Pretreatment Prediction of Pathologic Complete Response (pCR) to Neoadjuvant Chemotherapy (NAC) for ER+ HER2- Locally Advanced Breast Cancer (LABC)

Jiwoo Park

Department of Medicine

The Graduate School, Yonsei University

Pretreatment Prediction of Pathologic Complete Response (pCR) to Neoadjuvant Chemotherapy (NAC) for ER+ HER2- Locally Advanced Breast Cancer (LABC)

Directed by Professor Min Jung Kim

The Doctoral Dissertation
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Medical Science

Jiwoo Park

December 2023

This certifies that the Doctoral Dissertation of
Jiwoo Park is approved.

Thesis Supervisor : Min Jung Kim

Thesis Committee Member#1 : Joo Hyuk Sohn

Thesis Committee Member#2 : Eun-Kyung Kim

Thesis Committee Member#3 : Kyunghwa Han

Thesis Committee Member#4 : Yangmo Yoo

The Graduate School
Yonsei University

December 2023

ACKNOWLEDGEMENTS

Above all, I offer my sincere gratitude to God who has been my guiding light and shown me the road ahead.

I would also like to thank Professor Min Jung Kim, the supervisor for this study, for the full support throughout the entire process of completing this study for the past several years.

I would also like to give my warmest thanks to Professor Joo Hyuk Sohn, the chair of the committee, who helped to lead this study. Also, to Prof. Eun-Kyung Kim, the role model who I wanted to emulate since my residency years; Prof. Kyunghwa Han whose valuable advice has always been timely; and Prof. Yangmo Yoo who provided me with a learning opportunity.

Last but not least, I would like to give special thanks to my dad who is proudly watching over me from heaven; my mom and family who are always praying for me; and especially my husband, Hangil Kim for being my best friend forever and for his unconditional love and support. Without their understanding, it would be impossible for me to complete my study.

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

Abstract

Pretreatment Prediction of Pathologic Complete Response to Neoadjuvant Chemotherapy for ER+ HER2- Locally Advanced Breast Cancer : A Machine Learning Model with Comprehensive Radiomic Features from Tumoral and Peritumoral regions across MRI sequences

Jiwoo Park

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Min Jung Kim)

Objective: To predict the optimal patients who can achieve pathologic complete response (pCR) after neoadjuvant chemotherapy (NAC) before starting treatment for the ER+ HER2- locally advanced breast cancer (LABC) patient group, which is currently known to have a poor response to NAC.

Methods: This retrospective study was conducted on 265 patients who were diagnosed with ER+ HER2- LABC, underwent pretreatment magnetic resonance imaging (MRI) and performed NAC and whose final pathology was confirmed by surgery at our hospital between 2010 and 2020. Based on January 2016, the day of pretreatment MRI, the patients were divided into the training and validation cohorts. In this study, the volume of interest (VOI) was drawn for the tumoral and peritumoral regions in the pretreatment MRI, and three MRI sequences were used; T1-weighted fat-suppressed early and delayed post-contrast subtraction sequences (Ph2 and Ph6, respectively) and T2-weighted fat-suppressed sequence (T2FS). Seven machine learning models were constructed on the tumoral, peritumoral and tumoral + peritumoral texture features from each sequence, and models equally created by combining with one another across two different sequences. The same process was applied in constructing the models incorporating the clinical factors including patient age, tumor size and ER and PR expression rates. The pCR prediction performance

was evaluated and compared for all models based on receiver operator characteristic curve and area under the curve (AUC) values.

Results: A total of 7,533 texture features were obtained from the VOIs of three pretreatment MRI sequences. Among the models for a single sequence, the SVM model on the tumor + peritumor 1 mm VOIs in the Ph2 demonstrated superior performance (AUC=0.9447). And the K-Nearest Neighbor combination model on the tumor + peritumor 1 mm VOIs in Ph2 and on the peritumor 3 mm VOI in the T2FS exhibited the best performance (AUC=0.9631).

Conclusion: We suggest that the combination machine learning model incorporating tumoral and peritumoral texture features across the different MRI sequences can make more accurate pretreatment pCR prediction for NAC response of ER+HER2- LABC patients. Our results are also anticipated to make a potential contribution to the development of clinical therapeutic strategies.

Key words : ER+ HER2- locally advanced breast cancer, neoadjuvant chemotherapy (NAC), pathological complete response (PCR), tumoral radiomics, peritumoral radiomics, pretreatment MRI, radiomics, machine learning (ML)

Pretreatment Prediction of Pathologic Complete Response to Neoadjuvant Chemotherapy for ER+ HER2- Locally Advanced Breast Cancer : A Machine Learning Model with Comprehensive Radiomic Features from Tumoral and Peritumoral regions across MRI sequences

Jiwoo Park

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Min Jung Kim)

I. Introduction

Breast cancer is the most common female cancer worldwide [1,2]. Among them, ER+ HER2- breast cancer has consistently increased in number since surpassing the incidence of ER- breast cancer in 1950 [2]. Overall, ER+ HER2- breast cancer has a good prognosis compared to other breast cancer subtypes, but due to the high incidence, it is the main subtype that accounts for the highest proportion of breast cancer mortality [2].

The treatment of ER+HER2- breast cancer can be broadly divided into early stage with no lymph node (LN) involvement and advanced stage with LN involvement. Currently, for the early breast cancer without LN involvement, the oncotype DX breast recurrence score is used as a quantitative measurement using RT-PCR to predict the chemotherapy response for patients requiring adjuvant chemotherapy [3]. For the locally advanced breast cancer (LABC) with LN involvement, the standard protocol has so far been the operation after neoadjuvant chemotherapy (NAC) [4,5]. For NAC, first, the lesion size is reduced to improve the operability, and notably, the pathologic complete response (pCR) upon the operation after NAC has been proven as a powerful prognostic factor of the patients' long-

term outcomes [6,7,8]. Nonetheless, when the effect of NAC is compared with breast cancer of other molecular subtypes, it is rather known to be poor NAC response [4,5]. According to a meta-analysis study, the pCR rate of LABC upon the operation after NAC varied from 26.5% to 39.0% in other molecular subtypes, the ER+HER2- subtype showed a significant difference with a rate of 7.2% to 13.0% [9]. In this respect, we have focused on identifying, at an early stage, patients within the minority of approximately 10% of ER+ HER2- LABC patients who exhibit a favorable response to NAC and have the potential to achieve a pCR upon the operation. For this purpose, we aimed to utilize magnetic resonance imaging (MRI), a representative modality for treatment response evaluation, to classify these patients and potentially provide clinical assistance [6,7,10].

In the early days, MRI has been used to measure the lesion size, volume, morphologic features, and enhancement patterns in the assessment of treatment response [11]. Since then, with technological advancement, there have been attempts to predict NAC by multi-parametric MRI, where diffusion or perfusion is used, magnetic resonance spectroscopy (MRS) and fluorodeoxyglucose (FDG)-positron emission tomography (PET), etc. [7,12]. Nevertheless, it has been still a challenge to design a set of parameters with high accuracy and reproducibility. Recently, radiomics research using MRI texture features has been actively conducted [3, 10]. The MRI texture analysis (TA) is advantageous in allowing an objective MRI-based evaluation by quantifying a large amount of data reflecting the heterogeneity of internal tissue components that are difficult to observe by the naked eye [13]. Moreover, MRI radiomic features are applied to the field of machine learning models composed of key texture features, and it leads to attempts to increase the reproducibility of assessment [14].

Most previous studies that have utilized MRI to predict the pCR rate of breast cancer can be broadly divided into two general aspects based on the region of interest in the imaging. First, many studies have attempted to explain the treatment responses through the change in values in initial MRI and early or mid-term MRI after NAC [15,16]. However, in evaluating residual lesion in the follow-up MRI, it has been reported to either over- or

under-estimate the lesions due to various changes associated with the treatment response [17]. More importantly, the experience of drug toxicity or delay of suitable treatments by the patients who end up receiving unnecessary treatments should be considered. Thus, this study focused on how to refine the prediction of pCR in pretreatment MRI. Second, in many previous studies, when evaluating lesions on MRI, the focus was mainly on the tumor region, but in this study, the peritumoral region is also evaluated and analyzed [18,19,20]. Based on the several study results that the peritumoral region can be a critical to the response of NAC by reflecting angiogenic or lymphangiogenic activity, this study focused more on confirming the importance of the peritumoral region [21,22]. So far, a few recent studies have attempted to use pretreatment MRI only or include the peritumoral region as a consideration [21,23], but these studies considered the ER+HER2- subtype, the focus of this study, only as a part of the study population, and no study has yet investigated the pretreatment NAC response with a focus on the ER+HER2- subtype.

Therefore, this study aimed to develop and validate a reproducible practical machine learning model with the texture feature incorporating both the tumoral and peritumoral regions across initial MRI sequences before treatment in ER+ HER2- LABC patients whose NAC response is notably low. Through this study, we hope to provide practical help for clinicians to establish a tailored therapy strategies by stratifying this patient population prior to treatment.

II. Materials and Methods

1. Patient population and study design

This retrospective study was approved by the institutional review board of our hospital, and the requirement for informed consent was waived.

From January 2010 to December 2020, 2,349 patients with advanced breast cancer received neoadjuvant chemotherapy (NAC) at our hospital. Among them, 818 patients were diagnosed as ER+ HER2- locally advanced breast cancer (LABC) subtype. First, 403 patients were excluded due to the lack of raw data of dynamic study in Picture Archiving

and Communication System (PACS). Next, patients were excluded if they had a history of previous treatment, could not confirm the initial axillary LN metastasis cytology results or final pathology results, did not have a pretreatment MRI or all four MRI sequences focused on the study (T1-weighted fat-suppressed pre-contrast, early and delayed post-contrast subtraction sequences and T2-weighted fat-suppressed sequence) or had insufficient image quality to perform lesion segmentation. Lastly, the inclusion criteria were patients who (1) had a pretreatment MRI performed at our center, (2) completed all cycles of NAC and had surgery with final pathologic report whether achieving pCR or non-pCR, (3) can be confirmed that all four sequences with sufficient quality for segmentation, resulting in a total of 265 enrolled patients.

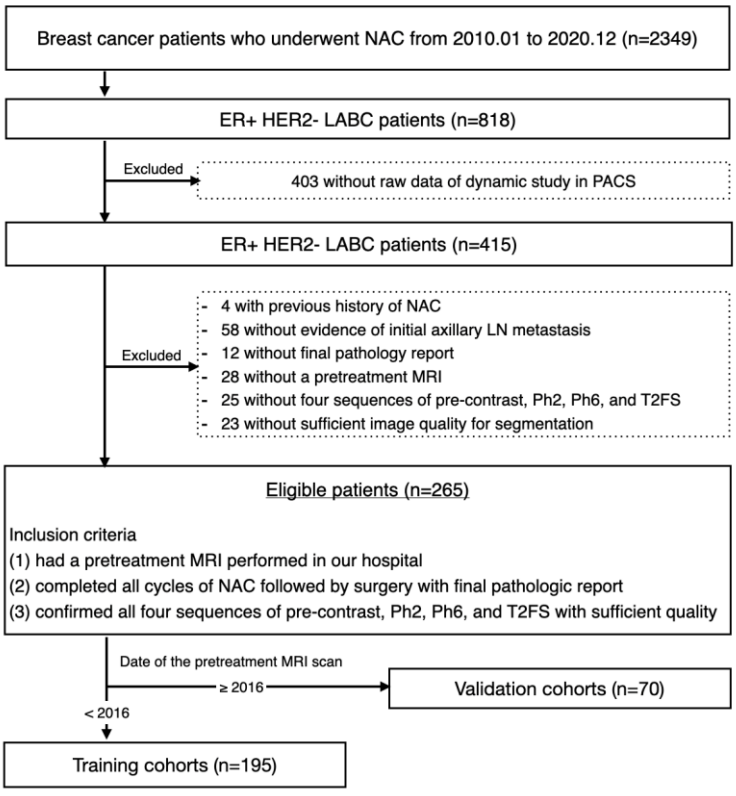


Figure 1. Flowchart of patient selection and data set

Abbreviations: NAC, neoadjuvant chemotherapy; LABC, locally advanced breast cancer;

MRI, magnetic resonance imaging; Ph2, T1-weighted fat-suppressed early post-contrast subtraction sequences; Ph6, T1-weighted fat-suppressed delayed post-contrast subtraction sequences; T2FS, T2-weighted fat-suppressed sequence; pCR, pathologic complete response.

Based on the date of the pretreatment MRI scans, it was divided into training and validation cohorts. The 195 patients who had MRI from 2010 to 2015 were included into the training cohort. Another 70 patients who had undergone MRI from 2016 to 2020 were included as the validation cohort. The patient selection process is shown in Figure 1.

2. MRI acquisition

The breast MRI examinations was performed with the patients in a prone position in a 3.0T scanner (MR750, GE Healthcare, Milwaukee, WI, USA or TrioTim, Siemens Healthcare, Erlangen, Germany using dedicated eight- or four-channel breast coil, respectively). The following images have been commonly obtained after the localizer images from one of the two types of scanners: T2 weighted fast spin echo axial images (TR/TE, 9100/100 ms; flip angle, 110°; matrix, 416×256 pixels; section thickness, 3 mm, or TR/TE, 4360/82ms; flip angle, 150°; matrix, 512×512 pixels; section thickness, 3 mm), T2 weighted short time inversion recovery (STIR) axial images (TR/TE, 5000/70 ms; inversion time, 200 ms; flip angle, 110°; matrix, 320×256 pixels; section thickness, 3 mm), and T1-weighted fat-suppressed pre-contrast and 3D dynamic post-contrast enhanced (DCE) axial images (TR/TE, 5.6/1.7 ms; flip angle, 12°; matrix, 280×512 pixels; section thickness, 3 mm, or TR/TE, 280/2.6 ms; flip angle, 65°; matrix, 343×512 pixels; section thickness, 3 mm) with one pre-contrast and six post-contrast dynamic series obtained before and after a bolus injection of 0.1 mmol/kg body weight of gadolinium-based contrast agent (Dotarem, Guerbet, Paris, France; Magnevist, Berlex Laboratories, Wayne, NJ, or Gadovist, Bayer Schering Pharma, AG, Berlin, Germany) at a rate of 2 mL/s, followed by 20 mL saline flush. Post-processing, image subtraction was

performed by subtracting the pre-contrast images from post-contrast images. The field of view was 32–34 cm for all of the MRI sequences.

3. Volume of interest (VOI) segmentation

The VOI segmentation of tumors was first semi-automatically performed along the margin of the tumor in the axial scan of T1-weighted fat-suppressed early post-contrast subtraction sequences (Ph2) by a radiologist (PJW with 5 years of experience in radiology) using a 3D-Slicer (version 5.0.2) software, and the accuracy of the image up to the 3D margin on the coronal and sagittal planes was checked with necessary modifications. For peritumoral VOI segmentation, the tumor mask was 3D dilated to a range of 1 mm and 3 mm, followed by the subtraction of the previous tumor mask (Figure 2). The same process was applied to the T1-weighted fat-suppressed delayed post-contrast subtraction sequences (Ph6) and T2-weighted fat-suppressed sequence (T2FS). In this way, 15 VOIs of tumoral, peritumoral (1mm, 3mm), and tumoral + peritumoral (1mm, 3mm) were obtained for Ph2, Ph6 and T2FS in each patient's pretreatment MRI. The process was evaluated by another senior radiologist (KMJ with 23 years of experience in radiology) to evaluate and revise the tumoral and peritumoral VOI segmentations so as to reconfirm the entire procedure.

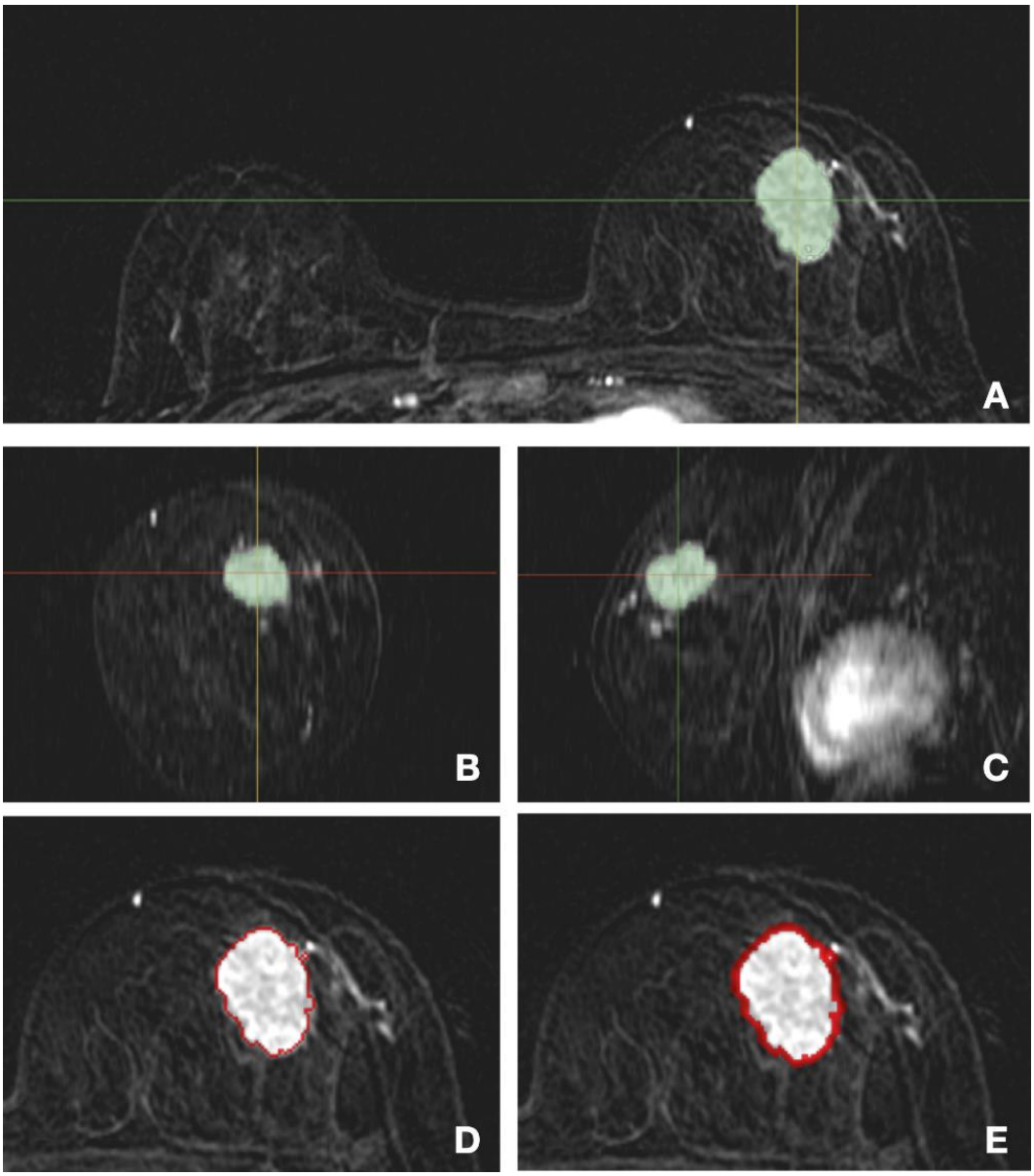


Figure 2. Segmentation on T1-weighted fat-suppressed early post-contrast subtraction sequence for a patient with histologically confirmed ER+ HER2- LABC. (A) axial (B) coronal (C) sagittal tumoral masks, (D) axial-1mm (E) axial-3mm peritumoral mask
Abbreviations: LABC, locally advanced breast cancer

4. MRI preprocessing and radiomic texture feature extraction

For the segmented VOIs, N4ITK MRI bias correction was applied to improve the non-uniformity of MR images between different patients [24], and the variation between data was minimized by normalizing the gray-level value as shown in this formula, $f(x)=\frac{s(x-\mu_x)}{\sigma_x}$ [25]. Here, x is the amplitude of the image, μ_x is the average of the image values, σ_x is the standard deviation of the image, and s is an optional scaling value set to 10 to prevent errors in the calculation of radiomic features that may occur due to a relatively large standard deviation. Then, after resampling the image with a 1×1×1mm iso-voxel, 863 radiomic features were extracted from each VOI of three sequences, respectively. Among the extracted features diagnostic features (n=12), which are information on the entire image, not VOI, and shape features among original features (n=14), which are information related to tumor size or volume measurable in conventional MRI, were excluded. The final feature set incorporated 2511 features for each sequence, and a total of 7533 features were extracted from each patient.

5. Dimension reduction

Python 3.8 was used from the data handling to the machine learning steps, and the key feature selection on the radiomic features extracted from each VOI was performed in two steps: First, the Mann-Whitney U test was used with statistical significance related to pCR or non-pCR prediction ($p<0.05$). Second, through the random forest (RF) algorithm, the top 30 features were selected for the radiomic feature importance in pCR prediction. Prior to data training, the Standard scaler was applied to adjust the deviating scales of radiomic features and reduce the influence of outliers. Additionally, the Synthetic Minority Over-sampling Technique (SMOTE) was performed to reduce the problem of overfitting toward non-pCR due to the numerical imbalance between the pCR and non-pCR groups, even if the number reflects the actual clinical pCR rate of ER+ HER2- LABC.

6. Development of pCR prediction model in the training cohort

A. Model development in each sequence

First, the process of the pCR prediction model development was performed for each sequence individually. The seven representative machine learning models were created with the key radiomic features for each of the 5 VOIs (tumor, peritumor 1 mm, peritumor 3 mm, area from tumor to peritumor 1 mm, and area from tumor to peritumor 3 mm) in MRI sequences of the training cohort; binary classification model, K-Nearest Neighbor model, Support Vector Machine (SVM), Decision Tree classifier, AdaBoost classifier, Random Forest (RF) classifier, and Light Gradient-Boosting Machine (LightGBM). Five-fold cross validation was conducted, and the optimal model for each VOI was selected based on the area under the curve (AUC) value.

B. Model development across different sequences

Next, to construct a more sophisticated pCR prediction model, the seven machine learning models were created with the sets of selected key radiomic features in combination for tumoral, peritumoral and tumoral + peritumoral VOIs across sequences from the training cohort. The training and testing processes were identical as above, and the AUC values were used to select the optimal model.

C. Model development using clinical factors

Lastly, a model incorporating clinical factors instead of radiomics features was created as a comparison group and its performance was evaluated. we selected patient age, tumor size, estrogen and progesterone receptor expression levels as clinical characteristics potentially associated with the prognosis of the disease, excluding the fixed variables; the molecular subtype of breast cancer and the presence of axillary lymph node metastasis.

7. Performance of the pCR prediction model in the validation cohort

We validated the predictive performance of the optimal models developed using radiomic features extracted from each sequence's VOIs, radiomic features combined from VOIs across different sequences, and clinical factors in the validation cohort. After calculating the AUC, precision, recall and F1 score, the predictive performance of the model was evaluated using the AUC value of the receiver operator characteristic (ROC) curve. The entire process of this study is summarized in Figure 3.

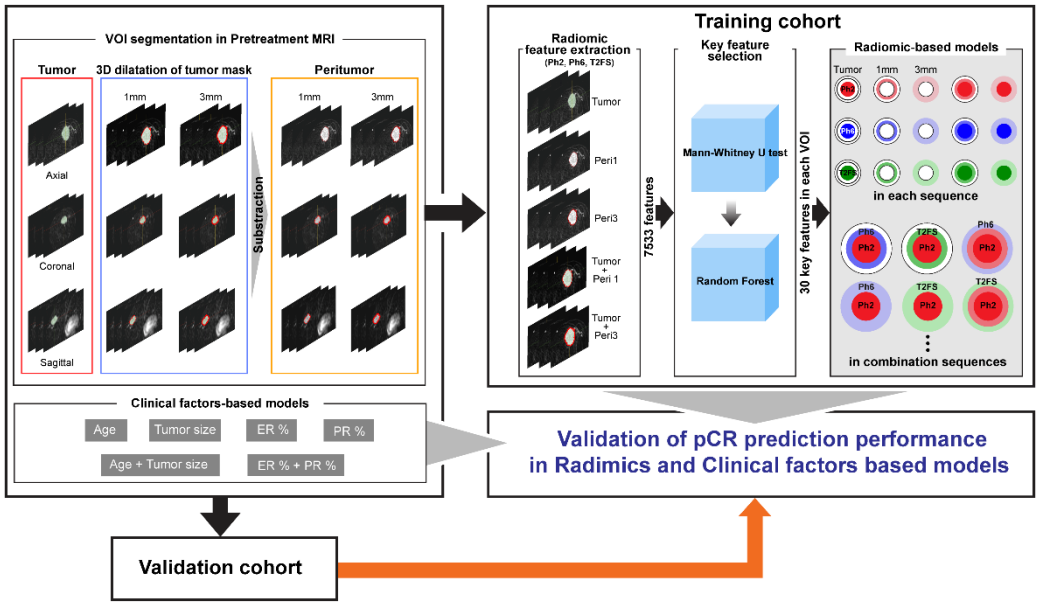


Figure 3. Radiomics workflow used in this study.

Abbreviations: Ph2, T1-weighted fat-suppressed early post-contrast subtraction sequence; Ph6, T1-weighted fat-suppressed delayed post-contrast subtraction sequences; T2FS, T2-weighted fat-suppressed sequence

8. Statistical analysis

The clinical characteristics were expressed as mean and standard deviation for continuous variables, and categorical variables were summarized as frequencies and

percentages. Continuous variables were tested using the Mann–Whitney U test, and categorical variables were compared by using the x2 test. The statistical significance was accepted where P values were <0.05.

III. Results

1. Patient characteristics

In this study, pretreatment MRIs of a total of 265 ER+ HER2- LABC patients with axillary LN metastasis were included. The clinical and histological factors of the pCR and non-pCR groups, considering the pCR which is the end point in this study, are shown in the table below (Table 1).

Table 1. Comparison of the patient characteristics between non-pCR and pCR groups.

	non-pCR	pCR	p value
	n=238 (89.8%)	n=27 (10.2%)	
Age, years	49.2 ± 9.1	48.6 ± 6.2	0.975
Tumor size, mm	14.5 ± 6.7	16.9 ± 9.3	0.210
ER expression, %	85.6 ± 20.5	75.3 ± 30.7	0.058
PR expression, %	34.3 ± 36.3	28.9 ± 38.9	0.386

Abbreviations: pCR, pathologic complete response; ER, estrogen receptor; PR, progesterone receptor.

Among the patients, 238 (89.8%) patients were non-pCR and 27 (10.2%) patients reached pCR. The mean age was 49.2 years in the non-pCR group and 48.6 years in the pCR group with no significant difference. Additionally, we could confirm from our hospital's patient cohort in 10 years, that there was no difference in tumor size and estrogen and progesterone receptor expression level between the pCR and non-pCR groups.

Next, the comparison of the training cohorts and validation cohorts is shown in the

Table 2. The two cohorts had no significant difference in the pCR rate, with 9.7% and 11.4%, respectively. And Table 2 provides other characteristics in two cohorts including patient age, tumor size and estrogen and progesterone receptor expression level.

Table 2. Comparison of the patient characteristics between the training and validation cohorts.

		Train	Validation	p value
		n=195 (73.6%)	n=70 (26.4%)	
Pathology	non-pCR	176 (90.3%)	62 (88.6%)	0.865
	pCR	19 (9.7%)	8 (11.4%)	
Age, years		48.5 ± 8.4	51.0 ± 9.9	0.043
Tumor size, mm		15.2 ± 7.1	13.0 ± 6.3	0.057
ER expression, %		85.2 ± 21.7	82.8 ± 22.5	0.408
PR expression, %		30.9 ± 35.7	41.7 ± 38.0	0.016

Abbreviations: pCR, pathologic complete response; ER, estrogen receptor; PR, progesterone receptor.

2. Radiomic texture feature composition and dimension reduction

As previously mentioned, excluding diagnostic features and shape features, 837 radiomic texture features per VOI were extracted from each patient's pretreatment MRI. These features consist of 93 original features and 744 wavelet features and Table 3 presents more details.

Table 3. The radiomic features extracted via 3D Slicer PyRadiomics

Image type	Feature Class	Num of features
original	First-order	18
original	GLCM	24
original	GLDM	14

original	GLRLM	16
original	GLSZM	16
original	NGTDM	5
wavelet-HHH	First-order	18
wavelet-HHH	GLCM	24
wavelet-HHH	GLDM	14
wavelet-HHH	GLRLM	16
wavelet-HHH	GLSZM	16
wavelet-HHH	NGTDM	5
wavelet-HHL	First-order	18
wavelet-HHL	GLCM	24
wavelet-HHL	GLDM	14
wavelet-HHL	GLRLM	16
wavelet-HHL	GLSZM	16
wavelet-HHL	NGTDM	5
wavelet-HLH	First-order	18
wavelet-HLH	GLCM	24
wavelet-HLH	GLDM	14
wavelet-HLH	GLRLM	16
wavelet-HLH	GLSZM	16
wavelet-HLH	NGTDM	5
wavelet-HLL	First-order	18
wavelet-HLL	GLCM	24
wavelet-HLL	GLDM	14
wavelet-HLL	GLRLM	16
wavelet-HLL	GLSZM	16
wavelet-HLL	NGTDM	5
wavelet-LHH	First-order	18

wavelet-LHH	GLCM	24
wavelet-LHH	GLDM	14
wavelet-LHH	GLRLM	16
wavelet-LHH	GLSZM	16
wavelet-LHH	NGTDM	5
wavelet-LHL	First-order	18
wavelet-LHL	GLCM	24
wavelet-LHL	GLDM	14
wavelet-LHL	GLRLM	16
wavelet-LHL	GLSZM	16
wavelet-LHL	NGTDM	5
wavelet-LLH	First-order	18
wavelet-LLH	GLCM	24
wavelet-LLH	GLDM	14
wavelet-LLH	GLRLM	16
wavelet-LLH	GLSZM	16
wavelet-LLH	NGTDM	5
wavelet-LLL	First-order	18
wavelet-LLL	GLCM	24
wavelet-LLL	GLDM	14
wavelet-LLL	GLRLM	16
wavelet-LLL	GLSZM	16
wavelet-LLL	NGTDM	5

Abbreviations: GLCM, gray-level co-occurrence matrix; GLDM, gray-level dependence matrix; GLRLM, gray-level run-length matrix; GLSZM, gray-level size- zone matrix; NGTDM, neighboring gray tone difference matrix.

837 radiomic texture features included 93 original (first-order, shape, gray-level

co-occurrence matrix (GLCM), gray-level dependence matrix (GLDM), gray-level run-length matrix (GLRLM), gray-level size-zone matrix (GLSZM), neighboring gray tone difference matrix (NGTDM)), and 744 wavelet features. And first, Mann-Whitney U test was used to remove 16 features showing no significant difference between pCR and non-pCR, then for the remaining 821 features are ranked by the importance values from the Random Forest (RF) algorithm, and the top 30 features were chosen.

3. Performance of the pCR prediction model in each sequence

Table 4 presents the final pCR prediction performance in the validation cohort that were confirmed by applying the optimal machine learning models developed from each of the five types of VOIs in a sequence. A general look at the table reveals that the models derived from the Ph2 and T2FS show relatively high AUC values, while even the best performing models in the Ph6 do not exceed an AUC value of 0.9. And the best model for pCR prediction of NAC in ER+ HER2- LABC in respective three sequences is SVM model of tumor + peritumor 1mm on Ph2 (AUC = 0.9447, recall = 91%, precision = 91%, and F1 score = 91%). The ROC curves and AUCs of the fifteen models in the validation cohorts are shown in Figure 4, and it can be confirmed once again that the overall high AUC value is shown in Ph2.

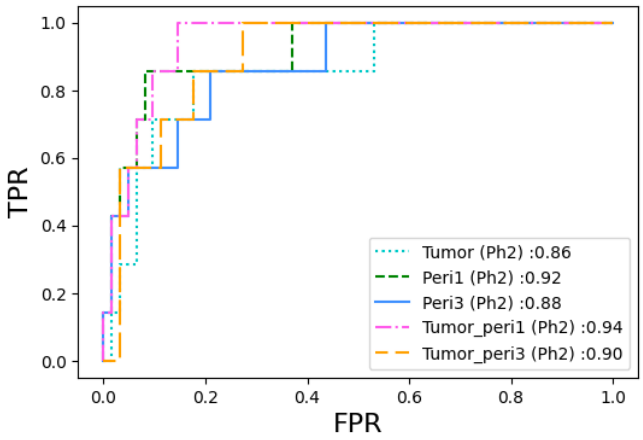
Table 4. The predictive performance of pCR in the validation cohort using the optimal machine learning models developed for tumoral, peritumoral, and tumoral + peritumoral VOIs in each sequence

Type	Rank	AUC	Precision	Recall	F1-score	Best Model
Tumor (Ph2)	8	0.8594	0.9093	0.8696	0.8840	SVC
Tumor (Ph6)	15	0.7005	0.8731	0.7826	0.8166	Random Forest Classifier
Tumor (T2FS)	12	0.8018	0.8984	0.6667	0.7327	SVC
Peri1 (Ph2)	3	0.9171	0.9335	0.8986	0.9098	Logistic Regression

Peri1 (Ph6)	9	0.8422	0.9051	0.8986	0.9015	K Neighbors Classifier
Peri1 (T2FS)	13	0.7995	0.8851	0.8406	0.8582	Random Forest Classifier
Peri3 (Ph2)	6	0.8756	0.8934	0.8696	0.8795	Random Forest Classifier
Peri3 (Ph6)	10	0.8249	0.8366	0.7826	0.8066	AdaBoost Classifier
Peri3 (T2FS)	11	0.8134	0.9215	0.6522	0.7203	Logistic Regression
Tumor_peri1 (Ph2)	1	0.9447	0.9130	0.9130	0.9130	SVC
Tumor_peri1 (Ph6)	5	0.8917	0.9193	0.9275	0.9185	AdaBoost Classifier
Tumor_peri1 (T2FS)	2	0.9240	0.9442	0.9275	0.9330	AdaBoost Classifier
Tumor_peri3 (Ph2)	4	0.9009	0.8061	0.8841	0.8433	Random Forest Classifier
Tumor_peri3 (Ph6)	14	0.7189	0.8450	0.7246	0.7718	Decision Tree Classifier
Tumor_peri3 (T2FS)	7	0.8710	0.9193	0.9275	0.9185	SVC

Abbreviations: Ph2, T1-weighted fat-suppressed early post-contrast subtraction sequence; Ph6, T1-weighted fat-suppressed delayed post-contrast subtraction sequences; T2FS, T2-weighted fat-suppressed sequence, Peri1, peritumoral region, 1mm; Peri3, peritumoral region, 3mm Tumor_peri1, tumoral + 1mm peritumoral region; Tumor_peri3, tumoral + 3mm peritumoral region; SVM, Support Vector Machine.

(A)



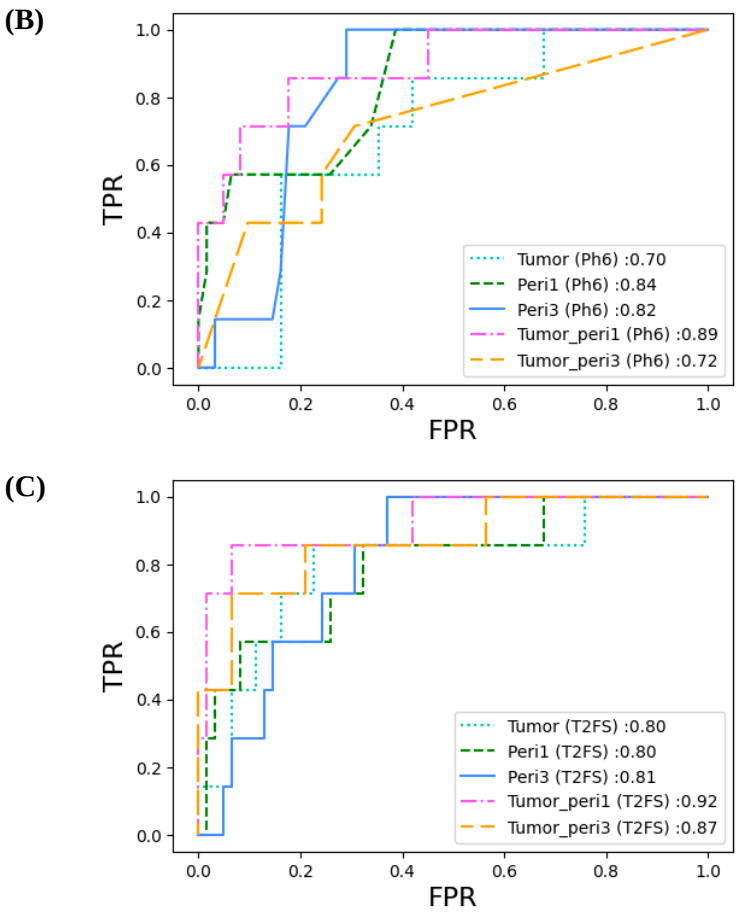


Figure 4. The ROC curve for the predictive performance of pCR in the validation cohort using the optimal machine learning models in each sequence.
(A) AUC of the tumor, Peri1, Peri3, Tumor_peri1, and Tumor_peri3 on Ph2, (B) AUC of the tumor, Peri1, Peri3, Tumor_peri1 and Tumor_peri3 on Ph6, (C) AUC of the tumor, Peri1, Peri3, Tumor_peri1, and Tumor_peri3 on T2FS.
Abbreviations: Ph2, T1-weighted fat-suppressed early post-contrast subtraction sequence; Ph6, T1-weighted fat-suppressed delayed post-contrast subtraction sequences; T2FS, T2-weighted fat-suppressed sequence, Peri1, peritumoral region, 1mm; Peri3, peritumoral region, 3mm, Tumor_peri1, tumoral + 1mm peritumoral region; Tumor_peri3, tumoral + 3mm peritumoral region; FPR false positive rate; TPR true positive rate

4. Performance of the pCR prediction in combination model across sequences

We confirmed the predictive performance of pCR for the optimal machine learning model developed from 75 VOIs combining tumoral and peritumoral regions in two different sequences for the validation cohort. The KNN model with key radiomics features derived from a combination of VOIs ranging from the tumor to peritumor 1mm in Ph2 and peritumor 3mm VOI in T2FS exhibited the best pCR prediction performance with an AUC of 0.96. The pCR prediction performances based on combination of tumoral and peritumoral regions of different sequences are shown in Table 5. Additionally, Figure 5 compares the ROC curve of the optimal model developed using the tumoral VOI, the peritumoral 1mm VOI, and the tumoral-peritumoral 1mm VOI of Ph2, and the peritumoral 3mm VOI of T2FS, which are the components of the combination model. And the Cochran Q test verified that there is a significant difference between these five models at the $p<0.001$ level.

Table 5. Predictive performance of pCR in the validation cohort using the optimal machine learning models based on combination of tumoral and peritumoral regions across sequences

Type1	Type2	Rank	AUC	Precision	Recall	F1-score	Best Model
Tumor_peri1 (Ph2)	Tumor (Ph6)	5	0.9378	0.9207	0.9130	0.8824	SVC
Tumor_peri1 (Ph2)	Tumor (T2FS)	9	0.9194	0.9335	0.8986	0.9098	SVC
Tumor_peri1 (Ph2)	Peri1 (Ph6)	6	0.9309	0.9329	0.9275	0.9088	AdaBoost
Tumor_peri1 (Ph2)	Peri1 (T2FS)	21	0.9032	0.9196	0.8406	0.8655	SVC
Tumor_peri1 (Ph2)	Peri3 (Ph6)	9	0.9194	0.8981	0.9130	0.8970	LGBM
Tumor_peri1 (Ph2)	Peri3 (T2FS)	1	0.9631	0.9234	0.9275	0.9250	K Neighbors
Tumor_peri1 (Ph2)	Tumor_peri1 (Ph6)	14	0.9101	0.8981	0.9130	0.8970	LGBM

Tumor_peri1 (Ph2)	Tumor_peri1 (T2FS)	12	0.9171	0.9137	0.8841	0.8950	SVC
Tumor_peri1 (Ph2)	Tumor_peri3 (Ph6)	6	0.9309	0.9234	0.9275	0.9250	LGBM
Tumor_peri1 (Ph2)	Tumor_peri3 (T2FS)	2	0.9505	0.9325	0.9275	0.9297	K Neighbors
Tumor_peri1 (T2FS)	Tumor (Ph2)	25	0.8963	0.8987	0.8841	0.8904	LGBM
Tumor_peri1 (T2FS)	Tumor (Ph6)	51	0.8594	0.9093	0.8696	0.8840	AdaBoost
Tumor_peri1 (T2FS)	Peri1 (Ph2)	9	0.9194	0.9420	0.9420	0.9420	Logistic
Tumor_peri1 (T2FS)	Peri1 (Ph6)	28	0.8940	0.9188	0.8986	0.9062	LGBM
Tumor_peri1 (T2FS)	Peri3 (Ph2)	17	0.9078	0.9420	0.9420	0.9420	LGBM
Tumor_peri1 (T2FS)	Peri3 (Ph6)	55	0.8525	0.9021	0.8406	0.8624	Logistic
Tumor_peri1 (T2FS)	Tumor_peri1 (Ph6)	41	0.8779	0.8963	0.8116	0.8411	AdaBoost
Tumor_peri1 (T2FS)	Tumor_peri3 (Ph2)	14	0.9101	0.8981	0.9130	0.8970	Random Forest
Tumor_peri1 (T2FS)	Tumor_peri3 (Ph6)	58	0.8364	0.8816	0.8261	0.8478	Logistic
Peri1 (Ph2)	Tumor (Ph6)	31	0.8894	0.9033	0.9130	0.9065	LGBM
Peri1 (Ph2)	Tumor (T2FS)	19	0.9055	0.9188	0.8986	0.9062	Logistic
Peri1 (Ph2)	Peri1 (Ph6)	38	0.8802	0.9021	0.8406	0.8624	Logistic
Peri1 (Ph2)	Peri1 (T2FS)	32	0.8871	0.9130	0.9130	0.9130	K Neighbors
Peri1 (Ph2)	Peri3 (Ph6)	43	0.8733	0.8922	0.8986	0.8951	Random Forest
Peri1 (Ph2)	Peri3 (T2FS)	17	0.9078	0.9130	0.9130	0.9130	Logistic
Peri1 (Ph2)	Tumor_peri1 (Ph6)	48	0.8641	0.9455	0.9420	0.9313	LGBM
Peri1 (Ph2)	Tumor_peri3 (Ph6)	32	0.8871	0.9130	0.9130	0.9130	Logistic
Peri1 (Ph2)	Tumor_peri3 (T2FS)	30	0.8917	0.9420	0.9420	0.9420	Random Forest
Tumor_peri3 (Ph2)	Tumor (Ph6)	71	0.7926	0.8032	0.8551	0.8284	Random Forest
Tumor_peri3 (Ph2)	Tumor (T2FS)	43	0.8733	0.9051	0.8986	0.9015	Random Forest

Tumor_peri3 (Ph2)	Peri1 (Ph6)	3	0.9401	0.9193	0.9275	0.9185	LGBM
Tumor_peri3 (Ph2)	Peri1 (T2FS)	56	0.8433	0.9033	0.9130	0.9065	Random Forest
Tumor_peri3 (Ph2)	Peri3 (Ph6)	21	0.9032	0.8507	0.8841	0.8627	Random Forest
Tumor_peri3 (Ph2)	Peri3 (T2FS)	38	0.8802	0.8346	0.8551	0.8441	LGBM
Tumor_peri3 (Ph2)	Tumor_peri1 (Ph6)	19	0.9055	0.8802	0.8986	0.8859	LGBM
Tumor_peri3 (Ph2)	Tumor_peri3 (Ph6)	43	0.8733	0.8611	0.8696	0.8651	Random Forest
Tumor_peri3 (Ph2)	Tumor_peri3 (T2FS)	8	0.9274	0.9137	0.8841	0.8950	K Neighbors
Tumor_peri1 (Ph6)	Tumor (Ph2)	12	0.9171	0.8074	0.8986	0.8505	SVC
Tumor_peri1 (Ph6)	Tumor (T2FS)	36	0.8825	0.9207	0.9130	0.8824	SVC
Tumor_peri1 (Ph6)	Peri1 (T2FS)	48	0.8641	0.9196	0.8406	0.8655	SVC
Tumor_peri1 (Ph6)	Peri3 (Ph2)	23	0.8986	0.8890	0.8551	0.8688	Logistic
Tumor_peri1 (Ph6)	Peri3 (T2FS)	41	0.8779	0.9130	0.9130	0.9130	Random Forest
Tumor_peri1 (Ph6)	Tumor_peri3 (T2FS)	38	0.8802	0.9130	0.9130	0.9130	Random Forest
Peri3 (Ph2)	Tumor (Ph6)	35	0.8836	0.9171	0.8261	0.8547	Decision Tree
Peri3 (Ph2)	Tumor (T2FS)	59	0.8318	0.8915	0.7826	0.8199	Random Forest
Peri3 (Ph2)	Peri1 (Ph6)	23	0.8986	0.9335	0.8986	0.9098	Random Forest
Peri3 (Ph2)	Peri1 (T2FS)	51	0.8594	0.8785	0.8116	0.8374	Random Forest
Peri3 (Ph2)	Peri3 (Ph6)	47	0.8652	0.8915	0.7826	0.8199	AdaBoost
Peri3 (Ph2)	Peri3 (T2FS)	16	0.9090	0.9281	0.7536	0.8016	AdaBoost
Peri3 (Ph2)	Tumor_peri3 (Ph6)	62	0.8249	0.8642	0.8261	0.8425	Random Forest
Peri3 (Ph2)	Tumor_peri3 (T2FS)	46	0.8687	0.8915	0.7826	0.8199	Random Forest
Tumor_peri3 (T2FS)	Tumor (Ph2)	28	0.8940	0.9329	0.9275	0.9088	SVC
Tumor_peri3 (T2FS)	Tumor (Ph6)	73	0.7857	0.8890	0.8551	0.8688	Random Forest
Tumor_peri3 (T2FS)	Peri1 (Ph6)	34	0.8848	0.9188	0.8986	0.9062	Random Forest

Tumor_peri3 (T2FS)	Peri3 (Ph6)	69	0.7995	0.8872	0.7536	0.7985	Random Forest
Tumor_peri3 (T2FS)	Tumor_peri3 (Ph6)	74	0.7834	0.8609	0.8116	0.8324	Random Forest
Tumor (Ph2)	Tumor (Ph6)	67	0.8018	0.8261	0.8261	0.8261	Random Forest
Tumor (Ph2)	Tumor (T2FS)	54	0.8571	0.8798	0.6957	0.7549	Logistic
Tumor (Ph2)	Peri1 (Ph6)	3	0.9401	0.9137	0.8841	0.8950	AdaBoost
Tumor (Ph2)	Peri1 (T2FS)	36	0.8825	0.8934	0.8696	0.8795	Logistic
Tumor (Ph2)	Peri3 (Ph6)	61	0.8272	0.8074	0.8986	0.8505	SVC
Tumor (Ph2)	Peri3 (T2FS)	25	0.8963	0.8981	0.9130	0.8970	SVC
Tumor (Ph2)	Tumor_peri3 (Ph6)	57	0.8410	0.8851	0.8406	0.8582	Random Forest
Peri1 (Ph6)	Tumor (T2FS)	51	0.8594	0.9207	0.9130	0.8824	SVC
Peri1 (Ph6)	Peri1 (T2FS)	59	0.8318	0.8915	0.7826	0.8199	SVC
Peri1 (Ph6)	Peri3 (T2FS)	25	0.8963	0.9021	0.8406	0.8624	Random Forest
Peri3 (Ph6)	Tumor (T2FS)	66	0.8111	0.8933	0.6087	0.6847	SVC
Peri3 (Ph6)	Peri1 (T2FS)	50	0.8618	0.9234	0.9275	0.9250	Random Forest
Peri3 (Ph6)	Peri3 (T2FS)	70	0.7972	0.8582	0.5217	0.6100	Logistic
Peri3 (T2FS)	Tumor (Ph6)	67	0.8018	0.8915	0.7826	0.8199	Random Forest
Peri3 (T2FS)	Tumor_peri3 (Ph6)	75	0.7097	0.8148	0.7681	0.7900	Random Forest
Tumor (T2FS)	Tumor (Ph6)	63	0.8226	0.8872	0.7536	0.7985	Random Forest
Tumor (T2FS)	Tumor_peri3 (Ph6)	72	0.7880	0.8725	0.8551	0.8629	LGBM
Peri1 (T2FS)	Tumor (Ph6)	64	0.8203	0.8922	0.8986	0.8951	Random Forest
Peri1 (T2FS)	Tumor_peri3 (Ph6)	65	0.8157	0.9207	0.9130	0.8824	SVC

Abbreviations: Ph2, second-post contrast subtraction image; Ph6, sixth-post contrast subtraction image; T2FS, T2 weighted fat-saturated image, Peri1, peritumoral region, 1mm; Peri3, peritumoral region, 3mm Tumor_peri1, tumoral + 1mm peritumoral region;

Tumor_peri3, tumoral + 3mm peritumoral region; SVM, Support Vector Machine; LGBM, Light Gradient-Boosting Machine.

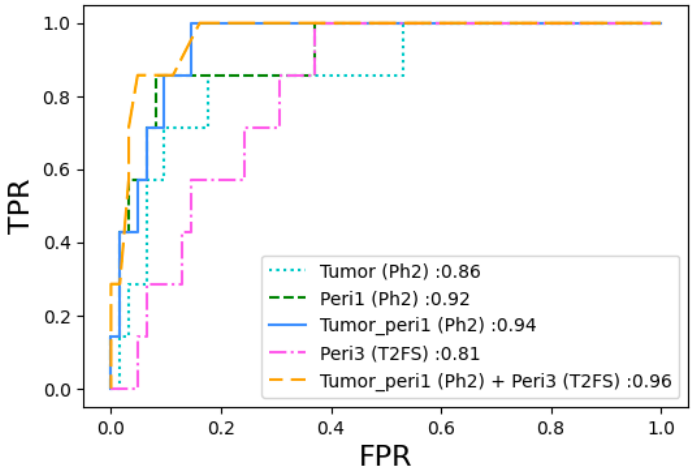


Figure 5. Comparison of the pCR prediction performance: The best combination model of the VOI from tumor to peritumor 1 mm in Ph2 and the peritumor 3 mm VOI in T2FS as well as the respective component VOI models.

Abbreviations: Ph2, T1-weighted fat-suppressed early post-contrast subtraction sequence; T2FS, T2-weighted fat-suppressed sequence; Peri1, peritumoral region, 1mm; Tumor_peri1, tumoral + 1mm peritumoral region; Peri3, peritumoral region, 3mm; FPR false positive rate; TPR true positive rate

5. Diagnostic performance of clinical model

Furthermore, we applied the same process to confirm the predictive performance of pCR for clinical factors that may be associated with the patient's prognosis in breast cancer, such as patient age, tumor size, and estrogen and progesterone expression levels.

In the validation cohort, the AUC values were generally low for pCR prediction performance when compared to the radiomics models. The AUC values for patient age, tumor size, and the combination model of patient age and tumor size were 0.63, 0.81, and 0.67, respectively. The AUC values for estrogen and progesterone expression levels and

their combination model were 0.68, 0.64, and 0.53, respectively. The results are summarized in Table 6 and Figure 6.

Table 6. pCR prediction performances of the clinical models in the validation cohort

Features	AUC	Precision	Recall	F1-score	Model
Age	0.6261	0.8766	0.7750	0.8180	K Neighbors Classifier
Tumor size	0.8063	0.9067	0.7500	0.8063	Logistic Regression
Age + Tumor size	0.6667	0.8936	0.6000	0.6933	Logistic Regression
ER expression	0.6847	0.8708	0.7250	0.7847	Decision Tree Classifier
PR expression	0.6396	0.8680	0.7000	0.7675	Random Forest Classifier
ER + PR expression	0.5315	0.8766	0.7750	0.8180	K Neighbors Classifier

Abbreviations: ER, estrogen receptor; PR progesterone receptor

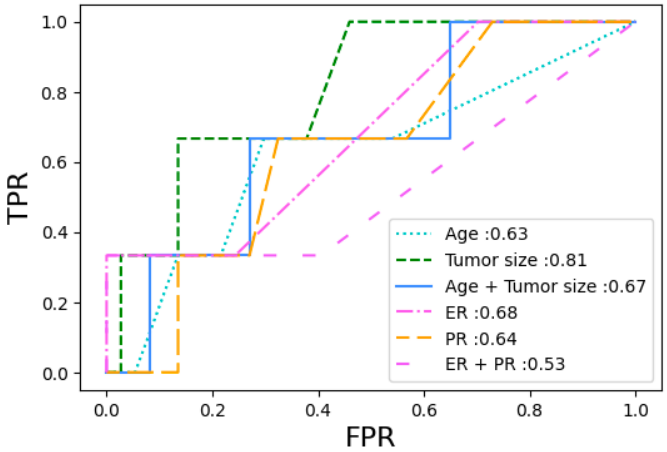


Figure 6. The ROC curve for pCR prediction performances of the clinical models in the validation cohort

Abbreviations: ER, estrogen receptor; PR progesterone receptor, FPR false positive rate; TPR true positive rate.

IV. Discussion

The ER+ HER2- locally advanced breast cancer (LABC) has a poor pathologic complete response (pCR) rate of around 10%, compared to the 3-40% pCR rates of other molecular subtypes upon the operation after neoadjuvant chemotherapy (NAC) [9]. Therefore, this study aimed to classify the ER+ HER2-LABC patients with a high probability of giving an effective response to the NAC, using pretreatment MRI, which is an important modality for non-invasive assessment of breast cancer [6,7,10]. Several recent studies have attempted to create a prognosis prediction model for breast cancer through the radiomic texture feature extraction with respect to the pretreatment MRI applied in this study [10,21,26]. However, all these studies were conducted on the heterogeneous molecular subtype that includes ER+ HER2- LABC, the focus in this study, only as a small part of the entire cohort.

To construct a sophisticated model for the pCR prediction after NAC in ER+ HER2- LABC patients, the radiomic texture features of MRI were extracted from the tumor, peritumor 1 mm, peritumor 3 mm, area from tumor to peritumor 1 mm and area from tumor to peritumor 3 mm, for each of early post-contrast subtraction image, delayed post-contrast subtraction image, T2 weighted fat-saturated image. In line with previous studies, it was also further established that early post-contrast subtraction image predominantly contains the most useful texture features in the machine learning models as a single sequence model evaluation [27,28,29]. The inclusion of the delayed post-contrast subtraction image in this study was based on the previous study by Jin et al., claiming that the texture heterogeneity is better reflected in the delayed enhanced phase for breast tumors [30]; however, the model incorporating the texture features of the tumor in the delayed phase did not produce more powerful data in comparison to other sequences in our study.

Furthermore, not only the tumoral region as the basis of determining the VOI for the radiomic feature extraction in MRI but also the peritumoral region reported to form a microenvironment that affects the NAC response [21,22], were included in this study. Until

recently, there have been studies to include the peritumoral region for investigating an extended area from the tumoral region to the peritumoral region on a single MRI sequence [22]. In this study, on the other hand, the tumoral and peritumoral regions in combination across sequences were examined to construct a more sophisticated model that reflects more important sequences related to each region. As a result, the model with the combination of the tumoral region in the early enhanced phase and the peritumoral region in the T2FS exhibited the highest AUC. This finding is significant as it coincides with the basic principle of MRI, that signal alterations for tumors generally arise greater in the T2FS [31,32]. In the future, more elaborate models need to be performed by combining the tumoral and peritumoral regions across different sequences and validated for other molecular subtypes of breast cancer.

There are several limitations in this study. First, there is a possibility of selection bias as the study was conducted based on a retrospective design at a single tertiary referral center. Second, it is difficult to generalize the results of this study to all breast cancer patients as this study intentionally focused on one molecular subtype ER+ HER2- LABC, the subtype with the lowest NAC response. For this, follow-up studies should be conducted. Third, the target patients were those who received the pretreatment MRI and operation and whose final pathological results were available for the 10 years, which is a long period of time. However, the pCR rates after NAC did not significantly vary in patients with ER+ HER2- breast cancer during this period and the lack of significant difference in the pCR rate based on 2016 was the first to be checked in this study. And techniques such as bias correction, gray scale normalization and standard scaler were applied in the preprocessing to reduce the non-uniformity of MRI images. Fourth, regarding the potential clinical utility, more time seems necessary for immediate clinical application of the findings through rapid and reliable automatic segmentation. To provide accurate key texture features to constitute machine learning models, an accurate VOI segmentation process on the tumor is a prerequisite. Although this study used a 3D slicer to produce images in a semi-automatic way, the reliability of the VOI produced by the program decreased as the irregularity of

tumor margin increased, which demanded the modification by a radiologist and reconfirmation by a senior radiologist for the process of refining the VOI segmentation. Lastly, the most fundamental limitation is found in the revision of treatment plans for ER+ HER2- LABC patients. Despite a mere 10% pCR rate after NAC, NAC is still given to ER+ HER2- LABC patients mainly because more effective and specific treatments for this patient group are still on the way. Nevertheless, the accumulation of such studies, where the NAC-effective and non-effective groups are distinguished, is thought to provide continuous support and assistance in the search for therapeutic strategies for clinicians.

V. Conclusion

To assess the NAC response of ER+ HER2- LABC patients in pretreatment MRI, this study applied the radiomic texture features in the tumoral and peritumoral regions across MRI sequences. We suggest that the combination machine learning model incorporating tumoral and peritumoral texture features across the different MRI sequences can make more accurate pCR prediction for NAC response of these patients. Our results are also anticipated to make a potential contribution to the development of clinical therapeutic strategies.

REFERENCES

1. Azadnajafabad S, Saeedi Moghaddam S, Keykhaei M, Shobeiri P, Rezaei N, Ghasemi E, et al. Expansion of the quality of care index on breast cancer and its risk factors using the Global Burden of Disease Study 2019. Cancer Medicine, 2023;12.2: 1729-1743.
2. Belete AM, Aynalem YA, Gemeda BN, Demelew TM, Shiferaw WS, et al. The Effect of Estrogen Receptor Status on Survival in Breast Cancer Patients in Ethiopia. Retrospective Cohort Study. Breast Cancer: Targets and Therapy, 2022;153-161.
3. Nam KJ, Park H, Ko ES, Lim Y, Cho HH, & Lee JE, et al. Radiomics signature on 3T dynamic contrast-enhanced magnetic resonance imaging for estrogen receptor-positive invasive breast cancers: preliminary results for correlation with Oncotype DX recurrence scores. Medicine, 2019;98.23.
4. von Minckwitz G, Raab G, Caputo A, Schütte M, Hilfrich J, Blohmer JU, et al. Doxorubicin with cyclophosphamide followed by docetaxel every 21 days compared with doxorubicin and docetaxel every 14 days as preoperative treatment in operable breast cancer: the GEPARDUO study of the German Breast Group. Journal of clinical oncology, 2005;23.12: 2676-2685.
5. Chica-Parrado MR, Godoy-Ortiz A, Jiménez B, Ribelles N, Barragan I, et al. Resistance to neoadjuvant treatment in breast cancer: clinicopathological and molecular predictors. Cancers, 2020;12.8: 2012.
6. Kim MJ, Kim EK, Park S, Moon HJ, Kim S, Park BW. Evaluation with 3.0-T MR imaging: predicting the pathological response of triple-negative breast cancer treated with anthracycline and taxane neoadjuvant chemotherapy. Acta Radiologica, 2015;56.9: 1069-1077.
7. Chen L, Yang Q, Bao J, Liu D, Huang X, Wang J. Direct comparison of PET/CT and MRI to predict the pathological response to neoadjuvant chemotherapy in breast cancer: a meta-analysis. Scientific reports, 2017;7.1: 1-10.
8. Shohdy KS, Almeldin DS, Fekry MA, Ismail MA, AboElmaaref NA, ElSadany EG, et

al. Pathological responses and survival outcomes in patients with locally advanced breast cancer after neoadjuvant chemotherapy: a single-institute experience. Journal of the Egyptian National Cancer Institute, 2021;33: 1-10.

9. Spring LM, Fell G, Arfe A, Sharma C, Greenup R, Reynolds KL, et al. Pathologic complete response after neoadjuvant chemotherapy and impact on breast cancer recurrence and survival: A comprehensive meta-analysispCR and association with clinical outcomes in breast cancer. Clinical cancer research, 2020;26.12: 2838-2848.
10. Bian T, Wu Z, Lin Q, Wang H, Ge Y, Duan S, et al. Radiomic signatures derived from multiparametric MRI for the pretreatment prediction of response to neoadjuvant chemotherapy in breast cancer. The British Journal of Radiology, 2020;93: 20200287.
11. Ko ES, Han BK, Kim RB, Ko EY, Shin JH, Hahn SY, et al. Analysis of factors that influence the accuracy of magnetic resonance imaging for predicting response after neoadjuvant chemotherapy in locally advanced breast cancer. Annals of surgical oncology, 2013;20: 2562-2568.
12. Overcast WB., Davis KM, Ho CY, Hutchins GD, Green MA, Graner BD, Veronesi MC, et al. Advanced imaging techniques for neuro-oncologic tumor diagnosis, with an emphasis on PET-MRI imaging of malignant brain tumors. Current Oncology Reports, 2021;23: 1-15.
13. Varghese BA, Cen SY, Hwang DH, Duddalwar VA. Texture analysis of imaging: what radiologists need to know. American Journal of Roentgenology, 2019;212.3: 520-528.
14. Wagner MW, Namdar K, Biswas A, Monah S, Khalvati F, Ertl-Wagner BB. Radiomics, machine learning, and artificial intelligence—what the neuroradiologist needs to know. Neuroradiology, 2021;1-11.
15. Parikh J, Selmi M, Charles-Edwards G, Glendenning J, Ganeshan B, Verma H, et al. Changes in primary breast cancer heterogeneity may augment midtreatment MR

imaging assessment of response to neoadjuvant chemotherapy. Radiology, 2014;272.1: 100-112.

16. Henderson S, Purdie C, Michie C, Evans A, Lerski R, Johnston M, et al. Interim heterogeneity changes measured using entropy texture features on T2-weighted MRI at 3.0 T are associated with pathological response to neoadjuvant chemotherapy in primary breast cancer. European radiology, 2017;27: 4602-4611.
17. Panico C, Ferrara F, Woitek R, D'Angelo A, Di Paola V, Bufi E, et al. Staging Breast Cancer with MRI, the T. A Key Role in the Neoadjuvant Setting. Cancers, 2022;14.23: 5786.
18. Umutlu L, Kirchner J, Bruckmann NM, Morawitz J, Antoch G, Ting S, et al. Multiparametric 18F-FDG PET/MRI-based Radiomics for prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer. Cancers, 2022;14.7: 1727.
19. Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari P, et al. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. Breast Cancer Research, 2017;19.1: 1-14.
20. Herrero Vicent C, Tudela X, Moreno Ruiz P, Pedralva V, Jiménez Pastor A, Ahicart D, et al. Machine Learning Models and Multiparametric Magnetic Resonance Imaging for the Prediction of Pathologic Response to Neoadjuvant Chemotherapy in Breast Cancer. Cancers, 2022;14.14: 3508.
21. Han X, Cao W, Wu L, Liang C. Radiomics assessment of the tumor immune microenvironment to predict outcomes in breast cancer. Frontiers in immunology, 2022;12: 5356.
22. Li C, Lu N, He Z, Tan Y, Liu Y, Chen Y, et al. A Noninvasive Tool Based on Magnetic Resonance Imaging Radiomics for the Preoperative Prediction of Pathological Complete Response to Neoadjuvant Chemotherapy in Breast Cancer. Annals of

Surgical Oncology, 2022;29.12: 7685-7693.

23. Granzier RW, Ibrahim A, Primakov SP, Samiei S, van Nijnatten TJ, de Boer M, et al. MRI-based radiomics analysis for the pretreatment prediction of pathologic complete tumor response to neoadjuvant systemic therapy in breast cancer patients: a multicenter study. Cancers, 2021;13.10: 2447.
24. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC. N4ITK: improved N3 bias correction. IEEE transactions on medical imaging, 2010;29.6: 1310-1320.
25. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magnetic resonance imaging, 2004;22.1: 81-91.
26. Chamming's F, Ueno Y, Ferré R, Kao E, Jannot AS, Chong J, et al. Features from computerized texture analysis of breast cancers at pretreatment MR imaging are associated with response to neoadjuvant chemotherapy. Radiology, 2018;286.2: 412-420.
27. Zhou J, Lu J, Gao C, Zeng J, Zhou C, Lai X, et al. Predicting the response to neoadjuvant chemotherapy for breast cancer: wavelet transforming radiomics in MRI. BMC cancer, 2020;20: 1-10.
28. Hussain L, Huang P, Nguyen T, Lone KJ, Ali A, Khan MS, et al. Machine learning classification of texture features of MRI breast tumor and peri-tumor of combined pre-and early treatment predicts pathologic complete response. BioMedical Engineering OnLine, 2021;20.1: 1-23.
29. Li Q, Xiao Q, Li J, Wang Z, Wang H, Gu Y. Value of machine learning with multiphases ce-mri radiomics for early prediction of pathological complete response to neoadjuvant therapy in her2-positive invasive breast cancer. Cancer Management and Research, 2021;5053-5062.
30. Kim JY, Kim JJ, Hwangbo L, Suh HB, Kim S, Choo KS, et al. Kinetic heterogeneity of breast cancer determined using computer-aided diagnosis of preoperative

MRI scans: relationship to distant metastasis-free survival. Radiology, 2020;295.3: 517-526.

31.Smith AS, Weinstein MA, Modic MT, Pavlicek W, Rogers LR, Budd TG, et al. Magnetic resonance with marked T2-weighted images: improved demonstration of brain lesions, tumor, and edema. American journal of neuroradiology, 1985;6.5: 691-697.

32.Kociemba A, Pyda M, Katulska K, Łanocha M, Siniawski A, Janus M, Grajek S. Comparison of diffusion-weighted with T2-weighted imaging for detection of edema in acute myocardial infarction. Journal of Cardiovascular Magnetic Resonance, 2013;15.1: 1-8.

APPENDICIES

Table 4-1. The predictive performance of pCR in the validation cohort using the optimal machine learning models developed for tumoral, peritumoral, and tumoral + peritumoral VOIs in each sequence

Type	Rank	AUC	Sensitivity	Specitivity	Precision	Recall	F1-score	Best Model
Tumor (Ph2)	8	0.8594	0.8571	0.8226	0.9093	0.8696	0.8840	SVC
Tumor (Ph6)	15	0.7005	0.8571	0.5806	0.8731	0.7826	0.8166	RandomForest
Tumor (T2FS)	12	0.8018	0.8571	0.7742	0.8984	0.6667	0.7327	SVC
Peri1 (Ph2)	3	0.9171	0.8571	0.9194	0.9335	0.8986	0.9098	Logistic
Peri1 (Ph6)	9	0.8422	1.0000	0.6129	0.9051	0.8986	0.9015	KNeighbors
Peri1 (T2FS)	13	0.7995	0.8571	0.6774	0.8851	0.8406	0.8582	RandomForest
Peri3 (Ph2)	6	0.8756	0.8571	0.7903	0.8934	0.8696	0.8795	RandomForest
Peri3 (Ph6)	10	0.8249	1.0000	0.7097	0.8366	0.7826	0.8066	AdaBoost
Peri3 (T2FS)	11	0.8134	1.0000	0.6290	0.9215	0.6522	0.7203	Logistic
Tumor_peri1 (Ph2)	1	0.9447	1.0000	0.8548	0.9130	0.9130	0.9130	SVC
Tumor_peri1 (Ph6)	5	0.8917	0.8571	0.8226	0.9193	0.9275	0.9185	AdaBoost
Tumor_peri1 (T2FS)	2	0.9240	0.8571	0.9355	0.9442	0.9275	0.9330	AdaBoost
Tumor_peri3 (Ph2)	4	0.9009	1.0000	0.7258	0.8061	0.8841	0.8433	RandomForest
Tumor_peri3 (Ph6)	14	0.7189	0.7143	0.6935	0.8450	0.7246	0.7718	DecisionTree
Tumor_peri3 (T2FS)	7	0.8710	0.7143	0.9355	0.9193	0.9275	0.9185	SVC

Abbreviations: Ph2, T1-weighted fat-suppressed early post-contrast subtraction sequence; Ph6, T1-weighted fat-suppressed delayed post-contrast subtraction sequences; T2FS, T2-weighted fat-suppressed sequence, Peri1, peritumoral region, 1mm; Peri3, peritumoral region, 3mm Tumor_peri1, tumoral + 1mm peritumoral region; Tumor_peri3, tumoral +

3mm peritumoral region.

Table 5-1. Predictive performance of pCR in the validation cohort using the optimal machine learning models based on combination of tumoral and peritumoral regions across sequences

Type1	Type2	Rank	AUC	Sensitivity	Specitivity	Precision	Recall	F1-score	Best Model
Tumor_peri1 (Ph2)	Tumor (Ph6)	5	0.9378	1.0000	0.8387	0.9207	0.9130	0.8824	SVC
Tumor_peri1 (Ph2)	Tumor (T2FS)	9	0.9194	0.8571	0.9355	0.9335	0.8986	0.9098	SVC
Tumor_peri1 (Ph2)	Peri1 (Ph6)	6	0.9309	0.8571	0.9194	0.9329	0.9275	0.9088	AdaBoost
Tumor_peri1 (Ph2)	Peri1 (T2FS)	21	0.9032	0.8571	0.8548	0.9196	0.8406	0.8655	SVC
Tumor_peri1 (Ph2)	Peri3 (Ph6)	9	0.9194	0.8571	0.8871	0.8981	0.9130	0.8970	LGBM
Tumor_peri1 (Ph2)	Peri3 (T2FS)	1	0.9631	1.0000	0.8387	0.9234	0.9275	0.9250	KNeighbors
Tumor_peri1 (Ph2)	Tumor_peri1 (Ph6)	14	0.9101	0.8571	0.8710	0.8981	0.9130	0.8970	LGBM
Tumor_peri1 (Ph2)	Tumor_peri1 (T2FS)	12	0.9171	0.8571	0.9032	0.9137	0.8841	0.8950	SVC
Tumor_peri1 (Ph2)	Tumor_peri3 (Ph6)	6	0.9309	1.0000	0.8065	0.9234	0.9275	0.9250	LGBMr
Tumor_peri1 (Ph2)	Tumor_peri3 (T2FS)	2	0.9505	1.0000	0.7581	0.9325	0.9275	0.9297	KNeighbors
Tumor_peri1 (T2FS)	Tumor (Ph2)	25	0.8963	0.8571	0.8871	0.8987	0.8841	0.8904	LGBM
Tumor_peri1 (T2FS)	Tumor (Ph6)	51	0.8594	0.8571	0.7581	0.9093	0.8696	0.8840	AdaBoost
Tumor_peri1 (T2FS)	Peri1 (Ph2)	9	0.9194	1.0000	0.7097	0.9420	0.9420	0.9420	Logistic
Tumor_peri1 (T2FS)	Peri1 (Ph6)	28	0.8940	0.8571	0.8226	0.9188	0.8986	0.9062	LGBM
Tumor_peri1 (T2FS)	Peri3 (Ph2)	17	0.9078	0.8571	0.8871	0.9420	0.9420	0.9420	LGBM
Tumor_peri1 (T2FS)	Peri3 (Ph6)	55	0.8525	0.7143	0.9032	0.9021	0.8406	0.8624	Logistic
Tumor_peri1 (T2FS)	Tumor_peri1 (Ph6)	41	0.8779	0.7143	0.9355	0.8963	0.8116	0.8411	AdaBoost
Tumor_peri1 (T2FS)	Tumor_peri3 (Ph2)	14	0.9101	1.0000	0.7097	0.8981	0.9130	0.8970	RandomForest

Tumor_peri1 (T2FS)	Tumor_peri3 (Ph6)	58	0.8364	0.5714	1.0000	0.8816	0.8261	0.8478	Logistic
Peri1 (Ph2)	Tumor (Ph6)	31	0.8894	0.8571	0.7903	0.9033	0.9130	0.9065	LGBM
Peri1 (Ph2)	Tumor (T2FS)	19	0.9055	0.7143	0.9677	0.9188	0.8986	0.9062	Logistic
Peri1 (Ph2)	Peri1 (Ph6)	38	0.8802	0.7143	0.9839	0.9021	0.8406	0.8624	Logistic
Peri1 (Ph2)	Peri1 (T2FS)	32	0.8871	0.8571	0.9355	0.9130	0.9130	0.9130	KNeighbors
Peri1 (Ph2)	Peri3 (Ph6)	43	0.8733	0.7143	0.9194	0.8922	0.8986	0.8951	RandomForest
Peri1 (Ph2)	Peri3 (T2FS)	17	0.9078	0.8571	0.9516	0.9130	0.9130	0.9130	Logistic
Peri1 (Ph2)	Tumor_peri1 (Ph6)	48	0.8641	0.7143	0.9839	0.9455	0.9420	0.9313	LGBM
Peri1 (Ph2)	Tumor_peri3 (Ph6)	32	0.8871	0.8571	0.8871	0.9130	0.9130	0.9130	Logistic
Peri1 (Ph2)	Tumor_peri3 (T2FS)	30	0.8917	0.7143	0.9677	0.9420	0.9420	0.9420	RandomForest
Tumor_peri3 (Ph2)	Tumor (Ph6)	71	0.7926	1.0000	0.5968	0.8032	0.8551	0.8284	RandomForest
Tumor_peri3 (Ph2)	Tumor (T2FS)	43	0.8733	0.8571	0.8387	0.9051	0.8986	0.9015	RandomForest
Tumor_peri3 (Ph2)	Peri1 (Ph6)	3	0.9401	0.8571	0.9516	0.9193	0.9275	0.9185	LGBM
Tumor_peri3 (Ph2)	Peri1 (T2FS)	56	0.8433	0.7143	0.8871	0.9033	0.9130	0.9065	RandomForest
Tumor_peri3 (Ph2)	Peri3 (Ph6)	21	0.9032	1.0000	0.7742	0.8507	0.8841	0.8627	RandomForest
Tumor_peri3 (Ph2)	Peri3 (T2FS)	38	0.8802	1.0000	0.7742	0.8346	0.8551	0.8441	LGBM
Tumor_peri3 (Ph2)	Tumor_peri1 (Ph6)	19	0.9055	0.8571	0.9032	0.8802	0.8986	0.8859	LGBM
Tumor_peri3 (Ph2)	Tumor_peri3 (Ph6)	43	0.8733	0.8571	0.8387	0.8611	0.8696	0.8651	RandomForest
Tumor_peri3 (Ph2)	Tumor_peri3 (T2FS)	8	0.9274	1.0000	0.6935	0.9137	0.8841	0.8950	KNeighbors
Tumor_peri1 (Ph6)	Tumor (Ph2)	12	0.9171	0.8571	0.8387	0.8074	0.8986	0.8505	SVC
Tumor_peri1 (Ph6)	Tumor (T2FS)	36	0.8825	0.8571	0.8871	0.9207	0.9130	0.8824	SVC
Tumor_peri1 (Ph6)	Peri1 (T2FS)	48	0.8641	0.8571	0.8710	0.9196	0.8406	0.8655	SVC
Tumor_peri1 (Ph6)	Peri3 (Ph2)	23	0.8986	0.8571	0.8065	0.8890	0.8551	0.8688	Logistic

Tumor_peri1 (Ph6)	Peri3 (T2FS)	41	0.8779	0.8571	0.9032	0.9130	0.9130	0.9130	RandomForest
Tumor_peri1 (Ph6)	Tumor_peri3 (T2FS)	38	0.8802	0.8571	0.7581	0.9130	0.9130	0.9130	RandomForest
Peri3 (Ph2)	Tumor (Ph6)	35	0.8836	0.8571	0.8226	0.9171	0.8261	0.8547	DecisionTree
Peri3 (Ph2)	Tumor (T2FS)	59	0.8318	0.8571	0.7097	0.8915	0.7826	0.8199	RandomForest
Peri3 (Ph2)	Peri1 (Ph6)	23	0.8986	0.8571	0.9194	0.9335	0.8986	0.9098	RandomForest
Peri3 (Ph2)	Peri1 (T2FS)	51	0.8594	0.8571	0.7581	0.8785	0.8116	0.8374	RandomForest
Peri3 (Ph2)	Peri3 (Ph6)	47	0.8652	1.0000	0.7581	0.8915	0.7826	0.8199	AdaBoost
Peri3 (Ph2)	Peri3 (T2FS)	16	0.9090	1.0000	0.7742	0.9281	0.7536	0.8016	AdaBoost
Peri3 (Ph2)	Tumor_peri3 (Ph6)	62	0.8249	0.8571	0.7258	0.8642	0.8261	0.8425	RandomForest
Peri3 (Ph2)	Tumor_peri3 (T2FS)	46	0.8687	1.0000	0.7258	0.8915	0.7826	0.8199	RandomForest
Tumor_peri3 (T2FS)	Tumor (Ph2)	28	0.8940	0.7143	0.9355	0.9329	0.9275	0.9088	SVC
Tumor_peri3 (T2FS)	Tumor (Ph6)	73	0.7857	0.7143	0.8065	0.8890	0.8551	0.8688	RandomForest
Tumor_peri3 (T2FS)	Peri1 (Ph6)	34	0.8848	0.7143	0.9516	0.9188	0.8986	0.9062	RandomForest
Tumor_peri3 (T2FS)	Peri3 (Ph6)	69	0.7995	0.8571	0.7097	0.8872	0.7536	0.7985	RandomForest
Tumor_peri3 (T2FS)	Tumor_peri3 (Ph6)	74	0.7834	0.8571	0.6452	0.8609	0.8116	0.8324	RandomForest
Tumor (Ph2)	Tumor (Ph6)	67	0.8018	1.0000	0.6129	0.8261	0.8261	0.8261	RandomForest
Tumor (Ph2)	Tumor (T2FS)	54	0.8571	1.0000	0.6452	0.8798	0.6957	0.7549	Logistic
Tumor (Ph2)	Peri1 (Ph6)	3	0.9401	0.8571	0.8710	0.9137	0.8841	0.8950	AdaBoost
Tumor (Ph2)	Peri1 (T2FS)	36	0.8825	0.8571	0.8065	0.8934	0.8696	0.8795	Logistic
Tumor (Ph2)	Peri3 (Ph6)	61	0.8272	0.8571	0.8226	0.8074	0.8986	0.8505	SVC
Tumor (Ph2)	Peri3 (T2FS)	25	0.8963	1.0000	0.7258	0.8981	0.9130	0.8970	SVC
Tumor (Ph2)	Tumor_peri3 (Ph6)	57	0.8410	0.8571	0.7742	0.8851	0.8406	0.8582	RandomForest
Peri1 (Ph6)	Tumor (T2FS)	51	0.8594	0.7143	0.9516	0.9207	0.9130	0.8824	SVC

Peri1 (Ph6)	Peri1 (T2FS)	59	0.8318	0.8571	0.7419	0.8915	0.7826	0.8199	SVC
Peri1 (Ph6)	Peri3 (T2FS)	25	0.8963	0.7143	0.9839	0.9021	0.8406	0.8624	RandomForest
Peri3 (Ph6)	Tumor (T2FS)	66	0.8111	0.8571	0.7581	0.8933	0.6087	0.6847	SVC
Peri3 (Ph6)	Peri1 (T2FS)	50	0.8618	0.7143	0.9516	0.9234	0.9275	0.9250	RandomForest
Peri3 (Ph6)	Peri3 (T2FS)	70	0.7972	0.7143	0.8548	0.8582	0.5217	0.6100	Logistic
Peri3 (T2FS)	Tumor (Ph6)	67	0.8018	0.7143	0.8387	0.8915	0.7826	0.8199	RandomForest
Peri3 (T2FS)	Tumor_peri3 (Ph6)	75	0.7097	1.0000	0.4355	0.8148	0.7681	0.7900	RandomForest
Tumor (T2FS)	Tumor (Ph6)	63	0.8226	0.8571	0.7419	0.8872	0.7536	0.7985	RandomForest
Tumor (T2FS)	Tumor_peri3 (Ph6)	72	0.7880	0.8571	0.7581	0.8725	0.8551	0.8629	LGBM
Peri1 (T2FS)	Tumor (Ph6)	64	0.8203	0.7143	0.9194	0.8922	0.8986	0.8951	RandomForest
Peri1 (T2FS)	Tumor_peri3 (Ph6)	65	0.8157	0.8571	0.8548	0.9207	0.9130	0.8824	SVC

Abbreviations: Ph2, second-post contrast subtraction image; Ph6, sixth-post contrast subtraction image; T2FS, T2 weighted fat-saturated image, Peri1, peritumoral region, 1mm; Peri3, peritumoral region, 3mm Tumor_peri1, tumoral + 1mm peritumoral region; Tumor_peri3, tumoral + 3mm peritumoral region.

ABSTRACT (IN KOREAN)

국소진행성유방암 (ER+ HER2- subtype)의 치료 전 선행 화학 요법 후 완전 관해 예측

<지도교수 김민정>

연세대학교 대학원 의학과

박 지 우

목적: 이 연구의 목적은 현재까지 선행화학요법에 저조한 성적을 보인다고 알려져 있는 ER+ HER2- 국소 진행성 유방암 환자 군의 더 효과적인 치료 전략 수립을 돕기 위해, 치료 시작 전 선행화학요법 후 병리학적 완전 관해를 달성할 수 있는 최적의 환자들을 예측해보는 것이다.

방법: 본 연구는 후향적으로 2010년부터 2020년까지 본원에서 ER+ HER2- 국소 진행성 유방암으로 진단된 818명의 환자중에서, 수술 전 자기공명영상을 시행 받고 선행화학요법을 모두 완료하였으며 수술을 통하여 최종 병리학적 결과가 확인된 환자 265명을 대상으로 하고 있고, 이 환자들을 치료 전 자기공명영상 시행일 2016년 1월을 기준으로 학습 코호트와 검증 코호트로 분류하였다. 본 연구에서는 치료 전 자기공명영상에서 종양과 종양 주변의 VOI를 그렸으며, 텍스처 특성 추출은 3종류의 자기공명영상 시퀀스; 초기 조영 증강 영상, 후기 조영증강 영상, 지방 포화 T2 강조 영상; 에서 이루어졌다. 일곱개의 머신 러닝 모델이 각각의 시퀀스에서 추출한 종양, 종양 주변, 종양과 종양 주변의 텍스처 특성으로 만들어졌고, 동일한 모델들은 서로 다른 두개의 시퀀스 각각에서 종양과 종양주변 텍스처 특성의 결합을 통하여서도 만들어졌다. 환자의 나이, 종양의 크기, ER and PR receptor expression rate을 포함하는 임상적인 요소를 포함한 모델도 동일한 방법으로 만들었다. 모든 모델의 pCR 예측 능력은 ROC 곡선의 AUC로 평가 및 비교되었다.

결과: 총 7533개의 텍스처 특성이 치료 전 자기공명영상에서 3종류의 시퀀스의 VOIs; 종양, 종양 주변, 종양과 종양 주변 영역; 으로부터 얻어졌다. 단독 시퀀스 모델로서는 초기 조영 증강 영상에서 종양부터 종양 주변부 1mm를 포함한 영역의 SVM 모델이 가장 우수한 성능을 보였다 (AUC =

0.9447). 그리고 앞서 언급한데로 서로 다른 시퀀스에서 종양, 종양 주변, 종양과 종양 주변 영역의 텍스처 특성을 결합해 보았을 때는, 초기 조영 증강 영상에서 종양부터 종양 주변 1mm까지와 지방 포화 T2 강조 영상의 종양 주변 3mm를 결합하였을 때 AUC 0.9631의 가장 우수한 모델이 개발되었다.

결론: 본 연구는 치료 전 자기공명영상에서 ER+ HER2- 국소 진행성 유방암 환자의 선행화학요법 반응을 평가해보고자 다양한 시퀀스에서 종양과 종양 주변 영역의 텍스처 특성을 활용하였다. 본 연구는 서로 다른 시퀀스에서의 종양과 종양 주변의 텍스처 특성의 결합 머신 러닝 모델로 이 환자군의 병리학적 완전 관해를 보다 정확하게 예측하고 분류할 수 있다는 것을 제안하며, 나아가 이 결과가 임상적인 치료 전략 수립에 직접적으로 도움을 줄 수 있을 것으로 기대한다.

핵심 단어: ER+ HER2- 국소 진행성 유방암, 선행화학요법, 병리학적 완전 관해, 종양의 텍스처 특성, 종양 주변의 텍스처 특성, 치료 전 자기공명영상, 머신 러닝

PUBLICATION LIST

Jiwoo P, Min Jung K, Jong-Hyun Y, Kyunghwa H, Eun-Kyung K, Joo Hyuk Sohn, et al. Machine Learning Predicts Pathologic Complete Response to Neoadjuvant Chemotherapy for ER+HER2- Breast Cancer: Integrating Tumoral and Peritumoral MRI Radiomic Features. Diagnostics, 2023; 13, 3031.