Enhancing pathological complete response prediction in breast cancer: the role of dynamic characterization of DCE-MRI and its association with tumor heterogeneity

Xinyu Zhang[1], Xinzhi Teng[1], Jiang Zhang[1], Qingpei Lai[1] and Jing Cai[1,2*]

Abstract

Background Early prediction of pathological complete response (pCR) is important for deciding appropriate treatment strategies for patients. In this study, we aimed to quantify the dynamic characteristics of dynamic contrast-enhanced magnetic resonance images (DCE-MRI) and investigate its value to improve pCR prediction as well as its association with tumor heterogeneity in breast cancer patients.

Methods The DCE-MRI, clinicopathologic record, and full transcriptomic data of 785 breast cancer patients receiving neoadjuvant chemotherapy were retrospectively included from a public dataset. Dynamic features of DCE-MRI were computed from extracted phase-varying radiomic feature series using 22 CAnonical Time-sereis CHaracteristics. Dynamic model and radiomic model were developed by logistic regression using dynamic features and traditional radiomic features respectively. Various combined models with clinical factors were also developed to find the optimal combination and the significance of each components was evaluated. All the models were evaluated in independent test set in terms of area under receiver operating characteristic curve (AUC). To explore the potential underlying biological mechanisms, radiogenomic analysis was implemented on patient subgroups stratified by dynamic model to identify differentially expressed genes (DEGs) and enriched pathways.

Results A 10-feature dynamic model and a 4-feature radiomic model were developed (AUC = 0.688, 95%CI: 0.635–0.741 and AUC = 0.650, 95%CI: 0.595–0.705) and tested (AUC = 0.686, 95%CI: 0.594–0.778 and AUC = 0.626, 95%CI: 0.529–0.722), with the dynamic model showing slightly higher AUC (train *p* = 0.181, test *p* = 0.222). The combined model of clinical, radiomic, and dynamic achieved the highest AUC in pCR prediction (train: 0.769, 95%CI: 0.722–0.816 and test: 0.762, 95%CI: 0.679–0.845). Compared with clinical-radiomic combined model (train AUC = 0.716, 95%CI: 0.665–0.767 and test AUC = 0.695, 95%CI: 0.656–0.714), adding the dynamic component brought significant improvement in model performance (train *p* < 0.001 and test *p* = 0.005). Radiogenomic analysis identified 297 DEGs, including CXCL9, CCL18, and HLA-DPB1 which are known to be associated with breast cancer prognosis or angiogenesis. Gene set enrichment analysis further revealed enrichment of gene ontology terms and pathways related to immune system.

*Correspondence:
Jing Cai
jing.cai@polyu.edu.hk

Conclusion Dynamic characteristics of DCE-MRI were quantified and used to develop dynamic model for improving pCR prediction in breast cancer patients. The dynamic model was associated with tumor heterogeniety in prognostic-related gene expression and immune-related pathways.

Keywords Breast cancer, Radiomics, DCE-MRI, Treatment response prediction

Introduction

Breast cancer is one of the most common malignant in women. In 2020, there were around 2.3 million women newly diagnosed with and over 600,000 women died of breast cancer worldwide [1]. Recently, neoadjuvant chemotherapy (NAC) has become increasingly used in breast cancer systemic treatment. NAC was initially used in inoperable breast cancer to enable surgical resection, and expanded to other types of breast cancer for increasing the chance of breast conservation owing to its remarkable efficacy [2]. Current NAC treatment schemes are determined by hormone receptor (HR) status and human epidermal growth factor receptor 2 (HER2) status as recommended by American Society of Clinical Oncology (ASCO) [3]. Pathological complete response (pCR), defined as no residual disease in breast and axillary region after NAC, is a validated prognostic factor to assess treatment response and associated with long-term outcome [4]. However, only 10-50% patients achieved pCR, varying according to their receptor subtypes [5]. Moreover, the assessment of pCR status is performed at surgery after completion of NAC, prior to which non-responders have suffered from the toxicity and side effects caused by NAC. Therefore, it is essential to identify patients who are likely to achieve pCR before NAC to avoid unnecessary complications and maximize potential benefits.

Dynamic contrast-enhanced MRI (DCE-MRI) is the clinical routine for breast cancer assessment. It has high sensitivity in diagnosis and treatment monitoring [6, 7]. Through acquisition of sequential images before, during, and after the administration of contrast agent, DCE-MRI provides valuable information about tissue perfusion and contrast agent enhancement dynamics associated with tumor angiogenesis [8]. Radiomics extracts high-dimensional image features that are imperceptible to human eyes to non-invasively quantify tumor characteristics [9]. Radiomic analysis of breast DCE-MRI has been used for pCR prediction in many previous studies, most of which only used radiomic features from one or several phases while ignoring the dynamic information [10, 11]. Recently, attempts have been made to leverage the dynamic information embedded in DCE-MRI for pCR prediction by combining radiomic features extracted from different DCE-MRI phases. For instance, Peng et al. calculated delta-features between two different phases for pCR prediction [12]; Li et al. employed simple statistical patterns of radiomic features extracted from

different phases for pCR prediction and achieved better performance compared to single-phase features, demonstrating the value of multi-phase information [13]. In BMMR2 challenge, radiomic features from kinetic maps, such as peak enhancement maps and signal enhancement ratio maps, were used to predict pCR [14]. However, the entire time series of radiomic features has not been fully explored and may contain additional information for tumor characterization. On the other hand, feature-based representation of time series data like 22 CAnonical Time-series CHaracteristics (Catch22) can capture the dynamic properties of time series data and was used in various tasks [15, 16]. Accordingly, there developed an assumption that the dynamics of radiomic feature series extracted by Catch22 can characterize the dynamic information in DCE-MRI and improve pCR prediction of breast cancer patients.

In this study, we aimed to systematically extract dynamic properties of radiomic feature series from DCE-MRI to improve treatment response prediction of breast cancer patients. To achieve this, a large number of dynamic features were extracted by Catch22 from DCE-MRI feature series, and a dynamic model was then built for pCR prediction. Various combinations of dynamic models and existing radiomic and clinical models were developed to find the optimal one as the final model. In addition, radiogenomic analysis of binarized dynamic model predictions was conducted to explore its association with tumor heterogeneity and biological process. Figure 1 shows the overall workflow of this study.

Materials and methods

Patient data

A total of 985 stage II/III locally advanced breast cancer patients enrolled in the multi-center I-SPY2 trial (clinical trial number: NCT01042379) during 2010 to 2016 were collected from the publicly available dataset on The Cancer Image Archive [17–19]. Institutional review board approval was waived due to the use of public data. The detailed descriptions of I-SPY2 trial have been reported by previous paper [20]. All the patients underwent MR examination and percutaneous biopsy before receiving NAC. After the completion of NAC, patients underwent surgical resection to assess residual disease. The exclusion criteria included: (1) incomplete image or clinicopathologic data; (2) deviations from the prescribed scanning protocol; (3) insufficient image quality.
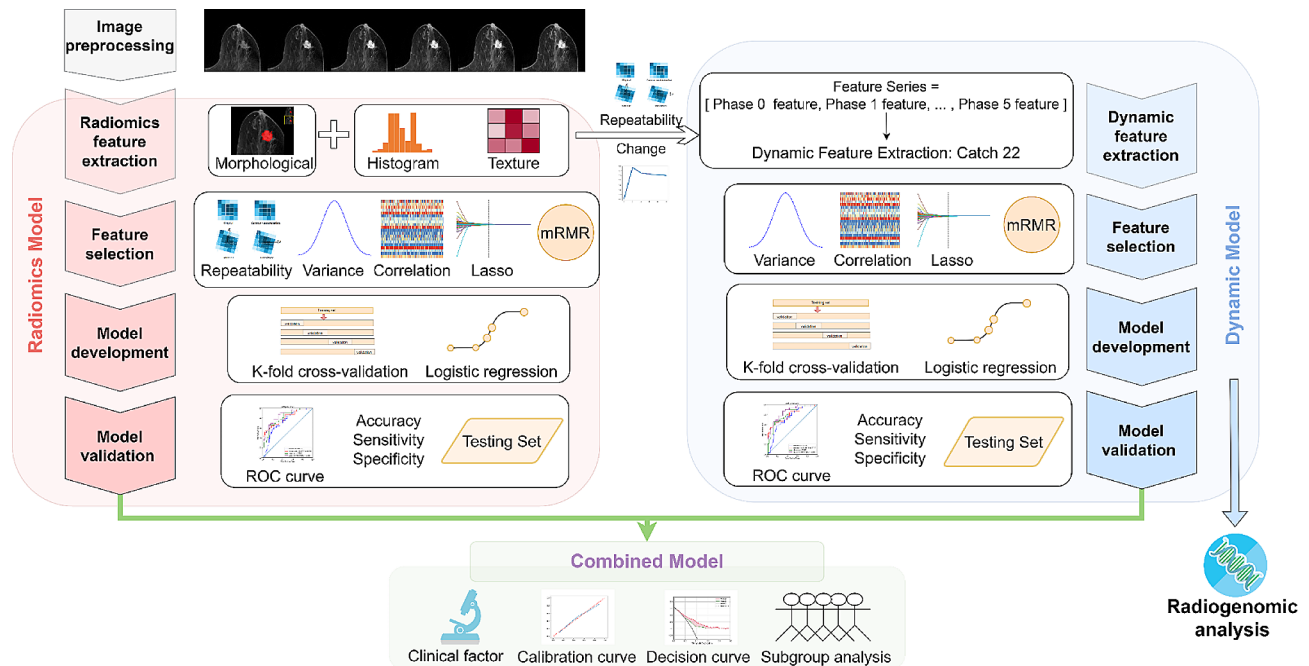
Fig. 1 Workflow of the study. Firstly, the collected DCE-MR images were preprocessed by normalization and discretization. Radiomic features were extracted from multiple phases of DCE-MRI, while dynamic features were extracted from radiomic feature series. Feature selection, model development, and model validation were then conducted separately for radiomic model and dynamic model. Subsequently, combined models were developed by integrating radiomic, dynamic, and clinical information and their performance were evaluated. In addition, radiogenomic analysis was performed on dynamic model to investigate potential biological mechanisms

Clinicopathologic data

Clinicopathologic data including HR, HER2, MammaPrint status (MP), pCR, and other patient characteristics was provided by the dataset. HR and HER2 were determnined by immunohistochemical (IHC) staining or fluorescence in-situ hybridization (FISH) of tissues obtained during pre-treatment biopsy. HR was determined as positive when ≥5% tumor staining for ER and/or PgR was seen. HER2 was determined as positive by IHC 3+or FISH overexpression [21]. The surrogate of treatment response pCR was defined as no residual disease in breast and axillary lymph nodes after NAC and obtained by post-treatment surgery [22].

Imaging data and tumor segmentation

The scanning process of DCE-MRI can be found on TCIA website [17, 18]. DCE-MRI scanning protocol details are provided in Supplementary Material Table S1. The pre-contrast phase and five post-contrast phases were used for radiomic feature extraction and subsequent analysis.

The region of interest was segmented by functional tumor volume (FTV) included in the dataset. The calculation of FTV involved background filtering, estimating signal enhancement ratio, and applying a peak enhancement threshold in a manual-defined 3D bounding box [23].

Image preprocessing and feature extraction

Image normalization was performed across different DCE-MRI phases of the same patient to preserve the dynamic information. All the images were isotropically resampled to $1*1*1mm^3$, and discretized by a fixed bin width of 5. More details of image preprocessing can be found in Supplementary Material Figure S1. Radiomic features were extracted from each phase of DCE-MRI using PyRadiomics package version 3.0.1 following the standardization and definitions in the image biomarker standadization initiative [24, 25]. The extracted features included morphological features ($n=14$), first-order features ($n=17$), and texture features ($n=79$). The repeatability of radiomic features was evaluated by perturbation which involved random translation, rotation, and contour randomization of original masks [26–28]. Features with high-repeatabliity (intraclass correlation coefficient, ICC≥0.9 [29]) were retained for better model repeatability. The fluctuation of radiomic features was measured by performing a single-sample t test on the variations between features from different phases. Radiomic feature series were constructed by concatenating the high-repeatable and phase-varying first-order and texture features. Dynamic features were extracted from radiomic feature series using the 22 CAnonical Time-series Characteristics (catch22) feature set, specifically designed for capturing the dynamic properties of time series data,

such as distributions and outliers, linear and non-linear autocorrelation, and so on [30].

Model development and evaluation

The development of radiomic model and dynamic model followed the same process including feature selection and model building. The features with low variances were removed first to retain those providing more information. Then, features with significant correlations with pCR were identified by MannWhitney U test, where a p value smaller than 0.05 was defined as significant. LASSO was subsequently used to select the independently discriminative features. Finally, features were ranked by minimum redundancy and maximum relevance (mRMR) algorithm considering the relevance to pCR and redundancy at the same time [31]. Clinical factors that are commonly used in clinical decision making and have significant associations with pCR were identified and used in developing clinical model. Combined models were constructed using logistic regression with clinical factors, prediction score of the dynamic model, and prediction score of radiomic model as variables. Different combination strategies were adopted, including combining two of the three variables respectively, as well as the combination of all three together. The independence of the components in the combined models were examined by their coefficients and p values. All the models were developed using Logistic Regression with 10-fold cross-validation in training set and tested in independent testing set.

The pCR prediction performance of the candidate models was assessed by various metrics, including area under receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity. AUCs were calculated by continuous prediction (the probability) and the other metrics were calculated by binary prediction (pCR or non-pCR) dichotomized by Youden index. An AUC typically ranges from 0 to 1 while AUC equal to one means a perfect descrimination ability. The optimal model was determined by the highest internal validation AUC in the training set. Heatmap was employed to visualize the relationships between different models and their association with clinical factors. SHapley Additive Explanations (SHAP), a method to interpret and explain the output of machine learning models, was employed to evaluate the importance of each component in the model with the highest AUC [32]. In our case, where the model output is the probability of achieving pCR, the SHAP values for each parameter ranges from -1 to 1 and a larger absolute value means a higher importance for model output. Calibration curves and Brier scores were used to further evaluate the alignment between model-predicted probabilities and actual probabilities. Brier score measures the accuracy of probabilistic predictions and takes the value from 0 to 1, for which 0 means a perfect prediction.

Decision curve analysis was performed to evaluate the clinical benefit obtained by the optimal model [33]. Besides, to further demonstrate the generalizability of the optimal model, its association with pCR in various pre-defined molecular subtypes, namely HR+HER2-, HR+HER2+, HR-HER2-, and HR-HER2+, and patients receiving different treatments were evaluated.

Radiogenomic analysis

To examine whether the dynamic model can reflect tumor heterogeneity and its association with gene expression, we collected paired total mRNA expression data from National Center for Biotechnology Information (NCBI) [34]. Patients were divided into DYN+ and DYN- groups according to the binary prediction of dynamic model. Student t test was performed to identify differentially expressed genes (DEGs) between the two groups. An absolute log-2 fold change larger than 0.25 and a p value smaller than 0.05 were used as cutoff. Enriched Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified by gene set enrichment analysis (GSEA) of DEGs [35–39]. A p value smaller than 0.05 and false discovery rate (FDR) smaller than 0.25 were considered statistically significant.

Statistical analysis and software

For statistical analysis, Chi-Square test or Fisher's exact test was used for categorical variables and MannWhitney U test was used for continuous variables. A two-tailed p-value smaller than 0.05 was considered statistically significant. The 95% CIs for AUCs were calculated according to DeLong's methods [40]. DeLong test was used to compare the AUCs of two independent models and likelihood ratio test was used to compare the model fit of nested models to demonstrate the improvements conferred by the additional factors in complex models. The statistical analysis was carried out on R4.2.2 [41] and Python3.7 [42]. Logistic regression was carried out by package scikit-learn 1.0.2 [43]. Radiogenomic analysis was conducted using packages scanpy 1.9.3 [44] and gseapy 1.1.0 [45].

Results

Patient characteristics

A total of 785 patients with complete imaging data and clinicopathological record constituted the entire patient cohort in this study and were divided into training set and testing set with a ratio of 3:1 (Fig. 2). As shown in Table 1, there is no significant difference in all of the patient characteristics between training set and testing set. The characteristics of pCR and non-pCR patients were tabulated in Supplementary Material Table S2. Significant association with pCR was observed in treatment,
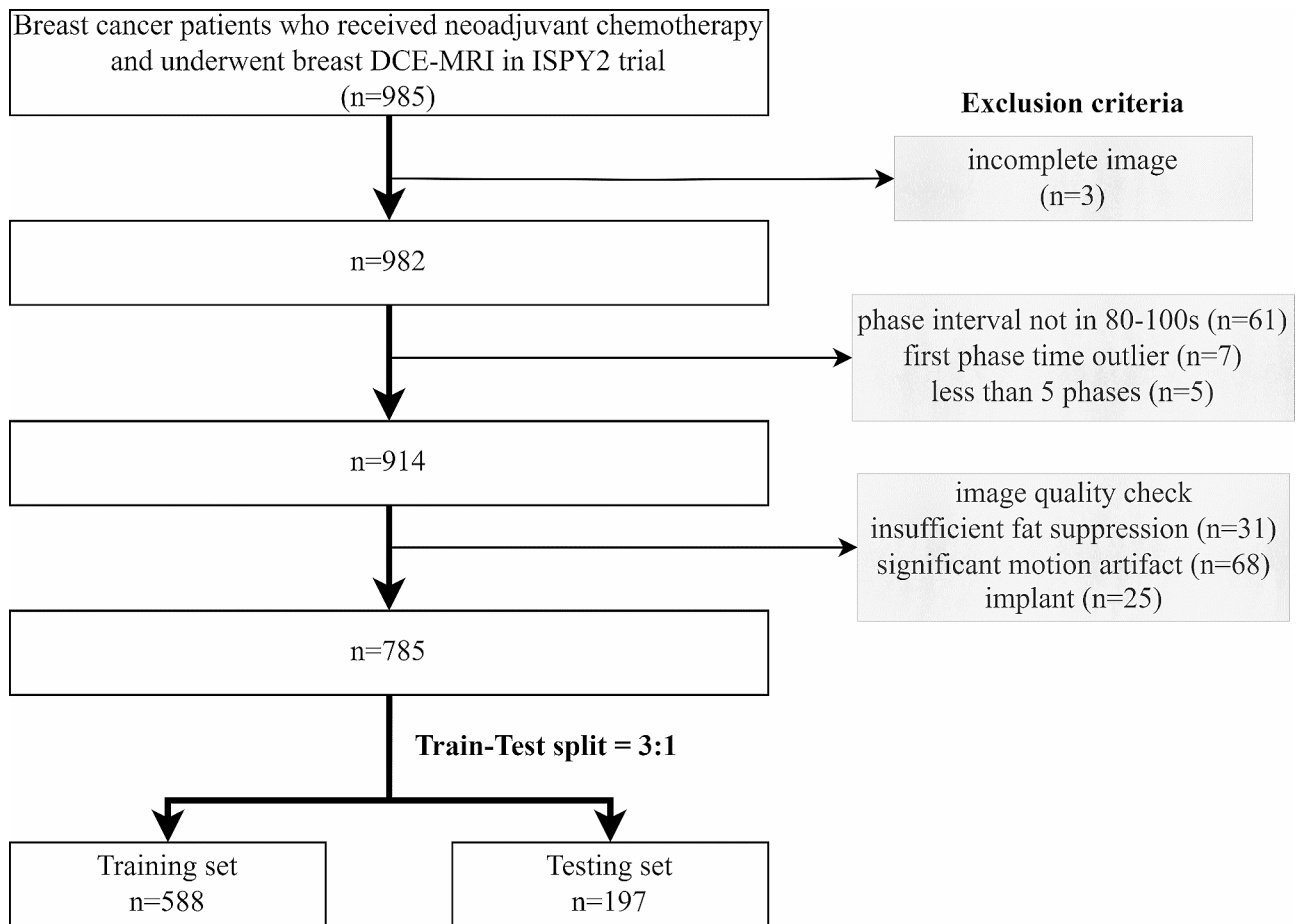
Breast cancer patients who received neoadjuvant chemotherapy
and underwent breast DCE-MRI in ISPY2 trial
(n=985)

Exclusion criteria

incomplete image
(n=3)

n=982

phase interval not in 80-100s (n=61)
first phase time outlier (n=7)
less than 5 phases (n=5)

n=914

image quality check
insufficient fat suppression (n=31)
significant motion artifact (n=68)
implant (n=25)

n=785

Train-Test split = 3:1

Training set
n=588

Testing set
n=197

Fig. 2 Patient cohort and train-test split

HR, HER2, and MP, while the other characteristics were independent of pCR.

Feature repeatability and feature change

The summarized results of feature repeatability and feature variation are shown in Supplementary Material Figure S2 and Figure S3. After removing the duplicate shape features and features with low repeatability, there were 480 radiomic features retained for each patient. For dynamic feature extraction, features repeatable in all the DCE-MRI phases and changing across different phases were retained. A total of 1232 dynamic features were extracted from 56 selected radiomic feature series for each patient and used in further analysis. An example of the dynamic feature is shown in Supplementary Material Table S3.

Different models in pCR prediction

A 10-feature dynamic model and a 4-feature radiomic model were developed separately (Supplementary Material Table S4). Dynamic model achieved higher AUC than radiomic model in both training set (0.688 vs. 0.650) and testing set (0.686 vs. 0.626) (Fig. 3(a)(b)), but the

difference was not statistically significant (p value=0.181 and 0.222). Dynamic model also had better performance in terms of accuracy and specificity, while the sensitivity was the same as radiomic model (Table 2). The significance of each feature in dynamic model and radiomic model was evaluated by the odds ratio and tabulated in Supplementary Material Table S5 and S6.

Among the clinicopathological variables provided in the dataset, treatment, HR, HER2, and MP were significantly associated with pCR (Supplementary Material Table S2). Since we intended to study biomarkers and MP requires expensive genomic test, only HR and HER2 were retained for further analysis. Table 3 summarized the metrics for pCR prediction performance of clinical model and combined models. Clinical-radiomic-dynamic (CRD) model achieved the highest training and testing AUC (Fig. 3(c)(d)), accuracy, and specificity, while clinical model had the highest sensitivity among all the models. The clinical factors, the dynamic model, and the radiomic model demonstrated independent value in CRD model as indicated by their coefficients and p values (Supplementary Material Table S7). Compared with clinical-radiomic (CR) model, CRD model shown significant improvement

Table 1 Patient characteristics in training set and testing set

		Train	Test	*P*-value
Treatment	Paclitaxel	95	47	0.448
	Paclitaxel + ABT 888 + Carboplatin	40	17	
	Paclitaxel + AMG 386	73	23	
	Paclitaxel + AMG 386 + Trastuzumab	12	4	
	Paclitaxel + Ganetespib	59	19	
	Paclitaxel + Ganitumab	69	20	
	Paclitaxel + MK-2206	32	14	
	Paclitaxel + MK-2206 + Trastuzumab	25	5	
	Paclitaxel + Neratinib	70	15	
	Paclitaxel + Pembrolizumab	44	10	
	Paclitaxel + Pertuzumab + Trastuzumab	23	10	
	Paclitaxel + Trastuzumab	16	4	
	T-DM1 + Pertuzumab	30	9	
HR	negative	265	89	1
	positive	323	108	
HER2	negative	454	157	0.322
	positive	144	40	
MP	negative	304	102	1
	positive	284	95	
pCR	negative	398	133	1
	positive	190	64	
Age (mean, range)		49 (25–77)	50 (23–72)	0.238
Race	American Indian or Alaska Native	4	0	0.434
	American Indian or Alaska Native; White	0	1	
	Asian	43	14	
	Asian; White	3	1	
	Black or African American	71	22	
	Native Hawaiian or Pacific Islander	4	0	
	Native Hawaiian or Pacific Islander; White	0	1	
	White	461	157	
	N/A	2	1	
Menopausal status	N/A	103	43	0.362
	Perimenopausal	21	7	
	Postmenopausal	177	64	
	Premenopausal	287	83	
Ethnicity	Hispanic or Latino	76	28	0.723
	Not Hispanic or Latino	511	169	
	N/A	1	0	

in both training and testing performance, indicating the additive value of dynamic model. Figure 3(e) shows the heatmap of the predicted probability by different models. Models containing dynamic features were clustered as similar models, showing the distinctive characteristic of dynamic features. The dynamic model was not associated with HR and HER2, demonstrating the independent value of dynamic model.

Evaluation of the optimal model
Overall, CRD model has the best performance in pCR prediction. The calibration curves of CRD model had Brier score of 0.174 and 0.180 in training and testing set respectively (Fig. 4(a)), indicating well-alignment between predicted probablities and actual probabilities. The decision curve analysis of CRD model demonstrated its clinical usefulness by higher net benefit gain compared to clinical model and the other combined models (Fig. 4(b)).

The SHAP value analysis of CRD model shown high importance of dynamic model, which was comparable to HR. The importance of radiomic model and HER2 was a little bit lower, but still had significant effect on the model output (Fig. 5(a)). The CRD model was also used in stratifying pCR and non-pCR patients under different pre-defined molecular subtypes. It shown a significant stratification ability in all the four molecular subtypes with odds ratio (OR) of 2.88–8.42. (Fig. 5(b)). In the analysis of patients receiving different drugs, except for the marginally significant performance in Pertuzumab arm, CRD model shown significant association with pCR with OR of 2.88–10.93 in the other treatment arms (Fig. 5(c)).

DEGs and enriched pathways
In DEG analysis, a total of 196 up-regulated genes and 101 down-regulated genes in DYN+group were identified. As compared with HR- group and HER2+group, which also associate with better pCR outcome in ISPY2 trial, there are 7 common up-regulated genes and 22 common down-regulated genes (Fig. 6(a)). In GSEA by GO terms, there are 36 biological processes, 3 cellular components, 2 molecular functions enriched in DYN+ (Fig. 6(b)), many of which are associated with immune system. There are 4 enriched pathways in GSEA by KEGG, in which 3 pathways are related to viral disease and 1 pathway is related to immune disease (Fig. 6(c)).

Discussion
While the dynamic information in DCE-MRI has shown potential in various clinical applications, the exploration of DCE-MRI-derived radiomic feature series has remained limited. This study systematically extracted dynamic features from DCE-MRI-derived radiomic feature series using feature-based time series analysis method and built dynamic model for pCR prediction. Adding dynamic model to exisitng clinical and radiomic model can improve pCR prediction. Radiogenomic analysis revealed correlations of dynamic model with some breast cancer prognosis-related genes and pathways,

Table 2 Results of radiomic model and dynamic model for pCR prediction

	Training AUC	*p* value	Testing AUC	*p* value	Accuracy	Sensitivity	Specificity
Dynamic model[#]	0.688 (0.635–0.741)		0.686 (0.594–0.778)		0.650	0.609	0.669
Radiomic model	0.650 (0.595–0.705)	0.181	0.626 (0.529–0.722)	0.222	0.589	0.609	0.579

P values were obtained by DeLong test

[#] indicate reference model for comparison

Accuracy, sensitivity, and specificity were obtained in testing set

Table 3 Results of clinical model and combined models for pCR prediction

	Training AUC	*p* value	Testing AUC	*p* value	Accuracy	Sensitivity	Specificity
CRD model[#]	0.769 (0.722–0.816)		0.762 (0.679–0.845)		0.736	0.578	0.812
CD model	0.754 (0.705–0.802)	<0.001[*]	0.755 (0.672–0.839)	0.112	0.660	0.734	0.624
CR model	0.716 (0.665–0.767)	<0.001[*]	0.695 (0.656–0.714)	0.005[*]	0.695	0.656	0.714
RD model	0.709 (0.658–0.761)	<0.001[*]	0.693 (0.602–0.784)	<0.001[*]	0.629	0.594	0.647
Clinical model	0.642 (0.586–0.697)	<0.001[*]	0.691 (0.600-0.782)	<0.001[*]	0.619	0.828	0.519

P values were obtained by likelihood ratio test

[*] indicate statistically significant [#] indicate reference model for comparison

Accuracy, sensitivity, and specificity were obtained in testing set

CRD model: Clinical-Radiomic-Dynamic model; CD model: Clinical-Dynamic model; CR model: Clinical-Radiomic model; RD model: Radiomic-Dynamic model
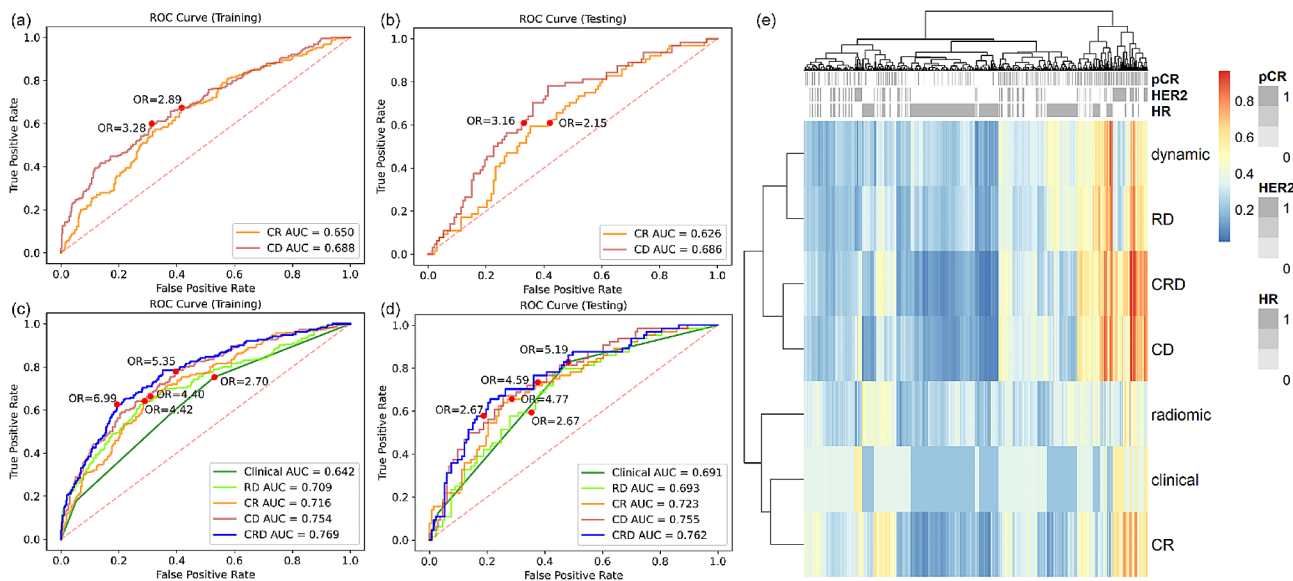

Fig. 3 Receiver operating characteristic (ROC) curve analysis of dynamic model and radiomic model in (**a**) training set and (**b**) testing set. ROC analysis of clinical model, radiomic-dynamic (RD) model, clinical-radiomic (CR) model, clinical-dynamic (CD) model, and clinical-radiomic-dynamic (CRD) model in (**c**) training set and (**d**) testing set. (**e**) Heatmap of predicted probability by different models

providing the potential biological explanations for the additive value.

The change in DCE-MR image appearances caused by the flow of contrast agent may contain valuable information for pCR prediction. Previous studies have employed delta features and statistical distributions to characterize the relevant dynamic information [12, 13]. However, the former method may provide limited information by utilizing only two of multiple DCE-MR phases, while the latter method disregards the temporal information that is crucial for reflecting the directional flow of contrast agent. A recently published paper implemented several classical time series analysis algorithms in DCE-MRI-derived radiomic feature series and achieved an accuracy
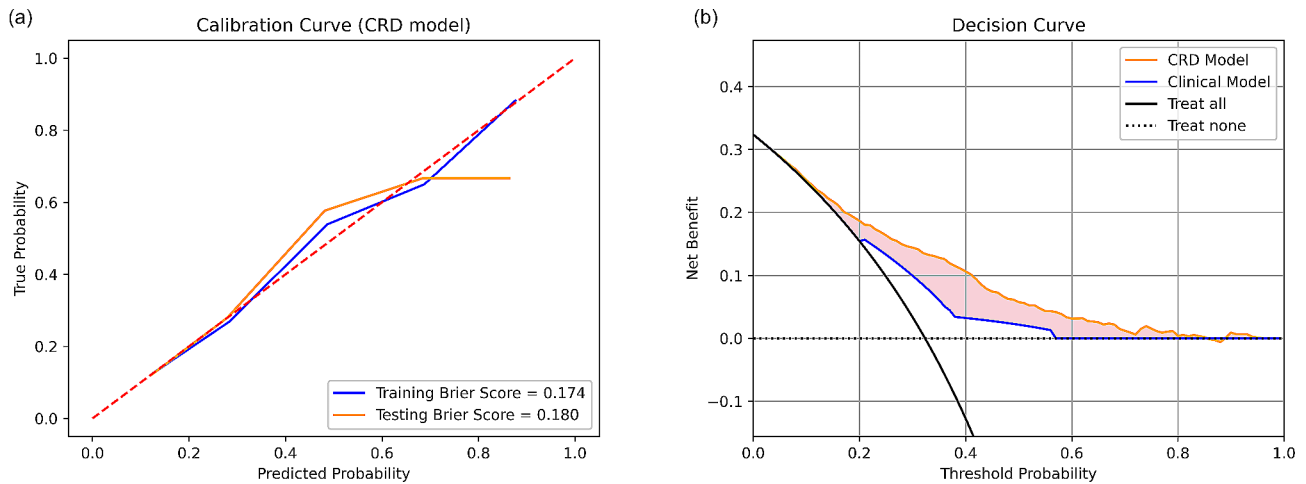
Fig. 4 Calibration curve and decision curve analysis

of 0.852 in breast cancer diagnosis, demonstrating the significance of serial information as well as the feasibility and efficacy of time series analysis [46]. In our study, we used radiomic features to comprehensively describe DCE-MR image appearance and adopted Catch22 to systematically analyze the dynamics of radiomic feature series. The Catch22 feature set takes into account both the temporal order and relative magnitude of series values. It has been successfully implemented in many time series analysis applications, such as breath signal and heart rate. To the best of our knowledge, it is the first study to apply a systematic feature-based time series analysis method to DCE-MRI for pCR prediction. Our results demonstrated the utility of the extracted dynamic features by showing a modestly higher AUC of dynamic model in comparison to the conventional radiomic model. Furthermore, the dynamic features provided additive value to the existing methods, as evidenced by a significantly improved model performance compared with both clinical model and CR model. Overall, we have demonstrated the feasibility and efficacy of extracting dynamic information through feature-based time series analysis and the potential of dynamic features in facilitating pCR prediction. Besides, our method offers the advantage of interpretability as Catch22 provides clear definition for each dynamic feature. And it is adaptable to different time series length which is frequently encountered in real-world clinical practice due to the variations of machines and scan settings. Our method demonstrates the potential to be implemented in real clinical practice, although further validation is required to confirm its performance in diverse settings.

Both single-modal and multi-modal models were developed in this study. While the imaging-based model and clinical model appeared to have similar performance, the combined models shown better performance than individual models. The CRD model achieved the highest

AUC, which is significantly better than RD model and clinical model alone, indicating that imaging features and clinical factors may provide distinct and complementary information for pCR prediction. Subgroup analysis of the CRD model was conducted to further explore the effectiveness of CRD model under various conditions. Breast cancer is a highly heterogeneous disease characterized by various HR and HER2 status, resulting in four molecular subtypes. Our results on molecular subtype analysis resulted in varying effect size by OR ranging from 2.88 to 8.42, where a larger OR indicates a stronger predictive ability. While CRD model is significantly associated with pCR in all the molecular subtypes, our results suggested that CRD model has stronger predictive ability for patients of HR+HER2-. The CRD model was also evaluated by its effect for patients receiving different drugs, resulting in the largest OR in Ganetespib and marginally significant OR in Pertuzumab. This indicates the various predictive value of CRD model for various treatment drugs and assists the clinicians to decide applicable scenarios. In general, CRD model shown generalizability acorss various molecular subtypes and various treatment drugs. However, due to the nature of trial data, the patient numbers are small in each subgroups and further validation on larger cohort is required to confirm the results.

It is believed that radiomics is able to detect the underlying biological processes in the human body by analyzing image textures that are imperceptible to human eyes. Moreover, pre-treatment radiomics mostly reflect the baseline tumor characteristics, which is the result of various biological processes and associated with treatment response. Previous studies indicated the representativeness of image phenotypes for the biological characteristics by demonstrating their similar predictive ability to pCR [47]. However, few radiomics study has linked the image phenotype to biological processes through
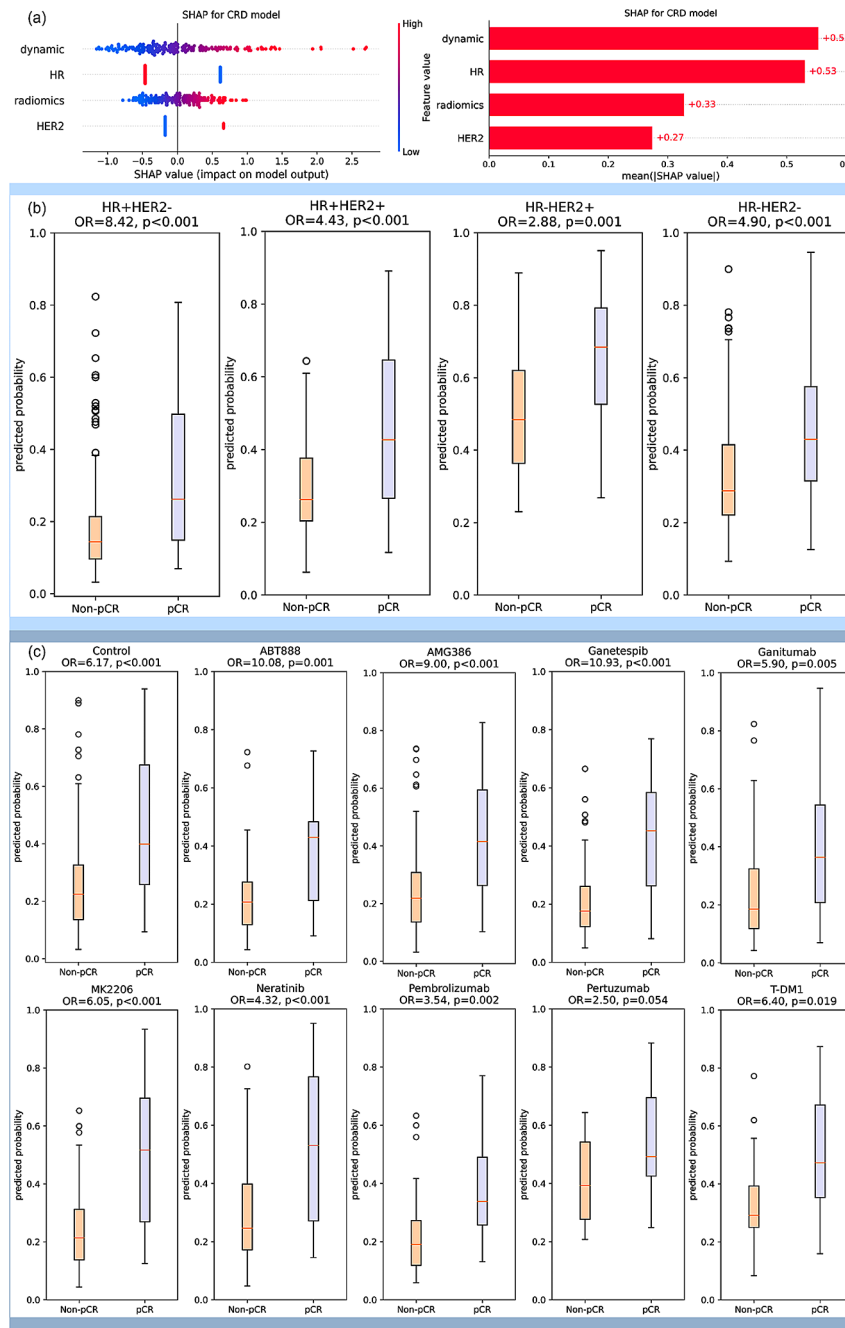
Fig. 5 (**a**) SHAP analysis for interpretable component importance of CRD model. The beaswarm plot shows how each variable influence model output on single data where one dot represents one patient (left). The mean absolute SHAP value reflects the global effect of each variable on model output (right). (**b**) Box plots showing the predictive ability of CRD model in patient subgroups of various molecular subtypes. The molecular subtypes were defined by the status of HR and HER2, namely HR + HER2-, HR + HER2+, HR-HER2+, and HR-HER2-. The box plots indicate that the CRD model yields significantly distinct prediction probabilities for patients with pCR and non-pCR in all the four molecular subtypes. (**c**) Box plots showing the predictive ability of CRD model in patients receiving various treatments. Patients received standard care (control) or standard care plus one trial agent (ABT888, AMG386, Ganetespib, Ganitumab, MK2206, Neratinib, Pembrolizumab, Pertuzumab, T-DM1) in the trial. The box plots suggest that CRD model demonstrates the capability to differentiate pCR and non-pCR patients across various treatment drugs in this trial, with the exception of a marginal significance observed in Pertuzumab. All the p values were obtained by student t test. OR: odds ratio
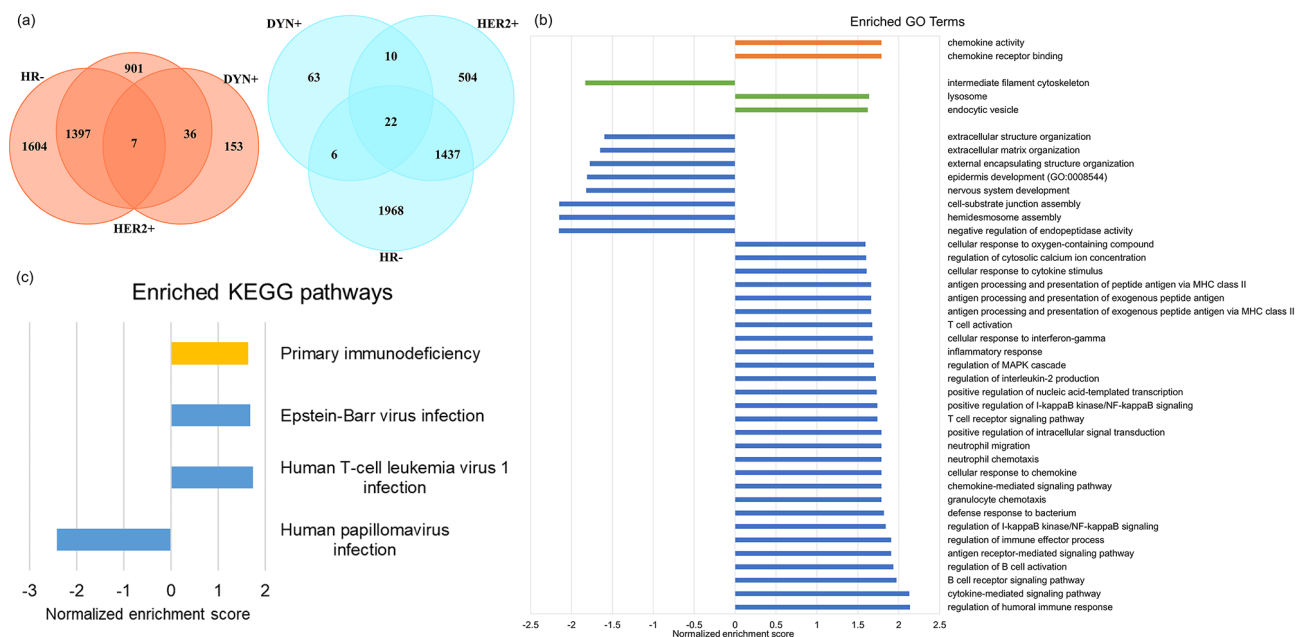
Fig. 6 (**a**) Up-regulated (left) and down-regulated (right) DEGs in DYN+, HR-, and HER2+group. (**b**) Enriched GO terms in DYN+group. (**c**) Enriched KEGG pathways in DYN+group

radiogenomic analysis [48]. In this study, we conducted a radiogenomic analysis to associate our dynamic model to the genomic profiles of breast tumors, providing insight into the underlying biological mechanisms of radiomics. Some DEGs in our DYN+subgroup is associated with better prognosis of breast cancer patients. For example, the DYN+subgroup has higher expression of CXCL9 which was demonstrated to associate with higher pCR rate in breast cancer patients receiving NAC in previous study [49]; HLA-DPB1 was up-regulated in DYN+subgroup and was also assocaited with more tumor infiltrating lymphcytes and thereby better prognosis [50]. On the other hand, DEGs such as CCL18 is associated with angiogenesis in breast cancer, demonstrating the potential of DYN to represent the dynamics in DCE-MRI [51].

Our study has several limitations. Due to the retrospective nature of the study, it is possible that our results suffer from spectrum bias and information bias, which may compromise the overall strength of evidence of our study. Besides, our study included a medium sample size ($n=785$) without external validation, which may not be representative enough for the large population of breast cancer patients. Therefore, further study is needed to externally validate our methods and conclusions in a prospective manner. Also, our study only employed DCE-MRI, while multi-parametric MR images could be available in clinics. Further exploration on incorporating other MR images is warranted.

Conclusions

In conclusion, this study quantified the dynamic characteristics of DCE-MRI by calculating dynamic properties of radiomic feature series and developed a dynamic model. The dynamic model can aid in improving pCR prediction of breast cancer patients receiving NAC. The potential biological underpinnings of the dynamic model was explored by demonstrating its association with tumor heterogeneity in gene expression. Further investigations on more biological associations and assisting treatment selection are warranted.

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s13058-024-01836-3.

Supplementary Material 1

Zhang *et al. Breast Cancer Research*   (2024) 26:77

Page 11 of 12

Data availability

Declarations

Ethics approval and consent to participate
Not applicable. A public dataset was used.

Consent for publication
Not applicable. A public dataset was used.

Competing interests
The authors declare no relevant financial or non-financial interests.

Author details
[1]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, China
[2]The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen, China

References

1. Breast cancer. Accessed December 6. 2023. https://www.who.int/news-room/fact-sheets/detail/breast-cancer.
2. Bonadonna G. Evolving concepts in the systemic adjuvant treatment of breast Cancer. Cancer Res. 1992;52(8):2127–37.
3. Korde LA, Somerfield MR, Carey LA, et al. Neoadjuvant chemotherapy, endocrine therapy, and targeted therapy for breast Cancer: ASCO Guideline. J Clin Oncol. 2021;39(13):1485–505. https://doi.org/10.1200/JCO.20.03399.
4. Wang-Lopez Q, Chalabi N, Abrial C, et al. Can pathologic complete response (pCR) be used as a surrogate marker of survival after neoadjuvant therapy for breast cancer? Crit Rev Oncol Hematol. 2015;95(1):88–104. https://doi.org/10.1016/j.critrevonc.2015.02.011.
5. Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet. 2014;384(9938):164–72. https://doi.org/10.1016/S0140-6736(13)62422-8.
6. Mann RM, Kuhl CK, Kinkel K, Boetes C. Breast MRI: guidelines from the European Society of Breast Imaging. Eur Radiol. 2008;18(7):1307–18. https://doi.org/10.1007/s00330-008-0863-7.
7. Fowler AM, Mankoff DA, Joe BN. Imaging neoadjuvant therapy response in breast Cancer. Radiology. 2017;285(2):358–75. https://doi.org/10.1148/radiol.2017170180.
8. Xiao J, Rahbar H, Hippe DS, et al. Dynamic contrast-enhanced breast MRI features correlate with invasive breast cancer angiogenesis. Npj Breast Cancer. 2021;7(1):1–9. https://doi.org/10.1038/s41523-021-00247-3.
9. Saha A, Harowicz MR, Grimm LJ, et al. A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. Br J Cancer. 2018;119(4):508–16. https://doi.org/10.1038/s41416-018-0185-8.
10. Granzier RWY, Ibrahim A, Primakov SP, 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. https://doi.org/10.3390/cancers13102447.
11. Fan M, Wu G, Cheng H, Zhang J, Shao G, Li L. Radiomic analysis of DCE-MRI for prediction of response to neoadjuvant chemotherapy in breast cancer patients. Eur J Radiol. 2017;94:140–7. https://doi.org/10.1016/j.ejrad.2017.06.019.
12. Peng S, Chen L, Tao J, et al. Radiomics Analysis of Multi-phase DCE-MRI in Predicting Tumor response to neoadjuvant therapy in breast Cancer. Diagnostics. 2021;11(11):2086. https://doi.org/10.3390/diagnostics11112086.
13. 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 Manag Res. 2021;13:5053–62. https://doi.org/10.2147/CMAR.S304547.
14. Li W, Partridge SC, Newitt DC, et al. Breast multiparametric MRI for prediction of Neoadjuvant Chemotherapy response in breast Cancer: the BMMR2 challenge. Radiol Imaging Cancer. 2024;6(1):e230033. https://doi.org/10.1148/rycan.230033.
15. Quicke P, Sun Y, Arias-Garcia M, et al. Voltage imaging reveals the dynamic electrical signatures of human breast cancer cells. Commun Biol. 2022;5(1):1–14. https://doi.org/10.1038/s42003-022-04077-2.
16. Orlova Y, Gorobtsov A, Sychev O, Rozaliev V, Zubkov A, Donsckaia A. Method for determining the Dominant type of human breathing using motion capture and Machine Learning. Algorithms. 2023;16(5):249. https://doi.org/10.3390/a16050249.
17. ISPY2 trial data 2. https://doi.org/10.7937/TCIA.KK02-6D95.
18. ISPY2 trial data1. https://doi.org/10.7937/TCIA.D8Z0-9T85.
19. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a Public Information Repository. J Digit Imaging. 2013;26(6):1045–57. https://doi.org/10.1007/s10278-013-9622-7.
20. Wang H, Yee D. I-SPY 2: a neoadjuvant adaptive clinical trial designed to improve outcomes in high-risk breast Cancer. Curr Breast Cancer Rep. 2019;11(4):303–10. https://doi.org/10.1007/s12609-019-00334-2.
21. Symmans WF, Peintinger F, Hatzis C, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol off J Am Soc Clin Oncol. 2007;25(28):4414–22. https://doi.org/10.1200/JCO.2007.10.6823.
22. Spring LM, Fell G, Arfe A, et al. Pathologic Complete Response after neoadjuvant chemotherapy and impact on breast Cancer recurrence and survival: a Comprehensive Meta-analysis. Clin Cancer Res off J Am Assoc Cancer Res. 2020;26(12):2838–48. https://doi.org/10.1158/1078-0432.CCR-19-3492.
23. Hylton NM, Gatsonis CA, Rosen MA, et al. Neoadjuvant chemotherapy for breast Cancer: functional tumor volume by MR Imaging predicts recurrence-free survival—results from the ACRIN 6657/CALGB 150007 I-SPY 1 TRIAL. Radiology. 2016;279(1):44–55. https://doi.org/10.1148/radiol.2015150013.
24. Zwanenburg A, Vallières M, Abdalah MA, et al. The image Biomarker Standardization Initiative: standardized quantitative Radiomics for High-Throughput Image-based phenotyping. Radiology. 2020;295(2):328–38. https://doi.org/10.1148/radiol.2020191145.
25. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic phenotype. Cancer Res. 2017;77(21):e104–7. https://doi.org/10.1158/0008-5472.CAN-17-0339.
26. Zwanenburg A, Leger S, Agolli L, et al. Assessing robustness of radiomic features by image perturbation. Sci Rep. 2019;9(1):614. https://doi.org/10.1038/s41598-018-36938-4.
27. Teng X, Zhang J, Ma Z, et al. Improving radiomic model reliability using robust features from perturbations for head-and-neck carcinoma. Front Oncol. 2022;12:974467. https://doi.org/10.3389/fonc.2022.974467.
28. Zhang J, Teng X, Zhang X, et al. Comparing effectiveness of image perturbation and test retest imaging in improving radiomic model reliability. Sci Rep. 2023;13(1):18263. https://doi.org/10.1038/s41598-023-45477-6.
29. Koo TK, Li MY. A Guideline of selecting and reporting Intraclass correlation coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
30. Lubba CH, Sethi SS, Knaute P, Schultz SR, Fulcher BD, Jones NS. catch22: CAnonical time-series CHaracteristics. Data Min Knowl Discov. 2019;33(6):1821–52. https://doi.org/10.1007/s10618-019-00647-x.
31. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. J Bioinform Comput Biol. 2005;03(02):185–205. https://doi.org/10.1142/S0219720005001004.
32. Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. In: *Advances in Neural Information Processing Systems*. Vol 30. Curran Associates, Inc.; 2017. Accessed December 6, 2023. https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.
33. Fitzgerald M, Saville BR, Lewis RJ, Decision Curve Analysis. JAMA. 2015;313(4):409–10. https://doi.org/10.1001/jama.2015.37.
34. NCBI. site.
35. Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000;25(1):25–9. https://doi.org/10.1038/75556.
36. The Gene Ontology Consortium, Aleksander SA, Balhoff J, et al. The Gene Ontology knowledgebase in 2023. Genetics. 2023;224(1):iyad031. https://doi.org/10.1093/genetics/iyad031.
37. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023;51(D1):D587–92. https://doi.org/10.1093/nar/gkac963.

38. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30. https://doi.org/10.1093/nar/28.1.27.
39. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. Protein Sci Publ Protein Soc. 2019;28(11):1947–51. https://doi.org/10.1002/pro.3715.
40. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a Nonparametric Approach. Biometrics. 1988;44(3):837–45. https://doi.org/10.2307/2531595.
41. Ihaka R, Gentleman R. R: a Language for Data Analysis and Graphics. J Comput Graph Stat. 1996;5(3):299–314. https://doi.org/10.1080/10618600.1996.10474713.
42. Van Rossum G, Drake FL. Python 3 reference Manual. CreateSpace; 2009.
43. Pedregosa F, Varoquaux G, Gramfort A et al. Scikit-learn: machine learning in Python. *Mach Learn PYTHON*.
44. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19(1):15. https://doi.org/10.1186/s13059-017-1382-0.
45. Fang Z, Liu X, Peltz G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. Bioinformatics. 2023;39(1):btac757. https://doi.org/10.1093/bioinformatics/btac757.
46. Prinzi F, Orlando A, Gaglio S, Vitabile S. Breast cancer classification through multivariate radiomic time series analysis in DCE-MRI sequences. Expert Syst Appl. 2024;249:123557. https://doi.org/10.1016/j.eswa.2024.123557.
47. Teng X, Zhang J, Zhang X, et al. Noninvasive imaging signatures of HER2 and HR using ADC in invasive breast cancer: repeatability, reproducibility, and association with pathological complete response to neoadjuvant chemotherapy. Breast Cancer Res. 2023;25(1):77. https://doi.org/10.1186/s13058-023-01674-9.
48. Bismeijer T, van der Velden BHM, Canisius S, et al. Radiogenomic analysis of breast Cancer by linking MRI phenotypes with Tumor Gene expression. Radiology. 2020;296(2):277–87. https://doi.org/10.1148/radiol.2020191453.
49. Liu H, Yang Z, Lu W, et al. Chemokines and chemokine receptors: a new strategy for breast cancer therapy. Cancer Med. 2020;9(11):3786–99. https://doi.org/10.1002/cam4.3014.
50. Lyu L, Yao J, Wang M et al. Overexpressed Pseudogene HLA-DPB2 Promotes Tumor Immune Infiltrates by Regulating HLA-DPB1 and Indicates a Better Prognosis in Breast Cancer. *Front Oncol.* 2020;10. Accessed December 19, 2023. https://www.frontiersin.org/articles/https://doi.org/10.3389/fonc.2020.01245.
51. Lin L, Chen YS, Yao YD, et al. CCL18 from tumor-associated macrophages promotes angiogenesis in breast cancer. Oncotarget. 2015;6(33):34758–73.

Publisher's Note