

A Dual Radiomic and Dosiomic Filtering Technique for Locoregional Radiation Pneumonitis  
Prediction in Breast Cancer Patients

by

Cheng Qian

DKU Medical Physics  
Duke University

Date: March 20<sup>th</sup>, 2025

Approved:

Zhenyu Yang, Co-Chair

Fang-Fang Yin, Co-Chair

Manju Liu

Jun Duan

Minbin Chen

Thesis submitted in partial fulfillment of the requirements for the degree of  
Master of Science in the DKU Medical Physics of  
Duke University  
2025

ABSTRACT

A Dual Radiomic and Dosiomic Filtering Technique for Locoregional Radiation Pneumonitis  
Prediction in Breast Cancer Patients

by

Cheng Qian

DKU Medical Physics  
Duke University

Date: March 20<sup>th</sup>, 2025

Approved:

Zhenyu Yang, Co-Chair

Fang-Fang Yin, Co-Chair

Manju Liu

Jun Duan

Minbin Chen

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of  
Master of Science in the DKU Medical Physics of  
Duke University  
2025

Copyright by  
Cheng Qian  
2025

## Abstract

**Purpose:** To develop a novel Explainable Dual-Omics Filtering (EDOF) model integrating dosiomic and radiomic filtering to predict locoregional radiation pneumonitis (RP) in breast cancer patients, and to explain the critical locoregional dosimetric indices and radiomic features that contribute to RP development.

**Materials and Methods:** This retrospective study collected 72 breast cancer patients treated with radiation therapy, and a total of 28 patients developed RP (including 5 grade II cases) within 4 months post-treatment. The lung volume was first segmented from pre-treatment CT, and 3D dose distribution was also collected from the treatment plan. A 3D sliding window kernel was implemented across the (1) lung CT to capture 70 spatial-encoded image texture information, and (2) lung dose distribution to capture 36 spatial-encoded dose intensity information. As such, each voxel coordinate of the original lungs was represented as a 106-dimensional dual-omics feature vector. A novel explainable boosting machine (EBM) model was employed to establish a voxel-wise association between extracted features with locoregional RP, as identified in follow-up CT. Comparative studies against (1) radiomic filtering-only (RF) and (2) dosiomic filtering-only (DF) models were also performed. The model performance was evaluated through voxel-wise AUC, accuracy, specificity, and sensitivity with 5-fold cross-validation. The dice coefficient was additionally calculated for 5 grade II cases. The mean absolute score from EBM was also extracted to rank the feature importance.

**Result:** The EDOF model showed highest voxel-wise RP prediction power (accuracy=0.93, sensitivity=0.93) than the RF (accuracy=0.89, sensitivity=0.01) and DF (accuracy=0.90, sensitivity=0.92) models. Specificity and AUC showed similar trends. Dice coefficient for 5 grade

II patients is 0.75 in our EDOF model. Based on the EBM score, the results suggested that heterogeneous lung tissue with high locoregional dose has high risk of RP.

**Conclusion:** The EDOF model accurately identified locoregional RP regions based on pre-treatment image and planning dose, offering a significant advancement in predictive analytics for radiation-induced complications.

# Contents

Abstract.....	iv
List of Tables .....	viii
List of Figures.....	ix
List of Abbreviations .....	x
Acknowledgements.....	xii
1. Introduction.....	1
1.1 Background .....	1
1.2 Radiomics and Dosiomics .....	3
1.2.1 Radiomics.....	3
1.2.2 Dosiomics.....	6
1.3 Machine Learning (ML).....	6
1.4 Current Study .....	8
1.4.1 RP Prediction Using Dosimetric Information .....	8
1.4.2 RP Prediction Using Dosimetric and Image Information.....	9
1.5 Limitations of Previous Study .....	10
1.5.1 Locoregional Risk Regions and Their Clinical Relevance.....	10
1.5.2 Challenges in Model Explainability for RP Prediction .....	10
1.6 Purpose .....	11
2. Materials and Methods.....	12
2.1 Patient Data .....	12
2.2 EDOF Model Design.....	14
2.2.1 Radiomic Filtering.....	15
2.2.2 Dosiomic Filtering.....	17

2.2.3 Explainable RP Prediction.....	19
2.2.4 Feature Importance and Partial Dependence Plot .....	21
2.3 Training Details.....	22
2.4 Evaluation Metrics .....	22
2.4.1 AUC .....	23
2.4.2 Sensitivity, Specificity, Accuracy .....	23
2.4.3 DSC .....	24
2.5 Comparison Study .....	24
3. Results.....	26
3.1 Evaluation results .....	26
3.2 Feature Importance and Partial Dependence Plot.....	28
4. Discussion.....	29
4.1 Technical Insights and Feature Analysis .....	29
4.2 Clinical Implications and Limitations .....	31
5. Conclusion .....	33
References.....	34

## List of Tables

Table 1. Patient clinical and treatment characteristics .....	12
Table 2. Radiomic and Dosiomic Features in EDOF.....	18
Table 3. Evaluation Results of Comparative Studies.....	26

## List of Figures

Figure 1. Interpretability and Accuracy of Machine Learning Models.....	8
Figure 2. The identification and alignment of RP-affected regions in follow-up and simulation CT scans.....	13
Figure 3. The overall design of the proposed EDOF model. (A) radiomic filtering, (B) dosiomic filtering, and (C) explainable voxel-wise RP prediction.....	14
Figure 4. Visualization of voxel-wise RP prediction from EBM model (image on the left) and RP ground truth (image on the right).....	19
Figure 5. ROC curves from Comparative Studies .....	26
Figure 6. Feature Importance and Partial Dependence Plot. (A) Feature Importance Plot. (B) R-SRLGLE PDP. (C) D-Mean PDP. (D) D-Mean Absolute deviation PDP.....	27

## List of Abbreviations

AI	Artificial Intelligence
AUC	Area Under the Curve
BC	Breast Cancer
CPU	Central Processing Unit
CT	Computed Tomography
CNN	Convolutional Neural Networks
DF	Dosimetric Filtering
DSC	Dice Similarity Coefficient
DVH	Dose-Volume Histogram
EBM	Explainable Boosting Machine
EDOF	Explainable Dual Radiomic and Dosimetric Filtering
FDG	Fluorodeoxyglucose
FP	False Positive
FAST	Fast Interaction Detection
GAM	Generalized Additive Model
GLCM	Gray-Level Co-occurrence Matrix
GLRLM	Gray-Level Run Length Matrix
GLSZM	Gray-Level Size Zone Matrix
IBSI	Imaging Biomarker Standardization Initiative
IMRT	Intensity-Modulated Radiation Therapy
LIME	Local Interpretable Model-agnostic Explanations
LRP	Layer-wise Relevance Propagation
MITK	Medical Imaging Interaction Toolkit

MLD	Mean Lung Dose
MRI	Magnetic Resonance Imaging
PDP	Partial Dependence Plot
PET	Positron Emission Tomography
RF	Radiomic Filtering
ROC	Receiver Operating Characteristic
ROI	Region of Interest
RP	Radiation Pneumonitis
SHAP	SHapley Additive exPlanations
SRLGLE	Short Run Low Gray Level Emphasis
TNBC	Triple-Negative Breast Cancer
TP	True Positive
VOI	Volume of Interest
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

## **Acknowledgements**

I extend my deepest gratitude to my advisor, Dr. Zhenyu Yang, for his unparalleled mentorship throughout my academic journey. His consistent guidance and unwavering support have been instrumental in shaping both my academic development and personal growth. For his dedication and encouragement, I am truly grateful.

I am also sincerely thankful to Dr. Fang-Fang Yin and Dr. Minbin Chen for providing me with the invaluable opportunity to collaborate with Duke Kunshan University and conduct research at The First People's Hospital of Kunshan. Their support has been critical in furthering my research career and enhancing my understanding of the field.

A special thanks goes to my teammates, Rihui Zhang, Haiming Zhu, and Tianyi Zhang. Their collaboration, insights, and friendship have made this journey both rewarding and memorable.

I would also like to express my appreciation to the team at Microsoft for their work on the Explainable Boosting Machine (EBM), which served as the foundation for the predictive model developed in this thesis. Their contributions to machine learning and explainability have significantly advanced my research.

Finally, I am forever grateful to my parents and friends for their endless love, encouragement, and belief in me. Their support has been a constant source of motivation, helping me through every challenge and driving me towards success.

# 1. Introduction

## 1.1 Background

Breast cancer is a major public health issue and the most diagnosed cancer among women globally (Giaquinto et al., 2022.; Xu et al., 2023). In 2022, around 2.3 million women were diagnosed with breast cancer, resulting in approximately 670,000 deaths (WHO, 2024). It is influenced by several factors, including genetic predispositions, hormonal influences, alcohol consumption, and environmental exposures (Momenimovahed & Salehiniya, 2019). The rate of breast cancer varies significantly across the globe. Developed countries tend to have higher incidence, while the mortality rate is higher in lower-income countries. These differences mainly stem from unequal access to CT screening, early detection, and treatment resources (Arnold et al., 2022). In China, there were approximately 303,600 new cases of breast cancer and the incidence and mortality rate continue to increase (Lei et al., 2021). The global burden of breast cancer emphasizes the need for personalized, multidisciplinary treatments.

Breast cancer treatment involves several standard approaches based on cancer stage, tumor characteristics, and patient-specific factors. The primary treatment methods include:

- 1) **Surgery:** This is typically the first-line treatment, aiming to remove the tumor. Options include mastectomy (removal of the entire breast) or breast-conserving surgery (lumpectomy), where only the tumor and some surrounding tissues are removed. The choice between these options depends on factors like tumor size, location, and patient preference (Shubeck et al., 2022).
- 2) **Chemotherapy:** This involves the use of medications to kill or prohibit the growth of cancer cells. It is commonly used either before surgery to shrink the tumor or after surgery to eliminate remaining cancer cells. However, chemotherapy can come with severe side effects like fatigue, hair loss, and nausea (P et al., 2022).

- 3) **Endocrine Therapy:** For hormone receptor-positive breast cancers, endocrine therapy is a critical treatment to block estrogen, which can fuel the growth of cancer cells
- 4) **Immunotherapy:** This method helps the body's immune system recognize and kill cancer cells. It has exhibited impressive anti-tumor activity and clinical benefit in different malignancies, with particularly promising outcomes recently observed in the treatment of triple-negative breast cancer (TNBC) (Zhang et al., 2022).
- 5) **Radiation Therapy (RT):** Radiation therapy is a crucial treatment modality in breast cancer, particularly valued for enabling breast-conserving surgery (lumpectomy). By delivering targeted radiation to the tumor, RT ensures oncologic safety comparable to mastectomy, allowing patients to preserve their breast while minimizing recurrence risk. Modern advancements, such as intensity-modulated radiation therapy (IMRT), enhance precision by conforming radiation doses to the tumor's 3D shape, thereby sparing adjacent healthy tissues like the heart and lungs. This accuracy reduces long-term radiation toxicity risks while maintaining therapeutic efficacy. Additionally, RT plays a critical role in eradicating residual cancer cells post-surgery, significantly lowering the likelihood of local recurrence. Studies show that approximately 70% of breast cancer patients receive RT (Maliko et al., 2022).

However, radiation pneumonitis (RP) is a major complication that can occur after RT, particularly in patients receiving treatment for breast cancer. RP occurs when radiation impacts healthy lung tissue within the treatment field, leading to inflammatory alveolar exudation. Symptoms typically include shortness of breath, dry cough, low-grade fever, and reduced oxygen levels in the blood. Pathologically, RP is characterized by alveolar septal edema, endothelial cell swelling, thickened vascular walls, and other tissue changes. On CT imaging, RP often appears as patchy, uniformly hazy regions in the radiated area, with thickened blood vessels and bronchial

markings that blend indistinctly with nearby healthy lung tissue. The duration of RP onset following radiotherapy varies; acute RP may develop within days of treatment, while chronic RP generally arises within 12 weeks to six months post-treatment (Hanania et al., 2019). Once present, RP can lead to irreversible fibrotic changes in the lung tissue, impairing respiratory function and significantly impacting patients' quality of life and survival (Chen et al., 2023). RP severity is graded using a modified version of the National Cancer Institute Common Toxicity Criteria (CTC), ranging from Grade 0 (no symptoms) to Grade 4 (life-threatening symptoms). Grades 2 and above indicate significant toxicity requiring clinical intervention.

With the development of technology, 90.8% of women with breast cancer survive for 5 years after diagnosis (American Cancer Society, 2019). Therefore, the quality of life for these patients is becoming more concerned, and radiation-induced RP has become a major focus in healthcare. Cases of higher-grade pneumonitis (grade  $\geq 2$ ) are linked to substantial quality-of-life reductions and potential progression to irreversible lung damage, including fibrosis and organizing pneumonia (Otani et al., 2017). Therefore, it is essential to establish a model for predicting RP during the initial assessment and treatment planning.

## ***1.2 Radiomics and Dosiomics***

### **1.2.1 Radiomics**

In recent years, radiomics has gained significant attention for its ability to correlate features in medical imaging with clinical outcomes and treatment responses. Radiomics is a technique that extracts high-dimensional quantitative features from medical images to reveal information about tissue characteristics, capturing texture, patterns, and voxel value distributions that are not visible to naked eyes (Gillies et al., 2016). These features extracted from lung CT images can be used to build models to predict RP (Y. Huang et al., 2022; Puttanawarut et al., 2022).

The radiomics workflow consists of five stages (Lambin et al., 2012):

1. **Image Acquisition and Pre-processing:** Radiomics can be applied to various imaging modalities, including PET, CT, and MRI. Different modalities and devices may introduce variations in image quality. Therefore, pre-processing is essential to standardize these images, ensuring consistency across datasets.
2. **Region of Interest (ROI) Segmentation:** ROIs are typically delineated by medical physicists, although automated or semi-automated software tools may also be used. Manually segmented ROIs are often considered the ground truth for subsequent analyses.
3. **Feature Extraction and Selection:** Advanced software tools, such as PyRadiomics and MITK, are used for feature extraction (Aerts et al., 2014). However, variations in computational methods across different platforms can lead to inconsistent results. To address this, the Image Biomarker Standardization Initiative (IBSI) has established a standardized set of 169 features to validate these tools (Zwanenburg et al., 2020). Additionally, feature selection is employed to eliminate correlated features, enhancing the accuracy of the analysis.
4. **Model Development:** Predictive models are constructed using statistical or machine learning techniques, integrating the selected features to forecast clinical outcomes based on the characteristics of the ROIs.
5. **Assessment and Prediction:** The performance of the models is rigorously evaluated for predictive accuracy, and they are subsequently applied to new datasets to forecast clinical outcomes, thus aiding in personalized treatment planning.

Radiomics features can be broadly divided into four categories: shape features, first-order statistics (intensity features), texture features, and higher-order features (van Griethuysen et al., 2017). Each category provides unique insights—shape features capture the external geometry of

the ROI; first-order statistics describe intensity distributions; texture features reveal spatial patterns in voxel intensities; and higher-order features offer advanced detail through mathematical transformations, unveiling fine structural characteristics within the tissue.

- 1) **Shape Features:** Shape features describe the geometric characteristics of a region of interest (ROI), such as a tumor, and include measurements like volume, surface area, sphericity, compactness, and elongation. These metrics provide information on the physical form and spatial structure of the tumor, independent of pixel intensity values. Since shape features focus on external morphology, they are particularly useful for understanding tumor growth patterns and tracking changes over time, especially in response to treatment.
- 2) **First-Order Statistics (Intensity Features):** First-order statistics, or intensity features, capture the distribution of voxel intensities within the ROI without considering spatial relationships. Common measures include mean intensity, standard deviation, skewness, kurtosis, median, and percentiles. These metrics describe the uniformity, brightness, and contrast of the tissue, reflecting its overall density and composition. Intensity features are valuable for identifying general tissue characteristics and changes in density, which could signal tumor heterogeneity.
- 3) **Texture Features:** Texture features assess the spatial relationships between pixel intensities within the image, capturing patterns that represent tissue structure and heterogeneity. Second-order texture metrics, like those derived from the Gray Level Co-occurrence Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM), measure aspects such as contrast, correlation, and homogeneity, which indicate the complexity and uniformity of tissue textures. Texture features provide insights into the tumor's

microenvironment, revealing details that may correlate with cellular or tissue-level characteristics not visible to the naked eye.

- 4) **Higher-order Features:** Higher-order features are obtained by applying mathematical transformations, such as wavelet or Laplacian filters, to the original medical image. These transformations emphasize complex, multi-scale patterns and highlight fine texture details within the tissue. By isolating certain frequency ranges or directional components, higher-order features provide a more detailed analysis of the tissue's microarchitecture. This advanced level of detail can potentially reveal structural and pathological characteristics related to genetic or molecular properties, offering deeper insights into the tissue's underlying biology.

### **1.2.2 Dosiomics**

Dosiomics has also gained attention in radiation therapy, focusing on extracting statistical features from 3D dose distributions. It includes shape features, first-order statistics, and various texture features that describe the spatial arrangement of voxel intensity levels within a specified volume of interest (Liang et al., 2019). The potential of dosiomic features to predict radiation-induced toxicity has been explored in organs like lung and head and neck, showing promising results (M. D. Anderson Cancer Center Head and Neck Quantitative Imaging Working Group, 2018). Since dosiomic features capture spatial information, models that incorporate them are expected to perform better than those based on dose-volume histograms (DVHs) (Feng et al., 2024).

### **1.3 *Machine Learning (ML)***

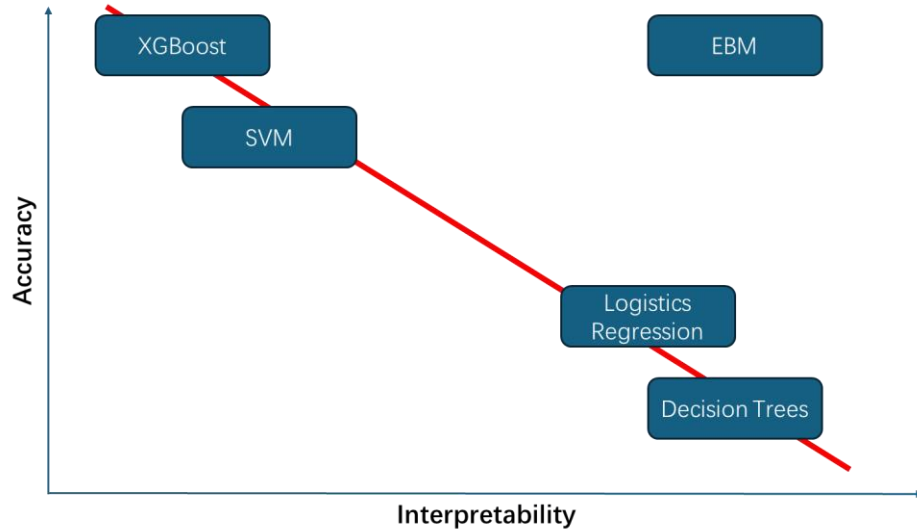
Machine learning (ML) has gained increasing attention in the field of medical imaging studies due to its capacity to analyze and interpret complex data, making it an invaluable tool for clinical applications such as disease detection, prognosis, and treatment planning (Asif et al., 2024;

Rana & Bhushan, 2023). ML refers to a subset of artificial intelligence (AI) that enables systems to learn patterns from data, making it particularly effective for handling the high dimensionality and variability inherent in medical imaging datasets. Despite their high accuracy in predictions, ML models are often criticized for their “black box” nature, which complicates the interpretation of their decision-making processes. To address this issue, various explanation techniques have been developed to provide insights into model behavior. These techniques include Shapley additive explanations (SHAP), local interpretable model-agnostic explanations (LIME), and permutation feature importance (PFI) (Lundberg & Lee, 2017; Ribeiro et al., 2016).

While these interpretation methods offer valuable tools for understanding ML models, their use requires careful consideration to avoid potential pitfalls. Common issues include misapplying these techniques in inappropriate contexts, analyzing models with poor generalization capabilities, overlooking feature dependencies and interactions, ignoring uncertainty estimates, and failing to address challenges posed by high-dimensional data. Moreover, unjustified causal interpretations remain a critical concern associated with these methods. Proper application of these interpretation techniques is essential to ensure that the insights drawn from ML models are both accurate and meaningful. (Alahmadi et al., 2023)

In contrast, “glass-box” ML models are specifically designed to provide inherent interpretability, ensuring that their outputs are both trustworthy and easily understood by users. One such model is the explainable boosting machine (EBM), a tree-based, cyclic gradient-boosting method that falls under the category of generalized additive models. EBM enables straightforward understanding of its internal workings by employing automatic interaction detection (Nori et al., 2021). Regarding performance, EBM has been shown to achieve results comparable to those of advanced ML models, including gradient boosting, support vector machines (SVM), extreme gradient boosting (XGBoost), and random forests, etc. A notable advantage of EBM is its use of

factor-specific shape functions, which contribute to producing inherently interpretable outcomes. (Maxwell et al., 2021). This model stands out from complex black-box ML approaches by offering decision-making processes and final predictions that are interpretable in both local and global contexts.



**Figure 1. Interpretability and Accuracy of Machine Learning Models**

## ***1.4 Current Study***

### **1.4.1 RP Prediction Using Dosimetric Information**

RP prediction with dosimetric information has been extensively studied. Dosimetric parameters, such as DVH, lung V20 (the percentage of lung volume receiving 20 Gy or more), and mean lung dose (MLD), are commonly used for RP prediction. Boonyawan et al. (2018) found that RP cases increase with V10 and V20. Bi et al. (2022) reported that V20 and MLD are independently associated with RP development. Ozgen et al. (2023) reported that low dose-volume metrics like V5 and V10 are important in predicting the risk of RP. Pinnix et al. (2015) concluded that the V5 parameter demonstrates a superior predictive ability for radiation pneumonitis compared to other dose-volume metrics, including V10, V15, and V20. These findings underscore that managing

specific dose thresholds is essential for reducing RP incidence. However, DVH parameters can only provide a summary of the two-dimensional dose distribution in the target area, and they do not convey the spatial dose distribution or the anatomical structure of the organs. Some studies have shown that the voxel-level dose is associated with RP development (Y. Huang et al., 2022).

To address the limitations of traditional dosimetric models, recent studies have explored dosiomic features that incorporate spatial information to improve RP prediction. Feng et al. (2024) developed a model using 3D dosiomic features extracted from dose distributions alongside DVH parameters and radiomic features from pre-treatment CT scans. This hybrid model outperformed traditional models, achieving an AUC of 0.920, indicating that dosiomic features could enhance the predictive accuracy.

#### **1.4.2 RP Prediction Using Dosimetric and Image Information**

Combining dosimetric and radiomic information has been shown to improve RP prediction accuracy (Puttanawarut et al., 2022). Dosimetric parameters estimate radiation exposure, while radiomic features provide insights into tissue characteristics and heterogeneity. This integration allows prediction models to leverage the strengths of both approaches for a more comprehensive assessment of RP risk.

Y. Huang et al. (2022) focused on non-small cell lung cancer patients and utilized dosiomic and deep learning-based radiomic features from 3D dose distributions, highlighting the enhanced predictive capabilities of these combined features. Puttanawarut et al. (2022) examined esophageal and lung cancer datasets, demonstrating that dosiomic and radiomic features outperform conventional DVH metrics in RP prediction, reinforcing the notion that advanced imaging techniques offer superior insights into treatment-related complications. Nie et al. (2024) further supported this approach by integrating various radiomic features and dosimetric parameters, demonstrating the advantages of a comprehensive modeling strategy in predicting symptomatic RP.

Similarly, Zhang et al. (2023) confirmed the robustness of combining radiomic and dosiomic features across multiple cancer types, underscoring their versatility in clinical applications. These studies affirm that models with radiomic and dosiomic features are more effective and accurate for RP prediction.

## ***1.5 Limitations of Previous Study***

### **1.5.1 Locoregional Risk Regions and Their Clinical Relevance**

Few studies have focused on identifying high-risk regions for locoregional RP, even though doing so could significantly refine treatment strategies. For instance, Petit et al. (2011) found that lung areas more susceptible to radiation-induced toxicity could be identified by analyzing [18F] fluorodeoxyglucose (FDG) uptake patterns before radiotherapy. Studies have demonstrated that elevated [18F]FDG uptake correlates with heightened metabolic response to radiation, increasing susceptibility to RP. Therefore, pre-treatment monitoring of [18F]FDG uptake could aid in identifying lung regions at greater RP risk, supporting a strategy of limiting radiation exposure in these high-uptake areas to better protect lung tissue. Also, J.-W. Huang et al., (2023) suggested that the percentage of RP volume to total lung volume can be used to identify severity of RP. The increased RP lesion volume tends to have more noticeable clinical symptoms and imaging abnormalities, which contribute to a higher RP grade. These studies underscore the promise of incorporating risk region identification into treatment planning.

### **1.5.2 Challenges in Model Explainability for RP Prediction**

One of the limitations with current machine learning-based RP prediction models is their lack of explainability. Machine learning analysis employs a variety of data mining algorithms across different types and formats to systematically characterize data features, enabling a more rigorous understanding of data patterns and recognized values (Zampieri et al., 2019). These

models often function as "black box". Although they can make accurate predictions, it is difficult to understand how and why these predictions are made. The "black box" nature of AI creates barriers to its adoption in clinical practice, as it limits clinicians' ability to evaluate the quality of training data and labels, conflicting with evidence-based medicine standards. Therefore, achieving clear explainability of a model's output is crucial for building user trust and identifying ways to refine the model (Ma et al., 2023).

The need for explainability in medical machine learning models is increasingly recognized by researchers and practitioners. As models become more complex, integrating explainable AI techniques will be crucial to ensure that these models are not only accurate but also reliable and useful in clinical practice.

### ***1.6 Purpose***

The purpose of this study is to develop an Explainable Dual Radiomic and Dosiomic Filtering (EDOF) model that leverages both imaging and dosimetric features to accurately predict locoregional RP in breast cancer patients. By incorporating explainable AI techniques, the model aims to provide transparent predictions, enabling clinicians to understand the key factors contributing to RP risk and facilitating more personalized treatment planning.

## 2. Materials and Methods

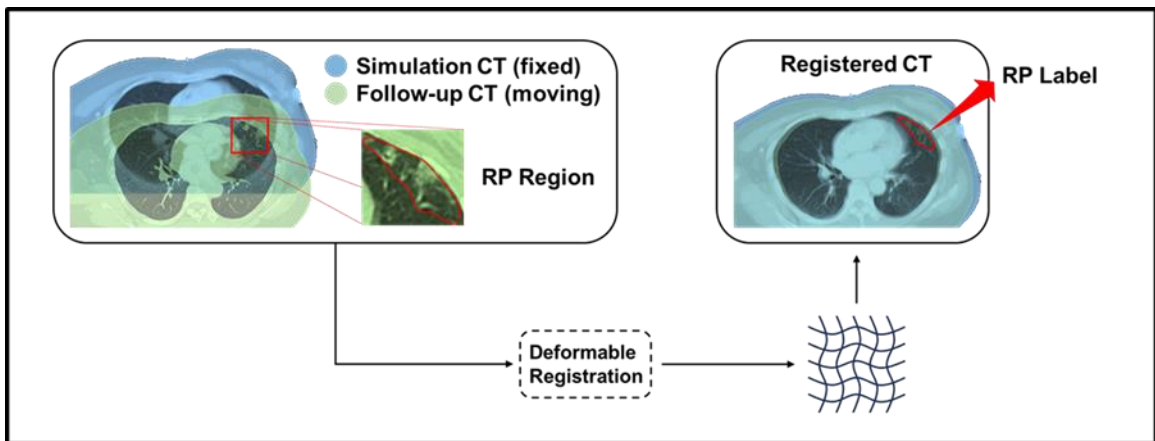
### 2.1 Patient Data

A total of 72 breast cancer patients treated with IMRT from 2019 to 2021 at the First People's Hospital of Kunshan were included in this study. Among these patients, 28 patients developed RP, with 23 cases classified as grade I RP and 5 cases classified as grade II RP. The details of the patient information were summarized in Table 1. The selection criteria included:

**Table 1. Patient clinical and treatment characteristics**

<b>Characters</b>	<b>RP</b>	<b>Non-RP</b>
Age median (years)	50 (26-79)	49 (30-68)
Gender		
Male	0	0
Female	28	44
Chemotherapy		
Yes	20	34
No	8	10
Tumor location		
Left	14	21
Right	14	23
Irradiation location		
Breasts, axillary, and infraclavicular lymph	13	18
Breasts only	15	26
Dose fractionations		
60-61 Gy/28 fractions	18	28
40-56 Gy/25 fractions	9	14
others	1	2
Volume dose		
V5	47.34±12.22	43.26±14.19
V10	27.72±6.86	25.48±7.86
V20	13.34±2.73	12.43±2.78
V30	8.85±1.87	7.93±2.12
MLD (Gy)	9.27±1.73	8.01±1.90

For each patient, simulation CT scans prior to treatment, lung segmentation, and corresponding IMRT planning dose distributions (Varian Eclipse™) were collected. Additionally, follow-up CT scans within four months post-treatment were obtained to identify the occurrence of RP. For the 28 patients who developed RP, the regions affected by RP were identified and contoured on the follow-up CT images by experienced medical physicists. These contoured RP regions were then aligned with the simulation CT images using deformable image registration. The obtained registration results provided voxel-wise RP labels, serving as the ground truth for RP regions.



**Figure 2. The identification and alignment of RP-affected regions in follow-up and simulation CT scans**

## 2.2 EDOF Model Design

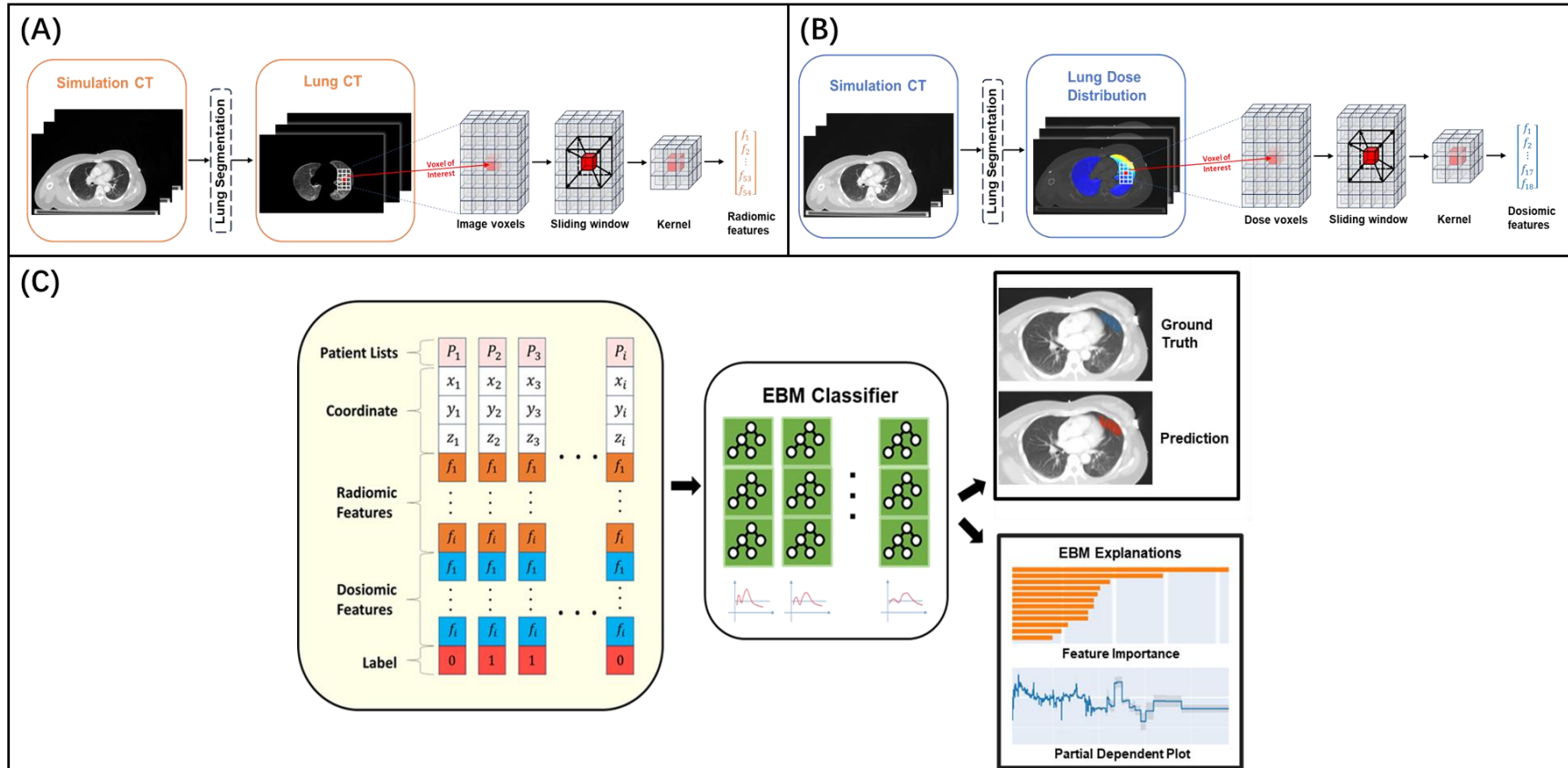


Figure 3. The overall design of the proposed EDOF model. (A) radiomic filtering, (B) dosiomic filtering, and (C) explainable voxel-wise RP prediction.

### 2.2.1 Radiomic Filtering

The lung volume for each patient was first obtained based on simulation CT and the corresponding lung segmentation. As shown in Figure 3(A), radiomic filtering employed a 3D sliding window kernel to capture the regional radiomic features throughout the entire lung volume. Specifically, a pre-defined 3D kernel systematically moves through the lung volume in a voxel-wise manner. For each position of the sliding window, a cubic sub-volume was defined, and radiomic features were extracted from this sub-volume. Consequently, each voxel coordinate within the original lung was represented as an n-dimensional feature vector. A set of 3D radiomic filtering maps can be obtained and visualized in the same reference frame as the original lung CT image. Following our pilot study, the kernel size for the 3D sliding window was set to  $7 \times 7 \times 7 \text{ mm}^3$ , which corresponds to the typical order of localized variations in lung texture. Formally, given  $N$  radiomic features and a CT image with dimensions  $I \times J \times K$ , a total of  $N$  radiomic filtering maps (denoted as  $\{R^n\}_{n=1}^N$ ) can be obtained. Each radiomic filtering map  $R^n$  is represented as a 3-dimensional tensor  $R^n \in \mathbb{R}^{I \times J \times K}$ , where the element  $R^n_{i,j,k}$  denotes the value of the  $n^{\text{th}}$  radiomic feature at the  $(i, j, k)^{\text{th}}$  tomographic coordinate.

Following previous lung radiomic studies and our pilot studies, a total of  $n=54$  radiomic texture features were included in this work to comprehensively capture local texture intensity and texture characteristics within the lungs. Table 2 summarizes all 54 features, which can be grouped into three types based on different joint-probability functions:

- 1) Gray Level Co-occurrence Matrix (GLCOM)-based features: These features measure the distribution of pair-wise gray level combinations within the image, providing insights into the textural patterns and heterogeneity of the lung tissue.

- 2) Gray Level Run-Length Matrix (GLRLM)-based features: These features assess the distribution of consecutive intensity values of the same gray level in specified directions, capturing the length and uniformity of texture runs within the lung parenchyma.
- 3) Gray Level Size Zone Matrix (GLSZM)-based features: These features measure the distribution of pixel patches of similar gray levels, indicating the size and extent of homogeneous zones within the lung tissue.

The multi-collinearity assessment was subsequently performed to remove the highly correlated radiomic filtering maps. Specifically, Pearson correlation analysis was conducted for each pair of radiomic filtering maps, resulting in a Pearson covariance matrix for each patient. To obtain a representative measure of multi-collinearity across the entire patient cohort, the mean Pearson covariance matrix was calculated by averaging the individual matrices. Subsequently, hierarchical clustering was applied to the average correlation matrix to categorize the radiomic features into well-separated clusters. This clustering approach utilized a dendrogram diagram to visualize the relationships among the features, where similar or highly correlated features were positioned closer together. A specific cut-off value was determined to delineate the clusters, and features whose distances fell below this threshold being grouped into the same cluster. This process yielded a total of 26 clusters, each representing a distinct set of correlated features. Within each cluster, one feature was arbitrarily selected and retained as the independent feature for subsequent modeling.

### 2.2.2 Dosiomic Filtering

The dosiomic filtering technique adopted the similar conceptual design of the radiomic filtering. As shown in Figure 3(B), the 3D lung dose distribution was first obtained from the IMRT treatment plan for each patient. A 3D sliding window kernel was implemented to capture the dosiomic intensity features throughout the lung dose. Similarly, given  $M$  dosiomic features and a lung dose distribution with dimensions  $I \times J \times K$ , a total of  $M$  dosiomic filtering maps (denoted as  $\{\mathbf{D}^m\}_{m=1}^M$ ) can be obtained. Each dosiomic filtering map  $\mathbf{D}^m$  is represented as a 3-dimensional tensor  $\mathbf{D}^m \in \mathbb{R}^{I \times J \times K}$ , where the element  $D_{i,j,k}^m$  denotes the value of the  $m^{\text{th}}$  dosiomic feature at the  $(i, j, k)^{\text{th}}$  tomographic coordinate.

In this work, a total of  $m = 18$  dosiomic features were included, as shown in Table 2. These features were chosen to capture the locoregional intensity characteristics within the lung dose. The multi-collinearity assessment was also performed in a manner identical to that used for the radiomic filtering maps. As such, 5 independent dosiomic filtering maps can be obtained.

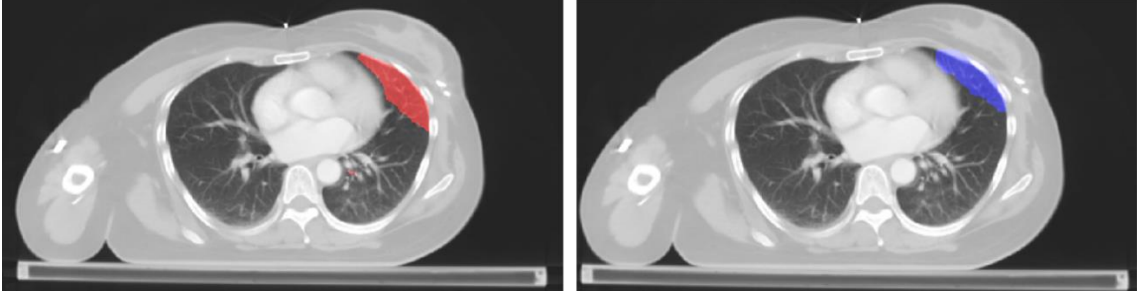
**Table 2. Radiomic and Dosiomic Features in EDOF**

R: Radiomic, D: Dosiomic

Class	Feature Name	Class	Feature Name
Intensity-based Features	D-Mean	GLRLM-based Features	R-Short Run Emphasis
	D-Variance		R-Long Run Emphasis
	D-Skewness		R-Gray Level Non-uniformity
	D-Intensity histogram kurtosis		R-Gray Level Non-uniformity Normalized
	D-Median		R-Run Length Non-uniformity
	D-Minimum grey level		R-Run Length Non-uniformity Normalized
	D-10th percentile		R-Run Percentage
	D-90th percentile		R-Low Gray Level Run Emphasis
	D-Maximum grey level		R-High Gray Level Run Emphasis
	D-Interquartile range		R-Short Run Low Gray Level Emphasis
	D-Range		R-Short Run High Gray Level Emphasis
	D-Mean absolute deviation		R-Long Run Low Gray Level Emphasis
	D-Robust mean absolute deviation		R-Long Run High Gray Level Emphasis
	D-Median absolute deviation		R-Grey Level Variance
	D-Coefficient of variation		R-Run Length Variance
	D-Quartile coefficient of dispersion		R-Run Entropy
	D-Energy		R-Small Zone Emphasis
D-Root mean square	R-Large Zone Emphasis		
GLCOM-based Features	R-Auto Correlation	GLSZM-based Features	R-Gray Level Non-uniformity
	R-Cluster Prominence		R-Gray Level Non-uniformity Normalized
	R-Cluster Shade		R-Size Zone Non-uniformity
	R-Cluster Tendency		R-Size Zone Non-uniformity Normalized
	R-Contrast		R-Zone Percentage
	R-Correlation		R-Low Gray Level Size Emphasis
	R-Differential Entropy		R-High Gray Level Size Emphasis
	R-Dissimilarity		R-Small Size Low Gray Level Emphasis
	R-Joint Energy / Angular Second Moment		R-Small Size High Gray Level Emphasis
	R-Joint Entropy		R-Large Size Low Gray Level Emphasis
	R-Homogeneity 1 / Inverse Difference		R-Large Size High Gray Level Emphasis
	R-Homogeneity 2 / Inverse Difference Moment		R-Gray Level Variance
	R-Info Measure Correlation 1		R-Zone Size Variance
	R-Info Measure Correlation 2		R-Zone Size Entropy
	R-Inverse Difference Moment Normalized		
	R-Inverse Difference Normalized		
	R-Inverse Variance		
	R-Joint maximum		
	R-Sum Average		
	R-Sum Entropy		
R-Sum Variance			
R-Joint Variance			

### 2.2.3 Explainable RP Prediction

To predict RP in breast cancer patients, the EBM model is summarized in Figure 3(C).



**Figure 4. Visualization of voxel-wise RP prediction from EBM model (image on the left) and RP ground truth (image on the right)**

Specifically, three-dimensional radiomic and dosiomic features were input into the EBM model. The predicted RP regions are visualized in red, while the ground truth RP regions are shown in blue, as illustrated in Figure 4.

EBM is a type of GAM that offers several advantages over traditional linear and multiple linear regression models. Unlike these conventional models, GAMs do not assume a linear relationship between predictor features  $x$  and the response variable  $y$ . Instead, GAMs predict  $y$  by learning an intercept  $\beta_0$  along with functions  $f_i(x_i)$  that describe the non-linear relationships between  $y$  and each predictor feature  $x_i$ , as shown in the equation (1).

$$E_y = \beta_0 + \sum f_i(x_i)$$

The coefficients in a multiple linear regression model are replaced with learned functions  $f_i$  that are not restricted to linear relationships. This approach allows GAMs to capture complex, non-linear dependencies between the predictors and the outcome.

In the EBM framework, the functions  $f_i$  are learned independently for each predictor feature  $x_i$ , enabling the model to separately explain the importance of each feature. This feature-specific learning facilitates a clear understanding of the overall contribution of each predictor to

the model. The mean value of the sum of  $|f_i(x_i)|$  across all samples, known as EBM's mean absolute score, which quantifies the importance of each predictor feature  $x_i$ . This scoring method provides insights into which features have the most significant impact on RP prediction. Additionally, EBM offers the feature  $x$  versus target  $y$  relationship (i.e., EBM's response function curve), allowing for a detailed examination of how changes in each predictor feature affect the predicted outcome. To derive the final binary classification results for RP, a link function was employed to convert the continuous predictions of the EBM into binary outcomes.

EBMs utilize a cyclic gradient boosting algorithm to iteratively learn shape functions, enabling accurate and interpretable models. The process starts by setting the model to use the average value of the response variable for regression tasks or classification, providing a baseline. The model then cycles through each predictor  $x_i$ , fitting a function  $f_i(x_i)$  to minimize the residual error from current predictions. These updates are incorporated into the model through boosting, where the fitted functions are added with a small learning rate to ensure gradual refinement. Periodically, the model detects significant interactions between predictors, fitting interaction terms  $f_{ij}(x_i, x_j)$  to capture their combined effects. This iterative process continues until the model converges, with performance stabilizing, indicating no further significant improvement. By employing this round-robin approach, EBMs mitigate issues like multicollinearity and ensure independent learning of each feature's contribution, balancing complexity and interpretability effectively. (Greenwell et al., 2023)

A notable method employed in EBMs for interaction detection is the Fast Interaction Detection (FAST) algorithm. The FAST algorithm in EBMs is designed to efficiently identify significant pairwise interactions between features. It evaluates all possible feature pairs, ranking them by their interaction strength, and incorporates only the most impactful interactions into the model. This selective approach reduces computational overfitting while maintaining the

interpretability and accuracy of the model. FAST uses a greedy block selection strategy within a block coordinate descent framework to achieve speed and scalability in high-dimensional datasets. A greedy block is the block that shows the greatest potential improvement to the objective function and is selected for optimization in the current iteration. By focusing on these impactful blocks, this approach enhances computational efficiency and accelerates convergence while maintaining accuracy. (Thanei et al., 2018)

Collectively, we proposed the EDOF model. For tomographic coordinate within the lung, 26 independent radiomic features and 5 independent dosiomics features were combined as a feature vector. The combined vector was subsequently fed into the EBM model to produce the enabling the locoregional segmentation of RP volume at a voxel-wise level. The voxel-wise RP volumes identified from follow-up CT scans serving as the classification ground truth.

#### **2.2.4 Feature Importance and Partial Dependence Plot**

In the context of the EBM model, feature importance is quantified by evaluating the mean absolute score of each predictor feature  $x_i$ . This score is calculated as the average of the absolute values of the learned function  $f_i(x_i)$  across all samples, providing a measure of each feature's contribution to the model's predictions. Features with higher mean absolute scores are considered more influential in predicting RP in breast cancer patients. (Nori et al., 2021)

Partial Dependence Plots (PDPs) are utilized to visualize the relationship between individual predictor features and the predicted outcome. A PDP illustrates how the predicted probability of RP changes as a specific feature  $x_i$  varies, while all other features are held constant. This visualization aids in understanding the effect of each feature on the model's predictions and can reveal non-linear dependencies that the EBM captures. In our study, the score of PDPs can be understood as the probability of developing RP. (Friedman, 2001)

### ***2.3 Training Details***

A cohort of 72 breast cancer patients was divided into training and testing sets using an 8:2 ratio through a 5-fold cross-validation method. In each fold, 80% of the patients were randomly assigned to the training set, while the remaining 20% constituted the testing set. This cross-validation process was repeated five times, ensuring that each patient was included in both the training and testing sets across different folds. This approach minimizes selection bias and allows for a robust assessment of the model's performance.

The radiomic and dosiomic filtering was implemented using an in-house developed toolbox with MATLAB (MATLAB R2023a; MathWorks, Natick, Mass). The toolbox was fully validated against the IBSI standardization as well as digital phantoms. (Zwanenburg et al., 2020) Additionally, the toolbox was specifically optimized for voxel-based, rotationally invariant, 3D calculations. The EBM was implemented in a Python environment using the InterpretML library. (Nori et al., 2019) All calculations were performed on a computational workstation equipped with a 32-core Intel Core i9-13700KF CPU at 4.0 GHz, 16 GB of RAM, and an Nvidia GeForce RTX 4090 graphics card.

### ***2.4 Evaluation Metrics***

The voxel-wise binarized RP prediction results were evaluated using Sensitivity, Specificity, Accuracy, and Area Under the ROC Curve (AUC). Performance metrics from each fold were aggregated to evaluate the overall effectiveness of the predictive model. Additionally, Grade-II RP patients typically present with larger RP volumes. The Dice Similarity Coefficient (DSC) was thus calculated for Grade-II RP patients to provide a more clinically meaningful evaluation of the model's performance.

### 2.4.1 AUC

The AUC is a widely adopted metric for assessing the performance of binary classifiers. Key terms include True Positives (TP) and False Positives (FP). True Positives are the cases where the model correctly identifies a positive outcome, while False Positives are the cases where the model incorrectly identifies a negative outcome as positive.

True Positive Rate (TPR), also known as Sensitivity or Recall, is the proportion of actual positive cases that are correctly identified by the model.

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

False Positive Rate (FPR) is the proportion of actual negative cases that are incorrectly identified as positive by the model.

$$FPR = \frac{FP}{FP + TN}$$

The Receiver Operating Characteristic (ROC) curve is plotted using TPR on the y-axis and FPR on the x-axis, across various threshold settings. The AUC measures the overall ability of the model to discriminate between positive and negative cases by calculating the area under this curve. A higher AUC value indicates better discriminative ability, with a value of 1 representing perfect discrimination and 0.5 indicating no discrimination.

### 2.4.2 Sensitivity, Specificity, Accuracy

Sensitivity (also known as recall) measures the proportion of true positive cases that are correctly identified by the model.

$$Specificity = \frac{TN}{TN + FP}$$

Specificity measures the proportion of true negative cases that are correctly identified.

$$Specificity = \frac{TN}{TN + FP}$$

Accuracy, on the other hand, evaluates the overall correctness of the model by calculating the proportion of all correctly classified cases (both true positives and true negatives) out of the total number of cases.

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

These metrics are crucial for understanding the model's ability to predict both positive and negative outcomes accurately.

### 2.4.3 DSC

The Dice Similarity Coefficient (DSC) is used to quantify the similarity between predicted RP regions and ground truth RP regions. It is particularly useful for evaluating segmentation tasks, such as predicting Grade-II RP, where accurate volumetric overlap is important. A higher DSC value indicates a better overlap between the predicted and actual RP regions, providing a more clinically relevant assessment of the model's segmentation performance.

$$Dice = \frac{2 \times TP}{TP + FP + TP + FN}$$

## 2.5 Comparison Study

To evaluate the contributions of each component of the EDOF model, ablation studies were conducted. These studies systematically removed dosiomic filtering, radiomic filtering, and EBM components from the EDOF, forming the following three variants:

**RF Model:** In the first variant, the dosiomic filtering component was excluded, and only the radiomic filtering features were used as input to the EBM. This variant determined the impact of removing the detailed locoregional dosimetric information.

**DF Model:** In the second variant, the radiomic filtering component was excluded, and only the dosiomic filtering features were utilized as input to the EBM. This variant assessed the significance of radiomic features derived from lung CT images in capturing the textural and structural variations that contribute to RP risk.

Each of these variants was evaluated and compared to the EDOF model to assess their relative performance. In these variants, the training settings, including five-fold cross-validation and training/test set assignment, were kept the same as the proposed EDOF model. The achieved voxel-wise RP prediction accuracy (sensitivity, specificity, accuracy, AUC, and DSC) was compared using the Wilcoxon signed-rank test. The statistical significance level was set at 0.05.

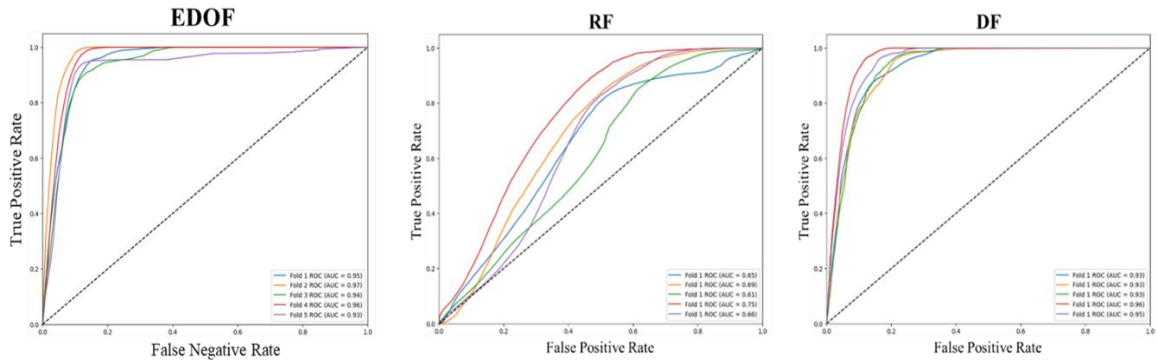
### 3. Results

#### 3.1 Evaluation results

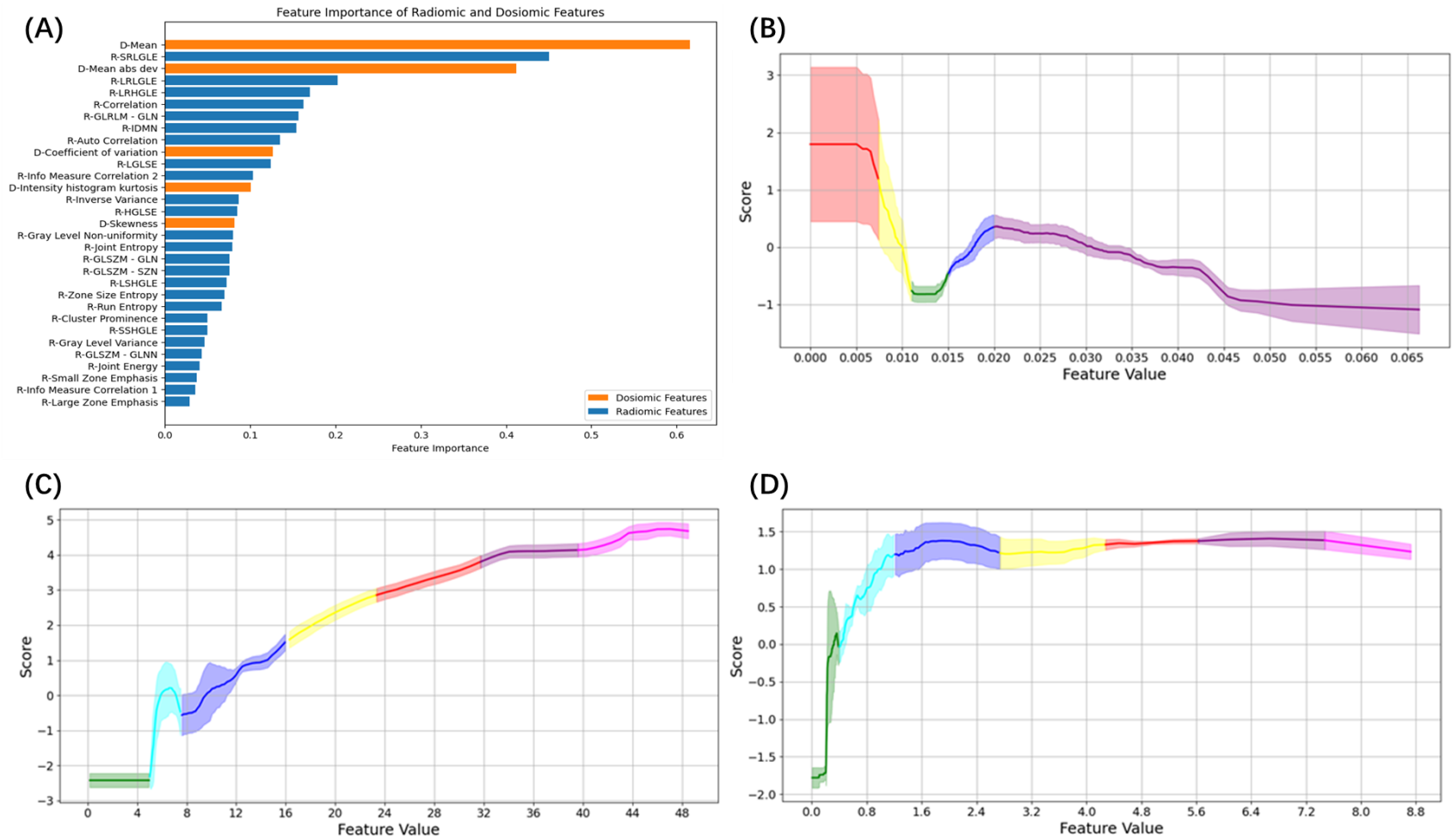
**Table 3. Evaluation Results of Comparative Studies**

Model	AUC	Accuracy	Sensitivity	Specificity	Dice Coefficient
EDOF Model	0.95±0.01	0.93±0.02	0.81±0.05	0.93±0.01	0.78±0.07
RF Model	0.68±0.07	0.89±0.03	0.01±0.01	0.85±0.02	0.04±0.03
DF Model	0.94±0.01	0.90±0.03	0.71±0.03	0.92±0.02	0.70±0.08

The EDOF model demonstrated the highest performance in RP prediction for breast cancer patients among all models in the comparative studies (Table 3). The ROC curves derived from these studies are presented in Figure 5. The EDOF model achieved a mean accuracy = 0.93,  $ROC_{AUC} = 0.95$ , mean sensitivity = 0.81, mean specificity = 0.93, and mean dice coefficient (grade II) = 0.78. The RF model achieved limited performance (with a mean accuracy = 0.89,  $ROC_{AUC} = 0.68$ , mean sensitivity = 0.01, mean specificity = 0.85) compared to the EDOF model. The DF model achieved a comparable AUC and specificity with the EDOF model, but still lower than the EDOF in accuracy, sensitivity, and Dice coefficient.



**Figure 5. ROC curves from Comparative Studies**



**Figure 6. Feature Importance and Partial Dependence Plot. (A) Feature Importance Plot. (B) R-SRLGLE PDP. (C) D-Mean PDP. (D) D-Mean Absolute deviation PDP.**

### ***3.2 Feature Importance and Partial Dependence Plot***

As shown in Figure 6(a), there are three features identified as key to RP development:

- 1) R-SRLGLE (Short Run Low Gray Level Emphasis) – a texture feature that describes the heterogeneity of tissues within the lung CT image kernel. The larger feature value corresponds to lower heterogeneity. The PDP shows that as R-SRLGLE values increase, the probability of RP generally decreases. However, A local minimum in the range of 0.011–0.014 reveals a non-linear relationship that highlights variations in RP probability within this specific range.
- 2) D-Mean – the mean value within the dose distribution kernel. It quantifies the average radiation dose delivered to a localized region. The analysis shows that between 0–5 Gy, RP probability remains very low, while from 5–8 Gy, there is a sharp increase in probability. Beyond 8 Gy, the RP probability rises approximately linearly with increasing doses, indicating a clear dose-response relationship.
- 3) D-Mean Absolute Deviation – the mean absolute deviation value within the dose distribution kernel. The RP probability increases sharply for dose deviations between 0–1 Gy, indicating a high sensitivity to small variations. Beyond 1 Gy, RP probability stabilizes, suggesting a consistent effect of larger dose deviations.

Collectively, these findings highlight the dose-dependent nature of RP risk and the influence of both dosimetric and radiomic features.

## 4. Discussion

### *4.1 Technical Insights and Feature Analysis*

Our voxel-wise RP prediction model leverages the interplay between locoregional dose distribution and radiomic tissue heterogeneity, providing a comprehensive profiling of RP risk regions. On pre-treatment CT scans, the proximity to high-dose areas and heterogeneous lung tissue was quantitatively characterized using spatially encoded dual-omics features. RP localization was subsequently refined through integrated dual filtering and voxel-wise EBM classification. On follow-up CT scans within four months post-treatment, our model successfully delineated locoregional RP regions that are consistent in location with radiologically confirmed lesions.

Traditional methods for RP prediction rely on dose-volume metrics such as V5, V20, and MLD. These metrics primarily provide global information, often overlooking localized high-risk regions. In contrast, our EDOF model leverages voxel-wise radiomic and dosiomic features to capture localized dose-response relationships, offering a more detailed understanding of RP risk. The integration of voxel-wise radiomic and dosiomic features allows the EDOF model to effectively identify spatial variations in tissue response to radiation. Radiomic features quantify tissue heterogeneity, while dosiomic features capture the variations in dose intensity. These combined features offer a framework for predicting RP by reflecting anatomical variability and dose distribution patterns. Our method uses deformable image registration to align with anatomical structures across patients, ensuring accurate extraction of these features, which is critical for understanding localized radiation effects.

The feature importance analysis indicates that both radiomic and dosiomic features significantly contribute to RP prediction and reveal key features. The PDPs of these key features provide important insights into their relationships with RP development. For the R-SRLGLE feature, RP tends to develop in heterogeneous lung regions. This radiomic feature reflects lung

texture patterns and may detect subtle inhomogeneous perfusion. These perfusion abnormalities may indicate pre-existing lung conditions, increasing the risk of RP. Another important factor is electron disequilibrium in low-density lung tissue. In these regions, electrons travel farther before depositing energy, leading to fewer electrons depositing dose in the intended area. This effect can result in lower actual dose than planned, potentially decreasing the RP probability (Disher et al., 2013). As for the local minimum of R-SRLGLE's PDP, tissues within such range are primarily near the lung margin. Respiratory motion during radiotherapy causes the lung margin to move up and down, leading to a lower actual dose than planned, thus reducing RP probability (van der Heyden et al., 2017). For the D-Mean feature, the probability of RP increases linearly with mean dose, with patients more likely to develop RP in high-dose regions. Higher radiation increases the risk of RP due to direct damage to lung cells and microvasculature, leading to inflammation and tissue fibrosis. High doses trigger the release of pro-inflammatory cytokines (e.g., TNF- $\alpha$ , IL-6, TGF- $\beta$ ) and immune cell activation, amplifying the inflammatory response. Additionally, the lung's repair capacity becomes overwhelmed at higher doses, resulting in irreversible damage and greater susceptibility to RP (Käsmann et al., 2020). Clinical studies consistently show a strong dose-response relationship, with parameters like MLD and V20 strongly correlated with RP risk (Hart et al., 2008; Wang et al., 2013). This model also highlights the importance of monitoring mean doses in the 5-8 Gy range, where the probability of RP sharply increases. In the D-Mean Absolute Deviation PDP, the probability of RP rises sharply with dose deviations from 0-1 Gy, indicating sensitivity to initial deviations. Beyond this range, the probability plateaus, suggesting stabilization in RP risk despite further deviations.

## ***4.2 Clinical Implications and Limitations***

The explainability of the EDOF model is a key strength, distinguishing it from traditional black-box machine learning models. The explainability is demonstrated through its ability to rank feature importance and explain how RP risk changes as a function of individual feature values. This relationship is illustrated by PDPs, which visually capture how specific dose metrics or tissue characteristics influence RP probability. The trends observed in PDPs highlight the relationship between specific features and RP risk, where sudden increases or decreases can provide critical clinical implications. For instance, abrupt surges in RP risk indicate regions where caution is needed during treatment planning. This insight ensures that high-risk zones are carefully evaluated, supporting more precise and personalized radiotherapy strategies.

The EDOF model demonstrates high practicality for clinical applications. Before undergoing radiotherapy, breast cancer patients routinely undergo pretreatment CT scans and dose distribution calculations. Radiomic and dosiomic features can be extracted from these datasets and fed into the EDOF model for RP risk prediction. Since this is a voxel-wise analysis, the model generates resolved RP risk maps, identifying high-risk regions and their extent. Clinicians can use this information to assess RP risk. Furthermore, medical physicists can refine treatment plans based on key predictors: dose mean, dose mean absolute deviation, and tissue heterogeneity. By iteratively adjusting these factors and re-evaluating the updated treatment plan with the EDOF model, they can achieve a more optimized balance between tumor control and RP risk reduction. Compared to traditional RP risk metrics such as V20 and MLD, this approach provides more modifiable parameters. As a result, RP risk can be improved without significantly compromising breast cancer treatment efficacy, ultimately improving patients' quality of life.

This study has several limitations that should be considered. Firstly, the small sample size of 72 breast cancer patients may limit the generalizability of the findings. A larger, more diverse

dataset, ideally from multiple institutions, would enhance the robustness and reliability of the EDOF model's predictive capabilities, and provide insights into how demographic and clinical factors, such as age or treatment history, may influence its performance. Secondly, the study lacks a direct comparison between the EDOF model and other well-established machine learning or deep learning models, such as SVMs, RF, or CNN. Comparative studies could clarify whether the EDOF model offers advantages in terms of accuracy, interpretability, or computational efficiency, particularly in the context of complex medical imaging data. Addressing these aspects would contribute to a better understanding of the EDOF model's role in the broader field of predictive modeling in radiation oncology.

## **5. Conclusion**

This study successfully developed the EDOF model to predict RP regions based on voxel-wise radiomic and dosiomic data, offering a precise approach to identifying areas at risk. The model also identified key features and revealed how these features impact RP development, providing valuable clinical implications for optimizing treatment planning. These insights underscore the potential of the EDOF model to support more targeted and personalized radiation therapy, advancing predictive capabilities in the management of RP in breast cancer patients.

## References

- Alahmadi, R., Almujiabah, H., Alotaibi, S., Elshekh, A., Alsharif, M., & Almakki, M. (2023). Explainable Boosting Machine: A Contemporary Glass-Box Model to Analyze Work Zone-Related Road Traffic Crashes. *Safety*, 9, 83. <https://doi.org/10.3390/safety9040083>
- Arnold, M., Morgan, E., Rungay, H., Mafra, A., Singh, D., Laversanne, M., Vignat, J., Gralow, J. R., Cardoso, F., Siesling, S., & Soerjomataram, I. (2022). Current and future burden of breast cancer: Global statistics for 2020 and 2040. *The Breast: Official Journal of the European Society of Mastology*, 66, 15–23. <https://doi.org/10.1016/j.breast.2022.08.010>
- Asif, S., Wenhui, Y., ur-Rehman, S., Ul-Ain, Q., Amjad, K., Yueyang, Y., Jinhai, S., & Awais, M. (2024). Advancements and Prospects of Machine Learning in Medical Diagnostics: Unveiling the Future of Diagnostic Precision. *Archives of Computational Methods in Engineering*, 1–31. <https://doi.org/10.1007/s11831-024-10148-w>
- Bi, J., Qian, J., Yang, D., Sun, L., Lin, S., Li, Y., Xue, X., Nie, T., Verma, V., & Han, G. (2022). Dosimetric Risk Factors for Acute Radiation Pneumonitis in Patients With Prior Receipt of Immune Checkpoint Inhibitors. *Frontiers in Immunology*, 12. <https://doi.org/10.3389/fimmu.2021.828858>
- Boonyawan, K., Gomez, D. R., Komaki, R., Xu, Y., Nantavithya, C., Allen, P. K., Mohan, R., & Liao, Z. (2018). Clinical and Dosimetric Factors Predicting Grade  $\geq 2$  Radiation Pneumonitis After Postoperative Radiotherapy for Patients With Non-Small Cell Lung Carcinoma. *International Journal of Radiation Oncology, Biology, Physics*, 101(4), 919–926. <https://doi.org/10.1016/j.ijrobp.2018.04.012>
- Chen, F., Niu, J., Wang, M., Zhu, H., & Guo, Z. (2023). Re-evaluating the risk factors for radiation pneumonitis in the era of immunotherapy. *Journal of Translational Medicine*, 21(1), 368. <https://doi.org/10.1186/s12967-023-04212-5>
- Disher, B., Hajdok, G., Gaede, S., Mulligan, M., & Battista, J. J. (2013). Forcing lateral electron disequilibrium to spare lung tissue: A novel technique for stereotactic body radiation therapy of lung cancer. *Physics in Medicine and Biology*, 58(19), 6641–6662. <https://doi.org/10.1088/0031-9155/58/19/6641>
- Feng, A., Huang, Y., Zeng, Y., Shao, Y., Wang, H., Chen, H., Gu, H., Duan, Y., Shen, Z., & Xu, Z. (2024). Improvement of Prediction Performance for Radiation Pneumonitis by Using 3-Dimensional Dosimetric Features. *Clinical Lung Cancer*, 25(4), e173-e180.e2. <https://doi.org/10.1016/j.clcc.2024.01.006>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Giaquinto, A. N., Sung, H., Miller, K. D., Kramer, J. L., Newman, L. A., Minihan, A., Jemal, A., & Siegel, R. L. (n.d.). *Breast Cancer Statistics, 2022*. <https://doi.org/10.3322/caac.21754>

- Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images Are More than Pictures, They Are Data. *Radiology*, 278(2), 563–577. <https://doi.org/10.1148/radiol.2015151169>
- Greenwell, B. M., Dahlmann, A., & Dhoble, S. (2023). *Explainable Boosting Machines with Sparsity—Maintaining Explainability in High-Dimensional Settings* (arXiv:2311.07452). arXiv. <https://doi.org/10.48550/arXiv.2311.07452>
- Hanania, A. N., Mainwaring, W., Ghebre, Y. T., Hanania, N. A., & Ludwig, M. (2019). Radiation-Induced Lung Injury: Assessment and Management. *CHEST*, 156(1), 150–162. <https://doi.org/10.1016/j.chest.2019.03.033>
- Hart, J. P., McCurdy, M. R., Ezhil, M., Wei, W., Khan, M., Luo, D., Munden, R. F., Johnson, V. E., & Guerrero, T. M. (2008). Radiation Pneumonitis: Correlation of Toxicity with the Pulmonary Metabolic Radiation Response. *International Journal of Radiation Oncology, Biology, Physics*, 71(4), 967. <https://doi.org/10.1016/j.ijrobp.2008.04.002>
- Huang, J.-W., Lin, Y.-H., Chang, G.-C., & Chen, J. J. W. (2023). A novel tool to evaluate and quantify radiation pneumonitis: A retrospective analysis of correlation of dosimetric parameters with volume of pneumonia patch. *Frontiers in Oncology*, 13. <https://doi.org/10.3389/fonc.2023.1130406>
- Huang, Y., Feng, A., Lin, Y., Gu, H., Chen, H., Wang, H., Shao, Y., Duan, Y., Zhuo, W., & Xu, Z. (2022). Radiation pneumonitis prediction after stereotactic body radiation therapy based on 3D dose distribution: Dosiomics and/or deep learning-based radiomics features. *Radiation Oncology*, 17(1), 188. <https://doi.org/10.1186/s13014-022-02154-8>
- Käsmann, L., Dietrich, A., Staab-Weijnitz, C. A., Manapov, F., Behr, J., Rimner, A., Jeremic, B., Senan, S., De Ruyscher, D., Lauber, K., & Belka, C. (2020). Radiation-induced lung toxicity – cellular and molecular mechanisms of pathogenesis, management, and literature review. *Radiation Oncology*, 15(1), 214. <https://doi.org/10.1186/s13014-020-01654-9>
- Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G. P. M., Granton, P., Zegers, C. M. L., Gillies, R., Boellard, R., Dekker, A., & Aerts, H. J. W. L. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer*, 48(4), 441–446. <https://doi.org/10.1016/j.ejca.2011.11.036>
- Lei, S., Zheng, R., Zhang, S., Chen, R., Wang, S., Sun, K., Zeng, H., Wei, W., & He, J. (2021). Breast cancer incidence and mortality in women in China: Temporal trends and projections to 2030. *Cancer Biology & Medicine*, 18(3), 900. <https://doi.org/10.20892/j.issn.2095-3941.2020.0523>
- Liang, B., Yan, H., Tian, Y., Chen, X., Yan, L., Zhang, T., Zhou, Z., Wang, L., & Dai, J. (2019). Dosiomics: Extracting 3D Spatial Features From Dose Distribution to Predict Incidence of Radiation Pneumonitis. *Frontiers in Oncology*, 9. <https://doi.org/10.3389/fonc.2019.00269>
- Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30.

[https://proceedings.neurips.cc/paper\\_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html](https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html)

- M. D. Anderson Cancer Center Head and Neck Quantitative Imaging Working Group. (2018). Investigation of radiomic signatures for local recurrence using primary tumor texture analysis in oropharyngeal head and neck cancer patients. *Scientific Reports*, 8(1), 1524. <https://doi.org/10.1038/s41598-017-14687-0>
- Ma, F., He, C., Yang, H., Hu, Z., Mao, H., Fan, C., Qi, Y., Zhang, J., & Xu, B. (2023). Interpretable machine-learning model for Predicting the Convalescent COVID-19 patients with pulmonary diffusing capacity impairment. *BMC Medical Informatics and Decision Making*, 23(1), 169. <https://doi.org/10.1186/s12911-023-02192-6>
- Maliko, N., Stam, M. R., Boersma, L. J., Vrancken Peeters, M.-J. T. F. D., Wouters, M. W. J. M., KleinJan, E., Mulder, M., Essers, M., Hurkmans, C. W., & Bijker, N. (2022). Transparency in quality of radiotherapy for breast cancer in the Netherlands: A national registration of radiotherapy-parameters. *Radiation Oncology*, 17(1), 73. <https://doi.org/10.1186/s13014-022-02043-0>
- Maxwell, A. E., Sharma, M., & Donaldson, K. A. (2021). Explainable Boosting Machines for Slope Failure Spatial Predictive Modeling. *Remote Sensing*, 13(24), Article 24. <https://doi.org/10.3390/rs13244991>
- Momenimovahed, Z., & Salehiniya, H. (2019). Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer: Targets and Therapy*, Volume 11, 151–164. <https://doi.org/10.2147/BCTT.S176070>
- Nie, T., Chen, Z., Cai, J., Ai, S., Xue, X., Yuan, M., Li, C., Shi, L., Liu, Y., Verma, V., Bi, J., Han, G., & Yuan, Z. (2024). Integration of dosimetric parameters, clinical factors, and radiomics to predict symptomatic radiation pneumonitis in lung cancer patients undergoing combined immunotherapy and radiotherapy. *Radiotherapy and Oncology*, 190, 110047. <https://doi.org/10.1016/j.radonc.2023.110047>
- Nori, H., Caruana, R., Bu, Z., Shen, J. H., & Kulkarni, J. (2021). Accuracy, Interpretability, and Differential Privacy via Explainable Boosting. *Proceedings of the 38th International Conference on Machine Learning*, 8227–8237. <https://proceedings.mlr.press/v139/nori21a.html>
- Nori, H., Jenkins, S., Koch, P., & Caruana, R. (2019). *InterpretML: A Unified Framework for Machine Learning Interpretability*. <https://www.microsoft.com/en-us/research/publication/interpretml-a-unified-framework-for-machine-learning-interpretability/>
- Otani, K., Seo, Y., & Ogawa, K. (2017). Radiation-Induced Organizing Pneumonia: A Characteristic Disease that Requires Symptom-Oriented Management. *International Journal of Molecular Sciences*, 18(2), 281. <https://doi.org/10.3390/ijms18020281>

- Ozgen, Z., Orun, O., Atasoy, B. M., Mega Tiber, P., Akdeniz, E., Cimsit, C., Eryuksel, E., & Karakurt, S. (2023). Radiation pneumonitis in relation to pulmonary function, dosimetric factors, TGFβ1 expression, and quality of life in breast cancer patients receiving post-operative radiotherapy: A prospective 6-month follow-up study. *Clinical and Translational Oncology*, 25(5), 1287–1296. <https://doi.org/10.1007/s12094-022-03024-1>
- P, D. N., C, L., M, G., S, B., M, A., R, M., & F, P. (2022). Chemotherapy in patients with early breast cancer: Clinical overview and management of long-term side effects. *Expert Opinion on Drug Safety*, 21(11). <https://doi.org/10.1080/14740338.2022.2151584>
- Petit, S. F., Van Elmpt, W. J. C., Oberije, C. J. G., Vegt, E., Dingemans, A.-M. C., Lambin, P., Dekker, A. L. A. J., & De Ruyscher, D. (2011). [18F]fluorodeoxyglucose Uptake Patterns in Lung Before Radiotherapy Identify Areas More Susceptible to Radiation-Induced Lung Toxicity in Non-Small-Cell Lung Cancer Patients. *International Journal of Radiation Oncology\*Biophysics*, 81(3), 698–705. <https://doi.org/10.1016/j.ijrobp.2010.06.016>
- Pinnix, C. C., Smith, G. L., Milgrom, S., Osborne, E. M., Reddy, J. P., Akhtari, M., Reed, V., Arzu, I., Allen, P. K., Wogan, C. F., Fanale, M. A., Oki, Y., Turturro, F., Romaguera, J., Fayad, L., Fowler, N., Westin, J., Nastoupil, L., Hagemester, F. B., ... Dabaja, B. (2015). Predictors of Radiation Pneumonitis in Patients Receiving Intensity Modulated Radiation Therapy for Hodgkin and Non-Hodgkin Lymphoma. *International Journal of Radiation Oncology, Biology, Physics*, 92(1), 175–182. <https://doi.org/10.1016/j.ijrobp.2015.02.010>
- Puttanawarut, C., Sirirutbunkajorn, N., Tawong, N., Jiarpinitnun, C., Khachonkham, S., Pattaranutaporn, P., & Wongsawat, Y. (2022). Radiomic and Dosiomic Features for the Prediction of Radiation Pneumonitis Across Esophageal Cancer and Lung Cancer. *Frontiers in Oncology*, 12. <https://doi.org/10.3389/fonc.2022.768152>
- Rana, M., & Bhushan, M. (2023). Machine learning and deep learning approach for medical image analysis: Diagnosis to detection. *Multimedia Tools and Applications*, 82(17), 26731–26769. <https://doi.org/10.1007/s11042-022-14305-w>
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- Shubeck, S. P., Morrow, M., & Dossett, L. A. (2022). De-escalation in breast cancer surgery. *Npj Breast Cancer*, 8(1), 1–4. <https://doi.org/10.1038/s41523-022-00383-4>
- Street, W. (n.d.). *Breast Cancer Facts & Figures 2019-2020*.
- Thanei, G.-A., Meinshausen, N., & Shah, R. D. (2018). The xyz algorithm for fast interaction search in high-dimensional data. *Journal of Machine Learning Research*, 19(37), 1–42.
- van der Heyden, B., van Hoof, S. J., Schyns, L. E. J. R., & Verhaegen, F. (2017). The influence of respiratory motion on dose delivery in a mouse lung tumour irradiation using the 4D

- MOBY phantom. *The British Journal of Radiology*, 90(1069), 20160419. <https://doi.org/10.1259/bjr.20160419>
- van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillion-Robin, J.-C., Pieper, S., & Aerts, H. J. W. L. (2017). Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*, 77(21), e104–e107. <https://doi.org/10.1158/0008-5472.CAN-17-0339>
- Wang, W., Xu, Y., Schipper, M., Matuszak, M. M., Ritter, T., Cao, Y., Haken, R. K. T., & Kong, F.-M. (Spring). (2013). Effect of Normal Lung Definition on Lung Dosimetry and Lung Toxicity Prediction in Radiation Therapy Treatment Planning. *International Journal of Radiation Oncology, Biology, Physics*, 86(5), 956. <https://doi.org/10.1016/j.ijrobp.2013.05.003>
- WHO. (n.d.). *Breast cancer*. Retrieved September 29, 2024, from <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>
- Xu, Y., Gong, M., Wang, Y., Yang, Y., Liu, S., & Zeng, Q. (2023). Global trends and forecasts of breast cancer incidence and deaths. *Scientific Data*, 10(1), 334. <https://doi.org/10.1038/s41597-023-02253-5>
- Z, Z., Z, W., M, Y., J, Y., A, D., L, Z., & L, W. (2023). Radiomics and Dosiomics Signature From Whole Lung Predicts Radiation Pneumonitis: A Model Development Study With Prospective External Validation and Decision-curve Analysis. *International Journal of Radiation Oncology, Biology, Physics*, 115(3). <https://doi.org/10.1016/j.ijrobp.2022.08.047>
- Zampieri, F. G., Salluh, J. I. F., Azevedo, L. C. P., Kahn, J. M., Damiani, L. P., Borges, L. P., Viana, W. N., Costa, R., Corrêa, T. D., Araya, D. E. S., Maia, M. O., Ferez, M. A., Carvalho, A. G. R., Knibel, M. F., Melo, U. O., Santino, M. S., Lisboa, T., Caser, E. B., Besen, B. A. M. P., ... the ORCHESTRA Study Investigators. (2019). ICU staffing feature phenotypes and their relationship with patients' outcomes: An unsupervised machine learning analysis. *Intensive Care Medicine*, 45(11), 1599–1607. <https://doi.org/10.1007/s00134-019-05790-z>
- Zhang, X., Ge, X., Jiang, T., Yang, R., & Li, S. (2022). Research progress on immunotherapy in triple-negative breast cancer (Review). *International Journal of Oncology*, 61(2), 95. <https://doi.org/10.3892/ijo.2022.5385>
- Zwanenburg, A., Leger, S., Vallières, M., & Löck, S. (2020). Image biomarker standardisation initiative. *Radiology*, 295(2), 328–338. <https://doi.org/10.1148/radiol.2020191145>