Journal for ImmunoTherapy of Cancer

Radiomic and clinical data integration using machine learning predict the efficacy of anti-PD-1 antibodies-based combinational treatment in advanced breast cancer: a multicentered study

Jianli Zhao,[1,2] Zhixian Sun,[1,2] Yunfang Yu,[3,4] Zhongyu Yuan,[5] Ying Lin,[6] Yujie Tan,[3] Xiaohui Duan,[7] Herui Yao,[1,2] Ying Wang,[1,2] Jieqiong Liu [1,2]

For numbered affiliations see end of article.

Correspondence to
Professor Jieqiong Liu;
liujieqiong01@163.com

Dr Ying Wang;
wangy556@mail.sysu.edu.cn

Dr Herui Yao;
yaoherui@mail.sysu.edu.cn

ABSTRACT

Background Immune checkpoint inhibitors (ICIs)-based therapy, is regarded as one of the major breakthroughs in cancer treatment. However, it is challenging to accurately identify patients who may benefit from ICIs. Current biomarkers for predicting the efficacy of ICIs require pathological slides, and their accuracy is limited. Here we aim to develop a radiomics model that could accurately predict response of ICIs for patients with advanced breast cancer (ABC).

Methods Pretreatment contrast-enhanced CT (CECT) image and clinicopathological features of 240 patients with ABC who underwent ICIs-based treatment in three academic hospitals from February 2018 to January 2022 were assigned into a training cohort and an independent validation cohort. For radiomic features extraction, CECT images of patients 1 month prior to ICIs-based therapies were first delineated with regions of interest. Data dimension reduction, feature selection and radiomics model construction were carried out with multilayer perceptron. Combined the radiomics signatures with independent clinicopathological characteristics, the model was integrated by multivariable logistic regression analysis.

Results Among the 240 patients, 171 from Sun Yat-sen Memorial Hospital and Sun Yat-sen University Cancer Center were evaluated as a training cohort, while other 69 from Sun Yat-sen University Cancer Center and the First Affiliated Hospital of Sun Yat-sen University were the validation cohort. The area under the curve (AUC) of radiomics model was 0.994 (95% CI: 0.988 to 1.000) in the training and 0.920 (95% CI: 0.824 to 1.000) in the validation set, respectively, which were significantly better than the performance of clinical model (0.672 for training and 0.634 for validation set). The integrated clinical-radiomics model showed increased but not statistical different predictive ability in both the training (AUC=0.997, 95% CI: 0.993 to 1.000) and validation set (AUC=0.961, 95% CI: 0.885 to 1.000) compared with the radiomics model. Furthermore, the radiomics model could divide patients under ICIs-therapies into high-risk and low-risk group with significantly different progression-free survival both in training (HR=2.705, 95% CI: 1.888 to 3.876, p<0.001) and validation set (HR=2.625, 95% CI: 1.506 to 4.574, p=0.001), respectively. Subgroup analyses showed that the radiomics model was not influenced by programmed death-ligand 1 status, tumor metastatic burden or molecular subtype.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Current biomarkers such as programmed death-ligand 1 (PD-L1), tumor mutation burden, stromal tumor infiltrating lymphocytes, and microsatellite instability/ defective mismatch repair for predicting the efficacy of immunotherapy require pathological slides, and their accuracy is limited. Radiomics have shown promising results to predict response to immunotherapy in some solid tumors. However, no radiomic biomarkers have been reported in predicting the response of immune checkpoint inhibitors (ICIs)-based therapies in breast cancer so far.

WHAT THIS STUDY ADDS

⇒ This multicentered study developed the first radiomics model for immunotherapy response prediction specifically for patients with advanced breast cancer. And this non-invasive prediction model performed effectively both in the training and validation cohort. This study presents an efficient, non-invasive, and reliable way to predict patients with advanced breast cancer responses to ICIs-based therapies via machine learning-based radiomics model.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ To our knowledge, this is the first study to apply radiomics to infer the clinical benefit of immunotherapy in patients with advanced breast cancer.

⇒ It confirms the potential of radiomics to identify patients with advanced breast cancer most likely to respond to ICIs-based therapies.

⇒ The accuracy of this radiomics model was not influenced by PD-L1 status, tumor metastatic burden, molecular subtype or combined treatment regimens, and could be widely used to aid breast oncologists in making decisions of ICIs-based therapies for patients with advanced breast cancer.

Conclusions This radiomics model provided an innovative and accurate way that could stratify patients with ABC who may benefit more from ICIs-based therapies.

INTRODUCTION

Among malignancies, breast cancer has been a significant cause of tumor-related deaths around the world.[1] Although the recurrence rate of early-stage breast cancer has gradually decreased with improved treatment, nearly 30% of patients with primary breast cancer still progress to incurable metastatic breast cancer in 10 years, according to our own database.[2]

Immune checkpoint inhibitors (ICIs) (ie, programmed death-ligand 1 (PD-L1) or cytotoxic T-lymphocyte antigen 4) bring new promise in cancer treatment by their ability to improve overall survival among patients, especially those with metastatic or locally advanced breast cancer (ABC).[3–6] Given the near doubling survival benefit of ICIs for some patients with metastatic triple-negative breast cancer (TNBC), ICIs is being widely explored in several clinical studies for different stages of breast cancer, and even is recommended in neoadjuvant setting of high-risk TNBC in order to improve the cure rate. However, it is worth noting that ICIs therapy is only beneficial to a small proportion of patients with breast cancer (20% to 40%), and also has non-negligible toxicity and even a 0.2% to 0.3% risk of treatment-related death. Therefore, it is of great clinical value to precisely target the beneficiary population of ICIs.

Although previous studies suggest that some biomarkers such as PD-L1,[4] tumor mutation burden (TMB),[6] stromal tumor infiltrating lymphocytes (TILs),[7] and microsatellite instability/defective mismatch repair (MSI/dMMR) can predict the population of immunotherapy benefit, invasive biopsies were needed to acquire these biomarkers and the accuracy are not ideal. Hence, it is important to explore innovative methods for more accurate and non-invasive prediction of their efficacy.

One of the non-invasive prediction modalities is to integrate the use of radiomics analysis. Since biomedical images could reflect the cellular and molecular properties of tissues, the radiomics could be used to analyze and translate medical images into quantitative data.[8] The quantitative imaging features could be extracted by machine learning in a high-throughput manner and further adopted to evaluate tumor microenvironment (TME) and heterogeneity.[9] Radiomics-based biomarkers have made achievements in auxiliary diagnosis and prognosis assessment recently, encouraging studies have been reported on the potential utility of radiomics for predicting response to ICIs in lung cancer and melanoma.[10–12] The non-invasive property and reliability of radiomics provide us with an innovative method to predict immunotherapy response. However, no radiomic biomarkers have been reported in predicting the response of ICIs-based therapies in breast cancer so far.

In view of this, this multicentered study was conducted to develop a machine learning based radiomics model that could accurately predict immunotherapy benefit and optimize treatment decisions for patients with ABC.

METHOD

Study design

In this study, a total of 240 patients with ABC who received ICIs-based therapy in three academic medical centers, as Sun Yat-sen Memorial Hospital (SYSMH), Sun Yat-sen University Cancer Center (SYSUCC), and the First Affiliated Hospital of Sun Yat-sen University (FAHSYSU), from February 2018 to January 2022 were retrospectively analyzed. To improve data quality and model robustness, 61 of these patients were recruited from two phase II prospective clinical studies (NCT03394287 and NCT04303741),[13 14] which were designed to explore the efficacy and safety of anti-programmed cell death protein-1 (PD-1) monoclonal camrelizumab plus anti-angiogenic therapy apatinib with or without chemotherapy in advanced TNBC. One study (NCT03394287) investigated the efficacy and safety of chemotherapy-free regimens 'camrelizumab in combination with apatinib' in the front-line (1–3 lines) treatment of advanced TNBC, showing an objective response rate (ORR) of 43.3%. The other study (NCT04303741) investigated the efficacy and safety of eribulin in combination with camrelizumab and apatinib in the second or later line treatment of advanced TNBC, which showed that the combination regimen still achieved an ORR of 37.0% and a median progression-free survival (PFS) of 8.1 m in these heavily pretreated patients with TNBC. The other 179 patients were enrolled from a multicentered retrospective database of patients with ABC receiving anti-PD-1 antibodies in combination with chemotherapy±anti-human epidermal growth factor receptor 2 (HER2) treatment±antiangiogenic therapy. These 240 patients were further divided into a training cohort and an independent validation cohort according to need. In order to better reflect the balance of multi-centered data and the principle of random data assignment, we randomly divided the patients with breast cancer from SYSUCC, the medical center with the largest sample size, into two groups. Then, 50% of the patients from SYSUCC together with the patients from SYSMH, totaling 171, were established as the training cohort, and the other 50% of the patients of SYSMH together with the patients from FAHSYSU, totaling 69, were established as the validation cohort (table 1, figure 1).

The inclusion criteria were (1) patients with pathologically proven ABC, (2) patients who received ICIs-based therapies and (3) patients who underwent CT scan within 1 month before ICIs-based treatment. Exclusion criteria were (1) patient lacks measurable lesions according to RECIST V.1.1 criteria; (2) samples of poor-quality or inadequate, and (3) incomplete data on baseline clinicopathological features or follow-up data (online supplemental appendix S1). This study's primary outcome was response to immunological combination treatment as measured by RECIST V.1.1 criteria,[15] with definition of complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). In this study, patients with PR and CR were categorized as 'responsive',

Table 1 Demographics and clinicopathologic characteristics of the training cohort and the testing cohort

Characteristics	Entire cohort (N=240)	Training cohort (N=171)			Testing cohort (N=69)		
		Non-response (N=110)	Response (N=61)	P value	Non-response (N=45)	Response (N=24)	P value
Age				0.294			0.139
Mean±SD	48.36±10.38	47.56±10.82	49.32±9.82		47.49±9.79	51.29±10.53	
Age group				0.398			0.713
<40	53 (22.1)	27 (24.5)	13 (21.3)		9 (20.0)	4 (16.7)	
40–50	77 (32.1)	37 (33.6)	16 (26.2)		17 (37.8)	7 (29.2)	
≥50	110 (45.8)	46 (41.8)	32 (52.5)		19 (42.2)	13 (54.2)	
BMI (kg/m²)				0.946			0.818
Mean±SD	22.91±3.23	23.01±3.36	22.98±3.36		22.76±3.29	22.59±2.27	
BMI group				0.788			0.569
<18.5	22 (9.2)	10 (9.1)	7 (11.5)		4 (8.9)	1 (4.2)	
18.5–24	133 (55.4)	59 (53.6)	34 (55.7)		24 (53.3)	16 (66.7)	
>24	85 (35.4)	41 (37.3)	20 (32.8)		17 (37.8)	7 (29.2)	
ECOG performance status				0.185			1.000
0–1	218 (90.8)	98 (89.1)	58 (95.1)		40 (88.9)	22 (91.7)	
≥2	22 (9.2)	12 (10.9)	3 (4.9)		5 (11.1)	2 (8.3)	
Menopausal status				0.470			0.285
Premenopause	82 (34.2)	32 (29.1)	21 (34.4)		21 (46.7)	8 (33.3)	
Postmenopause	158 (65.8)	78 (70.9)	40 (65.6)		24 (53.3)	16 (66.7)	
Prior operation				0.051			0.032
Surgery	183 (76.2)	93 (84.5)	44 (72.1)		34 (75.6)	12 (50.0)	
Without surgery	57 (23.8)	17 (15.5)	17 (27.9)		11 (24.4)	12 (50.00)	
Number of lines of prior therapies in the advanced setting				0.020			0.032
1	82 (34.2)	31 (28.2)	28 (45.9)		11 (24.4)	12 (50.0)	
≥2	158 (65.8)	79 (71.8)	33 (54.1)		34 (75.6)	12 (50.0)	
ER status				0.155			0.781
Negative	202 (84.2)	88 (80.0)	54 (88.5)		40 (88.9)	20 (83.3)	
Positive	38 (15.8)	22 (20.0)	7 (11.5)		5 (11.1)	4 (16.7)	
PR status				0.135			1.000
Negative	215 (89.6)	103 (93.6)	53 (86.9)		38 (84.4)	21 (87.5)	
Positive	25 (10.4)	7 (6.4)	8 (13.1)		7 (15.6)	3 (12.5)	
HER2 status				0.468			0.652
Negative	225 (93.8)	102 (92.7)	59 (96.7)		41 (91.1)	23 (95.8)	
Positive	15 (6.2)	8 (7.3)	2 (3.3)		4 (8.9)	1 (4.2)	
Molecular subtype				0.665			0.915
HR+/HER2−	43 (17.9)	18 (16.4)	10 (16.4)		10 (22.2)	5 (20.8)	
HER2+	15 (6.3)	8 (7.3)	2 (3.3)		4 (8.9)	1 (4.2)	
TNBC	182 (75.8)	84 (76.4)	49 (80.3)		31 (68.9)	18 (75.0)	
Molecular subtype group				0.550			0.594
Non-TNBC	58 (24.2)	26 (23.6)	12 (19.7)		14 (31.1)	6 (25.0)	
TNBC	182 (75.8)	84 (76.4)	49 (80.3)		31 (68.9)	18 (75.0)	
Ki67				0.708			1.000
≤20	18 (7.5)	6 (5.5)	5 (8.2)		5 (11.1)	2 (8.3)	
>20	222 (92.5)	104 (94.5)	56 (91.8)		40 (88.9)	22 (91.7)	

Table 1 Continued

Characteristics	Entire cohort (N=240)	Training cohort (N=171)			Testing cohort (N=69)		
		Non-response (N=110)	Response (N=61)	P value	Non-response (N=45)	Response (N=24)	P value
dNLR*				0.172			0.964
<2.38	104 (43.3)	44 (40.0)	31 (50.8)		19 (42.2)	10 (41.7)	
≥2.38	136 (56.7)	66 (60.0)	30 (49.2)		26 (57.8)	14 (58.3)	
LDH				0.197			
≤250	141 (58.8)	61 (55.5)	40 (65.6)		25 (55.6)	15 (62.5)	0.578
>250	99 (41.2)	49 (44.5)	21 (34.4)		20 (44.4)	9 (37.5)	
ALB				0.220			0.881
<40	96 (40.0)	54 (49.1)	24 (39.3)		12 (26.7)	6 (25.0)	
≥40	144 (60.0)	56 (50.9)	37 (60.7)		33 (73.3)	18 (75.0)	
Combined immunotherapy regimen				0.530			0.112
Immunotherapy+chemotherapy	127 (52.9)	56 (50.9)	28 (45.9)		25 (55.6)	18 (75.0)	
Immunotherapy+antiangiogenic therapy±chemotherapy	113 (47.1)	54 (49.1)	33 (54.1)		20 (44.4)	6 (25.0)	
Visceral metastasis				0.016			0.753
No	104 (43.3)	42 (38.2)	35 (57.4)		17 (37.8)	10 (41.7)	
Yes	136 (56.7)	68 (61.8)	26 (42.6)		28 (62.2)	14 (58.3)	
Number of metastatic sites				0.001			0.356
1, 2	122 (50.8)	46 (41.8)	41 (67.2)		21 (46.7)	14 (58.3)	
≥3	118 (49.2)	64 (58.2)	20 (32.8)		24 (53.3)	10 (41.7)	

Data were presented as number of patients; data in parentheses were percentages unless otherwise noted*value refer to the optimal threshold of dNLR using the maximum Youden Index method

ALB, serum albumin; dNLR, derived neutrophil-to-lymphocyte ratio; ECOG, Eastern Cooperative Oncology Group; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; LDH, lactate dehydrogenase; PR, progesterone receptor; TNBC, triple-negative breast cancer.

whereas patients with SD and PD were categorized as 'non-responsive'.
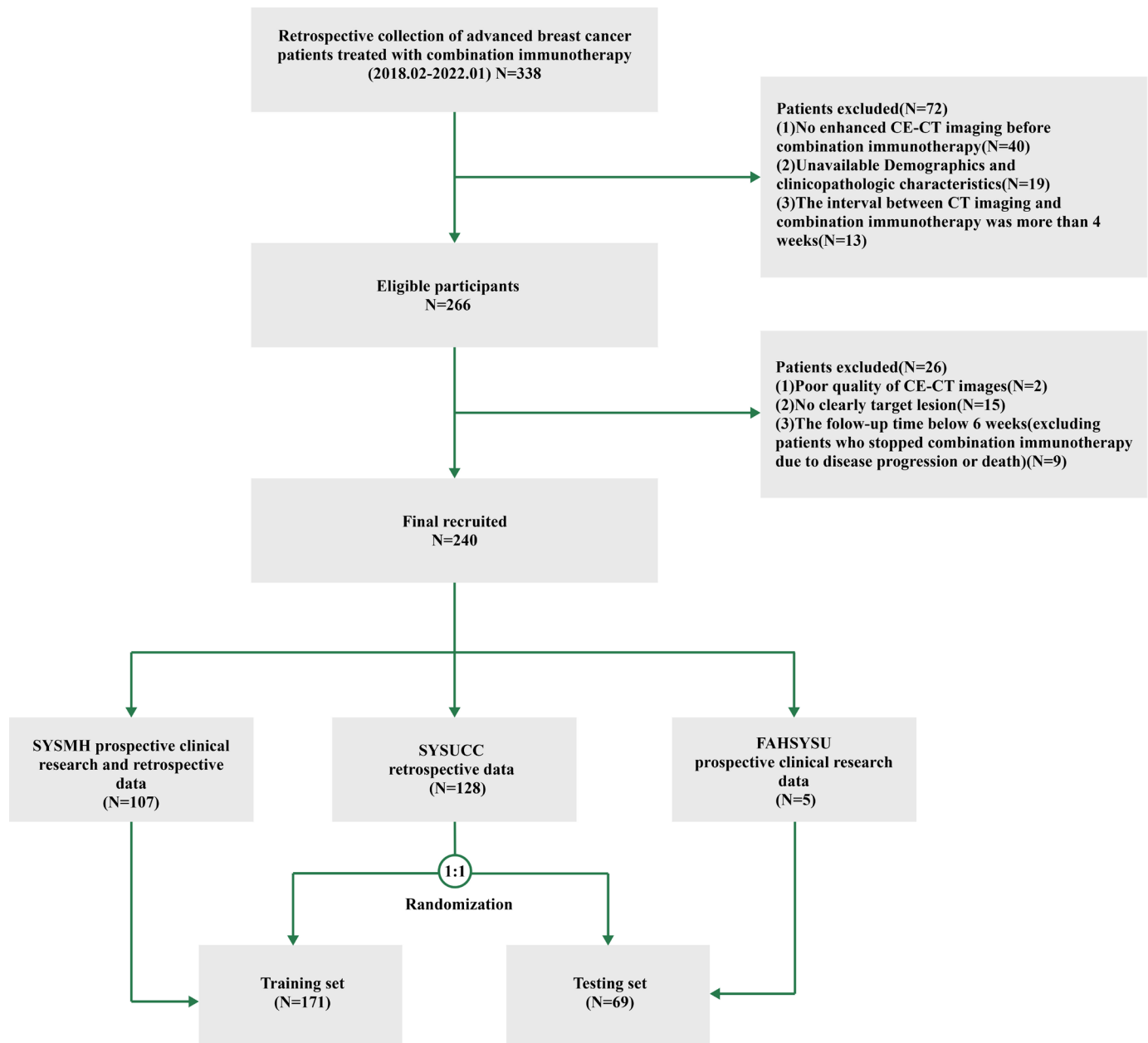
Clinicopathological data

Patients' clinicopathological information was acquired from the hospital information system at each study center. Clinical information includes patients' age, body mass index (BMI), menopausal status, prior treatment history, combination regimen with ICIs, lines of previous therapy in the context of metastatic disease, hematologic indicators, blood biochemical indices, and tumor metastatic burden. Pathological information included tumor pathological type, (estrogen receptor progesterone receptor and HER2 status), Ki67 proliferation index, and PD-L1 expression status.

In this work, PD-L1 expression was assessed using the Combined Positive Score (CPS), which has been used in previous study.[3] According to this protocol, CPS was defined as the number of PD-L1+ cells regardless of cell types (including tumor cells, lymphocytes, and macrophages) divided by the total number of tumor cells, then further multiplied by 100. PD-L1 positivity was defined as ≥10 of CPS. Quantitative and qualitative PD-L1 staining was conducted using the DAKO Link 48 platform and the Food and Drug Administration-approved DAKO 22C3 antibody.

CT imaging data acquisition

Patients underwent contrast-enhanced CT (CECT) examinations within 1 month prior to immunotherapy. The CT images were exported in DICOM format through image archiving and communication system (PACS) retrieval in the radiology department of each center. Patients were scanned with multi-slice spiral CT (GE Medical Systems, Siemens, Philips, Toshiba, United Imaging Healthcare, etc) of the neck, chest and abdomen. The majority of CECT scans tube voltage were 120 KeV, with automatic tube current modulation technique. Reconstruction using standard convolution kernel with 1.25 mm layer thickness (median 1.25 mm; range 1.0–2.0 mm). Each CT image is reconstructed in a 512×512 pixels image matrix. To eliminate image differences between images acquired by different CT instruments, all CT images were first resampled to the same image spacing of 1 mm × 1 mm × 1 mm using cubic spline interpolation to standardize the pixel size.[16] The details information of CT scan parameters is provided in supplementary material (online supplemental table S1).

Retrospective collection of advanced breast cancer patients treated with combination immunotherapy (2018.02-2022.01) N=338

Patients excluded(N=72)
(1)No enhanced CE-CT imaging before combination immunotherapy(N=40)
(2)Unavailable Demographics and clinicopathologic characteristics(N=19)
(3)The interval between CT imaging and combination immunotherapy was more than 4 weeks(N=13)

Eligible participants
N=266

Patients excluded(N=26)
(1)Poor quality of CE-CT images(N=2)
(2)No clearly target lesion(N=15)
(3)The folow-up time below 6 weeks(excluding patients who stopped combination immunotherapy due to disease progression or death)(N=9)

Final recruited
N=240

SYSMH prospective clinical research and retrospective data (N=107)

SYSUCC retrospective data (N=128)

FAHSYSU prospective clinical research data (N=5)

1:1
Randomization

Training set
(N=171)

Testing set
(N=69)

Figure 1 Flow chart of the study. Patients from Sun Yat-sen Memorial Hospital (SYSMH) prospective clinical trials and retrospective clinical trials and half of the patients from Sun Yat-sen University Cancer Center (SYSUCC) as the training set (128 patients from SYSUCC were randomized 1:1 into two groups). The First Affiliated Hospital of Sun Yat-sen University (FAHSYSU) prospective clinical trials and the other half of the SYSUCC retrospective data as the validation set. The ratio of training set and validation set is about 5:2. CECT, contrast-enhanced CT.

Tumor segmentation and image feature data extraction

The lesions of the patients were segmented in accordance with RECIST V.1.1 criteria. We defined criteria for the segmentation of several targets in ABC, selecting two maximum two-dimensional diameter visceral metastases, bone metastases soft tissue lesions or lymph node target lesions. With no more than two target lesions per organ or site, distinct sites are prioritized for tumor segmentation, beginning with visceral metastatic lesions.

Two experienced radiologists (Xiaohui Duan and Zhuo Wu, both of whom are senior doctors in the radiology department with over 10 years of professional experience)

independently viewed and segmented the above image using 3D slicer (V.4.11, https://www.slicer.org). In the event of any inconsistencies, a third senior radiologist consulted with both investigators and made the necessary adjustments.

The CT image features were extracted by radiomics extension module in 3D slicer,[17 18] including four aspects of intensity features, shape features, texture features, and image filtering features, with a total of 1130 features. The initial features consist of 107 original features, 279 LoG (Laplacian of Gaussian) features, and 744 wavelet features. The original features include 14 shape features,

18 histogram features and 75 texture features. Among the texture features, 24 Gray Level Co-occurrence Matrix (GLCM) texture features, 16 Gray Level Run Length Matrix (GLRLM) texture features, 16 Gray Level Size Zone Matrix (GLSZM) texture features, 14 Gray Level Difference Matrix (GLDM) texture features, and 5 Neighborhood Gray-tone Difference Matrix (NGTDM) texture features. LoG features and wavelet features consist of histogram features and texture features. The LoG filter with 1/2/3 kernel sizes is used to extract features, and the configured wavelet filter is used to calculate wavelet features. In the end, eight subwaves were decomposed in each layer, and the wavelet and LoG transforms yielded 744 wavelet features and 279 LoG features, respectively (online supplemental table S2).

Model construction and performance evaluation

Clinicopathological factors with statistical differences (p<0.05) in univariate and multifactorial analyses were used to construct clinicopathological models using logistic regression (LR). Multilayer perceptron (MLP)[19] was used to analyze imaging features using three hidden layer networks to construct imaging histology feature models. In the integrated model construction, the imaging features were imported in input layer 1, and after three layers of network, they were added to the fourth layer of network together with the clinical features in input layer 2 (figure 2). The receiver operating characteristic (ROC) curve, the area under the curve (AUC), positive predictive value and negative predictive value were analyzed to evaluate the accuracy of this radiomics model, and decision curve analysis (DCA) and clinical imaging curve (CIC) are applied to assess the net benefit.

Survival analysis

Risk ratings developed by classification models employing machine learning were used to predict the prognosis of patients with ABC. We based the scores of the clinical model and the deep histological model on log-rank tests to find the optimal cut-off value based on the data on both sides of a point. The optimal cut-off value was used to separate patients into low-risk and high-risk categories. We used Kaplan-Meier (KM) curves to measure PFS, along with the C statistic and risk ratio (HR) to determine the score's impact on PFS.

Statistical analysis

The MLP was built using Python V.3.7.7. Statistical data were analyzed in this study using R software (V.4.2.0, http://www.R-project.org) (R packages is detailed on page 1 of online supplemental file 1) and SPSS V.25.0 software (IBM SPSS V.25.0, Chicago, USA). In the analysis of baseline patient characteristics, independent t-tests and Wilcoxon rank-sum tests were used for continuous variables, and Pearson χ^2 tests or Fisher's exact tests were used for categorical variables. The performance of the categorical model was assessed through AUC calculation; and the corresponding 95% CI under ROC the

differences were compared between AUCs by DeLong test. As for the survival analysis, we used a log-rank test to compare the differences between KM curves. A two-sided p value<0.05 was considered to be statistically significant.

RESULT

Clinical characteristics

In this study, a total of 240 patients (171 patients in the training cohort, 69 patients in the external validation cohort) were retrospectively enrolled (figure 1). There were no statistical significant differences in the baseline characteristics between the training and validation cohort (table 1). Triple-negative disease was present in 75.8% of the enrolled patients. About half of the patients (56.7%) had visceral metastasis or more than three metastatic lesions. All patients underwent anti-PD-1 antibodies-based treatment. Combinatorial therapy of ICIs with chemotherapy were used in 49.1% (84/171) of patients in the training cohort, and 62.3% (43/69) of patients in the validation cohort, respectively. ICIs combined with anti-angiogenesis±chemotherapy were used in 50.9% (87/171) of patients in the training cohort, and 37.7% (26/69) of patients in the validation cohort, respectively. The above variables are equally distributed in the training and validation cohorts, and there is no statistically significant difference between these two cohorts.

Among the 240 patients, 6.7% patients (n=16) showed CR, 28.7% patients (n=69) showed PR, 42.1% patients (n=101) showed SD, and the rest of patients presented PD (n=54, 22.5%) (online supplemental figure S1). The overall disease control rate (DCR) was calculated at 77.5% (186 of 240). Immunotherapy-benefit was associated with lines of previous therapy in the context of metastatic disease (p=0.020), visceral metastasis status (p=0.016) and tumor metastasis burden (p=0.001) in the training cohort. No significant difference was detected in terms of age, BMI, menopausal status, molecular subtype, CPS, and immunotherapy regimen (p>0.05) between patients who responded to immunotherapy or otherwise. Multivariable LR analysis further revealed that lines of previous therapy (OR=0.35, 95% CI: 0.16 to 0.77, p=0.009) and tumor metastasis burden (OR=0.42, 95% CI: 0.18 to 0.99, p=0.047) were both significant predictors of immunotherapy-benefit in the training cohort (online supplemental figure S2). Hence, combined these two factors with visceral metastasis status and molecular subtype, the clinical model was constructed as a baseline. However, the performance of clinical model in predicting immunotherapy response status was unsatisfactory, with AUC values of 0.672 (95% CI: 0.588 to 0.756), and 0.634 (95% CI: 0.495 to 0.772) in the training cohort and validation cohort (figure 3E,F, table 2). Furthermore, the inclusion of PD-L1 status (CPS score) did not significantly improve the accuracy of clinical model (online supplemental figure S3 and table S3).
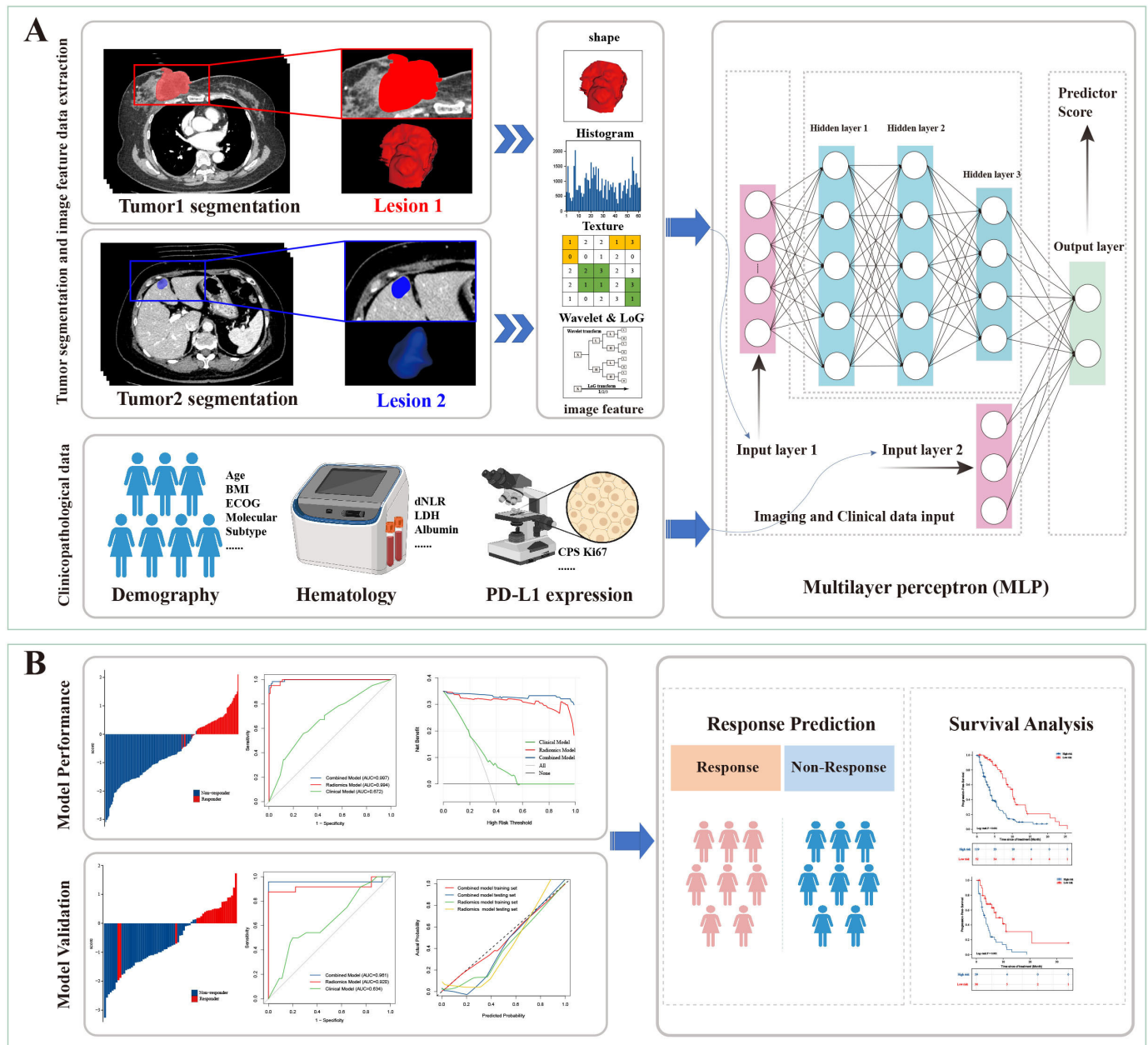
Figure 2 Radiomics workflow (A) The workflow includes data collection, design of the study, ROI delineation, image feature extraction and machine learning model construction. (B) Model performance evaluation and validation. AUC, area under the curve; BMI, body mass index; CPS, Combined Positive Score; ECOG, Eastern Cooperative Oncology Group; dNLR, derived neutrophil-to-lymphocyte ratio; LDH, lactate dehydrogenase; LoG, Laplacian of Gaussian; ROI, region of interest; PD-L1, programmed death-ligand 1.

Selection of radiomics feature and establishment of prediction model

After combining all extracted radiomics features, we used a MLP algorithm to select the features that are most closely correlated with immunotherapy-response in the training cohort. The following nine category of radiomics features were finally selected: Original_Shape, Original_First Order Histogram Features, Original_GLCM, Original_GLDM, Original_GLRLM, Original_GLSZM, Original_NGTDM, Wavelet transform, and LoG transform. The specific image features encapsulated in each category were listed in online supplemental table S2.

Then, the radiomic features were used to build a model to predict the immunotherapy response. The immunotherapy response status of patients with ABC predicted by the radiomics model was in good agreement with the actual clinical immunotherapy response status in both training and validation cohorts (figure 3A,C). Furthermore, the radiomics model also obtained a high accuracy with an AUC of 0.994 (95% CI: 0.988 to 1.000) in the training cohort, and 0.920 (95% CI: 0.824 to 1.000) in the validation cohort, respectively (figure 3E,F, table 2). The calibration curves of the radiomics model showed good agreements between the model prediction and actual
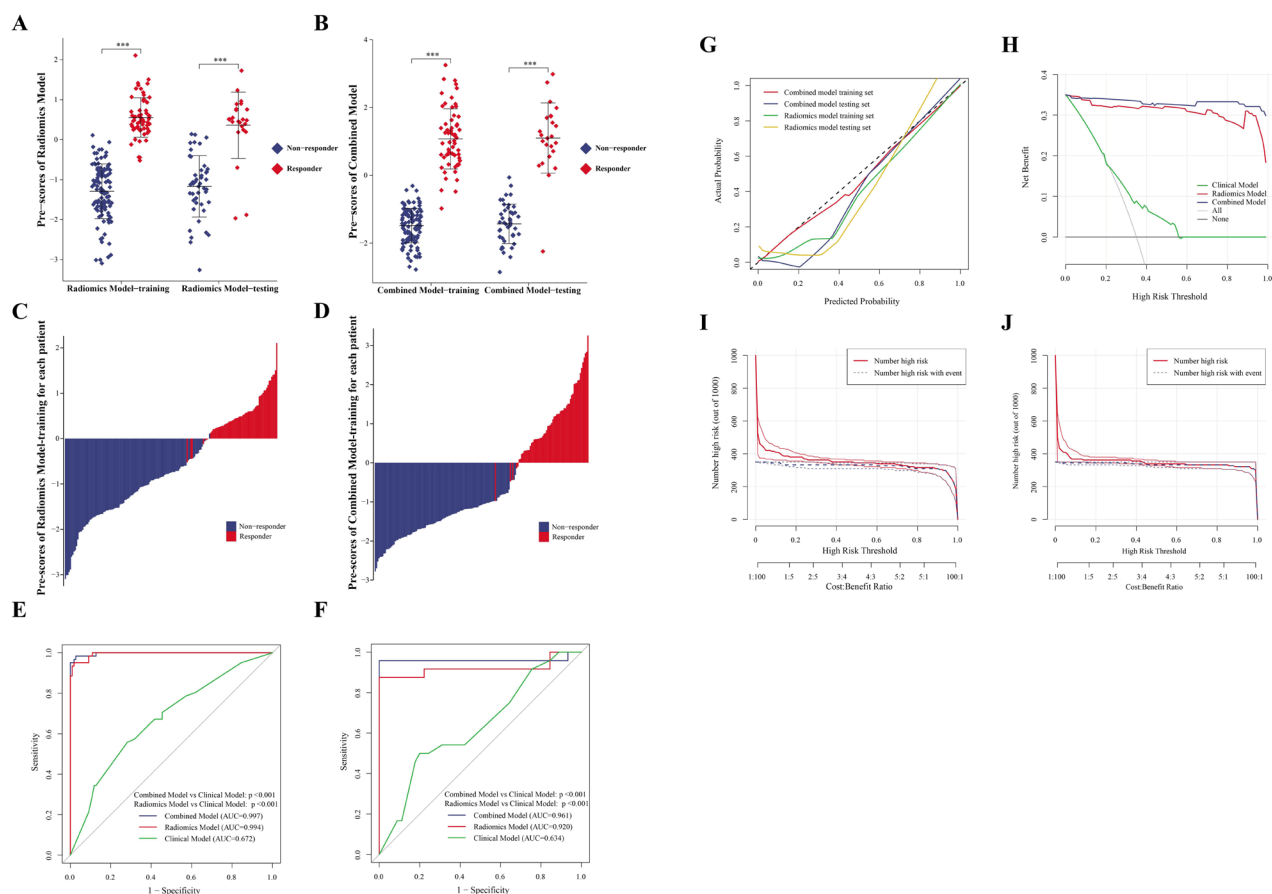
Figure 3 Performance of the predict models in training and validation sets. (A and B) Distribution of predictive score of radiomics and combined models among responders and non-responders in training set and testing set. The *** represents, p<0.001. (C) Waterfall of prediction score distribution between non-responders and responders in the training set of the radiomics model. (D) Waterfall of prediction score distribution between non-responders and responders in the training set of the combined model. (E and F) Receiver operating characteristic analysis of predict models for predicting response status in the training set and validation set, respectively. (G) The calibration curve of radiomics and combined models in the training cohort and validation cohort. (H) Decision curve analysis for the combined model (blue), image model (red) and clinical model (green) in the training set; the y-axis indicates the net benefit; x-axis indicates threshold probability. The gray line represents the assumption that all patients were responders. The black line represents the hypothesis that no patients were responders. (I and J) CIC showed the estimated number of image models that would be declared high risk for each risk threshold and the proportion of true positive patients, I represents the radiomics model training set CIC, J represents the combined model training set CIC. AUC, area under the curve; CIC, clinical imaging curve.

observation in the training cohort and validation cohort, respectively (figure 3G). DCA indicated that the highest overall net benefit was achieved in the radiomics model compared with the clinical model (figure 3H). The CIC showed that the radiomics model could target the immunotherapy response population accurately (figure 3I).

Furthermore, we integrated the clinical features into the radiomics model to develop a clinical-radiomics model in order to obtain better prediction performance. Interestingly, the clinical-radiomics model did not achieve significant improvement compared with the radiomics model, with AUCs of 0.997 (95% CI: 0.993 to 1.000) and 0.961 (95% CI: 0.885 to 1.000) in training and validation sets, respectively (figure 3B, D, E and F, table 2). The calibration curves, DCA and CIC analysis showed similar

agreement between radiomics model and the clinical-radiomics model, as well (figure 3G, H and J).

The prediction accuracy of the radiomics model and the clinical-radiomics model were both better than the clinical model (figure 4A,B, online supplemental table S4). Additionally, 108 of 110 (98.2%) patients with non-response in the training cohort, and 45 of 45 (100%) patients with non-response in the validation cohort were successfully identified by the radiomics model. Meanwhile, 58 of 61 (95.1%) patients with response in training cohort, and 21 of 24 (87.5%) patients with response in validation cohort were successfully identified by the radiomics model (figure 4C, table 2). Moreover, 107 of 110 (97.3%) patients with non-response in the training cohort, and 45 of 45 (100%) patients with non-response

Table 2 Performance of predict models for predicting efficacy of combined immunotherapy in patients with advanced breast cancer

Value	Training set			Testing set		
	Clinical model	Radiomics model	Combined model	Clinical model	Radiomics model	Combined model
AUC (95% CI)	0.672 (0.588 to 0.756)	0.994 (0.988 to 1.000)	0.997 (0.993 to 1.000)	0.634 (0.495 to 0.772)	0.920 (0.824 to 1.000)	0.961 (0.885 to 1.000)
SEN	0.557	0.951	0.984	0.500	0.875	0.958
SPE	0.718	0.982	0.973	0.800	1.000	1.000
ACC	0.661	0.971	0.977	0.696	0.957	0.986
PPV	0.523	0.967	0.952	0.571	1.000	1.000
NPV	0.745	0.973	0.991	0.750	0.938	0.978

ACC, accuracy; AUC, area under the receiver operating curve; NPV, negative predictive value; PPV, positive predictive value; SEN, sensitivity; SPE, specificity.

in the validation cohort were also identified by the clinical-radiomics model successfully. Meanwhile, 60 of 61 (98.4%) patients with response in training cohort, and 23 of 24 (95.8%) patients with response in validation cohort were successfully identified by the clinical-radiomics model (figure 4D, table 2).

The radiomics model was capable of dividing patients into two risk cohorts with significantly different PFS both in training set and validation set, with HR of 2.705 (cut-off point=0.209, 95% CI: 1.888 to 3.876, p<0.001, figure 5A) and 2.625 (cut-off point=−0.424, 95% CI: 1.506 to 4.574, p=0.001, figure 5B), respectively (online supplemental table S5). Moreover, the median PFS in the low-risk group was more than twice as much as that in the high-risk group (10.12 m vs 3.75 m for the training set, and 7.26 m vs 3.02 m for the validation set, (online supplemental table S6). Similarly, the clinical-radiomics model was also capable to distinguish the better PFS for the low-risk patients with an HR of 2.464 (cut-off point=−0.086, 95% CI: 1.720 to 3.529, p<0.001, figure 5C) in the training cohort, and 2.564 (cut-off point=−0.857, 95% CI: 1.469 to 4.475, p=0.001, figure 5D) in the validation cohort, respectively (online supplemental table S5).

Then, we conducted subgroup analyses based on PD-L1 status (CPS score), prior regimens, molecular subtype, tumor metastatic burden and immunotherapy regimens. The radiomics model achieved a good performance both in CPS-high and CPS-low groups, with AUC values of 1.000 (CPS-high) and 0.996 (CPS-low) in the training cohort, and 0.778 (CPS-high) and 1.000 (CPS-low) in the validation cohort (DeLong test p=0.305), respectively (figure 6A, online supplemental table S7). We also conducted a subgroup analysis based on the molecular subtype of these patients with breast cancer. Prediction within TNBC and non-TNBC subgroups all performed well, with AUC values of 0.997 (TNBC), and 0.994 (non-TNBC) in the training cohort, and 0.910 (TNBC), and 0.988 (non-TNBC) in the validation cohort (figure 6B, online supplemental table S8). The performances of the radiomics model in the different subgroups of lines of previous therapy, tumor metastatic burden and combined

immunotherapy regimens, were similar, too (figure 6C–F, online supplemental tables S9–S12). The immunotherapy regimens with or without antiangiogenic therapy did not affect the accuracy of the radiomics model, as well (figure 6F, online supplemental table S12). The performance of the clinical-radiomics model was not influenced by PD-L1 status, tumor metastatic burden or molecular subtype, as well (online supplemental figure S4 and tables S7–S12).

DISCUSSION

Immunotherapy, especially anti-PD-1/PD-L1 antibodies, brings 'breakthrough' improvements in treatments for patients with ABC,[13 14] but only a minority of patients can benefit from immunotherapy. How to accurately and non-invasively locate these immunotherapy-benefit-population is the current leading challenge in the field. In this study, we developed the first radiomics model for immunotherapy response prediction specifically for patients with ABC. And this non-invasive prediction model performed effectively both in the training and validation cohort with AUC of 0.994 and 0.920, respectively. Moreover, the accuracy of the radiomics model was not influenced by PD-L1 status, tumor metastatic burden, molecular subtype or combined regimens. This radiomics model provided an innovative, accurate and robust approach to stratify patients with ABC who may benefit from ICIs-based therapies and aid the personalized decision in the treatment of ABC.

Although previous studies suggest that some biomarkers such as PD-L1, TMB, TILs, and MSI/dMMR can predict response of ICIs-based therapies in ABC,[4–7] all these biomarkers assessment requires multiple tumor sampling, invasive tissue biopsy, high associated costs, and unsatisfactory accuracy, which limit their clinical applications. Therefore, development of a non-invasive prediction model of immunotherapy response is required. In our study, several imaging features showed association with ICIs response of patients with ABC, and these correlations indicated that radiomics features may be used to
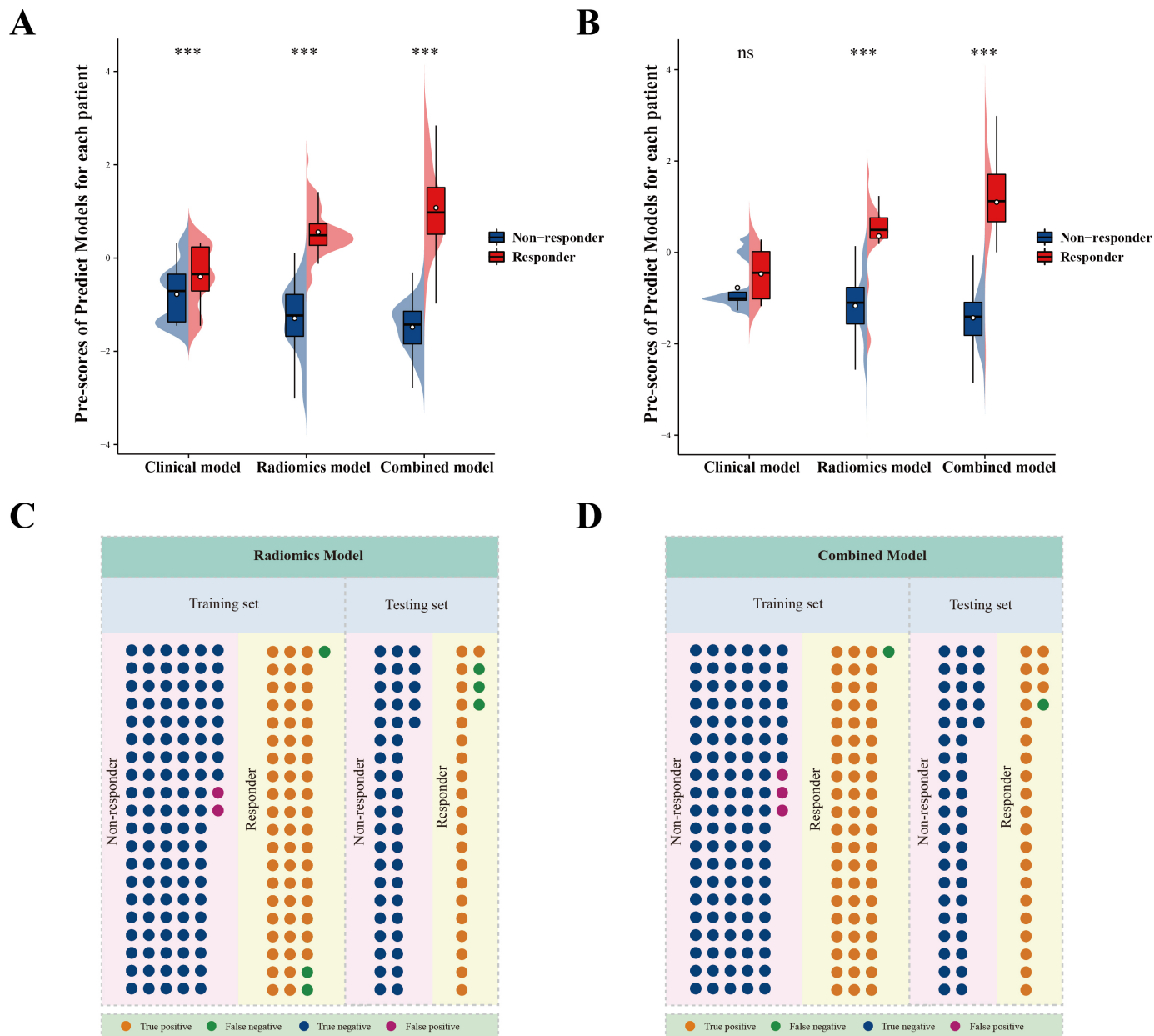
Figure 4 Comparison of prediction capabilities of various models. (A) Violin plot of predict models for non-responders and responders in the training set. The *** represents, p<0.001. (B) Violin plot of predict models for non-responders and responders in the validation set. The ns represents, p>0.05; the *** represents, p<0.001. (C) The number of events of true positive, false negative, true negative, and false positive for the radiomics model in the training cohort and external validation cohorts. (D) The number of events of true positive, false negative, true negative, and false positive for the combined model in the training cohort and external validation cohorts.

develop a non-invasive model to predict immunotherapy benefits in ABC.

To our knowledge, this is the first study to apply radiomics to infer clinical outcomes of patients with ABC with immunotherapy treatment. Previous studies have explored radiomics as a biomarker of immune status, or immunotherapy-prediction value among patients with lung cancer or melanoma, but no patients with breast cancer were included in these studies. Su *et al* established a TILs prediction radiomics model by retrospectively analyzing the preoperative dynamic contrast enhanced MRI (DEC-MRI), transcriptomics data and

postoperative TILs data of 139 patients with TNBC, and found that high Rad-TILs tumors was characterized with enriched immune-related pathways, hot immune microenvironment and relatively inflammatory TME.[20] This is an important work to explore radiomics and immune biomarkers in patients with breast cancer. Unfortunately, patients with breast cancer in this study did not receive immunotherapy, so the immune efficacy of patients could only be inferred indirectly through TILs. Coincidentally, He *et al* also explored associations between radiomic model and TMB, another immune biomarker, through CT images of 327 patients with lung cancer. The radiomic
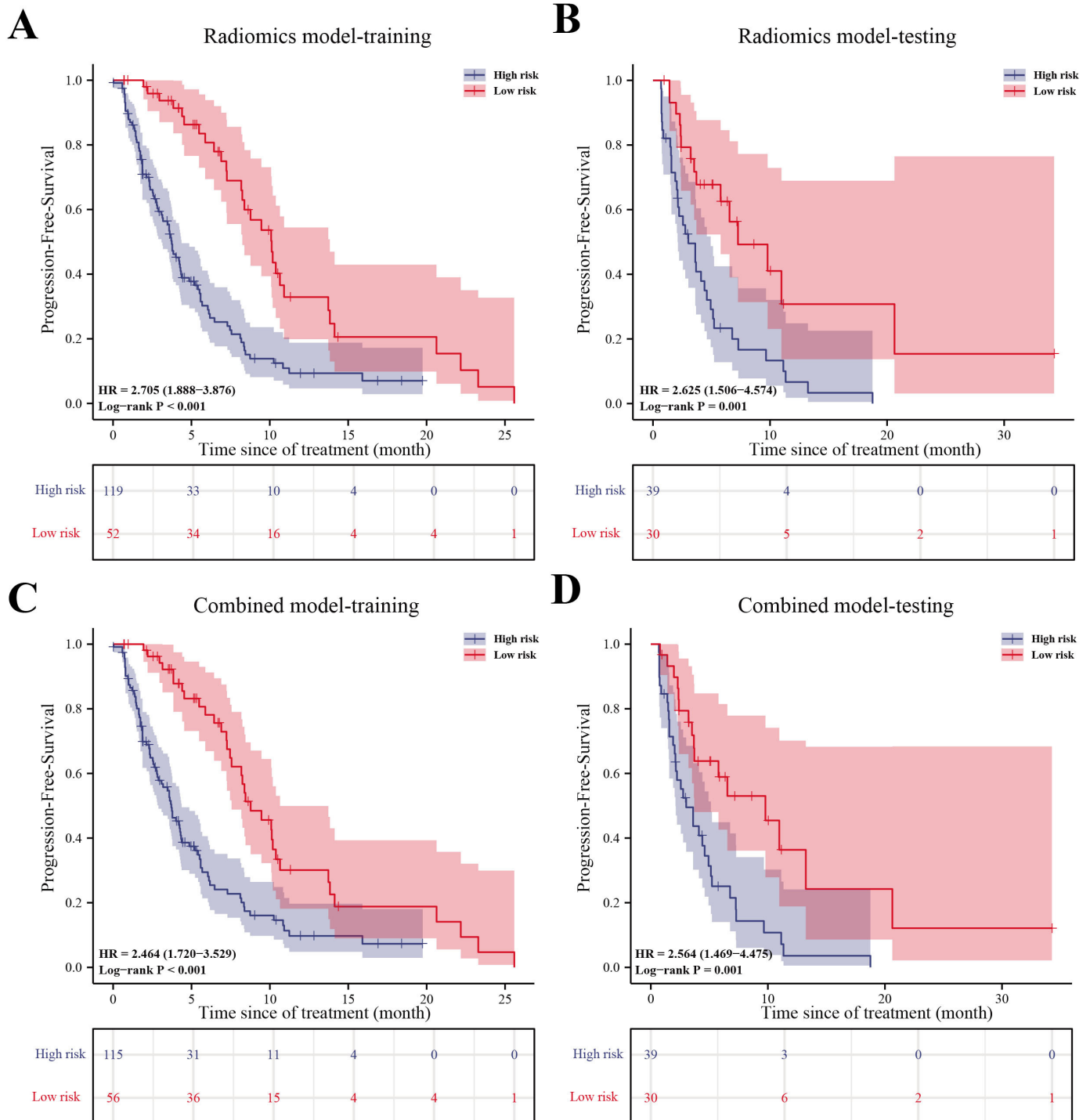
Figure 5 The prognostic value of the radiomics model and combined model in immunotherapy. (A) The Kaplan-Meier curves of the radiomics model depict PFS in high-risk and low-risk groups for training sets. (B)The Kaplan-Meier curves of the radiomics model depict PFS in high-risk and low-risk groups for validation sets. (C) The Kaplan-Meier curves of the combined model depict PFS in high-risk and low-risk groups for training sets. (D) The Kaplan-Meier curves of the combined model depict PFS in high-risk and low-risk groups for validation sets. PFS, progression-free survival.

model was validated in 123 patients with non-small cell lung cancer with ICIs treatment and was able to successfully distinguish high-risk patients from low-risk patients.[21] However, since the radiomics model was established based on TMB, it was not a direct prediction of immunotherapy response, and the results should be interpreted with caution. Sun *et al* developed a radiomics signature

to predict CD8 cell tumor infiltration by CT images and RNA sequencing data of patients with advanced solid tumors. The radiomic signature was further used to predict the immunotherapy response of 137 patients with advanced solid tumors (only 17 patients with breast cancer involved).[22] Notably, the study did not construct a predictive model for immunotherapy response and
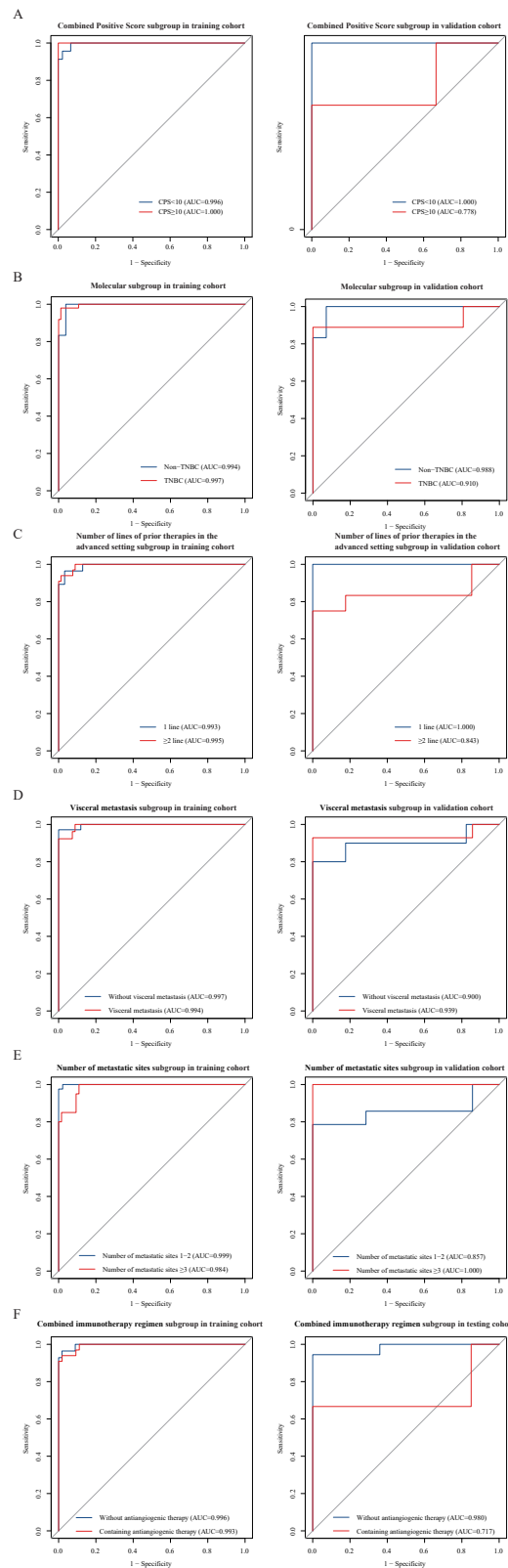
Figure 6 Subgroup analysis of the radiomics model in the training and validation set. (A) AUC of subgroup analysis stratified by comprehensive positive score (CPS <10 vs CPS ≥10) in the training and validation cohorts. (B) AUC of subgroup analysis stratified by molecular subgroups (TNBC vs non-TNBC) in the training and validation cohorts. (C) AUC of subgroup analysis stratified by number of lines of previous therapy in the context of metastatic disease in the training and validation cohorts. (D) AUC of subgroup analysis stratified by visceral metastasis in the training and validation cohorts. (E) AUC of subgroup analysis stratified by number of metastatic sites in the training and validation cohorts. (F) AUC of subgroup analysis stratified by combined immunotherapy regimen (with or without antiangiogenic therapy) in the training and validation cohorts. AUC, area under the curve; CPS, Combined Positive Score; TNBC, triple-negative breast cancer.

included very few patients with breast cancer. As a 'cold' tumor, breast cancer has a lower sensitive population to immunotherapy compared with lung cancer or melanoma. Therefore, the development of a predictive model for immunotherapy response in breast cancer by non-invasive imaging means is of great clinical value.

Regarding the clinical applicability, the radiomics model has the potential to assist the decision of ICIs-based therapies for patients with ABC. In the training cohort, of the 61 patients who were sensitive to immunotherapy, 58 patients (95.1%) were successfully identified by the radiomics model. While, of the 110 patients who were resistant to immunotherapy in this study, 108 patients (98.2%) were accurately identified by the radiomics model. Likewise, in the validation cohort, of the 24 patients who were sensitive to immunotherapy, 21 patients (87.5%) were successfully identified by the radiomics model. While, of the 45 patients who were resistant to immunotherapy in this study, 45 patients (100%) were accurately identified by the radiomics model. Thus, the radiomics model allowed 98.2% to 100% of patients to avoid immunotherapy-related toxicity and 87.5% to 95.1% of patients to receive immunotherapy accurately. This radiomics model may increase the precision of immunotherapy and enhance personalized decision in the treatment of ABC.

Furthermore, our study included the largest population of patients with breast cancer treated with immunotherapy at present. Two hundred and forty patients with ABC from three different academic hospitals were recruited for development and validation of our radiomics model. Among which, 61 patients were enrolled from two prospective clinical trials, which showed promising efficacy of anti-PD-1 antibody in combination with vascular endothelial growth factor receptor 2 (VEGFR2) tyrosine kinase inhibitors (TKI)±eribulin chemotherapy in treating advanced TNBC, regardless of PD-L1 status.[13 14] The characteristics of large sample size, multicentered patient recruitment and partial population from prospective trials greatly improved the data quality and reliability of this study, and are conducive to building a more precise and robust prediction model. The accuracy of the radiomics model was not influenced by PD-L1 status, tumor metastatic burden or molecular subtype. Therefore, this radiomics model could be used as a useful aid to assist clinicians make rational decisions on immunotherapy, so as to improve the benefits of immunotherapy for patients with breast cancer and avoid unnecessary adverse effects.

Our study had several limitations that should be acknowledged. First, the heterogeneity of the data from multicenter, especially for imaging parameters, could affect radiomics features, even though we have made some efforts to weaken this effect (all CT images were first resampled to the same image spacing using cubic spline interpolation to standardize the pixel size to eliminate image differences between images acquired by different CT instruments; MLP was used to analyze imaging features using three hidden layer networks to construct imaging histology feature models). Second, recent studies indicated that several specific gene mutations (such as TP53, MDM2, MDM3, STK11, among others) could affect the efficacy of immunotherapy. Due to lack of genomic sequencing data, we were unable to compare the predictive accuracy of the radiomics model with these immune-associated gene mutation scenarios.

In conclusion, our study suggests that this radiomics model could be an efficient, non-invasive, cost-effective, and reliable way to predict patients with ABC's responses to ICIs-based therapies. Our findings still need to be confirmed by large-scale prospective studies in the future.

Author affiliations
[1]Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China
[2]Breast Tumor Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China
[3]Department of Medical Oncology, Yat-sen Supercomputer Intelligent Medical Joint Research Institute, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China
[4]Faculty of Medicine, Macau University of Science and Technology, Taipa, Macao, China
[5]Department of Medical Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China
[6]Breast Disease Center, Sun Yat-sen University First Affiliated Hospital, Guangzhou, China
[7]Department of Radiology, Sun Yat-Sen Memorial Hospital, Guangzhou, China

ORCID iD
Jieqiong Liu http://orcid.org/0000-0001-7859-3764

REFERENCES

1 Sung H, Ferlay J, Siegel RL, *et al*. Global cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209–49.

2 Chen K, Pan Z, Zhu L, *et al*. Comparison of breast-conserving surgery and mastectomy in early breast cancer using observational data Revisited: A propensity score-matched analysis. *Sci China Life Sci* 2018;61:1528–36.

3 Schmid P, Cortes J, Pusztai L, *et al*. Pembrolizumab for early triple-negative breast cancer. *N Engl J Med* 2020;382:810–21.

4 Cortes J, Rugo HS, Cescon DW, *et al*. Pembrolizumab plus chemotherapy in advanced triple-negative breast cancer. *N Engl J Med* 2022;387:217–26.

5 Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med* 2017;377:2500–1.

6 Alva AS, Mangat PK, Garrett-Mayer E, *et al*. Pembrolizumab in patients with metastatic breast cancer with high tumor mutational burden: Results from the targeted agent and profiling utilization Registry (TAPUR) study. *JCO* 2021;39:2443–51.

7 Loibl S, Untch M, Burchardi N, *et al*. A randomised phase II study investigating Durvalumab in addition to an anthracycline Taxane-based Neoadjuvant therapy in early triple-negative breast cancer: Clinical results and biomarker analysis of Geparnuevo study. *Annals of Oncology* 2019;30:1279–88.

8 Gillies RJ, Kinahan PE, Hricak H. Hricak H: Radiomics: Images are more than pictures, they are data. *Radiology* 2016;278:563–77.

9 Lambin P, Leijenaar RTH, Deist TM, *et al*. Radiomics: The bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14:749–62.

10 Trebeschi S, Drago SG, Birkbak NJ, *et al*. Predicting response to cancer Immunotherapy using noninvasive Radiomic biomarkers. *Ann Oncol* 2019;30:998–1004.

11 Tunali I, Gray JE, Qi J, *et al*. Novel clinical and Radiomic predictors of rapid disease progression phenotypes among lung cancer patients treated with Immunotherapy: An early report. *Lung Cancer* 2019;129:75–9.

12 Mu W, Tunali I, Gray JE, *et al*. Radiomics of (18)F-FDG PET/CT images predicts clinical benefit of advanced NSCLC patients to Checkpoint blockade Immunotherapy. *Eur J Nucl Med Mol Imaging* 2020;47:1168–82.

13 Liu J, Liu Q, Li Y, *et al*. Efficacy and safety of Camrelizumab combined with Apatinib in advanced triple-negative breast cancer: An open-label phase II trial. *J Immunother Cancer* 2020;8:e000696.

14 Liu J, Wang Y, Tian Z, *et al*. Multicenter phase II trial of Camrelizumab combined with Apatinib and Eribulin in heavily pretreated patients with advanced triple-negative breast cancer. *Nat Commun* 2022;13.

15 Eisenhauer EA, Therasse P, Bogaerts J, *et al*. New response evaluation criteria in solid tumours. *Eur J Cancer* 2009;45:228–47.

16 Gong J, Bao X, Wang T, *et al*. A short-term follow-up CT based Radiomics approach to predict response to Immunotherapy in advanced non-small-cell lung cancer. *Oncoimmunology* 2022;11:2028962.

17 van Griethuysen JJM, Fedorov A, Parmar C, *et al*. Computational Radiomics system to Decode the radiographic phenotype. *Cancer Res* 2017;77:e104–7.

18 Zwanenburg A, Vallières M, Abdalah MA, *et al*. The image biomarker standardization initiative: Standardized quantitative Radiomics for high-throughput image-based Phenotyping. *Radiology* 2020;295:328–38.

19 Tolstikhin I, Houlsby N, Kolesnikov A, *et al*. MLP-mixer: An all-MLP architecture for vision. 2021.

20 Su GH, Xiao Y, Jiang L, *et al*. Radiomics features for assessing tumor-infiltrating lymphocytes correlate with molecular traits of triple-negative breast cancer. *J Transl Med* 2022;20:471.

21 He B, Dong D, She Y, *et al*. Predicting response to Immunotherapy in advanced non-small-cell lung cancer using tumor mutational burden Radiomic biomarker. *J Immunother Cancer* 2020;8:e000550.

22 Sun R, Limkin EJ, Vakalopoulou M, *et al*. A Radiomics approach to assess tumour-infiltrating Cd8 cells and response to anti-PD-1 or anti-PD-L1 Immunotherapy: An imaging biomarker, retrospective Multicohort study. *Lancet Oncol* 2018;19:1180–91.