Heliyon 10 (2024) e24371

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Multiparametric MRI-based radiomics combined with pathomics features for prediction of the efficacy of neoadjuvant chemotherapy in breast cancer

Nan Xu [a,1], Xiaobin Guo [b,1], Zhiqiang Ouyang [c], Fengming Ran [d], Qinqing Li [a], Xirui Duan [c], Yu Zhu [a], Xiaofeng Niu [a], Chengde Liao [c,**], Jun Yang [a,*]

[a] Department of Radiology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital/Center. No. 519 Kunzhou Road, Xishan District, Kunming 650118, Yunnan, PR China
[b] Department of Radiology, Fuwai Central China Cardiovascular Hospital, Fuwai Road, Zhengzhou, Henan, 461464, PR China
[c] Department of Radiology, Kunming Yan'an Hospital (Yan'an Hospital Affiliated to Kunming Medical University), Kunming, PR China
[d] Department of Pathology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital/Center. No. 519 Kunzhou Road, Xishan District, Kunming 650118, Yunnan, PR China

ARTICLE INFO

Keywords:
Breast cancer
Radiomics
MRI
Neoadjuvant chemotherapy
Pathology

ABSTRACT

Purpose: The aim of this study is to investigate a new method that combines radiological and pathological breast cancer information to predict discrepancies in pathological responses for individualized treatment planning. We used baseline multiparametric magnetic resonance imaging and hematoxylin and eosin-stained biopsy slides to extract quantitative feature information and predict the pathological response to neoadjuvant chemotherapy in breast cancer patients.
Methods: We retrospectively collected data from breast cancer patients who received neoadjuvant chemotherapy in our hospital from August 2016 to January 2018; multiparametric magnetic resonance imaging (contrast-enhanced T1-weighted imaging and diffusion-weighted imaging) and whole slide image of hematoxylin and eosin-stained biopsy sections were collected. Quantitative imaging features were extracted from the multiparametric magnetic resonance imaging and the whole slide image were used to construct a radiopathomics signature model powered by machine learning methods. Models based on multiparametric magnetic resonance imaging or whole slide image alone were also constructed for comparison and referred to as the radiomics signature and pathomics signature models, respectively. Four modeling methods were used to establish prediction models. Model performances were evaluated using receiver operating characteristic curve analysis and the area under the curve, accuracy, sensitivity, specificity, positive predictive value, and negative predictive value.
Results: The radiopathomics signature model had favourable performance for the prediction of pathological complete response in the training set (the best value: area under the curve 0.83, accuracy 0.84, and sensitivity 0.87), and in the test set (the best value: area under the curve 0.91, accuracy 0.90, and sensitivity 0.88). In the test set, the radiopathomics signature model also

* Corresponding author.
** Corresponding author.
E-mail addresses: xu_nan97@126.com (N. Xu), 455032074@qq.com (X. Guo), 980519215@qq.com (Z. Ouyang), 1045123166@qq.com (F. Ran), qinqing_81@139.com (Q. Li), 248190836@qq.com (X. Duan), 1751487800@qq.com (Y. Zhu), 18288976651@163.com (X. Niu), chengdeliao@qq.com (C. Liao), imdyang@qq.com (J. Yang).
[1] Contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e24371
Received 20 April 2023; Received in revised form 25 December 2023; Accepted 8 January 2024
Available online 12 January 2024

significantly outperformed the radiomics signature (the best value: area under the curve 0.83, accuracy 0.64, and sensitivity 0.62), pathomics signature (the best value: area under the curve 0.60, accuracy 0.74, and sensitivity 0.62) ($p > 0.05$). Decision curve analysis and calibration curves confirmed the excellent performance of these prediction models in discrimination, calibration, and clinical usefulness.

Conclusions: The results of this study suggest that radiopathomics, the combination of both radiological information regarding the whole tumor and pathological information at the cellular level, could potentially predict discrepancies in pathological response and provide evidence for rational treatment plans.

1. Introduction

Neoadjuvant chemotherapy (NAC) is used in the treatment of locally advanced breast cancer patients and breast-conserving patients with large tumors [1]. NAC can shrink the tumor, reduce the clinical stage, improve the success rate of surgery, and achieve pathological complete remission (pCR) in patients, resulting in a longer disease-free survival interval and overall survival time [2]. However, the response of breast cancer to NAC varies greatly based on many factors, and 2 %–30 % of patients do not benefit from this treatment [3] and are delayed in receiving local treatment. In addition, due to the delay of the operation time window, NAC may not only increase the drug resistance and metastasis risk of tumors but also cause overtreatment of breast cancer [4]. Therefore, it is particularly important to effectively and comprehensively evaluate and predict whether breast cancer patients can benefit from NAC before they receive chemotherapy.

Radiomics is an innovative method of image quantitative analysis that extracts a large number of high-dimensional features that cannot be recognized by the naked eye from medical images through data mining and machine learning; the obtained features are then applied to a clinical decision support system [5]. Radiomics is of great value in disease characterization, tumor staging, curative effect evaluation and prognosis prediction [6,7]. Multi-parameter magnetic resonance imaging (mp-MRI) is the most sensitive imaging method for breast cancer, and it has become a routine examination for patients with breast cancer [8]. DCE-MRI is mainly used to clarify the biological changes within the tumor through morphological, kinetic and perfusion parameters, so that the curative effect of NAT can be evaluated more early and accurately. DWI is widely used as an auxiliary means of DCE-MRI. The combination of them can improve the specificity of imaging. At the same time, ADC value, as a biomarker of malignant tumor, can help to evaluate the response of breast cancer to NAT [9,10]. With the continuous development of image processing technology and the rise of imaging science, multi-parameter MRI-based radiomics can better reflect tumor biological heterogeneity and assist treatment decision-making in the differentiation of benign and malignant breast cancer, classification and grading, NAC treatment response and risk prediction [11,12]. MRI-based radiomics can not only predict the efficacy of NAC in breast cancer patients by evaluating the imaging characteristics during different chemotherapy cycles but also use multiparameter MRI to quantify the dynamic process of tumor development and provide key specific information about tumor characteristics and heterogeneity, which improves the efficiency of evaluation and prediction of the NAC response [13].

The prediction efficacy of MRI-based radiomics often depends on the choice of algorithms and models [14,15]. Previous studies have mainly focused on the morphological analysis of tumors, but studies have shown that the pathological features of tumors also affect the efficacy of NAC [16,17]. For example, some pathological features, including tumor-associated lymphocytes, may be independent predictors of the efficacy of NAC [18]. Quantitative analysis of whole slide images is a new method for analyzing histopathological features. Compared with traditional methods, it avoids the errors caused by subjective factors in pCR and prognostic

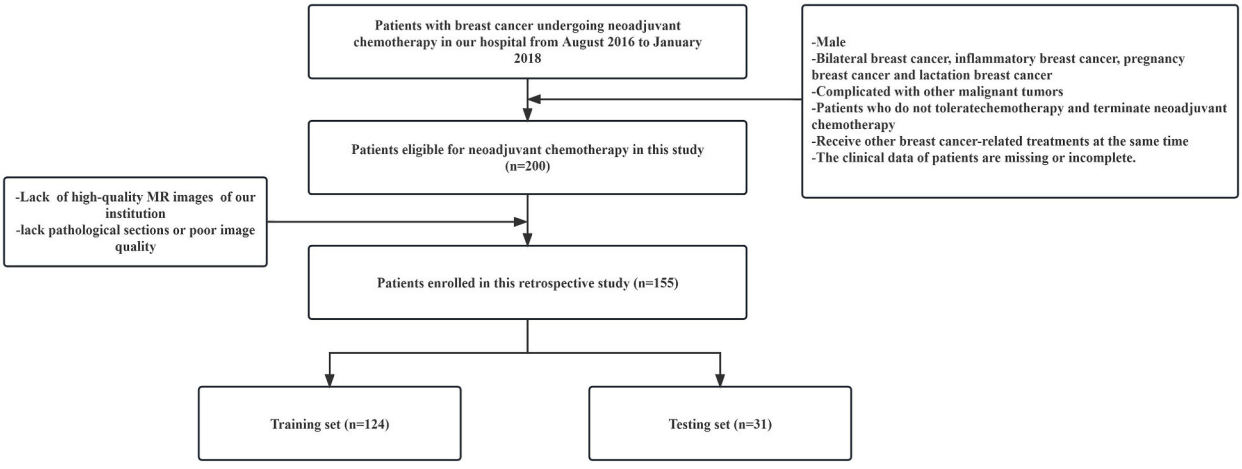

Fig. 1. Flowchart of the patient selection process.

ability evaluation and improves the accuracy of diagnosis. Studies have shown that artificial intelligence can be used to automatically analyze pathological features and provide relevant indicators for predicting the efficacy of neoadjuvant chemotherapy [19–21]. In this study, we extracted radiological features from multiparameter MRI of breast cancer patients, combined them with the histopathological characteristics identified by artificial intelligence, established a prediction model of breast cancer neoadjuvant chemotherapy, analyzed whether it can predict patients' pCR before treatment, and evaluated patients' response to chemotherapeutic drugs to better guide breast cancer neoadjuvant chemotherapy decision-making.

2. Materials and methods

2.1. Patients

This retrospective study was approved by the ethics committee of the ethics committee of our hospital, and the need for informed consent was waived. A total of 200 patients with breast cancer who underwent neoadjuvant chemotherapy in our hospital from August 2016 to January 2018 were consecutively enrolled in this study. The inclusion criteria (Fig. 1) were as follows: (1) breast cancer patients had completed standardized neoadjuvant chemotherapy; (2) before treatment, hematoxylin and eosin (H&E) staining biopsy was performed and confirmed as breast cancer; and (3) standardized baseline mp-MRI examination was performed. The exclusion criteria were as follows: (1) there were other tumors or diseases that affect the efficacy of neoadjuvant chemotherapy; (2) MRI images or pathological sections were of insufficient quality to obtain measurements; and (3) the clinical data of patients missing or incomplete. The decision to administer NAC was made cooperatively by the chief surgeon, oncologist, and patient. The chemotherapy regimen included cyclophosphamide + epirubicin + docetaxel, cyclophosphamide + docetaxel, or epirubicin + docetaxel, and all patients received complete neoadjuvant chemotherapy.

2.2. MRI image acquisition

All breast MRI examinations were performed using a 1.5 T MR scanner (Magnetom Avanto, Siemens, Germany) within 1 week before initiating NAC. Axial dynamic contrast enhanced (DCE) images and diffusion Weighted Imaging (DWI) with two b values (0 and 800 s/mm^2) were acquired for each patient. DCE sequences were collected after an initial fat-saturated T1-weighted precontrast scan. After an intravenous injection of 0.2 ml/kg gadolinium contrast, the first postcontrast scan was obtained, and then seven subsequent postcontrast images were acquired. Experimental details are summarized in Table 1 of the supplementary materials.

2.3. Whole slide image (WSI) acquisition

H&E-stained slides from biopsy formalin-fixed paraffin-embedded tissue were used for pathological diagnosis. WSIs for analysis

Table 1
Baseline characteristics of the patients.

Characteristics	Training Group (n = 124)		*p*	Testing Group (n = 31)		*p*
	pCR (n = 31)	Non-pCR (n = 93)		pCR (n = 8)	Non-pCR (n = 23)	
Age	44.161 ± 7.26	46.086 ± 8.533	0.262	51.250 ± 6.135	45.56 ± 7.347	0.059
cT stage			0.242			0.703
1	1 (0.032)	3 (0.032)		0 (0.000)	0 (0.000)	
2	27 (0.870)	67 (0.720)		8 (1.000)	20 (0.869)	
3	3 (0.096)	14 (0.150)		0 (0.000)	3 (0.130)	
4	0 (0.000)	9 (0.096)		0 (0.000)	0 (0.000)	
cN stage			0.616			0.704
0	4 (0.129)	22 (0.236)		2 (0.250)	7 (0.304)	
1	24 (0.774)	61 (0.655)		4 (0.500)	14 (0.608)	
2	2 (0.064)	7 (0.075)		1 (0.125)	1 (0.043)	
3	1 (0.032)	3 (0.032)		1 (0.125)	1 (0.043)	
ER status (%)			0.066			0.330
Positive	21 (0.677)	79 (0.849)		3 (0.375)	19 (0.826)	
Negative	10 (0.322)	14 (0.150)		5 (0.625)	4 (0.173)	
PR status (%)			1.000			0.008*
Positive	26 (0.838)	80 (0.860)		3 (0.375)	21 (0.913)	
Negative	5 (0.161)	13 (0.139)		5 (0.625)	2 (0.086)	
HER2 status (%)			0.916			0.925
Positive	19 (0.612)	54 (0.580)		5 (0.625)	12 (0.521)	
Negative	12 (0.387)	39 (0.419)		3 (0.375)	11 (0.478)	
Ki-67 status (%)			0.390			0.142
Positive	22 (0.709)	56 (0.602)		8 (1.000)	15 (0.652)	
Negative	9 (0.290)	37 (0.397)		0 (0.000)	8 (0.347)	

ER estrogen receptor, PR progesterone receptor, HER2 human epidermal growth factor receptor 2.
*$p < 0.05$.

were collected by panoramic digital image scanning technology (KF-PRO-005, Ningbo Konfoong Bioinformation Tech Co., Ltd.). Panoramic digital image scanning technology was used to collect WSIs at 20× magnification for analysis and generate 0.25 μm/pixel digital pathological images for pathological feature extraction.

2.4. Tumor segmentation and radiomics feature extraction from MRI

Two full-time radiologists with 10 years of working experience conducted the first phase of DCE subtraction image and DWI image processing in Siemens syngo. via workstation. The radiology application (Syngo.via, Frontier Radiomics) uses a semiautomatic segmentation algorithm to outline the region of interest (ROI), and automatically extracts quantitative image features, which are called radiomic features and include 18 first order statistics features, 75 texture features, 17 shape features and 744 wavelet features. The texture features included 24 Gy cooccurrence matrix features (GLCM), 14 Gy level dependence matrix features (GLDM), 16 Gy level run length matrix features (GLRLM), 16 Gy level size zone matrix features (GLSZM), and 5 neighbor gray tone difference matrix features (NGTDM). To obtain the wavelet features, we first applied wavelet filters to each input MR image, yielding 8 decompositions (all possible combinations of applying either a high or a low pass filter), which were used to calculate the textural features. Shape-based features quantitatively described the two-dimensional size and shape of the ROI.

2.5. ROI segmentation and pathomics feature extraction from WSIs

The pathologist applied a 20× magnification to the panoramic digital section, identified the tumor cells and intercepted the image. Then, CellProfiler (version 3.1.9, https://cellprofiler.org/) [22] was used to automatically extract the features of the intercepted 20× magnified pathological section images to obtain the heterotypic information of tumor cells. CellProfiler is an open-source tool widely used in the field of biological image analysis and can be used to analyze images quantitatively. In this study, features such as pixel intensity, morphology and nuclear texture were extracted and used as derive the pathological features.

2.6. Feature selection and forecasting model construction

The feature selection workflow is shown in Fig. 2. In the process of radiological feature extraction, intraclass correlation coefficients (ICCs) were used to evaluate interobserver consistency. The radiological features with ICCs greater than 0.75 were categorized as having good agreement with high reproducibility, and those with ICCs less than 0.75 were eliminated. The same process of ICC evaluation of pathological features was performed with the radiological features. Feature selection was performed separately for the MRI radiomic features, pathomic features, and radiopathomic features (pathomic features + radiomics features) of the training set. By

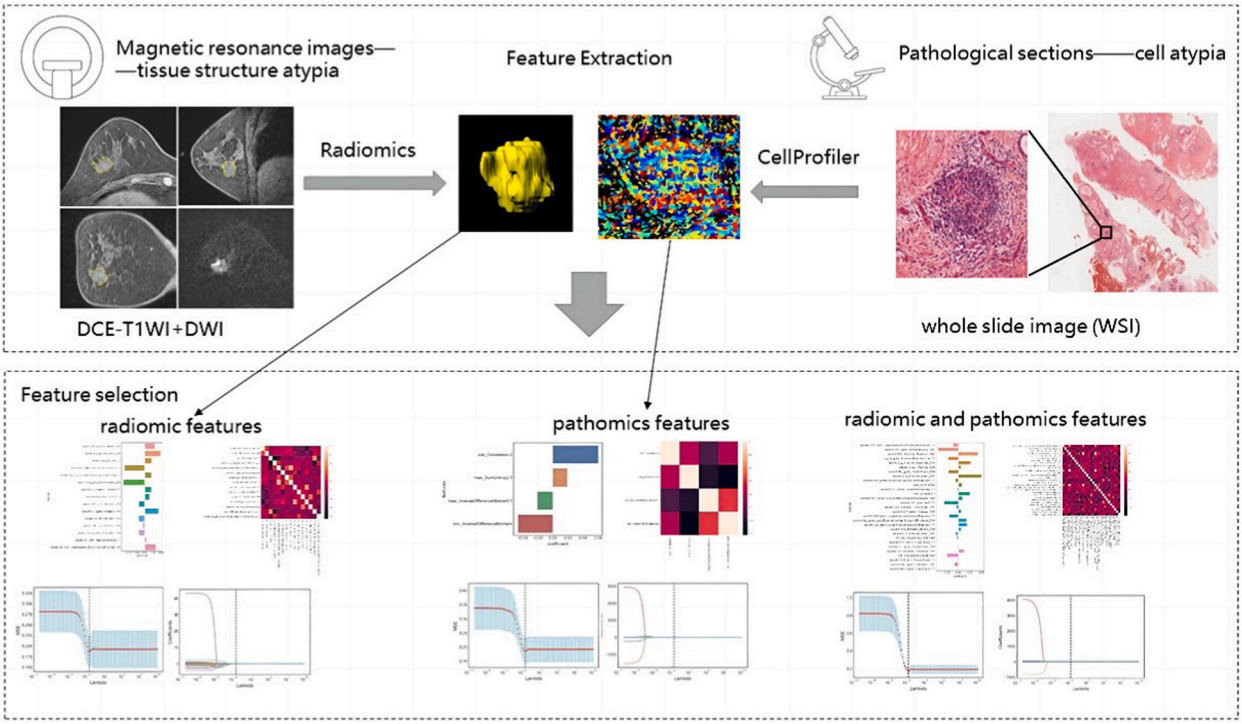

Fig. 2. Comprehensive schematics of the methodology developed and implemented in this study for radiomic feature and pathomic feature selection.

Pearson correlation analysis, features with correlation coefficients higher than 0.9 were first excluded. Then, using one-way analysis of variance (ANOVA), the top 50 features with the largest variances were retained. Finally, feature filtering was performed using LASSO (Fig. 1 in the Supplementary material). Detailed information on the features is presented in Tables 2, 3, and 4 of the Supplemental Material. Then, using the selected features, logistic regression, Naïve Bayesian, random forest regression, and XGBoost models were implement and their performance on different data sets was compared. A 5-fold cross-validation method was applied for the training process of all models, i.e., the training set was divided equally into five parts, with one part serving as the testing set and the remaining four parts as the training sets. The model established by using pathomic features alone is referred to as the pathomics signature (PS) model, the model established by using radiomic features alone is referred to as the radiomics signature (RS) model, and the model established by combining radiomic features and pathological features is referred to as the radiopathomics signature (RPS) model (As shown in Table 3).

2.7. Statistical analysis

Differences between groups were assessed using Mann–Whitney U tests or t tests for continuous variables and chi-squared tests for categorical variables. We assessed the predictive ability of the PS, RS and RPS models using receiver operating characteristic (ROC) curve analysis. The area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. The DeLong test was performed to determine whether the differences in predictive ability between the three models were statistically significant. The statistical analyses were performed using Python (version 3.6.5).

3. Results

3.1. Patient characteristics

A total of 155 breast cancer patients with an average age of 46 years (range 38–54 years, 46.00 ± 8.06) were enrolled and divided into the training (n = 124) and the testing sets (n = 31). Table 1 shows the baseline characteristics of the patients in the training set and testing set. The distributions of the patient characteristics in the two data sets were similar. In the training and testing data sets, the proportions of patients with pCR were 25.0 % and 25.8 %, respectively, and except for the PR status in the testing set, there were no significant differences in clinical characteristics such as age, stage and Ki-67 status between the pCR group and the non-pCR group ($p > 0.05$).

3.2. Feature selection

A total of 1708 radiomic features were extracted from the DCE and DWI images of each patient, including 854 DCE features and 854 DWI features, and a total of 260 pathomic features were extracted from pathological section images. The intraobserver ICCs of the MRI radiomic features ranged from 0.89 to 0.95, and the interobserver ICCs of the pathomic features ranged from 0.83 to 0.97. This result indicated that feature extraction had good repeatability between observers. With the LASSO feature coefficients, rad-scores for different data sets were constructed (Eqs. 1, 2, 3 in the Supplementary material). After performing ANOVA and LASSO, 15 MRI radiomic features, 4 pathomic features and 29 radiopathological features were selected, and the models were established.

The results of 5-fold cross-validations of different models using the set of radiomics features showed that Naïve Bayesian model and random forest model had the highest accuracy (ACC), which were 0.71 ± 0.12 and 0.71 ± 0.11, respectively, and the logistic regression model had the highest AUC (0.68 ± 0.11). The results of five cross-validations of different models using the pathomics features showed that the best model was logistic regression, with an ACC of 0.75 ± 0.04 and an AUC of 0.71 ± 0.04. Based on radiomic features and pathomic features, the parameters obtained of the RPS models were improved, and the optimal model was the naive Bayesian model, with an ACC of 0.89 ± 0.05 and AUC of 0.96 ± 0.03. Detailed information on the results is presented in Tables 5, 6, and 7 of the supplemental material.

Table 2
Prediction performance of different models for the training set.

Features	Models	AUC	Sensitivity	Specificity	Accuracy
Radiomics	Logistic Regression	0.68	0.68	0.74	0.73
	Naïve Bayesian	0.65	0.61	0.87	0.81
	Random forest	0.64	0.71	0.96	0.90
	XGBoost	0.58	0.90	0.97	0.95
Pathomics	Logistic Regression	0.71	0.03	1.0	0.76
	Naïve Bayesian	0.68	0.26	0.89	0.73
	Random forest	0.67	0.84	0.99	0.95
	XGBoost	0.67	0.90	0.92	0.92
Radiopathomics	Logistic Regression	0.83	0.87	0.83	0.84
	Naïve Bayesian	0.74	0.71	0.95	0.89
	Random forest	0.77	0.81	0.99	0.94
	XGBoost	0.79	0.90	1.0	0.94

Table 3
Prediction performance of different models for the testing set.

Features	Models	AUC	Sensitivity	Specificity	Accuracy
Radiomics	Logistic Regression	0.83	0.62	0.65	0.64
	Naïve Bayesian	0.79	0.62	0.78	0.74
	Random forest	0.81	0.50	0.95	0.83
	XGBoost	0.78	0.50	0.82	0.74
Pathomics	Logistic Regression	0.57	0.78	1.0	0.77
	Naïve Bayesian	0.59	0.25	0.86	0.70
	Random forest	0.60	0.62	0.78	0.74
	XGBoost	0.60	0.62	0.52	0.54
Radiopathomics	Logistic Regression	0.86	0.75	0.69	0.70
	Naïve Bayesian	0.91	0.88	0.91	0.90
	Random forest	0.79	0.37	0.91	0.77
	XGBoost	0.65	0.37	0.86	0.74

3.3. Performance comparison of the prediction models

As shown in Fig. 3a–c and Table 2 and in the training set, the highest AUC of the RS models was 0.68 with an ACC of 0.73, and the highest AUC of the PS models was 0.71 with an ACC of 0.76. The highest AUC value of the RPS models was 0.83 with an ACC of 0.84, which surpassed than that of the RS and PS models in the training set. In the test set (Fig. 3d–f and Table 2), the RS models had better AUCs than the PS models, and the highest AUC was 0.83 with an ACC of 0.64. Compared to the four RS models, the highest AUC for the PS models was only 0.60 with an ACC of 0.74. While the RPS obtained by Naïve Bayesian model is the best AUC of 0.91 in the test set, with ACC of 0.90.

Shown by the DeLong test, there was a significant difference between the AUC values of the models based on pathomic features only and those based on the combined features ($p = 0.03$). There were no significant differences between the AUC values of the models using radiomics features only and those using the combined features ($p = 0.32$), nor were there differences between the results of models using pathomic features only and those using MRI radiomics features only ($p = 0.41$). The decision curve analysis (DCA) results are shown in Fig. 4. In the test set, compared with the random forest model using MRI radiomic features and the random forest model using pathomic features, the naive Bayesian model based on radiopathomic features consistently exhibited higher overall net benefits (Fig. 4a). In addition, the calibration curve showed that, compared with the random forest model using MRI radiomic features and the random forest model using pathomic features, the prediction probability of the naive Bayesian model based on radiopathomic features showed good agreement with the real outcomes of neoadjuvant chemotherapy for breast cancer (Fig. 4b–d).

4. Discussion

NAC can reduce or even fully eliminate the primary tumor and regional lymph nodes, which has gradually become an important part of the standard treatment of breast cancer. It is very important to find feasible methods that can predict the sensitivity of patients

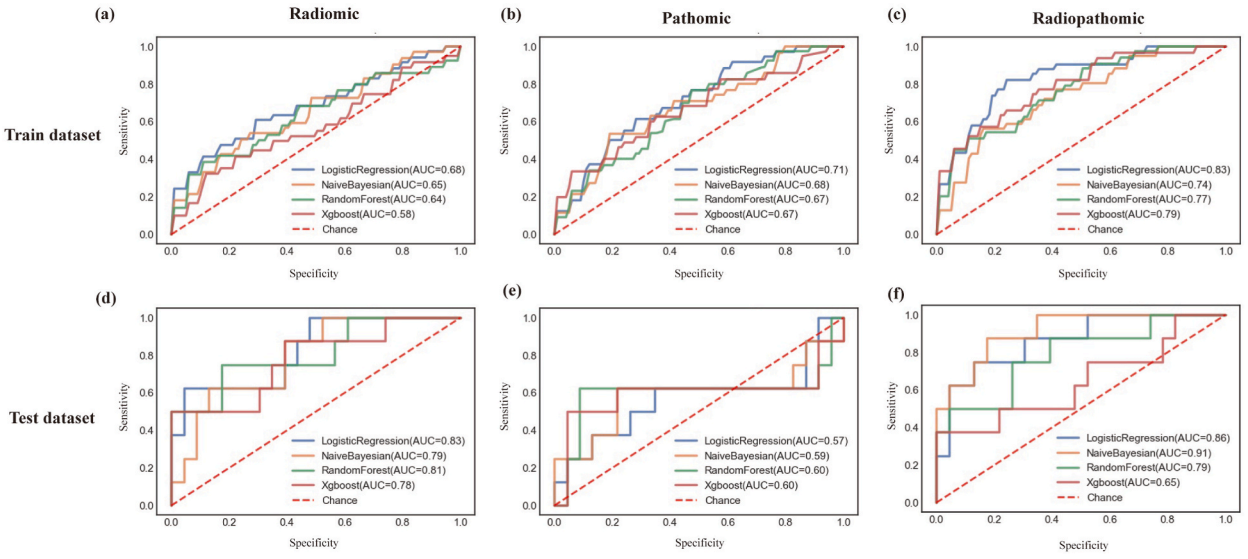

Fig. 3. The receiver operating characteristic (ROC) curves of the radiomics signature(a), (d), the pathomics signature(b), (e), and the radiopathomics signature(c) (f) in the training and test cohorts.

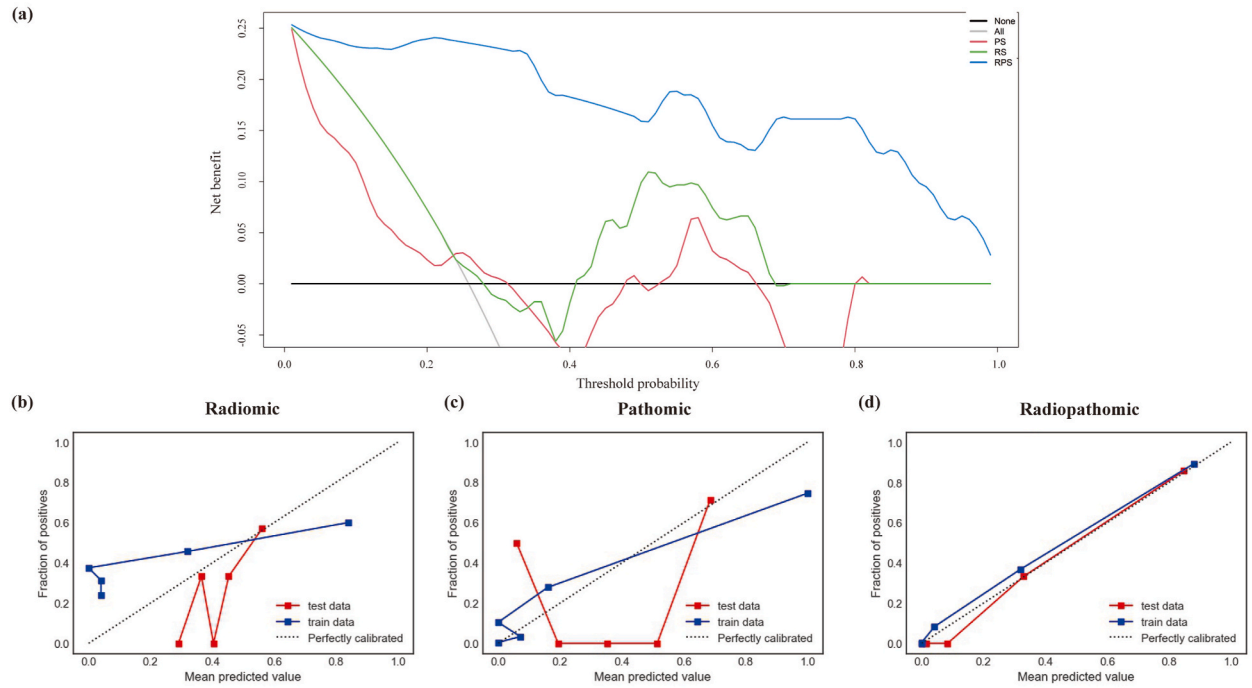

Fig. 4. Decision curve of the three models in the test data set (a) and calibration curves of the three models (b, c, d).

to NAC and whether they will achieve pCR because these predictions can help some insensitive patients avoid the toxicity of NAC treatment and avoid missing the best time for surgery; they can also guide the choice of follow-up chemotherapy drugs for treatment-sensitive patients. In this study, radiological and pathological features were extracted from mp-MRI and WSI, respectively, several feature prediction models were established, and the effectiveness of the different prediction models was compared. The results showed that the feature combination model established by Naïve Bayesian method had the best overall performance in prediction of pathological response after breast cancer patients received NAC; the optimal models of pathological features and radiological features alone were not as good as the combined model.

Radiomics is a rapidly developing research field [5], and many researchers have used radiomics methods to predict the pathological response to neoadjuvant chemotherapy in breast cancer patients. In this study, only the prediction model based on radiomics features was used, AUC values of 0.78–0.83 could also be obtained on the testing set. Many researchers have also confirmed that radiomics models based on MRI can predict the pathological response to neoadjuvant chemotherapy in patients with breast cancer [23–25]. Liu et al. [26] evaluated the performance of a new predictive model using multiparametric MRI in patients with breast cancer and achieved an AUC of 0.86. After that, researchers further improved the diagnostic efficiency by improving the algorithm and model development process [27–29]. However, radiomics only extracts the atypical features of tumor tissue, while the pathological characteristics of the tumor cell dimension also play an important role in the clinical decision-making of breast radiotherapy and medical oncology [30]. Therefore, this study also extracted the pathomics features of patient biopsy sections.

In contrast to radiomics, which largely relies on in vivo imaging, pathomics considers cell and subcellular biological information. This method has been applied and achieved good results in predicting the prognosis of renal cell carcinoma [31], diagnosing and predicting the survival of bladder cancer [32], and predicting the efficacy of neoadjuvant therapy in patients with rectal cancer [33]. However, pathomics is a recently established subfield, there are few relevant studies, and there is no standardized feature specification of radiomics systems for verification and calibration[34]. In this study, the AUC of the prediction model established by using pathomics alone was between 0.57 and 0.60, and the results were consistently not as good as those of the radiomics model, but the accuracy of the RPS model combining the radiomic and pathomics features was higher than those of the PS and RS models. Therefore, we speculate that more comprehensive information can be obtained by combining the tissue structure atypia of patients' MRI and the cell atypia of pathological sections to better predict the pathological responses of patients. However, there were few pathomics features extracted, so in this study, the performance of models using pathomics alone was not good. When using pathomics features alone, the selected features are not consistent with the pathomic features used in the radiopathomics model, which may be because when modeling alone, the selected features cannot accurately represent the characteristics of patients' tumor cells in the training set. After combining the two features, there are more types of features, and the selected features can better reflect the real data. Therefore, it is necessary to combine radiomic and pathomics features for joint modeling. In this study, the best AUC of the respective optimal models based on pathomic features only was 0.71, while using the combined features, it was 0.91. Meanwhile, the accuracy increased from 0.76 to 0.90.

In this study, four modeling methods were used in conjunction with the three feature sets to obtain the best prediction model. The random forest algorithm performed relatively well on both data sets using only pathological or only radiological features (both AUC

and ACC indicated high performance). This is because random forest is an integrated algorithm. Thus, it can handle both discrete data and continuous data. It is not sensitive to data and has strong adaptability to data sets. Additionally, random forest algorithm introduces randomness to avoid overfitting. The model that achieves the best performance in the combined use of radiomic features and pathomics features data sets was Naive Bayes algorithm. Naive Bayesian algorithm is not sensitive to noise in high-dimensional data [35,36]. Because it only uses the important features to reduce the influence of noise information, the Naive Bayesian algorithm can achieve the best performance.

In the RPS, the first three radiomics features with the highest coefficient were reference to as the wavelet-HHL_firstorder_Maximum_DWI, wavelet-HHL_glcm_InverseVariance_DWI, and wavelet-LHH_glcm_Imc2_T1. The first feature is the first-order feature extracted from the DWI sequence, which describes the distribution of gray-level pixel intensity in the image [37]. In this study, the larger the eigenvalue, the more likely it was for patients to achieve pCR. The latter two features are the GLCM (gray level cooccurrence matrix) features extracted from the DWI and T1 enhancement sequences, which are used to evaluate the combination of gray-level spatial dependence [38,39]. In this study, the coefficients of these two features were negative. With increases in these two feature values, the possibility of patients achieving pCR is smaller.

The limitation of this study lies in the small number of cases and the lack of external verification. This is because there are few inpatients who meet the enrollment conditions, so it was difficult to obtain data from other research centers for external verification. Due to the recent introduction of pathomics and less extractable features, more researchers are needed to develop and standardize the features.

5. Conclusion

Overall, our research indicates that the combined analysis of functional, morphological and biological characteristics facilitates a comprehensive information platform for accurate and noninvasive prediction of the therapeutic effects of neoadjuvant chemotherapy in breast cancer patients. Therefore, the predictive model using multiparameter MR imaging combined with histopathology is conducive to the stratification of patients before treatment and can support more personalized and targeted clinical decision-making.

Author Contributions

Xiaofeng Niu: Software, Methodology. Qinqing Li: Methodology, Investigation. Yu Zhu: Visualization, Software. Xirui Duan: Writing – review & editing, Software. Xiaobin Guo: Data curation. Nan Xu: Writing – original draft, Conceptualization. Fengming Ran: Methodology. Zhiqiang Ouyang: Formal analysis, Data curation. Chengde Liao: Project administration, Funding acquisition. Jun Yang: Writing – review & editing, Project administration, Funding acquisition

Funding

This study was supported by the National Natural Science Foundation of China (grant no. 82160340, 82060313), the joint project of basic research of Kunming Medical University and Department of Science and Technology of Yunnan Province (202301AY070001-243), the Outstanding Youth Science Foundation of Yunnan Basic Research Project (202201AW070002).

Ethical statement

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was approved by the ethics committee of Yunnan Cancer Hospital (Approval No.KYCS2022144).

Data availability statement

Data generated or analyzed during the study are available from the corresponding author by request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e24371.

References

[1] N.F. Pondé, D. Zardavas, M. Piccart, Progress in adjuvant systemic therapy for breast cancer[J], Nat. Rev. Clin. Oncol. 16 (1) (2019) 27–44.
[2] P. Cortazar, L. Zhang, M. Untch, et al., Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis, Lancet 384 (9938) (2014) 164–172.
[3] P. Cortazar, C.E. Geyer Jr., Pathological complete response in neoadjuvant treatment of breast cancer[J], Ann. Surg Oncol. 22 (5) (2015) 1441–1446.
[4] N. Harbeck, F. Penault-Llorca, J. Cortes, et al., Breast cancer[J], Nat. Rev. Dis. Prim. 5 (1) (2019) 66.
[5] M.E. Mayerhoefer, A. Materka, G. Langs, et al., Introduction to radiomics[J], J. Nucl. Med. 61 (4) (2020) 488–495.
[6] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data[J], Radiology 278 (2) (2016) 563–577.
[7] P. Lambin, R.T.H. Leijenaar, T.M. Deist, et al., Radiomics: the bridge between medical imaging and personalized medicine[J], Nat. Rev. Clin. Oncol. 14 (12) (2017) 749–762.
[8] A.K. Swayampakula, C. Dillis, J. Abraham, Role of MRI in screening, diagnosis and management of breast cancer[J], Expert Rev. Anticancer Ther. 8 (5) (2008) 811–817.
[9] A. Allarakha, Y. Gao, H. Jiang, et al., Prediction and prognosis of biologically aggressive breast cancers by the combination of DWI/DCE-MRI and immunohistochemical tumor markers[J], Discov. Med. 27 (146) (2019) 7–15.
[10] S.Y. Kim, N. Cho, I.A. Park, et al., Dynamic contrast-enhanced breast MRI for evaluating residual tumor size after neoadjuvant chemotherapy[J], Radiology 289 (2) (2018) 327–334.
[11] S. Chen, Z. Shu, Y. Li, et al., Machine learning-based radiomics nomogram using magnetic resonance images for prediction of neoadjuvant chemotherapy efficacy in breast cancer patients[J], Front. Oncol. 10 (2020) 1410.
[12] M. Caballo, W.B.G. Sanderink, L. Han, et al., Four-Dimensional Machine Learning Radiomics for the Pretreatment Assessment of Breast Cancer Pathologic Complete Response to Neoadjuvant Chemotherapy in Dynamic Contrast-Enhanced MRI[J], J Magn Reson Imaging, 2022.
[13] Y. Huang, L. Wei, Y. Hu, et al., Multi-parametric MRI-based radiomics models for predicting molecular subtype and androgen receptor expression in breast cancer[J], Front. Oncol. 11 (2021) 706733.
[14] B. Reig, L. Heacock, K.J. Geras, et al., Machine learning in breast MRI[J], J. Magn. Reson. Imag. 52 (4) (2020) 998–1018.
[15] S. Gitto, R. Cuocolo, K. Van Langevelde, et al., MRI radiomics-based machine learning classification of atypical cartilaginous tumour and grade II chondrosarcoma of long bones[J], EBioMedicine 75 (2022) 103757.
[16] A. Vörös, E. Csörgő, B. Kővári, et al., Different methods of pretreatment Ki-67 labeling index evaluation in core biopsies of breast cancer patients treated with neoadjuvant chemotherapy and their relation to response to therapy[J], Pathol. Oncol. Res. 21 (1) (2015) 147–155.
[17] G.M. Baker, T.A. King, S.J. Schnitt, Evaluation of breast and axillary lymph node specimens in breast cancer patients treated with neoadjuvant systemic therapy [J], Adv. Anat. Pathol. 26 (4) (2019) 221–234.
[18] C. Denkert, S. Loibl, A. Noske, et al., Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer, J. Clin. Oncol. 28 (1) (2010) 105–113.
[19] M. Peikari, S. Salama, S. Nofech-Mozes, et al., Automatic cellularity assessment from post-treated breast surgical specimens[J], Cytometry 91 (11) (2017) 1078–1087.
[20] A.C. Li, J. Zhao, C. Zhao, et al., Quantitative digital imaging analysis of HER2 immunohistochemistry predicts the response to anti-HER2 neoadjuvant chemotherapy in HER2-positive breast carcinoma[J], Breast Cancer Res. Treat. 180 (2) (2020) 321–329.
[21] S. Akbar, M. Peikari, S. Salama, et al., Automated and manual quantification of tumour cellularity in digital slides for tumour burden assessment, Sci. Rep. 9 (1) (2019) 14099.
[22] C. Mcquin, A. Goodman, V. Chernyshev, et al., CellProfiler 3.0: next-generation image processing for biology[J], PLoS Biol. 16 (7) (2018) e2005970.
[23] X. Chen, X. Chen, J. Yang, et al., Combining dynamic contrast-enhanced magnetic resonance imaging and apparent diffusion coefficient maps for a radiomics nomogram to predict pathological complete response to neoadjuvant chemotherapy in breast cancer patients[J], J. Comput. Assist. Tomogr. 44 (2) (2020) 275–283.
[24] K. Cao, B. Zhao, X.T. Li, et al., Texture analysis of dynamic contrast-enhanced MRI in evaluating pathologic complete response (pCR) of mass-like breast cancer after neoadjuvant therapy[J], JAMA Oncol. 2019 (2019) 4731532.
[25] A.G.V. Bitencourt, P. Gibbs, C. Rossi Saccarelli, et al., MRI-based machine learning radiomics can predict HER2 expression level and pathologic response after neoadjuvant therapy in HER2 overexpressing breast cancer[J], EBioMedicine 61 (2020) 103042.
[26] Z. Liu, Z. Li, J. Qu, et al., Radiomics of multiparametric MRI for pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study[J], Clin. Cancer Res. 25 (12) (2019) 3538–3547.
[27] Q. Xiong, X. Zhou, Z. Liu, et al., Multiparametric MRI-based radiomics analysis for prediction of breast cancers insensitive to neoadjuvant chemotherapy, Clin. Transl. Oncol. 22 (1) (2020) 50–59.
[28] J. Zhou, J. Lu, C. Gao, et al., Predicting the response to neoadjuvant chemotherapy for breast cancer: wavelet transforming radiomics in MRI, BMC Cancer 20 (1) (2020) 100.
[29] E.J. Sutton, N. Onishi, D.A. Fehr, et al., A machine learning model that classifies breast cancer pathologic complete response on MRI post-neoadjuvant chemotherapy[J], Breast Cancer Res. 22 (1) (2020) 57.
[30] A.R. Green, D. Soria, J. Stephen, et al., Nottingham Prognostic Index Plus: validation of a clinical decision making tool in breast cancer in an independent series [J], J Pathol Clin Res 2 (1) (2016) 32–40.
[31] S. Chen, L. Jiang, F. Gao, et al., Machine learning-based pathomics signature could act as a novel prognostic marker for patients with clear cell renal cell carcinoma[J], Br. J. Cancer 126 (5) (2022) 771–777.
[32] S. Chen, L. Jiang, X. Zheng, et al., Clinical use of machine learning-based pathomics signature for diagnosis and survival prediction of bladder cancer[J], Cancer Sci. 112 (7) (2021) 2905–2914.
[33] L. Shao, Z. Liu, L. Feng, et al., Multiparametric MRI and whole slide image-based pretreatment prediction of pathological response to neoadjuvant chemoradiotherapy in rectal cancer: a multicenter radiopathomic study[J], Ann. Surg Oncol. 27 (11) (2020) 4296–4306.
[34] C.K. Kuhl, D. Truhn, The long route to standardized radiomics: unraveling the knot from the end[J], Radiology 295 (2) (2020) 339–341.
[35] V. Giannini, S. Rosati, A. Defeudis, et al., Radiomics predicts response of individual HER2-amplified colorectal cancer liver metastases in patients treated with HER2-targeted therapy[J], Int. J. Cancer 147 (11) (2020) 3215–3223.
[36] C. Feng, H. Zhao, Y. Li, et al., Improved detection of focal cortical dysplasia in normal-appearing FLAIR images using a Bayesian classifier[J], Med. Phys. 48 (2) (2021) 912–925.
[37] T.W. Fan, H. Malhi, B. Varghese, et al., Computed tomography-based texture analysis of bladder cancer: differentiating urothelial carcinoma from micropapillary carcinoma[J], Abdom Radiol (NY) 44 (1) (2019) 201–208.
[38] X. Li, M. Guindani, C.S. Ng, et al., Spatial Bayesian modeling of GLCM with application to malignant lesion characterization[J], J. Appl. Stat. 46 (2) (2018) 230–246.
[39] W.A. Noortman, D. Vriens, C.H. Slump, et al., Adding the temporal domain to PET radiomic features[J], PLoS One 15 (9) (2020) e0239438.

Abbreviations

ACC: Accuracy
AUC: Area under the curve

DCA: Decision curve analysis
DCE: Dynamic contrast enhanced
DWI: Diffusion weighted imaging
ER: Estrogen receptor
GLCM: Gray cooccurrence matrix
GLDM: Gray level dependence matrix
GLRLM: Gray level run length matrix
GLSZM: Gray level size zone matrix
H&E: Hematoxylin and eosin
HER2: Human epidermal growth factor receptor 2
ICCs: Intraclass correlation coefficients
mp-MRI: Multiparametric magnetic resonance imaging
MRI: Magnetic resonance imaging
NAC: Neoadjuvant chemotherapy
NGTDM: Neighbor gray tone difference matrix
NPV: Negative predictive value
pCR: Pathological complete remission
PPV: Positive predictive value
PR: Progesterone receptor
PS: Pathomics signature
ROC: Receiver operating characteristic
ROI: Region of interest
RPS: Radiopathomics signature
RS: Radiomics signature
WSI: Whole slide image