scientific reports

OPEN

Radiomics early assessment of post chemotherapy cardiotoxicity in cancer patients using 2D echocardiography imaging an interpretable machine learning study

Masumeh Ahmadi[1,8], Erfan Barzegar-Golmoghani[2,9], Amir Ghaffari Jolfayi[3,9], Mobin Mohebi[2], Azin Alizadehasl[4,5], Mina Mohseni[6] & Ahmad Bitarafan-Rajabi[5,7,8]✉

Cardiotoxicity is the loss of the heart muscle's ability to contract effectively, often due to chemotherapy or radiation therapy. This study uses interpretable machine learning to predict post-chemotherapy cardiotoxicity using radiomics features extracted from the baseline echocardiography images. The study involved 100 cancer patients at Rajaei Cardiovascular Medical and Research Center. Baseline echocardiography images were used to extract radiomics features such as the left ventricular ejection fraction. According to the 12-month follow-up echocardiography, cardiotoxicity is defined as EF decline. Machine learning models predicted chemotherapy-induced cardiotoxicity based on radiomics features, with their significance confirmed via SHapley Additive exPlanations (SHAP) and Permutation-based Feature Importance Test (PermFIT). Among the 100 patients with a mean age of 54.5 ± 13.7, 41 patients (41%) experienced cardiotoxicity. For the short-axis view, the K-nearest neighbors (KNN) and Linear Support Vector Machine (SVM) models achieved accuracies of 0.92 and 0.90, respectively, with the best outcome of 92%. For the 4-chamber view, SVM and KNN reached accuracies of 0.88 and 0.83. These findings underscore the potential of machine learning, especially using short-axis echocardiography, to enhance early diagnosis of cardiotoxicity in chemotherapy patients.

Keywords Chemotherapy, Cardiotoxicity, Echocardiography, Radiomics, Machine learning

Abbreviations

RVID	Right Ventricular Internal Dimension
LVIDd	Left Ventricular Internal Dimension in Diastole
LVIDs	Left Ventricular Internal Dimension in Systole
IVSD	Interventricular Septal Thickness in Diastole
LA volume	Left Atrial Area, Left Atrial Diameter, Left Atrial Volume
EDV	End-Diastolic Volume
ESV	End-Systolic Volume
TRG max	Tricuspid Regurgitation Gradient Maximum

[1]Department of Medical Physics, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran. [2]Department of Biomedical Engineering, Tarbiat Modares University, Tehran, Iran. [3]Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Sciences, Tehran, Iran, Iran University of Medical Sciences, Tehran, Iran. [4]Cardio-Oncology Research Center, Rajaie Cardiovascular Institute, Tehran, Iran. [5]Echocardiography Research Center, Rajaie Cardiovascular Institute, Tehran, Iran. [6]Cardiovascular Research Center, Institute of Basic and Clinical Physiology Sciences, Kerman University of Medical Sciences, Kerman, Iran. [7]Cardiovascular Intervention Research Center, Rajaie Cardiovascular Institute, Tehran, Iran. [8]Rajaie Cardiovascular Institute, Iran University of Medical Sciences, Tehran, Iran. [9]Erfan Barzegar-Golmoghani and Amir Ghaffari Jolfayi contributed equally to this work. ✉email: bitarafan@hotmail.com; dr.bitarafan@gmail.com

1

PAP	Pulmonary Artery Pressure
IVC	Inferior Vena Cava Diameter
TAPSE	Tricuspid Annular Plane Systolic Excursion
RVSP	Right Ventricular Systolic Pressure
SPAP	Systolic Pulmonary Artery Pressure
First follow GLS	Global Longitudinal Strain
LVSD	Left ventricular systolic dysfunction
LVEF	Left ventricular ejection fraction
ATC	Anthracyclines
aaHER2	Humanized monoclonal antibodies targeting HER2
CTRCD	Cancer-related cardiac dysfunction
PACS	Picture Archiving and Communication System
ML	Machine learning
ACC	Accuracy
SEN	Sensitivity
SPE	Specificity
PR	Precision
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
RUS boosted	Random Under-Sampling Boosted
RF	Random Forest
FPR	False positive rate

The World Health Organization (WHO) reported that cancer was the second leading cause of death for individuals under 70 age in 112 out of 183 countries in 2019. The global burden of cancer is expected to rise by 47% to 28.4 million cases in 2040[1]. In pursuit of this goal to enhance treatment effectiveness, various approaches like chemotherapy, radiation therapy, and immunotherapy have been studied for treating cancer patients[2]. Despite advancements in cancer therapies leading to a decrease in the mortality rate[3], these treatments have long-lasting side effects on the heart, directly impacting its structure and function, leading to heart disease and cardiotoxicity in cancer patients[4,5]. Anthracyclines (ATC) are crucial chemotherapy treatments for hematologic and solid cancers, particularly breast cancer. ATC can cause irreparable heart damage, and as dosage increases, so does the risk of cardiotoxicity[6]. Trastuzumab and Pertuzumab, humanized monoclonal antibodies targeting HER2 (aaHER2), are also widely used in breast cancer and have been shown to cause cardiotoxicity[7,8]. Anti-HER2 cardiac effects are not dose-dependent, and left ventricular systolic dysfunction (LVSD) is usually reversible with TZ treatment suspension or withdrawal. However, anti-HER2 medications can worsen and accelerate the harm caused by earlier ATC therapy[9].

Chemotherapy-induced cardiotoxicity is defined as a 5–10% decrease in left ventricular ejection fraction (LVEF)[10] alongside other complications, including cardiac arrhythmias, pericarditis and irreversible heart damage[11].

Identifying cancer patients at high risk of cardiotoxicity is crucial for early detection and potential preventive measures, posing a significant challenge in clinical oncology[12]. Echocardiography, cardiac imaging, computed tomography of the heart, and cardiac magnetic resonance imaging are used to diagnose cardiac toxicity[13,14]. Echocardiography is the most widely used non-invasive method for screening cardiotoxicity[15,16].The inexpensive cost and widespread availability of this method are utilized to evaluate systolic and diastolic function heart valve abnormalities[17].

Two-dimensional echocardiography (2DE) is the preferred method for assessing cardiotoxicity in oncology patients, with LVEF being the most commonly used parameter[18].

Many studies have examined the effect of chemotherapy and radiotherapy on reduced LVEF[19] and reported that reduced LVEF before chemotherapy had a higher risk of subsequent cardiac events[20]. Despite the use of echocardiography in the diagnosis of cardiovascular diseases, some limitations such as the dependency on physician experience, low-quality and low-contrast images, and the estimation error of more than 14% can be noted[21,22].

Previous research suggests that Machine learning (ML)based approaches can improve models for predicting cancer therapy-related cardiac dysfunction[23,24] and automated echocardiography techniques using artificial intelligence (AI) can improve LVEF measurement accuracy and disease diagnosis[25,26]. However, its effectiveness can be hindered by interobserver variability and operator experience dependence. AI can address these challenges by enhancing consistency and reducing variance in image acquisition and interpretation. ML offers automated interpretations, reducing operator expertise, allowing echocardiography to accurately address cardiotoxicity and prompt early treatment initiation[27–29]. ML models have shown significant accuracy in predicting heart failure, atrial fibrillation, stroke, myocardial infarction, and cancer-related cardiac dysfunction (CTRCD)[18,24,30]. Studies explored radiomic features' role in disease prognosis and treatment response, aiming to enhance medical decisions using new image processing methods and ML data analysis[21,31–37] and there are limited radiomic studies to predict cardiotoxicity induced by chemotherapy.

Based on two study ML-based research should focus on the various complications that can occur after chemotherapy[38,39]. Two other studies have been performed on the clinical data of patients undergoing chemotherapy for colorectal cancer and breast cancer, respectively, showed that the development and comparison of different ML algorithms for predicting the risk of cardiotoxicity could predict better the decisions and prognosis of physicians[24,40].

Therefore, considering the current challenges in ML-based prediction of cardiotoxicity in patients undergoing cancer treatment, this study aimed to analyze the radiomic features of baseline pre-chemotherapy echocardiographic images to predict late cardiotoxicity 12 months post-chemotherapy. Echocardiographic assessments performed 12 months after chemotherapy were used to determine cardiotoxicity based on LVEF decline. Additionally, we developed an ML model to enhance the accuracy of cardiotoxicity prediction.

Materials and methods
Study population
Inclusion criteria

This retrospective cross-sectional study was designed to include a more homogeneous patient population by focusing on breast cancer patients who were not candidates for surgery and were receiving systemic chemotherapy as their primary treatment modality. This selection ensures that the study population represents individuals for whom chemotherapy is the standard-of-care treatment, thereby concentrating on the subgroup most at risk for chemotherapy-induced cardiotoxicity.

In addition, all eligible patients underwent standard chemotherapy regimens that commonly included anthracycline-based agents or anti-HER2 therapies (e.g., Herceptin). These treatments are widely used in the management of breast cancer and are known for their potential cardiotoxic effects, making them a critical focus for our radiomics assessment. Eligible records required documentation of at least one complete cycle of either anthracycline or anti-HER2 therapy, along with comprehensive clinical and imaging data to facilitate a robust analysis of treatment-related cardiac changes.

Exclusion criteria

The records of all patients newly diagnosed with cancer at the Cardio-Oncology Department of Rajaei Cardiovascular Medical and Research Center were evaluated. Patients with a previous history of chemotherapy-related cardiovascular disease, open-heart surgery, or taking medications that significantly impair heart function, as well as those with abnormal echocardiograms before initiating treatment, were excluded from the study.

Additionally, Patients who did not have complete echocardiography follow-up or whose echocardiographic images were of suboptimal quality were excluded from the study.

Furthermore, patients whose echocardiographic images quality were deemed suboptimal were also excluded. It is noteworthy that all these patients underwent echocardiography both before and 12 months after chemotherapy. The total number of patients in this investigation was 150. According to the inclusion and exclusion criteria mentioned in the Methods and Standards section, 100 patients were finally selected and evaluated for further analysis and processing.

Ejection fraction decline determination

Pre-chemotherapy, images were used to predict cardiotoxicity via radiomics features among the patients, while 12 months post-chemotherapy images were used to estimate the EF difference between two episodes of echocardiography to determine the EF decline and consequent cardiotoxicity.

Data acquisition

Transthoracic echocardiography was performed by two experienced Cardio-oncologists with at least 5 years of experience in this field, using a Phillips EPIQ7c ultrasound system. In instances of interobserver variability, the consensus among the observers resolved the conflicts. Common views utilized in echocardiographic imaging included short-axis and 4-chamber views.

LVEF was measured using the modified Simpson's biplane method[41]. Simpson's biplane method estimates left ventricular volumes by tracing the endocardial borders from apical two- and four-chamber views. The ventricle is conceptually divided into multiple disks, and the volume of each disk is calculated by measuring the area of the traced borders; these individual volumes are then summed to determine both the end-diastolic and end-systolic volumes. Finally, the ejection fraction is computed by subtracting the end-systolic volume from the end-diastolic volume and dividing the result by the end-diastolic volume.

Data stored in the Picture Archiving and Communication System (PACS) is used by an expert Cardio-oncologist and validated by an experienced echocardiologist. In addition to the patient's ECG characteristics, various echocardiographic parameters related to the heart valves, heart chambers, and myocardium during systole and diastole, including left and right ventricular and atrial volumes, were evaluated and recorded at each time point of imaging.

All patients underwent echocardiography immediately after completion of chemotherapy and also 12 months later. This evaluation was performed with the aim of assessing cardiac function and identifying any potential late cardiotoxicity.

Only the end-systole and end-diastole images of patients based on the ECG in the short-axis view and the 4-chamber view were manually selected and segmented. Study participants were divided into two groups: those with cardiotoxicity (41 patients) and those without (59 patients).

Preprocessing and segmentation

Echocardiography images before and after treatment in 4-chamber and short-axis views were extracted. The analysis is performed on cardiac images at end-systole and end-diastolic phases (frames), which were manually selected, segmented, and shown in Fig. 1. The area of the left ventricular myocardium for segmentation was selected by a specialist and identified in the Fig. 1. In addition, gray-level image intensities were normalized between 0 and 1and Eq. 1 defines this normalization.
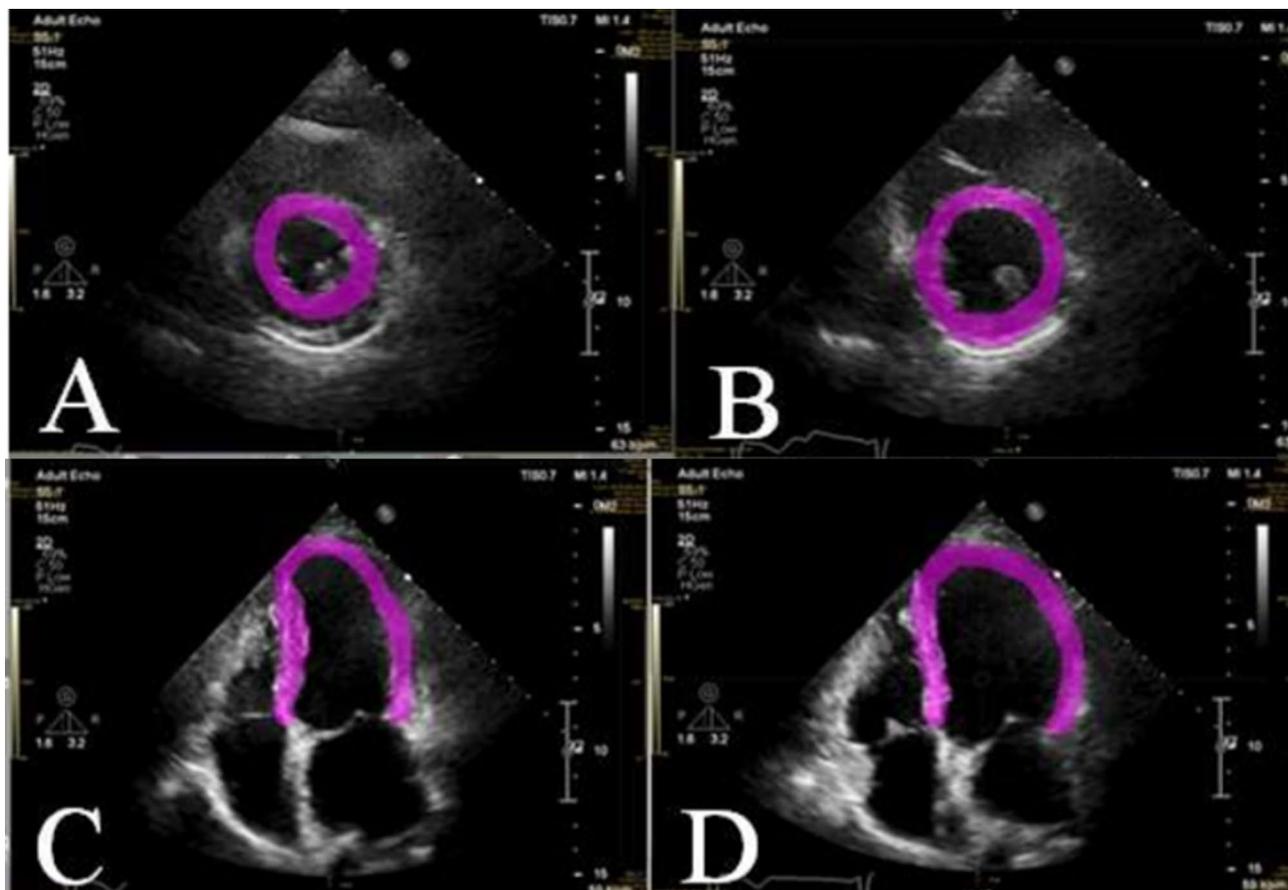
Fig. 1. Demonstrates the segmentation process for the left ventricular end-systolic (**A**) and end-diastolic (**B**) phases in the short-axis view. Panels (**C**) and (**D**) depict the corresponding left ventricular end-systolic and end-diastolic phases in the 4-chamber view echocardiography images, respectively. *P value <0.05, statistically significant.

$$I(x,y) = \frac{I'(x,y) - \min(I'(x,y))}{\max(I'(x,y)) - \min(I'(x,y))} \tag{1}$$

After preprocessing and segmenting the images, radiomics features were extracted and fed into various ML models programming languages.

Feature extraction and selection

Following the segmentation of the images, the next crucial step involved feature extraction[42], which aimed to identify and extract key characteristics or attributes from the segmented images, taking into account both the end-diastolic and end-systolic frames. This process encompassed a comprehensive array of 2D radiomics features, each contributing unique insights[43]. Specifically, the feature extraction incorporated prominent techniques such as Haralick, LAWS, wavelet-based texture analysis, SFTA (Segmentation Fractal texture analyze), Tamura, Shape, and power spectrum analysis. These diverse methods together yielded a total of 184 features for both the end-diastolic and end-systolic phases.

In this study, the chosen set of 2D radiomic features played a critical role in capturing essential patterns and textures present in the medical images. The Haralick features[44] allowed for the characterization of the textures' complexity and uniformity, while the LAWS feature[45] focused on capturing local anisotropic patterns. On the other hand, the wavelet-based texture analysis enabled the exploration of image details across different scales. Moreover, the SFTA[46] method facilitated the identification of significant local structures, while the Tamura features[47] emphasized describing the images' coarseness, contrast, and directionality. Furthermore, in addition to the aforementioned 2D radiomic features, Local Binary Pattern (LBP) features were incorporated into the feature extraction process. LBP[48] analysis provided valuable insights into the local texture patterns within the segmented images during the end-diastolic and end-systolic phases, further enriching the characterization of cardiac structures and assisting in identifying potential anomalies.

Additionally, the Shape features were instrumental in quantifying geometric attributes, such as size and compactness, providing valuable information about the morphological characteristics of the regions of interest.

Lastly, the power spectrum analysis allowed researchers to evaluate the frequency distribution within the images, aiding in the recognition of periodic patterns or repetitive structures.

This comprehensive suite of 2D radiomic features offered a holistic perspective on the cardiac images during both end-diastolic and end-systolic phases. By extracting and analyzing these 184 features, the study aimed to uncover intricate relationships between the identified characteristics and various cardiac conditions. The results of this feature-rich analysis were expected to significantly contribute to medical imaging and advance our understanding of cardiac health and pathology.

In the quest for optimal feature selection, the feature selection process encompassed several crucial steps. Initially, the Wilcoxon test was employed, with a significance level set at $p < 0.05$, to evaluate the significance of diverse features. The top 20% of features with the lowest p-values were then selected for further consideration. The sequential forward selection method was employed to refine the feature selection. These meticulous procedures resulted in identifying a set of promising features for the short-axis view, as detailed in Table 3. Similarly, for the 4-chamber view of the left ventricle, Table 3 presents the best-selected features derived from the rigorous feature selection process. These thoughtfully curated feature sets are anticipated to play a crucial role in augmenting the accuracy and effectiveness of subsequent analyses and predictive models. The steps to perform the work are demonstrated in Fig. 2.

Machine learning workflow
Images and data were randomly divided into two parts, with 80% of the data used for training and 20% for testing, and a 5-fold cross-validation was employed. The training and testing procedures were repeated 1000 times. Different ML algorithms, including K-nearest neighbor (KNN) with K = 13, Support vector machine (SVM), Random Under-Sampling Boosted (RUS), and Random Forest (RF), were implemented to perform better analysis and prediction[49].

Furthermore, we performed a Permutation-based Feature Importance Test (PermFIT) to strengthen the statistical validity of our findings. This rigorous test allowed us to assess the significance of our results by comparing them to a large number of randomly generated samples. Notably, a p-value of less than 0.001 was considered as the threshold for establishing meaningful statistical significance. This approach ensured that our conclusions were robust and reliable, reducing the likelihood of drawing spurious conclusions based on chance. The SHapley Additive exPlanations (SHAP) value serves as a powerful tool to assess the influence of various features on the outcome of a predictive model. It achieves this by quantifying the contribution of each feature to the final prediction, thereby providing a clearer and more transparent understanding of ML models. With the aid of SHAP values, we could interpret the significance of different features, unraveling the inner workings of the model and its decision-making process[50].

Ethical considerations
This retrospective study analyzed de-identified medical data that had been previously collected during routine patient care. The study protocol was approved by the ethical committee of Iran University of Medical Sciences with code number IR.IUMS.FMD.REC.1401.143. Informed consent was waived and privacy and confidentiality
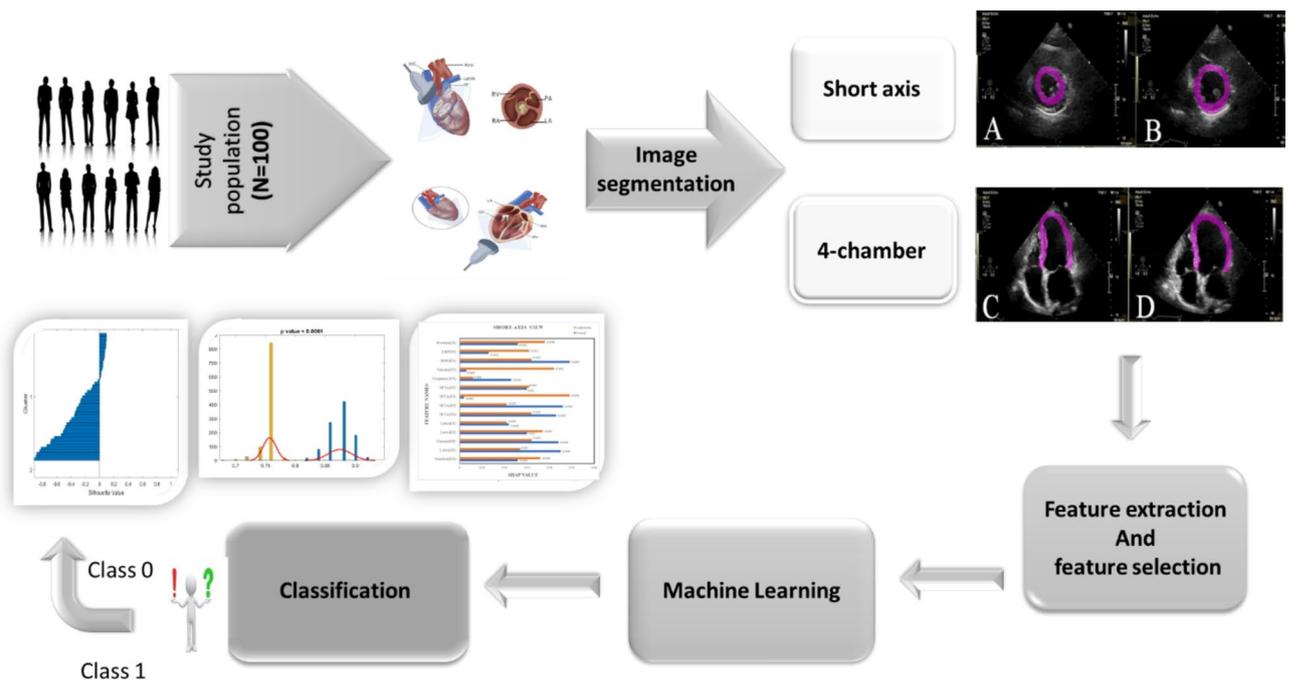
Fig. 2. Illustrates the sequential actions required to carry out the task. *(**A**, **C** ES represent end systolic and **B**, **D** ED represent end diastolic, for short axis and 4-chamber view respectively).

	Total population (n = 100)	Non-cardiotoxicity (n = 59)	Cardiotoxicity (n = 41)	P value
Baseline ejection fraction (%)	50.25 ± 9.19	47.46 ± 10.96	54.27 ± 2.64	0.001*
Age	54.53 ± 13.72	55.68 ± 14.19	52.88 ± 13.02	0.32
Height	162.33 ± 7.60	162.86 ± 8.23	161.56 ± 6.61	0.40
Weight	69.78 ± 12.70	70.96 ± 13.57	68.1 ± 11.28	0.27
BMI (Kg/m^2)	26.24 ± 4.44	26.46 ± 4.44	25.91 ± 4.48	0.54
Female Sex, n,(%)	84 (84%)	45 (76%)	39 (91%)	0.007*
Diabetes, n,(%)	18 (18%)	11 (18.64%)	8(19.51%)	0.29
Hypertension, n,(%)	21 (21%)	12 (20.33%)	9 (22%)	0.84
Dyslipidemia, n,(%)	25 (25%)	14 (23.7%)	11 (26.8%)	0.81
Previous Cardiovascular Disease, n,(%)	16 (16%)	13 (22%)	3 (7.3%)	0.056
Smoking, n,(%)	9 (9%)	6 (10.2%)	3 (7.3%)	0.45
Alcohol, n,(%)	0 (0%)	0 (0%)	0 (0%)	0.99
Metastasis, n,(%)	19 (19%)	11 (18.6%)	8 (19.5%)	0.91

Table 1. Patients' demographic and comorbidity characteristics. *P value <0.05, statistically significant.

	Total population (n = 100)	Non-cardiotoxicity (n = 59)	Cardiotoxicity (n = 41)	P value
RVID, cm	3.30 ± 3.06	3.56 ± 4	2.93 ± 0.28	0.37
LVIDd, cm	4.66 ± 0.62	4.78 ± 0.73	4.5 ± 0.42	0.03*
LVIDs, cm	3.84 ± 5.08	4.5 ± 6.8	3.02 ± 0.54	0.23
IVSD, cm	0.92 ± 0.99	1.03 ± 1.3	0.79 ± 0.15	0.28
Left atrium area, cm	14.95 ± 4.37	15.56 ± 4.77	14.15 ± 3.71	0.17
Left atrium dimeter, cm	3.28 ± 0.58	3.41 ± 0.59	3.13 ± 0.56	0.04
LA volume, cm^2	37.59 ± 16.35	39.4 ± 19.51	35.07 ± 10.4	0.26
EDV, mL	87.19 ± 29.17	91.81 ± 36.45	81.28 ± 14.04	0.08
ESV, mL	41.87 ± 23.09	45.25 ± 28.78	37.69 ± 12.28	0.13
TRG max, mmHg	22.91 ± 8.15	24 ± 9.05	21.34 ± 6.46	0.14
PAP, mmHg	27.16 ± 8.21	28.7 ± 9.04	25.38 ± 6.86	0.09
IVC, cm	1.50 ± 0.28	1.47 ± 0.28	1.53 ± 0.29	0.32
TAPSE	21.33 ± 3.54	20.71 ± 3.81	22.24 ± 2.94	0.04*
RVSP	26.95 ± 7.07	27 ± 6.82	26.88 ± 7.56	0.95
SPAP	26.72 ± 7.53	27.07 ± 8.19	26.19 ± 6.53	0.58
First follow GLS	− 0.36 ± 1.87	-0.49 ± 2.44	-0.19 ± 0.02	0.44

Table 2. Bassline echocardiographic findings of the patients among the two cardiotoxicity and non-cardiotoxicity groups. Abbreviations: Right Ventricular Internal Dimension (RVID), Left Ventricular Internal Dimension in Diastole (LVIDd), Left Ventricular Internal Dimension in Systole (LVIDs), Interventricular Septal Thickness in Diastole (IVSD), Left Atrial Area, Left Atrial Diameter, Left Atrial Volume (LA volume), End-Diastolic Volume (EDV), End-Systolic Volume (ESV), Tricuspid Regurgitation Gradient Maximum (TRG max), Pulmonary Artery Pressure (PAP), Inferior Vena Cava Diameter (IVC), Tricuspid Annular Plane Systolic Excursion (TAPSE), Right Ventricular Systolic Pressure (RVSP), Systolic Pulmonary Artery Pressure (SPAP), Global Longitudinal Strain (First follow GLS).

of their information were maintained by removing all protected health information and assigning anonymized data.

Result

The study group comprised 100 participants, with the majority being female (84%) and the remaining 16% identified as male. The mean age of the participants was 54.53 ± 13.72 years, and the mean body mass index (BMI) was 26.24 ± 4.44 Kg/m. Prevalence of comorbidities was included diabetes mellitus in 18% of participants, hypertension in 33%, dyslipidemia in 25%, smoking in 9%, and alcohol consumption in 0% (Table 1).

Table 2 provides the more detailed imagining findings of the echocardiography of the patients among the two groups. Most of the imaging variables were similar between the groups.

Table 3 provides a concise summary of selected features and their corresponding categories for both short-axis and 4-chamber views of patients. The features are categorized based on their types.

Name view	Number features	Feature category	Phase
Short axis	1	Haralick	ED
	2, 4, 5	Laws (1, 2, 3)	ED, ES, ES
	3,11	Tamura (1, 2)	ED, ES
	6,7,8,9	SFTA (1, 2, 3, 4)	ES
	10	Frequency	ES
	12	HOG	ES
	13	LBP	ES
	14	Wavelet	ES
4-chamber	1,2,6	Laws (1, 2, 3)	ED, ED, ES
	3	Frequency	ED
	4,7	LBP (1, 2)	ED, ES
	5,10	Shape (1, 2)	ED, ES
	8,9	Wavelet (1, 2)	ES

Table 3. Selected features and their categories for the short-axis and 4-chamber view of the patients.

Model	ACC	SEN	SPE	PR	FPR	F1 score	AUC
KNN	0.92	0.66	0.90	1.00	0	0.80	0.69
SVM	0.90	0.58	0.88	1.00	0	0.73	0.76
Bayesian	0.80	0.58	0.86	0.58	0.13	0.58	0.58
RF	0.82	0.50	0.85	0.66	0.07	0.57	0.59

Table 4. Performance metrics, including accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PR), false positive rate (FPR), F1 scores and ROC curve (AUC), for all models in the context of the short axis view. *Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Under-Sampling Boosted (RUS boosted), Random Forest (RF).

Model	ACC	SEN	SPE	PR	FPR	F1 score	AUC
SVM	0.88	0.67	0.85	1.00	0	0.80	0.68
KNN	0.83	0.60	0.81	0.89	0.03	0.72	0.58
RUS boosted	0.71	0.53	0.76	0.60	0.19	0.56	0.50
RF	0.75	0.64	0.80	0.64	0.19	0.64	0.46

Table 5. Illustrates performance metrics, including accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PR), false positive rate (FPR), F1 scores and ROC curve (AUC), for all models in the context of the 4-chamber view. *Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Under-Sampling Boosted (RUS boosted), Random Forest (RF).

Effectiveness of different machine learning models

In this study, using various ML algorithms, accuracy, sensitivity, specificity, precision, and F1 scores were assessed for each of the short axis and 4 chamber views of the left ventricle separately and for both views.

Among the 100 patients exclusively assessed with a short-axis view, the KNN and linear SVM models emerged as the most effective, achieving accuracies of 0.92 and 0.90, respectively. The best-reported outcome was 92%, whose values for each parameter in short axis view are reported in Table 4. The training and testing procedures with random sampling were repeated 1000 times.

Table 5 presents performance metrics for models in the context of 4-chamber views. Notably, in the 4-chamber view, the SVM model exhibits a high accuracy of 88%, achieving a perfect precision of 1.00 and an F1 score of 0.80. Additionally, the KNN model demonstrates strong overall performance with an accuracy of 83%, a precision of 0.89, and an F1 score of 0.72. These results highlight the effectiveness of these models in assessing both short-axis and 4-chamber views, providing valuable insights for potential clinical applications.

Table 6 its illustration the reported results are based on the mean of a 5-fold cross-validation. However, if the experiment is repeated, the results may vary due to the inherent randomness introduced by data shuffling.

Therefore, we did not conduct an extensive comparison between models. Instead, the various models highlight the power of radiomics, and the analysis of the permutation test confirms that the results are not due to chance. have implemented the DeLong test, which provides the p-values for pairwise comparisons between models. we have included the Area Under the ROC Curve (AUC) for 4chamber and values for each classifier as follows: KNN ($N = 13$) – 0.5896, SVM – 0.6801, RUSboosted – 0.5059, and Random Forest (RF) – 0.4649. the

View	Model	*p*-value
4-chamber	KNN vs. SVM	0.18
	KNN vs. RF	0.59
	KNN vs. Bayesian	0.78
	SVM vs. RF	0.23
	SVM vs. Bayesian	0.15
	RF vs. Bayesian	0.38
Short axis	KNN vs. SVM	0.14
	KNN vs. RF	0.30
	KNN vs. Bayesian	0.49
	SVM vs. RF	0.48
	SVM vs. Bayesian	0.44
	RF vs. Bayesian	0.75

Table 6. Shows p-values for all AUC values which will be the main metrics for 4-chamber and short axis views.

Area Under the ROC Curve (AUC) values for the classifiers on the short-axis view are as follows: KNN ($N = 13$) – 0.6919, SVM – 0.7637, RUSboosted – 0.5844, and Random Forest (RF) – 0.5983.

Figure 4 display influence the mean SHAP values of selected characteristics in short-axis samples in the normal and cardiotoxic classes. For class 0, which represents the healthy group (non-cardiotoxic), according to Fig. 4, Features SFTA3(ES), Tamura2(ES), and Wavelet (ES) are each most important to determine the toxicity or non-toxicity of each group.

Figure 5 shows the effect of the characteristics on the cardiotoxic and normal classes for the 4-chamber view using the average *SHAP* values. Features Wavelet1(ES), Wavelet 2(ES), and Shape 1 (ED) represent the most effective characteristics in the cardiotoxic group.

Discussion

The present study encompassed 100 participants, predominantly female (86%), with an average age of 54.53 ± 13.72 years. It aimed to employ interpretable machine learning techniques to predict cardiac toxicity following chemotherapy in cancer patients, utilizing echocardiographic imaging data. The objective was to gain insights into the intricate interplay between cancer treatment and cardiovascular complications within this patient population. The study design involved the analysis of echocardiography images obtained before and after chemotherapy administration to evaluate its impact on cardiac function. Parameters such as LVEF were assessed, and predictive models for chemotherapy-induced cardiotoxicity were developed. Additionally, feature extraction and selection methods were applied, with the significance of the extracted features validated through techniques such as SHAP and PermFIT testing.

In a systematic review[51], various research studies were reviewed which demonstrated several prevalent prognostic factors emerged, including age, initial left ventricular ejection fraction, hypertension, and diabetes. Moreover, scrutiny of four constructed models revealed significant apprehensions regarding bias susceptibility, notably concerning sample size, management of missing data, and the absence of suitable performance metric. An investigation[18] conducted retrospectively examined the predictive capabilities of ML algorithms in a cohort of 4309 patients with cancer. The study focused on leveraging clinically relevant variables for prediction purposes. Results yielded a range of area under the curve (AUC) values, spanning from 0.660 to 0.882, indicative of varying predictive performance across different models. Noteworthy variables associated with CTRCD were identified through the analysis, encompassing factors such as age, hypertension, glucose levels, LVEF, creatinine levels, and aspartate aminotransferase levels. In a study akin to our investigation[24], a range of parameters encompassing clinical, chemotherapy, and echocardiographic variables were employed to construct a RF model aimed at prognosticating chemotherapy-induced cardiac dysfunction and Heart failure with reduced ejection fraction (HFrEF). The study entailed a comprehensive evaluation of the cardiac impact of chemotherapeutic agents over a 3-year follow-up period. The Multi-Layer Perceptron (MLP) model emerging as the most effective in predicting CTRCD, yielding an AUC of 0.66, accompanied by sensitivity and specificity rates of 0.86 and 0.53, respectively. However, it is imperative to note that these performance metrics are notably lower compared to the findings of our own study. In a comparable investigation conducted by MARTINS et al.[52], the focus was on breast cancer patients undergoing anthracycline chemotherapy, with the utilization of the Extreme Gradient Booster (XGB) model to predict cardiotoxicity. Notable echocardiographic variables such as MPI (Myocardial Performance Index), ESV (End Systolic Volume), EDV (End Diastolic Volume), and SBP (sub-bottom profiler) emerged as significant predictors in their analysis. The XGB model demonstrated commendable performance, achieving an accuracy of 0.89 and an AUC of 0.91. Variable importance was assessed using the SHAP approach in their study. While bearing similarities to our present investigation in terms of accuracy, it is worth noting the differences in methodology. In our study, we opted for the PermFIT[53] for variable importance assessment. Additionally, our study diverges from the aforementioned one in terms of parameter selection. While the previous study solely relied on echocardiography parameters, our approach incorporated radiomic features, which are inherently

more consistent due to their automatic nature and are less susceptible to environmental errors such as user-dependency and inter-observer variability.

Figure 3 illustrates the outcomes of the PermFIT test designed to evaluate the model's accuracy. In Fig. 3a, the histogram distributions depict that the chance-level accuracy is concentrated around 0.49, whereas the model achieves an accuracy centered around 0.85 in the 4-chamber view. Additionally, Fig. 3b displays the PermFIT test results for the short-axis view, where the chance-level is approximately 0.76, and the model's accuracy is around 0.88. These results imply that the model's accuracy significantly exceeds the chance level, with a more pronounced effect observed in the 4-chamber view compared to the short-axis view, emphasizing its efficacy in the specified context.

The SHAP method investigates a local approximation to the predictions of regression or classification models, ensuring faithfulness and interpretability. Enhancing the understanding of the model's decision-making process[54]. A SHAP analysis was performed in Figs. 4 and 5 to assess the significance of features in short-axis and 4-chamber views across normal and cardiotoxic classes. This analysis elucidates the impact and relevance of selected features within the classification model for different classes. Figure 4 provides valuable insights into how various features influence the model's predictions in short-axis view. Among the radiomic features in the short-axis view, end-systolic SFTA and HOG exhibited the most substantial impact on the cardiotoxic and normal classes, respectively. Conversely, end-systolic SFTA and Tamura had minimal influence on the model's prediction for the normal class. Additionally, end-systolic Frequency demonstrated low values for the cardiotoxic class, signifying its limited impact. Alternatively, for the normal class, end-systolic HOG and SFTA were the most influential, while end-systolic SFTA and Tamura had the least impact on predictions. In the cardiotoxic class, end-systolic SFTA and Tamura had the most significant impact, while end-systolic Frequency and end-systolic SFTA and LAWS had the least effect on predictions. In the 4-chamber view (Fig. 5), end-systolic Wavelet and end-diastolic Laws had the most and least impact on predictions, respectively, in the cardiotoxic group. In the normal group, end-diastolic shape and Laws had the highest and lowest SHAP values. Generally, in the short-axis view, most features pertain to end-systolic features, whereas in the 4-chamber view, end-systolic and end-diastolic features contribute equally, suggesting a weaker influence of end-diastolic features in the short-axis view. Notably, in the 4-chamber view, the highest values are associated with the cardiotoxic class, underscoring the positive effects of features in the cardiotoxic class. In the present study, we employed the SHAP approach to identify selected features' value and importance. ML techniques can effectively capture intricate nonlinear connections. However, ML models' inherent complexity and "black box" nature pose challenges in interpreting their outputs, limiting their applicability in this domain.

Limitations
Financial and specific data unavailability, we had severe limitations in data acquisition and the conduction of the study due to financial and specific data unavailability. In addition, due to limited resources, we could not assess the prediction model based on external validation and used cross-validation with our self-data. Furthermore, 3D echocardiography with more accurate and precise parameters could be more reliable than 2D echocardiography. Furthermore, concerning personalized medicine and individual treatment plans for each patient, as well as dose
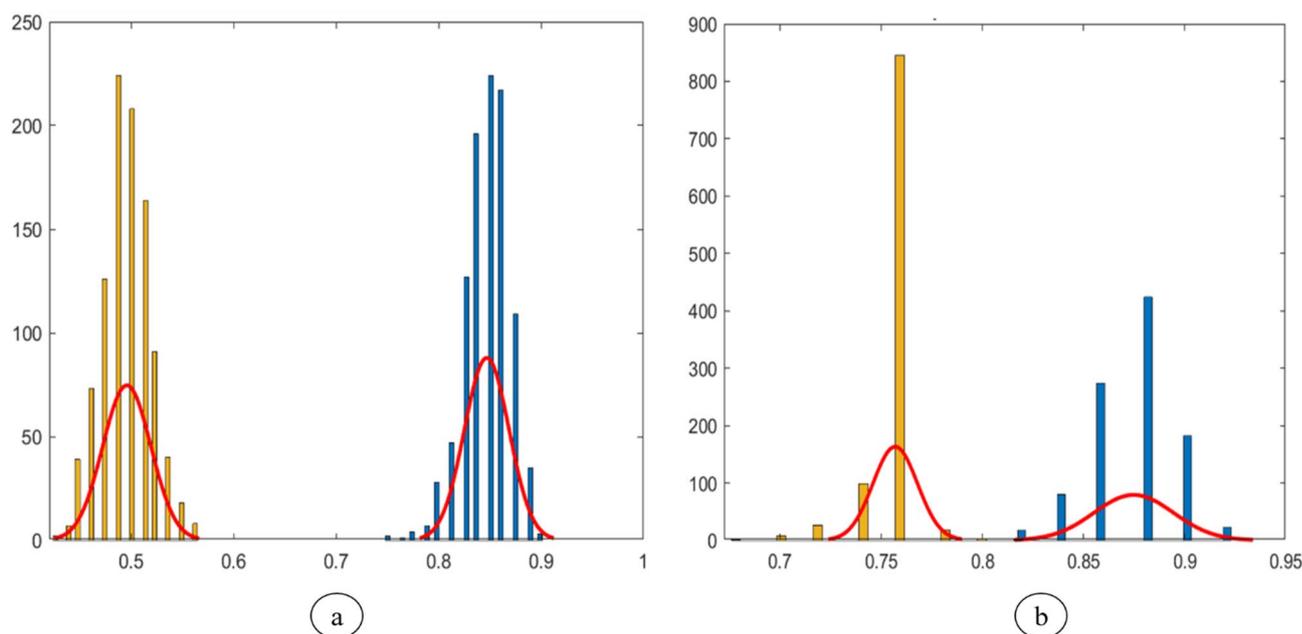
Fig. 3. Histogram representation of the results for PermFIT in the (**a**) 4-chamber and (**b**) short-axis view. The test was conducted by repeating it 1000 times with the correct label and another 1000 times with a random label. The horizontal axis represents accuracy (AUC), while the vertical axis denotes the frequency of result repeat over 1000 instances of correct labeling and 1000 instances of random labeling.
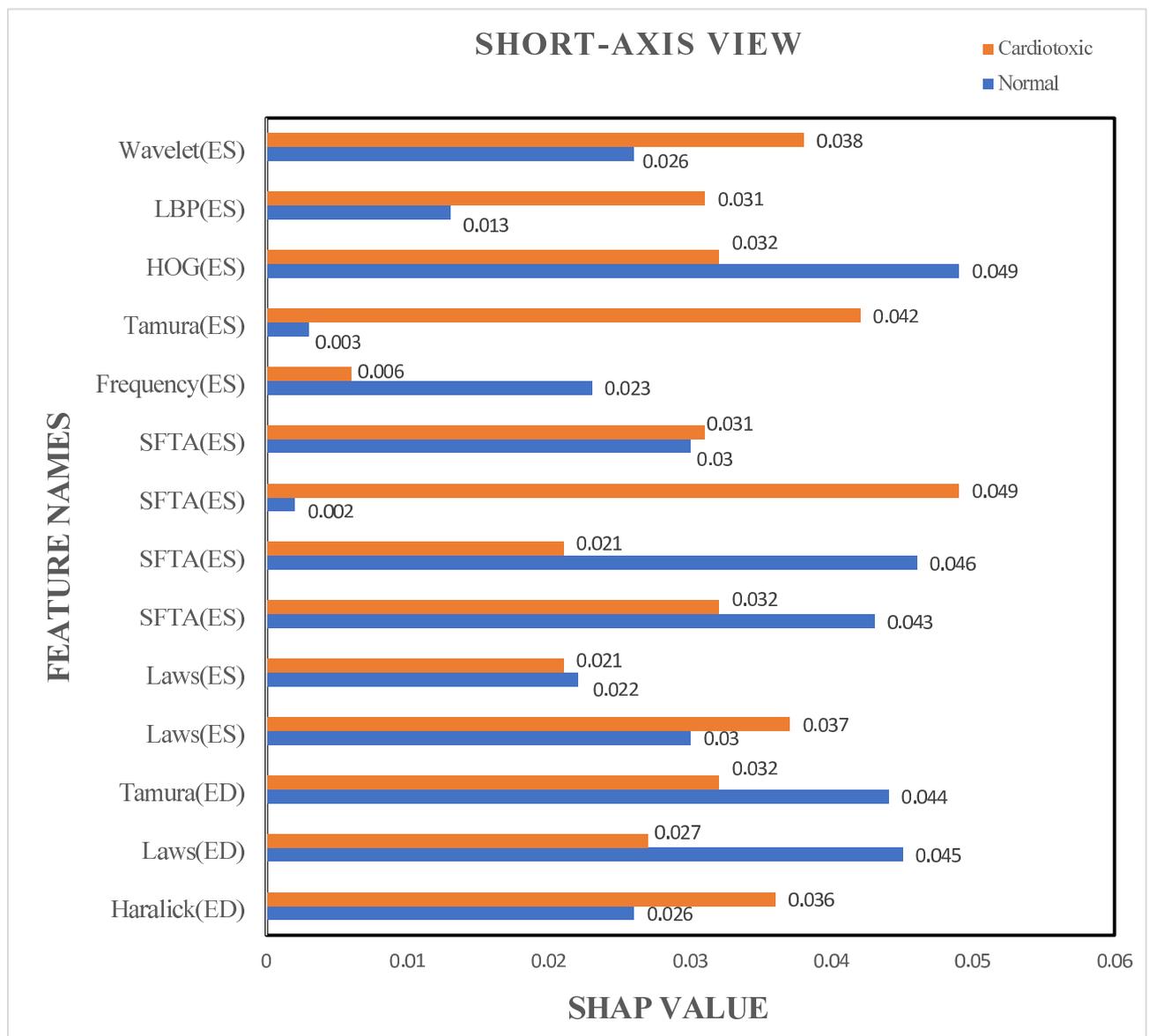
Fig. 4. Influence of features for the short-axis view on the normal classes and cardiotoxicity as determined by SHAP values.

adjustments and choice medication based on the patients' medical conditions, some heterogeneity was expected in the sample population. however, on average, all patients underwent standard chemoradiation therapy in accordance with the guidelines.

Future prospects

Based on our findings, future research could focus on larger, multi-center studies to validate and generalize our results, ensuring broader applicability. Additionally, integrating deep learning with radiomics could enhance predictive accuracy by capturing subtle features in echocardiographic images, while exploring early biomarkers for cardiotoxicity could enable earlier detection before measurable changes in LVEF occur. Future studies should also aim for personalized risk stratification, combining radiomics, clinical factors, and genetic data, to tailor chemotherapy and monitoring to individual patients. Long-term follow-up studies would further help understand the impact of early prediction on both cardiac health and cancer outcomes, advancing our approach to managing chemotherapy-induced cardiotoxicity.

Enrolling more related parameters like ECG and paraclinical laboratory data and use more sample sizes is suggested in future studies to achieve more sensitivity and specificity. Also, adding strain analysis (GLS, GRS and GCS) to the present parameters could enhance the accuracy of the prediction model. In the present study, we exclusively used imaging modalities with some demographic characteristics. Of course, it can be considered a strength point because of the limited resources needed.
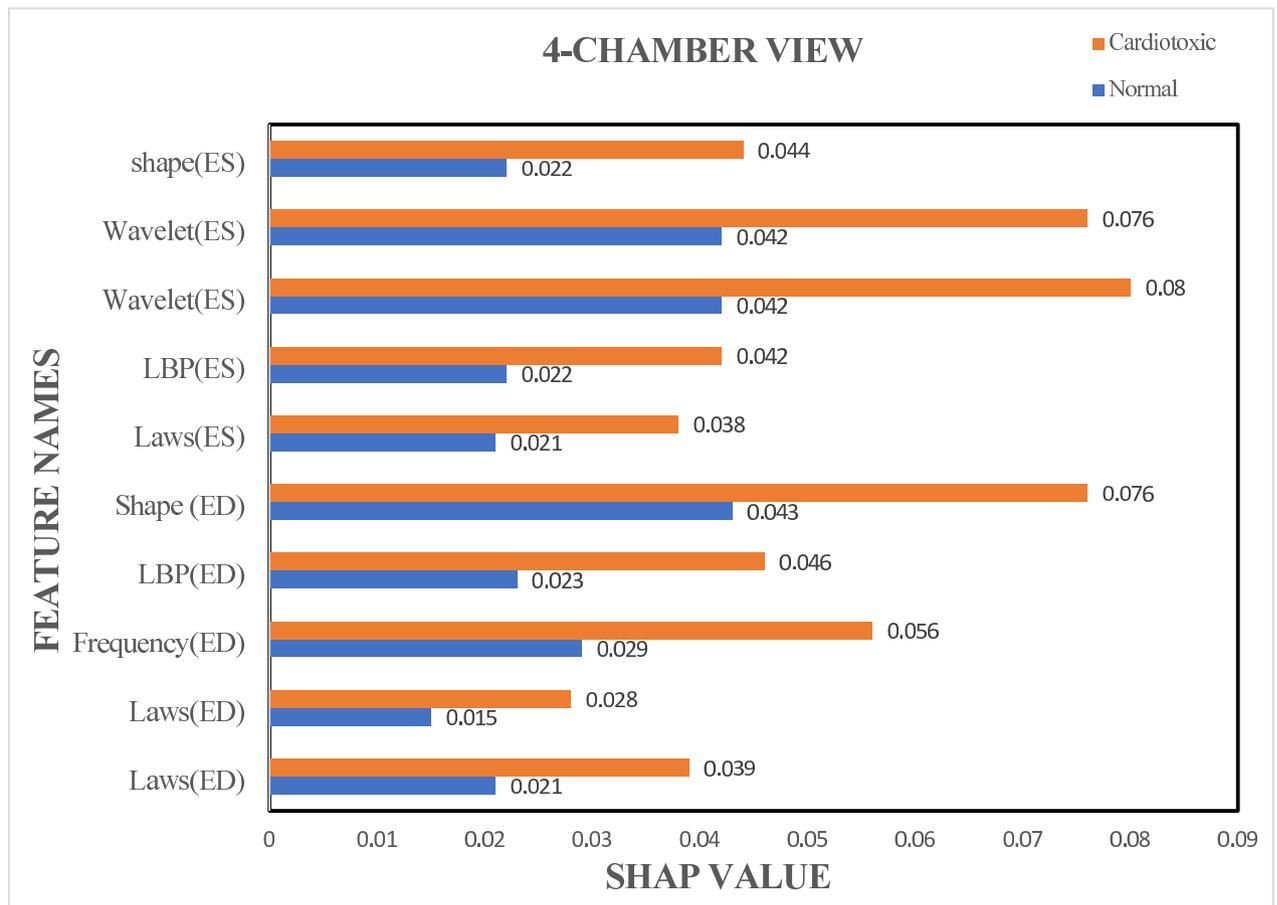
Fig. 5. The impact of features for the 4-chamber view on normal and cardiotoxic classes, as determined by SHAP values.

Conclusion

This study underscores the potential of ML in enhancing cardiac toxicity diagnosis, particularly in the context of short-axis echocardiography views. Given the importance of early detection, we advise clinicians to incorporate advanced imaging-based predictive models, such as ML-driven radiomics, into cardiac monitoring measures for chemotherapy patients, at least in some high-risk patients, if not all. These approaches may help identify early signs of cardiotoxicity, enabling timely interventions and personalized treatment adjustments. Future studies with larger cohorts will be necessary to validate these findings and further optimize risk stratification in clinical practice. The study faced limitations due to financial and data unavailability, limited resources, and cross-validation. 3D echocardiography could be more reliable than 2D echocardiography. Future studies should include more parameters and larger sample sizes to improve sensitivity and specificity. The study used imaging modalities with demographic characteristics, which could be a strength due to limited resources.

Data availability

The data supporting the findings of this study are available from the corresponding author(bitarafan@hotmail.com)upon reasonable request.

References

1. Sung, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J. Clin.* **71** (3), 209–249 (2021).
2. Senapati, S. et al. Controlled drug delivery vehicles for cancer treatment and their performance. *Signal. Transduct. Target. Therapy.* **3** (1), 7 (2018).
3. Clarke, M. et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. *Lancet* **366** (9503), 2087–2106 (2005).
4. Florescu, M., Cinteza, M. & Vinereanu, D. Chemotherapy-induced Cardiotoxicity. *Maedica (Bucur)* **8**(1): 59–67 (2013).
5. Taylor, C. et al. Estimating the risks of breast Cancer radiotherapy: evidence from modern radiation doses to the lungs and heart and from previous randomized trials. *J. Clin. Oncol.* **35** (15), 1641–1649 (2017).

6. Nicolazzi, M. A. et al. Anthracycline and trastuzumab-induced cardiotoxicity in breast cancer. *Eur. Rev. Med. Pharmacol. Sci.* **22** (7), 2175–2185 (2018).

7. Callahan, R. & Hurvitz, S. Human epidermal growth factor receptor-2-positive breast cancer: current management of early, advanced, and recurrent disease. *Curr. Opin. Obstet. Gynecol.* **23** (1), 37–43 (2011).

8. Qiu, S. et al. Risk factors for Anthracycline-Induced cardiotoxicity. *Front. Cardiovasc. Med.* **8**, 736854 (2021).

9. Zamorano, J. L. et al. [2016 ESC position paper on cancer treatments and cardiovascular toxicity developed under the auspices of the ESC committee for practice guidelines]. *Kardiol Pol.* **74** (11), 1193–1233 (2016).

10. Plana, J. C. et al. Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American society of echocardiography and the European association of cardiovascular imaging. *J. Am. Soc. Echocardiogr.* **27** (9), 911–939 (2014).

11. Albini, A. et al. Cardiotoxicity of anticancer drugs: the need for cardio-oncology and cardio-oncological prevention. *J. Natl. Cancer Inst.* **102** (1), 14–25 (2010).

12. Bergom, C. et al. Past, present, and future of Radiation-Induced cardiotoxicity: refinements in targeting, surveillance, and risk stratification. *JACC CardioOncol.* **3** (3), 343–359 (2021).

13. Jiji, R. S., Kramer, C. M. & Salerno, M. Non-invasive imaging and monitoring cardiotoxicity of cancer therapeutic drugs. *J. Nuclear Cardiol.* **19**, 377–388 (2012).

14. Celebi, A. S., Yalcin, H. & Yalcin, F. Current cardiac imaging techniques for detection of left ventricular mass. *Cardiovasc. Ultrasound.* **8**, 19 (2010).

15. Denlinger, C. S. et al. Survivorship, version 2.2018, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Canc. Netw.* **16** (10), 1216–1247 (2018).

16. Thavendiranathan, P. et al. Reproducibility of echocardiographic techniques for sequential assessment of left ventricular ejection fraction and volumes: application to patients undergoing cancer chemotherapy. *J. Am. Coll. Cardiol.* **61** (1), 77–84 (2013).

17. Zamorano, J. L. et al. ESC Position Paper on cancer treatments and cardiovascular toxicity developed under the auspices of the ESC Committee for Practice Guidelines: The Task Force for cancer treatments and cardiovascular toxicity of the European Society of Cardiology (ESC). *Eur. Heart J.* **37**(36), 2768–2801 (2016).

18. Zhou, Y. et al. Machine Learning-Based risk assessment for Cancer Therapy-Related cardiac dysfunction in 4300 longitudinal oncology patients. *J. Am. Heart Assoc.* **9** (23), e019628 (2020).

19. Gillespie, H. S., McGann, C. J. & Wilson, B. D. Noninvasive diagnosis of chemotherapy related cardiotoxicity. *Curr. Cardiol. Rev.* **7** (4), 234–244 (2011).

20. Ali, M. T. et al. Myocardial strain is associated with adverse clinical cardiac events in patients treated with anthracyclines. *J. Am. Soc. Echocardiogr.* **29** (6), 522–527e3 (2016).

21. Zhang, Z. et al. Artificial Intelligence-Enhanced echocardiography for systolic function assessment. *J. Clin. Med.*, **11**(10). (2022).

22. Kosaraju, A. et al. *Left Ventricular Ejection Fraction, in StatPearls* (StatPearls Publishing, 2023).

23. Yang, X. et al. *Identifying Cancer Patients at Risk for Heart Failure Using Machine Learning Methods.* AMIA Annu Symp Proc, 2019, pp. 933–941. (2019).

24. Chang, W. T. et al. An artificial intelligence approach for predicting cardiotoxicity in breast cancer patients receiving anthracycline. *Arch. Toxicol.* **96** (10), 2731–2737 (2022).

25. Liu, X. et al. Deep learning-based automated left ventricular ejection fraction assessment using 2-D echocardiography. *Am. J. Physiol. Heart Circ. Physiol.* **321** (2), H390–h399 (2021).

26. Zhou, J. et al. Artificial intelligence in echocardiography: detection, functional evaluation, and disease diagnosis. *Cardiovasc. Ultrasound.* **19** (1), 29 (2021).

27. Barry, T. et al. The role of artificial intelligence in echocardiography. *J. Imaging.* **9** (2), 50 (2023).

28. Khanna, M. et al. An enhanced and efficient approach for feature selection for chronic human disease prediction: A breast cancer study. *Heliyon* **10** (5), e26799 (2024).

29. Singh, L., Khanna, M. & Garg, H. Multimodal biometric based on fusion of ridge features with minutiae features and face features. *Int. J. Inf. Syst. Model. Des.* **11**, 37–57 (2020).

30. Sabouri, M. et al. Machine learning based readmission and mortality prediction in heart failure patients. *Sci. Rep.* **13** (1), 18671 (2023).

31. Mohebi, M. et al. Post-revascularization ejection fraction prediction for patients undergoing percutaneous coronary intervention based on myocardial perfusion SPECT imaging radiomics: a preliminary machine learning study. *J. Digit. Imaging.* **36** (4), 1348–1363 (2023).

32. Sabouri, M. et al. Myocardial perfusion SPECT imaging radiomic features and machine learning algorithms for cardiac contractile pattern recognition. *J. Digit. Imaging.* **36** (2), 497–509 (2023).

33. Srinivasan, S. et al. A hybrid deep CNN model for brain tumor image multi-classification. *BMC Med. Imaging.* **24** (1), 21 (2024).

34. Srinivasan, S. et al. Multimodal biomedical image segmentation using Multi-Dimensional U-Convolutional neural network. *BMC Med. Imaging.* **24** (1), 38 (2024).

35. A. G. B. et al. Robust brain tumor classification by fusion of deep learning and channel-wise attention mode approach. *BMC Med. Imaging.* **24** (1), 147 (2024).

36. Srinivasan, S. et al. An active learning machine technique based prediction of cardiovascular heart disease from UCI-repository database. *Sci. Rep.* **13** (1), 13588 (2023).

37. Singh, L., Pooja & Garg, H. Detection of Glaucoma in retinal fundus images using fast fuzzy C mean clustering. *Int. J. Fuzzy Syst. Adv. Appl.* **7**, 16–23 (2020).

38. Ardahan Sevgili, S. & Şenol, S. Prediction of chemotherapy-related complications in pediatric oncology patients: artificial intelligence and machine learning implementations. *Pediatr. Res.* **93** (2), 390–395 (2023).

39. Cuplov, V. & André, N. Machine learning approach to forecast chemotherapy-Induced haematological toxicities in patients with rhabdomyosarcoma. *Cancers (Basel)* **12**(7) (2020).

40. Li, C. et al. Using machine learning approaches to predict short-term risk of cardiotoxicity among patients with colorectal cancer after starting fluoropyrimidine-based chemotherapy. *Cardiovasc. Toxicol.* pp. 1–11 (2021).

41. Otterstad, J. E. et al. Accuracy and reproducibility of biplane two-dimensional echocardiographic measurements of left ventricular dimensions and function. *Eur. Heart J.* **18** (3), 507–513 (1997).

42. Moradi, S. et al. MFP-Unet: A novel deep learning based approach for left ventricle segmentation in echocardiography. *Physica Med.* **67**, 58–69 (2019).

43. Singh, L. et al. A novel hybridized feature selection strategy for the effective prediction of glaucoma in retinal fundus images. *Multimedia Tools Appl.* **83**, 1–73 (2023).

44. Haralick, R. M., Shanmugam, K. & Dinstein, I. H. Textural features for image classification. *IEEE Trans. Syst. Man. Cybern.* (6), 610–621. (1973).

45. Dash, S. & Jena, U. R. Multi-resolution laws' masks based texture classification. *J. Appl. Res. Technol.* **15** (6), 571–582 (2017).

46. Costa, A. F., Humpire-Mamani, G. & Traina, A. J. M. *An efficient algorithm for fractal analysis of textures.* In *2012 25th SIBGRAPI Conference on Graphics, Patterns and Images.* IEEE (2012).

47. Majtner, T. & Svoboda, D. *Extension of tamura texture features for 3d fluorescence microscopy.* In *Second international conference on 3D imaging, modeling, processing, visualization & transmission.* IEEE (2012).

48. Nanni, L., Lumini, A. & Brahnam, S. Local binary patterns variants as texture descriptors for medical image analysis. *Artif. Intell. Med.* **49** (2), 117–125 (2010).
49. Singh, L., Pooja & Garg, H. Detection of Glaucoma in retinal images based on multiobjective approach. *Int. J. Appl. Evolutionary Comput.* **11**, 15–27 (2020).
50. Wu, K. et al. A comprehensive texture feature analysis framework of renal cell carcinoma: pathological, prognostic, and genomic evaluation based on CT images. *Eur. Radiol.* pp. 1–11. (2021).
51. Kaboré, E. G. et al. Risk prediction models for cardiotoxicity of chemotherapy among patients with breast cancer: A systematic review. *JAMA Netw. Open.* **6** (2), e230569–230569 (2023).
52. Martins, H. C. et al. Abstract 16812: value of echocardiographic variables to predict cardiotoxicity in patients using machine learning. *Circulation* **148** (Suppl_1), A16812–A16812 (2023).
53. Mi, X. et al. Permutation-based identification of important biomarkers for complex diseases via machine learning models. *Nat. Commun.* **12** (1), 3008 (2021).
54. Nohara, Y. et al. Explanation of machine learning models using Shapley additive explanation and application for real data in hospital. *Comput. Methods Prog. Biomed.* **214**, 106584 (2022).

Acknowledgements

Author contributions

M.A. and A.G.J. contributed to data curation and wrote the original manuscript text and E.B.G. contributed to analyzing ML the images and writing part of drafting the original. M.M. and M.M. revised the draft. A.A.A and A.B.R. were involved in formal analysis and worked on visualization and finally also M.A. was responsible for conceptualization.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.B.-R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.