Fig. 3.2).

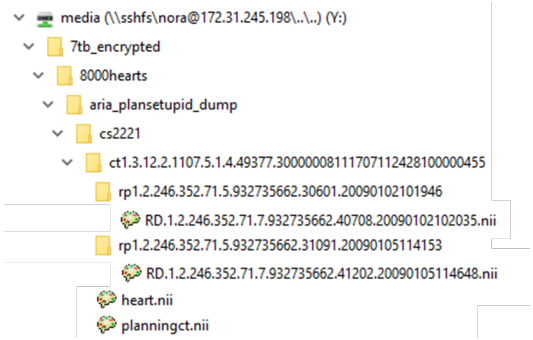


Figure 3.2: Example of the data structure for course 2221

3.3 Image Data

3.3.1 CT scans

CT scans are presented as a 3D grayscale image, with the voxel values representing Hounsfield units, which is a linearly standardized value representing radiodensity. For example, the air is represented by the color black at value -1000, water is represented by a medium gray color at value zero and cortical bone is represented by the color white at a value of >1000. The CT scans were initially required for us to create a segmentation model for the heart. Additionally, the dose matrix was presented in same coordinate system as CT scan.

In most cases, there is only one planning CT per course, but depending on the treatment adherence and disease progression, treatment adaptation and re-planning may occur and a new planning CT might be taken. In these cases, we used the CT scan associated with the majority of delivered fractions. Furthermore, if a patient had a secondary site treated simultaneously, there were additional scans and plans present. Non-thoracic scns and plans were omitted.

3.3.2 Whole Heart Segmentation

We used an open-source AI algorithm, RootPainter3D, which is a corrective annotation U-net model trained from manual delineations on 933 CT scans [5, 99]. This model was then executed on the entire population, yielding heart classification decisions (represented as a binary value; 1 = heart and 0 = not heart) for each of the voxels of the planning CT scans described above. This data is then used as a mask on top of the dose matrix to define which values to include for heart metric calculations.

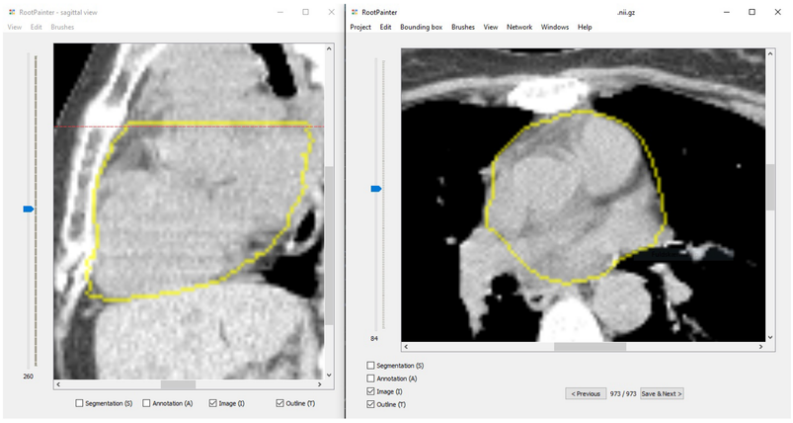


Figure 3.3: Screenshot of RootPainter3D software showing the outline view with segmentation and annotation hidden. The axial view is shown on the right and sagittal view on the left. The user is able to change what data is shown via keyboard shortcuts or with the checkboxes shown in each viewer. Source: Smith et al. 2022 [5]

3.3.3 Whole Heart Validation

We developed a software interface to validate the cardiac delineations in 2D 3.4. The user toggles through each CT scan overlaid with the segmentation and determines acceptability of the segmentation Undecided cases underwent secondary review in 3D and/or and the decision of usability was determined by an oncologist. All unacceptable cases were excluded from the analysis. The primary reasons for failures were anatomic abnormalities, artifacts, or incomplete hearts in the field of view. 6% of all images fed to the model did not contain a heart in the field of view and in 30% of these scans the model correctly identified no voxels as belonging to the heart. This is impressive because the purpose of the model is not classification but segmentation. Notably, 98.1% were accepted initially and 99.2% were included following secondary review. Unsurprisingly, the acceptance rate was a bit higher among patients with diagnoses matching the training data (98% vs 95%). When extending this model to out-of-distribution data we expected some failures, hence the need for thorough visual confirmation of the output, but the model did perform reasonably well.

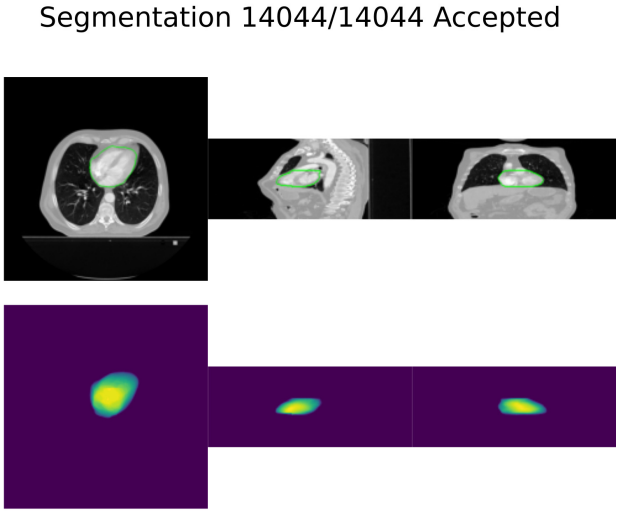


Figure 3.4: **(Study I: FigA1)** "Software interface used for validation of the cardiac delineations. The top row displays the CT scan overlayed by the cardiac delineations in all three anatomical planes – axial, sagittal, and coronal (left to right) – taken at the center of mass of the delineation. The bottom row represents a rendering of the 3D cardiac delineation, where a more intense yellow color indicates a thicker section of the heart." Source: Forbes et. al., 2024 [1]

3.3.4 Heart Substructure Segmentation

Detailed automated segmentation for medical image analysis is rapidly advancing. Multiple open source tools for cardiac substructure delineation have been published during the course of this PhD. Two of them caught our attention for the purposes of this research:

PlatiPy is an open-source Python toolkit for medical image processing and analysis, particularly useful in radiation therapy [100, 101]. The cardiac segmentation tool features an

advanced algorithm designed to automatically segment the heart and seventeen cardiac sub-structures in standard radiotherapy CT scans (Fig. 3.5). Utilizing a hybrid approach, the algorithm begins by segmenting the entire heart with a deep learning nnUNet model. It then employs multi-atlas mapping for contouring the great vessels and chambers of the heart, and concludes with geometric modeling to accurately capture the smaller cardiac structures (nodes, valves, and small vessles). The auto-segmentation algorithm generates cardiac structures following American Association of Physicists in Medicine (Task Group 263) guidelines [102].

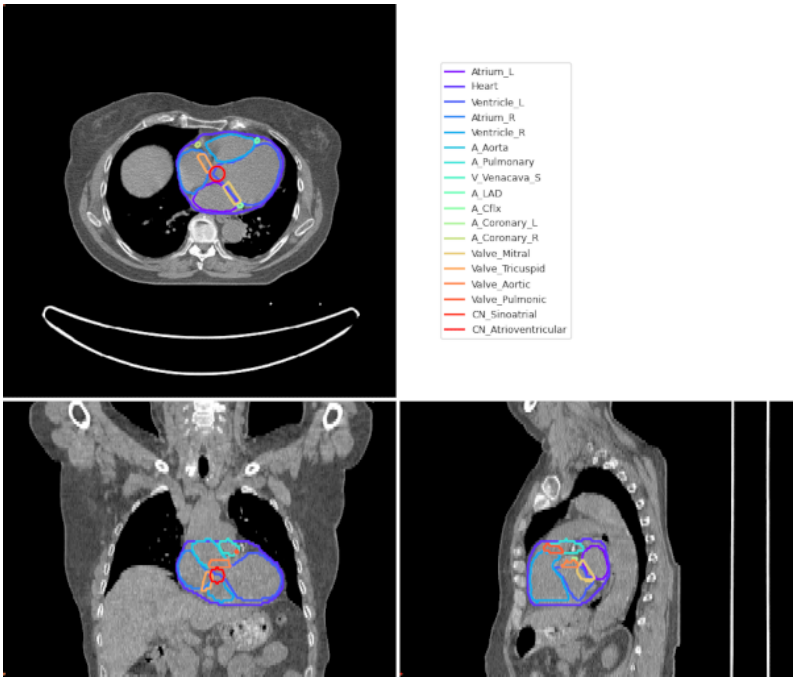


Figure 3.5: PlatiPy Segmentation

TotalSegmentator is an open-source tool designed to facilitate the automatic segmentation of medical images on CT scans [103]. It leverages a nnUNet architecture to identify and delineate various anatomical structures within these images [104]. The tool is known for its versatility, ease of use, and high accuracy in segmenting a wide range of tissues and organs. The Cardiovascular Segmentation Tools contain specialized modules that focus on the detailed segmentation of cardiovascular structures 3.6. These tools enable precise identification and delineation of key cardiovascular components such as:

- **Heart Chambers:** Segments the left and right atria and ventricles, allowing for detailed analysis of the heart's internal structure.
- **Great Vessels:** Identifies major blood vessels like the aorta, pulmonary arteries, and veins, which are crucial for diagnosing and treating cardiovascular diseases.
- **Coronary Arteries:** Segments the coronary arteries, providing valuable insights for assessing coronary artery disease and planning interventions like stent placements or bypass surgeries.
- **Valves:** Segments heart valves (aortic, mitral, pulmonary, and tricuspid), which are essential for diagnosing valvular heart diseases and planning surgical or interventional treatments.

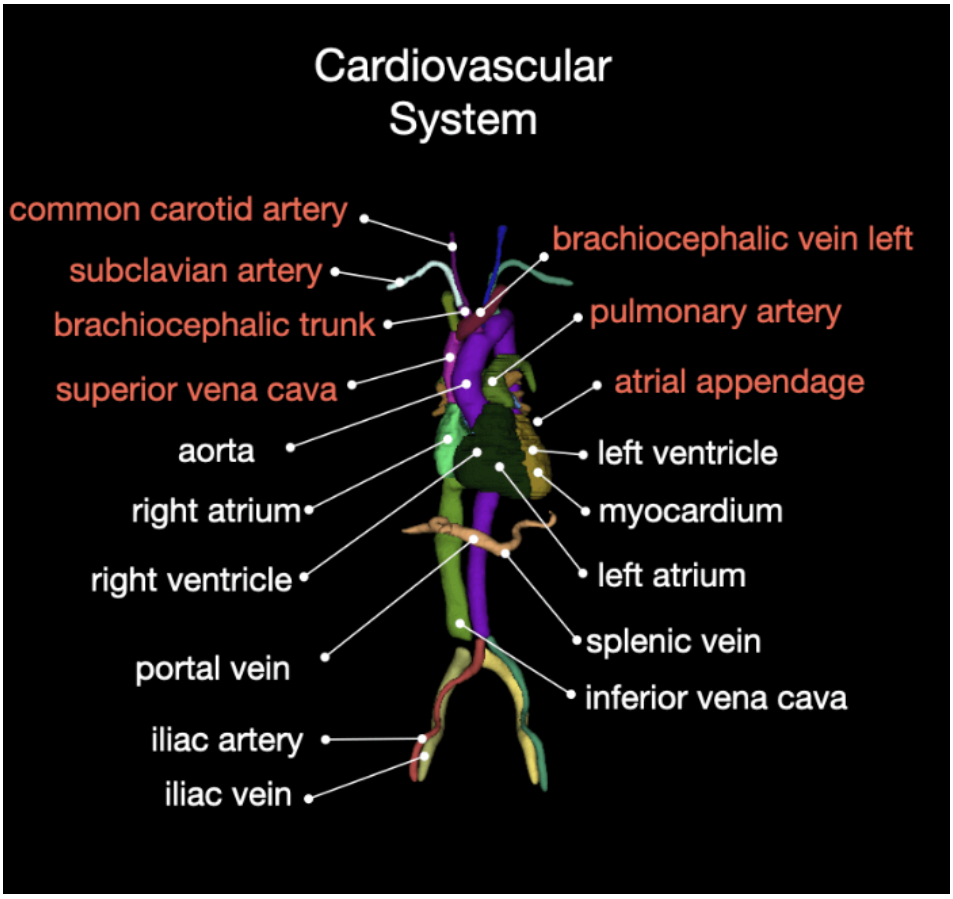


Figure 3.6: TotalSegmentator Segmentation

3.3.5 Heart Substructure Validation

Validation of the heart substructures is still underway and falls under prospective work. Notably, these two segmentation processes differ in their approach and in the cardiac substructures they include; however, they did both include the 4 chambers and 3 big vessels (aorta, pulmonary artery, and superior cava). While PlatiPy has been validated externally, we still must take steps to ensure that it is working well on our dataset [105, 106]. However, it was clear that the review process for the whole heart would not work, as there are now seventeen substructures to consider, which could not be visualized in 2D together and would take too much time to review individually. Additionally, since the substructures are generated using multi-atlas mapping and geometric modeling, the failures would be far more difficult to identify without proper medical knowledge of the anatomy. Therefore we decided that leveraging TotalSegmentator to Validate PlatiPy would allow us to identify potential outliers which required more scrutiny [107].

Note: We can also avoid some expected failures by excluding all hearts already determined to have a failed whole heart segmentation with Root Painter since these were often odd cases that would pose a challenge to automated AI tools. Additionally, we can exclude all hearts that are partially out of the field of view, since it is expected that the algorithms will struggle to accurately segment them.

3.4 Dose Data

3.4.1 3D dosimetry data

On the planning CT scan, a radiation technician delineates the target tumor and nearby OAR. Then a medical physicist will create a plan that defines the location, angles, and dose for each beam of the linear accelerator. This plan takes into account the varying density of tissues in the body and optimizes delivered target dose while preserving surrounding healthy tissue. This dataset does include multiple dose calculation methods, since we used the dose calculated by the software at the time of delivery, rather than recalculating. The plan itself is made up of continuous projections of the radiation beam, however, when we render this into an image it becomes discrete. Therefore, we get an image format where each voxel describes the quantity of dose delivered at that specific voxel volume. Doses are reported in Gy, the international standard unit for measuring the absorption of ionizing radiation amounting to one joule of energy per kilogram of tissue [108].

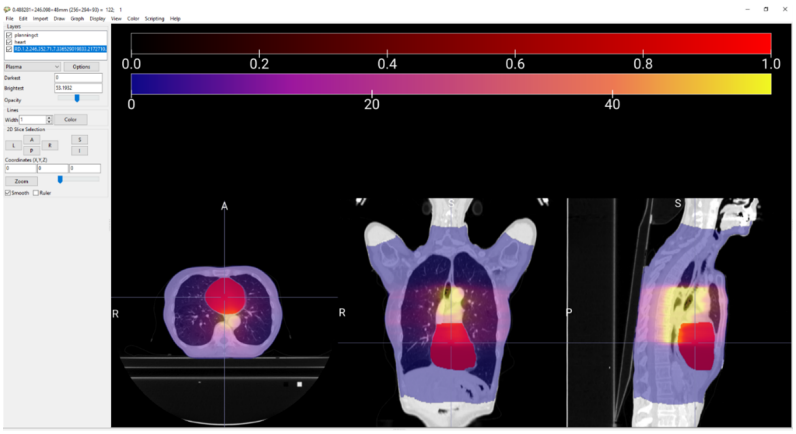


Figure 3.7: Example of a patient with a high cardiac dose. CT scan with whole heart segmentation (in red) overlaying a dose matrix (continuous scale with blue being low and yellow being high)

For cases with a single treatment plan, the total dose was determined by summing the doses across all delivered fractions. When multiple plans existed, it was possible to sum these together too yield a total dose, if they were made on the same planning CT. However, when different CT scans informed multiple plans, it was not feasible to simply add the fractions together. In these scenarios, we used the most delivered plan to estimate total dose by upscaling to the total delivered dose. We had to excluded cases where the treatment schedule was altered, since we could not be appropriately estimate the total dose via the above method. Additionally, courses deliver with electron plans, MR-linac, or using Ethos systems were excluded due to the inability to automate data extraction from these systems [1].

Note: Patients who experience a recurrence of cancer or a secondary malignancy may undergo multiple treatment courses. In Study I, primary treatment status was not a concern since outcome data was not analyzed, but in Study II, all non-primary treatments were excluded.

A mapping file was created to link the course-level data from the dose calculations to the plan-level information from the fraction and dose file. This file also has additional patient demographic information (e.g. sex, birthdate, breast cancer sidedness) acquired from the DICOM database. Dose statistics are summarized by course, making sure to aggregate cases with multiple plans and CT scans. In cases where a course contains multiple plans, we sum them across the fractions.

3.4.2 Dose Calculations

- **Clipped segmentation:** During validation, I observed some CT scans with hearts only partially in frame. While we expect no dose to be delivered outside the field of view, this could effect mean heart dose calculations and atlas-based substructure segmentation. Therefore clipped hearts, with positive voxels along the perimeter, were identified for possible exclusions or adjusted calculations.
- **Volume:** Voxel volume varies with image resolution, so measurement units were extracted from the NIfTI header's "pixdim" values, which define the spatial units. This is later used to cancluate mean dose.
- **EQD2:** As previously mentioned, plans vary in fractional dose and quantity of fractions delivered, even within the same diagnosis. Therefore, it is useful to standardize these plans to represent their relative biological effect. The way to do this is through Equivalent Dose at 2- Gray (EQD2) [36]:

$$EQD2 = D \cdot \frac{\frac{d}{\alpha/\beta} + 1}{\frac{2}{\alpha/\beta} + 1} \tag{3.1}$$

where:

$$D = \text{Total dose (at given voxel)}$$
$$d = \text{Fraction dose (at given voxel)}$$
$$\alpha/\beta = \text{Tissue-specific parameter (2 Gy)}$$

By adjusting values in each voxel, we will end up getting lower values in almost all voxels (since the D is almost always less than twice the fractions delivered) and thus lower dose statistics overall.

- **Mean Heart Dose:** The important step to enable simple calculation of the MHD was the resampling with affine matrix, described in the materials. Once the segmentation and dose are both at the resolution of the CT scan, we can simply overlay them and use the segmentation as a mask. In other words, for all cases where there is a heart segmentation, we sum up the dose only in voxels when the U-net has identified it as heart, and then divide it by the volume of the heart.
- **Near Max Heart Dose:** However, MHD is not very meaningful on its own, since you can conceptualize highly varied dose distributions that would yield the same mean (i.e. a uniform distribution at the mean value of one very high region in a mostly unirradiated heart). Therefore, we want to calculate a maximum value, since extremely high dose values can lead to serious damage to tissue. However, taking the single hottest voxel would lead to a lot of noise. Therefore, I took the distribution of

all the points (flattened from the 3D matrix and excluding non-heart regions) and selected the 95% percentile as a near-max value.

- **Dose Volume Histogram Metrics (VX):** Furthermore, we want to quantify the heart volume exposed to a given threshold of dose or higher. These types of metrics are used in Dose Volume Histograms, which are standard tools in medical physics for radiotherapy planning guidelines to quantify and control patient risks. It is typical to summarize these at the integer level, so I calculated the number of voxels and subsequent volume contained in each 1 Gy bin.

3.5 Registry Data

3.5.1 PERSIMUNE Data Warehouse

While the existence of national registries provided a promising opportunity to conduct national research, accessing and curating that data posed significant challenges. The Centre of Excellence for Personalized Medicine of Infectious Complications in Immune Deficiency (PERSIMUNE)is a multidisciplinary research group at Rigshospitalet. PERSIMUNE has developed a data warehouse providing access to patient records and registries for medical researchers at Rigshospitalet. It contains many data sources including clinical data across multiple departments, biomedical laboratory data, and imaging data. Researchers can access this data following a structured application process, which involves feasibility analysis, project proposal, data request, and data delivery. PERSIMUNE emphasizes collaboration, expecting researchers to contribute to data cleaning and validation efforts.

Through PERSIMUNE we were able to access critical Danish registry data to link to cardiovascular diseases before and after cancer treatment, as well as data on death and emigration. Cardiovascular events were defined using ICD-10 codes:

- **Ischemic heart disease**: I20-I25
- **Valvular disease**: I00-I09, I34-I39
- **Heart failure**: I50

3.5.2 Registries

For this research, we primarily used two registry databases:

- **The LPR** (Danish National Patient Register) in Denmark is a comprehensive health database that collects data on patient interactions within the national healthcare system. This includes information on hospital admissions, outpatient visits, and treatments. The LPR aims to support clinical research, healthcare planning, and quality assurance. It contains detailed records dating back several decades, making it a valuable resource for epidemiological studies and healthcare analyses.
- **The CPR** (Central Person Registry) in Denmark is a nationwide system that assigns a unique personal identification number to each resident. This number is used across various public and private sectors for identification purposes. The CPR contains vital information, such as name, address, birth date, and marital status, which facilitates the efficient administration of services, healthcare, and social benefits. The registry is a cornerstone of the Danish administrative system, ensuring seamless interaction between individuals and institutions.

3.5.3 Data approvals

Once we had identified where and how to access the necessary registry data, gaining approval to access it posed more challenging than expected. In Denmark, the NVEK (Nationalt Videnskabsetisk Komité) is responsible for the ethical review and approval of medical and health science research projects. It ensures that research on human data meets regulations and ethical standards, safeguarding participants' rights and welfare. Researchers must obtain NVEK approval before commencing studies that involve human subjects, biological material, or personal data.

The NVEK underwent significant changes in its approval rules in 2020. These changes were implemented to streamline the ethical review process and ensure more robust compliance with international standards and ethical considerations. The revisions aimed to enhance the efficiency of the review process and improve transparency and accountability in research involving human subjects. To access imaging data, researchers must follow a structured approval process that includes obtaining ethical and regulatory clearances. This process involves submitting a detailed project proposal, undergoing a feasibility analysis, and completing a data request form. The goal is to ensure compliance with data protection laws and ethical standards, safeguarding sensitive patient information.

These NVEK changes occurred at the onset of this PhD, so the approvals needed to conduct this research were not in place when I began. Additionally, since it was a new protocol it took significant effort to formulate the proposal and navigate the requirements. We went through numerous rounds of approval issues were made worse due to the ongoing COVID-19 crisis, which took priority. This resulted in a full year of navigating the system to gain access to the data needed to conduct this research.

4

STUDY I:
Time Trends in Cardiac Doses

This chapter contains a brief description of the research pertaining to the following prepared manuscript and associated ethics report:

> **Forbes, N. J., ..., Peterson, J.*, and Vogelius, I.***; (2024-02). *Time trends in cardiac doses in 10,000 patients receiving curative thoracic radiation therapy between 2009 and 2020.* medRxiv. DOI: 10.1101/2024.02.18.24303007. [1] ****Submitted to Journal of Clinical Oncology*
>
> **Forbes, N., and Vogelius, I.**; (2021). *Radiotherapy exposure and association with observed cardiovascular toxicity in patients treated for cancer at Rigshospitalet* [Report to the Scientific Ethics Committees]. National Videnskabsetisk Komité, 82427.

This manuscript is currently under review at the Journal of Clinical Oncology. The full manuscript, including supplementary material, is located in Appendix A and available online at MedRxiv. See references for DOI.

4.1 Brief Study Outline

In this study, we present methods for generating a large cohort data set of extracted treatment dose metrics using AI tools. This cohort study aims to determine how improvements in RT and cardiac irradiation reduction efforts have altered cardiac doses delivered over a 12-year study period. Eligible treatments include patients receiving curative intent RT in the thoracic region from 2009 to 2020. Patients were identified using treatment codes at Rigshospitalet; treatment exclusions are listed in the consort diagram (Fig. 3.1). Dose matrices and CT scans were included for system verification.

4.2 Key Findings

We included 10,215 treatments on 9,966 patients eligible for analysis. Treatments included four cancer diagnoses - breast cancer, lung cancer, lymphoma, and esophageal cancer. An in-house AI segmentation tool was leveraged to generate individually delineated hearts. A 2D manual review was conducted on every delineation; when needed, a secondary 3D review was conducted under medical guidance (Fig. 3.4). Dose metrics were calculated, including MHD and volumetric measures (e.g., V5, V20, and V40). Cardiac doses were quantified using time trends across the 12 years. A decrease in these cardiac dose measures was observed, and this was more evident

in high-dose exposures. Unexpectedly, in the most recent years, patients receiving breast cancer treatment increased in high-dose volume (V40Gy) to the heart.

4.3 Considerations

Increased awareness of cardiotoxic risks and advances in RT techniques that improved tumor targeting, likely contribute to a decrease in cardiac radiation exposure, as quantified in this large patient study. The application of custom AI methods allowed for the generation of high-quality, time-independent individual delineations. While critical for understanding observed treatment exposure, retrospective observational studies are unable to determine causal inference. New and unaccounted for factors, such as changing patient demographics, increase the complexities of real-world data. Although a decrease in cardiac dose has been observed, linking this treatment data to registry-based outcomes is necessary to understand the impact of these dose trends on patient outcomes.

4.4 Perspectives

AI models have the potential to provide high-quality patient-specific delineations and dose calculations for large cohorts. These methods are of particular interest in large-scale research settings, where manual delineation is impractical. Improvements in segmentation methods continue; segmentation tools are now widely available, including our open-source AI model. Validation of these AI tools with external patient cohorts is important to detect critical failures of outlier cases and out-of-distribution data. While rare, such errors are of particular concern in clinical processes informing diagnostics and treatment planning, and thus human-in-the-loop approaches will continue to be the standard.

5

STUDY II:
Cardiotoxicity Analysis

This chapter contains a brief description of the research pertaining to the following prepared manuscript and associated SAP:

> **Forbes, N. J., ..., and Vogelius, I.**; (2024-08). *Cardiac dose-volume analysis of 9,411 patients with registry-based outcome data for cardiotoxicity and overall survival.* medRxiv. DOI: 10.1101/2024.08.16.24312108. [2]
> ****Submitted to Journal of the European Society for Radiotherapy and Oncology*
>
> **Forbes, N.**; (2023). *Clinical and dosimetric risk factors for ischemic heart disease* [Statistical Analysis Plan]. DOI: 10.6084/m9.figshare.24247372.v1. [3]

This manuscript is currently under review at the Journal of the European Society for Radiotherapy and Oncology. The full pre-print manuscript, including supplementary material, is located in Appendix A. The manuscript is available online at MedRx. The SAP is available on Figshare.

5.1 Brief Study Outline

We aimed to assess the affect of cardiac exposure on long-term cardiotoxicity and overall survival. Study II builds on the work conducted in Study I by linking the calculated radiation doses to registry-based outcome data. This was made possible through access to the LPR and CPR registry databases via the PERSIMUNE data warehouse. All nonprimary treatments were excluded from the analysis.

5.2 Key Findings

9,411 patients were included in the 12-year retrospective cohort study. The cumulative incidences of patient outcomes with competing risks (cardiovascular events and death) were calculated using the Aalen-Johansen estimator. The effects of cardiac dose and patient factors on OS and the development of IHD were evaluated with the KM method and Cox models. The risk of cardiotoxicity was relatively low compared to death, especially for lung and esophageal cancers. Patient-specific factors (e.g. existing cardiac disease and age) were far stronger predictors of cardiotoxicity compared to dose. In dose-response studies, overall survival was not a reliable indicator of cardiotoxicity due to the confounding effects of the disease stage.

5.3 Considerations

This study could only include one site in Denmark over a 12-year period, due to the complex nature of these data, therefore generalizability of these results should be cautioned. Disease characteristics (e.g. lymph node involvement or overall extent) likely confound the impact of treatment-related toxicity. In future studies, additional cancer classification, such as subtypes of lymphoma (Hodgkins vs Non-Hodgkins) and NSCLC cancer (e.g., Adenocarcinoma, Squamous Cell Carcinoma), might control for dose- and prognosis-related heterogeneity; this data is advised for inclusion in smaller studies.

5.4 Perspectives

For most patients, contemporary high-quality approaches to radiotherapy effectively minimize the risk of heart toxicity, rendering it a minor clinical concern. However, the lack of a significant result does not indicate the absence of a dose relationship; still, a significant dose-response relationship was not established despite with a large cohort of 9,411 patients.

Discussion

6

> *If I have seen further it is by standing on the shoulders of giants.*
>
> **— Isaac Newton**

Through this work, we aim to better understand the negative side effects of RT through the application of modern computational tools to clinical and public health datasets. Machine learning segmentation tools applied to medical images enabled the quantification of RT dose to individual tissues. The impacts of radiation on patient outcomes were further quantified by linking image and dose data with negative side effects experienced by cancer patients.

6.1 Principal Findings

How has radiation delivered to the heart changed over time? Study I, presented in [Chapter 4], investigates the evolution of cardiac radiation dose delivery to the heart. Results support the understanding that advancements in RT technology, such as the adoption of IMRT and VMAT, significantly improve the precision of treatment plans which minimizes unnecessary radiation to surrounding healthy tissues. Increased awareness of cardiac radiosensitivity and its risks to cardiovascular health resulted in changes to existing guidelines reducing cardiac exposure. The sole exception to this trend was found in more recent breast cancer patients. This unexpected finding can be explained by new guidelines implementing a cardiac dose threshold, which were higher than standard practice. Let this finding serve as an example of potential inadvertent impacts of guideline changes and as a call to continue monitoring delivered doses to this cohort.

Is cardiac exposure associated with subsequent cardiotoxicity or increased mortality? Study II, presented in [Chapter 5], explores the correlation between cardiac radiation doses and the onset of cardiotoxicity and increased mortality rates. No clear dose-response relationship between cardiac doses and risks of adverse cardiac events was found. However, an absence of such a relationship should not be interpreted to contradict previous research findings; rather, the relationship was not found in this relatively low radiation exposure cohort. Years of awareness of cardiotoxic risks and ongoing advancements in RT have improved treatment efficiency and lowered observed cardiac doses. Dose-response relationships were observed for OS. The priority of target coverage is clear, as the risk of death far

outweighed that of cardiac events. This study highlights the need for continued care in planning and monitoring cardiac doses, especially in individuals with risk factors for heart disease, specifically existing cardiovascular events.

6.2 Research Impact

The integration of AI in RT, through tools like Varian's Ethos system, is offering greater precision and personalization of treatment planning [109]. Such advancements result in better sparing of healthy tissues and improved patient outcomes [110]. Existing studies on cardiotoxicity often lack consistent dosimetry data and use conventional methods for dose calculation, such as traditional ATLAS registration techniques. By individually segmenting hearts with an in-house AI tool and leveraging detailed 3D dosimetry data, ***Study I*** demonstrates that AI can enhance dose calculation accuracy and consistency in large cohorts. This offers improvements over prior studies and enabling the analysis of more precise dose-response relationships. AI-driven deep learning frameworks can further enhance automatic organ segmentation and provide more accurate dose calculations [111, 112].

Advancements in treatment planning and delivery techniques are enabling more acurate targeting of tumors, reducing radiation exposure to surrounding healthy tissues while maintaining adequate target coverage [25]. These improvements aim to minimize side effects without compromising treatment effectiveness. Current research in cardiotoxicity emphasize the risks of cardiac exposure and advancements in RT techniques, indicating cardiac sparing as a priority [113]. However, large-scale retrospective studies evaluating actual delivered doses remain limited. ***Study I*** demonstrates that radiation delivered to the heart is decreasing, highlighting this trend in high-dose exposure. This contributes to the growing body of work in radiation-induced cardiotoxicity by investigating trends in a modern large comprehensive high-quality dataset across multiple diagnoses over 12 years. Further, it identifies an interesting and unexpected recent increase in dose among breast cancer patients.

National registry data offers significant advantages in medical research, particularly in centralized healthcare systems like Denmark's. These systems enable more reliable tracking of patient outcomes, reducing the issues of incomplete or inconsistent follow-up data often seen in fragmented healthcare systems. Many cardiotoxicity studies lack detailed cardiovascular outcome data and tend to focus on survival rather than chronic conditions that may develop years after treatment [114, 115, 116, 53, 117]. In ***Study II***, utilized detailed diagnostic coding from Denmark's LPR registry to investigate prior and subsequent cardiac events — including IHD, VD, and HF — resulting in more accurate assessments of cardiotoxic risks. Additionally, detailed mortality and emigration data were extracted from the CPR registry to ensure high-quality censoring. These findings demonstrate the value of comprehensive registry

data in enhancing the quality and completeness of patient data, which is crucial for accurate outcome-focused research. Such insights would be more challenging to achieve in less integrated healthcare systems, where fragmented data and high population mobility hinder long-term follow-up.

As treatment precision improves, patient outcomes are expected to follow suit. However, long-term cardiovascular risks, especially for patients receiving radiation near the heart, require more focused attention [118, 119]. ***Study II*** leveraged Denmark's robust registry data, linking it with detailed dose data to assess the impact of cardiac exposure on long-term patient outcomes. Despite a cohort of nearly 10,000 individuals, no direct correlation was found between radiation dose and adverse cardiac outcomes, likely due to the lower doses achieved by modern RT techniques. This highlights the impact of recent RT advancements and the need for even larger sample sizes to detect potential trends. By minimizing data loss and ensuring comprehensive follow-up, the use of registry data enhances the reliability of research findings, ultimately leading to improved guidelines and better patient outcomes in RT.

This body of work contributes significantly to the field of radiotherapy safety and optimization. It underscores the critical balance between effective cancer treatment and the minimization of harmful side effects. The insights gained not only enhance our understanding of RT effects on cardiac health but also pave the way for future research focused on improving patient outcomes through comprehensive clinical data analysis and advanced computational techniques. The integration of big data analytics and AI is transforming RT, improving precision and personalization in treatment delivery. Continued validation of these advances is essential to ensure consistency with RT conventions, ultimately leading to better patient care and outcomes.

Prospective Work

7

While the existing work has enabled a deeper understanding of the cardiotoxic risks associated with contemporary RT, there are clear opportunities to build on the latest findings. This section outlines promising directions for expanding current research and refining methodologies, with the aim of further personalizing cancer treatment, resulting in improved patient outcomes.

7.1 Continued Monitoring and Cohort Expansion

This study successfully analyzed dose trends and dose-response relationships for a cohort of nearly 10,000 patients. Given the significant efforts in data extraction and analysis, continued monitoring of this cohort is feasible with minimal additional effort. Long-term follow-up, particularly for breast cancer and lymphoma patient subsets, is crucial as these groups tend to have better survival rates. Many of these patients are still alive and free from CVD after five years, but the risk of CVD increases with time, highlighting the importance of extended follow-up to detect longer-term late effects. Continuous monitoring under changed priorities is necessary to understand relevant trends.

Rigshospitalet is the largest research hospital in Denmark and therefore plays a significant role in providing highly specialized care. Expanding the cohort to include other hospitals in the capital region, or even those in Eastern Denmark, would strengthen the statistical power of the study. Further expanding the dataset to include international hospitals would significantly enhance the diversity of the cohort and improve the generalizability of the findings. Barriers to expanding beyond the current scope include data approval and access. The current NVEK approval lists only Rigshospitalet and the Danish Data Protection Authority (Datatilsynet) umbrella agreement currently terminates in Feb 2022, both of which would require an update. Extracting data from external EHR system always poses challenges, but established collaborations (e.g. Herlev Hospital) offer the lowest hurdle for accessing this data and expanding the cohort. Leveraging Denmark's national registry databases, combined with efforts to foster collaboration through initiatives like the Danish Comprehensive Cancer Center (DCCC) and data infrastructure projects such as DcmCollab could aid in broadening the dataset [120, 121, 122]. Additionally, international collaborations, such as those with Swedish institutions through the Öresund Workshop on Radiotherapy, present opportunities for data expansion and a more comprehensive understanding of the study outcomes [123].

7.2 Detailed Dose Distribution Analysis

7.2.1 Volumetric Analysis

Volumetric measures provide a more detailed understanding of radiation dose distribution across tissues compared to summary statistics like mean dose. The DVHs allow for the evaluation of OAR volume (absolute or relative) receiving specific dose levels, aiding in risk stratification and toxicity prediction. While some aspects of volumetric analysis were explored in this study, much remains to be analyzed.

A notable challenge with volumetric analysis is the high correlation between dose metrics. For instance, different dose-volume parameters (e.g. MHD, V5, and V30) can be highly interdependent, which may introduce multicollinearity issues in statistical models. This high level of correlation complicates efforts to identify independent predictors of toxicity and can obscure the relationship between specific dose levels and clinical outcomes. Principal component analysis (PCA) could be employed to help decrease the data dimensions by transforming the correlated dose metrics into fewer uncorrelated components. This reduces redundancy and allows for a more robust analysis of the underlying relationships between dose distributions and outcomes. Similarly, least absolute shrinkage and selection operator (LASSO) regression can handle multicollinearity by performing variable selection and regularization, helping to identify the most relevant dose metrics while penalizing less important ones. Due to limitations of time and resources, we decided to prioritize the work required for completing Studies I and II, rather than further investigation of these volumetric analyses, especially given the weak preliminary evidence for a dose-response relationship.

Further, these volumetric measures do not offer spatial information, making it challenging to assess how the location of high-dose regions effects outcomes. This is important because OARs are not uniform with certain substructures being potentially more vulnerable to radiation-induced damage. To further address the limitations of volumetric measures, future research should combine volumetric analysis with spatial techniques. These approaches may provide a more precise understanding of dose-response relationships potentially improving patient care by targeting radiation more effectively [124].

7.2.2 Substructure Analysis

Background and Motivation. Building on Studies I and II, we hypothesize that a more granular dose-response could be uncovered with the segmentation of the heart into substructures. Determining the radio-sensitivity of cardiac tissue could in turn lead to opportunities for clinically meaningful optimizations that selectively avoid high-risk regions while maintaining appropriate coverage of target areas [95, 125].

Whole heart dose calculations, while informative, fail to capture the detailed anatomical variations within the heart. These calculations can mask important substructure-level information that may be critical for understanding the risk of radiation-induced cardiac damage [126]. This proposed research seeks to elucidate the role of radiation dose to specific substructures and its impact on patient outcomes.

The objective of this prospective analysis is to examine the variation in dose distribution across different regions of the heart, assess how these have changed over time, and investigate whether specific locations of radiation exposure correlate with negative patient outcomes. Further, this study aims to demonstrate the utility of automated segmentation in cardiac substructure analysis, while emphasizing the need for rigorous validation and cross-verification.

Preliminary Methods and Results. We have started an analysis and provide some preliminary results to show feasibility. As described in Chapter 3 - Data Materials, we were able to generate cardiac substructure segmentations for this cohort by leveraging an open-source AI tool known as PlatiPy. Dose metrics on 8,643 patient hearts were successfully calculated for all seventeen substructures provided. The dose distribution trends observed in these substructures paralleled those seen in the whole heart, with significant decreases over time across the diagnoses. We identified significantly higher doses in the left anterior descending Artery (LAD) of left-sided breast cancer, right coronary artery (RCA) of right-sided breast cancer, pulmonary artery (PA) of lymphoma, left atrium (LA) of esophageal cancer, and superior vena cava (SVC) of lung cancer (Fig. 7.1). The extent to which these substructure exposures relate to an increased risk of IHD or OS is still being investigated.

Future Work. Cardiac substructure delineation was performed using the open-source AI tool PlatiPy, which segments seventeen distinct cardiac substructures. However, further validation is required to ensure robustness. To validate these results, we plan to compare PlatiPy's segmentations with those obtained from TotalSegmentator, another open-source AI tool with broader anatomical segmentation capabilities. We will segment the four cardiac chambers (LA, RA, LV, RV) and major vessels (AA, PA, and SVC) using both tools. The agreement between segmentations will be assessed using Dice similarity coefficients. Additionally, we will utilize the in-house quick review tool previously used for validating whole-heart segmentations to evaluate any major discrepancies (Fig. 3.4). This will be followed by a comprehensive dose-response analysis to further assess the impact of these segmentations on dose calculations. Our ongoing manuscript explores substructure dose trends and outcome analysis, demonstrating the potential of AI tools in refining dose calculations.

Considerations and Perspectives. Conventional RT planning often quantifies cardiac sparing across the whole heart, overlooking anatomical differences that could explain varying substructure sensitivities. Literature suggests that whole heart dose-response relationships may serve as proxies for substructure effects, but further

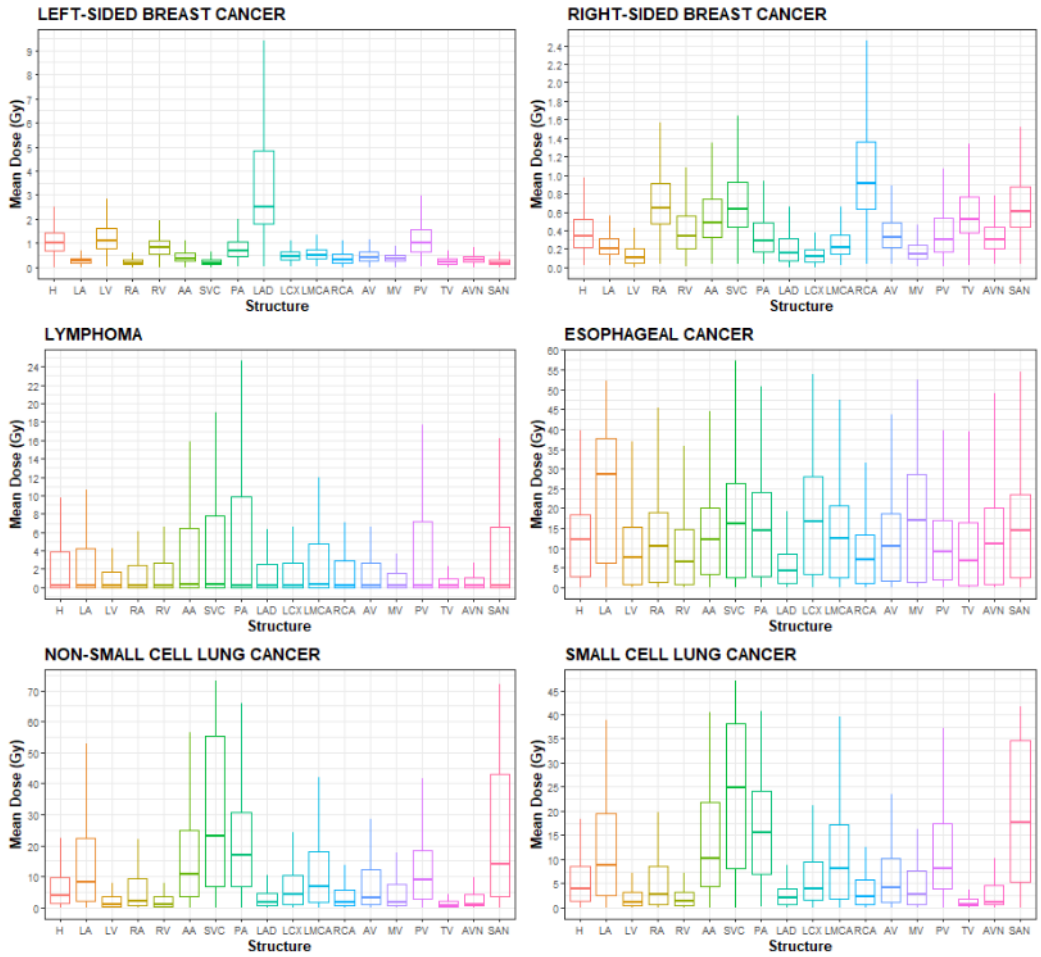


Figure 7.1: Boxplots of mean heart structure dose by cancer type. Outliers not included. (H = Whole Heart, LV = Left Ventricle, RV = Right Ventricle, LA = Left Atrium, RA = Right Atrium, AA = Aorta Artery, PA = Pulmonary Artery, LAD = Left Anterior Descending Artery, LCX = Circonflex Artery, LMCA = Left Coronary Artery, RCA = Right Coronary Artery, SVC = Superior Vena Cava, MV = Mitral Valve, TV = Tricuspid Valve, AV = Aortic Valve, PV = Pulmonic Valve, SAN = Sinoatrial Conduction Node, AVN = Atrioventricular Conduction Node)

study is needed. This research explores substructure doses across multiple cancer types — breast, esophageal, lymphoma, and lung cancers. Building on findings like the correlation between LAD dose and heart disease in breast cancer [126], we explore relative doses between substructures and with the whole heart, as well as time trends and outcome analysis.

The integration of open-source AI tools in RT is advancing rapidly, providing accessible resources for enhancing treatment planning and analysis. Tools such as PlatiPy for automated contouring and TotalSegmentator for multi-organ delineation show promise in refining dose-response analyses [100, 103]. However, the adoption of these tools requires thorough validation to ensure robustness, especially with outlier cases or out-of-distribution data [127]. Dose calculation errors, particularly in smaller substructures, underscore the need for continued scrutiny of AI-based segmentation methods.

Findings from substructure dose analysis may call for more detailed anatomical data in normal tissue exposure studies, potentially informing updates to RT threshold guidelines. Standardizing AI tool applications across clinical practice will require careful comparison and validation to ensure consistent accuracy and reliability, ultimately improving RT planning and reducing cardiac complications in thoracic RT patients.

In parallel with refining AI segmentation tools, integrating complementary methods like image-based data mining (IBDM) could help address tissue-specific radiosensitivity [114]. As imaging technology advances, AI segmentation is anticipated to increasingly contribute to the fields of radiation oncology, supporting personalized treatment strategies that minimize adverse effects while maintaining therapeutic efficacy.

7.2.3 Image-Based Data Mining

Background and Motivation. The field of IBDM has evolved significantly, with early studies laying the foundation for its current use in RT. Acosta et al. (2013) introduced voxel-based population analysis, linking dose distributions with clinical outcomes such as rectal toxicity [128]. This was followed by improvements in statistical methods by Chen et al. (2013) and Sörensen et al. (2017), who incorporated permutation tests and false discovery rate corrections to ensure the reliability of IBDM results [129, 130]. These advances allow for more sophisticated analyses, such as those by Green et al. utilizing per-voxel Cox regression to map dose-response relationships and predict patient outcomes more accurately (Fig. 7.2) [114].

Methods in IBDM offer a unique advantage over conventional methods, which often treat anatomical structures as if they are uniformly sensitive, overlooking the heterogeneity of tissue responses. By analyzing dose-response relationships at a voxel

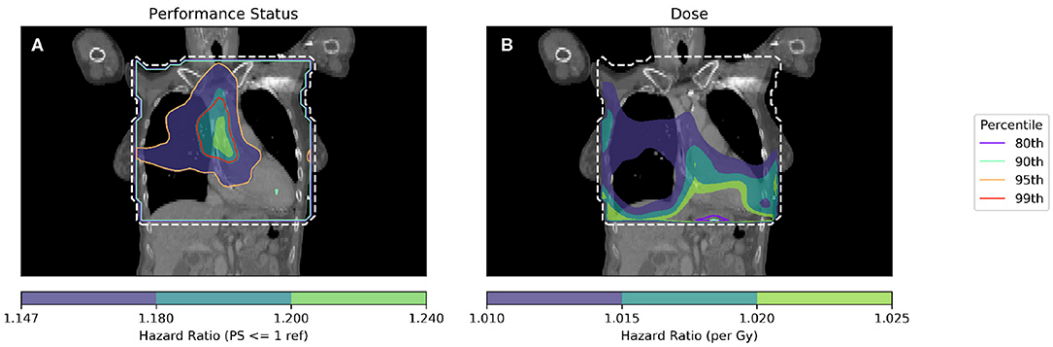


Figure 7.2: "Cox-IBDM performed considering only PS (A) and dose (B). Shown by the dashed white line is the region inside the mask, where hazard ratio calculations were performed. In unfilled contours, hazard ratio percentiles are shown, while hazard ratios are shown in filled contours... Note that the hazard ratio scale is different in each subfigure." Source: Green et. al., 2020 [114]

level, IBDM provides a more granular understanding of how radiation affects specific subregions within organs. This approach uncovers complex spatial relationships that conventional models might miss, allowing for a deeper understanding of the effects of RT on both target and adjacent tissues. As radiation therapy techniques evolve, IBDM becomes increasingly valuable for identifying subtle patterns of toxicity or damage that may emerge only under specific spatial configurations or dose distributions.

Future Work. The next steps in this project will involve applying IBDM techniques to our cohort, focusing on analyzing dose-response relationships at a voxel level within specific substructures. We could use patient imaging data to assess radiation exposure across different tissue types and substructures, particularly in complex thoracic cases. Our goal is to map these dose distributions to clinical outcomes, identifying regions of interest that correlate with adverse effects. Validation of the IBDM results will be crucial, requiring thorough testing against external datasets to ensure reliability.

Considerations and Perspectives. Despite its potential, IBDM requires careful validation to avoid misinterpretations. As Shortall et al. (2021) caution, voxel-based analyses can sometimes produce misleading results if the complexity of spatial data is not adequately addressed [131]. The heterogeneity within substructures and the possibility that radio-sensitive tissues extend beyond traditional anatomical boundaries highlight the importance of rigorous evaluation and external validation. Moreover, dose-response behavior is unlikely to be uniform within organs or tissues. This complexity underscores the need for complementary approaches, such as deep learning, which can enhance the capabilities of IBDM by providing richer datasets and more robust predictive models. These advancements will help tailor RT to individual patients, improving treatment precision and contributing to a deeper understanding of how RT impacts various tissue structures [19].

7.2.4 Surrounding organs delineation

Future studies could include additional OARs in the analysis. This expansion would allow for a more comprehensive assessment of the potential side effects and risks associated with radiation therapy, particularly about cumulative doses received by various critical structures. When balancing dose to the heart and lungs, normal tissue sparing becomes a complex trade-off in thoracic radiotherapy. The goal is to minimize radiation exposure to both organs, but reducing dose to one can inadvertently increase dose to the other. For instance, reducing heart dose in left-sided breast cancer using DIBH or other techniques may increase lung exposure, potentially leading to pulmonary complications. Similarly, prioritizing lung sparing could elevate cardiac risks, such as ischemic heart disease. This delicate balance requires careful consideration of organ sensitivities and potential long-term toxicities.

The integration of IBDM techniques can aid in navigating these trade-offs by identifying subtle dose-response patterns in both cardiac and pulmonary tissues. By delineating critical structures within both organs and analyzing how different treatment plans affect these structures, clinicians can make more informed decisions about dose allocation. Future studies could expand the analysis to include a broader range of OARs and integrate machine learning to develop predictive models that optimize normal tissue sparing while maintaining effective tumor coverage.

7.3 Additional Late Effects

Expanding the scope of outcomes to include additional thoracic RT-related late effects could contribute to a more thorough understanding of long-term consequences for cancer patients. While IHD, HF, and VD were explored in this work, there are more potential impacts of RT on the ***cardiovasular system***. Radiation can lead to acute or chronic complications due to inflammation of the pericardium (pericarditis), with the potential to cause pericardial effusion or even constrictive pericarditis. Additionally, radiation can induce damage to the major blood vessels increasing the risk of stroke and peripheral vascular disease years after treatment. These vascular issues, although less frequently discussed than direct cardiac damage, represent significant long-term risks for thoracically irradiated patients.

In the ***pulmonary system***, radiation pneumonitis can result from thoracic RT, particularly in lung cancer patients. This acute inflammation of the lung tissue can progress to pulmonary fibrosis (poor functioning and scarring lungs). Chronic obstructive pulmonary disease (COPD) or other preexisting lung conditions can increase a patients risk of exacerbated symptoms or the development of new chronic lung diseases following radiation. Long-term monitoring of these patients is critical to detect and manage the onset of chronic pulmonary conditions that may reduce quality of life and overall survival.

Immune system effects are an key consideration. Radiation-induced lymphopenia is common during and after thoracic RT, with studies estimating that a significant portion of circulating lymphocytes can receive substantial radiation doses. This depletion of lymphocytes compromises the immune system, leading to an increased risk of infections, potentially reducing the effectiveness of concurrent or subsequent immunotherapy [6]. Further, the immunosuppressive effects of lymphopenia can exacerbate other treatment-related toxicities and negatively impact OS, underscoring the importance of monitoring immune function during and after treatment.

7.4 Additional Sources of Multi-Modal Data

7.4.1 Treatment Details

This research has focused primarily on the cardiotoxic effects of RT, but it's important to acknowledge that cancer patients also face significant cardiac risks from chemotherapy. Approximately 30-50% of irradiated cancer patients receive chemotherapy, including adjuvant, neoadjuvant, or concomitant regimens. In the case of breast cancer, 50-60% of patients undergoing chemotherapy are treated with anthracyclines, a class of drugs known for their efficacy but also for their heightened risk of cardiotoxicity, particularly when combined with RT [132]. Incorporating chemotherapy data into the analysis could have provided better control for these confounding factors. However, challenges with pharmaceutical database integration prevented its inclusion within the timeframe of interest. Expanding this analysis to include chemotherapy data could yield a more thorough understanding of the cumulative cardiac risks across various cancer diagnoses [133].

Further, incorporating details about advanced treatments could have provided a more comprehensive understanding of cardiotoxicity risks. For example, the use of targeted therapies like trastuzumab, which is known for its cardiotoxic effects, and immunotherapies like checkpoint inhibitors, could have influenced cardiac outcomes when combined with RT.

7.4.2 Disease Characteristics

Expanding the available data with additional pathology and tumor staging information could have significantly enhanced the precision of this analysis, particularly in addressing the heterogeneity in diagnostic groupings. Tumor staging, such as TNM classification, and detailed tumor delineations could have provided essential insights into disease progression and its impact on treatment outcomes. For example, understanding whether patients had early-stage localized tumors versus more advanced metastatic disease would have allowed for more accurate adjustments in the analysis, particularly for confounding factors like regional nodal irradiation, which can influence both treatment planning and dose distributions.

The absence of staging and receptor status data limits our ability to fully control for these potential confounders, which could obscure important relationships between radiotherapy and cardiac outcomes. Furthermore, when exploring the relationship between OS and heart dose, preliminary findings from Study II suggest that OS does not serve as a good surrogate for RT-related heart damage, as it is the disease that dominates OS. This could be due to the multifactorial nature of OS, which is influenced by a variety of factors beyond cardiotoxicity, such as disease progression, response to therapy, and other treatment-related toxicities. Therefore, more focused endpoints, such as cardiac-specific morbidity or mortality, may provide a clearer picture of the cardiotoxic effects of RT.

7.4.3 Patient Characteristics

A comprehensive analysis of CVD risk factors could greatly enhance the understanding of variability in treatment responses among cancer patients undergoing radiotherapy. In addition to conventional risk factors like hypertension, smoking status, and obesity, detailed patient characteristics such as genetic profiles, lifestyle factors, and pre-existing co-morbidities should be considered. For example, coronary artery calcium (CAC) scoring is a valuable predictor of increased CVD risk, particularly in patients already exposed to radiation [134]. CAC scoring, when integrated with other imaging and biomarker data, could provide a more nuanced assessment of cardiotoxicity risk. Further, there are emerging automated methods which would make it possible to acquire this measure for a large cohort.

Additionally, inclusion of more comprehensive data on risk factors would be valuable. We did have access to indices like the Charlson Comorbidity Index (CCI) to help in quantifying the burden of co-morbidities and accounts for various conditions (e.g. diabetes, prior myocardial infarction, and chronic pulmonary disease) that increase risks of treatment-related issues [135, 136] . We hoped it could have allowed for better stratification of patients based on their overall health status and pre-existing conditions, which directly impact their response to cancer treatment and their risk of developing treatment-related CVD. However, it was insufficient and unreliable and lacked the detail needed for such an investigation, therefore further metrics are needed.

However, much of this detailed information was not readily available for the cohort under study. Acquiring data on genetic profiles, lifestyle factors, and co-morbidities, as well as scoring indices like CAC, would require a significant review of patient charts and would likely be more feasible in a smaller, more focused subset of the population. This detailed data collection could ultimately enhance the predictive models for cardiotoxicity, allowing for more personalized and targeted interventions that take into account both the cancer treatment and the overall patient health status.

Conclusion

8

> *. . . the disciplines are the place where we begin, but not where we end.*
>
> **— Allen F. Repko**

This thesis demonstrates the immense potential of collaborations between computer science and oncology in addressing complex healthcare challenges. The integration of advanced computational techniques into oncological research and clinical practice exemplifies the benefits of multidisciplinary work. Modern methods, including automated segmentation and dose calculation, can significantly improve the accuracy and efficiency of RT planning and delivery, enhancing clinical outcomes. AI models can provide high-quality, patient-specific delineations for large cohorts, a task that would be impractical through manual efforts alone. This collaboration highlights the power of observational studies, which can provide valuable real-world insights and contribute to the continuous improvement of treatment strategies.

As segmentation methods evolve, the availability of open-source AI tools, such as the models leveraged in this work, is expanding. These tools show great promise but also highlight the critical need for ongoing validation, particularly with external patient cohorts. Validation is essential to identify and address potential failures in outlier cases or when dealing with out-of-distribution data. While these occurrences are rare, they are particularly concerning in clinical processes that inform diagnostics and treatment planning. Therefore, human-in-the-loop approaches will remain a vital standard, ensuring that AI tools become robust and reliable assets in clinical settings, where human oversight plays a crucial role in preventing errors.

The findings from this research contribute to the growing body of knowledge on RT-related cardiotoxicity. Although a significant dose-response relationship was not observed within the studied cohort, the work emphasizes the importance of careful cardiac dose management in reducing long-term cardiovascular risks for cancer patients. As the field progresses towards more refined, individualized treatment strategies, contemporary high-quality RT approaches are effectively minimizing heart toxicity, rendering it a minor clinical concern. Future research that delves into more granular dose-response relationships, leveraging substructure delineation or IBDM methods, can further enhance our understanding of the radiosensitivity of cardiac and surrounding tissues. As tools continue to advance, the application of AI in RT is poised to become an increasingly valuable and ubiquitous tool.

Looking forward, the field promises exciting advancements that could revolutionize cancer care. As imaging analysis technologies and computational power and models continue to evolve, the integration of AI in RT is expected to further enhance precision and personalization in treatment planning. The potential for AI to play a pivotal role not only in delineation and dose calculation but also in diagnostics and predictive analytics is becoming more apparent. This would enable more tailored treatments based on individual patient profiles. Additionally, the growing availability of large-scale data will allow for more comprehensive analyses, improving our knowledge about the long-term impacts of RT and other treatments. This progression holds the potential to deliver more effective, safer, and patient-centered cancer care, with AI and data-driven insights leading the way.

Glossary

Source: Encyclopedia Britannica [137]

Aalen-Johansen Estimator A method used in survival analysis to estimate cumulative incidence functions in the presence of competing risks.
Atlas-Based Segmentation A semi-automated technique using pre-segmented anatomical atlases to guide the segmentation of structures in medical images.
Beta-Blockers Medications that reduce blood pressure and heart rate, often used to manage heart conditions.
Cardiac Sparing Techniques and strategies used to minimize radiation exposure to the heart during radiotherapy.
Central Person Register (CPR) A national registry in Denmark containing personal data for statistical and administrative use.
Dose-Volume Histogram (DVH) A graphical representation commonly used in radiation therapy to evaluate the distribution of radiation doses within a target volume (such as a tumor) and surrounding organs at risk (OAR).
Echocardiograms An ultrasound of the heart used to monitor cardiac function and detect abnormalities.
Fully Automated Segmentation Uses advanced algorithms, including machine learning, to automatically segment organs and tumors without manual input.
Kaplan-Meier (KM) A statistical method used to estimate survival probabilities over time.
Landspatientregisteret (LPR) The Danish National Patient Register, containing data on all patients admitted to hospitals in Denmark.
Manual Segmentation The process of manually outlining the boundaries of organs and tumors in medical imaging.
Mean Heart Dose (MHD) The mean dose of radiation received by cardiac tissue during radiotherapy.
Neuroimaging Informatics Technology Initiative (NIfTI) A file format used for storing neuroimaging data.
Organs at Risk (OAR) Healthy tissues that are sensitive to radiation and must be protected during radiotherapy.
Overall Survival (OS) The duration of time from baseline to death from any cause.
Pericarditis Inflammation of the pericardium, the sac surrounding the heart, which can cause chest pain, fever, and abnormal heart rhythms.

Proton Therapy A type of radiotherapy that uses protons instead of X-rays, allowing for more precise targeting of tumors with minimal damage to surrounding tissues.
Semi-Automated Segmentation Combines manual input with automated algorithms to delineate structures in medical imaging.
Statistical Analysis Plan (SAP) A document detailing the statistical methods to be used in analyzing data from a study.
Valvular Disease Disorders involving the heart valves, including conditions such as aortic stenosis and mitral regurgitation.
V5 and V30 The volume of the cardiac tissue receiving 5 or 30 Gy, respectively. of radiation at least

Bibliography

[1] Nora Forbes, Cynthia Terrones-Campos, Abraham Smith, et al. „Time trends in cardiac doses in 10,000 patients receiving curative thoracic radiation therapy between 2009 and 2020“. In: *medRxiv* (2024) (cit. on pp. ix, 30, 33, 37).

[2] Nora Forbes, Cynthia Terrones-Campos, Abraham Smith, et al. „Cardiac dose volume analysis of 9,411 patients with registry-based outcome data for cardiotoxicity and overall survival“. In: *medRxiv* (2024). eprint: `https://www.medrxiv.org/content/10.1101/2024.08.16.24312108` (cit. on pp. ix, 39).

[3] Nora Forbes. „Clinical and dosimetric risk factors for ischemic heart disease [Statistical Analysis Plan]“. In: *figshare* (Oct. 2023) (cit. on pp. ix, 39).

[4] N Forbes, A Smith, J Petersen, et al. „MO-0716 Radiotherapy exposure and association with observed cardiovascular toxicity in over 5000 patients [Abstract]“. In: *Radiotherapy and Oncology* 170 (2022), S627–S628 (cit. on p. ix).

[5] Abraham George Smith, Jens Petersen, Cynthia Terrones-Campos, et al. „RootPainter3D: Interactive-machine-learning enables rapid and accurate contouring for radiotherapy“. In: *Medical Physics* 49 (1 Jan. 2022), pp. 461–473 (cit. on pp. ix, 29).

[6] C. Terrones-Campos, B. Ledergerber, N. Forbes, et al. „Prediction of Radiation-induced Lymphopenia following Exposure of the Thoracic Region and Associated Risk of Infections and Mortality“. In: *Clinical Oncology* 35 (7 July 2023), e434–e444 (cit. on pp. ix, 52).

[7] Freddie Bray, Mathieu Laversanne, Hyuna Sung, et al. „Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries“. In: *CA: A Cancer Journal for Clinicians* 74 (3 May 2024), pp. 229–263 (cit. on p. 1).

[8] Sophie Pilleron, Enrique Soto-Perez-de-Celis, Jerome Vignat, et al. „Estimated global cancer incidence in the oldest adults in 2018 and projections to 2050“. In: *International Journal of Cancer* 148 (3 Feb. 2021), pp. 601–608 (cit. on p. 1).

[9] Kimberly D. Miller, Leticia Nogueira, Theresa Devasia, et al. „Cancer treatment and survivorship statistics, 2022“. In: *CA: A Cancer Journal for Clinicians* 72 (5 Sept. 2022), pp. 409–436 (cit. on p. 1).

[10] Raymond J. Chan, Vivienne E. Milch, Fiona Crawford-Williams, et al. „Patient navigation across the cancer care continuum: An overview of systematic reviews and emerging literature“. In: *CA: A Cancer Journal for Clinicians* 73 (6 Nov. 2023), pp. 565–589 (cit. on p. 2).

[11]Peter Barrett, Louise E Mullen, and Triona Mccarthy. „Survivorship after Childhood Cancer in Ireland - NCCP National Needs Assessment“. In: *National Cancer Control Programme* (2019) (cit. on p. 2).

[12]Anthony Edwards, Brendan Mcclean, Dave Meehan, Donal Murphy, and Andres Poveda. *NATIONAL CANCER STRATEGY: 2017 - 2026* (cit. on p. 2).

[13]Douglas Hanahan and Robert A. Weinberg. „Hallmarks of Cancer: The Next Generation“. In: *Cell* 144 (5 Mar. 2011), pp. 646–674 (cit. on pp. 2, 3).

[14]Farhad Islami, Emily C. Marlow, Blake Thomson, et al. „Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States, 2019“. In: *CA: A Cancer Journal for Clinicians* (July 2024) (cit. on p. 2).

[15]*Centers for Disease Control and Prevention (CDC). Cancer Risk Factors. CDC, 2024.* (Cit. on p. 2).

[16]Yiwey Shieh, Martin Eklund, George F. Sawaya, et al. „Population-based screening for cancer: hope and hype“. In: *Nature Reviews Clinical Oncology* 13 (9 Sept. 2016), pp. 550–565 (cit. on p. 3).

[17]Nigel Hawkes. „Cancer survival data emphasise importance of early diagnosis“. In: *BMJ* (Jan. 2019), p. l408 (cit. on p. 3).

[18]*Types of cancer*. June 2024. URL: https://www.cancerresearchuk.org/about-cancer/what-is-cancer/how-cancer-starts/types-of-cancer (cit. on p. 3).

[19]F. L. Greene and L. H. Sobin. „The Staging of Cancer: A Retrospective and Prospective Appraisal“. In: *CA: A Cancer Journal for Clinicians* 58 (3 May 2008), pp. 180–190 (cit. on pp. 3, 50).

[20]Mahul B. Amin, Frederick L. Greene, Stephen B. Edge, et al. „The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging“. In: *CA: A Cancer Journal for Clinicians* 67 (2 Mar. 2017), pp. 93–99 (cit. on p. 3).

[21]Carolyn Compton. *Cancer: The Enemy from Within*. Springer International Publishing, 2020 (cit. on p. 3).

[22]Jr. DeVita Vincient T., Theodore S. Hellman, and Steven A. Rosenberg. *Cancer: Principles and Practice of Oncology*. 11th. Wolters Kluwer, 2018 (cit. on p. 3).

[23]L. J. Kleinsmith. *Principles of Cancer Biology*. 1st. Pearson, 2013 (cit. on p. 3).

[24]Seymour H. Levitt, James A. Purdy, Carlos A. Perez, and Philip Poortmans. *Radiation Oncology Series Technical Basis of Radiation Therapy: Practical Clinical Applications*. 2012 (cit. on p. 4).

[25]Ben G. L. Vanneste, Evert J. Van Limbergen, Emile N. van Lin, Joep G. H. van Roermund, and Philippe Lambin. „Prostate Cancer Radiation Therapy: What Do Clinicians Have to Know?“ In: *BioMed Research International* 2016 (2016), pp. 1–14 (cit. on pp. 5, 42).

[26]Bozica Vujosevic and Bojana Bokorov. „Radiotherapy: Past and present“. In: *Archive of Oncology* 18 (4 2010), pp. 140–142 (cit. on p. 5).

[27]Eric J. Hall and Amato J. Giaccia. *Radiobiology for the radiologist: Seventh edition*. 2012 (cit. on p. 5).

[28] Stephen G. Chun, Chen Hu, Hak Choy, et al. „Impact of Intensity-Modulated Radiation Therapy Technique for Locally Advanced Non–Small-Cell Lung Cancer: A Secondary Analysis of the NRG Oncology RTOG 0617 Randomized Clinical Trial“. In: *Journal of Clinical Oncology* 35 (1 Jan. 2017), pp. 56–62 (cit. on p. 5).

[29] Francesco Sciacca. „Organs at risk“. In: Radiopaedia.org, July 2020 (cit. on p. 5).

[30] Leonard L. Gunderson and Joel E. Tepper. *Clinical Radiation Oncology*. Elsevier, Inc., 2021 (cit. on p. 5).

[31] Feng-Ming (Spring) Kong, Timothy Ritter, Douglas J. Quint, et al. „Consideration of Dose Limits for Organs at Risk of Thoracic Radiotherapy: Atlas for Lung, Proximal Bronchial Tree, Esophagus, Spinal Cord, Ribs, and Brachial Plexus“. In: *International Journal of Radiation Oncology*Biology*Physics* 81 (5 Dec. 2011), pp. 1442–1457 (cit. on p. 5).

[32] Meng Xu-Welliver, William T. C. Yuh, Julia R. Fielding, et al. „Imaging across the Life Span: Innovations in Imaging and Therapy for Gynecologic Cancer“. In: *RadioGraphics* 34 (4 July 2014), pp. 1062–1081 (cit. on p. 6).

[33] Anders Ahnesjö and Maria Mania Aspradakis. „Dose calculations for external photon beams in radiotherapy“. In: *Physics in Medicine and Biology* 44 (11 Nov. 1999), R99–R155 (cit. on p. 6).

[34] Yun Zhang, Yuling Huang, Shenggou Ding, et al. „A dosimetric and radiobiological evaluation of VMAT following mastectomy for patients with left-sided breast cancer“. In: *Radiation Oncology* 16 (Sept. 2021) (cit. on p. 7).

[35] Nan Li, Ruben Carmona, Igor Sirak, et al. „Highly Efficient Training, Refinement, and Validation of a Knowledge-based Planning Quality-Control System for Radiation Therapy Clinical Trials“. In: *International Journal of Radiation Oncology*Biology*Physics* 97 (1 Jan. 2017), pp. 164–172 (cit. on p. 6).

[36] Gordon Steel. *Basic Clinical Radiobiology*. 3rd. Hodder Education Publisher, 2002 (cit. on pp. 6, 7, 34).

[37] Lawrence B. Marks, Ellen D. Yorke, Andrew Jackson, et al. „Use of Normal Tissue Complication Probability Models in the Clinic“. In: *International Journal of Radiation Oncology*Biology*Physics* 76 (3 Mar. 2010), S10–S19 (cit. on p. 7).

[38] *Danske Multidisciplinære Cancer Grupper*. Accessed: 2024-07-31. 2024. URL: `\url{https://www.dmcg.dk/}` (cit. on p. 7).

[39] F. M. Khan and J. P. Gibbons. *The Physics of Radiation Therapy*. Lippincott Williams and Wilkins, 2014 (cit. on p. 8).

[40] M. Durante and J. S. Loeffler. „Particle therapy for cancer“. In: *Nature Reviews Clinical Oncology* 7 (1 2010), pp. 37–43 (cit. on p. 8).

[41] Björn Dahlöf. „Cardiovascular Disease Risk Factors: Epidemiology and Risk Assessment“. In: *The American Journal of Cardiology* 105 (1 Jan. 2010), 3A–9A (cit. on p. 11).

[42] Carrie G. Lenneman and Douglas B. Sawyer. „Cardio-Oncology“. In: *Circulation Research* 118 (6 Mar. 2016), pp. 1008–1020 (cit. on p. 12).

[43] Joerg Herrmann, Amir Lerman, Nicole P. Sandhu, et al. „Evaluation and Management of Patients With Heart Disease and Cancer: Cardio-Oncology“. In: *Mayo Clinic Proceedings* 89 (9 Sept. 2014), pp. 1287–1306 (cit. on p. 12).

[44] Adriana Albini, Giuseppina Pennesi, Francesco Donatelli, et al. „Cardiotoxicity of Anticancer Drugs: The Need for Cardio-Oncology and Cardio-Oncological Prevention“. In: *JNCI: Journal of the National Cancer Institute* 102 (1 Jan. 2010), pp. 14–25 (cit. on p. 12).

[45] David G Gent and Rebecca Dobson. „The 2022 European Society of Cardiology Cardio-oncology Guidelines in Focus“. In: *European Cardiology Review* 18 (Apr. 2023) (cit. on p. 12).

[46] Vivek Verma, Charles Simone, and Maria Werner-Wasik. „Acute and Late Toxicities of Concurrent Chemoradiotherapy for Locally-Advanced Non-Small Cell Lung Cancer“. In: *Cancers* 9 (12 Sept. 2017), p. 120 (cit. on p. 12).

[47] Kyle Wang, Kevin A. Pearlstein, Nicholas D. Patchett, et al. „Heart dosimetric analysis of three types of cardiac toxicity in patients treated on dose-escalation trials for Stage III non-small-cell lung cancer“. In: *Radiotherapy and Oncology* 125 (2 Nov. 2017), pp. 293–300 (cit. on p. 12).

[48] Carmen Bergom, Julie A. Bradley, Andrea K. Ng, et al. „Past, Present, and Future of Radiation-Induced Cardiotoxicity: Refinements in Targeting, Surveillance, and Risk Stratification“. In: *JACC: CardioOncology* 3 (3 Sept. 2021), pp. 343–359 (cit. on pp. 12, 13).

[49] KEITH E. COHN, J ROBERT STEWART, LUIS F. FAJARDO, and E WILLIAM HANCOCK. „Heart disease following radiation“. In: *Medicine* 3 (46 May 1967), pp. 281–298 (cit. on p. 12).

[50] E. William Hancock. „Heart Disease after Radiation“. In: *New England Journal of Medicine* 308 (10 Mar. 1983), pp. 588–588 (cit. on p. 12).

[51] J Cuzick, H Stewart, L Rutqvist, et al. „Cause-specific mortality in long-term survivors of breast cancer who participated in trials of radiotherapy.“ In: *Journal of Clinical Oncology* 12 (3 Mar. 1994), pp. 447–453 (cit. on p. 12).

[52] G Gagliardi, I Lax, A Ottolenghi, and L E Rutqvist. „Long-term cardiac mortality after radiotherapy of breast cancer—application of the relative seriality model“. In: *The British Journal of Radiology* 69 (825 Sept. 1996), pp. 839–846 (cit. on p. 12).

[53] Sarah C Darby, Paul McGale, Carolyn W Taylor, and Richard Peto. „Long-term mortality from heart disease and lung cancer after radiotherapy for early breast cancer: prospective cohort study of about 300000 women in US SEER cancer registries“. In: *The Lancet Oncology* 6 (8 Aug. 2005), pp. 557–565 (cit. on pp. 12, 42).

[54] Sarah C Darby, Marianne Ewertz, Paul McGale, et al. „Risk of ischemic heart disease in women after radiotherapy for breast cancer.“ In: *The New England journal of medicine* 368 (11 Mar. 2013), pp. 987–998 (cit. on pp. 12, 13, 23).

[55] Frederika A. Van Nimwegen, Michael Schaapveld, David J. Cutter, et al. „Radiation dose-response relationship for risk of coronary heart disease in survivors of Hodgkin lymphoma“. In: *Journal of Clinical Oncology* 34 (3 Jan. 2016), pp. 235–243 (cit. on pp. 12, 23).

[56] Maja V. Maraldo, Francesco Giusti, Ivan R. Vogelius, et al. „Cardiovascular disease after treatment for Hodgkin's lymphoma: An analysis of nine collaborative EORTC-LYSA trials". In: *The Lancet Haematology* 2 (11 2015) (cit. on pp. 12, 23).

[57] Kareena M. Menezes, Huichen Wang, Megumi Hada, and Premkumar B. Saganti. „Radiation Matters of the Heart: A Mini Review". In: *Frontiers in Cardiovascular Medicine* 5 (July 2018) (cit. on p. 12).

[58] Mirjana Josipovic, Marianne C Aznar, Jakob B Thomsen, et al. „Deep inspiration breath hold in locally advanced lung cancer radiotherapy: validation of intrafractional geometric uncertainties in the INHALE trial". In: *The British Journal of Radiology* 92 (1104 Dec. 2019), p. 20190569 (cit. on p. 13).

[59] Carolyn W. Taylor, Zhe Wang, Elizabeth Macaulay, et al. „Exposure of the Heart in Breast Cancer Radiation Therapy: A Systematic Review of Heart Doses Published During 2003 to 2013". In: *International Journal of Radiation Oncology*Biology*Physics* 93 (4 Nov. 2015), pp. 845–853 (cit. on pp. 13, 23).

[60] Lara Barazzuol, Rob P. Coppes, and Peter van Luijk. „Prevention and treatment of radiotherapy-induced side effects". In: *Molecular Oncology* 14 (7 July 2020), pp. 1538–1554 (cit. on p. 13).

[61] Chirag Mehta, Puneet Singh, and Jess Brar. „Contemporary review: recognition, management, and screening for radiation-induced heart disease". In: *Exploration of Medicine* (Oct. 2023), pp. 772–781 (cit. on pp. 13, 14).

[62] Michael C. Joiner and Albert J. van der Kogel. *Basic clinical radiobiology*. CRC press, 2018 (cit. on p. 13).

[63] Drew Latty, Kirsty E. Stuart, Wei Wang, and Verity Ahern. „Review of deep inspiration breath-hold techniques for the treatment of breast cancer". In: *Journal of Medical Radiation Sciences* 62 (1 Mar. 2015), pp. 74–81 (cit. on p. 13).

[64] Lloyd M. Smyth, Kellie A. Knight, Yolanda K. Aarons, and Jason Wasiak. „The cardiac dose-sparing benefits of deep inspiration breath-hold in left breast irradiation: a systematic review". In: *Journal of Medical Radiation Sciences* 62 (1 Mar. 2015), pp. 66–73 (cit. on p. 13).

[65] V. Salvestrini, G. C. Iorio, P. Borghetti, et al. „The impact of modern radiotherapy on long-term cardiac sequelae in breast cancer survivor: a focus on deep inspiration breath-hold (DIBH) technique". In: *Journal of Cancer Research and Clinical Oncology* 148 (2 Feb. 2022), pp. 409–417 (cit. on pp. 13, 23).

[66] Anders N. Pedersen, Stine Korreman, Håkan Nyström, and Lena Specht. „Breathing adapted radiotherapy of breast cancer: Reduction of cardiac and pulmonary doses using voluntary inspiration breath-hold". In: *Radiotherapy and Oncology* 72 (1 July 2004), pp. 53–60 (cit. on p. 13).

[67] Henrik D. Nissen and Ane L. Appelt. „Improved heart, lung and target dose with deep inspiration breath hold in a large clinical series of breast cancer patients". In: *Radiotherapy and Oncology* 106 (1 Jan. 2013), pp. 28–32 (cit. on p. 13).

[68] Robert W. Mutter, J. Isabelle Choi, Rachel B. Jimenez, et al. „Proton Therapy for Breast Cancer: A Consensus Statement From the Particle Therapy Cooperative Group Breast Cancer Subcommittee". In: *International Journal of Radiation Oncology*Biology*Physics* 111 (2 Oct. 2021), pp. 337–359 (cit. on p. 13).

[69] L.J. Boersma, M.G.A. Sattler, J.H. Maduro, et al. „Model-Based Selection for Proton Therapy in Breast Cancer: Development of the National Indication Protocol for Proton Therapy and First Clinical Experiences“. In: *Clinical Oncology* 34 (4 Apr. 2022), pp. 247–257 (cit. on p. 13).

[70] Line Bjerregaard Stick, Ebbe Laugaard Lorenzen, Esben Svitzer Yates, et al. „Selection criteria for early breast cancer patients in the DBCG proton trial – The randomised phase III trial strategy“. In: *Clinical and Translational Radiation Oncology* 27 (Mar. 2021), pp. 126–131 (cit. on p. 13).

[71] Charles B. Simone, Jeffrey Bradley, Aileen B. Chen, et al. „ASTRO Radiation Therapy Summary of the ASCO Guideline on Management of Stage III Non-Small Cell Lung Cancer“. In: *Practical Radiation Oncology* 13 (3 May 2023), pp. 195–202 (cit. on p. 14).

[72] Manish A. Shah, Erin B. Kennedy, Daniel V. Catenacci, et al. *Treatment of locally advanced esophageal carcinoma: ASCO guideline*. 2020 (cit. on p. 14).

[73] Larissa A. Korde, Mark R. Somerfield, Lisa A. Carey, et al. „Neoadjuvant Chemotherapy, Endocrine Therapy, and Targeted Therapy for Breast Cancer: ASCO Guideline“. In: *Journal of Clinical Oncology* 39 (13 2021) (cit. on p. 14).

[74] Megan E. Daly, Navneet Singh, Nofisat Ismaila, et al. „Management of Stage III Non-Small-Cell Lung Cancer: ASCO Guideline“. In: *Journal of Clinical Oncology* 40 (12 2022) (cit. on p. 14).

[75] Saro H. Armenian, Christina Lacchetti, Ana Barac, et al. *Prevention and monitoring of cardiac dysfunction in survivors of adult cancers: American society of clinical oncology clinical practice guideline*. 2017 (cit. on p. 14).

[76] G. Curigliano, D. Lenihan, M. Fradley, et al. „Management of cardiac disease in cancer patients throughout oncological treatment: ESMO consensus recommendations“. In: *Annals of Oncology* 31 (2 Feb. 2020), pp. 171–190 (cit. on p. 14).

[77] Jinzhong Yang, Harini Veeraraghavan, Wouter van Elmpt, et al. „CT images with expert manual contours of thoracic cancer for benchmarking auto-segmentation accuracy“. In: *Medical Physics* 47 (7 July 2020), pp. 3250–3255 (cit. on p. 15).

[78] Aurélie Isambert, Frédéric Dhermain, François Bidault, et al. „Evaluation of an atlas-based automatic segmentation software for the delineation of brain organs at risk in a radiation therapy clinical context“. In: *Radiotherapy and Oncology* 87 (1 Apr. 2008), pp. 93–99 (cit. on p. 15).

[79] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. „U-net: Convolutional networks for biomedical image segmentation“. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 9351. 2015 (cit. on p. 15).

[80] Xuming Chen, Shanlin Sun, Narisu Bai, et al. „A deep learning-based auto-segmentation system for organs-at-risk on whole-body computed tomography images for radiation therapy“. In: *Radiotherapy and Oncology* 160 (July 2021), pp. 175–184 (cit. on p. 15).

[81] Malvika Ashok and Abhishek Gupta. „A Systematic Review of the Techniques for the Automatic Segmentation of Organs-at-Risk in Thoracic Computed Tomography Images“. In: *Archives of Computational Methods in Engineering* 28 (4 2021) (cit. on p. 16).

[82] Carsten B Pedersen. „The Danish Civil Registration System“. In: *Scandinavian journal of public health* 39.7_suppl (2011), pp. 22–25 (cit. on p. 16).

[83] Sissel Toft Sørensen, Frederik Pagh Kristensen, Frederikke Schønfeldt Troelsen, Morten Schmidt, and Henrik Toft Sørensen. *Health registries as research tools: a review of methodological key issues*. 2023 (cit. on p. 17).

[84] Ivan R. Vogelius, Jens Petersen, and Søren M. Bentzen. „Harnessing data science to advance radiation oncology“. In: *Molecular Oncology* 14 (7 July 2020), pp. 1514–1528 (cit. on pp. 17, 18).

[85] Carrol Gamble, Ashma Krishan, Deborah Stocken, et al. „Guidelines for the Content of Statistical Analysis Plans in Clinical Trials“. In: *JAMA* 318 (23 Dec. 2017), p. 2337 (cit. on p. 18).

[86] Sushmita Rai, Prabhakar Mishra, and Uday C. Ghoshal. „Survival analysis: A primer for the clinician scientists“. In: *Indian Journal of Gastroenterology* 40 (5 Oct. 2021), pp. 541–549 (cit. on p. 18).

[87] Loai Albarqouni. *Tutorial about hazard ratios*. 2018. URL: https://s4be.cochrane.org/blog/2016/04/05/tutorial-hazard-ratios/ (cit. on p. 19).

[88] E. L. Kaplan and Paul Meier. „Nonparametric Estimation from Incomplete Observations“. In: *Journal of the American Statistical Association* 53 (282 June 1958), p. 457 (cit. on p. 19).

[89] David G. Kleinbaum and Mitchel Klein. *Survival Analysis*. Springer New York, 2012 (cit. on p. 19).

[90] Martin J. Crowder. *Multivariate Survival Analysis and Competing Risks*. Chapman and Hall/CRC, Apr. 2012 (cit. on p. 19).

[91] Søren M. Bentzen and Ivan R. Vogelius. „Using and Understanding Survival Statistics - or How We Learned to Stop Worrying and Love the Kaplan-Meier Estimate“. In: *International Journal of Radiation Oncology*Biology*Physics* 115 (4 Mar. 2023), pp. 839–846 (cit. on pp. 20, 21).

[92] D. R. Cox. „Regression Models and Life-Tables“. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 34 (2 Jan. 1972), pp. 187–202 (cit. on p. 20).

[93] Jason P. Fine and Robert J. Gray. „A Proportional Hazards Model for the Subdistribution of a Competing Risk“. In: *Journal of the American Statistical Association* 94 (446 1999) (cit. on p. 20).

[94] Søren M. Bentzen, Ivan R. Vogelius, David Hodgson, et al. „Radiation Dose-Volume-Response Relationships for Adverse Events in Childhood Cancer Survivors: Introduction to the Scientific Issues in PENTEC“. In: *International Journal of Radiation Oncology*Biology*Physics* 119 (2 June 2024), pp. 338–353 (cit. on p. 22).

[95] David J. Cutter, Michael Schaapveld, Sarah C. Darby, et al. „Risk for valvular heart disease after treatment for hodgkin lymphoma“. In: *Journal of the National Cancer Institute* 107 (4 Apr. 2015) (cit. on pp. 23, 46).

[96] Takafumi Nemoto, Natsumi Futakami, Masamichi Yagi, et al. „Efficacy evaluation of 2D, 3D U-Net semantic segmentation and atlas-based segmentation of normal lungs excluding the trachea and main bronchi“. In: *Journal of Radiation Research* 61 (2 Mar. 2020), pp. 257–264 (cit. on p. 23).

[97] Jinhan Zhu, Jun Zhang, Bo Qiu, et al. „Comparison of the automatic segmentation of multiple organs at risk in CT images of lung cancer between deep convolutional neural network-based and atlas-based techniques“. In: *Acta Oncologica* 58 (2 Feb. 2019), pp. 257–264 (cit. on p. 23).

[98] Søren M. Bentzen, Louis S. Constine, Joseph O. Deasy, et al. „Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC): An Introduction to the Scientific Issues“. In: *International Journal of Radiation Oncology*Biology*Physics* 76 (3 Mar. 2010), S3–S9 (cit. on p. 25).

[99] Cynthia Terrones-Campos, Bruno Ledergerber, Ivan Richter Vogelius, et al. „Lymphocyte Count Kinetics, Factors Associated with the End-of-Radiation-Therapy Lymphocyte Count, and Risk of Infection in Patients with Solid Malignant Tumors Treated with Curative-Intent Radiation Therapy“. In: *International Journal of Radiation Oncology Biology Physics* 105 (4 Nov. 2019), pp. 812–823 (cit. on p. 29).

[100] Robert Finnegan, Jason Dowling, Eng-Siew Koh, et al. „Feasibility of multi-atlas cardiac segmentation from thoracic planning CT in a probabilistic framework“. In: *Physics in Medicine and Biology* 64 (8 Apr. 2019), p. 085006 (cit. on pp. 30, 49).

[101] Robert N. Finnegan, Vicky Chin, Phillip Chlap, et al. „Open-source, fully-automated hybrid cardiac substructure segmentation: development and optimisation“. In: *Physical and Engineering Sciences in Medicine* 46 (1 Mar. 2023), pp. 377–393 (cit. on p. 30).

[102] Charles S. Mayo, Jean M. Moran, Walter Bosch, et al. „American Association of Physicists in Medicine Task Group 263: Standardizing Nomenclatures in Radiation Oncology“. In: *International Journal of Radiation Oncology*Biology*Physics* 100 (4 Mar. 2018), pp. 1057–1066 (cit. on p. 31).

[103] Jakob Wasserthal, Hanns-Christian Breit, Manfred T. Meyer, et al. „TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images“. In: *Radiology: Artificial Intelligence* 5 (5 Sept. 2023) (cit. on pp. 31, 49).

[104] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. „nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation“. In: *Nature Methods* 18 (2 Feb. 2021), pp. 203–211 (cit. on p. 31).

[105] V. Chin, R.N. Finnegan, P. Chlap, et al. „Validation of a Fully Automated Hybrid Deep Learning Cardiac Substructure Segmentation Tool for Contouring and Dose Evaluation in Lung Cancer Radiotherapy“. In: *Clinical Oncology* 35 (6 June 2023), pp. 370–381 (cit. on p. 32).

[106] Fabian Isensee, Tassilo Wald, Constantin Ulrich, et al. „nnU-Net Revisited: A Call for Rigorous Validation in 3D Medical Image Segmentation“. In: (Apr. 2024) (cit. on p. 32).

[107] Alexander Jaus, Constantin Seibold, Kelsey Hermann, et al. „Towards Unifying Anatomy Segmentation: Automated Generation of a Full-body CT Dataset via Knowledge Aggregation and Anatomical Guidelines“. In: (July 2023) (cit. on p. 32).

[108] Barry N Taylor, Ambler Thompson, A K Cd Mol Kg, and S Si. „The International System of Units (SI) NIST Special Publication 330 2008 Edition“. In: *NIST Special Publication 330* 330 (1 2008) (cit. on p. 33).

[109] varian. *Ethos*. URL: `https://www.varian.com/products/adaptive-therapy/ethos` (cit. on p. 42).

[110] Guillaume Landry, Christopher Kurz, and Alberto Traverso. „The role of artificial intelligence in radiotherapy clinical practice“. In: *BJR|Open* 5 (1 2023) (cit. on p. 42).

[111] Stanislav Nikolov, Sam Blackwell, Alexei Zverovitch, et al. „Clinically applicable segmentation of head and neck anatomy for radiotherapy: Deep learning algorithm development and validation study“. In: *Journal of Medical Internet Research* 23 (7 2021) (cit. on p. 42).

[112] Bulat Ibragimov and Lei Xing. „Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks:“ in: *Medical Physics* 44 (2 2017) (cit. on p. 42).

[113] Joshua R. Niska, Cameron S. Thorpe, Sorcha M. Allen, et al. „Radiation and the heart: systematic review of dosimetry and cardiac endpoints“. In: *Expert Review of Cardiovascular Therapy* 16 (12 Dec. 2018), pp. 931–950 (cit. on p. 42).

[114] Andrew Green, Eliana Vasquez Osorio, Marianne C. Aznar, Alan McWilliam, and Marcel van Herk. „Image Based Data Mining Using Per-voxel Cox Regression“. In: *Frontiers in Oncology* 10 (July 2020) (cit. on pp. 42, 49, 50).

[115] Carsten Brink, Uffe Bernchou, Anders Bertelsen, et al. „Causal relation between heart irradiation and survival of lung cancer patients after radiotherapy“. In: *Radiotherapy and Oncology* 172 (July 2022), pp. 126–133 (cit. on p. 42).

[116] Alan McWilliam, Jason Kennedy, Clare Hodgson, et al. „Radiation dose to heart base linked with poorer survival in lung cancer patients“. In: *European Journal of Cancer* 85 (Nov. 2017), pp. 106–113 (cit. on p. 42).

[117] F. Sun, K. Franks, L. Murray, et al. „Cardiovascular mortality and morbidity following radical radiotherapy for lung cancer: Is cardiovascular death under-reported?“ In: *Lung Cancer* 146 (Aug. 2020), pp. 1–5 (cit. on p. 42).

[118] Robert T. Dess, Yilun Sun, Martha M. Matuszak, et al. „Cardiac Events After Radiation Therapy: Combined Analysis of Prospective Multicenter Trials for Locally Advanced Non–Small-Cell Lung Cancer“. In: *Journal of Clinical Oncology* 35 (13 May 2017), pp. 1395–1402 (cit. on p. 43).

[119] Kathryn Banfill, Meredith Giuliani, Marianne Aznar, et al. „Cardiac Toxicity of Thoracic Radiotherapy: Existing Evidence and Future Directions“. In: *Journal of Thoracic Oncology* 16 (2 Feb. 2021), pp. 216–227 (cit. on p. 43).

[120] *Danish Comprehensive Cancer Center*. URL: `https://www.dccc.dk/` (cit. on p. 45).

[121] Simon Long Krogh, Carsten Brink, Ebbe Laugaard Lorenzen, et al. „A national repository of complete radiotherapy plans: design, Results, and experiences“. In: *Acta Oncologica* 62 (10 2023) (cit. on p. 45).

[122] *The DICOM Collaboration system*. URL: `https://dcmcollab.rsyd.dk/` (cit. on p. 45).

[123] *Previous Öresund Workshop Meetings*. June 2023. URL: `www.msf.lu.se/previous-oresund-workshop-meetings` (cit. on p. 45).

[124] Bradford S. Hoppe, James E. Bates, Nancy P. Mendenhall, et al. „The Meaningless Meaning of Mean Heart Dose in Mediastinal Lymphoma in the Modern Radiation Therapy Era“. In: *Practical Radiation Oncology* 10 (3 2020) (cit. on p. 46).

[125]Barbara Stam, Heike Peulen, Matthias Guckenberger, et al. „Dose to heart substructures is associated with non-cancer death after SBRT in stage I–II NSCLC patients“. In: *Radiotherapy and Oncology* 123 (3 June 2017), pp. 370–375 (cit. on p. 46).

[126]Katelyn M. Atkins, Bhupendra Rawal, Tafadzwa L. Chaunzwa, et al. „Cardiac Radiation Dose, Cardiac Disease, and Mortality in Patients With Lung Cancer“. In: *Journal of the American College of Cardiology* 73 (23 June 2019), pp. 2976–2987 (cit. on pp. 47, 49).

[127]Robert Finnegan, Ebbe Laugaard Lorenzen, Jason Dowling, et al. „Analysis of cardiac substructure dose in a large, multi-centre danish breast cancer cohort (the DBCG HYPO trial): Trends and predictive modelling“. In: *Radiotherapy and Oncology* 153 (Dec. 2020), pp. 130–138 (cit. on p. 49).

[128]Oscar Acosta, Gael Drean, Juan D Ospina, et al. „Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy“. In: *Physics in Medicine and Biology* 58 (8 Apr. 2013), pp. 2581–2595 (cit. on p. 49).

[129]Chun Chen, Marnix Witte, Wilma Heemsbergen, and Marcel V. Herk. „Multiple comparisons permutation test for image based data mining in radiotherapy“. In: *Radiation Oncology* 8 (1 2013) (cit. on p. 49).

[130]Arnd Sörensen, Ganna Blazhenets, Gerta Rücker, et al. „Prognosis of conversion of mild cognitive impairment to Alzheimer's dementia by voxel-wise Cox regression based on FDG PET data“. In: *NeuroImage: Clinical* 21 (2019), p. 101637 (cit. on p. 49).

[131]Jane Shortall, Giuseppe Palma, Hitesh Mistry, et al. „Flogging a Dead Salmon? Reduced Dose Posterior to Prostate Correlates With Increased PSA Progression in Voxel-Based Analysis of 3 Randomized Phase 3 Trials“. In: *International Journal of Radiation Oncology*Biology*Physics* 110 (3 July 2021), pp. 696–699 (cit. on p. 50).

[132]Lesley A. Smith, Victoria R. Cornelius, Christopher J. Plummer, et al. „Cardiotoxicity of anthracycline agents for the treatment of cancer: Systematic review and meta-analysis of randomised controlled trials“. In: *BMC Cancer* 10 (2010) (cit. on p. 52).

[133]Hellen Gelband, Prabhat Jha, Rengaswamy Sankaranarayanan, and Susan Horton, eds. *Cancer: Disease Control Priorities*. 2015th ed. Vol. 3rd. World Bank Group (cit. on p. 52).

[134]Agon Olloni, Carsten Brink, Ebbe L. Lorenzen, et al. „Does coronary artery calcium score have an impact on overall survival for locally advanced non-small cell lung cancer treated with definitive radiotherapy“. In: *Radiotherapy and Oncology* 185 (Aug. 2023), p. 109719 (cit. on p. 53).

[135]Mary E. Charlson, Peter Pompei, Kathy L. Ales, and C.Ronald MacKenzie. „A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation“. In: *Journal of Chronic Diseases* 40 (5 Jan. 1987), pp. 373–383 (cit. on p. 53).

[136]Charlotte Rotbøl Bøje, Susanne O. Dalton, Therese K. Grønborg, et al. „The impact of comorbidity on outcome in 12 623 Danish Head and Neck Cancer Patients: A population based study from the DAHANCA database“. In: *Acta Oncologica* 52 (2 Feb. 2013), pp. 285–293 (cit. on p. 53).

[137]*Encyclopedia Britannica*. Accessed: 2024-07-31. 2024. URL: `\url{https://www.britannica.com/}` (cit. on p. 57).

Title

Time trends in cardiac doses in 10,000 patients receiving curative thoracic radiation therapy between 2009 and 2020

Author names and affiliations

Nora Forbes[1,2], Cynthia Terrones-Campos[3], Abraham Smith[1,2], Joanne Reekie[3], Sune Darkner[2], Maja Maraldo[1], Mette Pøhl[1], Signe Risumlund[1], Lena Specht[1,4], Soren M. Bentzen[5], Jens Petersen*[1,2], Ivan R. Vogelius*[1,4]

[1] *Dept of Clinical Oncology, Center for Cancer and Organ Diseases, Copenhagen University Hospital – Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark*

[2] *University of Copenhagen, Department of Computer Science, Universitetsparken 1, 2100 Copenhagen, Denmark*

[3] *Center of Excellence for Health, Immunity, and Infections, Copenhagen University Hospital – Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark*

[4] *Dept. of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark*

[5] *Department of Epidemiology and Public Health, University of Maryland School of Medicine, 655 W. Baltimore Street, Baltimore, MD 21201, USA*

Highlights

- 10,215 radiation therapy plans extracted and analyzed
- Hearts were re-delineated with machine learning on the therapy CT-scan to ensure time-independent delineations
- Cardiac doses were found to decrease over time, especially high dose exposure of intrathoracic lesions
- An unexpected increase in cardiac high dose volume (V40Gy) for breast cancer treatment plans in recent years was observed

Abstract

Background and purpose

In the last 20 years, it has become well-documented that incidental cardiac exposure to ionizing radiation is associated with a clinically relevant increased risk of cardiovascular morbidity. In parallel, radiation therapy technologies have been developed to provide target dose coverage with less exposure to adjacent organs at risk. In the current work, we investigate trends in cardiac exposure among patients treated with curative intent radiotherapy from a single institution between 2009 to 2020.

Materials and methods

10,215 treatment courses were analyzed from 9,966 patients treated with curative intent for intrathoracic or breast cancers in the period 2009-2020. All hearts were re-delineated using an AI model to ensure consistency over time. Cardiac doses were extracted in 3D from the record-and-verify system and converted, voxel-by-voxel, to equi-effective doses in 2 Gy fractions (EQD2) using $\alpha/\beta=2$ Gy. Mean heart dose (in EQD2) and volume exposed to 5 and 40 Gy (V5 and V40Gy), respectively, were extracted. Time trends in these cardiac dose-volume metrics were investigated for each diagnosis.

Results

Patients with esophageal cancer had the highest mean heart dose (median = 11.67 Gy; IQR = 2.85, 18.18), while the lowest was observed in patients with breast cancer (median = 0.60 Gy; IQR = 0.30, 1.08) and lymphoma (median = 0.01 Gy; IQR = 0.00, 0.38). A decreasing trend over time was seen most clearly for patients with esophageal and lung cancers ($p<0.05$). Among patients with breast cancer, V40 decreased from 2009-2015 after which we observed an increase.

Conclusion

There has been a significant reduction in radiation exposure to the heart in patients treated in the period 2009-2020, likely due to increased awareness of cardiovascular toxicity and technological developments. The study also found a significant increase in V40 from 2015 to 2020 for patients with breast cancers, possibly related to increased prioritization of target coverage.

Introduction

The heart used to be considered a relatively radioresistant muscular structure, but publications in the modern era have demonstrated a clear association between cardiac exposure and risk of subsequent cardiovascular morbidity (Cutter et al., 2015; Darby et al., 2005, 2013; Van Nimwegen et al., 2016; van Velzen et al., 2022).

As a consequence, a number of technical developments have been introduced with the aim of minimizing the exposure of the heart to radiation. Examples of techniques with a documented dosimetric benefit during conventional exposure include proton therapy (Hassan et al., 2023) and deep inspiration breath hold (Nissen & Appelt, 2013; Pedersen et al., 2004; Petersen et al., 2015), whereas Intensity modulated radiotherapy (IMRT) has been shown to decrease cardiac exposure, but possibly at the expense of an increase in low dose bath (Chun et al., 2017).

Even for patients with a high risk of cancer recurrence, there have been reports of survival detriments with higher cardiac exposures (Atkins et al., 2021; Brink et al., 2022). Conversely, the Danish Breast Cancer Collaborative group showed that omitting an elective target to decrease cardiac dose may be unacceptable in terms of treatment efficacy (Thorsen et al., 2016a).

It remains to be defined to what extent increased awareness and technological advances allowing for more conformal treatment have impacted cardiac exposure in clinical practice. This study investigates trends in the cardiac dose received by patients treated with curative intent radiotherapy at a tertiary hospital in Denmark over the last decade (2009 to 2020). This study utilizes a previously presented method and pipeline for analyzing large cohorts of patients with detailed individual 3D dosimetry data and consistently delineate hearts (Smith et al., 2022). The study aim is to elucidate changes in dose delivered to the heart over time, to display the cumulative impact of gradual improvements in radiotherapy techniques and concerted efforts to reduce heart irradiation.

Methods

We included all patients receiving curative intent radiotherapy at Copenhagen University Hospital - Rigshospitalet, Denmark between 2009 and 2020 with breast cancer, esophageal cancer, lymphoma, non-small cell lung cancer (NSCLC), or small cell lung cancer (SCLC), cf. Figure 1. Patients were identified in the ARIA record and verify system by local radiation therapy codes for treatments with curative intent. We extracted CT scans and 3D dose matrices using the Digital Imaging and Communications in Medicine (DICOM) protocol for automated access. Dose matrices obtained from the system represented dose calculations based on Anisotropic Analytical Algorithm or Acuros as employed at the time of treatment. Finnegan et al. showed that the guidelines and practice of contouring the heart have changed between 2009 and 2014, and this in turn will affect the dose-volume metrics for heart (Finnegan et al., 2020). Therefore, we used an in-house and open-source AI software, RootPainter3D, which facilitates the training of a 3D U-net model via corrective-annotation (Smith et al., 2022), to redelineate all hearts. Our cardiac delineations model was trained on 933 manually corrected CT scans from a prior immune toxicity study (Terrones-Campos et al., 2023). This model was then used to delineate all 14,044 CT scans within the entire population, yielding cardiac delineation on each CT scan. For quality assurance, all delineations were reviewed for gross errors in 2D with an in-house review tool (Figure A1). The average time to review each delineation was less than 5 seconds, which equated to approximately 16 hours of work. Questionable cases were extracted in 3D and assessed with a physician to determine inclusion.

Fraction delivery times and dates were extracted from the record and verify system to aid dose assessment. When only one plan existed, the total dose was calculated as the sum over the fractions delivered. If two or more plans existed based on the same CT scan, the total dose was calculated as the sum over the total delivered fractions for each plan and then added together. If multiple plans existed based on different CT scans, it was not possible to directly sum fractions. In such cases, the plan with the most delivered fractions was used and we scaled up the dose to correspond to the number of fractions delivered in total across all CT scan/treatment plans. Cases with multiple plans, changes in treatment

schedule, or otherwise not possible to reliably process as above were excluded. Similarly, we excluded electron plans and a small number of cases treated on an MR accelerator or Ethos as these systems could not be accessed automatically.

Fractionation correction was performed using Withers formula for equi-effective dose in 2 Gy fractions (EQD2) voxel-wise on the 3D dose matrix:

$$EQD_2 = D\left(\frac{d + \frac{\alpha}{\beta}}{2 + \frac{\alpha}{\beta}}\right)$$

where D was the total dose at that voxel, d was the fraction dose at that voxel (D/fractions), and α/β was set to 2 Gy.

The mean heart dose (MHD) was calculated in EQD2. The dose volume histogram (DVH) values for each 1-Gy bin were extracted using EQD2 and calculated in cubic centimeters. We focused on one low dose-volume metric (V5 – volume receiving at least 5 Gy) and one high dose-volume metric (V40 - volume receiving at least 40 Gy).

Outliers in dose (several metrics) for all diagnoses were reviewed manually and compared to the record and verify system doses together with a medical physicist and/or oncologist to ensure that the automated dose extraction was accurate. This led to some necessary exclusion criteria to avoid erroneous cases (for example changes from curative to palliative intent).

The density of MHD was plotted for each of the patient groups. Scatter plots and trend plots were generated for each metric. Spearman's Rho correlation was calculated for each metric and for each group.

R (R Core Team, 2023) was used within the RStudio environment (Posit team, 2023) to conduct analysis and generate figures along with the ggplot2 (Wickham, 2016), tidyverse (Wickham et al., 2019), and gridextra (Auguie, 2015) packages.

Results

14,826 treatment courses (denoted "treatments" from here) on 13,855 patients in ARIA were eligible for this study, as shown in the consort flow diagram (Figure 1). Of these, 10,215 treatments on 9,966 patients could be included in the final analysis. Details of reasons for exclusion are given in the appendix, but in short, a combination of multiple scans/plans, "empty" courses and miscodings were common reasons.

10,215 treatments were included in the final analysis (Figure 1). Of these, the majority of treatments were patients with breast cancer (70%) and the smallest group was patients with SCLC (2%) (Table 1). The most common fractionation schemes are reported in Table 2.

We observed a decrease over time in the annual number of patients receiving radiotherapy at Rigshospitalet for the indications considered here. This change reflects the transition of a satellite center to become a separate clinic causing a change in uptake area.

The dosimetric results are shown in Figures 2-4 with further details in supplementary figures. Starting with breast cancer, 92% of patients had a MHD below 2 Gy (Figure 2). The median MHD, V5, and V40 was second lowest across all analysis groups, while the means were the lowest (Table 1). The median MHD decreased significantly across the timeframe ($p < 0.001$) (Figure 3, Table 1). Neither the median V5 nor V40 decreased significantly ($p = 0.37$ and $p = 0.065$, respectively) (Figure 4, Table 1). High dose outlier cases decreased for both V5 and V40 (Figure A2). However, in the second half of the time period, the median MHD, V5, and V40 of breast cancer patients increased significantly ($p < 0.001$).

Patients with lymphoma did not have a significant decrease in heart exposure over time (Table 2, Figure 3). Rather, there as a weak trend toward an increasing dose, which was driven by a decrease in cases with no heart exposure. Specifically, in the first 3 years 57% of MHD were zero, while in the last 3 years this fell to 38%. Mean MHD was 1.5 Gy, with 82% of the patients exposed to less than 2 Gy and 95% of the patients exposed <10 Gy.

Patients with Esophageal cancer had the highest MHD, V5, and V40. Only 22% of esophageal patients received less than 2Gy MHD. This is where we see the largest improvement in cardiac doses over time, especially V40 is reduced substantially. Specifically, median MHD decreased from 18.1 Gy in the first 4 years to 12.0 Gy in the last for years for the esophageal group. For NSCLC, median MHD decreased from 7.0 Gy in the first 4 years to 3.5 Gy in the last 4 years. The corresponding decrease for SCLC was from 6.9 Gy to 3.0 Gy. The dose reduction in MHD is driven by the decrease in the high dose exposure. We also observe that the occurrence of high dose outliers decreasing over time (Figure 3, Figure A2). Only patients with SCLC and NSCLC had a significant decrease in V5. Only patients with Esophageal and NSCLC had a significant decrease in V40.

Discussion

Using an AI method, we were able to assess the MHD, and cardiac V5 and V40 of 10,000 patients treated with curative intent radiotherapy from 2009 to 2020 in key diagnostic groups. This real-world data analysis provides the end-result of cardiac exposures because of all combined changes of treatment in the study period, with a time-independent cardiac delineation method.

Starting with breast cancer the key findings were that cardiac exposures among patients with breast cancer in the Danish cohort were much lower compared to historical data, including the patients studied by Darby et al. (Darby et al., 2013). According to Darby et al., a MHD increase of 2 Gy is associated with a 14.8% excess relative risk of major coronary events. For a patient exposed to 2 Gy at age 50, this excess relative risk converts to an excess absolute risk of death from ischemic heart disease of just 0.1 % at attained age of 70 years (Calculated from cumulative incidence data derived from Darby et al supplemental table 13). Given that 92% of patients have lower exposure than 2 Gy, our data supports that cardiac risk pertaining to modern radiotherapy for most breast cancer patients is extremely low.

The Danish Breast Cancer Collaborative Group, DBCG, published the results of a population-based cohort study on the effect of internal mammary node irradiation in early node-positive breast cancer in 2016 with the main conclusion that omission of irradiation to spare the heart may negatively affect survival and disease control (Thorsen et al., 2016b). The DBCG data was also included in a recent seminal work of the early breast cancer trialists' collaborative group (EBCTCG) individual patient data meta-analysis of 14,324 participants in clinical trials. The EBCTCG data shows an excess non-cancer related mortality in older trials when adding regional node irradiation, but this effect was absent in trials accruing from 1989-2013. In the more recent trials, a benefit was seen in reduction of breast cancer mortality without an adverse effect on non-breast cancer mortality. More specifically, the EBCTCG analysis report an absolute risk reduction in breast cancer mortality of 3% (95% CI: 1.1-4.9%) at 15 years and a risk reduction of 3% (95% CI: 1.0-5.0 %) for overall mortality when adding regional node radiotherapy. These numbers compared to the estimated 0.1% excess absolute risk of cardiac mortality clearly support for the increased prioritization of coverage of the target for breast cancer patients introduced in national and departmental guidelines after the IMN study in 2016. This clinical prioritization may explain the slight increase in cardiac exposure observed in recent years.

Nevertheless, continued monitoring, as presented here, can still point to continued optimization potential. Future work includes omission of radiotherapy for select patients (e.g., NCT03646955) or advanced techniques, such as proton therapy (e.g., ISRCTN14220944). It should be noted that the continued potential to gradually reduce high doses in outlier cases in recent years may imply that patient volume estimates for proton therapy trials using historical data, as in the DBCG proton trial, may overestimate the frequency of proton therapy indications for future patients (Stick et al., 2021).

The lymphoma patients are a quite different cohort, where the disease localization is highly heterogeneous. Lymphoma cases with high dose likely have a mediastinal target at the level of the heart, and this will effectively drive the incidental dose to the heart. DIBH was available as a treatment option from the early years for this treatment group (Rechner et al., 2017) which reduces cardiac exposure, but it appears that further improvement is difficult with current available methods. However, increased conformity to target should still be sought, including adaptation of the plan to daily patient anatomy and proton therapy as possible avenues for dose reductions to the heart.

We now turn to discuss lung and esophageal cancers. Among the intrathoracic solid cancers cardiac exposures vary considerably depending on target position. The observed reduction in MHD is driven by the reduction in high dose regions which is consistent with the increased high dose conformity and increased low dose bath associated with intensity modulated techniques (Bergom et al., 2021). Additionally, the reduction is accompanied by reduced number and severity of high dose outlier cases. IMRT and volumetric arc techniques are used routinely for these patients in the later years. Deep inspiration breath-hold techniques have been evaluated for these patient groups but are not yet uniformly offered and do depend on patient compliance (Josipovic et al., 2019). Thus, there appears to be potential for further improvement in these patient groups. Central versus non-central location and mediastinal lymph node involvement will be important factors for individual exposure. Knowledge based planning is one way to individualize the expected achievable dose plan and the present work could be used as an avenue to increase the training data from the current standard of ~100 patients (Li et al., 2017).

The decrease in patient volume at Rigshospitalet is due to a change in treatment center capacity in our historical uptake area. While there may be some differences in the patient population (e.g., socioeconomic status and lifestyle) it is not expected that this would have an association with the severity of cancer nor prescription of dose. Therefore, we do not anticipate this change in demographics to yield relevant bias to our conclusions.

Machine learning models for structure delineation combined with modern archival systems means that large-scale dosimetric analyses have been made possible as also demonstrated previously in breast cancer trial datasets (Finnegan et al., 2020). In the current study we used such methods to documents trends in incidental cardiac irradiation in a variety of diagnoses and over a period of 12 years in a single institution and in a very large patient cohort. Unfortunately, it is not possible to draw a specific causal relationship between individual changes in guidelines and the decrease in cardiac dose, however this serves as an overarching description of the combined impact of all these changes. Data points to improved cardiac sparing over time. A formal health economics perspective is outside of the scope of this paper, but we still find it relevant to emphasize that the employed techniques in later years (IMRT, dose planning expertise, awareness, and breath hold) come at a very modest – if any – excess resource expenditure per patient once the investment in equipment has been made.

The scalability of ML algorithms, of which the present study is one example of several, is a potential game changer in several areas (Vogelius et al., 2020). Here we demonstrated the ability to closely monitor doses to organs at risk across disease sites and over time in an unselected patient cohort. We were able to monitor detailed dose-volume data, including low dose bath and high dose exposure to the heart. The methods could be combined with machine learning techniques to account for the 3D heart exposure, one could generate much larger “knowledge-based planning” cohorts or the data could be coupled to registry or claims-based outcome data to elucidate new knowledge of rare, high-grade toxicities.

In conclusion, the analysis of a large consecutive dataset of patients documented continuous improvements in cardiac sparing over time consistent with continued technical improvements as well as clinical awareness.

Acknowledgments

We report institutional research and teaching contracts with Brainlab, Varian, and ViewRay. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 801199. The authors acknowledge support from Kirsten and Freddy Johansen's award and National Cancer Institute (Grant No. P30 CA 134274-04). SMB was supported by funds through the Maryland Department of Health's Cigarette Restitution Fund Program. IRV and JP received research funding from Varian Medical Systems. IRV and JP are supported by Danish Cancer Society (Grant No. R231-A13976). This work is supported by Danish National Research Foundation (DNRF) grant number 126 (PERSIMUNE).

References

Atkins, K. M., Chaunzwa, T. L., Lamba, N., Bitterman, D. S., Rawal, B., Bredfeldt, J., Williams, C. L., Kozono, D. E., Baldini, E. H., Nohria, A., Hoffmann, U., Aerts, H. J. W. L., & Mak, R. H. (2021). Association of Left Anterior Descending Coronary Artery Radiation Dose With Major Adverse Cardiac Events and Mortality in Patients With Non–Small Cell Lung Cancer. *JAMA Oncology*, *7*(2), 206. https://doi.org/10.1001/jamaoncol.2020.6332

Auguie, B. (2015). *gridExtra: Miscellaneous Functions for "Grid" Graphics*. http://CRAN.R-project.org/package=gridExtra

Bergom, C., Bradley, J. A., Ng, A. K., Samson, P., Robinson, C., Lopez-Mattei, J., & Mitchell, J. D. (2021). Past, Present, and Future of Radiation-Induced Cardiotoxicity: Refinements in Targeting, Surveillance, and Risk Stratification. *JACC: CardioOncology*, *3*(3), 343–359. https://doi.org/10.1016/j.jaccao.2021.06.007

Brink, C., Bernchou, U., Bertelsen, A., Hansen, O., Schytte, T., Hjelmborg, J. v. B., Holloway, L., van Herk, M., Johnson-Hart, C., Price, G. J., Aznar, M. C., McWilliam, A., Faivre-Finn, C., & Hansen, C. R. (2022). Causal relation between heart irradiation and survival of lung cancer patients after radiotherapy. *Radiotherapy and Oncology*, *172*, 126–133. https://doi.org/10.1016/j.radonc.2022.05.002

Chun, S. G., Hu, C., Choy, H., Komaki, R. U., Timmerman, R. D., Schild, S. E., Bogart, J. A., Dobelbower, M. C., Bosch, W., Galvin, J. M., Kavadi, V. S., Narayan, S., Iyengar, P., Robinson, C. G., Wynn, R. B., Raben, A., Augspurger, M. E., MacRae, R. M., Paulus, R., & Bradley, J. D. (2017). Impact of intensity-modulated radiation therapy technique for locally advanced non-small-cell lung cancer: A secondary analysis of the NRG oncology RTOG 0617 randomized clinical trial. *Journal of Clinical Oncology*, *35*(1), 56–62. https://doi.org/10.1200/JCO.2016.69.1378

Cutter, D. J., Schaapveld, M., Darby, S. C., Hauptmann, M., Van Nimwegen, F. A., Krol, A. D. G., Janus, C. P. M., Van Leeuwen, F. E., & Aleman, B. M. P. (2015). Risk for valvular heart disease after treatment for hodgkin lymphoma. *Journal of the National Cancer Institute*, *107*(4). https://doi.org/10.1093/jnci/djv008

Darby, S. C., Ewertz, M., McGale, P., Bennet, A. M., Blom-Goldman, U., Brønnum, D., Correa, C., Cutter, D., Gagliardi, G., Gigante, B., Jensen, M.-B., Nisbet, A., Peto, R., Rahimi, K., Taylor, C., & Hall, P. (2013). Risk of Ischemic Heart Disease in Women after Radiotherapy for Breast Cancer. *New England Journal of Medicine*, *368*(11), 987–998. https://doi.org/10.1056/nejmoa1209825

Darby, S. C., McGale, P., Taylor, C. W., & Peto, R. (2005). Long-term mortality from heart disease and lung cancer after radiotherapy for early breast cancer: prospective cohort study of about 300 000 women in US SEER cancer registries. *The Lancet Oncology*, *6*(8), 557–565. https://doi.org/10.1016/S1470-2045(05)70251-5

Finnegan, R., Lorenzen, E. L., Dowling, J., Jensen, I., Berg, M., Thomsen, M. S., Delaney, G. P., Koh, E.-S., Thwaites, D., Brink, C., Offersen, B. V., & Holloway, L. (2020). Analysis of cardiac substructure dose in a large, multi-centre danish breast cancer cohort (the DBCG HYPO trial): Trends and predictive modelling. *Radiotherapy and Oncology*, *153*, 130–138. https://doi.org/10.1016/j.radonc.2020.09.004

Josipovic, M., Aznar, M. C., Thomsen, J. B., Scherman, J., Damkjaer, S. M., Nygård, L., Specht, L., Pøhl, M., & Persson, G. F. (2019). Deep inspiration breath hold in locally advanced lung cancer radiotherapy: validation of intrafractional geometric uncertainties in the INHALE trial. *The British Journal of Radiology*, *92*(1104), 20190569. https://doi.org/10.1259/bjr.20190569

Li, N., Carmona, R., Sirak, I., Kasaova, L., Followill, D., Michalski, J., Bosch, W., Straube, W., Mell, L. K., & Moore, K. L. (2017). Highly Efficient Training, Refinement, and Validation of a Knowledge-based Planning Quality-Control System for Radiation Therapy Clinical Trials. *International Journal of Radiation Oncology*Biology*Physics*, *97*(1), 164–172. https://doi.org/10.1016/j.ijrobp.2016.10.005

Posit team. (2023). *RStudio: Integrated Development Environment for R*. http://www.posit.co/

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. https://www.R-project.org/

Rechner, L. A., Maraldo, M. V., Vogelius, I. R., Zhu, X. R., Dabaja, B. S., Brodin, N. P., Petersen, P. M., Specht, L., & Aznar, M. C. (2017). Life years lost attributable to late effects after radiotherapy for early stage Hodgkin lymphoma: The impact of proton therapy and/or deep inspiration breath hold. *Radiotherapy and Oncology*, *125*(1), 41–47. https://doi.org/10.1016/j.radonc.2017.07.033

Smith, A. G., Petersen, J., Terrones-Campos, C., Berthelsen, A. K., Forbes, N. J., Darkner, S., Specht, L., & Vogelius, I. R. (2022). RootPainter3D: Interactive-machine-learning enables rapid and accurate contouring for radiotherapy. *Medical Physics*, *49*(1), 461–473. https://doi.org/10.1002/mp.15353

Stick, L. B., Lorenzen, E. L., Yates, E. S., Anandadas, C., Andersen, K., Aristei, C., Byrne, O., Hol, S., Jensen, I., Kirby, A. M., Kirova, Y. M., Marrazzo, L., Matías-Pérez, A., Nielsen, M. M. B., Nissen, H. D., Oliveros, S., Verhoeven, K., Vikström, J., & Offersen, B. V. (2021). Selection criteria for early breast cancer patients in the DBCG proton trial – The randomised phase III trial strategy. *Clinical and Translational Radiation Oncology*, *27*, 126–131. https://doi.org/10.1016/j.ctro.2021.01.012

Terrones-Campos, C., Ledergerber, B., Forbes, N., Smith, A. G., Petersen, J., Helleberg, M., Lundgren, J., Specht, L., & Vogelius, I. R. (2023). Prediction of Radiation-induced Lymphopenia following Exposure of the Thoracic Region and Associated Risk of Infections and Mortality. *Clinical Oncology*, *35*(7), e434–e444. https://doi.org/10.1016/j.clon.2023.04.003

Thorsen, L. B. J., Offersen, B. V., Danø, H., Berg, M., Jensen, I., Pedersen, A. N., Zimmermann, S. J., Brodersen, H. J., Overgaard, M., & Overgaard, J. (2016a). DBCG-IMN: A population-based cohort study on the effect of internal mammary node irradiation in early node-positive breast cancer. *Journal of Clinical Oncology*, *34*(4), 314–320. https://doi.org/10.1200/JCO.2015.63.6456

Thorsen, L. B. J., Offersen, B. V., Danø, H., Berg, M., Jensen, I., Pedersen, A. N., Zimmermann, S. J., Brodersen, H.-J., Overgaard, M., & Overgaard, J. (2016b). DBCG-IMN: A Population-Based Cohort Study on the Effect of Internal Mammary Node Irradiation

in Early Node-Positive Breast Cancer. *Journal of Clinical Oncology*, *34*(4), 314–320. https://doi.org/10.1200/JCO.2015.63.6456

Van Nimwegen, F. A., Schaapveld, M., Cutter, D. J., Janus, C. P. M., Krol, A. D. G., Hauptmann, M., Kooijman, K., Roesink, J., Van Der Maazen, R., Darby, S. C., Aleman, B. M. P., & Van Leeuwen, F. E. (2016). Radiation dose-response relationship for risk of coronary heart disease in survivors of Hodgkin lymphoma. *Journal of Clinical Oncology*, *34*(3), 235–243. https://doi.org/10.1200/JCO.2015.63.4444

van Velzen, S. G. M., Gal, R., Teske, A. J., van der Leij, F., van den Bongard, D. H. J. G., Viergever, M. A., Verkooijen, H. M., & Išgum, I. (2022). AI-Based Radiation Dose Quantification for Estimation of Heart Disease Risk in Breast Cancer Survivors After Radiation Therapy. *International Journal of Radiation Oncology*Biology*Physics*, *112*(3), 621–632. https://doi.org/10.1016/j.ijrobp.2021.09.008

Vogelius, I. R., Petersen, J., & Bentzen, S. M. (2020). Harnessing data science to advance radiation oncology. In *Molecular Oncology* (Vol. 14, Issue 7, pp. 1514–1528). John Wiley and Sons Ltd. https://doi.org/10.1002/1878-0261.12685

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., Francois, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Muller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Tables and Figures

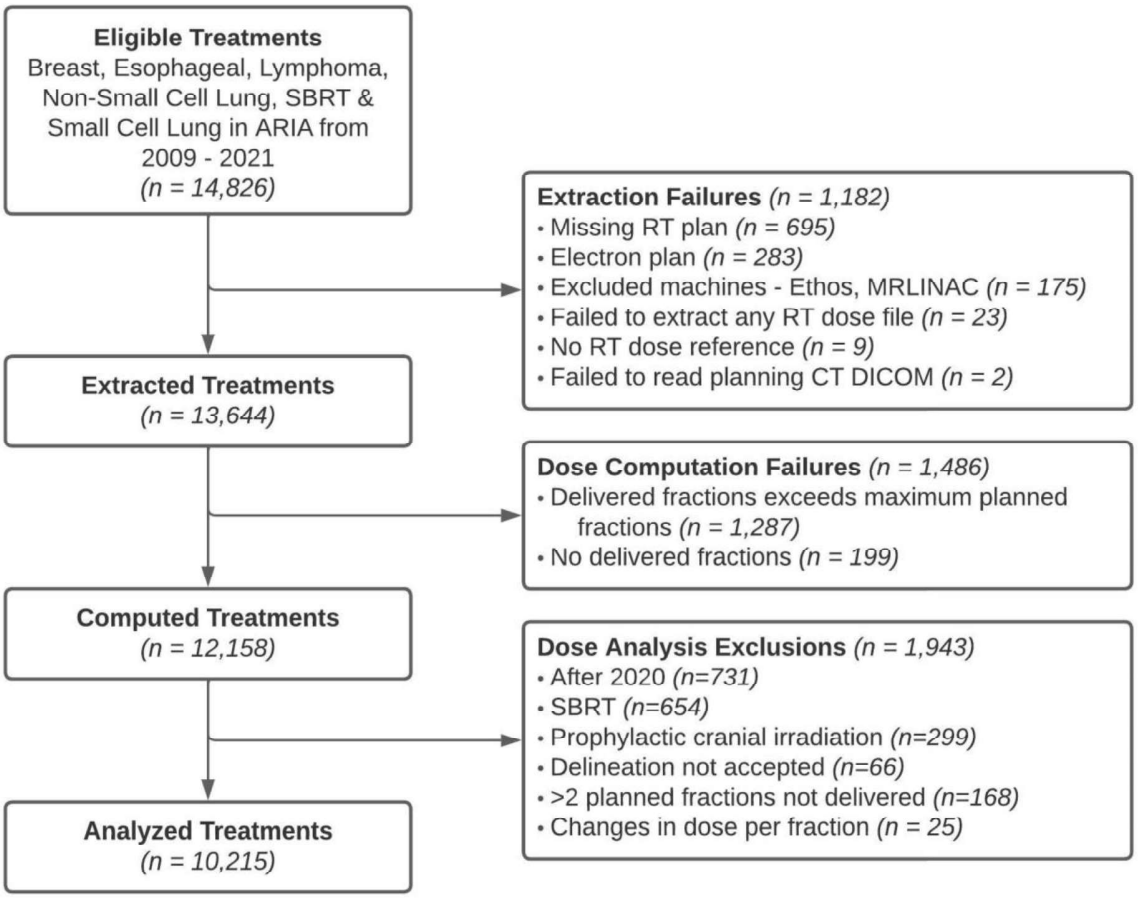


Figure 1. Consort diagram.

Table 1. Demographics and dose statistics, including mean heart dose (MHD), heart volume receiving at least 5 Gy (V5), and heart volume receiving at least 40 Gy (V40).

	Breast	**Esophageal**	**Lymphoma**	**Non-Small Cell Lung**	**Small Cell Lung**
Treatments (N)	7113	663	1624	623	192
Male, n (%)	27 (0.38)	461 (69.5)	859 (52.9)	339 (54.4)	89 (46.4)
Age, median (IQR)	62.5 (54.4, 68.8)	66.3 (60.7, 71.7)	63.7 (50.0, 72.4)	67.0 (60.3, 71.9)	65.7 (60.2, 71.0)
MHD in Gy, median (IQR) rho (P)	0.60 (0.30, 1.08) -0.071 (<0.001)*	11.67 (2.85, 18.18) -0.29 (<0.001)*	0.010 (0.00, 0.38) 0.12 (<0.001)*	4.28 (1.26, 9.95) -0.28 (<0.001)*	4.43 (1.65, 8.13) -0.23 (0.0011)*
V5 in cubic cm, median (IQR) rho (P)	2.245 (0.00, 13.92) 0.011 (0.37)	466.55 (93.46, 681.42) -0.081 (0.37)	0.00 (0.00, 0.00) 0.054 (0.029)*	132.8 (33.9, 314.0) -0.17 (<0.001)*	156.62 (48.39, 329.76) -0.17 (0.021)*
V40 in cubic cm, median (IQR) rho (P)	0.00 (0.00, 0.012) -0.022 (0.065)	41.74 (4.99, 84.72) -0.36 (<0.001)*	0.00 (0.00, 0.00) -0.004 (0.86)	11.86 (0.00, 50.02) -0.39 (<0.001)*	0.84 (0.00, 7.81) -0.015 (0.84)

Table 2. Fractionation schemes with at least 5% of each analysis group

Group	Fractionation Scheme	N	Proportion
Breast	15 X 2.7 Gy	3184	44.8%
	25 X 2 Gy	2742	38.5%
Esophageal	25 X 2 Gy	551	83.1%
Lymphoma	20 X 2 Gy	268	16.5%
	10 X 3 Gy	215	13.2%
	15 X 2 Gy	187	11.5%
	17 X 1.8 Gy	139	8.6%
	13 X 2 Gy	124	7.6%
	5 X 5 Gy	123	7.6%
	12 X 2 Gy	107	6.6%
Non-Small Cell Lung	33 X 2 Gy	455	73.0%
	32 X 2 Gy	41	6.6%
Small Cell Lung	30 X 1.5 Gy	157	81.8%

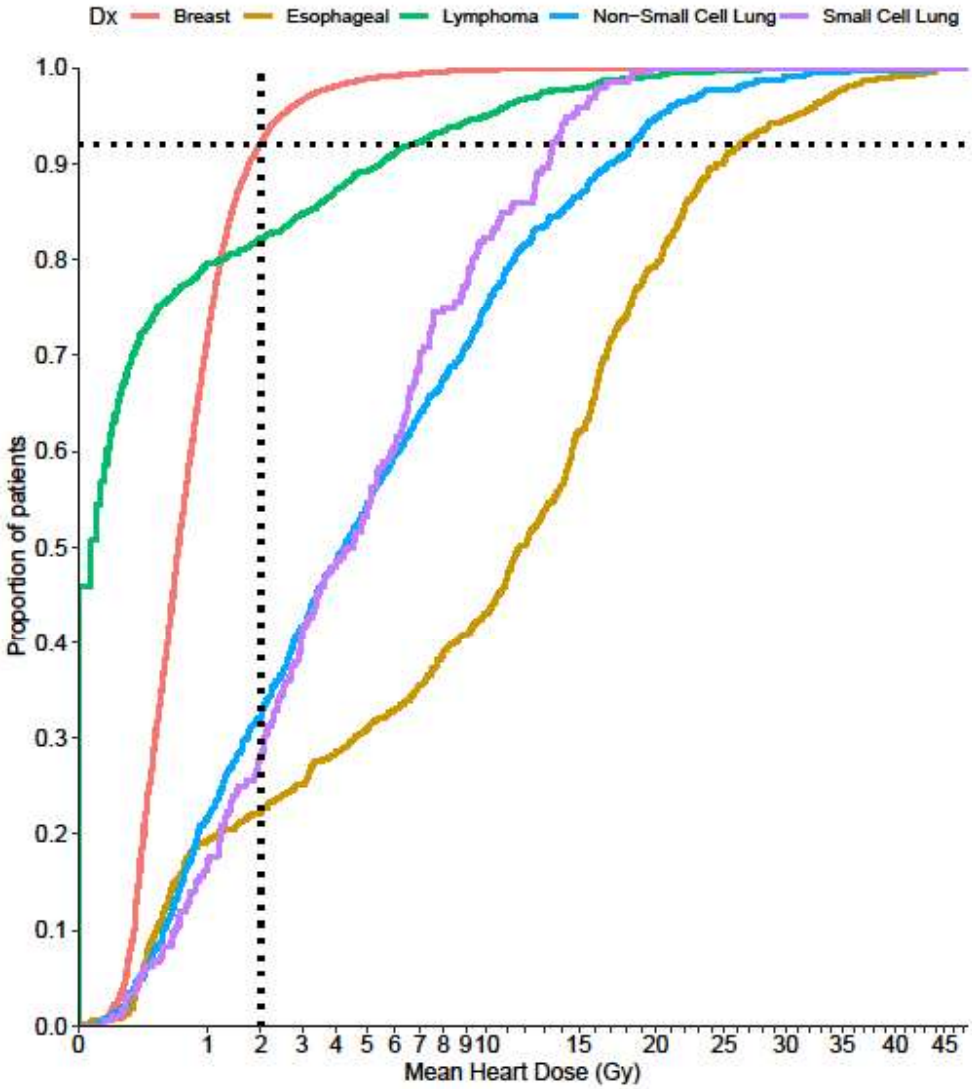


Figure 2. Cumulative density plot showing the proportion of patients by mean heart dose (MHD) in each analysis group. For example, approximately 92% of treatments in the breast cancer group received less than 2 Gy in MHD. The X-axis is scaled by a square root transform for clarity across the range of doses. Dose is reported in equivalent dose in 2 Gy fractions.

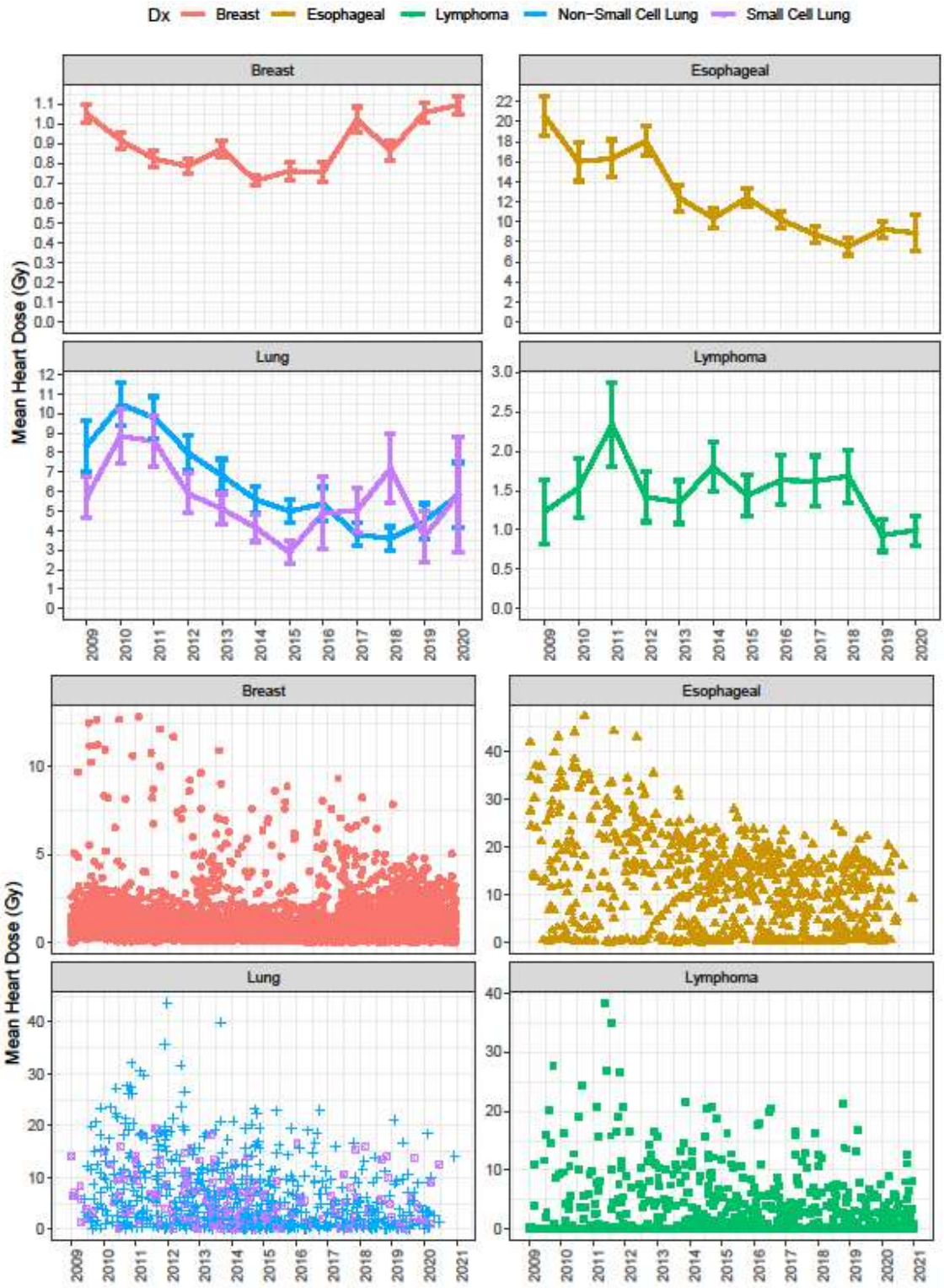


Figure 3. Annual trend plots (above) of median mean heart dose (MHD) for each analysis group, with lower (Q1) and upper (Q3) bounds. Scatter plots (below) of MHD for each analysis group. Note that each panel has an independent y-axis scale, and in particular the breast cancer scale is relatively small.

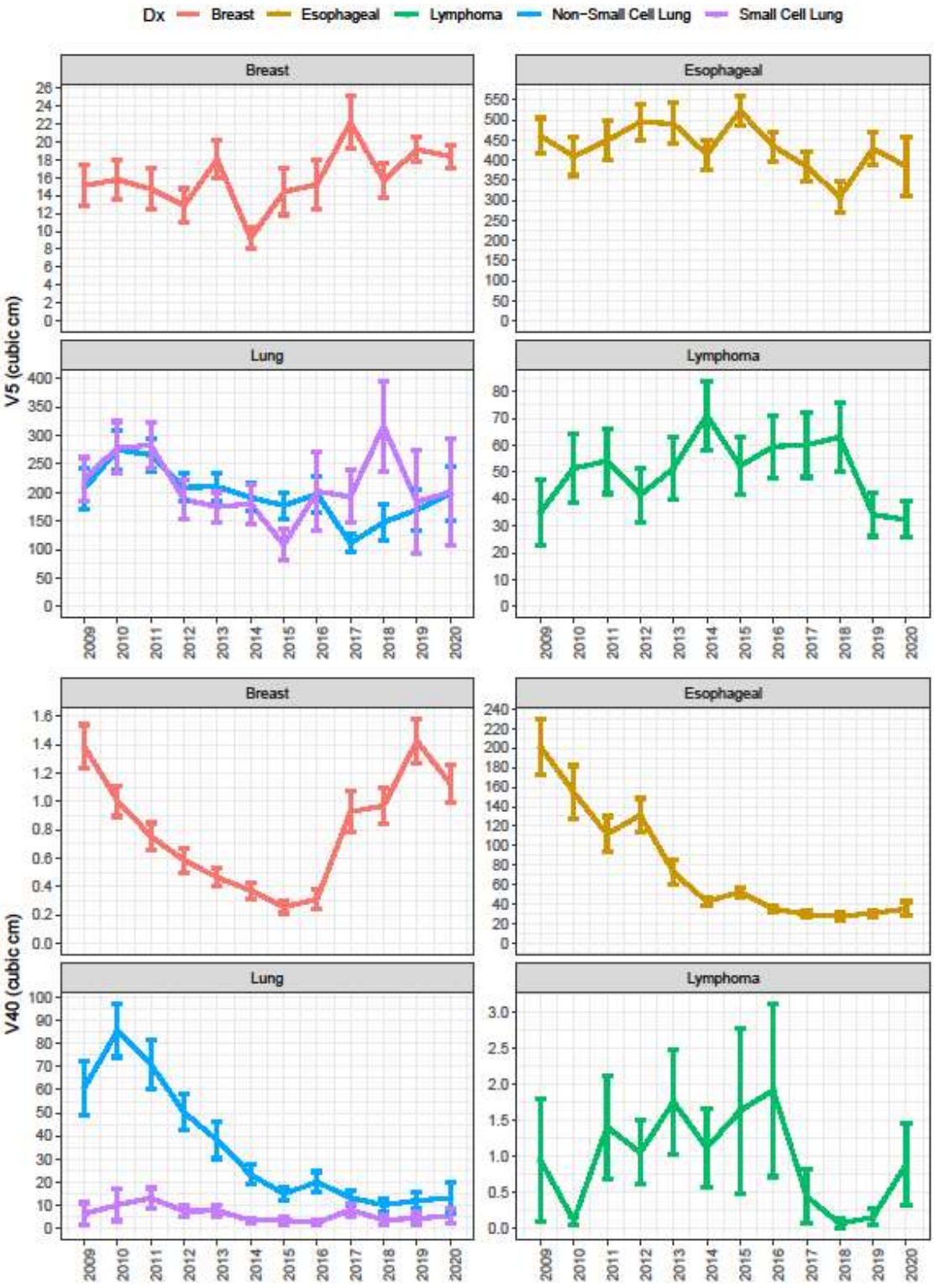


Figure 4. Annual trend plots of median heart volume receiving at least 5 Gy (V5) and heart volume receiving at least 40 Gy (V40) for each analysis group, with lower (Q1) and upper (Q3) bounds. Note that each panel has an independent y-axis scale, and in particular the breast cancer scale is relatively small.

Appendix

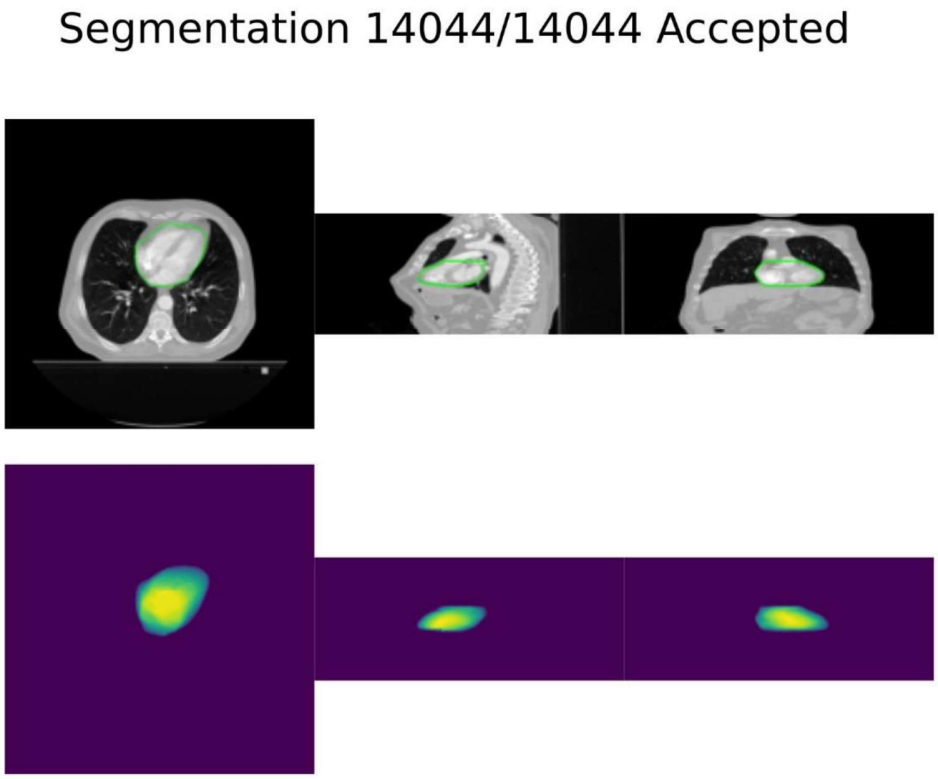


Figure A1. Example of software interface used for validation of the cardiac delineations. The top row displays the CT scan overlayed by the cardiac delineations in all three anatomical planes – axial, sagittal, and coronal (left to right) – taken at the center of mass of the delineation. The bottom row represents a rendering of the 3D cardiac delineation, where a more intense yellow color indicates a thicker section of heart. The user toggles through all patients and decides whether the delineation is acceptable, unacceptable, or undecided (which required further review with an oncologist).

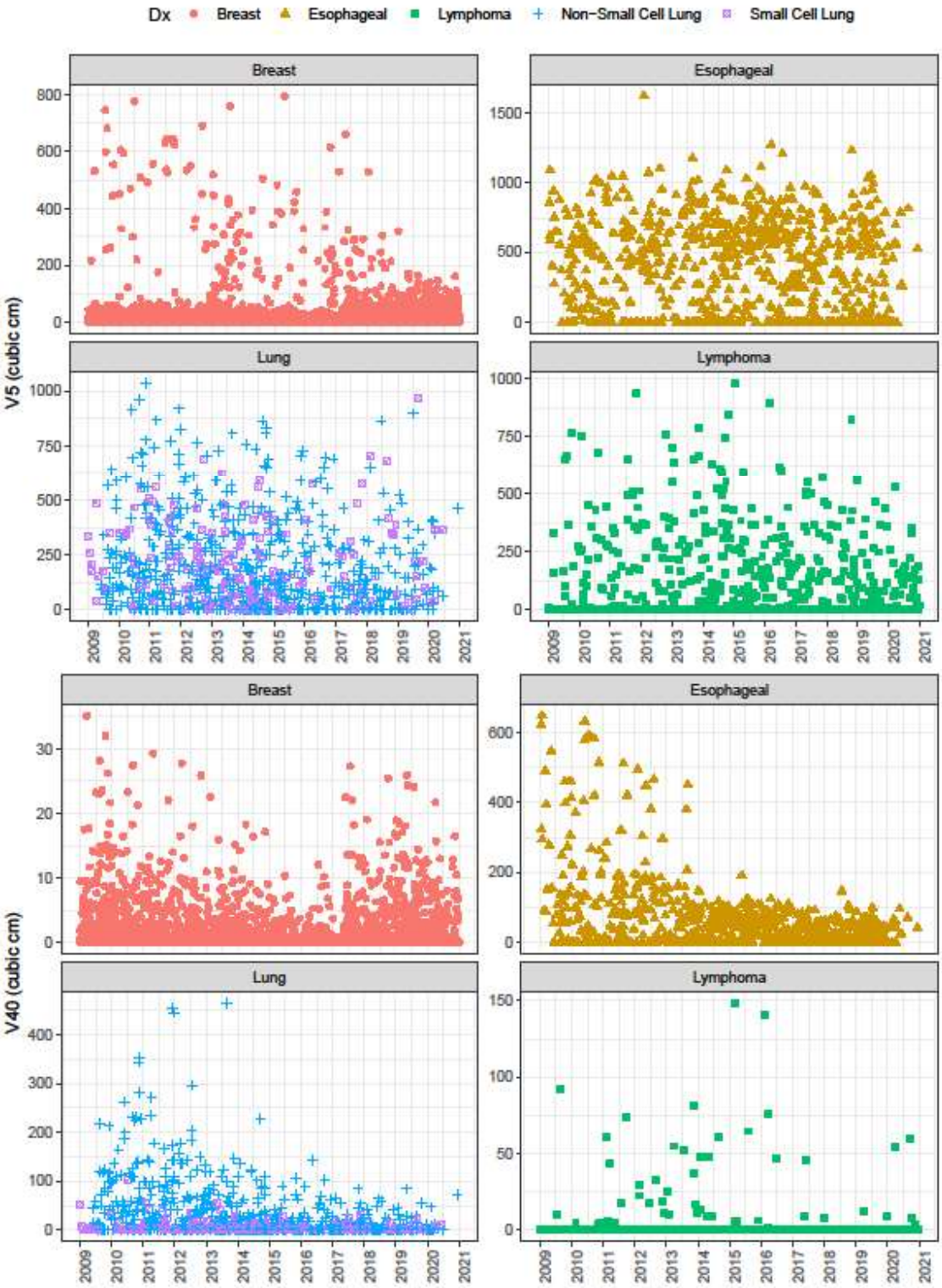


Figure A2. Scatter plots of heart volume receiving at least 5 Gy (V5) and heart volume receiving at least 40 Gy (V40) for each analysis group. Note that each panel has an independent y-axis scale, and in particular the breast cancer scale is relatively small.

Title

Cardiac dose-volume analysis of 9,411 patients with registry data for cardiovascular disease and overall survival.

Short Title

Cardiac dose-volume analysis for cardiovascular disease and overall survival

Author names and affiliations

Nora Forbes[1,2], Cynthia Terrones-Campos[3], Abraham George Smith[1,2], Joanne Reekie[3], Sune Darkner[2], Maja Maraldo[1,4], Mette Pøhl[1], Signe Risumlund[1], Lena Specht[1,4], Soren M. Bentzen[5,] Jens Petersen[1,2], Ivan R. Vogelius[1,4]

[1] *Dept of Clinical Oncology, Center for Cancer and Organ Diseases, Copenhagen University Hospital – Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark*

[2] *University of Copenhagen, Department of Computer Science, Universitetsparken 1, 2100 Copenhagen, Denmark*

[3] *Center of Excellence for Health, Immunity, and Infections, Copenhagen University Hospital – Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark*

[4] *Dept. of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark*

[5] *Department of Epidemiology and Public Health, University of Maryland School of Medicine, 655 W. Baltimore Street, Baltimore, MD 21201, USA*

Highlights

- Age and existing heart disease far outweighed heart dose as predictors of ischemic heart disease
- Overall survival is not a useful surrogate for cardiac toxicity in dose-response studies due to confounding by disease stage
- With modern RT techniques, the excess absolute risk attributable to radiotherapy is so small that a statistically significant dose-response could not be observed even in 9,411 patients
- For most patients, good quality contemporary radiotherapy is sufficient to limit heart toxicity as a clinically relevant concern

Abstract

Background and purpose

Radiation therapy (RT) to the thorax poses risks of radiation-induced cardiotoxicity, potentially increasing cardiovascular diseases (CVD) incidence. Advances in RT strive to minimize these risks by reducing heart radiation dose exposure.

This study integrates detailed 3D dosimetry on individually delineated hearts with registry-based outcome data to assess the impact of radiation dose on cardiovascular morbidity and overall survival (OS) across multiple cancer types. It also examined the influence of patient-specific factors on cardiotoxicity risk and survival outcomes.

Materials and methods

We analyzed data from 9,411 patients receiving RT at Rigshospitalet between 2009 and 2020 for breast, esophageal, lymphoma, and lung cancers. Cumulative incidence of CVD and death in the presence of competing risks was calculated with the Aalen-Johansen estimator. The impact of radiation dose and patient characteristics on ischemic heart disease (IHD) onset and OS were assessed using Kaplan-Meier and Cox Proportional-Hazards Models.

Results

Higher mean heart dose (MHD) was associated with poorer OS in breast and lung cancer patients (Hazard ratio 2.8 and 1.2), but no significant relationship was found between MHD and IHD. Established cardiac risk factors (age, sex, and existing IHD) outweighed cardiac dose as a risk factor for subsequent cardiac events for all diagnoses. The risk of death was greater than subsequent CVD, especially in esophageal and lung cancers (cumulative incidence 60% versus 17% and 60% versus 14%), despite comparatively high heart doses.

Conclusion

The study demonstrates that risk of death from primary cancer is of far greater concern than risk of subsequent cardiac events from cardiac radiation dose exposure in the range achievable with contemporary RT techniques, especially for lung and esophageal cancer patients. Further sparing of the heart should not be prioritized at the expense of adequate treatment of the index cancer.

Introduction

Radiation therapy (RT) is a critical treatment modality for cancer, offering effective control for many tumor types. Despite therapeutic benefits, RT can pose significant risks of adverse events from unavoidable irradiation of normal tissue surrounding the target. Radiation-induced cardiotoxicity is a concern among thoracically irradiated patients [1,2]. It can manifest as ischemic heart disease (IHD) among other cardiovascular diseases (CVD).

Radiation-induced cardiotoxic late-effects are an increasing concern for cancer patients [3–10]. The focus on cardiotoxicity mitigation is increasing with the establishment of guidelines including thresholds on delivered cardiac dose [11]. Darby et. al quantified the risk of CVD in breast cancer patients in 2013; recent work has extended findings into lymphoma and lung cancers [10,12]. Ongoing advances in radiation technology and techniques aim to reduce CVD risks by minimizing the mean heart dose (MHD) and other critical dosimetric parameters. These measures resulted in significant decreases in cardiac dose [13].

Many studies highlight the complex interplay between radiation dose, patient characteristics, and cancer stage on cardiac outcomes [14]. Particularly, the relationship between radiation-induced cardiotoxicity and overall survival (OS) has garnered attention, with evidence suggesting higher radiation doses to the heart correlate with poorer survival outcomes in certain cancers [15–18]. Therefore, there is a need for precise dosimetric assessments and comprehensive outcome data to understand the dose effect on overt cardiotoxicity.

The current study integrates detailed 3D dosimetry data on 9,411 individually delineated hearts with outcomes from national registry systems to analyze cardiotoxicity and mortality following standard cancer treatment with RT. We aimed to quantify and contextualize the impact of radiation dose and patient-specific factors on cardiotoxic side effects and OS. These findings can inform optimal directions for future refinement of treatment planning and delivery.

Methods

Data Curation

Eligible patients included those treated with curative intent at Rigshospitalet in 2009-2020. Patients with breast cancer, esophageal cancer, lymphoma, and non-small cell (NSCLC) or small cell (SCLC) lung cancer were included. We identified patients by departmental codes used in the record and verify system.

Associated computed tomography (CT) scans and 3D dose matrices were acquired through the record and verify system. Heart segmentation on all planning CT scans were performed using an open-source AI software [19]. All heart delineations were validated by manual review in 2D. Questionable cases were secondarily reviewed in 3D with a physician and unacceptable cases were omitted. Calculated dose metrics included mean heart dose (MHD), absolute volume receiving at least 5 Gy (V5), and absolute volume receiving at least 30 Gy (V30) - all converted to equivalent doses in 2 Gy (EQD2) fractions (α/β ratio = 2 Gy). Sex and birthdate were collected from patient records, as previously detailed in Forbes et al. 2024 [13].

Outcome data were extracted through electronic health records (EHR) maintained at the Centre of Excellence for Personalized Medicine of Infectious Complications in Immune Deficiency (PERSIMUNE) data warehouse. Diagnostic codes were sourced from The National Patient Register (Landspatientregisteret (LPR)) database. Death and emigration records were sourced from the Central Person Register (CPR).

Statistical Analysis

A statistical analysis plan (SAP) was established and published prior to analysis [20]. Modifications were made to the original SAP (Table A1).

The primary endpoint was onset of IHD following RT, defined by ICD-10 codes I20-I25 occurring after baseline, defined as the date of the last fraction of RT. Additionally, a broader outcome of CVD including valvular disease (VD; I00-I09, I34-I39) and heart failure (HF; I50) was assessed. OS was also analyzed.

All analyses were conducted separately on each cancer diagnosis due to heterogeneity among patient populations. The reverse Kaplan-Meier (KM) method was used to report follow-up in the absence of an event [21]. Absolute risk estimates of the competing events of the three CVD outcomes and death from other causes were analyzed using Aalen-Johansen estimates.

To predict risk of IHD, KM and Cox Proportional-Hazards Model (PHM) were used. Univariable KM curves were generated for each independent variable. Survival distributions were compared between factor levels with a log rank test (for variables with 2 levels) or test for trend analysis (for variables with >2 levels). Multivariable analysis of these variables was conducted with the Cox PHM. Patients were censored at the date of emigration, death, or the last potential follow-up (December 31st, 2020). These KM and Cox PHM were then replicated with OS as the outcome.

Independent variables of interest included sex, age, existing IHD, and MHD. Sex was omitted for the breast cancer cohort as the very few males were excluded. Age was calculated at baseline and fit as a categorical variable with three levels (<60, 60-70, and >70). Existing IHD was defined as the same diagnostic codes as the outcome occurring prior to baseline. For the KM analyses, MHD was presented as below or above the group specific median for visual simplicity. For the Cox PHM, MHD was fitted as a linear predictor.

Sub-analyses were conducted on populations with and without existing IHD as a secondary analysis and among lymphoma patients receiving high MHD (>5 Gy). Young patients in the breast and lymphoma cohorts were further subdivided into age groups (<40, 40-50, and 50-60), due to their abundance and heterogeneity. The KM and Cox PHM were replicated with the broader definition of diagnosis codes to represent CVD in place of IHD. This was for both the outcome and independent variable of pre-existing disease. The impact of including a low (V5) and high (V30) volumetric dose measure on model performance was assessed with log-rank tests comparing nested Cox PHMs.

All analyses were conducted using R statistical software [22,23]. Packages ‘prodlim’, ‘survival’, and ‘rms’ were used to conduct analysis [24–26]. Additionally, we leveraged ‘ggplot2’, ‘ggfortify’, ‘sjPlot’, and ‘cowplot’ to generate figures [27–30]. $p<0.05$ was considered significant.

Results

9,411 patients were included in the outcome analysis (Figure 1). Prior cancer was the main reason for exclusion. Breast cancer formed the largest proportion with 6808 patients, followed by lymphoma with 1251, esophageal with 627, and lung with 729. Breast cancer and lymphoma had the youngest populations, see Table 1. In terms of dosimetric data, MHD was less than 1.74 Gy in 90% of breast cancer cases.

The median follow-up time was 6.3 years (IQR 3.1-8.9). The 3-year cumulative incidence [95% CI] of CVD was greatest at 17% [13-21%] in esophageal, followed by 14% [11-17%] in lung, 9% [7-11%] in lymphoma, and 5.3% [5-6%] breast cancer patients (Figure 2). The 3-year cumulative incidence of death far outweighed that of CVD, especially in esophageal at 60% [56-64%] and in lung at 60% [56-63%], and to a lesser extent at 20% [17-22%] in lymphoma and 5% [4-5%] in breast cancer patients. Patients with breast cancer had a steady linearly increasing risk of death across time, while the other patient groups had a higher initial risk of death in the first year.

Existing IHD had a large impact on subsequent IHD across all groups (Figure 3) and was associated with poor OS in all but lung cancer patients (Figure 4). In contrast, the effect of MHD on IHD was small and non-significant, except for patients with lymphoma where the relationship was inverse, see Figure 3. A high MHD was significantly associated with poorer OS in breast and lung cancer patients, see Figure 4. Patients with lymphoma and a MHD above 5 Gy were also significantly younger than those with a lower MHD (median age of 44.9 compared to 63.3; $P < 0.001$ (Table A2)). Age had a significant impact on IHD in patients with breast cancer and lymphoma, and on OS in all cancers except esophageal.

The multivariable analysis demonstrated how the magnitude of existing IHD far outweighed the other predictors for subsequent IHD, see Figure 5. MHD was not significantly associated with IHD in any group. There was a large significant effect of MHD on OS in patients with breast cancer (HR 2.8) and a smaller, but still significant effect of high MHD on OS in patients with lung cancer (HR 1.2). Age was significantly associated with OS for all groups apart from esophageal cancer, while the effect was particularly strong in lymphoma patients. There was no significant effect of sex in any group for either outcome, though the direction of the effect was as expected with males having higher hazards than females (Figure A1).

As sensitivity analyses, we separately analyzed populations with and without existing IHD, but this did not vary greatly from the primary results (Figure A2; Figure A3; Figure A4). Younger patients with breast cancer and lymphoma did show distinct trends (Figure A5; Figure A6). Results for the broader definition of CVD did not differ significantly from IHD (Figure A7; Figure A8). There was substantial correlation observed between dose parameters (Figure A9; Table A3). Adding V5 or V30 to models including MHD therefore did not add significant explainability, and thus results and discussion focus on models without volumetric dose measures.

Discussion

This study first quantifies mortality and cardiotoxic side effects in thoracically irradiated cancer patients. It then analyzes the relative impact of radiation dose and other patient-specific factors on outcomes. We found the risk of death from cancer far exceeded the risk of cardiac disease, particularly among patients with lung and esophageal cancer. Contrary to earlier findings, no significant relationship was found between cardiac irradiation parameters and subsequent IHD [8–10]. Subsequent IHD risk was dominated by known background population factors of existing IHD, age, and gender in this study. Our findings support that modern RT cardiac sparing advancements limit radiation-related cardiotoxicity and tumor control remains the main clinical problem across all diagnoses studied. Known patient-specific prognostic risk factors impact outcomes significantly more than dose in the achievable cardiac dose range in this series. While these factors are not controllable, they may help identify high-risk patients and inform treatment planning [21].

Overall Survival

There is a dose-related association between high MHD and poorer OS in breast and lung cancer patients in accordance with other reports in literature [15]. For breast cancer patients, the guidelines define the extent of radiation, particularly parasternal lymph node irradiation, according to disease stage**.** Therefore, the observed dose relationship with OS may be confounded by patients receiving regional node irradiation having a higher dose to the heart and a less favorable survival of the index cancer. Similarly, different subtypes of lung cancer may tend to occur in different parts of the lungs. Additionally, the size of the primary tumor will influence the risk of irradiating the heart. Therefore, an endpoint specific analysis is necessary for further understanding.

It should surprise no-one that age influences the survival in patients with a favorable cancer prognosis (Figure 4; Figure A5). In addition to the obvious association with general mortality, there may be age-related staging effects and challenges such as poor adherence to treatment protocol. For breast cancer, mammography screening is not recommended after age 70 in Denmark, meaning a higher proportion of patient-detected and thus poorer-prognosis cases [31–33]. For lymphoma, the magnitude of this age effect is much larger than what we observed for breast cancer. The lymphoma data are complex due to confounding with lymphoma subtypes with very different prognoses being diagnosed at different ages and may be treated with radiotherapy at primary treatment or relapse. Nevertheless, the data shows that the prognosis of elderly patients is poor. The systemic treatment for lymphomas, typically developed on younger trial patients, is generally quite aggressive and older patients often do not tolerate this treatment well. The role of radiotherapy is therefore even more important in older patients where the detrimental effects of radiation to the heart need to be carefully balanced against the necessity for adequate treatment of the active cancer.

Ischemic Heart Disease

Existing IHD was found to increase the risk of subsequent IHD as the strongest predictor in the model (Figure 3). The scale of existing IHD's impact can be quantified relative to the dose effect. For breast cancer, the presence of existing IHD is equivalent to a MHD increase of 12.5 Gy – doses that should not be seen after contemporary treatment. Similarly, the HR between the oldest age group compared to the youngest is equivalent to an additional 5.4 Gy. Pre-existing disease status and age must therefore be

considered when selecting patient interventions, such as late effects surveillance or proton therapy [34,35].

MHD is inversely related to IHD in lymphoma. This result is confounded by the association of age and lymphoma subtype. Younger patients more often have mediastinal disease, typically Hodgkin lymphoma or primary mediastinal large cell lymphoma, which have good prognoses. The combination of young age and good prognoses, even paired with a greater risk of irradiating the heart due to mediastinal RT, seems to confer a lower risk of IHD. However, in a multivariable analysis, this MHD effect on IHD is non-significant.

For breast cancer patients, the observed impact of MHD on OS is unlikely to be mediated by an association with subsequent IHD or CVD . This suggests the dominant effect of OS may be related to the higher disease stage associated with the wider tangential RT fields to cover regional lymph nodes. Prior studies have shown a clear link between MHD and CVD. However, the doses being observed in this contemporary cohort are lower than the doses in previous studies [36]. Guidelines in Denmark were updated to emphasize target coverage over cardiac dose following the 2016 DBCG-IMN's findings [37]. Our data support this decision. The primary disease should be the main concern and sparing of normal tissue should largely be achieved through technological advances rather than compromising the target dose.

For patients with lung cancer, this study supports the detrimental effect on OS from higher MHD reported by the Manchester group [15]. Yet, we did not observe a corresponding increase in risk of IHD, again putting a question mark over cause and effect. This suggests that survival impacts may be more related to disease characteristics rather than treatment toxicity, as is the case with breast cancer patients. This could include confounding by association between cardiac exposure and mediastinal lymph node involvement or disease extent. Additionally, most lung cancer patients experience recurrence and therefore the competing risk of death due to cancer is high compared to risk of death due to cardiac disease. Access to the specific endpoints of cardiotoxicity improves the interpretability of the dose-response findings compared to assessing OS alone. Our data contributes to quantifying the discussion of the prioritization between target coverage and cardiac risk in lung cancer patients. Still, there is a possibility of significant detrimental effect of radiation to the heart, given the large confidence intervals for the estimate of the effect of dose. Nevertheless, disease control should remain the top priority in lung cancer patients (Figure 5).

Strengths & Limitations

This study integrates detailed 3D dosimetry on 9,411 patients from a single institution with individually delineated hearts and linked outcome data from Denmark's national registry systems. The dataset enables the investigation and quantification of cardiotoxic side effects of radiation doses over a 12-year period. Such data can inform the clinically important discussion of balancing exposure of normal tissue against the need to treat the active cancer in patients undergoing radiotherapy with greater confidence than smaller studies with manual data analysis.

A limitation is the study's lack of chemotherapy data. Cardiotoxicity of some chemotherapy drugs, such as anthracyclines, are well established [38]. Unfortunately, the registries did not allow reliable extraction

of chemotherapy data. Similarly, other data, such as diabetes status or BMI, could have been valuable; this data was not reliably available for such a large cohort, but could potentially in the future be extracted using natural language processing of patient charts [39,40]. Additionally, inclusion of staging and tumor characteristics would have added substantial information to a predictive model. Inclusion of coronary artery calcium (CAC) scoring may also contribute to better predictions [41]. Despite these limitations, the model captures important patient related factors combined with dose.

Another limitation is that the median follow-up time of 6.5 years is limited due to data availability within EHR systems. The statistical power is sufficient to solidly resolve conventional risk factors of age and pre-existing disease, but despite 540 events of IHD we do not observe a significant dose response relationship. Supplementary Figure A10 converts our results to excess relative risk (ERR) per Gy for comparison with Darby and Nimwegen and demonstrates overlapping but much wider confidence intervals in this cohort study [9,10]. It may be the narrow range of heart doses in the contemporary patients which challenges the fitting process and limits the power of the current study. Extended follow-up would be advantageous, and the method here could be expanded to larger collaborations for further confirmation.

Future studies could consider target coverage in addition to cardiac dose. For lung cancers, the competing risk of death is so high that a much larger population is needed to detect the effect of MHD on IHD, due to endpoint rarity. The point estimate is still trending as expected, suggesting a trend consistent with observed toxic effects of cardiac exposure [1]. However, the scale of the impact of this aggressive disease significantly overshadows the detrimental impact of radiation on cardiovascular health.

An alternative approach to using whole heart dose metrics is a focus on dose effects on cardiac substructures. Some studies investigating the relationship between cardiac dose, cardiac events and mortality suggest that dose to cardiac substructures is more important than whole heart dose [12,18,42,43]. Zhang et al performed a systematic review of eighteen studies published before 2018 which included locally advanced NSCLC patients treated with concurrent chemo-radiotherapy [44]. All twenty parameters were found to be associated significantly with OS and IHD, the most frequent being MHD, V5, and V30. However, unlike this study, consistent heart dose-volume parameters associated with OS of patients with NSCLC were not identified. Several heart contouring atlases have been published aiming at consistent dose reporting to cardiac substructures [15,16,42,45]. Additionally, multiple open-access AI tools have been published which may allow us to investigate substructure dose trends[46,47] . Furthermore, localization of radio-sensitive regions can be assessed with image-based data mining tools [48–50]. Despite these future prospects, the small magnitude of effect of radiation dose to the whole heart compared to patient related risk factors would suggest the probability of finding a strong association with a substructure of the heart is probably modest.

In conclusion, we have analyzed a large dataset for registry-based assessment of radiation dose effect for the heart. Within the confidence intervals, our dose-effect estimates are consistent with Darby's famous ERR=7.4 %/Gy, but the risk of cardiac radiation exposure with contemporary treatment techniques is dwarfed by patient-related risk factors and the competing risk of death for most patients. The primary cancer diagnosis remains the single greatest health risk across studied populations, pointing to the importance of reducing cardiac radiation *without* compromising the target coverage.

References

1. Atkins KM, Rawal B, Chaunzwa TL, et al. Cardiac Radiation Dose, Cardiac Disease, and Mortality in Patients With Lung Cancer. *J Am Coll Cardiol*. 2019;73(23):2976-2987. doi:10.1016/j.jacc.2019.03.500

2. Sun F, Franks K, Murray L, et al. Cardiovascular mortality and morbidity following radical radiotherapy for lung cancer: Is cardiovascular death under-reported? *Lung Cancer*. 2020;146:1-5. doi:10.1016/j.lungcan.2020.05.004

3. Verma V, Simone C, Werner-Wasik M. Acute and Late Toxicities of Concurrent Chemoradiotherapy for Locally-Advanced Non-Small Cell Lung Cancer. *Cancers (Basel)*. 2017;9(12):120. doi:10.3390/cancers9090120

4. Wang K, Eblan MJ, Deal AM, et al. Cardiac Toxicity After Radiotherapy for Stage III Non–Small-Cell Lung Cancer: Pooled Analysis of Dose-Escalation Trials Delivering 70 to 90 Gy. *Journal of Clinical Oncology*. 2017;35(13):1387-1394. doi:10.1200/JCO.2016.70.0229

5. Dess RT, Sun Y, Matuszak MM, et al. Cardiac Events After Radiation Therapy: Combined Analysis of Prospective Multicenter Trials for Locally Advanced Non–Small-Cell Lung Cancer. *Journal of Clinical Oncology*. 2017;35(13):1395-1402. doi:10.1200/JCO.2016.71.6142

6. Darby SC, McGale P, Taylor CW, Peto R. Long-term mortality from heart disease and lung cancer after radiotherapy for early breast cancer: prospective cohort study of about 300 000 women in US SEER cancer registries. *Lancet Oncol*. 2005;6(8):557-565. doi:10.1016/S1470-2045(05)70251-5

7. van Velzen SGM, Gal R, Teske AJ, et al. AI-Based Radiation Dose Quantification for Estimation of Heart Disease Risk in Breast Cancer Survivors After Radiation Therapy. *International Journal of Radiation Oncology*Biology*Physics*. 2022;112(3):621-632. doi:10.1016/j.ijrobp.2021.09.008

8. Cutter DJ, Schaapveld M, Darby SC, et al. Risk for valvular heart disease after treatment for hodgkin lymphoma. *J Natl Cancer Inst*. 2015;107(4). doi:10.1093/jnci/djv008

9. Van Nimwegen FA, Schaapveld M, Cutter DJ, et al. Radiation dose-response relationship for risk of coronary heart disease in survivors of Hodgkin lymphoma. *Journal of Clinical Oncology*. 2016;34(3):235-243. doi:10.1200/JCO.2015.63.4444

10. Darby SC, Ewertz M, McGale P, et al. Risk of ischemic heart disease in women after radiotherapy for breast cancer. *N Engl J Med*. 2013;368(11):987-998. doi:10.1056/NEJMoa1209825

11. Gent DG, Dobson R. The 2022 European Society of Cardiology Cardio-oncology Guidelines in Focus. *European Cardiology Review*. 2023;18. doi:10.15420/ecr.2022.63

12. Banfill K, Giuliani M, Aznar M, et al. Cardiac Toxicity of Thoracic Radiotherapy: Existing Evidence and Future Directions. *Journal of Thoracic Oncology*. 2021;16(2):216-227. doi:10.1016/j.jtho.2020.11.002

13. Forbes N, Terrones-Campos C, Smith A, et al. Time trends in cardiac doses in 10,000 patients receiving curative thoracic radiation therapy between 2009 and 2020. *medRxiv*. Published online 2024. doi:10.1101/2024.02.18.24303007

14. Niska JR, Thorpe CS, Allen SM, et al. Radiation and the heart: systematic review of dosimetry and cardiac endpoints. *Expert Rev Cardiovasc Ther*. 2018;16(12):931-950. doi:10.1080/14779072.2018.1538785

15. McWilliam A, Kennedy J, Hodgson C, Vasquez Osorio E, Faivre-Finn C, van Herk M. Radiation dose to heart base linked with poorer survival in lung cancer patients. *Eur J Cancer*. 2017;85:106-113. doi:10.1016/j.ejca.2017.07.053

16. McWilliam A, Khalifa J, Vasquez Osorio E, et al. Novel Methodology to Investigate the Effect of Radiation Dose to Heart Substructures on Overall Survival. *International Journal of Radiation Oncology*Biology*Physics*. 2020;108(4):1073-1081. doi:10.1016/j.ijrobp.2020.06.031

17. van den Bogaard VAB, Ta BDP, van der Schaaf A, et al. Validation and Modification of a Prediction Model for Acute Cardiac Events in Patients With Breast Cancer Treated With Radiotherapy Based on Three-Dimensional Dose Distributions to Cardiac Substructures. *Journal of Clinical Oncology*. 2017;35(11):1171-1178. doi:10.1200/JCO.2016.69.8480

18. Atkins KM, Chaunzwa TL, Lamba N, et al. Association of Left Anterior Descending Coronary Artery Radiation Dose With Major Adverse Cardiac Events and Mortality in Patients With Non–Small Cell Lung Cancer. *JAMA Oncol*. 2021;7(2):206. doi:10.1001/jamaoncol.2020.6332

19. Smith AG, Petersen J, Terrones-Campos C, et al. RootPainter3D: Interactive-machine-learning enables rapid and accurate contouring for radiotherapy. *Med Phys*. 2022;49(1):461-473. doi:10.1002/mp.15353

20. Forbes N. SAP_v1_Clinical and dosimetric risk factors for ischemic heart disease.pdf. Published online October 4, 2023.

21. Sathish N, Wu CLA. *Let's Flip: An Approach to Understand Median Follow-up by the Reverse Kaplan-Meier Estimator from a Statistical Programmer's Perspective*.

22. Posit team. RStudio: Integrated Development Environment for R. Published online 2023. http://www.posit.co/

23. R Core Team. R: A Language and Environment for Statistical Computing. Published online 2023. https://www.R-project.org/

24. Estimation H. Standard Survival Analysis. *Analysis*. Published online 2010.

25. Therneau Terry. A Package for Survival Analysis in R. *R package version 35-0*. Published online 2023.

26. Harrell Jr FE. rms: Regression Modeling Strategies. R package version 5.0-0. *CRAN*. Published online 2016.

27. Wilkinson L. ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H. *Biometrics*. 2011;67(2). doi:10.1111/j.1541-0420.2011.01616.x

28. Tang Y, Horikoshi M, Li W. ggfortify: Data Visualization Tools for Statistical Analysis Results. *R J*. 2016;8(2).

29. Lüdecke D. sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.10. *Parasitol Res*. 2021;115(ttps://CRAN.R-project.org/package=sjPlot.).

30. Wilke CO. cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2". R package version 0.7.0. *URL: https://CRANR-project.org/package=cowplot Acessed 18 October 2018*. Published online 2016.

31. Screening for breast cancer. Sundhedsstyrelsen. https://www.sst.dk/en/english/General-public/Screening-for-cancer/Screening-for-breast-cancer

32. Frank S, Carton M, Dubot C, et al. Impact of age at diagnosis of metastatic breast cancer on overall survival in the real-life ESME metastatic breast cancer cohort. *The Breast*. 2020;52:50-57. doi:10.1016/j.breast.2020.04.009

33. Fernandes JO, Cardoso-Filho C, Kraft MB, et al. Differences in breast cancer survival and stage by age in off-target screening groups: a population-based retrospective study. *AJOG Global Reports*. 2023;3(2):100208. doi:10.1016/j.xagr.2023.100208

34. Treanor C, Donnelly M. The late effects of cancer and cancer treatment: a rapid review. *J Community Support Oncol*. 2014;12(4):137-148. doi:10.12788/jcso.0035

35. Mutter RW, Choi JI, Jimenez RB, et al. Proton Therapy for Breast Cancer: A Consensus Statement From the Particle Therapy Cooperative Group Breast Cancer Subcommittee. *International Journal of Radiation Oncology*Biology*Physics*. 2021;111(2):337-359. doi:10.1016/j.ijrobp.2021.05.110

36. Taylor CW, Wang Z, Macaulay E, Jagsi R, Duane F, Darby SC. Exposure of the Heart in Breast Cancer Radiation Therapy: A Systematic Review of Heart Doses Published During 2003 to 2013. *International Journal of Radiation Oncology*Biology*Physics*. 2015;93(4):845-853. doi:10.1016/j.ijrobp.2015.07.2292

37. Thorsen LBJ, Offersen BV, Danø H, et al. DBCG-IMN: A Population-Based Cohort Study on the Effect of Internal Mammary Node Irradiation in Early Node-Positive Breast Cancer. *Journal of Clinical Oncology*. 2016;34(4):314-320. doi:10.1200/JCO.2015.63.6456

38. Volkova M, Russell R. Anthracycline Cardiotoxicity: Prevalence, Pathogenesis and Treatment. *Curr Cardiol Rev*. 2012;7(4):214-220. doi:10.2174/157340311799960645

39. Wu S, Roberts K, Datta S, et al. Deep learning in clinical natural language processing: a methodical review. *Journal of the American Medical Informatics Association*. 2020;27(3):457-470. doi:10.1093/jamia/ocz200

40. Luo Y, Thompson WK, Herr TM, et al. Natural Language Processing for EHR-Based Pharmacovigilance: A Structured Review. *Drug Saf*. 2017;40(11):1075-1089. doi:10.1007/s40264-017-0558-6

41. Olloni A, Brink C, Lorenzen EL, et al. Does coronary artery calcium score have an impact on overall survival for locally advanced non-small cell lung cancer treated with definitive radiotherapy. *Radiotherapy and Oncology*. 2023;185:109719. doi:10.1016/j.radonc.2023.109719

42. Stam B, Peulen H, Guckenberger M, et al. Dose to heart substructures is associated with non-cancer death after SBRT in stage I–II NSCLC patients. *Radiotherapy and Oncology*. 2017;123(3):370-375. doi:10.1016/j.radonc.2017.04.017

43. Piroth MD, Baumann R, Budach W, et al. Heart toxicity from breast cancer radiotherapy. *Strahlentherapie und Onkologie*. 2019;195(1):1-12. doi:10.1007/s00066-018-1378-z

44. Zhang TW, Snir J, Boldt RG, et al. Is the Importance of Heart Dose Overstated in the Treatment of Non-Small Cell Lung Cancer? A Systematic Review of the Literature. *International Journal of Radiation Oncology*Biology*Physics*. 2019;104(3):582-589. doi:10.1016/j.ijrobp.2018.12.044

45. Chun SG, Hu C, Choy H, et al. Impact of Intensity-Modulated Radiation Therapy Technique for Locally Advanced Non–Small-Cell Lung Cancer: A Secondary Analysis of the NRG Oncology RTOG 0617 Randomized Clinical Trial. *Journal of Clinical Oncology*. 2017;35(1):56-62. doi:10.1200/JCO.2016.69.1378

46. Finnegan RN, Chin V, Chlap P, et al. Open-source, fully-automated hybrid cardiac substructure segmentation: development and optimisation. *Phys Eng Sci Med*. 2023;46(1):377-393. doi:10.1007/s13246-023-01231-w

47. Wasserthal J, Breit HC, Meyer MT, et al. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiol Artif Intell*. 2023;5(5). doi:10.1148/ryai.230024

48. Ibragimov B, Toesca DAS, Chang DT, et al. Deep learning for identification of critical regions associated with toxicities after liver stereotactic body radiation therapy. *Med Phys*. 2020;47(8):3721-3731. doi:10.1002/mp.14235

49. Green A, Vasquez Osorio E, Aznar MC, McWilliam A, van Herk M. Image Based Data Mining Using Per-voxel Cox Regression. *Front Oncol*. 2020;10. doi:10.3389/fonc.2020.01178

50. Ebert MA, Gulliford S, Acosta O, et al. Spatial descriptions of radiotherapy dose: normal tissue complication models and statistical associations. *Phys Med Biol*. 2021;66(12):12TR01. doi:10.1088/1361-6560/ac0681

Tables and Figures

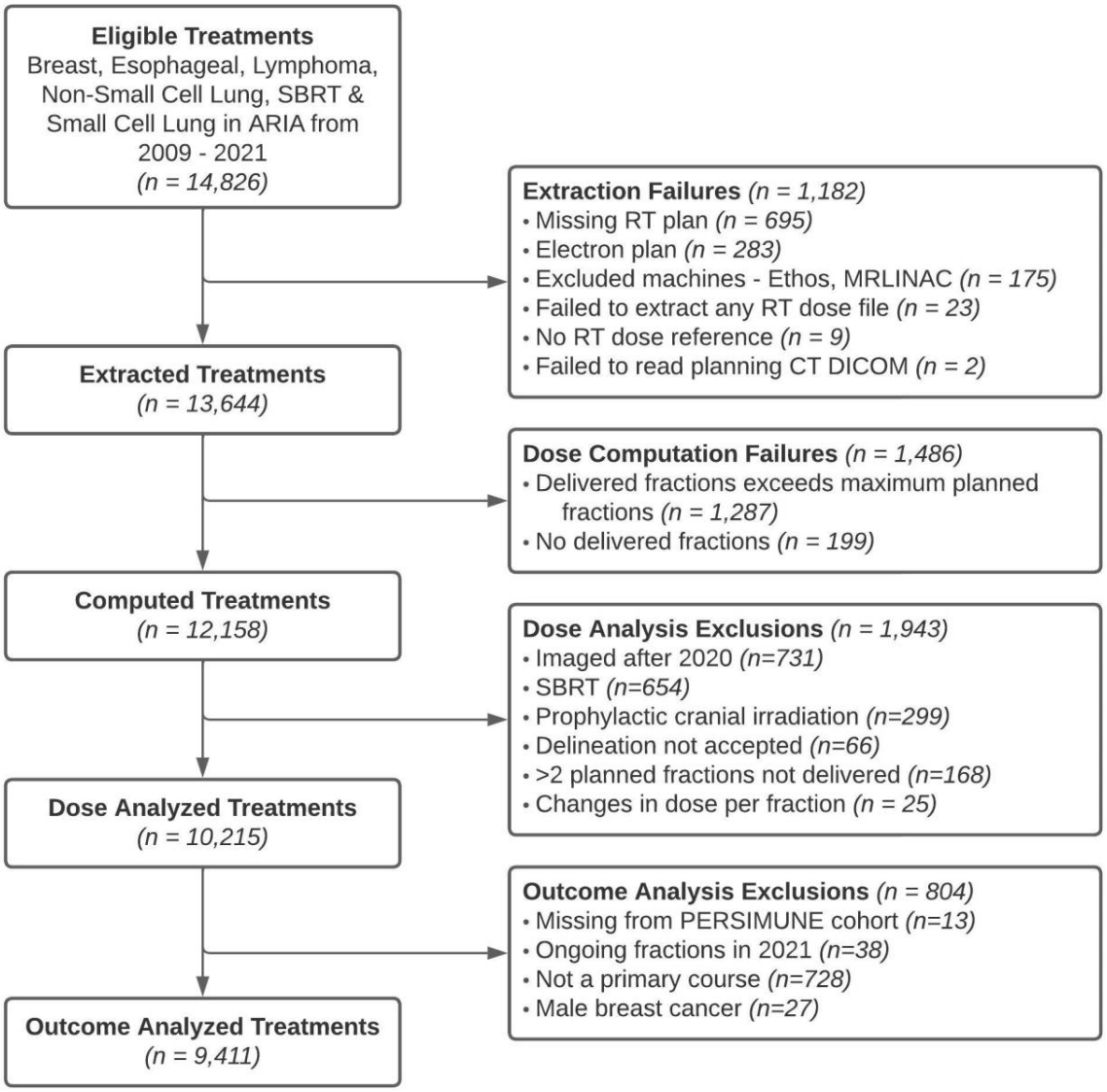


Figure 1. Consort diagram.

Table 1. Patient characteristics presented as independent variables – sex, age, existing cardiovascular disease (CVD), and dose metrics – and outcomes – subsequent CVD, death, and follow-up time. CVD includes ischemic heart disease (IHD), valvular disease (VD), and heart failure (HF). Existing CVD is defined as occurring before baseline (delivery date of last fraction) and subsequent CVD is defined as occurring after baseline. Dose metrics include mean heart dose (MHD), absolute volume receiving at least 5 Gy (V5), and absolute volume receiving at least 30 Gy (V30). Doses are presented in equivalent dose in 2 Gy fractions (EQD2). Continuous variables are presented as the median (MDN) and interquartile range (IQR). Categorical variables are presented as the number (N) and percentage (%) of patients. Follow-up time to first event is calculated using the reverse Kaplan-Meier estimator.

	BREAST CANCER (N = 6804)	**ESOPHAGEAL CANCER** (N = 627)	**LYMPHOMA** (N = 1251)	**NON-SMALL CELL LUNG CANCER** (N = 553)	**SMALL CELL LUNG CANCER** (N = 176)
INDEPENDENT VARIABLES					
MALE, N (%)	0 (0%)	437 (70%)	666 (53%)	304 (55%)	82 (47%)
AGE *(YRS)*, MDN (IQR)	62 (54,69)	66 (61,72)	63 (47,72)	67 (60,72)	66 (60,71)
< 60, N (%)	2820 (41%)	141 (22%)	555 (44%)	135 (24%)	43 (24%)
60-70, N (%)	2634 (39%)	282 (45%)	310 (25%)	220 (40%)	81 (46%)
>70, N (%)	1350 (20%)	204 (33%)	386 (31%)	198 (36%)	52 (30%)
EXISTING CVD, N (%)	618 (9%)	136 (22%)	166 (13%)	111 (20%)	28 (16%)
EXISTING IHD, N (%)	506 (7%)	119 (19%)	134 (11%)	98 (18%)	26 (15%)
EXISTING VD, N (%)	100 (1%)	13 (2%)	28 (2%)	15 (3%)	4 (2%)
EXISTING HF, N (%)	119 (2%)	29 (5%)	47 (4%)	30 (5%)	7 (4%)
MHD *(GY)* MDN (IQR)	0.61 (0.31,1.1)	12.1 (2.8,18.3)	0.02 (0,0.8)	4.2 (1.3,10.0)	4.8 (1.6,8.5)
V5 *(CM^3)*, MDN (IQR)	2 (0,14)	471 (95,689)	0 (0,9)	132 (39,310)	167 (48,330)
V30 *(CM^3)*, MDN (IQR)	0 (0,1)	72 (10,148)	0 (0,0)	19 (0.2,74)	22 (2, 69)
OUTCOMES					
FOLLOW-UP *(YRS)*, MDN (IQR)	6.5 (3.2,9.1)	5.02 (3.0,8.3)	5.3 (2.3,8.1)	5.7 (3.7,7.3)	6.7 (4.4,9.5)
DEATH, N (%)	874 (13%)	424 (68%)	339 (27%)	393 (71%)	125 (71%)
SUBSEQUENT CVD, N (%)	603 (9%)	94 (15%)	127 (10%)	72 (13%)	18 (10%)
SUBSEQUENT IHD, N (%)	342 (5%)	61 (10%)	73 (6%)	49 (9%)	15 (9%)
SUBSEQUENT VD, N (%)	184 (3%)	17 (3%)	28 (2%)	9 (2%)	2 (1%)
SUBSEQUENT HF, N (%)	210 (3%)	41 (7%)	63 (5%)	35 (6%)	5 (3%)

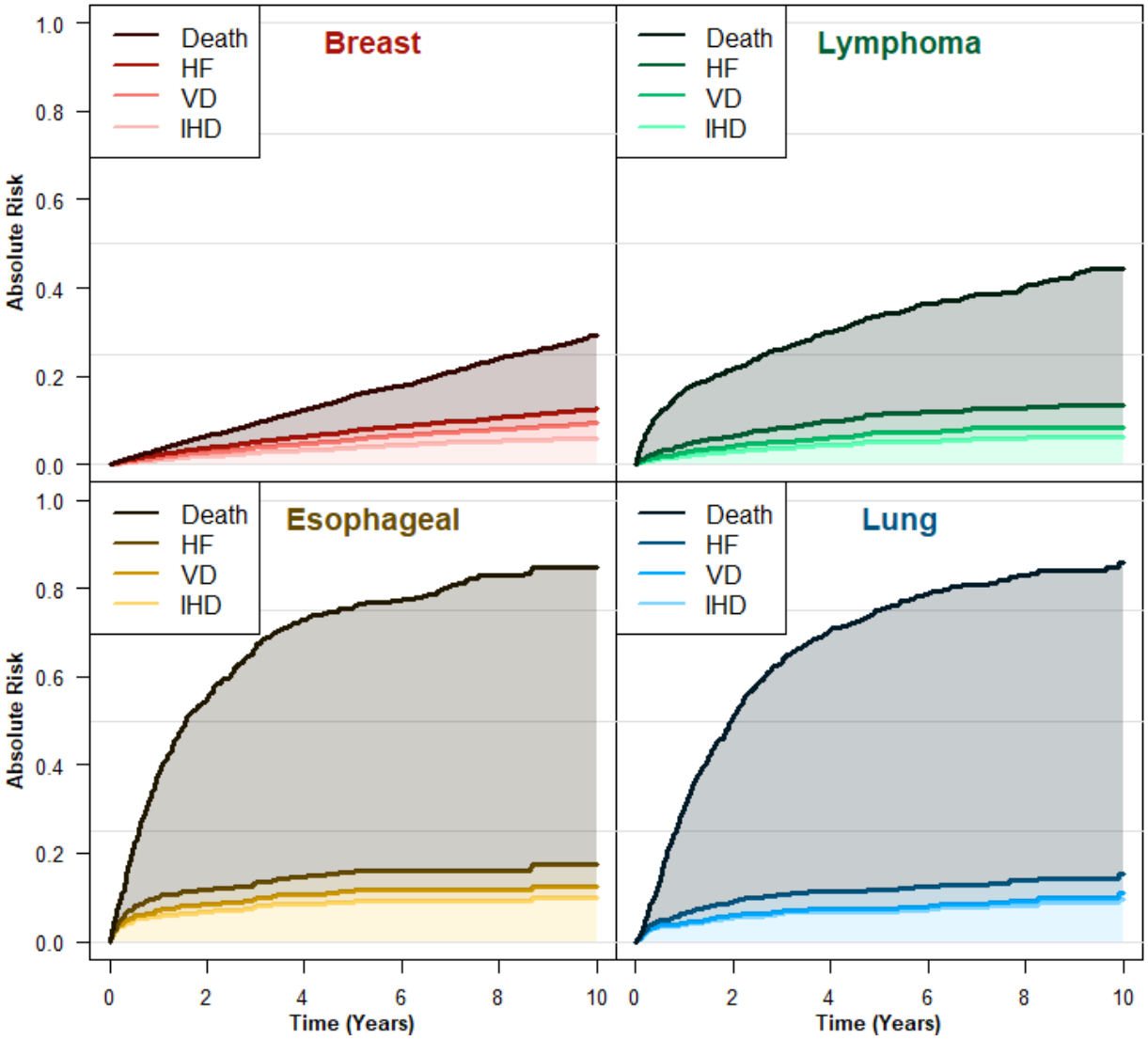


Figure 2. Stacked cumulative incidence plots for each outcome – death, heart failure (HF), valvular disease (VD), and ischemic heart disease (IHD), separated by diagnosis. Numbers at risk available in supplementary Tables A4-6.

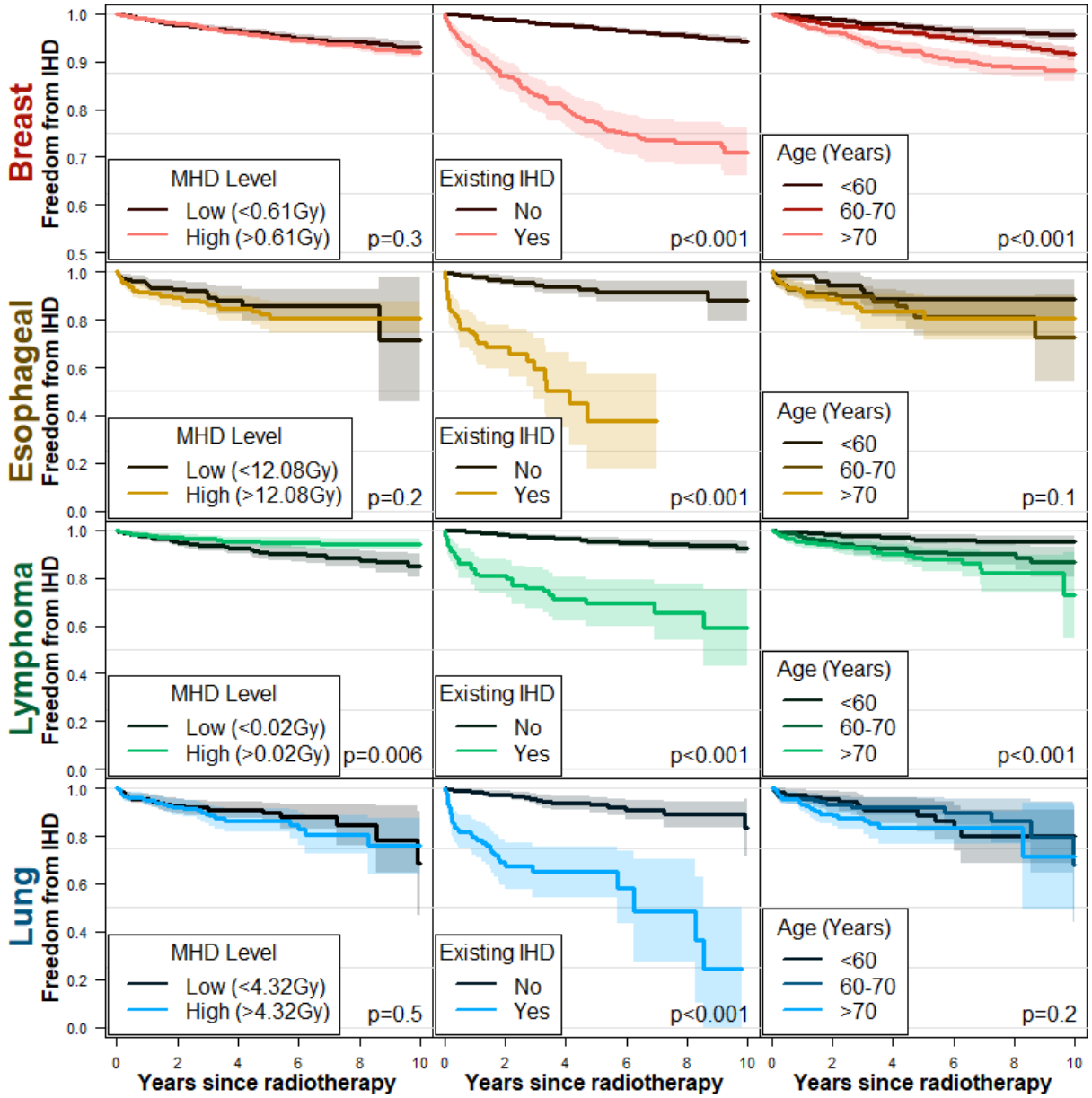


Figure 3. Kaplan-Meier plots for freedom from ischemic heart disease (IHD), by mean heart dose (MHD) level (below or above median), existing IHD, and age group (<60, 60-70, and >70 years old), separated by diagnosis. P-values generated from log-rank test (MHD level and existing IHD) and test for trend (age group). *Note that the Breast cancer y-axis is truncated at 0.5 to better show separation.* Numbers at risk available in supplementary Table A4. KM plots for overall CVD available in supplement Figure A7. For the effect of sex on freedom from IHD see supplementary Figure A1.

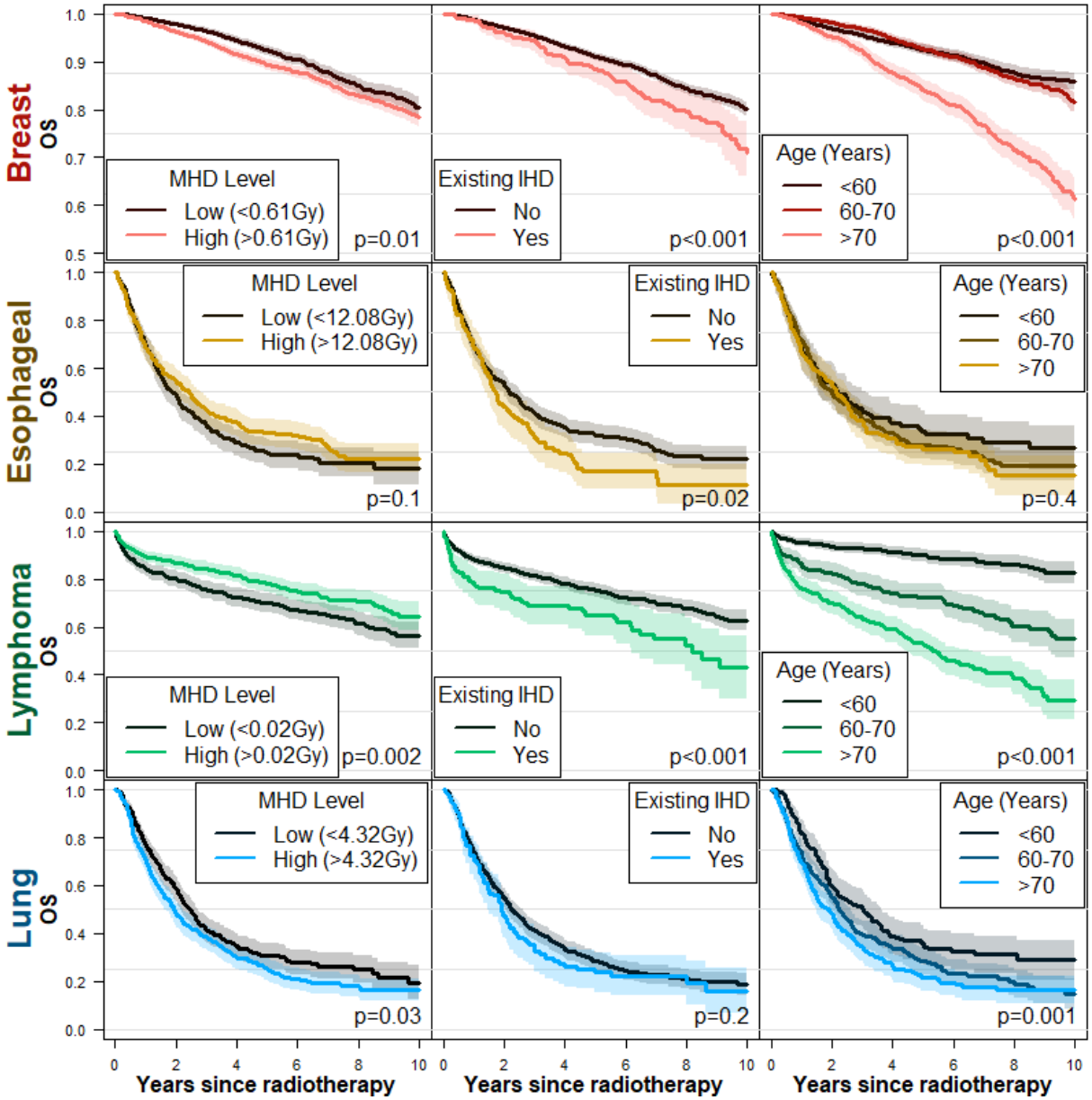


Figure 4. Kaplan-Meier plots for overall survival (OS), by mean heart dose (MHD) level (below or above median), existing ischemic heart disease (IHD), and age group (<60, 60-70, and >70 years old), separated by diagnosis. *Note that the Breast cancer y-axis is truncated at 0.5 to better show separation*. P-values generated from log-rank test (MHD level and existing IHD) and test for trend (age group). Numbers at risk available in supplement Table A5. For the effect of sex on OS see supplementary Figure A1.

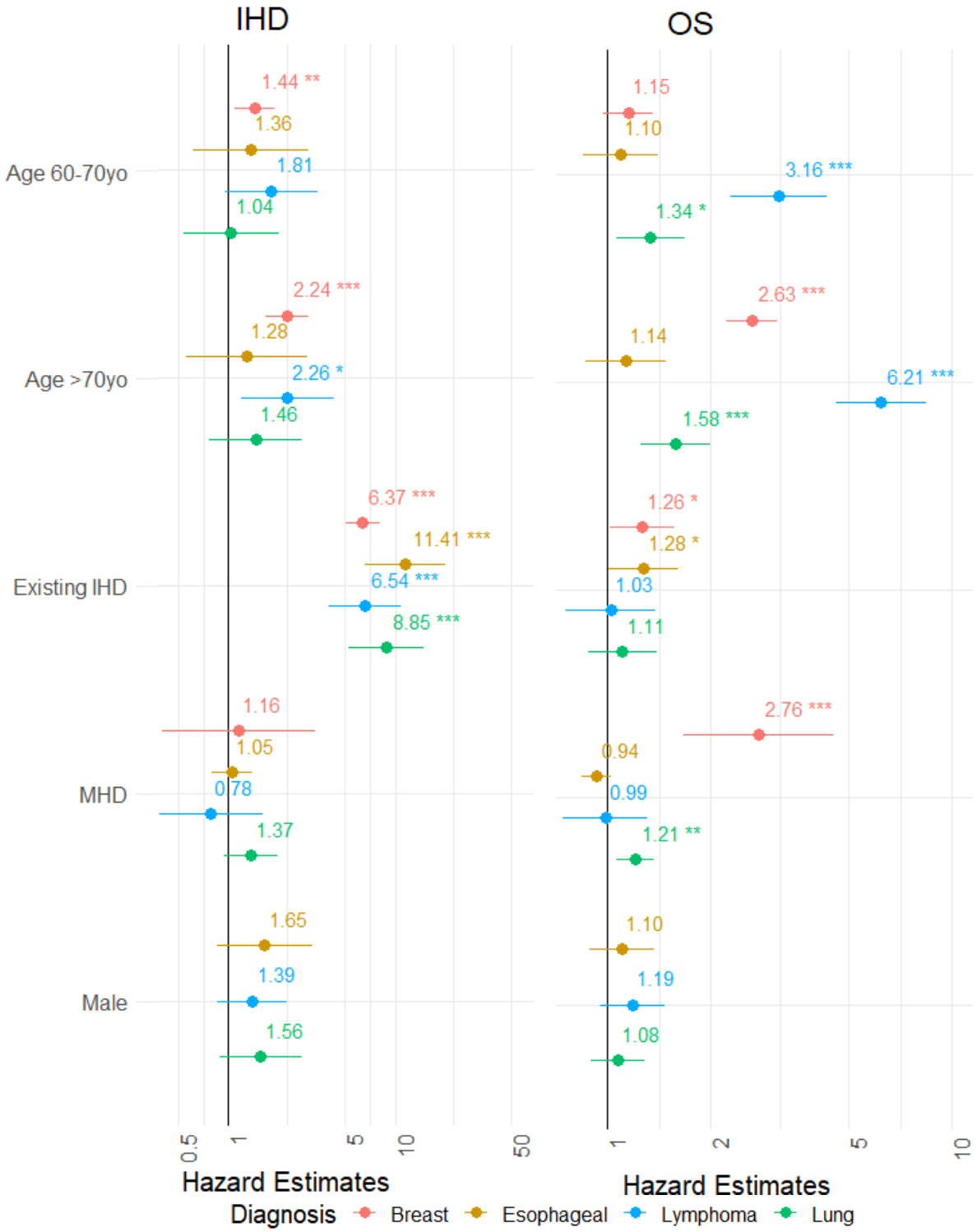


Figure 5. Forest plots of hazard estimates with 95% CI for predictors in eight Cox proportion hazards models on two outcomes - ischemic heart disease (IHD) and overall survival (OS) - separated by four diagnoses. The Breast cancer cohort only includes females. A hazard estimate above one indicates an increased risk of the outcome. Forest plots for overall CVD in supplement Figure A8. P-value thresholds are 0.05(*), 0.01 (**), and 0.001(***).

Appendix A – Supplementary Tables and Figures

Table A1. Rationale for changes in methodology from statistical analysis plan (SAP) to manuscript

Aspect	SAP	Manuscript	Rationale
Cohort size	Estimated cohort size is 10,492 patients.	Final cohort analyzed includes 9,411 patients.	Cohort size was reduced due to limiting ICD code list to I20-I25. This was not implemented at the time of estimation.
High Dose Metric	Plans to use volume receiving at least 40 Gray (V40) as the high dose metric.	Analysis used volume receiving at least 30 Gray (V30) as the high dose metric.	V40 was zero-dominant for the largest cohorts - breast and lymphoma. After discussing with physicians, we decided to use a slightly more informative variable. Upon testing, this did not seem to have an impact.
Volumetric dose measure analysis	Plans to investigate mean heart dose (MHD) and considers additional low and high volumetric dose measures.	Analyzes MHD as the primary metric. Limited assessment of volumetric measures included.	These volumetric measures did not significantly enhance the model. This is likely due to the high amount of correlation. They were excluded to prioritized simplicity and explainability. Future work may explore this further.
Stratification	The analysis was intended to be stratified by diagnostic group.	Separate models were run for each diagnostic group.	It was not reasonable to assume that predictors had a similar effect across different strata due to distinct characteristics of each diagnostic group.
Splines	Proposed using spline models for MHD and age, depending on AIC performance.	Used only linear predictors for MHD and a categorical variable for age.	Testing revealed no significant improvement of models with inclusion of splines. Prioritized simplification of the statistical model to avoid overfitting and aid interpretation.
Competing Risk Model	Fine & Gray method proposed for competing risk analysis.	Aalen-Johansen Estimator used for competing risk analysis.	Preference for a non-parametric approach to handle competing risks, which provides a method for cumulative incidence estimation. Covariates were assessed with the Cox proportional hazards model.
Exploratory Analyses	Suggested potential time-dependent Cox regression and separate analyses for esophageal and lung cancers.	Includes sensitivity analyses and sub-analyses, including separation by age and existing IHD status.	The additional analyses provided further granularity based on specific patient characteristics, improving the depth of findings.
Endpoints Scope	Focuses on Ischemic Heart Disease (IHD) as the primary endpoint.	Analyzes both IHD and a broader category of cardiovascular disease (CVD), including valvular disease and heart failure.	Expanded focus on a broader range of cardiovascular outcomes reflects a more comprehensive analysis.
Cause of death	Death from IHD considered an event. Death from other causes considered a censor.	Only death from IHD coded correctly by ICD codes in the LPR were captured in events.	Cause of death is notoriously unreliable. Therefore, we stuck to ICD codes from LRP for defining CVD events and censored deaths from CPR.
Outcome Measures	Primary analysis centered around Cox Proportional Hazards Models for time to IHD.	Expands to include overall survival (OS) as a critical endpoint in addition to IHD and CVD.	Inclusion of OS allows for a more holistic view of patient outcomes, beyond just IHD.

Table A2. Number of lymphoma patients with mean heart dose (MHD) > 5 Gy by age group. Chi-squared test for independence yields a p-value < 0.001.

	MHD <5 Gy	MHD >5 Gy
<60 years old	449	103
60-70 years old	285	23
>70 years old	359	32

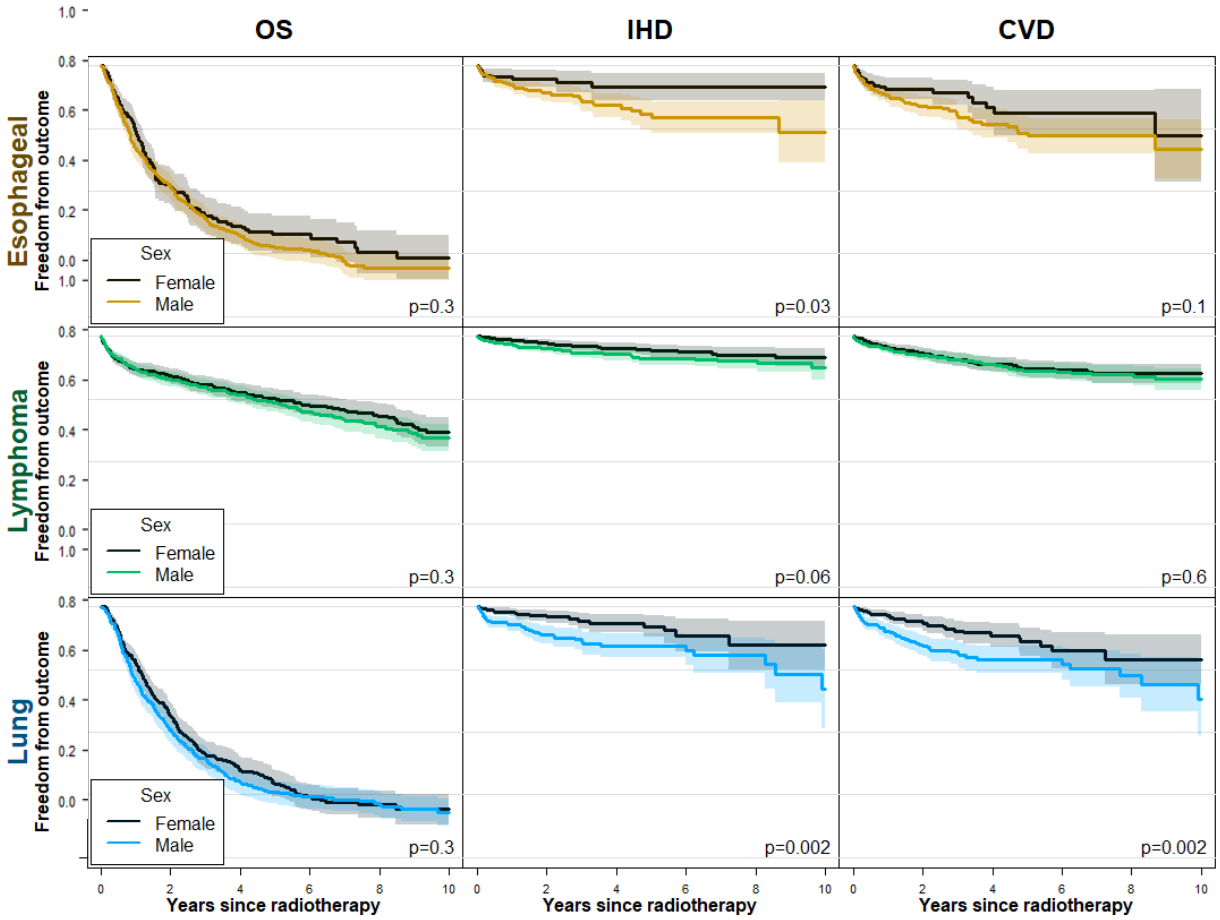


Figure A1. Kaplan-Meier plots for overall survival (OS), ischemic heart disease (IHD), and cardiovascular disease (CVD) by sex in three diagnosis groups. Note: Breast cancer was excluded since only females were analyzed. P-values generated from test for trend.

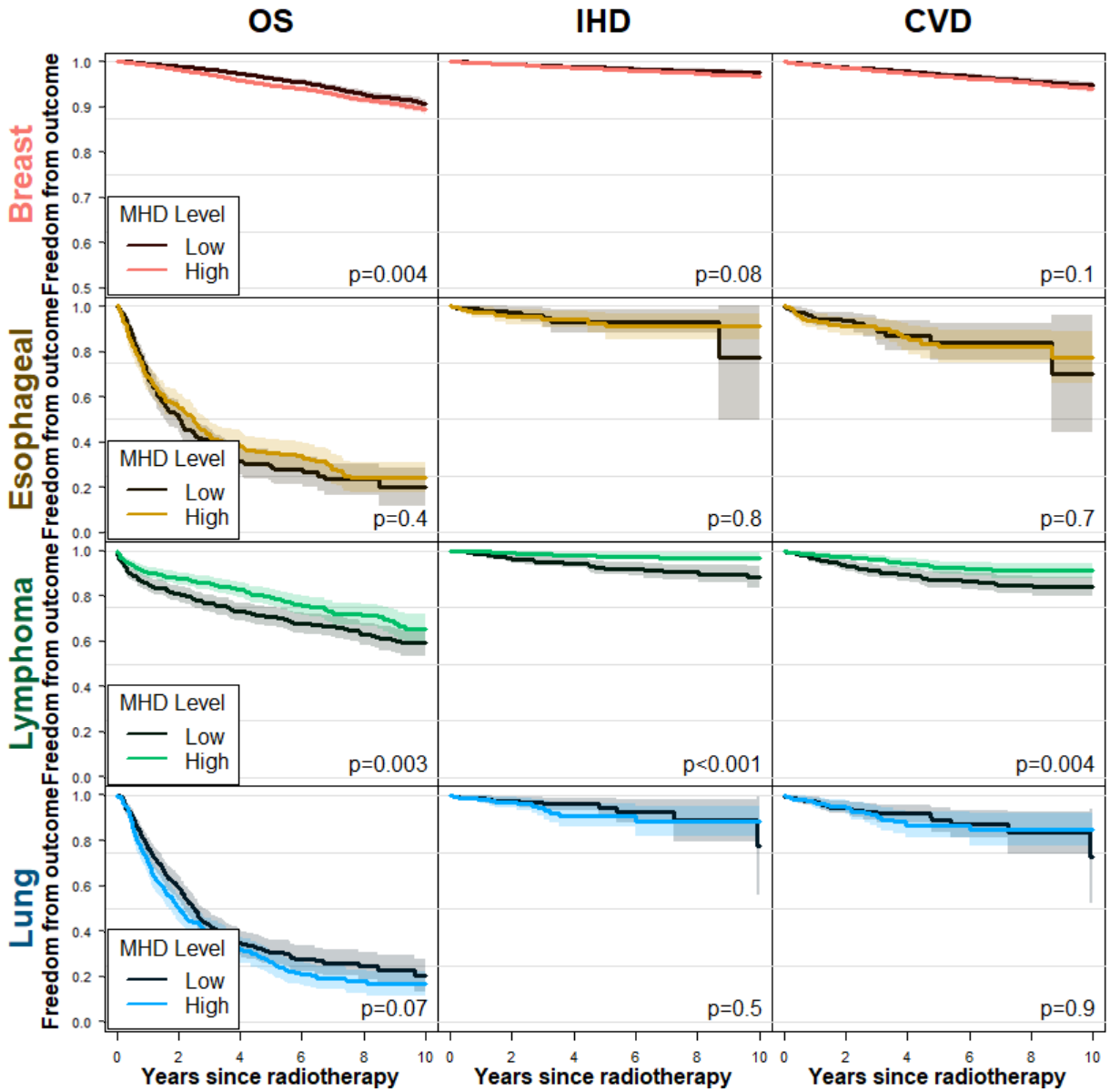


Figure A2. Secondary sub-analyses on populations without existing IHD by diagnosis group. Kaplan-Meier plots for overall survival (OS), ischemic heart disease (IHD), and cardiovascular disease (CVD) by mean heart dose (MHD) level – above and below the group median. P-values generated from test for trend.

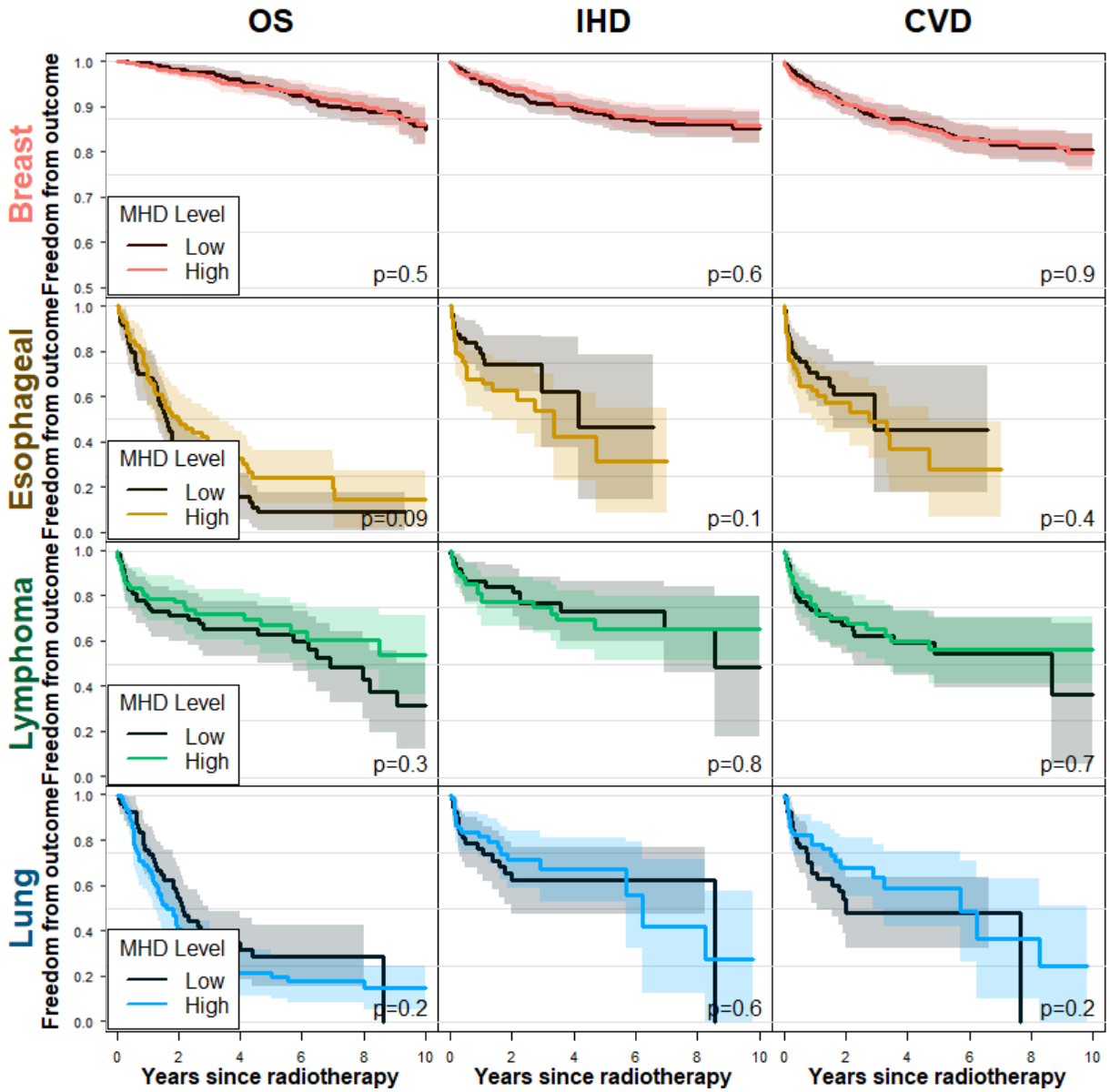


Figure A3. Secondary sub-analyses on populations with existing IHD, by diagnosis group. Kaplan-Meier plots for overall survival (OS), ischemic heart disease (IHD), and cardiovascular disease (CVD) by mean heart dose (MHD) level – above and below the group median. P-values generated from test for trend.

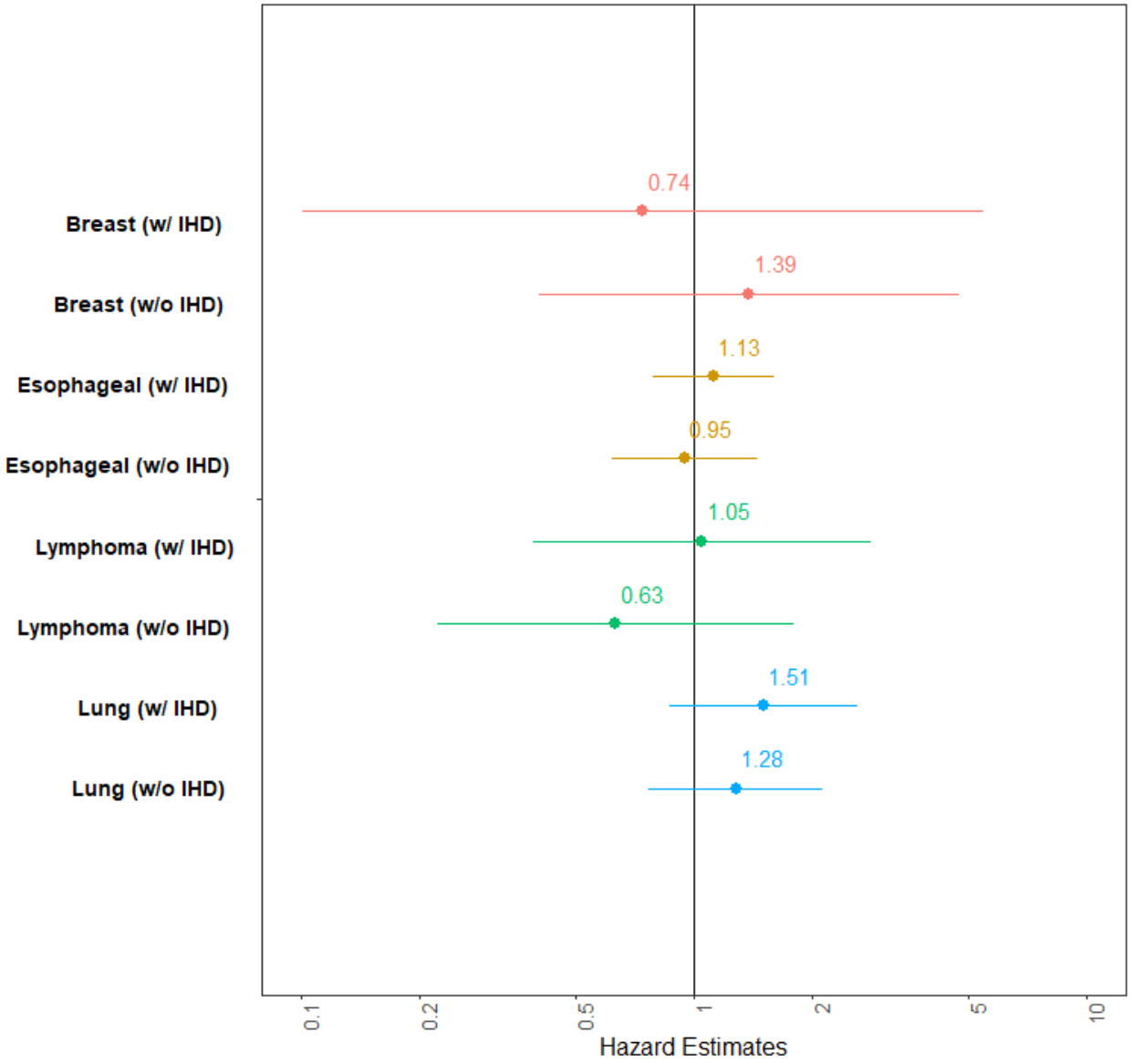


Figure A4. Forest plots hazard estimates with 95% CI for per Gy mean heart dose (MHD) in eight Cox proportion hazards models separated by four diagnoses and existing IHD status (with or without). A hazard estimate above one indicates an increased risk of subsequent IHD.

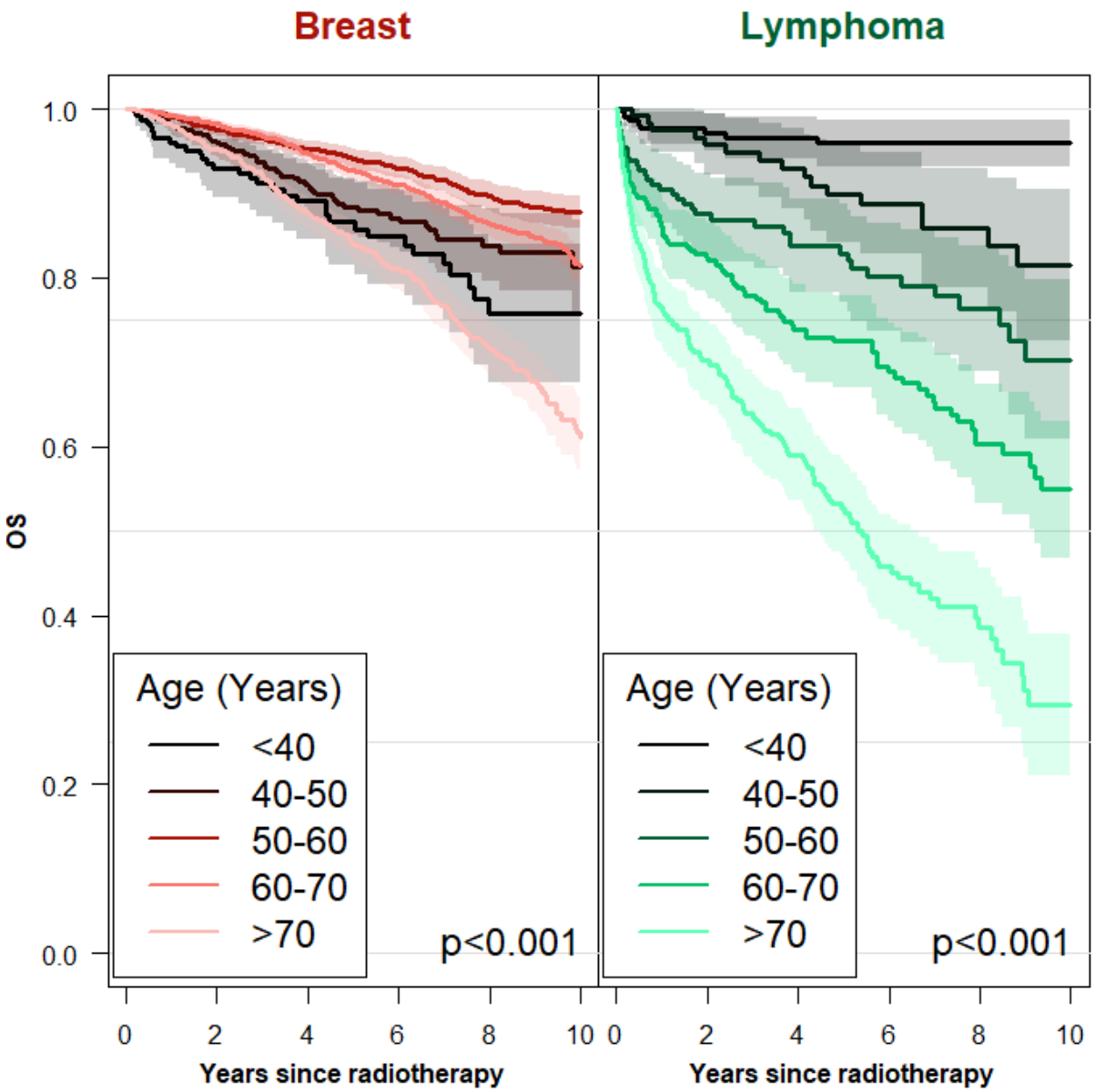


Figure A5. Kaplan-Meier plots for overall survival (OS) by detailed age group (<40, 40-50, 50-60, 60-70, and >70 years old) for breast and lymphoma. P-values generated from test for trend.

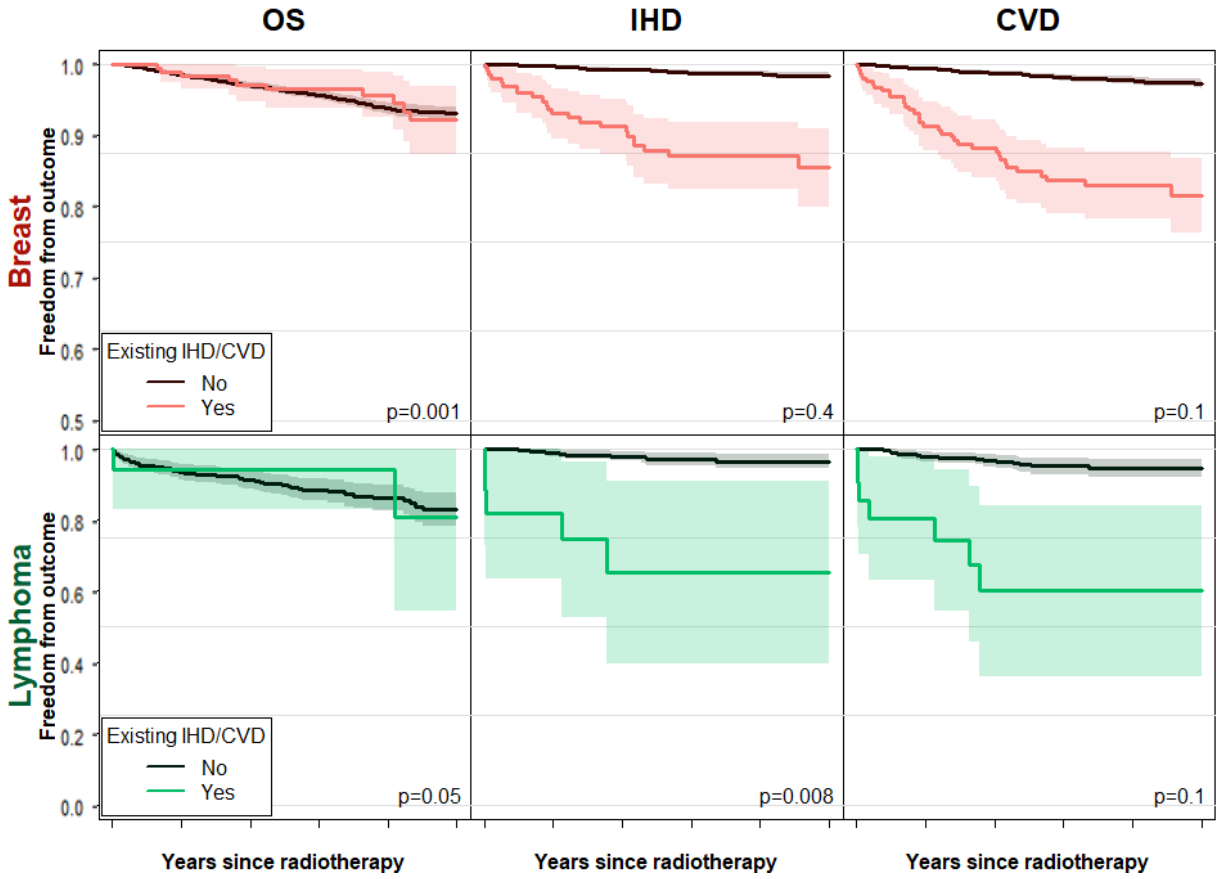


Figure A6. Secondary sub-analysis of youngest patients in the breast and lymphoma cohorts (<60 years old). Kaplan-Meier plots for overall survival (OS) and ischemic heart disease (IHD) by existing IHD as well as cardiovascular disease (CVD) by existing CVD. *Note that the Breast cancer y-axis is truncated at 0.5 to better show separation.* P-values generated from test for trend.

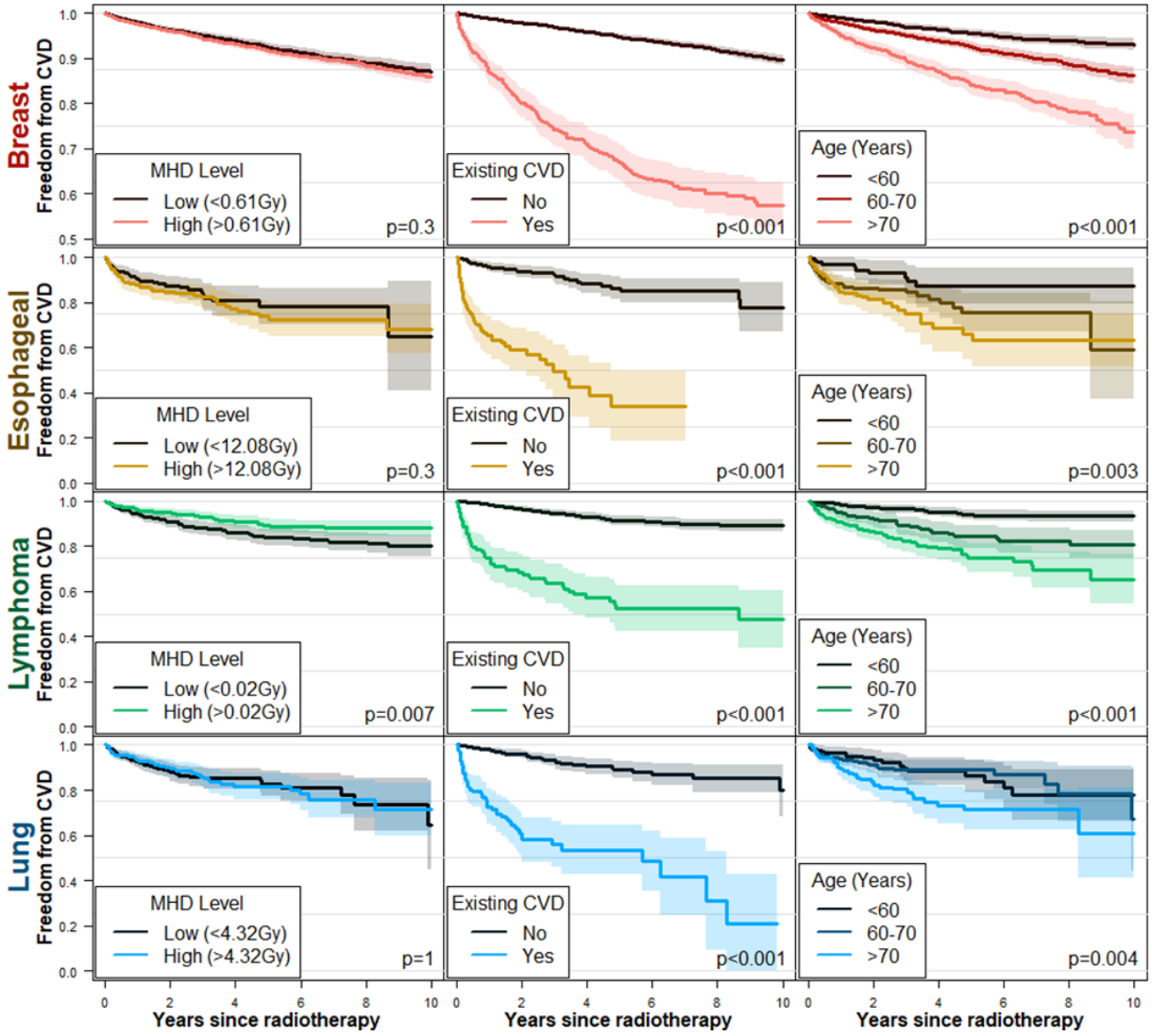


Figure A7. Kaplan-Meier plots for CVD, by mean heart dose (MHD) level (below or above median), existing ischemic heart disease (IHD), and age group (<60, 60-70, and >70 years old), separated by diagnosis. P-values generated from log-rank test (MHD level and existing IHD) and test for trend (age group). Note that the Breast cancer y-axis is truncated at 0.5 to better show separation. Numbers at risk available in supplementary Table A6. For the effect of sex on freedom from CVD see supplementary Figure A1.

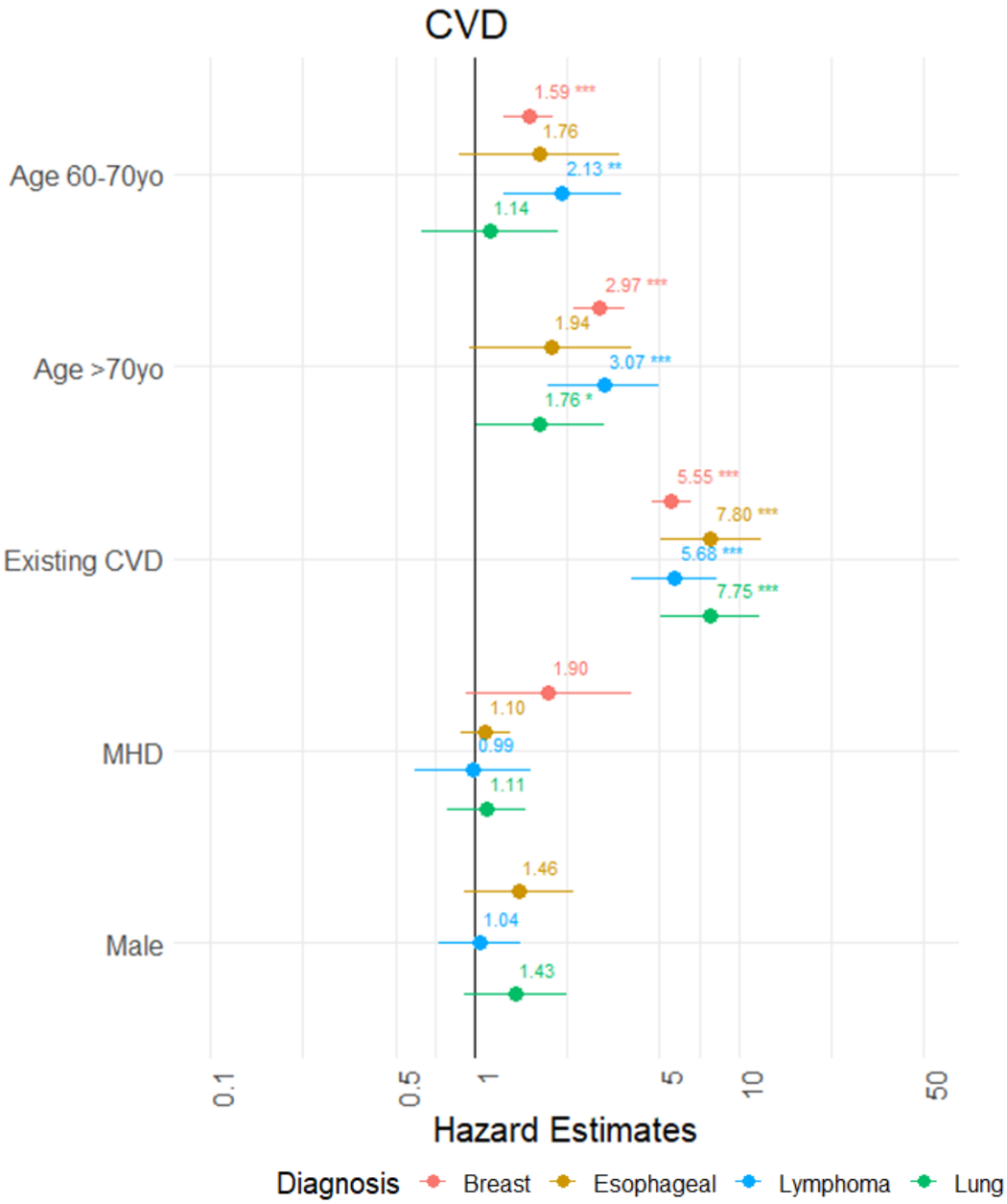


Figure A8. Forest plots of hazard estimates with 95% CI for predictors in Cox proportion hazards models on CVD separated by four diagnoses. The Breast cancer cohort only includes females. A hazard estimate above one indicates an increased risk of the outcome. P-value thresholds are 0.05(*), 0.01 (**), and 0.001(***).

Table A3. Kendall correlations between mean heart dose (MHD) and low (V5) and high (V30) volumetric dose measures, by diagnosis group.

	MHD X V5	MHD X V30	V5 X V30
Breast	0.775	0.636	0.745
Esophageal	0.684	0.833	0.661
Lung	0.812	0.808	0.703
Lymphoma	0.727	0.581	0.748

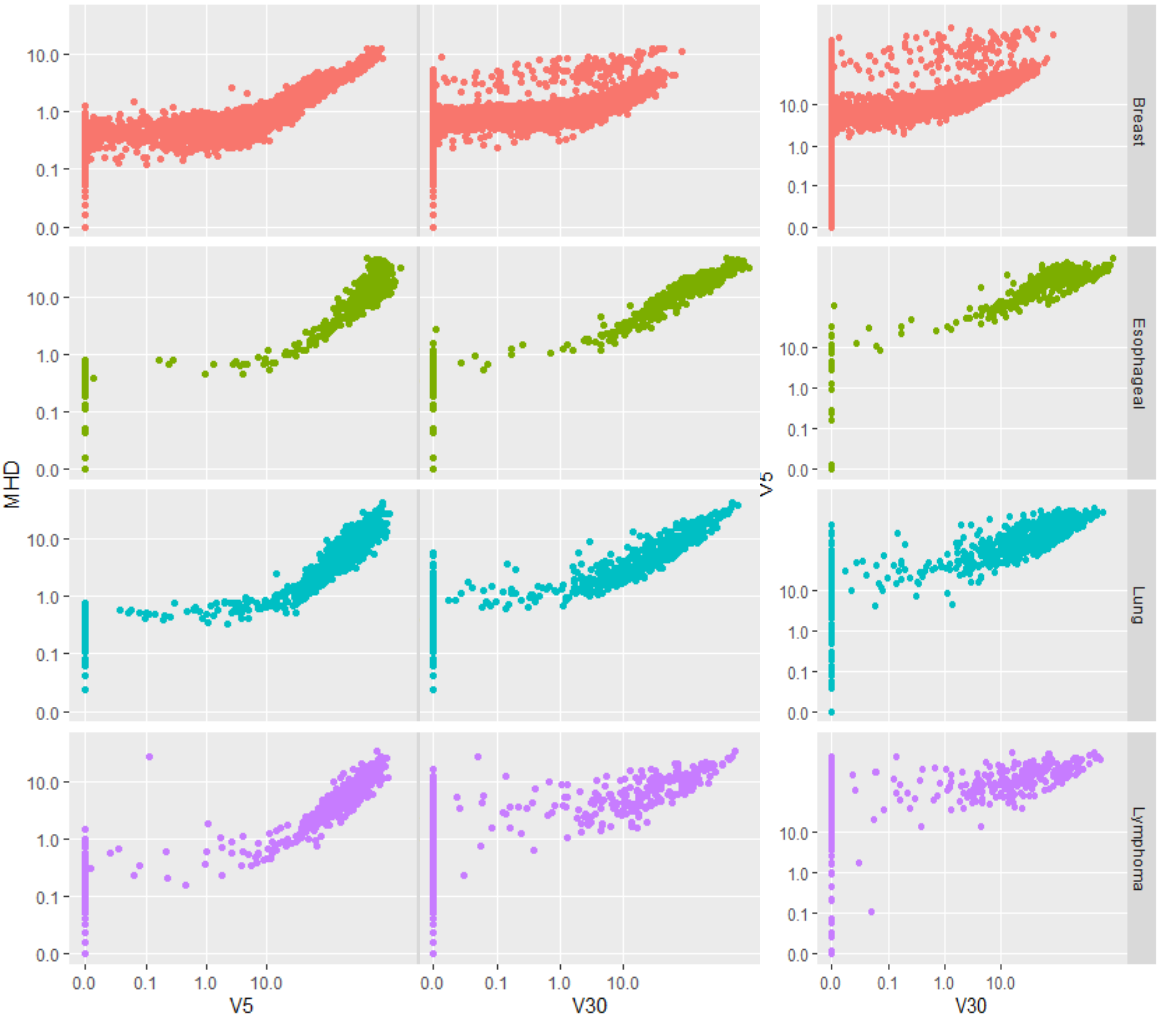


Figure A9. Scatter plot of mean heart dose (MHD) and low (V5) and high (V30) volumetric dose measures, by diagnosis group

Table A4. Numbers at risk for ischemic heart disease (IHD) shown in Figure 3

Dx	Subset	0 yrs	1 yr	2 yrs	3 yrs	4 yrs	5 yrs	6 yrs	7 yrs	8 yrs	9 yrs	10 yrs
Breast	**Total**	**6804**	**6216**	**5579**	**4994**	**4465**	**4003**	**3426**	**2727**	**2018**	**1484**	**929**
	MHD Level: Low	3394	3130	2830	2508	2237	1972	1673	1318	963	678	399
	MHD Level: High	3410	3086	2749	2486	2228	2031	1753	1409	1055	806	530
	Existing IHD: No	6298	5761	5187	4652	4168	3738	3207	2553	1897	1399	882
	Existing IHD: Yes	506	455	392	342	297	265	219	174	121	85	47
	Age: <60	2788	2491	2202	1947	1749	1558	1330	1069	805	611	382
	Age: 60-70	2638	2480	2296	2118	1911	1745	1530	1225	923	683	440
	Age: >70	1378	1245	1081	929	805	700	566	433	290	190	107
Esophageal	**Total**	**627**	**394**	**257**	**176**	**125**	**92**	**71**	**48**	**31**	**20**	**16**
	MHD Level: Low	313	199	123	76	45	33	25	15	8	5	5
	MHD Level: High	314	195	134	100	80	59	46	33	23	15	11
	Existing IHD: No	508	332	228	160	115	87	67	47	31	20	16
	Existing IHD: Yes	119	62	29	16	10	5	4	1	0	0	0
	Age: <60	136	89	65	49	37	27	22	16	13	8	8
	Age: 60-70	283	184	111	83	53	37	29	21	12	8	5
	Age: >70	208	121	81	44	35	28	20	11	6	4	3
Lymphoma	**Total**	**1251**	**949**	**823**	**706**	**605**	**493**	**399**	**297**	**224**	**157**	**88**
	MHD Level: Low	584	438	381	325	283	238	200	154	119	84	49
	MHD Level: High	667	511	442	381	322	255	199	143	105	73	39
	Existing IHD: No	1117	869	754	651	564	460	373	279	211	148	82
	Existing IHD: Yes	134	80	69	55	41	33	26	18	13	9	6
	Age: <60	552	472	419	376	331	280	240	179	139	99	56
	Age: 60-70	308	236	204	172	144	127	99	79	58	43	27
	Age: >70	391	241	200	158	130	86	60	39	27	15	5
Lung	**Total**	**729**	**500**	**340**	**233**	**168**	**125**	**82**	**54**	**35**	**26**	**14**
	MHD Level: Low	364	258	184	121	87	64	40	26	16	12	7
	MHD Level: High	365	242	156	112	81	61	42	28	19	14	7
	Existing IHD: No	605	429	300	209	153	114	74	50	31	24	14
	Existing IHD: Yes	124	71	40	24	15	11	8	4	4	2	0
	Age: <60	176	137	92	74	53	39	27	16	12	10	6
	Age: 60-70	295	199	140	92	68	55	35	24	15	12	5
	Age: >70	258	164	108	67	47	31	20	14	8	4	3

Table A5. Numbers at risk for overall survival (OS) shown in Figure 4

		0 yrs	1 yr	2 yrs	3 yrs	4 yrs	5 yrs	6 yrs	7 yrs	8 yrs	9 yrs	10 yrs
Breast	**Total**	**6804**	**6292**	**5698**	**5157**	**4638**	**4199**	**3624**	**2907**	**2160**	**1591**	**1005**
	MHD Level: Low	3394	3171	2894	2588	2318	2063	1761	1395	1019	723	435
	MHD Level: High	3410	3121	2804	2569	2320	2136	1863	1512	1141	868	570
	Existing IHD: No	6298	5797	5252	4750	4274	3860	3336	2674	1992	1473	937
	Existing IHD: Yes	506	495	446	407	364	339	288	233	168	118	68
	Age: <60	2788	2508	2228	1989	1789	1607	1381	1113	843	646	408
	Age: 60-70	2638	2509	2351	2180	1983	1823	1611	1302	983	730	477
	Age: >70	1378	1275	1119	988	866	769	632	492	334	215	120
Esophageal	**Total**	**627**	**418**	**281**	**199**	**145**	**104**	**79**	**53**	**35**	**23**	**18**
	MHD Level: Low	313	209	131	85	53	37	28	16	9	7	6
	MHD Level: High	314	209	150	114	92	67	51	37	26	16	12
	Existing IHD: No	508	338	236	168	121	90	70	47	31	21	17
	Existing IHD: Yes	119	80	45	31	24	14	9	6	4	2	1
	Age: <60	136	91	70	53	42	29	23	16	13	8	8
	Age: 60-70	283	195	121	92	63	44	35	25	16	11	7
	Age: >70	208	132	90	54	40	31	21	12	6	4	3
Lymphoma	**Total**	**1251**	**974**	**853**	**743**	**640**	**528**	**433**	**323**	**242**	**167**	**96**
	MHD Level: Low	584	451	399	347	303	261	223	170	130	91	55
	MHD Level: High	667	523	454	396	337	267	210	153	112	76	41
	Existing IHD: No	1117	877	769	671	582	480	392	295	222	153	87
	Existing IHD: Yes	134	97	84	72	58	48	41	28	20	14	9
	Age: <60	552	477	428	385	340	292	251	189	148	103	59
	Age: 60-70	308	242	212	185	157	138	110	88	62	46	29
	Age: >70	391	255	213	173	143	98	72	46	32	18	8
Lung	**Total**	**729**	**522**	**360**	**250**	**179**	**135**	**93**	**62**	**41**	**30**	**17**
	MHD Level: Low	364	269	194	127	92	69	44	26	17	12	8
	MHD Level: High	365	253	166	123	87	66	49	36	24	18	9
	Existing IHD: No	605	434	306	216	155	116	78	52	34	25	15
	Existing IHD: Yes	124	88	54	34	24	19	15	10	7	5	2
	Age: <60	176	141	95	78	55	42	31	19	14	12	7
	Age: 60-70	295	207	148	98	73	59	39	28	18	13	7
	Age: >70	258	174	117	74	51	34	23	15	9	5	3

Table A6. Numbers at risk for cardiovascular disease (CVD) shown in Figure A7.

		0 yrs	1 yr	2 yrs	3 yrs	4 yrs	5 yrs	6 yrs	7 yrs	8 yrs	9 yrs	10 yrs
Breast	**Total**	**6804**	**6216**	**5579**	**4994**	**4465**	**4003**	**3426**	**2727**	**2018**	**1484**	**929**
	MHD Level: Low	3394	3103	2789	2464	2190	1916	1627	1275	926	653	385
	MHD Level: High	3410	3044	2697	2434	2175	1968	1694	1363	1014	775	511
	Existing CVD: No	6186	5621	5045	4523	4043	3601	3091	2455	1813	1339	849
	Existing CVD: Yes	618	526	441	375	322	283	230	183	127	89	47
	Age: <60	2788	2479	2183	1926	1726	1533	1306	1046	787	599	376
	Age: 60-70	2638	2457	2262	2081	1871	1696	1479	1182	884	652	420
	Age: >70	1378	1211	1041	891	768	655	536	410	269	177	100
Esophageal	**Total**	**627**	**394**	**257**	**176**	**125**	**92**	**71**	**48**	**31**	**20**	**16**
	MHD Level: Low	313	190	117	71	41	30	23	14	8	5	5
	MHD Level: High	314	186	129	96	74	56	44	33	23	14	11
	Existing CVD: No	491	312	214	150	104	80	63	46	31	19	16
	Existing CVD: Yes	136	64	32	17	11	6	4	1	0	0	0
	Age: <60	136	88	65	49	37	27	22	16	13	8	8
	Age: 60-70	283	176	104	76	48	35	27	20	12	7	5
	Age: >70	208	112	77	42	30	24	18	11	6	4	3
Lymphoma	**Total**	**1251**	**949**	**823**	**706**	**605**	**493**	**399**	**297**	**224**	**157**	**88**
	MHD Level: Low	584	427	365	309	268	227	193	149	114	80	48
	MHD Level: High	667	504	433	371	308	241	188	138	101	72	39
	Existing CVD: No	1085	843	727	623	534	437	357	268	202	142	80
	Existing CVD: Yes	166	88	71	57	42	31	24	19	13	10	7
	Age: <60	552	469	414	371	324	273	235	176	136	98	55
	Age: 60-70	308	230	197	164	133	117	91	75	55	42	27
	Age: >70	391	232	187	145	119	78	55	36	24	12	5
Lung	**Total**	**729**	**500**	**340**	**233**	**168**	**125**	**82**	**54**	**35**	**26**	**14**
	MHD Level: Low	364	251	176	115	85	62	40	26	15	12	7
	MHD Level: High	365	240	153	112	79	59	41	27	19	14	7
	Existing CVD: No	590	418	289	202	149	110	72	49	31	24	14
	Existing CVD: Yes	139	73	40	25	15	11	9	4	3	2	0
	Age: <60	176	136	89	72	53	39	27	16	12	10	6
	Age: 60-70	295	195	137	90	67	54	35	24	14	12	5
	Age: >70	258	160	103	65	44	28	19	13	8	4	3

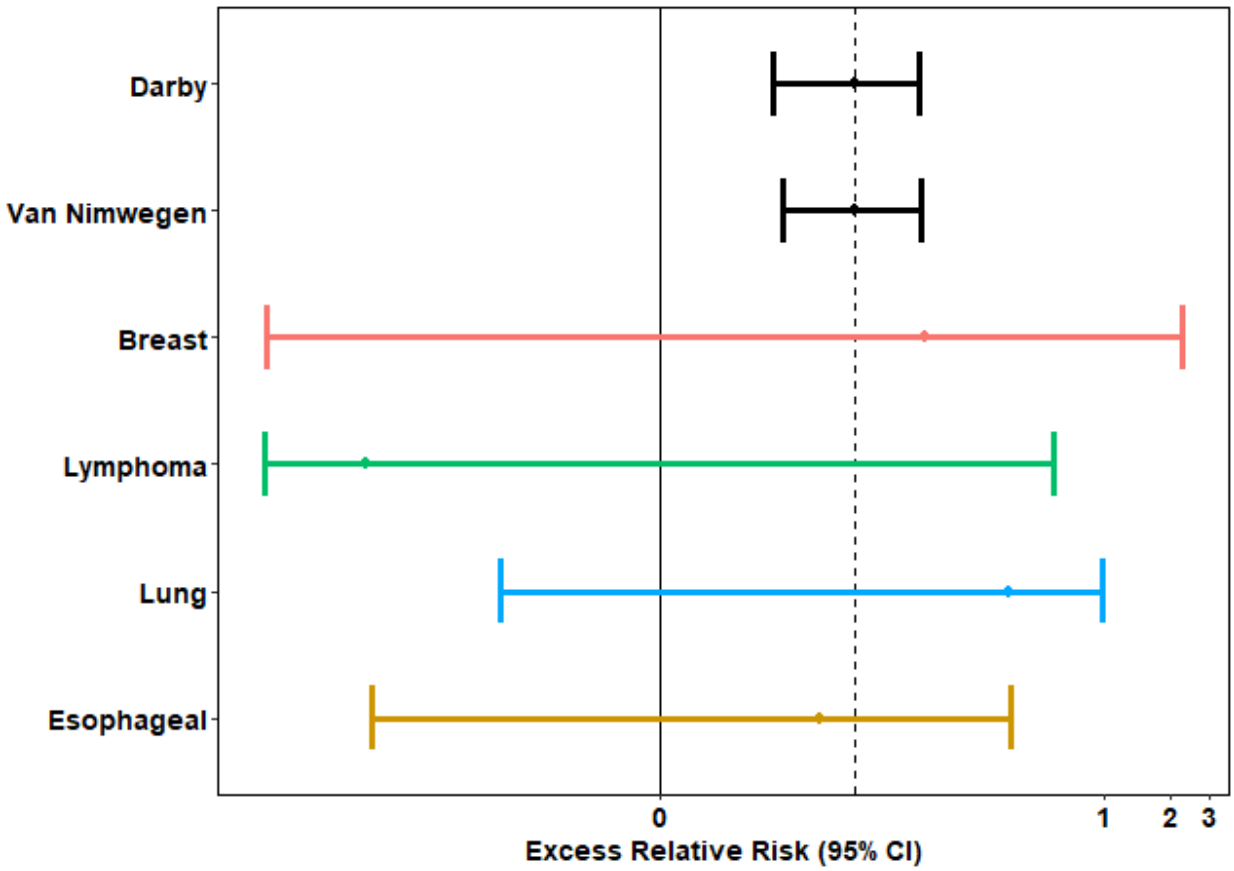


Figure A10. Converted excess relative risk (and 95% confidence intervals) of per Gy mean heart dose by diagnosis and compared with existing literature (Darby and Van Nimwegen) [9,10]. The x-axis is displayed on a pseudo logarithmic scale.

Ørestads Boulevard 5, Bygning 37 K, st.

2300 København S

København Ø, den 5.7.2021

Anmeldelse til de Videnskabsetiske Komitéer

Alle oplysninger på denne blanket kan blive offentliggjort.

Komité

Primærkomité:	National Videnskabetisk Komité
Sekundærkomitéer:	De Videnskabsetiske Komiteer for Region Hovedstaden
Anmeldelsesnr.:	82427

A. Forsøgsansvarlig

1. Titel:	Professor
2. Navn:	Ivan Richter Vogelius
3. Hospital/institution:	Rigshospitalet
4. Afdeling/institut	Kræftbehandling
5. Vejnavn og nr.	Blegdamsvej 9
6. Postnummer/7. by:	2100 København Ø
8. Telefonnr.:	35459885
9. E-mail:	ivan.richter.vogelius@regionh.dk

B. Evt. anden kontaktperson

1. Titel:	PhD student
2. Navn:	Nora J. Forbes
3. Hospital/institution:	Rigshospitalet
4. Afdeling/institut	Kræftbehandling
5. Vejnavn og nr.	Belgdamsvej 9
6. Postnummer/7. by:	2100 København Ø
8. Telefonnr.:	
9. E-mail:	nora.jarrett.forbes@regionh.dk

C. Projektinformation

1. Projekttitel:	Radiotherapy exposure and association with observed cardiovascular toxicity in patients treated for cancer at Rigshospitalet
2. Projektets hovedformål:	Forsøgets overordnede formål er at forbedre vores evne til at forudsige risikoen for hjerte-kar bivirkninger fra en stråleterapi planlægning. Dette vil gøre os bedre i stand til at vælge den mest optimale strålebehandlingsstrategi for fremtidige patienter. Moderne stråleterapiudstyr giver os stærke muligheder for at forme dosis meget præcist. Uanset præcisionen af den afgivne dosis er man dog nødt til at "få dosis ind til svulsten", hvilket betyder medbestråling af rask væv i strålens vej. Forsøget vil gøre

	os bedre i stand til at prioritere de bedste veje ind til svulsten, der giver lavest risiko for bivirkninger.
3. Sted(er) for projektets gennemførelse:	Hospitalsafdeling Sygehusklassifikationer: De Videnskabsetiske Komiteer for Region Hovedstaden: Rigshospitalet
- Adresse(r) / forsøgsansvarlige	Afdeling for kræftbehandling Bledamsvej 9 2100 København Ø
4. Forsøgsgrupper:	Myndige habile
5. Design:	Åben
6. Forskning i biologisk materiale:	Ingen
7. Lægemiddelforsøg:	Nej
8. Medicinsk udstyr:	Nej
9. Projektøkonomi:	Fondsstøtte
10. Udbetales der vederlag til forsøgspersoner:	Nej
11. Projekt iværksættes den:	01-09-2021
12. Projekt afsluttes den:	01-09-2030
13. Forventet antal forsøgspersoner i DK:	44000
14. Heraf raske:	0
15. Læge/sundhedsfagligt område:	Onkologi
16. Sygdommens art/navn:	Kræftsygdomme med indikation for strålebehandling
17. ICD10 kode:	C00-D49
18. Multistat projekt:	Nej

_____

Underskrift

Professor Ivan Richter Vogelius
Rigshospitalet
Kræftbehandling
Blegdamsvej 9
2100 København Ø

Protokolresume version 1.1 Juli 23 2021. Anmeldelsesnummer 82427

1: Projektets originale titel

Radiotherapy exposure and association with cardiovascular toxicity in patients treated for cancer at Rigshospitalet

Dansk titel:

Stråledosis og sammenhæng med observeret frekvens af hjerte-kar-bivirkninger i patienter behandlet for kræftsygdomme på Rigshospitalet

2: Forsøgsansvarliges navn og forsøgssted

Forsøgsansvarlig er Professor, PhD, Dr. Med. Ivan Richter Vogelius, ansat på Rigshospitalets Onkologiske Klinik. Adressen for forsøget er Onkologisk Klinik, Rigshospitalet, Blegdamsvej 9, 2100 København Ø

3: Hvilke typer af data og hvorfra de stammer

Billedata og stråledoser
Data for stråledosis og registrering af hvilke område der er blevet bestrålet med en givet dosis haves fra de rutinemæssige scanninger taget i forbindelse med planlægning af strålebehandlingen. Dette involverer CT og somme tider også MR og PET scanninger. Disse scanninger og data om stråledosis haves fra planlægningen af strålebehandlingen inden første behandling.

Data for den daglige dosis haves fra scanninger taget på behandlingsapparaterne under behandling, typisk hver dag i 4-6 uger. Disse er som oftest CT scanninger, men kan også være 2D røntgenbilleder og i senere år MR scanninger for de relativt få patienter, der bliver behandlet på afdelingens MR accelerator

Data for opfølgende billeddiagnostik haves fra de rutinemæssige scanninger taget i patientens opfølgningsprogram.

Bivirkningsdata
Registrerede bivirkningsdata vil komme fra registerkoder fra landspatientregistret samt overlevelseskoder fra CPR registret. I enkelte tilfælde kan disse data blive suppleret med eller kontrolleret ved opslag i patientjournal.

Protokolresume version 1.1 Juli 23 2021. Anmeldelsesnummer 82427

4 Forsøgets formål

Forsøgets overordnede formål er at forbedre vores evne til at forudsige risikoen for hjerte-kar bivirkninger fra en stråleterapi planlægning. Dette vil gøre os bedre i stand til at vælge den mest optimale strålebehandlingsstrategi for fremtidige patienter.
Moderne stråleterapiudstyr giver os stærke muligheder for at forme dosis meget præcist. Uanset præcisionen af den afgivne dosis er man dog nødt til at "få dosis ind til svulsten", hvilket betyder medbestråling af rask væv i strålens vej. Forsøget vil gøre os bedre i stand til at prioritere de bedste veje ind til svulsten, der giver lavest risiko for bivirkninger.

5 Forsøgets metode

I forsøget søger vi at sammenholde dosis til normalt væv med risikoen for bivirkninger. Dosis til patienten registreres ved analyse af scanningerne brugt til planlægning af behandlingen - disse scanninger og den tilgængelige dosisplan giver os fordelingen af stråledosis i kroppen. Bestråling af hjerte og de store kar registreres ved manuel eller automatiseret analyse af disse data.

Dosis til de enkelte områder i patienten sammenholdes med forekomsten af bivirkninger i et stort antal patienter for at give os mulighed for at observere sjældne, men alvorlige bivirkninger.

Statistiske metoder vil inkludere "traditionelle" metoder (logistisk regression, Cox proportional hazards modeller mv), men også data drevne metoder, herunder metoder baseret på kunstig intelligens.

De oplysninger, der skal bruges i projektet videregives til forsker. Forsker har således ikke direkte adgang til journaloplysninger.

6 forsøgspersoner, herunder inklusions- og eksklusionskriterier

Forsøget vil som udgangspunkt inkludere alle patienter behandlet med stråleterapi på Rigshospitalet i perioden **2009-2020** med forventet forlængelse hvis forsøget er succesfuldt.

7 Økonomiske forhold

Studiet er igangsat af forskere og klinisk personale på Rigshospitalet. Økonomisk støtte søges løbende og haves i skrivende stund fra to bevillinger fra Kræftens Bekæmpelse samt en bevilling fra den kommercielle partner i radioterapi-udstyr (Varian Medical Systems, Palo Alto, USA – i skrivende stund under opkøb af Siemens Healthineers, Erlangen, Tyskland). Bevillingerne håndteres som forskningsbevillinger underlagt revision i Rigshospitalets system. Bevillingerne udbetales aldrig personligt til forskere, men går primært til finansiering af lønudgifter via de involverede forskeres institutioner. Et mindre beløb går til udstyr (fx. højtydende computere) samt rejse- og konferenceaktivitet.

8 Offentliggørelse af forsøgsresultater

Forsøgets resultater vil løbende blive publiceret i internationale faglige tidsskrifter og indsendt til faglige konferencer. Det vil ikke være muligt for tilskudsgivere at blokere for offentliggørelse og resultaterne vil blive offentliggjort uanset om de er “positive” eller “negative” eller ”inkonklusive” for såvel forskere som tilskudsgivere.

Protocol version 1.2 August 23 2021. Anmeldelsesnummer: 82427

1: Title

Radiotherapy exposure and association with observed cardiovascular toxicity inpatients treated for cancer at Rigshospitalet

2: Data origin

The data used in this study is imaging data acquired as part of routine treatment planning and clinical care of patients diagnosed with cancer and with indication for radiotherapy and acceptance of this offered treatment at Rigshospitalet between **Jan 2009- Dec 2020**. Three levels of data sources will be used in the study:

2.1 Treatment planning imaging and radiotherapy dose exposure

Treatment planning imaging data are acquired as part of the clinical routine before radiotherapy and include CT scans and possibly MR and PET scans as requested by the treatingphysician in the individual case. Dose exposure is calculated as part of the treatment planning process and stored as a 3D matrix in the same coordinate system as the treatment planning images. This data is stored for the lifetime of the patient in the clinical systems since the knowledge of prior exposure is critical to safely deliver a possible second (or later) exposure to a patient on clinical indication.

2.2 Treatment verification imaging

Treatment verification imaging is acquired during radiotherapy to verify accuracy of thedaily treatment fraction as part of the clinical routine. This will involve 3D Cone-beam CT imaging, 2D portal imaging or 3D MR scan depending on the treatment machine. Consent to these imaging procedures at the time of treatment are an integral part of the consent to receive radiotherapy.

2.3 Follow-up imaging

Follow-up imaging by CT MR or PET as acquired as part of the routine clinical follow up program and requested by the treating physician. In the clinical routine consent for this procedure is taken in connection with the follow-up visit, most likely verbal.

2.4 Outcome data

Outcome data will originate from electronic patient records and the Danish registries as described in detail below in section 8 and 9

Data will be provided to the researcher, such that the researcher does not have direct access to the patient records.

3: Research

We will publish the results in international peer reviewed publications and new protocols for prospective testing of the results produced under this project will be implemented as separate prospective clinical investigations. Reproducibility and generalizability are core discussion of all our research and we also expect methodological publications including a "levels of evidence" ladder for risk models together with international collaborators as a spinoff (without direct access to data).

4: Purpose of the research

The purpose of the research is to estimate the association between radiotherapy exposure and cardiovascular toxicity after radiotherapy. In particular the aim is to establish quantitative associations between normal tissue exposure and incidence/prevalence of side effects. Side effects can be estimated by

a. Manual assessment of patient charts
b. Coupling with registry data with vital status, procedural codes or laboratory values (e.g. ICD10 codes or laboratory test results excluding genetic sequencing)

The hypothesis is that there is an association between radiotherapy exposure and the risk of cardiovascular sideeffects. We furthermore seek to quantify the association between radiotherapy exposure and theprobability of the above-mentioned endpoints of toxicity.

The rationale for the project is to improve our understanding of such associations, which will in turn allow us to better tailor our radiotherapy to minimize the risk of the observed cardiovascular toxicity while maintaining or improving the probability of durable disease control.

Literature
The importance of the research question is documented through the impact of the 2010 special issue of the primary radiotherapy research journal *Int. J. Radiation Oncology Biol. Phys.* focused on Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC)(1-3). The effect of radiation exposure on cardiovascular disease has been studied in several publications, across many patient diagnoses and with varying definitions of cardiovascular disease (4). For example, in Breast cancer the rate of major coronary events increased linearly by 7.4% per gray of mean radiation dose to the heart (5).

An effort to update the 2010 QUANTEC approach has been initiated with a *Beyond Quantec* meeting as initiator held in Maryland Feb 2020 with the research responsible for the current project (Ivan Richter Vogelius) as member of steering committee.

5 Method

This is a non-randomized observational study with a large sample size to estimate radiation dose-effect relationships. All patients diagnosed with cancer who receive radiotherapyat Rigshospitalet between 2009-2020 will be eligible for inclusion in the study. This is an observational study and thus will use only data collected a part of routine treatment and clinical care.

Detailed statistical analysis plan will be made prior to anyanalysis being conducted.

5.1 Assessment of organ specific radiation exposure

Organ-specific (i.e. heart or substructures thereof) exposure will be assessed through delineation of structures using a combination of manual annotation of the structures on the CT scan and interactive deep learning models to identify the same structures. Manual delineations of organs at risk are time consuming and is thus effectively limited to populations in the <200 patient range. Interactive deep learning annotation is available and is expected to be used when data are in the 100-1500 patient range, see (6) for a description of a 2D interactive annotation algorithm which will be expanded to a 3Dalgorithm for medical images. Finally, fully automated organ annotation will be applied in patient cohorts significantly exceeding 1000 patients and will generally be based on convolutional neural nets architectures optimized for the data in question, see (7-9) for details.

5.2 Correlation with outcome with traditional statistics

Correlation between cardiovascular toxicity and important clinical outcome data (i.e. overall survival and disease control) will be assessed using traditional statistical methods such as multivariatelogistic regression, Cox proportional hazards modeling or the Fine and Gray methods and derivatives for competing risk assessments(10,11).

For such methods, the covariables to include will be defined in detail depending on the baseline rates when data has been extracted and after a consideration of the plausible cause-effects and collinearities in the data in a review session with both statisticians and medical doctors. A statistical analysis plan will be made accordingly.

However, we will stratify according to the index cancer treatment and to exemplify we expect to use patient age, gender, comorbidity status as covariables of interest together with a measure of hematological toxicity. In addition, the exposure to the organs delineated above will be included after appropriate considerations of collinearities in the data. Further covariables will depend on the above-mentioned consensus between statisticians and medical doctors when reviewing baseline data.

Classic statistical survivapproaches to survival analysis often require strict assumptions, such as proportional hazards, and have limitations when dealing with missingness, heterogeneous cohorts, and high dimensionality (12). Therefore, we will also develop machine learning approaches, which have advantages in modeling complex survival data (. This will include application of random survival forests, a non-parametric ensemble approach that can robustly

handle non-linear effects and multiple variable interactions (13). Additionally, we will explore Cox-based penalized regression models (i.e. LASSO, Ridge, and Elastic Net), which introduce regularization and improve variable selection in high dimensional data thereby providing a supplement to the consensus selection of predictor variables above (14). We will also explore deep survival neural networks with the same objective (15). Regardless of methods, the structures from 5.1 will be included as a predicting variable together with patient baseline data.

5.3 Inverse learning methods

As an exploratory part of the study, we will attempt reversing the association process by conducting machine learning analysis *without* the prior annotation of scans as described in 5.1. Such methods are exploratory and in development, with the goal to identify substructures that have an increased influence on toxicity prediction (16). For example, Ibragimov et al. has recently used CNNs, supplemented with transfer learning, on CT images and RT dose plan data to identify radiosensitive regions of the liver which should be spared to reduce toxicity (17).

In our case, input to the methods will be a subset of treatment planning images, radiotherapy dose exposure, treatment verification images, and possibly patient related factors as in 5.2 and follow-up images. Output will be hematological toxicity, overall survival and disease control and possibly hospitalizations suspected to be associated with hematological toxicity. The goal is to be able to predict patient and population-level saliency maps, a 3D matrix of probabilistic values indicating important regions within the image, for a given outcome. Methods will be based on convolution and/or transformer neural networks (18), in addition to exploring techniques based on large-scale hypothesis testing (19). To allow learning from limited amounts of outcome data, techniques such as self-supervision may be used to build useful representations (20).

6 Statistical considerations

This is a large-sample non-randomized study, so the primary concern is bias(7). With approximately 4000 patients treated per year over a 11-year period of inclusion, sample size will only be an issue in rare outcomes or sub-indications.

7 Study subjects

a) Inclusion criteria
    i) All patients referred for radiotherapy at Rigshospitalet between 2009 and 2020
b) Exclusion criteria
    i) Explicit requests not to allow use of medical data for research as reported by treatment staff when implemented
        1) A procedure to register a denial of use is currently under implementation with inspiration from *vævsanvendelsesregistret* but only covering radiotherapy at Rigshospitalet.

Up to 4,000 patients per year will be included starting in 2009 and continuing to 2020. This yields a total number of 44,000 patients.

8 Secondary findings

It should be stressed that the risk of secondary findings is very small: The image workup of patients treated with radiotherapy for cancer at Rigshospitalet involves both the image assessment performed during diagnosis but also the subsequent treatment planning image assessment. Both of these steps in the clinical routine involve specialized and time-consuming image assessment by experienced radiologists, nuclear medicine physicians and oncologists in collaboration and subsequent image annotation by experienced radiographers. The probability of a severe finding passing undetected at the original treatment stage and being detected by a researcher at the analysis stage is deemed highly unlikely to occur even once despite the relatively large data set studied here.

In the case that secondary findings do occur, a committee will be settled consisting of the responsible physician for the radiotherapy department (at the time of writing Jeppe Friborg, MD) and head of department (At time of writing prof. Ulrik Lassen). The two will be responsible for establishing a committee adapted to the clinical case, which will consist of Jeppe Friborg, Ulrik Lassen and Anne Kiil Berthelsen. The two first are oncologists and the last is radiologist. All specialists and senior consultants "Overlæge".

Reporting of secondary findings to the participant can occur if it is certain or very likely that the finding is a severe or life-threatening condition, which

- Can be prevented or treated with aim of cure or symptom relief
- The disease has significant impact on the patient and
- The secondary finding can be validated clinically and
- The method used to assess the secondary finding is considered certain

9 Patient chart information

Manual review of patient charts may be required in individual cases to resolve incongruency or missingness in data sources. This may involve registration of precise disease type (i.e. diagnosis, stage, and intent to treat), patient comorbidities (i.e. BMI, smoking status, history of COPD, history of diabetes) and related parameters (i.e. medication) and at the time of treatment.See below for registry assessment.

Data extracted from the patient charts from dates prior to the date of treatment for the cancer diagnosis in question will consist of the items corresponding to a Charlson comorbidity index assessment, see (21) for details. The patient chart information prior to treatment is necessary to control for confounders associated with patient comorbidity and stage and extend of disease.

The data will be forwarded to the researcher through electronic registration without direct chart access.

10 Access to registries

Access to the registries will be used to obtain data necessary for analysis. From the CPR registry we will extract birth date and vital status/mors date. From Landspatientregister (and automated extraction from electronic patient records at hospital level) we will extract data on prior disease history corresponding to an established comorbidity scoring (by procedural or ICD10 codes).

- We have typically used the Charlson comorbidity index.
- Unfortunately, the Charlson index has shown limitations in our prior research. We will therefore also use other scoring systems. Such variation will rely on the same conditions to be extracted from landspatientregistret as the Charlson index, but add data for a severity of each event.

The registry coupling will be performed through an already established protocol at PERSIMUNE research center at Rigshospitalet with appropriate approvals. PERSIMUNE is the Danish National Research Foundation Centre of Excellence for Personalized Medicine of Infectious Complications in Immune Deficiency. The data warehouse contains prospective and historic data derived from patients treated at clinical departments at Rigshospitalet including data from para-clinical departments, laboratories and imaging. Full details can be found at www.persimune.dk .

Data extracted from the registries up to the date of treatment are corresponding to the Charlson comorbidity index assessment as in the case of chart review (21) although the patient registry datais less detailed they will often suffice (22) and thus be preferred to manual review when deemed of sufficient quality.

Data extracted from registries *after* the date of treatment of the index cancer will be procedural codes for the toxicity under analysis and date of occurrence of such toxicity. For example, the following ICD10 codes are of interest for the analysis of the risk of cardiac toxicity: I00-I50.

Only “dry” data will be accessed meaning that no new analyses of biological data will be made for this project.

Only data relevant for assessment of treatment related cardiovascular toxicity will be extracted for the current project. Procedural codes of relevance to the research question will be identified prior to data extraction and only those data will be accessed. Data will be pseudoanonymized.

Information prior to treatment is necessary to control for confounders associated with patient comorbidity and stage and extend of disease. Date acquired from later dates than the date of treatment of the index cancer is necessary for scoring the toxicity outcome.

The data will be forwarded to the researcher through electronic registration without direct chart access.

11 General data protection rules

General Data Protection Rules will be respected and we will adhere to the Danish "Databeskyttelsesforordningen" and the law of "databeskyttelsesloven". Approvals are already in place for the coupling of radiotherapy data from Rigshospitalet to PERSIMUNE. Investigator is responsible for adhering to these regulations.

12 Economy

The project is investigator-initiated by researchers and clinical staff at Rigshospitalet.
The project receives financial support from a number of foundations with continuous efforts to secure further funding. At present funding bodies include:

Danish Cancer Society
Evidensbaseret dosisplanlægning af strålebehandling af lymfom under hensyntagen til alle potentielle senfølger og individuelle patientfaktorer. Principal investigator: Professor, dr.med. **Lena Specht**. Supported in 2021 - 2023 by 2.325.000 kr.

Brug af registerdata til at modellere sammenhæng mellem dosis og effekt af strålebehandling.
Principal investigator: Ivan Richter Vogelius. Supported in 2020 - 2022 with 1.600.000 kr.

Support from Danish Comprehensive Cancer Center for PhD salary: 400.000 DKK

Varian Medical systems
Machine learning for detailed normal tissue exposure assessment, longitudinal imaging and associated damage of normal tissue substructures. Principal Investigator: Ivan Vogelius. Supported by 835.000 DKK

Computer Science, Copenhagen University
PhD salary support for typically one of the three years, currently 900.000 DKK in value.

12.c Use of economical support
The funding from all grants is used predominantly for salary support to investigators and PhD candidates with a minor fraction (currently less than 100.000 DKK) spent on supportive equipment and also a fraction of the budget for conference participation and associated costs.

Economical support will always follow Rigshospitalet's rules for research funding and consequently be paid to accounts at Rigshospitalet and never directly to researchers. The management of the grants are subject to review by Rigshospitalet's accounting.

The investigators have no financial interests in the research. Varian Medical Systems is an industrial partner with some financial interest in the success of the project. The collaboration with Varian is regulated by Rigshospitalet's legal department and none of the funding bodies can prevent publication of negative results. We collaborate with Varian Medical Systems, but Varian Medical systems does not have access to personally identifiable data or other privileged access to data based on this protocol

13 Publication of results

Study results will be published in reputable international medical journals whether they are deemed "positive" or "negative" or "inconclusive". Similarly, results will be presented at international scientific meetings.

14 Research ethics

14.1 Protection of the participant's right to privacy and integrity
All regulations related to general data protection rules will be followed. The integration with PERSIMUNE has all relevant approvals for the secure communication and de-identification within the database.

Access to data will always follow an approach where the minimum identifiable information is used or stored on local hard-drives, also including cache.

14.2 Burden to the participant

As described above, it is highly unlikely that a secondary finding is identified and, consequently, it is highly unlikely that there will be any excess burden on the participants.

14.3 Justification of project in terms of potential to improve therapy to future

Problem addressed
Modern radiotherapy equipment represents investments of several hundred mio DKK at Rigshospitalet alone, at it provides us with the opportunity to "sculpt" the dose distribution in patients with very high precision. The current biggest barrier to better use of these investments are the gaps in knowledge about the biological impact of dose for normal tissue and tumor exposure. This is what we aim to improve with the current effort.

Value of project to future patients
The project will facilitate more gentle radiotherapy for future patients by providing more detailed information of regions to avoid during radiotherapy planning and delivery and by providing more solid bases for clinical decision making when comparing therapeutic options. This includes the generation of more robust models for shared decision making in difficult clinical cases where risk of toxicity and risk of recurrence are challenging to manage.

Furthermore, the results will potentially allow a new paradigm of radiotherapy optimization based on robust risk models, see (23) for a conceptual example limited by the access to reliablerisk models.

Considering the minimal burden on patients we therefor believe that the project is justified ethically.

References

1. Bentzen SM, Constine LS, Deasy JO, et al. Quantitative analyses of normal tissue effects in theclinic (quantec): An introduction to the scientific issues. *Int J Radiat Oncol Biol Phys* 2010;76:S3-9.
2. Marks LB, Yorke ED, Jackson A, et al. Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys* 2010;76:S10-9.
3. Marks LB, Ten Haken RK, Martel MK. Guest editor's introduction to quantec: A users guide. *Int J Radiat Oncol Biol Phys* 2010;76:S1-2.
4. Niska JR, Thorpe CS, Allen SM, Daniels TB, Rule WG, Schild SE, Vargas CE, Mookadam F. Radiation and the heart: systematic review of dosimetry and cardiac endpoints. Expert Rev Cardiovasc Ther. 2018 Dec;16(12):931-950.
5. Darby SC, Ewertz M, McGale P, Bennet AM, Blom-Goldman U, Brønnum D, Correa C, Cutter D, Gagliardi G, Gigante B, Jensen MB, Nisbet A, Peto R, Rahimi K, Taylor C, Hall P. Risk of ischemic heart disease in women after radiotherapy for breast cancer. N Engl J Med. 2013 Mar 14;368(11):987-98.
6. Smith AG, Han E, Petersen J, et al. Rootpainter: Deep learning segmentation of biological images with corrective annotation. *bioRxiv* 2020:2020.04.16.044461.
7. Vogelius IR, Petersen J, Bentzen SM. Harnessing data science to advance radiation oncology. *Mol Oncol* 2020;14:1514-1528.
8. Francolini G, Desideri I, Stocchi G, et al. Artificial intelligence in radiotherapy: State of the

artand future directions. *Med Oncol* 2020;37:50.
9. Cardenas CE, Yang J, Anderson BM, et al. Advances in auto-segmentation. *Semin Radiat Oncol* 2019;29:185-197.
10. Chappell R. Competing risk analyses: How are they different and why should you care? *Clin Cancer Res* 2012;18:2127-9.
11. Dignam JJ, Zhang Q, Kocherginsky M. The use and interpretation of competing risks regression models. *Clin Cancer Res* 2012;18:2301-8.
12. Spooner, Annette, et al. "A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction." Scientific reports 10.1 (2020): 1-10.
13. Ishwaran, Hemant, et al. "Random survival forests." The annals of applied statistics 2.3 (2008): 841-860.
14. Lai, Yeuntyng, Morihiro Hayashida, and Tatsuya Akutsu. "Survival analysis by penalized regression and matrix factorization." The Scientific World Journal 2013 (2013).
15. Hao, Lin, et al. "Deep learning-based survival analysis for high-dimensional survival data." Mathematics 9.11 (2021): 1244.
16. Ebert, Martin A., et al. "Spatial descriptions of radiotherapy dose: normal tissue complication models and statistical associations." Physics in Medicine & Biology 66.12 (2021): 12TR01.
17. Ibragimov B, Toesca DAS, Chang DT, et al. Deep learning for identification of critical regions associated with toxicities after liver stereotactic body radiation therapy. Med Phys 2020;47:3721-3731.
18. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).
19. Efron, Bradley. "Large-scale simultaneous hypothesis testing: the choice of a null hypothesis." Journal of the American Statistical Association 99.465 (2004): 96-104.
20. Jing, L., & Tian, Y. (2020). Self-supervised visual feature learning with deep neural networks: A survey. IEEE transactions on pattern analysis and machine intelligence.
21. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis* 1987;40:373-83.
22. Thygesen SK, Christiansen CF, Christensen S, et al. The predictive value of icd-10 diagnosticcoding used to assess charlson comorbidity index conditions in the population-based danish national registry of patients. *BMC Med Res Methodol* 2011;11:83.
23. Rechner LA, Modiri A, Stick LB, et al. Biological optimization for mediastinal lymphoma radiotherapy - a preliminary study. *Acta Oncol* 2020;59:879-887.

Statistical Analysis Plan version 1.0

Administrative Information

Title

Clinical and dosimetric risk factors for ischemic heart disease following radiotherapy - an electronic health record based study of a large consecutive patient cohort at a single institution.

Roles and Responsibilities

Principal Investigator

Nora Forbes, MS

Dept of Clinical Oncology, Copenhagen University Hospital – Rigshospitalet

Department of Computer Science, University of Copenhagen

Investigators (responsibilities)

Dept of Clinical Oncology, Copenhagen University Hospital – Rigshospitalet

Maja Maraldo, MD, PhD (Breast responsible MD)

Mette Pøhl, MD, PhD (Lung responsible MD)

Signe Risumlund, MD, PhD (Esophagus responsible MD)

Lena Specht, MD, DMSc (Lymphoma responsible MD)

Ivan Vogelius, PhD, DMSc (Supervisor and statistics)

CHIP, Copenhagen University Hospital – Rigshospitalet

Joanne Reekie, PhD (Supervisor and statistics, PERSIMUNE data)

Cynthia Terrones Campos, PhD (Data extraction, PERSIMUNE)

Department of Epidemiology and Public Health, University of Maryland

Søren Bentzen, MSc, PhD, DMSc (statistics)

Introduction

The *Guidelines for the Content of Statistical Analysis Plans in Clinical Trials* were used to guide the generation of this statistical analysis plan.

The treatment of cancer with radiotherapy has been known to yield negative long-term effects, including cardiotoxicity. By combining machine learning on medical images, detailed 3D dosimetry data, and registry-based outcomes in a large cohort, we can better quantify the effect of radiation on cancer patients.

The purpose of this document is to present a brief analysis plan. The analysis shall describe the incidence and trends in ischaemic heart disease (IHD) as a function of cardiac exposure adjusting for a number of clinical factors. The objective is to generate an IHD risk model based on mean heart dose (MHD). Additionally, we will determine if other dose volume histogram (DVH) parameters can improve the prediction of IHD.

The overall objective is to use a comparatively large dataset with detailed exposure data from the record and verify system coupled with outcome data from the Danish registries to get a contemporary model able to predict the risk of cardiotoxicity from radiotherapy in the thoracic region.

Study Methods and Trial Population

This is an observational retrospective cohort study. A current estimate of crude numbers of treatments and events are as follows:

	Breast	Esophageal	Lung	Lymphoma
Treatments	7230	683	871	1708

	No IHD After RT	IHD After RT	Total
No Prior IHD	8,696	499	9,195
Prior IHD	894	403	1,297
Total	9,590	902	10,492

The study cohort will include all curative intent radiotherapy (RT) patients for Breast, Esophageal, Lung (NSCLC and SCLC) or Lymphoma cancers at Rigshospitalet from 2009 to 2020. These will be identified by treatment code in Aria. Exclusions in the extraction, computation, and analysis of the data will be detailed in a CONSORT flowchart. Patients with missing data will be excluded, as no imputation will be conducted.

Patients will be followed up in the Landspatientregisteret (LPR) and the Centrale Personregister (CPR) for the endpoint date. Baseline characteristics, including sex, age, and prior IHD, will be acquired in the LPR as well. Radiotherapy data will be acquired in the Aria database. This data will be acquired through the PERSIMUNE Data warehouse.

Statistical Principles and Analysis

R Statistical Software or Python will be used to conduct analyses. We will use an alpha level of 0.5 to determine statistical significance. Any modifications to this statistical analysis plan will be detailed in the corresponding manuscript.

The primary endpoint for this study will be Ischemic Heart Disease (IHD) as identified by ICD-10 codes I20-I25 recorded in the national patient register after the baseline date.

The baseline date will be the date of the last RT fraction. The endpoint date will be the first of the follow: the date of hospital admission for IHD (event), the date of death from IHD (event), date of death from other causes (censoring), date of emigration (censoring), or the last follow-up date (censoring).

Independent variables will include prior IHD (defined as the above ICD10 codes occurring before the baseline date), age at baseline date, heart dose as extracted in the study *Time trends in cardiac doses in a real-world data series of 11,000 curative thoracic* ***radiation therapy*** *courses from 2009-2020* by Forbes et al (in preparation) and patient sex as identified in the registry.

Firstly a descriptive analysis of cumulative incidence for each diagnosis group will be performed as follows: Cumulative Incidence will be calculated for each diagnosis (Breast, Lung, Esophageal, and Lymphoma), using the Fine & Gray method for competing risk, with the three outcomes being IHD, death, or still living at 31-12-2020. Data will be presented as stacked plots of incidences.

The primary analysis will be a Cox Proportional Hazards Regression to evaluate the time until IHD. The analysis will be stratified by diagnostic group (Breast, Lymphoma, Esophageal, and Lung cancer, any histology). Mean Heart Dose (MHD), Sex, Age, and prior IHD will be included in the model. Age will be fit with a spline model. MHD will either be included as a linear predictor or a spline model - an initial analysis will be conducted to determine which is a more appropriate fit as determined by AIC with preference towards linear predictor for simplicity in case of similar performance.

The fitted model will be made available as a R structure.

Following an analysis of MHD, we will separately add in a low (V5) and a high (V40) dose metric. We will then run a log rank test on the nested models to determine whether either of these variables add a significant amount of explainability to the model.

Planned exploratory analyses

Depending on the results of the primary analysis above, we intend the following extensions:

1) Methodological exploration: Time dependent Cox regression from age 0 with time dependent exposure variable and/or use attained age as the time unit
2) Separate analysis of Esophageal and Lung cancers according to the method described in primary analysis. Here the purpose will be to clarify if the proposed stratification for diagnosis in the primary analysis is sufficient or if there is a need to handle the diagnostic groups completely separately.

Title

Radiotherapy exposure and association with observed cardiovascular toxicity in over 5000 patients

Authors

Nora Forbes[1], Abraham Smith[2], Jens Petersen[2], Cynthia Terrones-Campos[3], Joanne Reekie[4], Sune Darkner[2], Lena Specht[5], Ivan Vogelius[5]

Authors Affiliations

[1]Rigshospitalet, Radiation Oncology, Copenhagen , Denmark; [2]University of Copenhagen, Department of Computer Science, Copenhagen, Denmark; [3]Rigshospitalet, Department of Infectious Diseases, Copenhagen, Denmark; [4]Rigshospitalet, PERSIUNE, Copenhagen, Denmark; [5]Rigshospitalet, Department of Oncology, Copenhagen, Denmark

Purpose or Objective

The effect of radiation exposure on cardiovascular disease has been studied in several publications, across many patient diagnoses and with varying definitions of cardiovascular disease. For example, in Breast cancer the rate of major coronary events increased linearly by 7.4% per gray of mean radiation dose to the heart.

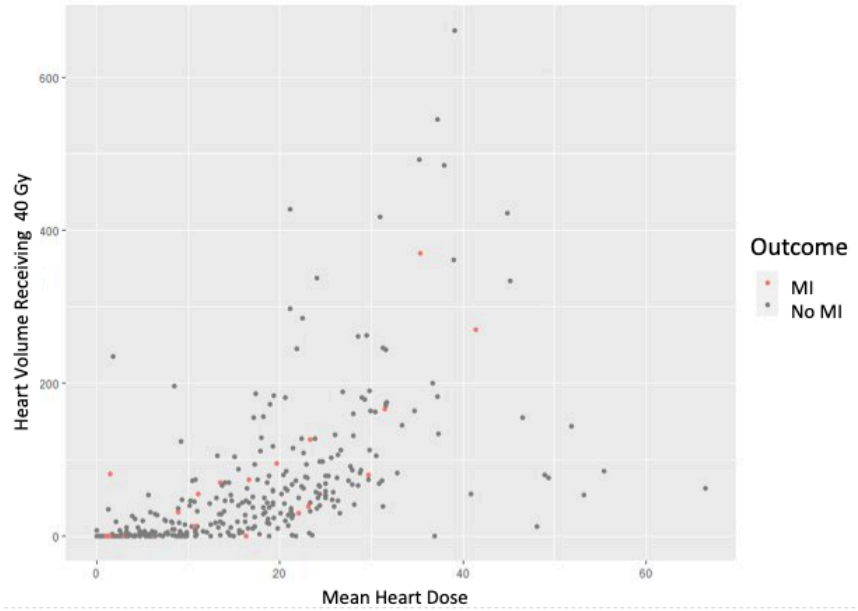
The purpose of the research is to estimate the association between radiotherapy exposure and cardiovascular toxicity after radiotherapy. In particular the aim is to establish quantitative associations between normal tissue exposure and incidence/prevalence of side effects. The rationale for the project is to improve our understanding of such associations, which will in turn allow us to better tailor our radiotherapy to minimize the risk of the observed cardiovascular toxicity while maintaining or improving the probability of durable disease control.

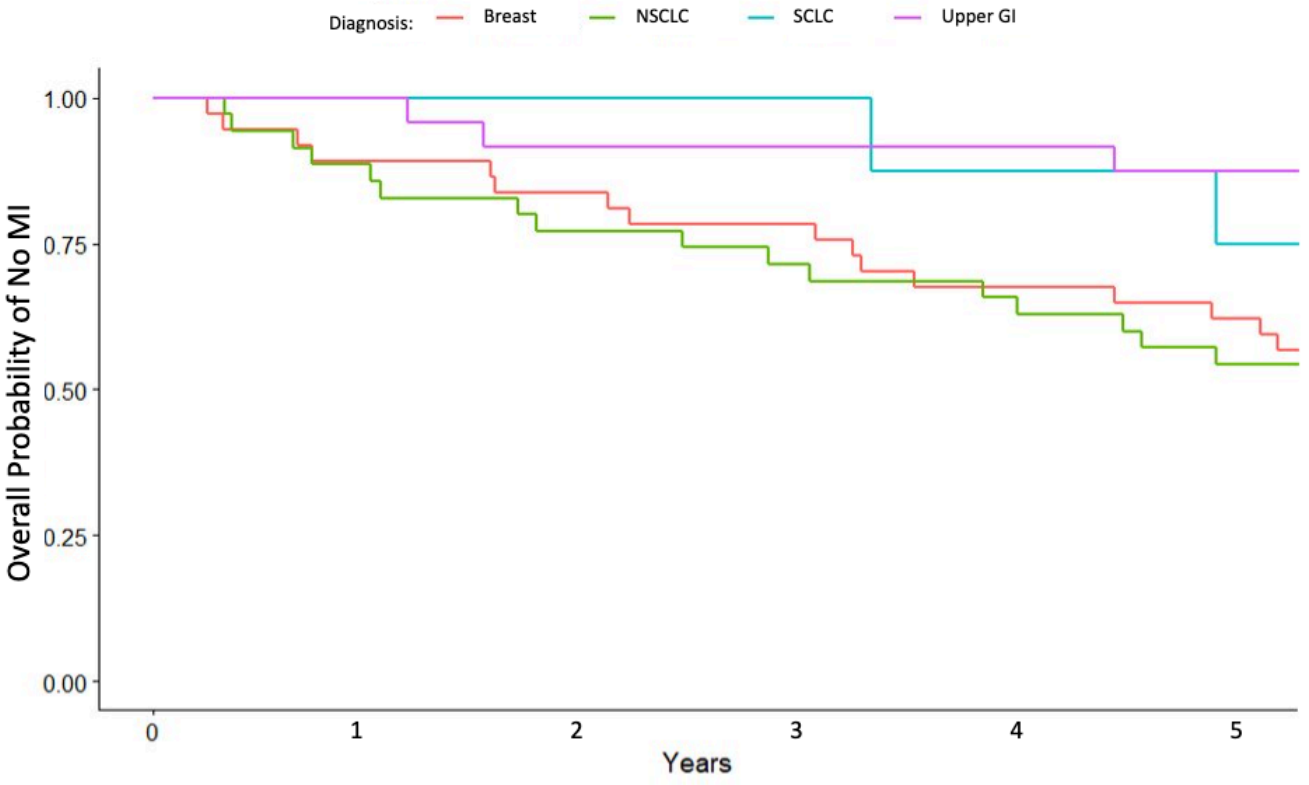
Materials and Methods

This is a non-randomized retrospective study with a large sample size to estimate radiation dose-effect relationship with cardiotoxicity, specifically myocardial infarction (MI) defined by ICD-10 code I21. Eligible patients were diagnosed with cancer and received curative intent radiotherapy at Rigshospitalet between 2009 and 2016. Manual delineations of organs at risk are time consuming and thus effectively limited to small populations. Therefore, heart annotation was conducted using RootPainter3D trained from 933 manually delineated CT scans in a hematological toxicity study. All delineations were reviewed by a researcher and questionable cases were discussed with a physician.

Results

Of the patients considered for inclusion, 5544 were eligible for analysis. The diagnoses of greatest interest with regard to heart exposure are Breast (n = 2139), NSCLC (n = 411), Upper GI (n = 406), and SCLC (n = 106). Of those 2.6% of Breast (n = 55), 10.1% of NSCLC (n = 45), 7.1% of Upper GI (n = 29), and 8.5% SCLC (n = 9) experienced MI following radiotherapy treatment.





Conclusion

This abstract represents preliminary findings for a large retrospective cohort study. This study supports a direction away from mean organ at risk analysis and toward distributional dose quantification. Additional organs at risk should be segmented and accounted for. Furthermore, quantification of dose to substructures of the heart may provide further insights. While this is a large study compared to the existing literature on this topic, there is an opportunity to expand the timeframe as well as location, which may improve generalizability and power of the study.