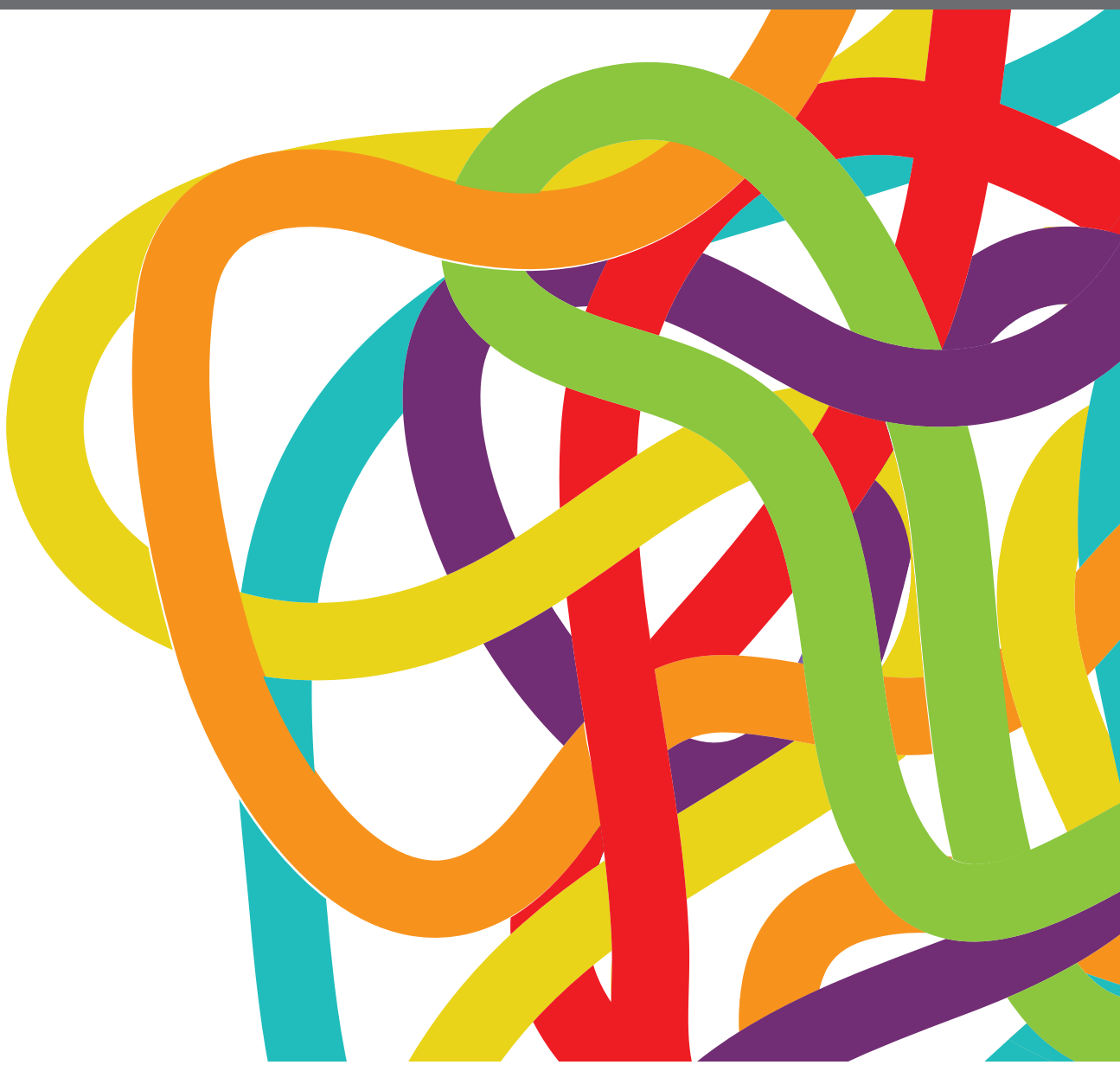
INSIGHTS IN BREAST CANCER: 2021

EDITED BY: Kara Britt and Paula R. Pohlmann
PUBLISHED IN: Frontiers in Oncology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-376-4
DOI 10.3389/978-2-88976-376-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

INSIGHTS IN BREAST CANCER: 2021

Topic Editors:
Kara Britt, Peter MacCallum Cancer Centre, Australia
Paula R. Pohlmann, University of Texas MD Anderson Cancer Center, United States

Citation: Britt, K., Pohlmann, P. R., eds. (2022). Insights in Breast Cancer: 2021. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-376-4

Table of Contents

frontiers in Oncology

ORIGINAL RESEARCH
published: 01 July 2021
doi: 10.3389/fonc.2021.675955

Factors Predicting Locoregional Recurrence After Neoadjuvant Chemotherapy and Nipple-Sparing/Skin-Sparing Mastectomy With Immediate Breast Reconstruction

*Zhen-Yu Wu[1,2], Hee Jeong Kim[2], Jong Won Lee[2], Il Yong Chung[2], Jisun Kim[2], Sae Byul Lee[2], Byung-Ho Son[2], Jin Sup Eom[3], Jae Ho Jeong[4], Gyungyub Gong[5], Hak Hee Kim[6], Sei-Hyun Ahn[2] and BeomSeok Ko[2]**

[1] Department of Breast Surgery, Shanghai East Hospital, Tongji University School of Medicine, Shanghai, China, [2] Division of Breast Surgery, Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea, [3] Department of Plastic Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea, [4] Department of Oncology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea, [5] Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea, [6] Department of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea

Background: Few data are available on the risk factors of locoregional recurrence (LRR) after neoadjuvant chemotherapy (NACT) and immediate breast reconstruction (IBR) in breast cancer. Herein, we evaluated the factors predicting LRR in a large series of patients who underwent either nipple- (NSM) or skin-sparing mastectomy (SSM) with IBR after NACT.

Methods: We retrospectively analyzed 609 breast cancer patients who underwent NACT and NSM/SSM with IBR between February 2010 and June 2017. Factors associated with an increased risk of LRR were analyzed by univariate (chi-square or Fisher's exact test) and multivariate (Cox proportional hazard regression model) analyses.

Results: During a median follow-up of 63 months, LRR as the first event occurred in 73 patients, and the 5-year cumulative LRR rate was 10.8%. Multivariate analysis revealed post-NACT Ki67 ≥ 10% [hazard ratio (HR), 2.208; 95% confidence interval (CI), 1.295-3.765; $P = 0.004$], high tumor grade (HR, 1.738; 95% CI, 1.038-2.908; $P = 0.035$), and presence of lymphovascular invasion (LVI) (HR, 1.725; 95% CI, 1.039-2.864; $P = 0.035$) as independently associated with increased LRR risk. The 10-year LRR rate was 8.5% for patients with none of the three associated risk factors, 11.6% with one factor, 25.1% with two factors, and 33.7% with all three factors ($P < 0.001$).

OPEN ACCESS

Edited by:
San-Gang Wu,
First Affiliated Hospital of Xiamen University, China

Reviewed by:
Gilles Houvenaeghel,
Institut Paoli-Calmettes (IPC), France
Paolo Veronesi,
European Institute of Oncology (IEO), Italy

***Correspondence:**
BeomSeok Ko
spdoctorko@gmail.com

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: 04 March 2021
Accepted: 18 June 2021
Published: 01 July 2021

Citation:
Wu Z-Y, Kim HJ, Lee JW, Chung IY, Kim J, Lee SB, Son B-H, Eom JS, Jeong JH, Gong G, Kim HH, Ahn S-H and Ko B (2021) Factors Predicting Locoregional Recurrence After Neoadjuvant Chemotherapy and Nipple-Sparing/Skin-Sparing Mastectomy With Immediate Breast Reconstruction. Front. Oncol. 11:675955. doi: 10.3389/fonc.2021.675955

Conclusions: Post-NACT Ki67 ≥ 10%, high tumor grade, and presence of LVI are independently associated with an increased risk of developing LRR after NACT and NSM/SSM with IBR. Future prospective trials are warranted to decrease the risk of LRR in patients with associated risk factors.

Keywords: breast cancer, immediate breast reconstruction, skin-sparing mastectomy, nipple-sparing mastectomy, neoadjuvant chemotherapy, locoregional recurrence, risk factor

INTRODUCTION

Neoadjuvant chemotherapy (NACT) has been established as the standard of care for locally advanced breast cancer and is now being used more often as a treatment in early-stage breast cancer (1). NACT aims to increase the rate of breast conservation; however, a large proportion of patients receiving NACT undergo mastectomy as the surgical treatment, either because breast-conserving surgery is not feasible or because of patient preference. Over the last decade, patients have begun to prefer nipple- (NSM) or skin-sparing mastectomy (SSM) combined with immediate breast reconstruction (IBR) in the treatment of breast cancer, as it provides improved aesthetic results and quality of life (2, 3). Several non-randomized studies have demonstrated that the oncologic outcomes of NSM/SSM with IBR are comparable to those of conventional mastectomy alone (4–6). Recently, NSM/SSM with IBR has also been performed in patients who receive NACT; however, data related to the long-term safety of such treatments in this patient population are still insufficient (7). In addition, locoregional recurrence (LRR) following NSM/SSM with IBR remains clinically challenging, not only because it may indicate poor prognosis (8), but also because the oncologic management of LRR may lead to loss of the initial reconstruction (9). In patients who receive NACT and breast reconstruction, the predictive value of clinicopathologic features or treatment-associated factors for LRR is unclear due to a lack of data.

In this study, we aimed to identify the factors associated with an increased risk of LRR in a large series of breast cancer patients who underwent NSM/SSM with IBR after NACT.

MATERIALS AND METHODS

This study was approved by the institutional review board (IRB) of Asan Medical Center, Seoul, Republic of Korea (No. 2017-1341). This study is a retrospective study conducted with the exemption of consent under IRB deliberation using a platform for extracting unidentified clinical information for research purposes. The medical records of all patients who underwent IBR with NSM/SSM after NACT for primary breast cancer between January 2010 and June 2017 at the Asan Medical Center, Seoul, Republic of Korea, were reviewed from a prospectively maintained database. Patients presenting with inflammatory breast cancer or synchronous distant metastasis were excluded. Patient and tumor characteristics were collected and analyzed, including age at diagnosis, tumor stage, grade, molecular subtype, histotype, lymphovascular invasion (LVI) status, presence of extensive intraductal component, post-NACT Ki67 status, and pathological multifocality/multicentricity. Tumor staging was conducted according to the 8th American Joint Committee on Cancer Staging Manual (10). Pathological complete response (pCR) was defined as no evidence of invasive cancer in the breast or axillary lymph nodes.

All patients included in this study received NACT after breast cancer diagnosis. The NACT regimens were selected at the discretion of the treating oncologist. NSM/SSM was performed by breast surgeons, and IBR was performed by plastic surgeons using autologous flaps or implants. NSM or SSM was performed according to the indications of conventional mastectomy, regardless of tumor size or tumor-to-nipple distance, as long as there was no evidence of tumor involvement in the breast skin and nipple-areola complex, clinically or on imaging. In cases of NSM, retroareolar frozen-section biopsy specimens were collected and examined intraoperatively. The nipple-areola complex was preserved if the shape, color, and palpated features of the nipple were normal, and if the nipple margin was confirmed to be tumor free on frozen-section biopsy. In cases in which the retroareolar tissue was positive for malignancy in the frozen section or permanent biopsy, the nipple with or without the areola was removed, and these cases were considered SSM. The decision to undergo adjuvant radiotherapy was made by the treating radiation oncologist after consideration of pre- and post-NACT disease stages, tumor response to NACT, and other tumor biomarkers in patients. Most patients who required adjuvant radiotherapy after evaluation underwent simultaneous irradiation of the chest wall and supraclavicular region. Adjuvant hormonal therapy was applied in patients with hormone receptor-positive disease.

Postoperatively, patients were regularly followed up every 3–6 months for the first 5 years and annually thereafter. Recurrence and metastasis were identified based on the results of the clinical examination, chest radiography, and tumor marker (CA15–3) measurements, which were taken every follow-up visit. In some cases, abnormal clinical findings were further evaluated using chest computed tomography (CT), a bone scan, ultrasonography, and/or positron emission tomography-CT. In patients suspected of LRR, fine needle aspiration, core needle, or excisional biopsy was performed for pathological confirmation. Lesions with clear evidence of distant metastasis on imaging evaluation were considered as recurrence without pathological examination.

LRRs were classified as local or regional recurrence. Local recurrence was defined as biopsy-proven recurrences in the ipsilateral skin/subcutaneous layer, chest wall, or nipple-areola

TABLE 1 | Patient, tumor, and treatment characteristics (N=609).

Characteristics		N	%
Age at diagnosis, years	Median	42 (23–72)	
	≤40	254	41.7
	>40	355	58.3
Clinical T stage	cT1	38	6.2
	cT2	355	58.3
	cT3-4	216	35.5
Clinical N stage	cN0	212	34.8
	cN1	313	51.4
	cN2-3	84	13.8
Pathological T stage	ypT0/ypTis	87	14.3
	ypT1	217	35.6
	ypT2	220	36.1
	ypT3	85	14.0
Pathological N stage	ypN0	287	47.1
	ypN1	221	36.3
	ypN2-3	101	16.6
Molecular subtype	HR+/HER2-	323	53.0
	HR+/HER2+	159	26.1
	HR-/HER2+	64	10.5
	TN	63	10.3
pCR	Yes	79	13.0
	No	530	87.0
Pathological MF/MC	Yes	206	33.8
	No	403	66.2
Histotype	Ductal	533	87.5
	Lobular	26	4.3
	Mixed/Others	50	8.2
Tumor grade	1	15	2.5
	2	436	71.6
	3	158	25.9
LVI	Yes	227	37.3
	No	382	62.7
Extensive intraductal component	Yes	170	27.9
	No	439	72.1
Post-NACT Ki67	<10%	281	46.1
	≥10%	255	41.9
	Unknown	73	12.0
NACT regimens	AC/AC+T	546	89.7
	T	51	8.4
	Others	12	2.0
Mastectomy type	NSM	370	60.8
	SSM	239	39.2
Axillary surgery	SLNB alone	359	58.9
	ALND	250	41.1
Adjuvant radiotherapy	Yes	316	51.9
	No	293	48.1
Adjuvant hormonal therapy	Yes	482	79.1
	No	127	20.9
Adjuvant chemotherapy	Yes	70	11.5
	No	539	88.5
Trastuzumab in HER2+	Yes	219	98.2
	No	4	1.8
Reconstructive surgery	Autologous flaps	420	69.0
	Implants	189	31.0

AC, anthracycline; ALND, axillary lymph node dissection; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; LVI, lymphovascular invasion; MF/MC, multifocality/multicentricity; NACT, neoadjuvant chemotherapy; NSM, nipple-sparing mastectomy; pCR, pathological complete response; SLNB, sentinel lymph node biopsy; SSM, skin-sparing mastectomy; T, taxane; TN, triple negative.

complex, and regional recurrence was defined as carcinoma metastases in the ipsilateral axillary, supraclavicular, or internal mammary lymph node. Any other site of recurrence was considered distant metastasis. Patients with initial distant metastasis were excluded from the LRR group. In cases of concurrent LRR and distant metastasis, each recurrence was counted as an event. Occurrence of contralateral breast cancer was considered a new primary cancer and was not counted as a recurrence. Follow-up was calculated from the date of diagnosis.

The 5- and 10-year cumulative LRR rates were calculated using the Kaplan-Meier method and compared using the log-rank test between subgroups. The clinicopathological factors that were significant in univariate analyses (Chi-square or Fisher's exact test) of LRR were included in the multivariate analysis using the Cox proportional hazards regression model. All statistical analyses were performed using IBM SPSS Statistics software version 24.0 for Windows (IBM Corp., Armonk, NY, USA). Two-tailed *P*-values < 0.05 were considered significant.

RESULTS

A total of 609 patients who underwent NACT and IBR with NSM/SSM for primary breast cancer were included. Patient, tumor, and treatment characteristics are shown in **Table 1**.

The median age at diagnosis was 42 years (range, 23-72 years). The majority (89.7%) of patients received anthracycline-based (with or without taxane) NACT. NSM was performed in 370 (60.8%) patients and SSM in 239 (39.2%). Four hundred and twenty (69%) patients underwent autologous flap reconstruction, and 189 (31%) patients underwent implant-based reconstruction. Adjuvant radiotherapy was administrated in 316 (51.9%) patients. Among the 223 patients with human epidermal growth factor receptor 2 (HER2)-positive disease, 219 (98.2%) received adjuvant trastuzumab. On follow-up, pCR was observed in 79 (13%) patients.

The median follow-up period was 63 months (range, 11-135 months). LRR as the first event occurred in 73 patients, and the 5-year cumulative LRR rate was 10.8%. Among these, isolated LRR occurred in 55 patients (75.3%) and concurrent LRR with distant metastasis occurred in 18 (24.7%). **Table 2** summarizes the oncologic outcomes of the entire cohort. The median time to LRR was 35 months (range, 7-76 months). Patients with isolated LRR as the first event showed a significantly lower 10-year overall survival rate than those without LRR (64.7% *vs.* 90.2%; log-rank $P = 0.035$). **Table 3** shows the incidence rates of LRR according to various clinicopathological and treatment factors. The following factors were significantly associated with increased rates of LRR in the univariate analysis: age at diagnosis ≤ 40 years, pathological T stage, pathological nodal status, pCR status, tumor grade, LVI, and post-NACT Ki67 status. Of these, post-NACT Ki67 ≥ 10% [hazard ratio (HR), 2.208; 95% confidence interval (CI), 1.295-3.765; $P = 0.004$], high tumor grade (HR, 1.738; 95% CI, 1.038-2.908; $P = 0.035$), and presence of LVI (HR, 1.725; 95% CI, 1.039-2.864; $P = 0.035$) were independently associated with reduced LRR-free survival in the multivariate analysis (**Table 4**).

Figure 1 shows the Kaplan-Meier curves for LRR risk, according to the number of independent risk factors. The 10-year rate of LRR was 8.5% for patients with none of the three independent risk factors (n = 197, 32.3%), 11.6% for those with one risk factor (n = 226, 37.1%), 25.1% for those with two risk factors (n = 144, 23.6%), and 33.7% for those with all three risk factors (n = 42, 6.9%; log-rank $P < 0.001$).

TABLE 2 | Oncologic outcomes.

	N	%
Locoregional recurrence	73[a]	12
Skin/chest wall	27	4.4
Nipple-areola complex	7	1.9[b]
Regional lymph nodes	45	7.4
Distant metastasis	99	16.3
Any first recurrence	138	22.7
Death	57	9.4
5-y locoregional recurrence-free survival		87.6
5-y disease-free survival		77.5
5-y distant metastasis-free survival		83.6
5-y overall survival		92.3

[a]*Including 5 cases of concurrent local and regional recurrence without distant metastasis, and 18 cases of concurrent local and/or regional recurrence with distant metastasis as the first event.*

[b]*Calculated in 370 cases of nipple-sparing mastectomy.*

DISCUSSION

Although previous studies have investigated predictive factors of LRR after NACT in conventional mastectomy or breast-conserving surgery (11–15), little data regarding the risk factors of LRR after NACT for NSM/SSM with IBR exists. In this study, we identified the 5-year LRR rate (10.8%) and factors predicting LRR in breast cancer patients who underwent NSM/SSM with IBR after receiving NACT. Post-NACT Ki67 ≥ 10%, high tumor grade, and presence of LVI were independent risk factors for LRR in the current setting. Notably, the 10-year LRR rate reached 33.7% in patients with all three risk factors and was 8.5% in patients with none of these factors.

NSM/SSM with IBR has become an important surgical strategy in modern breast cancer care. This surgical procedure, particularly NSM with IBR, can provide significantly improved aesthetic results, patient satisfaction, and/or psychosocial/sexual well-being (2, 3, 16). A recent analysis from the National Cancer Database of the American College of Surgeons and the American Cancer Society showed an increasing trend toward the application of NSM in patients with advanced disease, particularly in those who received NACT, and highlighted the importance of further prospective trials to validate the evidence of oncologic safety of this procedure (7). The current National Comprehensive Cancer Network (NCCN) guidelines recommend that NSM/SSM should be performed by an experienced breast surgery team working in a multidisciplinary fashion, according to specific clinical features and selected criteria (17). In case of NSM, NCCN guidelines include some cases of locally advanced invasive breast cancers, provided there is complete clinical response after NACT and no nipple involvement. Furthermore, assessment of nipple margin during surgery is mandatory (17). Several studies have reported on the feasibility of this approach in patients who receive NACT, and the LRR rates ranged between 3.2% and 10.3% (18–22). However, the majority of the studies involved a relatively small sample size and short follow-up durations. In the current study, with a median follow-up of 63 months, we found a 5-year cumulative LRR rate of 10.8% for the entire cohort. The LRR rate of our cohort appears acceptable in consideration of the

TABLE 3 | Univariate analysis of factors associated with LRR.

Variables		N	LRR rate, %	*P*-value
		73	12.0	
Age at diagnosis, years	≤40	39	15.4	0.030
	>40	34	9.6	
Clinical T stage	cT1	4	10.5	0.530
	cT2	47	13.2	
	cT3-4	22	10.2	
Clinical N stage	cN0	22	10.4	0.669
	cN1	40	12.8	
	cN2-3	11	13.1	
Pathological T stage	ypT0/ypTis	4	4.6	0.045
	ypT1	33	15.2	
	ypT2	29	13.2	
	ypT3	7	8.2	
Pathological nodal status	ypN-	25	8.7	0.019
	ypN+	48	14.9	
Molecular subtype	HR+/HER2-	35	10.8	0.362
	HR+/HER2+	17	10.7	
	HR-/HER2+	11	17.2	
	TN	10	15.9	
pCR	Yes	3	3.8	0.015
	No	70	13.2	
Pathological MF/MC	Yes	30	14.6	0.162
	No	43	10.7	
Tumor grade	1, 2	45	10.0	0.010
	3	28	17.7	
LVI	Yes	40	17.6	0.001
	No	33	8.6	
Extensive intraductal component	Yes	22	12.9	0.652
	No	51	11.6	
Post-NACT Ki67	<10%	24	8.5	0.001
	≥10%	47	18.4	
	Unknown	2	NA	
Mastectomy type	NSM	46	12.4	0.674
	SSM	27	11.3	
Axillary surgery	SLNB alone	37	10.3	0.126
	ALND	36	14.4	
Adjuvant radiotherapy	Yes	31	9.8	0.086
	No	42	14.3	
Adjuvant hormonal therapy	Yes	52	10.8	0.076
	No	21	16.5	
Adjuvant chemotherapy	Yes	10	14.3	0.529
	No	63	11.7	
Trastuzumab in HER2+	Yes	28	12.6	1.000
	No	0	0.0	
Reconstructive surgery	Autologous flaps	50	11.9	0.926
	Implants	23	12.2	

ALND, axillary lymph node dissection; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; LRR, locoregional recurrence; LVI, lymphovascular invasion; MF/MC, multifocality/multicentricity; NA, not applicable; NACT, neoadjuvant chemotherapy; NSM, nipple-sparing mastectomy; pCR, pathological complete response; SLNB, sentinel lymph node biopsy; SSM, skin-sparing mastectomy; TN, triple negative.

previously reported LRR rates, which ranged from 6.0% to 21.0% after NACT and mastectomy with or without reconstruction (14, 23–26).

The occurrence of breast cancer LRR is an important determinant of adverse survival outcomes (8, 27–29). In our study, isolated LRR as the first event in patients who underwent NSM/SSM with IBR after NACT was associated with a poor 10-year overall survival rate. In addition, patients with isolated LRR often required oncologic management, including surgical excision of the recurrent tumor, which could result in loss of the initial reconstruction (9). Therefore, identifying risk factors for LRR in the current setting is imperative for optimal locoregional management and patient surveillance strategies. However, investigating risk factors for recurrence after NACT remains a challenge because of the high frequency of inconsistent disease status in patients between before and after neoadjuvant treatment. Previous studies have described several clinical and pathological factors of LRR after NACT. The National Surgical Adjuvant Breast and Bowel Project (NSABP) study, including NSABP B-18 and NSABP B-27 data, identified that young age (< 50 years), clinical tumor size (> 5 cm), clinical node status (cN+), and pCR status (ypT+ or ypN+) were predictive of an increased risk of LRR after NACT in patients who underwent mastectomy and breast conservation therapy (11). The authors developed a

TABLE 4 | Multivariate analysis of risk factors associated with LRR.

Variables		HR	95% CI	*P*-value
Age at diagnosis, years	>40	1 (reference)		0.427
	≤40	1.214	0.752-1.959	
Pathological T stage	ypT0/ypTis	1 (reference)		
	ypT1	1.888	0.367-9.708	0.447
	ypT2	2.017	0.867-4.692	0.103
	ypT3	1.857	0.806-4.278	0.146
Pathological nodal status	ypN-	1 (reference)		0.097
	ypN+	1.589	0.920-2.745	
pCR	Yes	1 (reference)		0.452
	No	2.971	0.174-50.602	
Tumor grade	1, 2	1 (reference)		0.035
	3	1.738	1.038-2.908	
LVI	No	1 (reference)		0.035
	Yes	1.725	1.039-2.864	
Post-NACT Ki67	<10%	1 (reference)		0.004
	≥10%	2.208	1.295-3.765	

CI, confidence interval; HR, hazard ratio; LRR, locoregional recurrence; LVI, lymphovascular invasion; NACT, neoadjuvant chemotherapy; pCR, pathological complete response.

nomogram using these factors to predict the risk of LRR and guide optimal administration of adjuvant radiotherapy (11); however, histopathological characteristics such as molecular subtype, tumor grade, LVI, and Ki67 index were not analyzed in that study (11). One study by the European Organization for Research and Treatment of Cancer 10994/BIG 1-00 revealed that triple-negative or HER2-positive subtype and lack of pathologic response were associated with increased LRR after NACT (12). However, Ki67 index, tumor grade, and LVI were not analyzed in that study (12). Our current study investigated the risk factors of LRR exclusively in patients who underwent NSM/SSM with IBR after NACT and involved several prognostic factors not included in the aforementioned studies that used prospective data. Moreover, in our multivariate analysis, post-NACT Ki67 index, tumor grade, and LVI independently influenced LRR. In our univariate analysis, factors including age at diagnosis, pathological T stage, pathological node stage, and pCR status were associated with LRR rates; however, after multivariate analysis these factors were no longer significant. Notably, the role of post-NACT Ki67, tumor grade, and LVI in LRR risk has

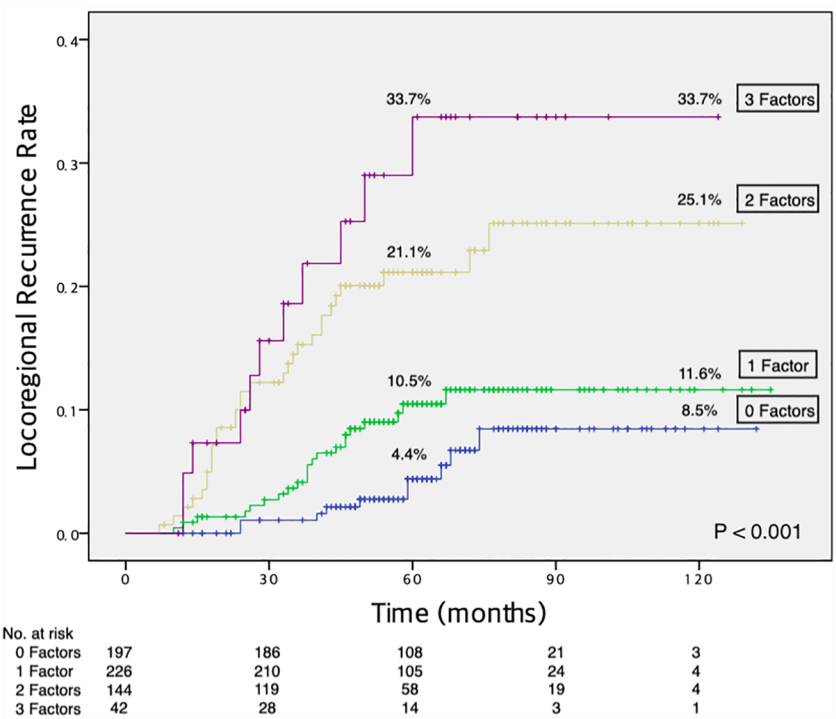


FIGURE 1 | Increased risk of LRR with an increasing number of independent risk factors. LRR, locoregional recurrence.

previously been suggested in smaller retrospective studies (13–15). In a study by Yamazaki et al., 217 patients who underwent NACT and breast-conserving surgery were analyzed, and post-NACT Ki67 > 20%, triple-negative subtype, the presence of LVI, and high tumor grade were found to be significant prognostic factors of LRR (13). However, these factors were identified in a univariate analysis, and no multivariate analysis was conducted (13). In another retrospective study by Wang et al. that included 217 patients with cT1-2N0-1 who underwent NACT and mastectomy, the 5-year LRR rate was 12%, and LVI, tumor grade, and ypN stage were independent prognostic factors of LRR in multivariate analysis (14). However, no data on the Ki67 index were presented in that study (14). In a previous retrospective study including 319 NSM cases after NACT conducted at our center demonstrated that post-NACT Ki67 index was the only independent risk factor for LRR in multivariate analysis (30). Our results on factors correlated with higher LRR risk after NACT are in line with those of previous reports (13–15, 30). In addition, we quantified LRR risk according to the number of independent risk factors and found that the 10-year LRR rate was 8.5% in patients with none of the three independent risk factors, while patients with one, two, or, three of these factors had 10-year LRR rates of 11.6%, 25.1%, and 33.7%, respectively. This risk stratification of LRR may aid in selecting patients who can benefit from further investigation of locoregional management (i.e., adjuvant radiotherapy) strategies in the current setting.

The current study was limited by its retrospective, single-center design, and the study population was heterogeneous for clinicopathological and treatment characteristics. Detailed analysis of the relationship between different adjuvant radiotherapy regimens and LRR, as well as the rate of reconstruction failure, could not be conducted in this study because relevant data were not available. In addition, a relatively small number of patients and LRR events were included in certain subgroups of interest, which might have affected the statistical power of the results.

In conclusion, post-NACT Ki67 ≥ 10%, high tumor grade, and presence of LVI are independently associated with a high risk of developing LRR after NACT and NSM/SSM with IBR. Future prospective trials are warranted to decrease the risk of LRR in patients with associated risk factors.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

This study was approved by the institutional review board of Asan Medical Center, Seoul, Korea (No. 2017-1341). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Z-YW and BK: conception and design. S-HA and BK: administrative support. Z-YW, HK, JL, IC, JK, SL, B-HS, JE, JJ, GG, HK, and BK: data collection. Z-YW: data processing, analysis and manuscript writing. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Early Breast Cancer Trialists' Collaborative Group. Long-Term Outcomes for Neoadjuvant Versus Adjuvant Chemotherapy in Early Breast Cancer: Meta-Analysis of Individual Patient Data From Ten Randomised Trials. *Lancet Oncol* (2018) 19:27–39. doi: 10.1016/S1470-2045(17)30777-5
2. Bailey CR, Ogbuagu O, Baltodano PA, Simjee UF, Manahan MA, Cooney DS, et al. Quality-Of-Life Outcomes Improve With Nipple-Sparing Mastectomy and Breast Reconstruction. *Plast Reconstr Surg* (2017) 140:219–26. doi: 10.1097/PRS.0000000000003505
3. Galimberti V, Vicini E, Corso G, Morigi C, Fontana S, Sacchini V, et al. Nipple-Sparing and Skin-Sparing Mastectomy: Review of Aims, Oncological Safety and Contraindications. *Breast* (2017) 34 Suppl 1:S82–4. doi: 10.1016/j.breast.2017.06.034
4. Lee SB, Lee JW, Kim HJ, Ko BS, Son BH, Eom JS, et al. Long-Term Outcomes of Patients With Breast Cancer After Nipple-Sparing Mastectomy/Skin-Sparing Mastectomy Followed by Immediate Transverse Rectus Abdominis Musculocutaneous Flap Reconstruction: Comparison With Conventional Mastectomy in a Single Center Study. *Med (Baltimore)* (2018) 97:e0680. doi: 10.1097/MD.0000000000010680
5. Shimo A, Tsugawa K, Tsuchiya S, Yoshie R, Tsuchiya K, Uejima T, et al. Oncologic Outcomes and Technical Considerations of Nipple-Sparing Mastectomies in Breast Cancer: Experience of 425 Cases From a Single Institution. *Breast Cancer (Auckl)* (2016) 23:851–60. doi: 10.1007/s12282-015-0651-6
6. Adam H, Bygdeson M, de Boniface J. The Oncological Safety of Nipple-Sparing Mastectomy - a Swedish Matched Cohort Study. *Eur J Surg Oncol* (2014) 40:1209–15. doi: 10.1016/j.ejso.2014.07.037
7. Wong SM, Chun YS, Sagara Y, Golshan M, Erdmann-Sager J. National Patterns of Breast Reconstruction and Nipple-Sparing Mastectomy for Breast Cancer, 2005-2015. *Ann Surg Oncol* (2019) 26:3194–203. doi: 10.1245/s10434-019-07554-x
8. van Tienhoven G, Voogd AC, Peterse JL, Nielsen M, Andersen KW, Mignolet F, et al. Prognosis After Treatment for Loco-Regional Recurrence After Mastectomy or Breast Conserving Therapy in Two Randomised Trials (EORTC 10801 and DBCG-82tm). EORTC Breast Cancer Cooperative Group and the Danish Breast Cancer Cooperative Group. *Eur J Cancer* (1999) 35:32–8. doi: 10.1016/s0959-8049(98)00301-3
9. Mirzabeigi MN, Rhemtulla IA, McDonald ES, Sataloff DM, Kovach SJ, Wu LC, et al. Locoregional Cancer Recurrence After Breast Reconstruction: Detection, Management, and Secondary Reconstructive Strategies. *Plast Reconstr Surg* (2019) 143:1322–30. doi: 10.1097/PRS.0000000000005522
10. Amin MB, Edge SB, Greene FL. *AJCC Cancer Staging Manual*. New York: Springer (2017).
11. Mamounas EP, Anderson SJ, Dignam JJ, Bear HD, Julian TB, Geyer CEJr., et al. Predictors of Locoregional Recurrence After Neoadjuvant Chemotherapy: Results From Combined Analysis of National Surgical

Adjuvant Breast and Bowel Project B-18 and B-27. *J Clin Oncol* (2012) 30:3960–6. doi: 10.1200/JCO.2011.40.8369
12. Gillon P, Touati N, Breton-Callu C, Slaets L, Cameron D, Bonnefoi H. Factors Predictive of Locoregional Recurrence Following Neoadjuvant Chemotherapy in Patients With Large Operable or Locally Advanced Breast Cancer: An Analysis of the EORTC 10994/BIG 1-00 Study. *Eur J Cancer* (2017) 79:226–34. doi: 10.1016/j.ejca.2017.04.012
13. Yamazaki N, Wada N, Yamauchi C, Yoneyama K. High Expression of Post-Treatment Ki-67 Status is a Risk Factor for Locoregional Recurrence Following Breast-Conserving Surgery After Neoadjuvant Chemotherapy. *Eur J Surg Oncol* (2015) 41:617–24. doi: 10.1016/j.ejso.2015.01.036
14. Wang X, Xu L, Yin Z, Wang D, Wang Q, Xu K, et al. Locoregional Recurrence-Associated Factors and Risk-Adapted Postmastectomy Radiotherapy for Breast Cancer Staged in Ct1-2N0-1 After Neoadjuvant Chemotherapy. *Cancer Manag Res* (2018) 10:4105–12. doi: 10.2147/CMAR.S173628
15. Werutsky G, Untch M, Hanusch C, Fasching PA, Blohmer JU, Seiler S, et al. Locoregional Recurrence Risk After Neoadjuvant Chemotherapy: A Pooled Analysis of Nine Prospective Neoadjuvant Breast Cancer Trials. *Eur J Cancer* (2020) 130:92–101. doi: 10.1016/j.ejca.2020.02.015
16. Romanoff A, Zabor EC, Stempel M, Sacchini V, Pusic A, Morrow M. A Comparison of Patient-Reported Outcomes After Nipple-Sparing Mastectomy and Conventional Mastectomy With Reconstruction. *Ann Surg Oncol* (2018) 25:2909–16. doi: 10.1245/s10434-018-6585-4
17. National Comprehensive Cancer Network. Breast Cancer (Version 3.2020). Available at: http://www.nccn.org/professionals/physician_gls/pdf/breast.pdf (Accessed May 7, 2020).
18. Wengler CA, Valente SA, Al-Hilli Z, Woody NM, Muntean JH, Abraham J, et al. Determinants of Short and Long Term Outcomes in Patients Undergoing Immediate Breast Reconstruction Following Neoadjuvant Chemotherapy. *J Surg Oncol* (2017) 116:797–802. doi: 10.1002/jso.24741
19. Santoro S, Loreti A, Cavaliere F, Costarelli L, La Pinta M, Manna E, et al. Neoadjuvant Chemotherapy is Not a Contraindication for Nipple Sparing Mastectomy. *Breast* (2015) 24:661–6. doi: 10.1016/j.breast.2015.08.001
20. Agresti R, Sandri M, Gennaro M, Bianchi G, Maugeri I, Rampa M, et al. Evaluation of Local Oncologic Safety in Nipple-Areola Complex-Sparing Mastectomy After Primary Chemotherapy: A Propensity Score-Matched Study. *Clin Breast Cancer* (2017) 17:219–31. doi: 10.1016/j.clbc.2016.12.003
21. Burdge EC, Yuen J, Hardee M, Gadgil PV, Das C, Henry-Tillman R, et al. Nipple Skin-Sparing Mastectomy is Feasible for Advanced Disease. *Ann Surg Oncol* (2013) 20:3294–302. doi: 10.1245/s10434-013-3174-4
22. Peled AW, Wang F, Foster RD, Alvarado M, Ewing CA, Sbitany H, et al. Expanding the Indications for Total Skin-Sparing Mastectomy: Is it Safe for Patients With Locally Advanced Disease? *Ann Surg Oncol* (2016) 23:87–91. doi: 10.1245/s10434-015-4734-6
23. Krug D, Lederer B, Seither F, Nekljudova V, Ataseven B, Blohmer JU, et al. Post-Mastectomy Radiotherapy After Neoadjuvant Chemotherapy in Breast Cancer: A Pooled Retrospective Analysis of Three Prospective Randomized Trials. *Ann Surg Oncol* (2019) 26:3892–901. doi: 10.1245/s10434-019-07635-x
24. Sun Y, Liao M, He L, Zhu C. Comparison of Breast-Conserving Surgery With Mastectomy in Locally Advanced Breast Cancer After Good Response to Neoadjuvant Chemotherapy: A PRISMA-Compliant Systematic Review and Meta-Analysis. *Med (Baltimore)* (2017) 96:e8367. doi: 10.1097/MD.0000000000008367
25. Wright JL, Takita C, Reis IM, Zhao W, Saigal K, Wolfson A, et al. Predictors of Locoregional Outcome in Patients Receiving Neoadjuvant Therapy and Postmastectomy Radiation. *Cancer* (2013) 119:16–25. doi: 10.1002/cncr.27717
26. Aurilio G, Bagnardi V, Graffeo R, Nolè F, Petit JY, Locatelli M, et al. Does Immediate Breast Reconstruction After Mastectomy and Neoadjuvant Chemotherapy Influence the Outcome of Patients With non-Endocrine Responsive Breast Cancer? *Anticancer Res* (2014) 34:6677–83.
27. Rouzier R, Extra JM, Carton M, Falcou MC, Vincent-Salomon A, Fourquet A, et al. Primary Chemotherapy for Operable Breast Cancer: Incidence and Prognostic Significance of Ipsilateral Breast Tumor Recurrence After Breast-Conserving Surgery. *J Clin Oncol* (2001) 19:3828–35. doi: 10.1200/JCO.2001.19.18.3828
28. Wapnir IL, Anderson SJ, Mamounas EP, Geyer CEJr., Jeong JH, Tan-Chiu E, et al. Prognosis After Ipsilateral Breast Tumor Recurrence and Locoregional Recurrences in Five National Surgical Adjuvant Breast and Bowel Project Node-Positive Adjuvant Breast Cancer Trials. *J Clin Oncol* (2006) 24:2028–37. doi: 10.1200/JCO.2005.04.3273
29. Anderson SJ, Wapnir I, Dignam JJ, Fisher B, Mamounas EP, Jeong JH, et al. Prognosis After Ipsilateral Breast Tumor Recurrence and Locoregional Recurrences in Patients Treated by Breast-Conserving Therapy in Five National Surgical Adjuvant Breast and Bowel Project Protocols of Node-Negative Breast Cancer. *J Clin Oncol* (2009) 27:2466–73. doi: 10.1200/JCO.2008.19.8424
30. Wu ZY, Kim HJ, Lee JW, Chung IY, Kim JS, Lee SB, et al. Oncologic Outcomes of Nipple-Sparing Mastectomy and Immediate Reconstruction After Neoadjuvant Chemotherapy for Breast Cancer. *Ann Surg* (2020). doi: 10.1097/SLA.0000000000003798

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 12 July 2021
doi: 10.3389/fonc.2021.706428

Single Nucleotide Polymorphisms in *HOTAIR* Are Related to Breast Cancer Risk and Prognosis in the Northeastern Chinese Population

Zheng Lv[1], Changgui Kou[2], Naifei Chen[1], Lin Jia[1], Xu Sun[1], Yangyang Gao[1], Rilan Bai[1], Ming Yang[3] and Jiuwei Cui[1]**

[1] Cancer Center, The First Hospital of Jilin University, Changchun, China, [2] Department of Epidemiology and Biostatistics, School of Public Health, Jilin University, Changchun, China, [3] Shandong Provincial Key Laboratory of Radiation Oncology, Cancer Research Center, Shandong, Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China

OPEN ACCESS

Edited by:
Mohammad Taheri,
Shahid Beheshti University of Medical Sciences, Iran

Reviewed by:
Rezvan Noroozi,
Jagiellonian University, Poland
Atefe Abak,
Tabriz University of Medical Sciences, Iran

***Correspondence:**
Ming Yang
aaryoung@yeah.net
Jiuwei Cui
cuijw@jlu.edu.cn

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: 07 May 2021
Accepted: 25 June 2021
Published: 12 July 2021

Citation:
Lv Z, Kou C, Chen N, Jia L, Sun X, Gao Y, Bai R, Yang M and Cui J (2021) Single Nucleotide Polymorphisms in HOTAIR Are Related to Breast Cancer Risk and Prognosis in the Northeastern Chinese Population. Front. Oncol. 11:706428. doi: 10.3389/fonc.2021.706428

Background: The long noncoding RNA HOX transcript antisense RNA (*HOTAIR*) is highly expressed in breast cancer (BC) tissues and is associated with the recurrence and metastasis of BC. Until now, the results of studies on associations between several functional single nucleotide polymorphisms(SNPs) (rs920778, rs1899663, and rs4759314) in *HOTAIR* with BC susceptibility carried out in different regions of China are still inconsistent. There is no study on correlation between *HOTAIR* SNPs and prognosis of Chinese population. Therefore, we investigated the relationship between *HOTAIR* SNPs and susceptibility to and prognosis of BC.

Method: We conducted a population-based case-control study involving 828 BC cases and 905 healthy controls. Peripheral blood DNA was used for genotyping. The association between *HOTAIR* genotypes and BC risk were estimated by odds ratios (ORs) computed using the binary logistic regression model. The relationships between *HOTAIR* SNPs and clinicopathological features were tested by Pearson's chi-square test or Fisher's exact test. Survival was analyzed using the Kaplan-Meier method.

Results: The functional rs920778 genetic variant increased BC risk in the codominant model. Individuals with the rs920778 GG genotype had an OR of 2.426 (95% confidence interval [CI] = 1.491–3.947, $P < 0.001$) for developing BC compared to individuals with the AA genotype. Individuals with the AG genotype had an OR of 1.296 (95% CI = 1.040–1.614, $P = 0.021$) for developing BC compared to individuals with the AA genotype. Individuals with the rs4759314 GA genotype had a lower BC risk than individuals with the rs4759314 AA/GG genotype (OR = 0.566, 95% CI = 0.398–0.803, $P = 0.001$). The rs1899663 genotype had no correlation with BC susceptibility. Haplotypes composed of rs920778–rs1899663 and rs920778–rs1899663–rs4759314 could increase BC risk (all $P < 0.001$). There were no statistically significant associations between *HOTAIR* SNPs and clinicopathological characteristics. The rs920778 GG/AG genotypes were associated

with worse disease-free survival (DFS) (p = 0.012), and the rs4759314 GA genotype was associated with worse DFS and overall survival (OS) (p = 0.011).

Conclusion: *HOTAIR* SNPs(rs920778 and rs4759314) are significantly related to BC susceptibility and prognosis in the northeastern Chinese population, indicating the significance in the occurrence and development of BC.

Keywords: *HOTAIR*, breast cancer, susceptibility, prognosis, single nucleotide polymorphisms

INTRODUCTION

Breast cancer (BC) is one of the most common cancers among women, and its morbidity and mortality have continued to increase worldwide in recent years, reflecting its strong invasive and metastatic characteristics (1, 2). In China, the incidence of BC is increasing annually and is currently the most common malignant tumor in women (3, 4).

Long noncoding RNAs are non-protein-coding transcripts longer than 200 nt and play important roles in the epigenetic regulation of gene expression. One such RNA, HOX transcript antisense RNA (*HOTAIR*), is transcribed from the antisense strand of the *HOXC* locus and mainly regulates *HOXD* genes. *HOTAIR* can guide the polycomb repressor complex 2/lysine-specific histone demethylase 1 complex to a specific target gene, where the complex then trimethylates lysine 27 of histone H3 and dimethylates lysine 4 of histone H3, causing chromatin remodeling (5–7). This can block some metastasis suppressor genes, such as junctional adhesion molecule 2, protocadherin beta 5, and protocadherin 10 (6).

HOTAIR is overexpressed in BC and is related to the occurrence, development, recurrence, and metastasis of BC. A large number of researches indicate that *HOTAIR* has oncogenic impacts. In the diagnosis of gastric cancer, pancreatic cancer, and colorectal cancer, the expression of *HOTAIR* is used to distinguish benign and malignant tissues, compared with benign tissues, the expression of *HOTAIR* in tumor tissues is higher. *HOTAIR* is a biomarker of therapeutic response and poor prognosis (8). In our previous studies, we identified several single nucleotide polymorphisms (SNPs) in *HOTAIR* (rs920778, rs4759314, and rs1899663). These SNPs are located in the intronic region of *HOTAIR* and can regulate its expression (9–11). Therefore, these SNPs are expected to be related to the occurrence, development, recurrence, and metastasis of BC. These SNPs may have the potential to be a new therapeutic target. Further research demonstrated that these sites are related to gastric cancer, esophageal cancer, and papillary thyroid cancer susceptibility. Several meta-analyses showed that these SNPs are associated with the susceptibility of gastrointestinal cancer and estrogen-dependent tumors (12–17), especially in Asian populations. However, these SNPs have different prevalences in different regions and races and are more common in Asian populations than in Caucasian populations. There are also different prevalences in different parts of Asia (12, 17). Few studies have reported a relationship between *HOTAIR* SNPs and BC susceptibility. The participants of the current study were mainly Chinese, Turkish, Iranian, and Indian. The results of the research on populations in different regions are inconsistent and controversial. There are obvious regional differences in the distribution of *HOTAIR* genetic polymorphisms in gastrointestinal cancer. The GG genotype of rs920778 in northeastern population is higher than in middle or southern population, the GG genotype of rs4759314 in southeastern population is higher than in middle and northern population, the GG genotype of rs1899663 in southeastern population is lower than in middle and northern population. Therefore, it is of great significance for us to study the role of *HOTAIR* gene polymorphisms in the occurrence, development, and prognosis of BC in the Northeast population for the first time. This can provide research basis for discovering new pathogenic targets of BC.Therefore, we retrospectively analyzed the relationship between *HOTAIR* SNPs (rs920778, rs1899663, and rs4759314) and BC clinicopathological features and prognosis in the northeastern Chinese population.

Abbreviations: BC, breast cancer; HOTAIR, HOX transcript antisense RNA; SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval; DFS, disease-free survival; OS, overall survival.

MATERIALS AND METHODS

Ethics

This study was approved by the Institutional Ethics Committee of our hospital (ethical approval number 2014-031). Written informed consent was obtained from each participant at recruitment. The study methods were carried out in accordance with the relevant guidelines.

Study Design

Selection and Description of Participants

We investigated the relationship between *HOTAIR* SNPs (rs920778, rs1899663, and rs4759314) and the risk of BC in a case-control study. All of the participants were genetically unrelated Han Chinese individuals from northeast China. This study enrolled 828 BC patients and 905 age-matched healthy control individuals from The First Affiliated Hospital of Jilin University (Changchun, Jilin Province, China) between April 2013 and September 2016. The median follow-up time was 6.7 years. The participants' clinical characteristics were collected through medical records. The inclusion criteria were female patients with early breast cancer diagnosed by pathology.

HOTAIR SNP Genotyping

DNA was extracted from peripheral blood samples. Genotypes were detected using the MassArray system (Agena, San Diego, CA, USA) by the matrix-assisted laser desorption ionization-time of flight mass spectrometry method. *HOTAIR* was selected and genotyped as described previously (9–11). SNP genotyping was performed without knowledge of case status. Reciprocal testing was performed in a random sample of 15%, and the reproducibility was 99.7%.

Statistics

SPSS 24.0 (IBM Corp., Armonk, NY, USA) and the online SNPStats program (https://www.snpstats.net/start.htm, developed by the Institut Català d'Oncologia) were used to analyze BC risk. Variables are characterized as percentages. The Hardy-Weinberg equilibrium test was conducted to test whether the allele frequency distribution of the case group and the control group is biased. Pearson's chi-square test was used to examine differences in demographic variables and *HOTAIR* htSNP genotype distributions between BC cases and controls. Associations between *HOTAIR* genotypes and BC risk were estimated by odds ratios (ORs) and their 95% confidence intervals (CIs), which were computed using the binary logistic regression model. All ORs were adjusted by age whenever appropriate. Pearson's chi-square test or Fisher's exact test were used to evaluate the relationships between *HOTAIR* SNPs and clinicopathological features. The effects of the *HOTAIR* SNPs on disease-free survival (DFS) and overall survival (OS) were evaluated using the Kaplan-Meier method and the univariate Cox model. All statistical tests were two-sided. P values < 0.05 were considered statistically significant.

RESULTS

Participant Characteristics

The control group was composed of healthy women who had undergone routine physical examination in our hospital who did not have a family history of cancer. The median age of the control group was 38 years (range 32–53 years). There were 678 premenopausal women and 226 postmenopausal women. The median age of the case group was 51 years (range 44–58 years), in which there were 398 premenopausal women and 430 postmenopausal women. Only 32 cases had a family history of cancer. Among 828 BC cases, 793 were of an invasive ductal carcinoma and 35 were of other types. Detailed information on the characteristics of the BC patients can be found in **Table 1**.

TABLE 1 | Clinical characteristics of breast cancer patients.

Characteristics	Cases No. (%)
Median age (years)	51 (44-58)
Menstrual status	
Premenopause	398 (48.07)
Postmenopause	430 (51.93)
Family history	
Negative	796 (96.14)
Positive	32 (3.86)
Pathological type	
Invasive ductal carcinoma	793 (95.77)
Other types	35 (4.23)
Histological grade	
I	31 (3.74)
II	511 (61.71)
III	286 (34.54)
Tumor size	
T1	422(50.97)
T2	365(44.08)
T3	27(3.26)
T4	14(1.69)
Lymph node	
N0	396 (47.83)
N1	285 (34.42)
N2	101 (12.20)
N3	46 (5.58)
Lymphovascular invasion	
Negative	471 (56.88)
positive	357 (43.12)
Total	828 (100.00)

Relationship Between *HOTAIR* SNPs and Risk of BC

The genotype distribution of cases and controls showed no deviation for different *HOTAIR* SNPs either in controls or in cases (**Table 2**). The functional rs920778 genetic variant was associated with an increased risk of BC in three genetic models. We used the Akaike Information Criterion to select the optimal genetic model, and the lowest AIC was found in the codominant genetic model. We discovered that the rs920778 GG genotype had an OR for BC development of 2.426 (95% CI = 1.491–3.947, P < 0.001) compared to the AA genotype. The rs920778 AG genotype was also associated with an increased BC risk compared to the rs920778 AA genotype (OR = 1.296, 95% CI = 1.040–1.614, P = 0.021). The functional rs4759314 genetic variants had different associations with BC risk in different genetic models (i.e., the codominant model, dominant model, and overdominant model). The AIC was the lowest in the overdominant model; therefore, using that model, the rs4759314 GA genotype was associated with a lower risk of BC development (OR = 0.566, 95% CI = 0.398–0.803, P = 0.001) than the AA/GG genotype. The rs1899663 SNP did not show an association with BC risk (**Table 3**).

Haplotype Analysis

In order to analyze the influence of different haplotype systems composed of three *HOTAIR* SNP sites on the occurrence of BC, We explored the correlation between haplotypes and BC risk by comparing the distribution of each haplotype in the case group and the control group. There were significant differences between the case and control groups in the distributions of the following haplotypes: rs920778–rs1899663 and rs920778–rs1899663–rs4759314 (all P < 0.001). However, rs1899663–rs4759314 was not related to BC risk (**Table 4**). Haplotype 1 is composed of wild-type genotypes of three SNPs. Haplotype 2 increased BC risk compared with haplotype 1 (OR=1.39, 95%CI=1.13-1.70, P=0.002). Haplotype 4,5, and 6 reduced BC risk compared with haplotype 1 (all P < 0.001) (**Table 5**).

TABLE 2 | Hardy-Weinberg equilibrium test for different *HOTAIR* SNPs.

SNPs	Cases				Controls			
	$^1H_0$	^1He	χ^2	*P*	$^1H_0$	^1He	χ^2	*P*
rs920778	0.3321	0.3545	3.2952	0.0695	0.3105	0.3072	0.1054	0.7455
rs1899663	0.2923	0.3043	1.2960	0.2549	0.3127	0.2989	1.9426	0.1634
rs4759314	0.0743	0.0870	17.5992	<0.001	0.1149	0.1142	0.0398	0.8419

[1]H0:observed value of heterozygote frequency; [1]He:expected value of heterozygote frequency.

TABLE 3 | Association between *HOTAIR* SNPs and breast cancer risk.

SNP	Genotype	Model	Cases No.(%)	Controls No.(%)	OR (95%CI)[a]	*P* value	AIC
rs920778	AA	Codominant	498(60.36)	593(65.52)	1.000	<0.001	2174.4
	AG		274(33.21)	281(31.05)	1.296(1.040-1.614)	0.021	
	GG		53(6.43)	31(3.43)	2.426(1.491-3.947)	<0.001	
	AA	Dominant	498(60.36)	593(65.52)	1.000	0.001	2180.6
	AG/GG		327(39.64)	312(34.48)	1.406(1.140-1.735)		
	AA/AG	Recessive	772(93.58)	874(96.57)	1.000	0.001	2177.2
	GG		53(6.42)	31(3.43)	2.220(1.373-3.588)		
	AA/GG	Overdominant	551(66.79)	624(68.95)	1.000	0.078	2188.1
	AG		274(33.21)	281(31.05)	1.215(0.979-1.509)		
rs1899663	CC	Codominant	552(66.67)	598(66.08)	1.000	0.163	2190.6
	CA		242(29.23)	283(31.27)	0.910(0.729-1.134)	0.400	
	AA		34(4.10)	24(2.65)	1.586(0.900-2.793)	0.111	
	CC	Dominant	552(66.67)	598(66.08)	1.000	0.721	2194.3
	CA/AA		276(33.33)	307(33.92)	0.962(0.778-1.190)		
	CC/CA	Recessive	794(95.89)	881(97.35)	1.000	0.087	2190.1
	AA		34(4.11)	24(2.65)	1.633(0.931-2.864)		
	CC/AA	Overdominant	586(70.77)	622(68.73)	1.000	0.296	2192.6
	CA		242(29.23)	283(31.27)	0.890(0.715-1.108)		
rs4759314	AA	Codominant	756(91.64)	798(88.18)	1.000	0.004	2182.8
	GA		62(7.52)	104(11.49)	0.568(0.400-0.807)	0.002	
	GG		7(0.84)	3(0.33)	1.930(0.459-8.119)	0.370	
	AA	Dominant	756(91.64)	798(88.18)	1.000	0.004	2184
	GA/GG		69(8.36)	107(11.82)	0.609(0.434-0.855)		
	AA/GA	Recessive	818(99.15)	902(99.67)	1.000	0.331	2189.3
	GG		7(0.85)	3(0.33)	2.039(0.486-8.560)		
	AA/GG	Overdominant	763(92.48)	801(88.51)	1.000	0.001	2182
	GA		62(7.52)	104(11.49)	0.566(0.398-0.803)		

[a]OR and 95%CI were analyzed by logistic regression and adjusted by age. Common genotype was taken as reference.

TABLE 4 | Association between haplotypes in *HOTAIR* and breast cancer risk.

Haplotypes	df	Global P
rs920778-rs1899663	3	<0.001
rs1899663-rs4759314	3	0.100
rs920778-rs1899663-rs4759314	7	<0.001

adjusted by age.

Relationship Between *HOTAIR* SNPs and Prognosis of BC

We did not find any significant associations between *HOTAIR* SNPs and clinicopathological characteristics of BC, including tumor size, lymph node metastasis, lymphovascular invasion, molecular type, histological grade, family history, menstrual status, and pathological type (**Table 6**). We then assessed the correlation between *HOTAIR*

TABLE 5 | Haplotype distribution analysis.

Haplotype	rs920778	rs1899663	rs4759314	Controls: Case frequency		*OR*(95%*CI*)	*p*
1				0.7353:	0.7657	1.00	
2	G	A	A	0.1313:	0.1854	1.39(1.13-1.70)	0.002
3	G	C	G	0.0354:	0.0453	1.32(0.91-1.92)	0.140
4	A	A	A	0.0498:	0.0018	0.02(1.01-0.07)	<0.001
5	A	C	G	0.0237:	0.0006	0.01(0.00-0.10)	<0.001
6	G	C	A	0.0228:	0.0012	0.03(0.01-0.12)	<0.001
rare	A	A	G	0.0017:	NA	0.00(-)	–

TABLE 6 | Association between *HOTAIR* SNPs and BC clinical characteristics.

Characteristic	Grouping	rs920778 genotype			P	rs4759314 genotype			P	rs1899663 genotype			P
		AA n(%)	GA n(%)	GG n(%)		AA n(%)	GA n(%)	GG n(%)		AA n(%)	GAn(%)	GG n(%)	
Molecular type	**luminalA**	62 (12.45)	37 (13.50)	5 (9.44)	0.650	95 (12.57)	8 (12.90)	1 (14.29)	0.649	70 (12.68)	30 (12.40)	4 (11.76)	0.527
	luminalB	324 (65.06)	161 (58.76)	33 (62.26)		473 (62.57)	42 (67.74)	3 (42.86)		358 (64.86)	142 (58.68)	20 (58.82)	
	HER2	58 (11.65)	38 (13.87)	7 (13.21)		94 (12.43)	7 (11.29)	1 (14.29)		65 (11.78)	34 (14.05)	4 (11.76)	
	Triple negative	54 (10.84)	38 (13.87)	8 (15.09)		94 (12.43)	5 (8.07)	2 (28.56)		5 9(10.68)	36 (14.87)	6 (17.66)	
Lymph node	**N0**	251 (50.40)	120 (43.80)	23 (43.40)	0.258	363 (48.02)	25 (40.32)	6 (85.71)	0.200	273 (49.46)	111 (45.87)	12 (35.29)	0.407
	N1	168 (33.73)	94 (34.31)	22 (41.51)		260 (34.39)	24 (38.71)	0 (0.00)		187 (33.88)	82 (33.88)	16 (47.06)	
	N2	53 (10.64)	41 (14.96)	7 (13.21)		93 (12.30)	7 (11.29)	1 (14.29)		60 (10.86)	36 (14.88)	5 (14.71)	
	N3	26 (5.23)	19 (6.93)	1 (1.88)		40 (5.29)	6 (9.68)	0 (0.00)		32 (5.80)	13 (5.37)	1 (2.94)	
Tumor size	**T1**	264 (53.01)	137 (50.00)	19 (35.85)	0.221	390 (51.59)	30 (48.39)	1 (14.29)	0.397	288 (52.17)	119 (49.17)	15 (44.12)	0.505
	T2	213 (42.77)	121 (44.16)	31 (58.49)		327 (43.25)	30 (48.39)	6 (85.71)		240 (43.48)	109 (45.04)	16 (47.06)	
	T3	14 (2.81)	11 (4.01)	1 (1.89)		26 (3.44)	1 (1.61)	0 (0.00)		16 (2.90)	10 (4.13)	1 (2.94)	
	T4	7 (1.41)	5 (1.83)	2 (3.77)		13 (1.72)	1 (1.61)	0 (0.00)		8 (1.45)	4 (1.66)	2 (5.88)	
Menstrual status	**Premenopause**	244 (49.00)	124 (45.26)	29 (54.72)	0.372	359 (47.49)	34 (54.84)	4 (57.14)	0.475	276 (50.00)	102 (42.15)	20 (58.82)	0.055
	Postmenopause	254 (51.00)	150 (54.74)	24 (45.28)		397 (52.51)	28 (45.16)	3 (42.86)		276 (50.00)	140 (57.85)	14 (41.18)	
Family history	**negative**	474 (95.18)	266 (97.08)	53 (100.00)	0.136	726 (96.03)	60 (96.77)	7 b(100.00)	1.000	526 (95.29)	236 (97.52)	34 (100.00)	0.159
	positive	24 (4.82)	8 (2.92)	0 (0.00)		30 (3.97)	2 (3.23)	0 (0.00)		26 (4.71)	6 (2.48)	0 (0.00)	
Histological grade	**Ⅰ**	20 (4.02)	8 (2.92)	3 (5.66)	0.877	28 (3.70)	3 (4.84)	0 (0.00)	0.861	21 (3.80)	9 (3.72)	1 (2.94)	0.992
	Ⅱ	305 (61.24)	172 (62.77)	32 (60.38)		464 (61.38)	40 (64.52)	5 (71.43)		342 (61.96)	149 (61.57)	20 (58.82)	
	Ⅲ	173 (34.74)	94 (34.31)	18 (33.96)		264 (34.92)	19 (30.64)	2 (28.57)		189 (34.24)	84 (34.71)	13 (38.24)	
Total		498 (100.00)	274 (100.00)	53 (100.00)		756 (100.00)	62 (100.00)	7 (100.00)		552 (100.00)	242 (100.00)	34 (100.00)	

SNPs and survival in Cox regression analysis. GA genotype of rs920778 and GA genotype of rs4759314 could predict poor prognosis both in univariate analysis and multivariate analysis (**Tables 7** and **8**).

For the rs920778 SNP, there were many significant differences in DFS (P = 0.012) after comparing all three genotype of rs920778, the GG genotype was associated with the worst DFS of the three genotypes (GG, AG, and AA) in univariate analysis (HR = 1.909, P = 0.048). The AG genotype was associated with worse DFS than the AA genotype (HR = 1.48, P = 0.037). However, there was no significant difference in OS (P = 0.13). (**Figure 1** and **Table 8**).

There was no difference in DFS or OS between individuals with the rs1899663 CC or CA genotypes and those with the AA genotype in multivariate analysis. (**Figure 2** and **Table 8**).

When comparing all three rs4759314 genotypes, the GA genotype had worse DFS and OS than those with the AA genotype (P = 0.008). The OS was significantly different when comparing all three genotypes (P = 0.011); individuals with the GA genotype had the worst OS(P=0.001). However, individuals with the GG genotype and those with the AA genotype had similar OS (P = 0.968) (**Figure 3** and **Table 8**).

TABLE 7 | HR in different genotypes of *HOTAIR* SNPs in univariate Cox regression analysis.

Genotype	DFS		OS	
	HR(95%CI)	P-value	HR(95%CI)	P-value
rs920778				
AA	1.000	–	1.000	–
GA	1.646(1.141-2.375)	0.007	1.786(1.008-3.163)	0.047
GG	1.909(1.005-3.625)	0.048	1.209(0.364-4.014)	0.757
rs1899663				
AA	1.000	–	1.000	–
CA	0.768(0.346-1.703)	0.007	1.011(0.232-4.397)	0.989
CC	0.537(0.247-1.166)	0.062	0.892(0.214-3.724)	0.876
rs4759314				
AA	1.000	–	1.000	–
GA	1.850(1.078-3.173)	0.026	2.792(1.357-5.745)	0.007
GG	0.981(0.137-7.028)	0.985	<0.001(0–∞)	0.996

TABLE 8 | HR in different genotypes of *HOTAIR* SNPs in multivariate Cox regression analysis.

Genotype	DFS		OS	
	HR(95%CI)	P-value	HR(95%CI)	P-value
rs920778				
AA	1.000	–	1.000	–
GA	1.480(1.024-2.139)	0.037	1.547(0.871-2.746)	0.138
GG	1.795(0.939-3.429)	0.077	1.133(0.340-3.780)	0.838
rs1899663				
AA	1.000	–	1.000	–
CA	0.819(0.365-1.835)	0.627	1.019(0.233-4.460)	0.980
CC	0.648(0.296-1.416)	0.276	1.091(0.260-4.570)	0.905
rs4759314				
AA	1.000	–	1.000	–
GA	2.076(1.206-3.571)	0.008	3.472(1.675-7.197)	0.001
GG	0.844(0.116-6.130)	0.867	<0.001(0–∞)	0.968

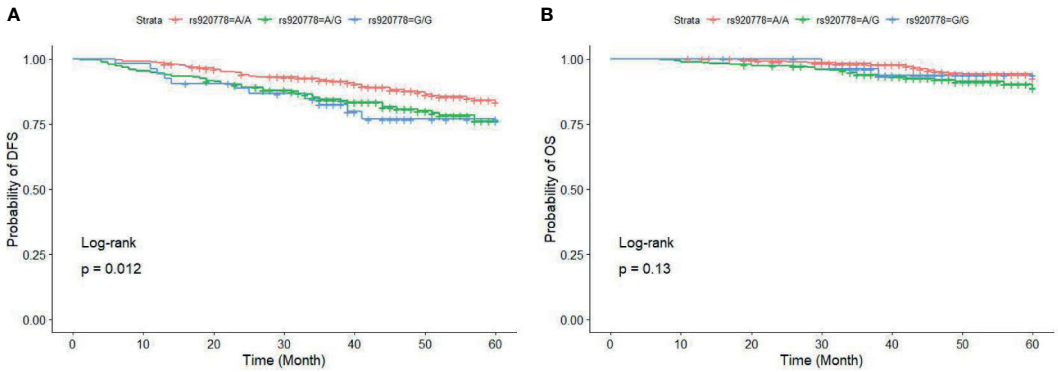


FIGURE 1 | DFS **(A)** and OS **(B)** for BC patients with different genotypes of HOTAIR rs920778. BC, breast cancer; *HOTAIR*, HOX transcript antisense RNA; DFS, disease-free survival; OS, overall survival. The three curves of DFS are statistically significant (P = 0.012), subjects with GG genotype had a worst DFS (P = 0.048); the three curves of OS are not statistically different (P = 0.13), however, subjects with GA genotype had a worst OS than subjects with AA genotype (P = 0.047).

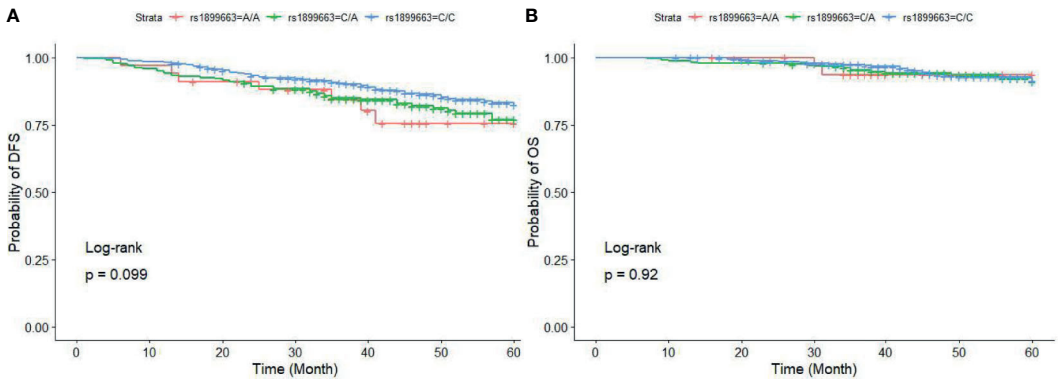


FIGURE 2 | DFS **(A)** and OS **(B)** for BC patients with different genotypes of HOTAIR rs1899663. BC, breast cancer; *HOTAIR*, HOX transcript antisense RNA; DFS, disease-free survival; OS, overall survival. The three curves of DFS/OS are not statistically different, the P value is 0.099 and 0.92 respectively. However, subjects with CA genotype had a worse DFS compared to subjects with AA genotype (P = 0.007).

DISCUSSION

HOTAIR is widely studied as an oncogene, and functional SNPs of *HOTAIR* have been related to cancer risk, including lung cancer, gastric cancer, esophageal cancer, cervical cancer and, prostate cancer, among others. Due to the difference in sample size and population characteristics, the relationship between the *HOTAIR* SNPs and BC risk is still contradictory. Our study may help to identify the significance of these three functional SNPs in BC susceptibility. Over-expression of *HOTAIR* is correlated with poor tumor prognosis, The expression of *HOTAIR* is regulated by multiple factors at the transcriptional and post-transcriptional levels, including estrogen receptors and estrogen receptor coregulators such as histone methylases MLL1 and MLL3 and CBP/p300 binding to the promoter of *HOTAIR* and regulating *HOTAIR* expression (18) and Pumilio homolog 1 regulating *HOTAIR* expression *via* a post-transcriptional mechanism (19). Three functional SNPs of *HOTAIR* can regulate *HOTAIR* expression (20–22), which may influence the BC prognosis. Our present study explored the relationship between the *HOTAIR* SNPs and BC prognosis.

The rs920778 SNP (G > A) is located in the intronic enhancer region of *HOTAIR*, and the AA genotype can increase the expression of *HOTAIR*. In our study, this SNP increased BC risk, which is consistent with the results of Bayram et al. (23), Rajagopal et al. (24), and Hassanzarei et al. (25) (**Table 9**). However, Yan et al. (26) found that the A allele is the most common genotype in the central Chinese population and could increase BC risk, which is contrary to the findings for northeast Chinese, southeast Iranian, South Indian, and Turkish populations (the present study, Hassanzarei et al.'s study, Rajagopal et al.'s study, and Bayram et al.'s study, respectively). We found that the G allele is rare and can increase BC risk. The distributions of rs920778 genotypes in BC patients in these five BC studies differ slightly. However, in these five BC studies, the AA genotype is more common whereas the GG genotype is rare.

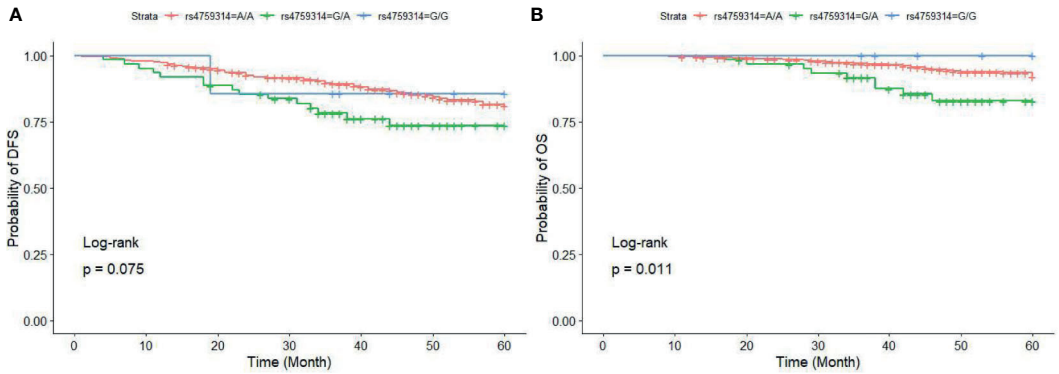


FIGURE 3 | DFS **(A)** and OS **(B)** for BC patients with different genotypes of HOTAIR rs4759314. BC, breast cancer; *HOTAIR*, HOX transcript antisense RNA; DFS, disease-free survival; OS, overall survival. The overall three curves of DFS are insignificant different (P = 0.075), however, subjects with GA genotype had a worst DFS than subjects with AA genotype(P = 0.026); the overall three curves of OS are statistically significant, subjects with GA genotype had a worst OS in the three genotypes, however, subjects with GG genotype and AA genotype had a similar OS.

TABLE 9 | Comparison of previous studies with our study in the association of HOTAIR SNPs and BC risk.

HOTAIR SNPs	Ethnicity	Source of Control	No. of Cases	Assay methods	Genotype			Genetic model	OR (P value)
Author rs920778					**CC%**	**CT%**	**TT%**		
Bayram et al. (23)	Turkish	Hospital[a]	123	TaqManSNP Genotyping	25.2	42.3	32.5	Recessive[a]	2.4 P=0.01
Yan et al. (26)	Chinese	Population	502	PCR-RFLP CRS–RFLP	2.4	30.1	67.5	T Allele	1.41 P=0.02
Hassanzarei et al. (25)	Southeast Iranian	Population	220	PCR-RFLP	15.0	54.1	30.9	Dominant[a]	2.64 P<0.0001
Rajagopal et al. (24)	Indian	Population	502	PCR-RFLP	17.3	50.2	32.5	Over-dominant[a]	1.31 P=0.031
Present Study 2021	Chinese	Population	828	MassAray system	6.4	33.2	60.4	Codominant[a]	2.426 P<0.001
rs1899663					GG%	GT&	TT%		
Yan et al. (26)	Chinese	Population	502	PCR-RFLP CRS–RFLP	67.53	31.35	3.97	T Allele	0.88 P=0.25
Hassanzarei et al. (25)	Southeast Iranian	Population	220	PCR-RFLP	37.7	55.0	7.3	Over-dominant[a]	0.38 P<0.0001
Khorshidi et al. (28)	Iranian	Population	122	ARMS-PCR	30.0	52.0	18.0		1.433 P=0.118
Lin et al. (21)	southeast Chinese	Population	969	PCR-RFLP	82.7	16.2	0.01		2.08 P=0.027
Rajagopal et al. (24)	Indian	Population	502	PCR-RFLP	38.5	45.4	16.1	dominant[a]	1.32 P=0.03
Present Study 2021	Chinese	Population	828	MassAray system	66.67	29.23	4.1	Recessive[a]	1.633 P=0.087
rs4759314					**AA%**	**AG%**	**GG%**		
Yan et al. (26)	Chinese	Population	502	PCR-RFLP CRS–RFLP	89.84	10.71	0.40	G Allele	0.9 P=0.57
Hassanzarei et al. (25)	Southeast Iranian	Population	220	PCR-RFLP	93.2	6.8	0	Codominant[a]	2.31 P=0.0808
Khorshidi et al. (28)	Iranian	Population	122	ARMS-PCR	79.0	21.0	1,0		0.755 P=0.316
Lin et al. (21)	southeast Chinese	Population	969	PCR-RFLP	82.7	16.2	0.01		1.12 P=0.52
Present Study 2021	Chinese	Population	828	MassAray system	91.64	7.52	0.84	overdominant[a]	0.566 P=0.001

[a]*optimal model.*

They differ from the distributions observed in other tumor studies [Yan et al. (26) found that the GG genotype is more common, the AA genotype is rare, and the A allele carries disease risk]. One possible reason for this is differences in tumor type and gender. Further, the study by Yan et al. has limitations in terms of sample size, detection methods, research results, and population. Therefore, we think that the rs920778 GG/AG genotypes can increase BC risk.

The rs1899663 SNP (C > A) is located in the intronic region of *HOTAIR*, and the AA genotype can increase the expression of *HOTAIR* by altering the binding affinity of various transcription factors, such as paired box 4, spermatogenic leucine zipper 1, and zinc finger protein 281 (ZFP281) (28) to *HOTAIR*. The results of studies on the relationship between the rs1899663 SNP and BC susceptibility remain controversial. It has been observed that rs1899663 polymorphism is associated with BC risk in the South Indian population (Rajagopal et al.'s study), the southeast Chinese population (Lin et al.'s study) (21), and the Southeast Iranian population (Hassanzarei et al.'s study) (**Table 9**). However, no relationships were observed in the central Chinese population (Yan et al.'s study), northeastern Chinese population (the present study), or in the Iranian population (Khorshidi et al.'s study) (27). Two smaller studies from Iran (Hassanzarei et al. and Khorshidi et al.) have inconsistent results, as do three larger Chinese studies (Lin et al., Yan et al., and the present study). In Taheri et al.'s study, the relationship between rs1899663 SNP and prostate cancer susceptibility was not observed due to the sample size, however, they compared prostate hyperplasia tissues and prostate cancer tissues and identified that the risk of AA alleles in tumor tissues was higher than CC alleles, This result suggests that AA alleles might increase prostate cancer susceptibility (28).The P value of 0.087 in the present study is close to 0.05. Therefore, we think that SNP has a weak relationship with BC risk when increasing the sample size due to the weak effect of rs1899663 SNP on BC risk.

The rs4759314 SNP (A > G) is located in intronic region of *HOTAIR*, and the GG genotype can increase the expression of *HOTAIR* by enhancing the promoter activity of *HOXC11*. Of five studies examining the relationship between rs4759314 and BC susceptibility (**Table 9**), only two Chinese studies [Yan et al. (26) and this study] have shown a significantly decreased risk of BC in individuals with at least one G allele (GA or GG) compared to individuals with homozygous A alleles. The other three studies did not show any association of rs4759314 with BC risk. Two studies in the Iranian population [Hassanzarei et al (25). and Khorshidi et al. (27)] are too small to draw such conclusions, and another Chinese study in southeast China (Lin et al.'s study) showed that rs4759314 has no correlation with the risk of BC (21). This may be because BC has a population bias, and the population in the other two studies are in middle and northeast China.

We also examined the haplotypes of these three SNPs. We found that the rs920778–rs1899663 and rs920778–rs1899663–rs4759314 haplotypes significantly increase BC risk (P < 0.001). We believe that the gene effect of rs920778 affects the gene effects of the other two SNPs, which leads to an increase in breast cancer susceptibility.

In Bayram's study, researchers found an association between the rs920778 SNP and clinicopathological features in the Turkish population, including advanced TNM stage, larger tumor size, distant metastasis, perineural invasion, and poor histological grade (23). In Hassanzarei's study, they found that the rs920778 SNP was only significantly associated with ER status (25). In Rajagopal's study, they found that the rs920778 variant (AG + GG genotype) increased BC risk in premenopausal women (OR = 5.86, 95% CI = 3.87–8.88, P < 0.0001) (24). However, we did not find any relationship between the rs920778 SNP and any clinicopathological features. This may be because all of these studies were retrospective and there might be an inherent selection bias. Because of the low distribution frequency of the GG genotype (about 3–8% among common populations), a large sample size is needed to analyze the relationship between the GG genotype and clinical characteristics.

We initially found that the rs920778 SNP is associated with the prognosis of BC patients. Our study found that the DFS of patients with the AG/GG genotypes was much shorter than that of patients with the AA genotype (P = 0.012). However, we did not find similar results for OS. Our result is consistent with the result of Weng et al's (29) study showing that subjects with GG genotype of rs920778 had a poor OS, however Xavier-Magalhhães et al's study (30) had the opposite result that subjects with the AG genotype of rs920778 had a longer overall survival than GG subjects in glioma patients. The sample size and tumor type might result the inconsistent results. *HOTAIR* is regarded as an oncogene involved in both the initiation and progression of cancer. The rs920778 SNP is located in the intronic enhancer region of *HOTAIR*, and polymorphism of rs920778 could alter the activity of this enhancer and lead to overexpression of *HOTAIR*. Elevated expression of *HOTAIR* has been reported to be associated with reduced DFS and OS in cervical cancer patients (31). Therefore, we infer that the influence of the rs920778 SNP on BC prognosis is mediated by the resultant increased expression of *HOTAIR*. We need to prove this hypothesis further in BC tissue.

The rs1899663 SNP had no effect on DFS. However, in subgroup analysis, individuals with the CA genotype had worse DFS than those with the AA genotype (P = 0.007), which could provide references for future research. Individuals with the rs4759314 GA genotype had worse DFS and OS than patients with other genotypes(P=0.008 and P=0.001 respectively), which was also interesting and needed further study. Because of the low distribution frequency of the rare genotypes AA of rs1899663 and GG of rs4759314 (no more than 2.4%), a larger sample size is needed to assess their associations with prognosis. Because the rs1899663 and rs4759314 SNPs can increase the expression of *HOTAIR*, their effect on BC prognosis appears to be mediated by the increased expression of *HOTAIR*. However, we need to prove this hypothesis further in BC tissue. Although all the results of survival analysis have not been verified in multivariate analysis, our results suggest that some gene loci may play a role in the occurrence and development of BC.

In summary, this study demonstrates, for the first time, that functional *HOTAIR* SNPs rs920778 and rs4759314 are related to the

risk and prognosis of BC in the northeastern Chinese population, suggesting that these two SNP sites may be involved in the occurrence, development, and metastasis of BC by regulating the expression of *HOTAIR*. This may have certain significance for future diagnosis, drug development, and prognostic judgment of BC. The distribution of gene frequency of the three functional *HOTAIR* SNP loci has a certain correlation with regions and populations. This study only examined the northeast Chinese population as its research object, and it therefore cannot explain why these three *HOTAIR* SNP loci are responsible for the occurrence and development of BC in the overall Chinese population. Therefore, we need a more large prospective multi-center, multi-regional, multi-ethnic population to analyze the significancy of *HOTAIR* SNP in BC development and find a target of treatment.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary files, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Ethics Committee of our hospital (ethical approval number 2014-031). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JC and MY conceived and designed the experiment. NC, LJ, and RB performed the experiment. CK analyzed the data. SX and GY collected clinical information. ZL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was financially supported by the Natural Science Foundation of Jilin Province (20200201474JC).

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer Incidence and Mortality Worldwide: Sources, Methods and Major Patterns in GLOBOCAN 2012. *Int J Cancer* (2015) 136(5):E359–86. doi: 10.1002/ijc.29210
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin* (2021) 71(1):7–33. doi: 10.3322/caac.21654
3. Fan L, Strasser-Weippl K, Li JJ, St Louis J, Finkelstein DM, Yu KD, et al. Breast Cancer in China. *Lancet Oncol* (2014) 15(7):e279–89. doi: 10.1016/S1470-2045(13)70567-9
4. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer Statistics in China, 2015. *CA Cancer J Clin* (2016) 66(2):115–32. doi: 10.3322/caac.21338
5. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* (2007) 129(7):1311–23. doi: 10.1016/j.cell.2007.05.022
6. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-Coding RNA HOTAIR Reprograms Chromatin State to Promote Cancer Metastasis. *Nature* (2010) 464(7291):1071–6. doi: 10.1038/nature08975
7. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science* (2010) 329(5992):689–93. doi: 10.1126/science.1192002
8. Ghafouri-Fard S, Dashti S, Farsi M, Taheri M. HOX Transcript Antisense RNA: An Oncogenic lncRNA in Diverse Malignancies. *Exp Mol Pathol* (2021) 118:104578. doi: 10.1016/j.yexmp.2020.104578
9. Zhang X, Zhou L, Fu G, Sun F, Shi J, Wei J, et al. The Identification of an ESCC Susceptibility SNP Rs920778 That Regulates the Expression of lncRNA HOTAIR via a Novel Intronic Enhancer. *Carcinogenesis* (2014) 35(9):2062–7. doi: 10.1093/carcin/bgu103
10. Pan W, Liu L, Wei J, Ge Y, Zhang J, Chen H, et al. A Functional lncRNA HOTAIR Genetic Variant Contributes to Gastric Cancer Susceptibility. *Mol Carcinog* (2016) 55(1):90–6. doi: 10.1002/mc.22261
11. Zhu H, Lv Z, An C, Shi M, Pan W, Zhou L, et al. Onco-lncRNA HOTAIR and its Functional Genetic Variants in Papillary Thyroid Carcinoma. *Sci Rep* (2016) 6:31969. doi: 10.1038/srep31969
12. Tian T, Li C, Xiao J, Shen Y, Lu Y, Jiang L, et al. Quantitative Assessment of the Polymorphisms in the HOTAIR lncRNA and Cancer Risk: A Meta-Analysis of 8 Case-Control Studies. *PloS One* (2016) 11(3):e0152296. doi: 10.1371/journal.pone.0152296
13. Ge Y, Jiang R, Zhang M, Wang H, Zhang L, Tang J, et al. Analyzing 37,900 Samples Shows Significant Association Between HOTAIR Polymorphisms and Cancer Susceptibility: A Meta-Analysis. *Int J Biol Markers* (2017) 32(2):e231–42. doi: 10.5301/jbm.5000235
14. Li J, Cui Z, Li H, Lv X, Gao M, Yang Z, et al. Long non-Coding RNA HOTAIR Polymorphism and Susceptibility to Cancer: An Updated Meta-Analysis. *Environ Health Prev Med* (2018) 23(1):8. doi: 10.1186/s12199-018-0697-0
15. Liu X, Zhao Y, Li Y, Lin F, Zhang J. Association Between HOTAIR Genetic Polymorphisms and Cancer Susceptibility: A Meta-Analysis Involving 122,832 Subjects. *Genomics* (2020) 112(5):3036–55. doi: 10.1016/j.ygeno.2020.05.018
16. Moazeni-Roodi A, Aftabi S, Sarabandi S, Karami S, Hashemi M, Ghavami S. Genetic Association Between HOTAIR Gene and the Risk of Cancer: An Updated Meta-Analysis. *J Genet* (2020) 99(1):48. doi: 10.1007/s12041-020-01214-w
17. Qi Q, Wang J, Huang B, Chen A, Li G, Li X, et al. Association of HOTAIR Polymorphisms Rs4759314 and Rs920778 With Cancer Susceptibility on the Basis of Ethnicity and Cancer Type. *Oncotarget* (2016) 7:38775–84. doi: 10.18632/oncotarget.9608
18. Bhan A, Hussain I, Ansari KI, Kasiri S, Bashyal A, Mandal SS. Antisense Transcript Long Noncoding RNA (lncRNA) HOTAIR is Transcriptionally Induced by Estradiol. *J Mol Biol* (2013) 425(19):3707–22. doi: 10.1016/j.jmb.2013.01.022
19. Zhang Y, He XY, Qin S, Mo HQ, Li X, Wu F, et al. Upregulation of PUM1 Expression in Preeclampsia Impairs Trophoblast Invasion by Negatively Regulating the Expression of the lncRNA HOTAIR. *Mol Ther* (2020) 28(2):631–41. doi: 10.1016/j.ymthe.2019.11.025
20. Hajjari M, Rahnama S. Association Between SNPs of Long Non-Coding RNA HOTAIR and Risk of Different Cancers. *Front Genet* (2019) 10:113. doi: 10.3389/fgene.2019.00113
21. Lin Y, Guo W, Li N, Fu F, Lin S, Wang C. Polymorphisms of Long non-Coding RNA HOTAIR With Breast Cancer Susceptibility and Clinical Outcomes for a Southeast Chinese Han Population. *Oncotarget* (2018) 9(3):3677–89. doi: 10.18632/oncotarget.23343
22. Du M, Wang W, Jin H, Wang Q, Ge Y, Lu J, et al. The Association Analysis of lncRNA HOTAIR Genetic Variants and Gastric Cancer Risk in a Chinese Population. *Oncotarget* (2015) 6(31):31255–62. doi: 10.18632/oncotarget.5158

23. Bayram S, Sumbul AT, Batmaci CY, Genc A. Effect of HOTAIR Rs920778 Polymorphism on Breast Cancer Susceptibility and Clinicopathologic Features in a Turkish Population. *Tumour Biol* (2015) 36(5):3863–70. doi: 10.1007/s13277-014-3028-0
24. Rajagopal T, Seshachalam A, Akshaya RL, Rathnam KK, Talluri S, Jothi A, et al. Association of HOTAIR (Rs920778 and Rs1899663) and NME1 (Rs16949649 and Rs2302254) Gene Polymorphisms With Breast Cancer Risk in India. *Gene* (2020) 762:145033. doi: 10.1016/j.gene.2020.145033
25. Hassanzarei S, Hashemi M, Sattarifard H, Hashemi SM, Bahari G, Ghavami S. Genetic Polymorphisms of HOTAIR Gene Are Associated With the Risk of Breast Cancer in a Sample of Southeast Iranian Population. *Tumour Biol* (2017) 39(10):1010428317727539. doi: 10.1177/1010428317727539
26. Yan R, Cao J, Song C, Chen Y, Wu Z, Wang K, et al. Polymorphisms in lncRNA HOTAIR and Susceptibility to Breast Cancer in a Chinese Population. *Cancer Epidemiol* (2015) 39(6):978–85. doi: 10.1016/j.canep.2015.10.025
27. Khorshidi HR, Taheri M, Noroozi R, Soudyab M, Sayad A, Ghafouri-Fard S. Investigation of the Association of HOTAIR Single Nucleotide Polymorphisms and Risk of Breast Cancer in an Iranian Population. *Int J Cancer Manage* (2017) 10(5):e7498. doi: 10.5812/ijcm.7498
28. Taheri M, Habibi M, Noroozi R, Rakhshan A, Sarrafzadeh S, Sayad A, et al. HOTAIR Genetic Variants are Associated With Prostate Cancer and Benign Prostate Hyperplasia in an Iranian Population. *Gene* (2017) 613:20–4. doi: 10.1016/j.gene.2017.02.031
29. Weng SL, Wu WJ, Hsiao YH, Yang SF, Hsu CF, Wang PH. Significant Association of Long non-Coding RNAs HOTAIR Genetic Polymorphisms With Cancer Recurrence and Patient Survival in Patients With Uterine Cervical Cancer. *Int J Med Sci* (2018) 15(12):1312–9. doi: 10.7150/ijms.27505
30. Xavier-Magalhaes A, Oliveira AI, de Castro JV, Pojo M, Goncalves CS, Lourenco T, et al. Effects of the Functional HOTAIR Rs920778 and Rs12826786 Genetic Variants in Glioma Susceptibility and Patient Prognosis. *J Neurooncol* (2017) 132 (1):27–34. doi: 10.1007/s11060-016-2345-0
31. Huang L, Liao LM, Liu AW, Wu JB, Cheng XL, Lin JX, et al. Overexpression of Long Noncoding RNA HOTAIR Predicts a Poor Prognosis in Patients With Cervical Cancer. *Arch Gynecol Obstet* (2014) 290(4):717–23. doi: 10.1007/s00404-014-3236-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 22 September 2021
doi: 10.3389/fonc.2021.738330

Non-Mass Enhancements on DCE-MRI: Development and Validation of a Radiomics-Based Signature for Breast Cancer Diagnoses

Yan Li[1], Zhenlu L. Yang[1], Wenzhi Z. Lv[2], Yanjin J. Qin[1], Caili L. Tang[1], Xu Yan[3], Yihao H. Guo[4], Liming M. Xia[1] and Tao Ai[1]**

[1] Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2] Department of Artificial Intelligence, Julei Technology Company, Wuhan, China, [3] Scientific Marketing, Siemens Healthcare Ltd., Shanghai, China, [4] Magnetic Resonance (MR) Collaboration, Siemens Healthcare, Guangzhou, China

OPEN ACCESS

Edited by:
Nicola Fusco,
University of Milan, Italy

Reviewed by:
Elham Sajjadi,
University of Milan, Italy
Umberto Malapelle,
University of Naples Federico II, Italy

****Correspondence:***
Liming M. Xia
xialiming2017@outlook.com
Tao Ai
aitao007@hotmail.com

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: *08 July 2021*
Accepted: *07 September 2021*
Published: *22 September 2021*

Citation:
Li Y, Yang ZL, Lv WZ, Qin YJ, Tang CL, Yan X, Guo YH, Xia LM and Ai T (2021) Non-Mass Enhancements on DCE-MRI: Development and Validation of a Radiomics-Based Signature for Breast Cancer Diagnoses. Front. Oncol. 11:738330. doi: 10.3389/fonc.2021.738330

Purpose: We aimed to assess the additional value of a radiomics-based signature for distinguishing between benign and malignant non-mass enhancement lesions (NMEs) on dynamic contrast-enhanced breast magnetic resonance imaging (breast DCE-MRI).

Methods: In this retrospective study, 232 patients with 247 histopathologically confirmed NMEs (malignant: 191; benign: 56) were enrolled from December 2017 to October 2020 as a primary cohort to develop the discriminative models. Radiomic features were extracted from one post-contrast phase (around 90s after contrast injection) of breast DCE-MRI images. The least absolute shrinkage and selection operator (LASSO) regression model was adapted to select features and construct the radiomics-based signature. Based on clinical and routine MR features, radiomics features, and combined information, three discriminative models were built using multivariable logistic regression analyses. In addition, an independent cohort of 72 patients with 72 NMEs (malignant: 50; benign: 22) was collected from November 2020 to April 2021 for the validation of the three discriminative models. Finally, the combined model was assessed using nomogram and decision curve analyses.

Results: The routine MR model with two selected features of the time-intensity curve (TIC) type and MR-reported axillary lymph node (ALN) status showed a high sensitivity of 0.942 (95%CI, 0.906 - 0.974) and low specificity of 0.589 (95%CI, 0.464 - 0.714). The radiomics model with six selected features was significantly correlated with malignancy ($P<0.001$ for both primary and validation cohorts). Finally, the individual combined model, which contained factors including TIC types and radiomics signatures, showed good discrimination, with an acceptable sensitivity of 0.869 (95%CI, 0.816 to 0.916), improved specificity of 0.839 (95%CI, 0.750 to 0.929). The nomogram was applied to the validation cohort, reaching good discrimination, with a sensitivity of 0.820 (95%CI, 0.700 to 0.920), specificity of 0.864 (95%CI,0.682 to 1.000). The combined model was clinically helpful, as demonstrated by decision curve analysis.

Conclusions: Our study added radiomics signatures into a conventional clinical model and developed a radiomics nomogram including radiomics signatures and TIC types. This radiomics model could be used to differentiate benign from malignant NMEs in patients with suspicious lesions on breast MRI.

Keywords: breast cancer, non-mass enhancement, radiomics, differential diagnosis, magnetic resonance imaging

1 INTRODUCTION

According to the American College of Radiology (ACR) BI-RADS® Atlas, 5th edition (1), breast lesions with abnormal enhancement variables on dynamic contrast-enhanced breast magnetic resonance imaging (breast DCE-MRI) include foci, masses, and non-mass enhancement lesions (NMEs). In 2020, breast cancer became the most common cancer of women worldwide (2), and the differentiation between benign and malignant breast lesions using MRI-based diagnostics was found to be critical for breast cancer treatments. However, distinguishing benign and malignant breast lesions on DCE-MRI is challenging, especially when NMEs are present (3).

NMEs are associated with a wide-ranging spectrum of different pathologic findings (4–6), with an overlap in the imaging findings between malignant and benign lesions. NMEs remain a diagnostic challenge for radiologists despite the frequent attempts to distinguish benign from malignant NMEs using different methodologies, including conventional morphologic comparisons (6–8) and the measurement of different parameters, such as ADC values and the initial slope of kinetic curves (9–11). Baltzer et al. reported that the primary cause for false positive results of breast MRI may due to NMEs, resulting in unnecessary biopsies (12). Studies have shown that morphologic assessments are disputable in attempting to differentiate benign *vs*. malignant NMEs. Some studies have demonstrated that morphologic assessments are more useful than kinetic assessments in distinguishing NMEs (13–15), while other studies have reported that morphologic assessments have a relatively low specificity and sensitivity to distinguish NMEs (16–18). In addition, morphologic assessments depend on the human eye are subjective with limitations; thus, substantial inter- and intra-observer variability is seen with these assessments (19). A meta-analysis (20) showed heterogeneity among studies with sensitivities from 0% to 100% and specificities from 48% to 100%. These factors underscore the complexity of the diagnostic phase and simultaneously present a therapeutic challenge. For example, idiopathic granulomatous mastitis, a benign inflammatory disease, can mimic breast cancer, both clinically and radiologically (21, 22).

In recent years, radiomics, a technology of transforming digital medical images into quantifiable data to improve medical decisions (23), has been found to have a potential benefit in increasing the knowledge base of diagnostic oncology and predicting the accuracy of medical imaging. Radiomics is partially based on the hypothesis that medical images contain much more information than can be visually deciphered by radiologists (24). According to our best knowledge, there is little research reported the additional value of radiomics to differentiate benign *vs*. malignant NMEs on DCE-MRI. Additionally, to date, a model that combines a radiomics signature and conventional analysis to produce superior diagnostic performance in diagnosing malignant NMEs has yet to be reported.

In this study, we developed and validated a nomogram that combined radiomics and conventional analytic clinical factors to evaluate the additional value of radiomics in differentiating benign from malignant NMEs. We also compared the diagnostic performance of the nomogram with the radiomics score and analytic clinical factors alone.

2 MATERIALS AND METHODS

2.1 Patients

We retrospectively reviewed 3352 consecutive patients who underwent breast MRI in our hospital between December 2017 and October 2020. In total, 232 female patients with 247 lesions were selected and comprised the primary training cohort (mean age, 44.8 ± 10.6 years). Among these patients, 14 had additional lesions in the contralateral breast and 1 patient had two lesions in different quadrants of her left breast. The inclusion criteria were as follows: (a) histologically confirmed benign or malignant breast lesions on DCE-MRI examinations; (b) no previous treatments or breast implants; (c) no pregnancy or lactation; and (d) NMEs found on DCE images. Patients were excluded if image quality was poor, hemorrhage was present after biopsy, lesions did not involve parenchyma on the DCE images, or the lesion sizes were <5 mm. Using this inclusion and exclusion criteria, a validation cohort of 72 consecutive female patients (mean age, 47.9± 11.2 years) was selected from 908 consecutive patients between November 2020 and April 2021 in our hospital. A flowchart of this study is presented in **Figure 1**. For each patient, conventional clinical data, including age and menopause status, were obtained from electronic medical records.

2.2 Magnetic Resonance Image Acquisition

MR examinations for both the validation cohort and training cohort were obtained on a 3T scanner (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany) in our hospital. All scans were performed with a dedicated 16-channel phased-array breast coil in the prone position using the same protocol.

For breast diffusion-weighted imaging (DWI), multi-b-value DWI was applied with a readout-segmented technique (RESOLVE DWI), similar to our previous works (25): repetition time (TR) = 5000 ms, echo time (TE) = 70 ms, field of view (FOV) = 169 x 280 mm^2, matrix size = 114 x 188, slice thickness = 5.0 mm, readout

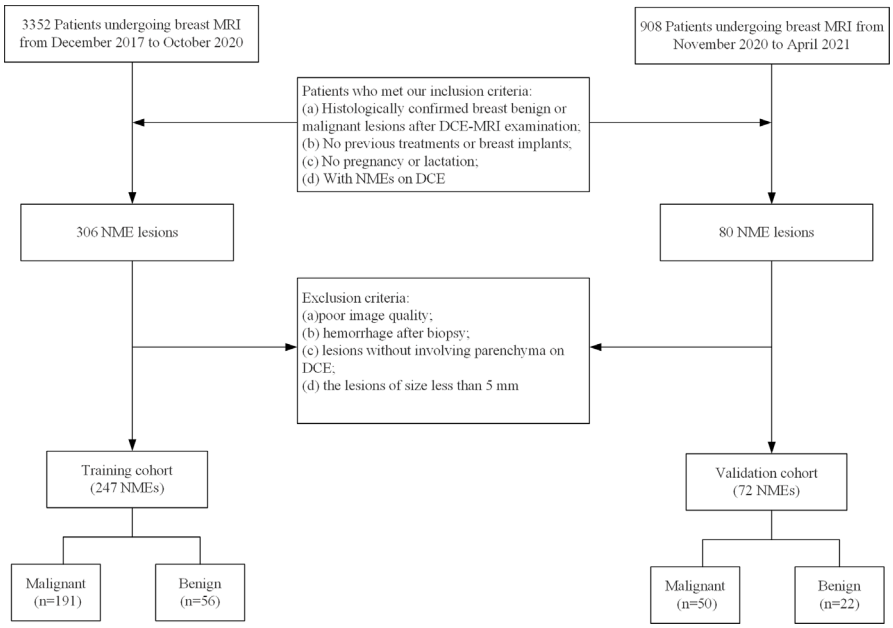


FIGURE 1 | Flowchart of the study population enrollment. NME, non-mass enhancement lesion.

segment = 5, average = 1, diffusion gradient mode = 3-scan-trace, b values = 0, 50, 1000s/mm^2, and acquisition time = 4:27 (min: sec).

For breast DCE-MRI, a protocol based on time-resolved angiography was used with a stochastic trajectory, volume-interpolated breath-hold examination sequence (TWIST-VIBE). The detailed scan parameters were as follows: TR = 5.24 ms, TE = 2.46 ms, matrix size = 182 x 320, FOV = 260 x 320 mm^2, slice thickness = 1.5 mm without gap, flip angle = 10°, temporal resolution = 5.74 s/phase, and acquisition time = 5:57(min: sec).

The contrast medium (Omniscan, GE Healthcare, Milwaukee, WI) was intravenously injected with a power injector at the end of the third acquisition phase. The dose was 0.1 mmol/kg body weight, with an injection rate of 2.5 mL/s, which was followed by a 20 mL saline flush.

2.3 Image Interpretation

For each patient in the training cohort and the validation cohort, two radiologists (Y.L. and T.A. with 8 and 10 years of experience in breast MRI, respectively), were blinded to the pathologic results. Each radiologist reviewed all breast MR images from the 304 patients, assessing breast density, the degree of background parenchymal enhancement, and MR-reported lymph node status by consensus. The maximal diameter, internal enhancement, and distribution were recorded in the very early phase (about 90 seconds) after contrast media injection according to the BI-RADS 5th edition (1). Of these, the maximal diameter was assessed on multiplanar reformatted images using a Siemens clinical workstation. The type of time-intensity curve (TIC) for each case was drawn based on DCE-MRI with a region of interest (ROI) of approximately 0.2-0.4 cm^2 placed on each slice at the brightest part of the lesions on images obtained in the early phase after the contrast injection. We recorded the high-level TIC curve types when different types were present in each lesion. On all slices of the apparent diffusion coefficient (ADC) maps, multiple ROIs were carefully placed on the darkest areas, which were confirmed by agreement by the two radiologists. Thus, the lowest ROI ADC value was regarded as the minimum ADC value for each lesion. If no lesions could be evaluated with DWI or the ADC maps, we copied ROIs on the DCE-MRI image and pasted them on the ADC maps. We defined the axillary lymph node (ALN) with a maximal short diameter of ≥10mm, an absent fatty hilum, or a long axis/short axis of <2 as MR-reported ALN positive. Vasodilation of the surrounding feeding artery was defined as positive on maximum intensity projection images (MIPs) and was included based on our experience. The above-mentioned factors were all initial clinical candidate predictors for NME differentiation.

2.4 Features Extraction and Radiomics Signature

The radiomics signature was applied to the clinical analyses, and a diagnostic model for differentiation was developed using the training cohort. The radiomics analysis was performed on the very early phase (90 seconds) images after contrast media injection, as was the morphologic evaluation. Prior to the

radiomics analysis, the images of each case were transferred into the open-source software, ITK-SNAP (Version 3.8.0), to perform semi-automatically ROI segmentation. ROIs were drawn with care to include the whole lesion, avoiding normal glandular tissue, fat, vessels, and necrosis. Pyradiomics open-source software (https://pyradiomics.readthedocs.io/en/latest/index.html) was used to automatically extract tissue intensities and textural, morphologic, and wavelet features. We used the least absolute shrinkage and selection operator (LASSO) method, an appropriate tool for high-dimensional data regression (26), to select the most effective features from the training cohort data set. For each lesion, a radiomics score (Rad-score) was calculated weighting by the respective coefficients of selected features.

2.5 Nomogram in the Training Cohort and Validation

Initial clinical multivariate logistic regression analysis included age, menopause status, maximal diameter, fibrotic gland tissue, background parenchymal enhancement, morphologic assessment, ALN status, and TIC assessment on DCE-MRI and the minimum ADC values on DWI. We added radiomics features into the clinical multivariable logistic regression analysis and built the radiomics nomogram to supply the radiologists and clinicians with an effective tool for differentiating benign and malignant NMEs. The calibration curve and Hosmer & Lemeshow test (27) were adapted to evaluate the radiomics nomogram calibration. Nomogram performance was evaluated using the area under the curve (AUC) analysis.

2.5.1 Consistency Validation

In the data set of the training cohort, consistency validation was performed by comparing the first measurement and second measurement one month later of reader 1 (Y.L.) for intra-observer agreement. The second measurement of reader 1 and the extraction of reader 2 (Z.L.Y) in 60 patients were compared to produce inter-observer agreement. The interclass correlation coefficient (ICC) was applied to assess the feature extraction agreement, which was greater than 0.80 and considered excellent.

2.5.2 Data Validation

We applied the same method as that of the training cohort to calculate the Rad-score in the validation cohort. We applied the logistic regression equation produced in the training cohort to all lesions of the validation cohort. We tested the performance of the nomogram using calibration and AUC analyses.

2.6 Statistical Analysis

R (RStudio, Version 3.6.3) software was used for algorithms and statistical analyses. For continuous variates, Student's t-tests were performed. For categorical variates, the chi-square test or Wilcoxon rank-sum test were applied. We used univariate logistic regression analysis to determine potential factors affecting differentiation. Then, logistic regression models containing the above-mentioned potential factors were used for multivariate analysis. A nomogram was built on the logistic regression model as a graphical presentation. The area under the receive operating characteristic (AUC-ROC) curve, accuracy, sensitivity, and specificity were applied to indicate the discriminative ability of each factor and nomogram. P-values <0.05 (two-tailed) was considered statistically significant.

3 RESULTS

3.1 Conventional Clinical Analysis

3.1.1 Training Cohort

In the training cohort, of the 247 lesions, 191 malignant and 56 benign lesions were confirmed pathologically by either biopsy, lumpectomy, or mastectomy. For the patient who had two lesions in the left breast, the lesion in the upper outer quadrant was confirmed as adenosis, while the lesion in the medial area was ductal cancer *in situ*. Specific pathologic results are shown in **Table 1**. Internal enhancement patterns, background parenchymal enhancements (BPEs), and MRI reported-fibroglandular tissue (FGT) were not different between malignant and benign lesions (P=0.397, 0.760, 0.139). The mean age of the patients with malignant lesions was older than that of the benign cases (P=0.035). The maximal diameter of the malignant lesions was significantly longer than that of the benign lesions (P<0.001). A higher proportion of postmenopausal women were found in the malignant group than in the benign group (P=0.034). The constituent ratio of distribution was significantly different between malignant and benign cases (P<0.001). Of these, the proportion with linear distributions was higher in the benign group than in the malignant group (P=0.046). The minimum ADC value of the malignant lesions was significantly lower than that of the benign lesions (P<0.001). The malignant group had a significantly higher percentage of higher-level TIC pattern types and MR-reported ALN-positive and MIP-positive cases (all P<0.001). Specific results are shown in **Table 2**. Age, menopause status, maximal diameters, distributions, TIC patterns, minimum ADC values, MRI reported-ALN status, and MIP status were potential factors influencing differentiation according to the univariate logistic

TABLE 1 | Pathologic findings for all non-mass enhancement (NME) lesions.

Pathological results	Training Cohort (n = 247)	Validation Cohort (n = 72)
Benign	56	22
Adenosis	49	22
Papilloma	1	0
Chronic inflammation	5	0
Fibroadenoma/fibroadenomatous change	1	0
Malignant	191	50
IDC	88	21
ILC	11	0
Pure DCIS	31	9
Invasive cancer with CIS	50	14
CIS with invasive component	24	6
Mucinous carcinoma	1	0

IDC, Invasive ductal carcinoma; ILC, Invasive lobular carcinoma; DCIS, ductal carcinoma in situ; CIS, cancer in situ.

TABLE 2 | Characteristics of patients in the training and validation cohorts.

Characteristic	Training Cohort			Validation Cohort		
	Malignant (n = 191)	Benign (n = 56)	P	Malignant (n = 50)	Benign (n = 22)	P
Age, mean ± SD, years	45.4 ± 10.2	42.0 ± 11.8	0.035	51.4 ± 10.2	39.9 ± 9.5	<0.001
Menopause status, No (%)						
Postmenopausal	41 (21.5)	5 (8.9)	0.034	20 (40)	2 (9.1)	0.011
premenopausal	150 (78.5)	51 (91.1)		30 (60)	20 (90.9)	
MRI reported-FGT, No (%)						
a	1 (0.5)	0 (0)		2 (4)	1 (4.5)	
b	37 (19.4)	8 (14.3)		11 (22)	3 (13.6)	
c	140 (73.3)	44 (78.6)		34 (68)	15 (68.1)	
d	13 (6.8%)	4 (7.1)	0.395	3 (6)	3 (13.6)	0.331
MRI reported-BPE, No (%)						
Minimal-Mild	143 (74.9)	36 (64.3)		24 (48)	7 (31.8)	
Moderate	43 (22.5)	19 (33.9)		23 (46)	10 (45.5)	
Marked	5 (2.6)	1 (1.8)	0.099	3 (6)	5 (22.7)	0.079
Maximal diameter, mean ± SD, mm	47.7 ± 21.4	32.9 ± 18.9	<0.001	44.3 ± 16.6	29.6 ± 10.6	<0.001
NME Enhancement patterns						
Distribution, No (%)						
Focal	21 (11.0)	16 (28.6)		0 (0)	0 (0)	
Linear	2 (1.0)	6 (10.7)		0 (0)	4 (18.2)	
Segmental	38 (19.9)	11 (19.6)		13 (26)	5 (22.7)	
Regional	82 (42.9)	14 (25.0)		26 (52)	13 (59.1)	
Multiple regions	35 (18.3)	8 (14.3)		9 (18)	0 (0)	
Diffuse	13 (6.8)	1 (1.8)	<0.001	2 (4)	0 (0)	0.016
Internal enhancement patterns, No (%)						
Homogeneous	11 (5.8)	7 (12.5)		2 (4)	4 (18.2)	
Heterogeneous	127 (66.5)	34 (60.7)		36 (72)	12 (54.5)	
Clumped	46 (24.1)	13 (23.2)		9 (18)	6 (27.3)	
Clustered ring	7 (3.7)	2 (3.6)	0.438	3 (6)	0 (0)	0.474
TIC pattern, No (%)						
Persistent	11 (5.8)	33 (58.9)		3 (6)	12 (54.5)	
Plateau	82 (42.9)	18 (32.1)		27 (54)	9 (40.9)	
Washout	98 (51.3)	5 (8.9)	<0.001	20 (40)	1 (4.5)	<0.001
Minimum ADC value, mean ± SD, 10^-6 mm2/s	769.3 ± 173.4	914.2 ± 247.8	<0.001	730.1 ± 147.8	898.5 ± 118.1	<0.001
MRI reported- ALN status, No (%)						
ALN-positive	70 (36.6)	4 (7.1)		15 (30)	0 (0)	
ALN-negative	121 (63.4)	52 (92.9)	<0.001	35 (70)	22 (100)	0.003
MIP						
positive	117 (61.3)	17 (30.4)		16 (32)	8 (36.4)	
negative	74 (38.7)	39 (69.6)	<0.001	34 (68)	14 (63.6)	0.789
Radiomics score, median (interquartile range)	1.833 (1.320to 2.391)	0.368 (-0.335 to 0.977)	<0.001	1.453 (1.108 to 2.074)	0.376 (-0.161 to 0.925)	<0.001

Percentages may not add up to 100 because of rounding. TIC, Time Intensity Curve; BPE, Background Parenchymal Enhancement; FGT, Fibro glandular Tissue; ALN, Axillary Lymph Nodes; MIP, Maximum Intensity Projection.

regression analysis. From the multivariate analysis results, higher-level TIC pattern types, and MR-reported ALN-positive statuses were significantly associated with malignancy (all P<0.001). The AUCs, sensitivities, and specificities of the clinical multivariate regression model developed using TIC types and MR-reported ALN status were 0.852 (95%CI: 0.799-0.906), 0.942 (95%CI: 0.906-0.974), and 0.589 (95%CI:0.446-0.714), respectively, to differentiate between malignant and benign NME lesions. The specific results are shown in **Table 3**.

3.1.2 Validation Cohort

In the validation dataset, there were 50 malignant lesions and 22 benign lesions. Like the training dataset, internal enhancement patterns, MRI reported-FGT, and BPE were not significantly different between malignant and benign lesions. Moreover, no significant differences were found between the two cohorts regarding the MIP status. When applying the clinical multivariate logistic regression equation of the primary cohort to the validation dataset, the AUCs, sensitivities, and specificities were 0.842(95%CI: 0.758-0.926), 0.940(95%CI: 0.860-1.000), and 0.545(95%CI: 0.364-0.727), respectively (**Table 3**).

3.2 Radiomics Analysis and the Combined Model

3.2.1 Training Cohort

Of all features extracted from the lesions in the primary cohort, six features were selected as potentially effective factors for differentiation and were applied in the Rad-score calculation (**Figure 2**). The final computation of the model coefficients led to the following differentiation model for NMEs:

TABLE 3 | Risk factors for malignancy and the performance of the clinical and combined models for breast non-mass enhanced (NME) lesions.

Intercept and Variable	Clinical model			Radiomics model	Combined model		
	β	Odds Ratio (95% CI)	P	NA	β	Odds Ratio (95% CI)	P
Intercept	-3.064		<0.001	NA	-3.167		<0.001
TIC types	2.049	7.761 (4.225 to14.259)	<0.001	NA	1.463	4.319 (2.310 to 8.074)	<0.001
MR-reported ALN status	1.399	4.052 (1.262 to13.009)	0.019	NA	0.680	1.975 (0.526 to 7.419)	0.314
Radiomics signature	NA	NA	NA	NA	1.173	3.233 (1.963 to 5.325)	<0.001
AUC							
Training cohort		0.852 (0.798 to 0.906)		0.864 (0.805 to 0.923)		0.908 (0.864 to 0.952)	
Validation cohort		0.842 (0.758 to 0.926)		0.876 (0.791 to 0.962)		0.901 (0.827 to 0.974)	
Sensitivity							
Training cohort		0.942 (0.906 to 0.974)		0.827 (0.770 to 0.880)		0.896 (0.817 to 0.916)	
Validation cohort		0.940 (0.860 to 1.000)		0.800 (0.680 to 0.900)		0.820 (0.700 to 0.920)	
Specificity							
Training cohort		0.589 (0.464 to 0.714)		0.804 (0.696 to 0.911)		0.839 (0.750 to 0.862)	
Validation cohort		0.545 (0.364 to 0.727)		0.863 (0.727 to 1.000)		0.864 (0.682 to 1.000)	

β is the regression coefficient; NA, not applicable; TIC, time-intensity curve; ALN, axillary lymph node; AUC, area under the curve.

$$\begin{aligned} Rad-score = & -0.594(original\_shape\_SurfaceVolumeRatio) \\ & +0.061(wavelet.HLL\_glcm\_ldn) \\ & -0.176(original\_firstorder\_Skewness) \\ & +0.343(wavelet.LLH\_glcm\_ldmn) \\ & -0.017(original\_glszm\_SmallAreaEmphasis) \\ & +0.110(wavelet.LLL\_firstorder\_Kurtosis) \\ & +1.468 \end{aligned}$$

Of the six features, the biggest weight was given to the shape feature (Surface Area to Volume Ratio). A significant difference in the Rad-score between benign and malignant NMEs was found in the training cohort (P<0.001). The AUC, sensitivity, and specificity of the radiomics multivariable logistic regression alone for NME differentiation was 0.864 (95%CI: 0.805-0.923), 0.827 (95%CI: 0.770-0.880), and 0.804 (95%CI: 0.696-0.893) (**Figure 3**, **Table 3**). After adding the radiomics analysis into the clinical multivariate regression model, MR-reported ALN status was no longer an independent factor of malignancy. We built a nomogram for the training cohort based on the TIC types and the radiomics signature (**Figure 4**), the specificity of which was improved from 0.589 (95%CI: 0.464- 0.714) in the clinical model to 0.839 (95%CI: 0.750- 0.862) in the combined model (**Table 3**). The final regression equation and correlation coefficients were calculated. In **Table 3**, the parameters in detail are reported. Using ROC curve analysis, the optimal cutoff value of the final regression equation was 0.772. Lesions with values below the cutoff value are judged as benign, while those with values exceeding the cutoff value are judged as malignant.

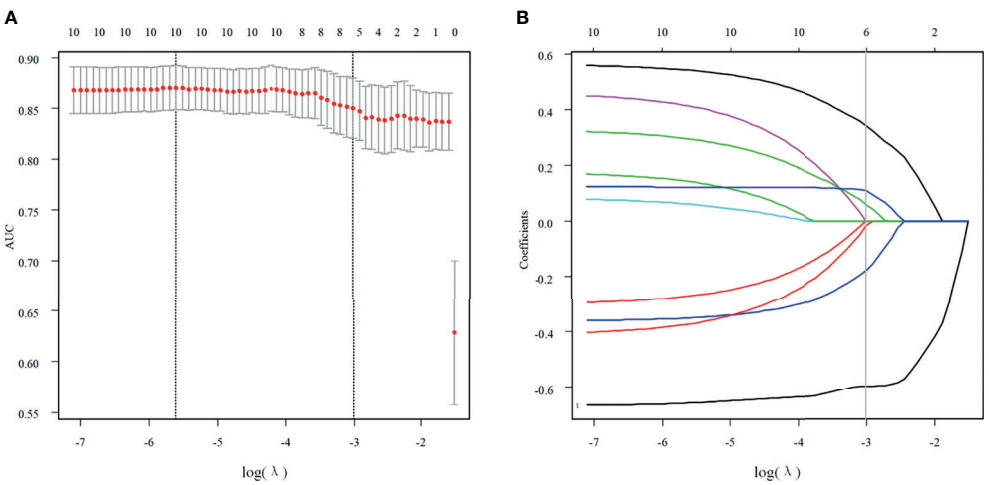


FIGURE 2 | Texture feature selection. **(A)** Using the LASSO model, tuning parameter (λ) selection was according to a 5-fold cross-validation. Using the minimum criteria and the 1 standard error of the minimum criteria, dotted vertical lines were drawn for the optimal values. A λ value of 0.0495 with a log (λ) of -3.005783 was chosen for the 5-fold cross-validation. **(B)** According to the log (λ) sequence, a coefficient profile plot was produced. At the value selected with the 5-fold cross-validation, a vertical line was drawn, where the optimal λ resulted in six non-zero coefficients.

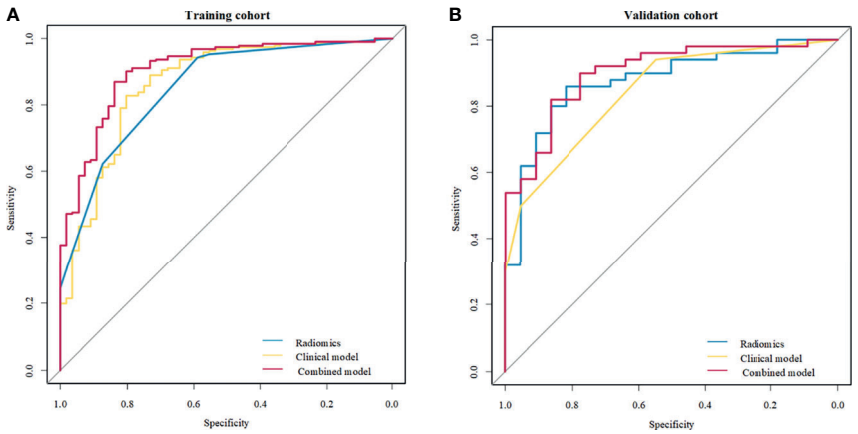


FIGURE 3 | Receiver operating characteristic (ROC) curves of the clinical model, radiomics signature, and combined model to differentiate benign from malignant non-mass enhancement (NME) lesions. **(A)** Three methods in the training cohort; **(B)** Three methods in the validation cohort.

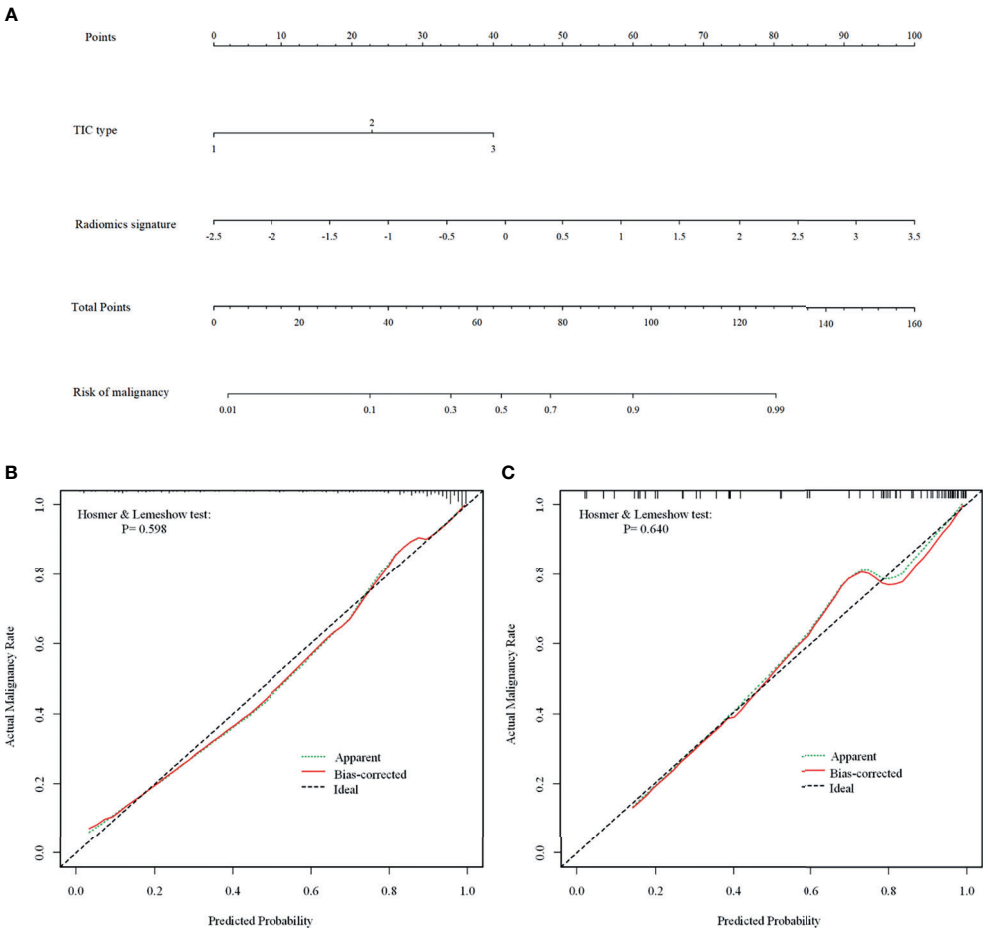


FIGURE 4 | The combined nomogram for differentiating benign and malignant non-mass enhancement (NME) lesions. **(A)** The radiomics nomogram developed with the training cohort included time-intensity curve (TIC) types and radiomics signatures. **(B, C)** Calibration curves of the combined model in the training **(B)** and validation **(C)** cohorts. The Bias-corrected line represents the nomogram performance. The closer the red Bias-corrected line is to the diagonal dotted (ideal) line indicates a better differentiation performance.

3.2.2 Validation Cohort

In the validation cohort, there was also a significant difference in the Rad-score between benign and malignant NMEs (P<0.001). After adding the Rad-score analysis into the clinical model, the specificity increased from 0.545 (95%CI: 0.364- 0.727) to 0.864 (95%CI: 0.682- 1.000) (**Table 3**).

For the differentiation between benign and malignant NMEs, the calibration curve of the combined model demonstrated excellent agreement between the prediction and real pathologic results in the training cohort as well as the validation cohort (**Figure 4**). In clinical medicine, the decision curve analysis for the combined model was developed according to a previous study (28) and is showed in **Figure 5**. The decision curve demonstrated that if the threshold probability was >19%, the nomogram could add more benefit to the discrimination of benign and malignant NMEs than the clinical model.

3.3 Consistency Validation

Based on the comparisons of radiomics feature measurements assessed one month apart by reader 1, the intra-observer agreement was excellent (ICC value=0.936, 95%CI: 0.929 to 0.942). Using the second measurements of the 60 patients assessed by reader 1 and the features extraction of the same data set assessed by reader 2, inter-observer was also excellent (ICC value =0.887, 95%CI: 0.876 to 0.898).

Figures 6 and **7** show two cases in detail.

3.4 Specificity Changes

Considering the low specificity in the conventional clinical analysis, we conducted an analysis for the false positive (FP) lesions (n=33) and the true negative (TN) lesions (n=45) on the basis of the conventional clinical analysis in the whole cohort (78 benign NMEs). The results showed that compared to the TN lesions, the FP lesions had a significant larger proportion of moderate or marked BPE (P=0.004), plateau or washout type of TIC (P<0.001), and positive MIP sign (P<0.001). Of the 33 FP NMEs, 30 (90.9%) lesions were confirmed as adenosis, and the other 3 lesions were chronic inflammation. In addition, 21 of 33 (63.6%) FP lesions were categorized as malignancy applying the final combined model.

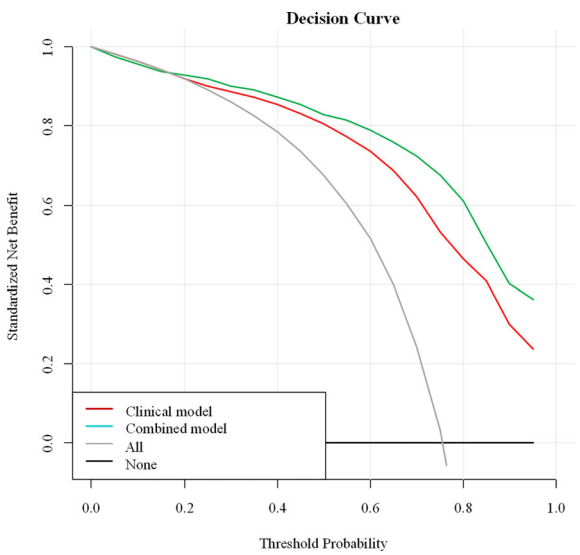


FIGURE 5 | Decision curve analysis of the combined model. The Y-axis demonstrates the net benefit to patients. As indicated in the curve, the net benefit of using the combined model to differentiate benign and malignant NME lesions is greater than when the clinical model is used at a threshold probability of > 0.19.

4 DISCUSSION

In this study, we developed a clinical model that consisted of clinical characteristics, morphologic lesion assessments, the ALN status, TIC assessments on DCE-MRI, and minimum ADC values on DWI to differentiate benign and malignant NMEs. This model showed high sensitivity and low specificity in both the training (0.942, 0.589) and validation (0.940, 0.545) cohorts. To investigate the added value of the radiomics signature for NME differentiation, we added radiomics features derived from early phase DCE-MRI to the clinical model and built the combined model. The combined model achieved a higher specificity in the training (0.839) and validation (0.864) cohorts.

For the morphologic analysis, we used early phase images after contrast agent injection for NME evaluations because NMEs can be affected and obscured by more pronounced BPEs on the delayed phase images (29). Remarkably, although morphologic assessments, including distribution and internal enhancement patterns, were reported effective in previous studies (13–15), our study demonstrated that these morphologic features were not independently associated with NME differentiation, which is consistent with the results of a study by Naoko Mori et al. (10). Conversely, this lack of an independent association with morphologic features could be explained by decision-making pitfalls caused by the subjective judgment of visual examinations and by the variance of morphologic proportions contained in different study cohorts. In China, this can happen because the national breast cancer screening program is largely lacking compared with other countries; therefore, the lesions in the cohort of our study had larger sizes and a higher proportion of regional distributions and heterogeneous enhancement patterns. Thus, considering the potential role and subjective nature of morphologic assessments, we drew ROIs covering the whole lesion in each image plane and investigated the performance of the radiomics signatures alone, achieving a high sensitivity (82.7%) and specificity (80.4%). Of the six selected radiomics features, the surface area to volume ratio was given a maximum negative correlation (-0.594); lower ratios indicated a greater likelihood of NME malignancy, which is hard to identify with the human eye. Overall, these results indicated an important role for morphologic assessments in differentiating benign and malignant NMEs. However, it also indicated that histological patterns enrolled in the study may impact on the sensitivity and specificity of the model. The number of lesions in this study is

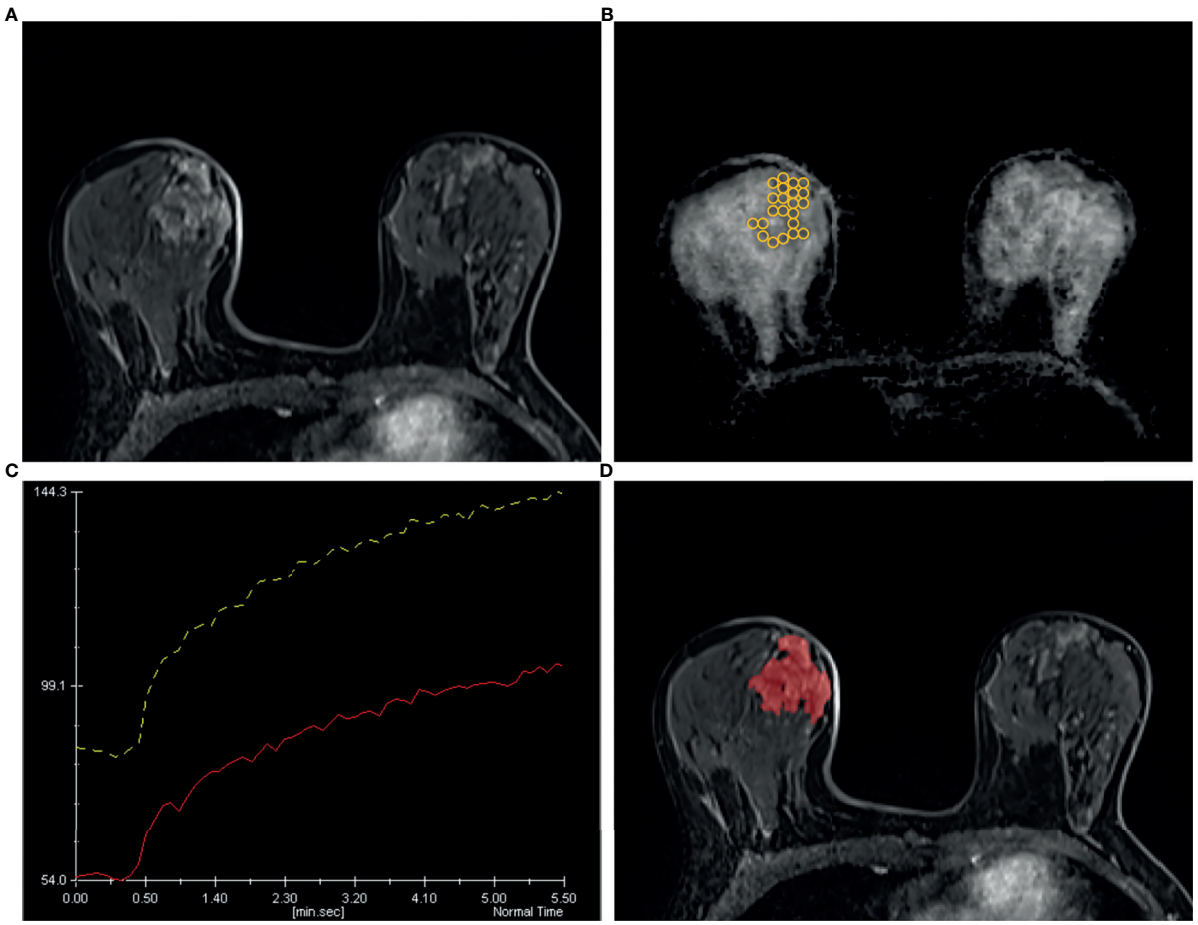


FIGURE 6 | A 49 years old woman diagnosed as BIRADS 4 preoperatively by radiologists and confirmed as adenosis by operation. **(A)** Axial dynamic contrast-enhancement images obtained in the very early phase (about 90 seconds) show a non-mass enhancement lesion with segmental distribution in the right breast. **(B)** On the ADC map, multiple ROIs are placed to cover the whole area of the lesion. The ADC map shows the minimum ADC value of the ROIs is 1056 $\times 10^{-6}$ mm^2/s. **(C)** After drawing the TIC curves for all ROIs at the brightest part on each slice, the high-level TIC curve type of this lesion is persistent type. **(D)** Using the ITK-SNAP software, the whole lesion was segmented. Finally, the logistic regression equation of the combined model for this lesion was calculated as 0.669, which was lower than the cut-off value 0.772 and adjudicated as benign lesion, consistent with the pathological results.

relatively small, and further research should be undertaken in a large cohort to investigate the impact of different histological patterns on the differentiation performance of the model.

A previous study observed that minimum ADC values potentially suggested the presence of an invasive component in ductal carcinoma *in situ* (DCIS) (30). In our study, we applied the same approach for malignant component detection. To perform this approach, we assumed that the area with minimum ADC values corresponded to the region with the highest tumor cell density, reflecting malignancy. However, we demonstrated that malignant lesions had significantly lower minimum ADC values than benign lesions. The multivariate analysis indicated that the minimum ADC value was not an independent factor for the discrimination of benign and malignant lesions, suggesting a limited role for DWI. These results are consistent with those of some recent studies (9, 31).

Naoko Mori et al. reported that kinetic assessments might be more important than the morphologic assessments in differentiating benign from malignant NMEs on the ultrafast DCE-MRI (10). In this study, we employed a similar ultrafast DCE-MRI approach and achieved similar results. Comparatively, malignant lesions tended to have more neovascularization (32). Thus, it is reasonable to set the ROI on the brightest areas of the images during the very early phase after contrast injection to obtain TIC curves. The selection of higher TIC curve types could provide greater detection of malignant components in the lesion enhancements. The TIC type alone gave a higher sensitivity (94.2%) and lower specificity (58.9%) for NME differentiation.

Our results showed that MR-reported ALN alone offered a higher specificity (92.9%) and lower sensitivity (36.6%) than conventional DCE-MRI assessments, which could be explained since less axillary lymphadenopathy was detected on the MRI images of most patients with malignant or benign lesions in this study. However, this situation was not consistent with what is seen in clinical practice.

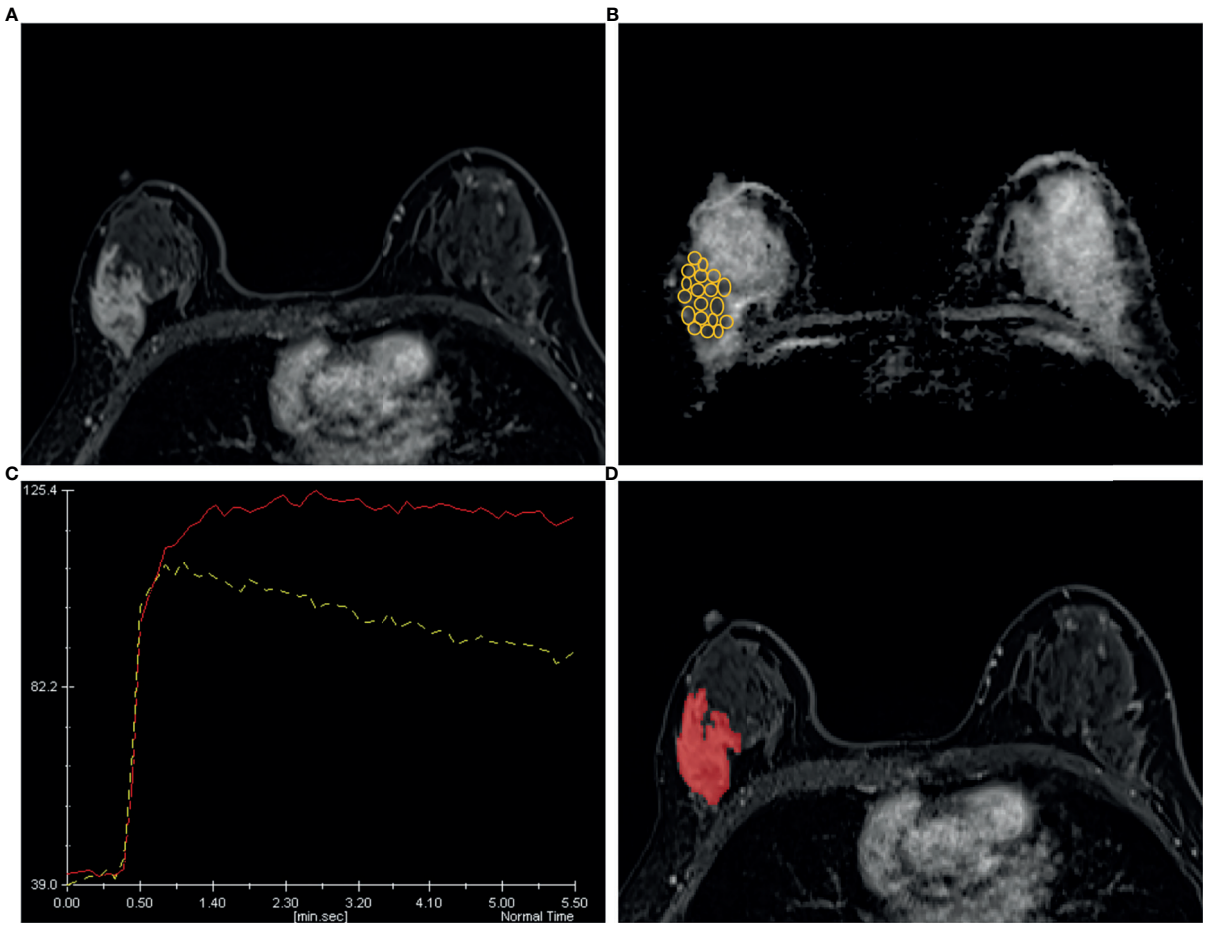


FIGURE 7 | A 44 years old woman diagnosed as BIRADS 4b preoperatively by radiologists and confirmed as invasive ductal carcinoma by operation. **(A)** Axial dynamic contrast-enhancement images obtained in the very early phase (about 90 seconds) show a non-mass enhancement lesion with segmental distribution in the right breast. **(B)** The ADC map shows the minimum ADC value of the ROIs is 745 $\times 10^{-6}$ mm^2/s. **(C)** After drawing the TIC curves for all ROIs at the brightest part on each slice, the high-level TIC curve type of this lesion is washout type. **(D)** Using the ITK-SNAP software, the whole lesion was segmented. Finally, the logistic regression equation of the combined model for this lesion was calculated as 0.989, which was higher than the cut-off value 0.772 and adjudicated as malignant lesion, consistent with the pathological results.

The analysis of low specificity showed that moderate or marked BPE, plateau or washout TIC, and MIP positive status may be prone to yield false positive results for NMEs in the conventional clinical analysis. It further indicated the difficulty and complexity of differentiation in clinical practice. Finally, the combined model of clinical features with added radiomics signature features improved the specificity in both the training (0.839) and validation cohorts (0.864). Given the comparable proportion of benign and malignant lesions and the good agreement between observers, the improved performance indicated that the radiomics signature was robust for the differentiation of benign and malignant NME lesions. The nomogram was primarily used to improve personalized diagnostics. The results of our study might suggest that additional radiomics signatures could help improve the specificity of differentiating benign and malignant NME lesions and avoid unnecessary biopsies. However, further studies with larger sample sizes are needed.

There were several limitations in our study. A primary limitation was the retrospective nature of the analysis, making potential selection bias difficult to avoid. Second, most of the patients in our hospital underwent breast MRI scans for two possible indications; preoperative staging for known breast cancer and further scanning for suspicious lesions in high-risk patients. Thus, the proportion of malignant lesions in our cohort was high, and there was a difference in the malignant/benign ratio between the training and validation cohorts. Third, the morphologic assessments and parameter measurements were accomplished by two radiologists using a consensus, and further research is needed to validate the repeatability of inter- and intra-observer. Fourth, the maximal diameters and morphologic assessments were recorded in the early phase to avoid being affected by BPEs; thus, some lesions with progressive enhancements might not have been evaluated accurately. Optimal timing needs to be determined in future studies.

In conclusion, the clinical multivariate regression analysis indicated that TIC patterns and ALN status were independent factors for the differentiation of benign and malignant NME lesions. Our results demonstrated that a radiomics nomogram combining clinical factors with radiomics signatures derived from early phase DCE-MRI could achieve high sensitivity and specificity for NME differentiation. Additional radiomics signatures could be used to improve specificity and avoid unnecessary biopsies. We believe that our model may not substitute but could improve conventional diagnostic workflow. However, a more extensive analysis with large samples is needed.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The ethics committee of Tongji Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

YL, ZY, TA, and LX participated in the conception and design of the study. YL, YQ, and CT collected the clinical and imaging data. YL and WL performed the statistical analyses. YL, ZY, XY, YG, TA, and LX coordinated, drafted, revised and finalized the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

1. *ACR BI-RADS® ATLAS — BREAST MRI* (2013). Available at: https://www.acr.org/-/media/ACR/Files/RADS/BI-RADS/MRI-Reporting.pdf.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, 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(3):209–49. doi: 10.3322/caac.21660
3. Tan Y, Mai H, Huang Z, Zhang L, Li C, Wu S, et al. Additive Value of Texture Analysis Based on Breast MRI for Distinguishing Between Benign and Malignant non-Mass Enhancement in Premenopausal Women. *BMC Med Imaging* (2021) 21(1):48.
4. Dratwa C, Jalaguier-Coudray A, Thomassin-Piana J, Gonin J, Chopier J, Antoine M, et al. Breast MR Biopsy: Pathological and Radiological Correlation. *Eur Radiol* (2016) 26(8):2510–9. doi: 10.1007/s00330-015-4071-y
5. Torous VF, Resteghini NA, Phillips J, Dialani V, Slanetz PJ, Schnitt SJ, et al. Histopathologic Correlates of Nonmass Enhancement Detected by Breast Magnetic Resonance Imaging. *Arch Pathol Lab Med* (2021). doi: 10.5858/arpa.2020-0266-OA
6. Thomassin-Naggara I, Trop I, Chopier J, David J, Lalonde L, Darai E, et al. Nonmasslike Enhancement at Breast MR Imaging: The Added Value of Mammography and US for Lesion Categorization. *Radiology* (2011) 261(1):69–79. doi: 10.1148/radiol.11110190
7. Asada T, Yamada T, Kanemaki Y, Fujiwara K, Okamoto S, Nakajima Y. Grading System to Categorize Breast MRI Using BI-RADS 5th Edition: A Statistical Study of non-Mass Enhancement Descriptors in Terms of Probability of Malignancy. *Jpn J Radiol* (2018) 36(3):200–8. doi: 10.1007/s11604-017-0717-9
8. Aydin H. The MRI Characteristics of non-Mass Enhancement Lesions of the Breast: Associations With Malignancy. *Br J Radiol* (2019) 92(1096). doi: 10.1259/bjr.20180464
9. Avendano D, Marino MA, Leithner D, Thakur S, Bernard-Davila B, Martinez DF, et al. Limited Role of DWI With Apparent Diffusion Coefficient Mapping in Breast Lesions Presenting as non-Mass Enhancement on Dynamic Contrast-Enhanced MRI. *Breast Cancer Res* (2019) 21(1):136. doi: 10.1186/s13058-019-1208-y
10. Mori N, Sheth D, Abe H. Nonmass Enhancement Breast Lesions: Diagnostic Performance of Kinetic Assessment on Ultrafast and Standard Dynamic Contrast-Enhanced MRI in Comparison With Morphologic Evaluation. *AJR Am J Roentgenol* (2020) 215(2):511–8. doi: 10.2214/AJR.19.21920
11. Yabuuchi H, Matsuo Y, Kamitani T, Setoguchi T, Okafuji T, Soeda H, et al. Non-Mass-Like Enhancement on Contrast-Enhanced Breast MR Imaging: Lesion Characterization Using Combination of Dynamic Contrast-Enhanced and Diffusion-Weighted MR Images. *Eur J Radiol* (2010) 75(1):e126–132. doi: 10.1016/j.ejrad.2009.09.013
12. Baltzer PA, Benndorf M, Dietzel M, Gajda M, Runnebaum IB, Kaiser WA. False-Positive Findings at Contrast-Enhanced Breast MRI: A BI-RADS Descriptor Study. *AJR Am J Roentgenol* (2010) 194(6):1658–63. doi: 10.2214/AJR.09.3486
13. Goto M, Ito H, Akazawa K, Kubota T, Kizu O, Yamada K, et al. Diagnosis of Breast Tumors by Contrast-Enhanced MR Imaging: Comparison Between the Diagnostic Performance of Dynamic Enhancement Patterns and Morphologic Features. *J Magn Reson Imaging* (2007) 25(1):104–12. doi: 10.1002/jmri.20812
14. Uematsu T, Kasami M. High-Spatial-Resolution 3-T Breast MRI of Nonmasslike Enhancement Lesions: An Analysis of Their Features as Significant Predictors of Malignancy. *AJR Am J Roentgenol* (2012) 198(5):1223–30. doi: 10.2214/AJR.11.7350
15. Newell D, Nie K, Chen JH, Hsu CC, Yu HJ, Nalcioglu O, et al. Selection of Diagnostic Features on Breast MRI to Differentiate Between Malignant and Benign Lesions Using Computer-Aided Diagnosis: Differences in Lesions Presenting as Mass and non-Mass-Like Enhancement. *Eur Radiol* (2010) 20(4):771–81. doi: 10.1007/s00330-009-1616-y
16. Illan IA, Ramirez J, Gorriz JM, Marino MA, Avendano D, Helbich T, et al. Automated Detection and Segmentation of Nonmass-Enhancing Breast Tumors With Dynamic Contrast-Enhanced Magnetic Resonance Imaging. *Contrast Media Mol Imaging* (2018). doi: 10.1155/2018/5308517
17. Gutierrez RL, DeMartini WB, Eby PR, Kurland BF, Peacock S, Lehman CD. BI-RADS Lesion Characteristics Predict Likelihood of Malignancy in Breast MRI for Masses But Not for Nonmasslike Enhancement. *AJR Am J Roentgenol* (2009) 193(4):994–1000. doi: 10.2214/AJR.08.1983
18. Sakamoto N, Tozaki M, Higa K, Tsunoda Y, Ogawa T, Abe S, et al. Categorization of non-Mass-Like Breast Lesions Detected by MRI. *Breast Cancer* (2008) 15(3):241–6. doi: 10.1007/s12282-007-0028-6
19. Pinker K, Chin J, Melsaether AN, Morris EA, Moy L. Precision Medicine and Radiogenomics in Breast Cancer: New Approaches Toward Diagnosis and Treatment. *Radiology* (2018) 287(3):732–47. doi: 10.1148/radiol.2018172171
20. Shao Z, Wang H, Li X, Liu P, Zhang S, Cao S. Morphological Distribution and Internal Enhancement Architecture of Contrast-Enhanced Magnetic Resonance Imaging in the Diagnosis of non-Mass-Like Breast Lesions: A Meta-Analysis. *Breast J* (2013) 19(3):259–68. doi: 10.1111/tbj.12101
21. Donn W, Rebbeck P, Wilson C, Gilks CB. Idiopathic Granulomatous Mastitis. A Report of Three Cases and Review of the Literature. *Arch Pathol Lab Med* (1994) 118(8):822–5.
22. Zhao Q, Xie T, Fu C, Chen L, Bai Q, Grimm R, et al. Differentiation Between Idiopathic Granulomatous Mastitis and Invasive Breast Carcinoma, Both Presenting With non-Mass Enhancement Without Rim-Enhanced Masses:

The Value of Whole-Lesion Histogram and Texture Analysis Using Apparent Diffusion Coefficient. *Eur J Radiol* (2020) 123:108782. doi: 10.1016/j.ejrad.2019.108782
23. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
24. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding Tumour Phenotype by Noninvasive Imaging Using a Quantitative Radiomics Approach. *Nat Commun* (2014) 5:4006. doi: 10.1038/ncomms5006
25. Li Z, Ai T, Hu Y, Yan X, Nickel MD, Xu X, et al. Application of Whole-Lesion Histogram Analysis of Pharmacokinetic Parameters in Dynamic Contrast-Enhanced MRI of Breast Lesions With the CAIPIRINHA-Dixon-TWIST-VIBE Technique. *J Magn Reson Imaging* (2018) 47(1):91–6. doi: 10.1002/jmri.25762
26. Sauerbrei W, Royston P, Binder H. Selection of Important Variables and Determination of Functional Form for Continuous Predictors in Multivariable Model Building. *Stat Med* (2007) 26(30):5512–28. doi: 10.1002/sim.3148
27. Kramer AA, Zimmerman JE. Assessing the Calibration of Mortality Benchmarks in Critical Care: The Hosmer-Lemeshow Test Revisited. *Crit Care Med* (2007) 35(9):2052–6. doi: 10.1097/01.CCM.0000275267.64078.B0
28. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to Decision Curve Analysis, a Novel Method for Evaluating Diagnostic Tests, Prediction Models and Molecular Markers. *BMC Med Inform Decis Mak* (2008) 8:1–17. doi: 10.1186/1472-6947-8-53
29. Chikarmane SA, Michaels AY, Giess CS. Revisiting Nonmass Enhancement in Breast MRI: Analysis of Outcomes and Follow-Up Using the Updated BI-RADS Atlas. *AJR Am J Roentgenol* (2017) 209(5):1178–84. doi: 10.2214/AJR.17.18086
30. Mori N, Ota H, Mugikura S, Takasawa C, Tominaga J, Ishida T, et al. Detection of Invasive Components in Cases of Breast Ductal Carcinoma in Situ on Biopsy by Using Apparent Diffusion Coefficient MR Parameters. *Eur Radiol* (2013) 23(10):2705–12. doi: 10.1007/s00330-013-2902-2
31. Clauser P, Krug B, Bickel H, Dietzel M, Pinker K, Neuhaus VF, et al. Diffusion-Weighted Imaging Allows for Downgrading MR BI-RADS 4 Lesions in Contrast-Enhanced MRI of the Breast to Avoid Unnecessary Biopsy. *Clin Cancer Res* (2021) 27(7):1941–8. doi: 10.1158/1078-0432.CCR-20-3037
32. Hulka CA, Smith BL, Sgroi DC, Tan L, Edmister WB, Semple JP, et al. Benign and Malignant Breast Lesions: Differentiation With Echo-Planar MR Imaging. *Radiology* (1995) 197(1):33–8. doi: 10.1148/radiology.197.1.7568850

Conflict of Interest: Author XY and YG were employed by the Siemens Healthcare company. WL was employed by Julei Technology Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 27 October 2021
doi: 10.3389/fonc.2021.722624

Knockdown of Oligosaccharyltransferase Subunit Ribophorin 1 Induces Endoplasmic-Reticulum-Stress-Dependent Cell Apoptosis in Breast Cancer

Jiajun Ding[1,2], *Jiahui Xu*[1], *Qiaodan Deng*[1], *Wei Ma*[1], *Rui Zhang*[1], *Xueyan He*[1], *Suling Liu*[1*] and *Lixing Zhang*[1*]

[1] Fudan University Shanghai Cancer Center & Institutes of Biomedical Sciences, Cancer Institutes, Key Laboratory of Breast Cancer in Shanghai, The Shanghai Key Laboratory of Medical Epigenetics, Key Laboratory of Medical Epigenetics and Metabolism, Shanghai Medical College, Fudan University, Shanghai, China, [2] Breast Surgery, Obstetrics and Gynecology Hospital of Fudan University, Shanghai, China

OPEN ACCESS

Edited by:
Jay Brewster,
Pepperdine University, United States

Reviewed by:
Vidya Sethunath,
Dana–Farber Cancer Institute, United States
Lin Zhang,
National Institutes of Health (NIH), United States

***Correspondence:**
Lixing Zhang
Zhang_lx@fudan.edu.cn
Suling Liu
suling@fudan.edu.cn

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: 09 June 2021
Accepted: 07 October 2021
Published: 27 October 2021

Citation:
Ding J, Xu J, Deng Q, Ma W, Zhang R, He X, Liu S and Zhang L (2021) Knockdown of Oligosaccharyltransferase Subunit Ribophorin 1 Induces Endoplasmic-Reticulum-Stress-Dependent Cell Apoptosis in Breast Cancer. Front. Oncol. 11:722624. doi: 10.3389/fonc.2021.722624

Ribophorin 1 (RPN1) is a major part of Oligosaccharyltransferase (OST) complex, which is vital for the N-linked glycosylation. Though it has been verified that the abnormal glycosylation is closely related to the development of breast cancer, the detail role of RPN1 in breast cancer remains unknown. In this study, we explored the public databases to investigate the relationship between the expression levels of OST subunits and the prognosis of breast cancer. Then, we focused on the function of RPN1 in breast cancer and its potential mechanisms. Our study showed that the expression of several OST subunits including RPN1, RPN2, STT3A STT3B, and DDOST were upregulated in breast cancer samples. The protein expression level of RPN1 was also upregulated in breast cancer. Higher expression of RPN1 was correlated with worse clinical features and poorer prognosis. Furthermore, knockdown of RPN1 suppressed the proliferation and invasion of breast cancer cells *in vitro* and induced cell apoptosis triggered by endoplasmic reticulum stress. Our results identified the oncogenic function of RPN1 in breast cancer, implying that RPN1 might be a potential biomarker and therapeutic target for breast cancer.

Keywords: breast cancer, ribophorin 1, oligosaccharyltransferase complex, endoplasmic reticulum stress, apoptosis

INTRODUCTION

Breast cancer (BC) is one of the leading causes for the mortality of women all over the world. It accounted for 24.2% of the 8.6 million new cases of female cancer and 15.0% of 4.2 million cancer-related deaths in women worldwide in 2018 (1). BC is a complex and heterogeneous disease. Four major intrinsic molecular subtypes, which are Luminal A, Luminal B, HER2-enriched, and basal-like breast cancer (BLBC), have been identified (2). Among which, BLBC is recognized as the worst subtype due to the lack of effective treatment. Although great progress has been made in the diagnosis and treatment of BC, finding new targets for early diagnosis and treatment remains a

challenge. Recently, large accessible databases like Oncomine have become efficient and economic tools for identifying targets for BC (3, 4). And they may play an important role in identifying novel genes associated with BC.

N-linked glycosylation is a vital protein modification in eukaryotic cells. Proteins are N-glycosylated in the endoplasmic reticulum lumen by Oligosaccharyltransferase (OST) complex. Although the exact structure of OST in eukaryotes is largely unknown, it has been found that OST complex consists 12 subunits, including STT3 OST complex catalytic subunit A and B (STT3A, STT3B), Ribophorin 1 (RPN1), Ribophorin 2 (RPN2), dolichyl-diphosphooligosaccharide-protein glycosyltransferase (DDOST), defender against cell death 1 (DAD1), oligosaccharyl-transferase complex subunit 4 (OST4), transmembrane protein 258 (TMEM258), oligosaccharyltransferase complex (OSTC) and keratinocyte associated protein 2 (KRTCAP2), magnesium transporter 1 (MAGT1), and tumor suppressor candidate 3 (TUSC3) (5–8).

The abnormality of OST subunits can lead to the hypoglycosylation of proteins, which account for the misfolding of proteins. The accumulation of misfolded proteins would affect the homeostasis of endoplasmic reticulum, ultimately inducing an imbalance between protein folding load and capacity. This abnormality is known as endoplasmic reticulum stress (ERS) (9), which is associated with the development and prognosis of cancers (10–13). At first, ERS initiates unfolded protein response (UPR) to improve the adaptability and reestablish the homeostasis. With the persistent ERS, the UPR could turn from a pro-survival to a pro-death response, playing a biswitch role in homeostasis maintenance (14).

RPN1, which is only found in the rough endoplasmic reticulum, facilitates the N-glycosylation by selecting the specific substrates (15). Though it is a critical subunit of OST, the association between RPN1 and cancers has rarely been reported. In this study, we analyzed the relationship between OST subunits, especially RPN1, and BC by several accessible databases, and then explored the effects of RPN1 knockdown on the proliferation, migration, and invasion of BC cells. Finally, we found that the ERS-induced cell apoptosis was responsible for the inhibition of cell proliferation and invasion after RPN1 knockdown.

MATERIALS AND METHODS

Oncomine Database Analysis

Expression level of the OST subunits in various cancer types was retrieved from Oncomine (http://www.oncomine.org, accessed on February 28, 2019) (3). Thresholds were set as the following: p-value: 0.0001; fold change: 1.5; gene rank: top 10%; and data type: mRNA. After analyzing the mRNA expression level in different cancers, we additionally performed a meta-analysis with the providing 13 datasets, which contained 43 analyses of 3,555 samples on different kinds of BC (**Supplementary Table S1**), aiming to compare the over-expression variation of different subunits. p-value<0.01 was considered statistically significant.

BC Gene-Expression Miner v4.5 Analysis

Bc-GenExMiner v4.5 (bcgenex.centregauducheau.fr/, accessed on August 3, 2020) was used to measure the correlation between the OST subunits and the clinicopathologic features in BC (4). P-value<0.01 was considered statistically significant. The Pearson correlation coefficient between the expression level of candidate genes and RPN1 was computed to determine the co-expressed genes of RPN1. We identified genes as the co-expressed genes of RPN1 when the Pearson correlation coefficient > 0.4.

Survival Analysis

Kaplan-Meier Plotter (kmplot.com, accessed on March 3, 2019) was used to identify the prognostic genes among OST subunits in BC (16). We also identified the prognostic genes among OST subunits in each subtype of BC, and the subtype of BC was determined by the 2013 St Gallen criteria. The patients were divided into two groups (high expression and low expression) by the median value of gene expression level, and only the best probe for each gene was selected. The hazard ratio (HR) with 95% confidence intervals (CI) and log-rank Genes with P-value<0.05 was considered as prognostic genes.

The Cancer Genome Atlas and Gene Expression Omnibus Database Analysis

The TCGA and GEO database (GSE42568) were used to explore the expression of RPN1 in BC tissues and normal breast tissues or para-tumor tissues. The expression level of RPN1 in each subtype was also analyzed in TCGA and GEO database (GSE47561). In addition, using the transcriptome data from TCGA, we evaluated the co-expression level between two of OST subunits by custom R scripts.

GeneMANIA Analysis

As a prediction server for gene prioritization and predicting gene function (17), GeneMANIA database (http://genemania.org/, accessed on January 17, 2020) was used in our study to construct an interactive functional-associated network for OST subunits in terms of physical interactions, predictions, pathways, shared protein domains, co-expression, co-localization, and genetic interactions, as well as to find their functions.

The Human Protein Atlas Database Analysis

We used the HPA (https://www.proteinatlas.org/, accessed on December 19, 2019) to explore the immunohistochemical (IHC) staining of RPN1 (18–20). The images of normal breast tissues were gotten from the TISSUE ALTAS, while the images of BC tissues were gotten from the PATHOLOGY ALTAS. Both normal breast and BC tissues were stained by antibody CAB009748.

UALCAN Database Analysis

UALCAN (http://ualcan.path.uab.edu/, accessed on December 20, 2019) is an interactive database for analyzing cancer omics data, including TCGA data and the Clinical Proteomic Tumor Analysis Consortium (CPTAC) data (21). We used UALCAN to

analyze the protein level of RPN1 in BC tissues compared to the normal breast tissues in CPTAC samples and the methylation level on the promoter region of RPN1 in TCGA samples.

Functional Enrichment Analysis

Gene Ontology (GO) enrichment analysis for gene lists from BcGenExMiner v4.5 database was conducted using the R package "clusterProfiler", "org.Hs.eg.db", "enrichplot" (https://bioconductor.org/, accessed on January 18, 2020), and "ggplot2" (https://cran.rproject.org/web/packages/, accessed on January 18, 2020). Only the top five significant enriched GO terms were plotted.

Cell Culture and Reagents

The human BC cell lines SUM149 and SUM159 (purchased from Asterland Bioscience, MI, USA) were confirmed without mycoplasma and then cultured in Han's F12 medium with 5% fetal bovine serum (FBS, Thermo Fisher), 1% streptomycin/penicillin (Beyotime), 1 mg/ml hydrocortisone (Sigma-Aldrich), 10 ug/ml gentamicin (Life Technologies), and 5 mg/ml insulin (Sigma-Aldrich). All cells were incubated under 37°C with 5% CO_2. Sodium phenylbutyrate (4-PBA) was purchased from MCE and dissolved in DMSO.

Virus Infection and Cell Lines Construction

The effective sequences of shRNAs were bought from Sigma-Aldrich (**Supplementary Table S2**). The RPN1 knockdown lentiviruses were produced by transfecting 293T cell in the University of Michigan Vector Core Facility. SUM159 and SUM149 cells were infected in the presence of polybrene (8 ug/ml, Millipore) for 24 h, then the medium was discharged and replaced with the fresh medium. And knockdown cells were selected by Puromycin (Invitrogen) for 14 days.

RNA Extraction and Real-Time qRT-PCR

Total RNA was extracted using Trizol (Takara) and reverse-transcribed into cDNA with the HiScript II 1st Strand cDNA Synthesis Kit (Vazyme Biotech). The primers for qRT-PCR were provided in **Supplementary Table S3**. And qRT-PCR was carried out using AceQ qPCR SYBR Green Master Mix (Vazyme Biotech) in a real-time PCR system (7300, Applied Biosystem). TATA-box binding protein (TBP) was used as a reference gene.

MTT Assay

One thousand cells of SUM159 and 3,000 cells of SUM149 were seeded in per well of 96-well plates and cultured for 1 day for eliminating the counting error. Two hundred cells of SUM159 and 500 cells of SUM149 were did the same at the same time but cultured for 3, 5, and 7 days. Then 20 ul MTT (5 mg/ml, Biosharp) was added in each well, and the plates were incubated at 37°C for 4 h. After removing the supernatant, 100 ul DMSO was added in per well, and the optical density (OD) was measured at 490 nm with microplate reader (Elx800, BioTek). Each group had six parallel wells and was performed in triplicate.

Colony Formation Assay

One thousand cells of SUM159 and 3,000 cells of SUM149 were seeded and cultured in six-well plates under 37°C for 2 weeks. Ten percent formaldehyde was used for fixing for 30 min, and the cell colonies were stained with 0.1% crystal violet for another 30 min. After washing and drying, the number of colonies was calculated. Each group had three parallel wells and was performed in triplicate.

Wound Healing Test

One million cells of SUM159 were seeded in six-well plates and grew to approximately total confluence. Then the wounds were created by a 200 ul pipette tip. The wells were washed by PBS for two times, and none-serum medium was added. Wound healing within the scrape lines were then observed and photographed at 0, 6, 18, and 24 h. Each group had more than three parallel positions and was performed in triplicate.

Invasion Assay

Transwell chambers (#3422, Corning, USA) precoated with matrigel (354234, Corning, USA) were placed in 24-well plates at 37°C for 4 h. Then 5×10^4 cells of SUM159 were plated on chambers without serum and medium containing 5% FBS offered in the bottom well. After 36 h of incubation in normal condition, the chambers were fixed (methyl alcohol: glacial acetic acid = 3:1) and stained with 0.1% crystal violet. After washing and drying, the invaded cells were photographed for statistical analysis. Each group had three parallel wells and was performed in triplicate.

Western Blot

Cells were lysed in RPRA buffer (Beyotime, China), and protein concentration was measured by BCA Kit (Pierce, USA). Protein samples were separated by SDS-PAGE and subsequently transferred onto PVDF membranes. Membrane was blocked in 5% de-fat milk and incubated with primary antibody at 4°C for a night and sequentially HRP-conjugated secondary antibody at room temperature for 1 h. ImageQuant LAS 4000 mini imaging system (GE, Fairfield, USA) and Western HRP Substrate (WBLUF0500, Millipore) were used in chemiluminescent detection. The antibody used in this study are as following: anti-GAPDH (M017, TransGen), anti-PERK (5683, CST), anti-IRE1α (3294, CST), anti-ATF6 (24169-1-AP, Proteintech), anti-BiP (3177, CST), anti-Bax (2774, CST), anti-Bcl-2 (4223, CST), goat anti-mouse IgG-HRP (sc-2005, Santa Cruz), and goat anti-rabbit IgG-HRP (sc-2004, Santa Cruz).

Flow Cytometry

For apoptosis analysis, the cells were stained with annexin V-fluorescein isothiocyanate (BD) and propidium iodide (Sigma-Aldrich) for 15 min at room temperature. The stained cells were measured by flow cytometer according to the manufacturer's instructions. For cell cycle analysis, cells were fixed in 70% ethyl alcohol at 4°C for 6 h and then stained with propidium iodide containing 1% RNase A (Takara) at 37°C for 30 min. CytoFLEX (Beckman Coulter) was used for the detection and acquisition of data, and the analysis was performed in CytoExpert software.

Statistical Analysis

All data were present as the mean ± standard deviation. Difference between two groups was analyzed by Student t test with GraphPad Prism 6. Some results of statistical analysis were download from the websites directly. $P<0.05$ was considered statistically significant unless otherwise indicated.

RESULTS

Correlations of Transcriptional Expression Among OST Subunits and Construction of a Protein-Protein Interaction Network

According to the transcriptional data from TCGA database, the Pearson correlations among OST subunits in BC patients were analyzed (**Figure 1**), and Pearson correlation coefficient exceeding 0.40 indicated a good correlation. It could be found that there was a significant positive correlation between RPN1 and DDOST, STT3B and MAGT1, TMEM258 and OSTC, as well as among OST4, TMEM258, and KRTCAP2.

Then, GeneMANIA database was used to construct a protein-protein interaction network for the OST subunits and to analyze their potential functions (**Supplementary Figure S1**). The 12 central nodes represented the OST subunits, and the 20 nodes surrounding represented the top 20 genes that correlated to the OST subunits in terms of physical interactions, predictions, pathways, shared protein domains, co-expression, co-localization, and genetic interactions. And further functional analysis showed that the 12 central genes we focused on were as expected greatest related to the OST complex and the function of glycosylation.

mRNA Expression Profiles of the Subunits of OST in BC Patients

We first analyzed the mRNA expression level of OST subunits in different human cancers, especially in BC, compared to the normal breast tissues (**Figure 2**). Analyses that met the threshold were listed in **Supplementary Table S4**. According to **Figure 2** and **Supplementary Table S4**, it could be found that the expression level of RPN1 was upregulated in various subtypes of breast cancer including invasive breast carcinoma, mucinous breast carcinoma, medullary breast carcinoma, invasive ductal breast carcinoma, and ductal breast carcinoma in Curtis's dataset (22). And the mRNA expression of RPN1 in ductal breast carcinoma was 1.684-fold higher than normal tissues in Sorlie's

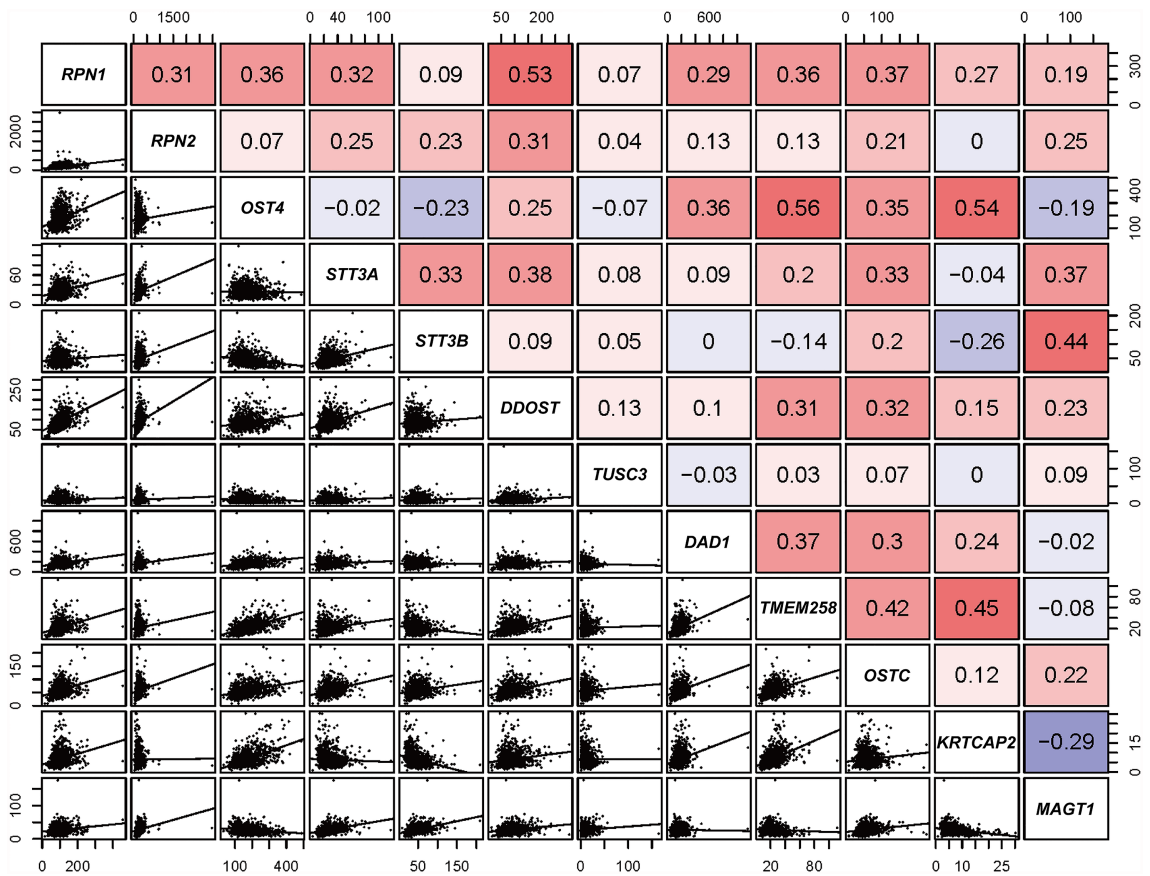


FIGURE 1 | Correlation analysis of OST family members (data from TCGA). Pearson's correlation of OST subunits. In the upper right, red and blue cells represent positive and negative relationship, respectively. In the lower left, correlation scatter diagram of the two genes is listed.

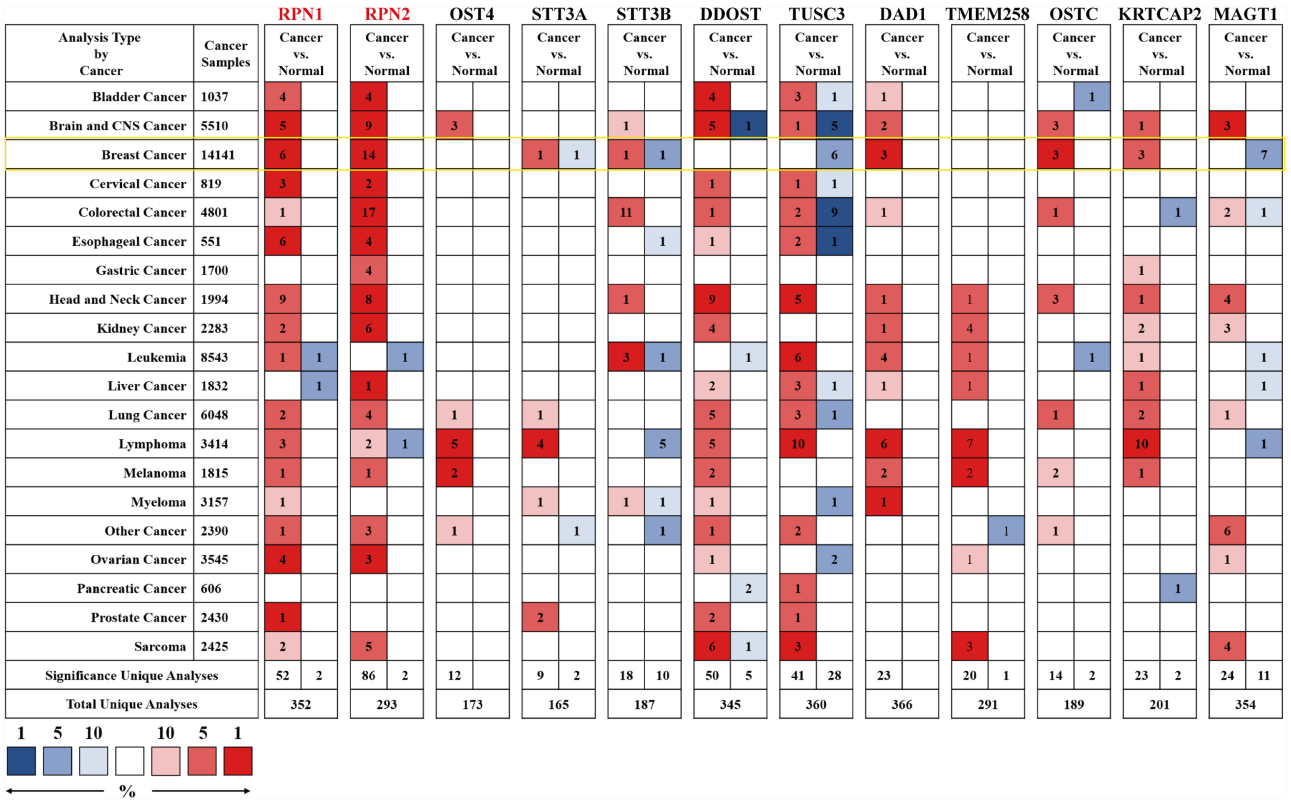


FIGURE 2 | The mRNA expression levels of the OST subunits in different types of human cancers (Oncomine). The figure was generated from Oncomine database with the thresholds that p-value, 0.0001; fold change, 1.5; gene rank: top 10%. The cell number represented the dataset number that met all of the thresholds with the color blue for low expression while the color red for high expression, and the cell color was determined by the best gene rank percentile for the analyses within the cell. An analysis might be counted in more than one cancer type. mRNA expression levels of OST subunits in breast cancer are delineated with yellow highlight. CNS, central nervous system.

dataset (23). The other OST subunit expression patterns were also analyzed in BC tissues. Higher mRNA expression levels of RPN2, DAD1, OSTC, KRTCAP2 and lower expression levels of TUSC3, MAGT1 could be found in different types of BC compared to the normal breast tissues in Curtis's dataset (22), Zhao's dataset (24), Ma's dataset (25), Finak's dataset (26), Karnoub's dataset (27), and TCGA dataset.

We carried out a meta-analysis by Oncomine and found that only RPN1, RPN2, STT3A, STT3B, and DDOST significantly upregulated in BC tissues according to the 43 analyses of 13 datasets (**Figure 3** and **Supplementary Table S1**).

The Relationships Between OST Subunits and the Clinicopathologic Features of BC

Then we analyzed the correlations between the mRNA expression of OST subunits especially RPN1, RPN2, STT3A, STT3B, and DDOST, and the clinicopathologic features of BC patients according to Bc-GenExMiner v4.5. The results were presented in **Table 1** and **Supplementary Table S5**. For age character, RPN2 (P = 0.0019) was found to have significantly higher expression in the group not more than 51 years old. For the patient samples with negative estrogen receptor (ER) status, the expression of RPN1, RPN2, STT3A, DDOST were upregulated. The expression levels of RPN1 (P<0.0001), RPN2 (P=0.0003), and DDOST (P<0.0001) were also significantly higher in the BC patient samples with negative progesterone receptor (PR) status. Moreover, compared to the patients with positive human epidermal growth factor receptor 2 (HER2) expression, the mRNA levels of RPN2 (P<0.0001) and STT3B (P<0.0001) were significantly upregulated in the negative ones. In addition, in the BC patients with nodal metastasis, only STT3B (P<0.0001) mRNA expression increased significantly.

Intrinsic molecular subtype is one of the most important clinicopathologic characteristics of BC. The expression levels of all five subunits were significantly higher in basal-like and HER2-enriched patients compared with Luminal A patients, while all five subunits except STT3A expressed higher in Luminal B patients than Luminal A ones. RPN1 and DDOST could be found upregulated in basal-like patients compared with Luminal B and HER2-enriched patients, while RPN2, STT3A, and STT3B were in the opposite. Additionally, it could be found that the expression level of RPN2, STT3A, STT3B, and DDOST increased significantly in Luminal B patients compared with HER2-enriched patients (**Table 1**, **Supplementary Figure S2**).

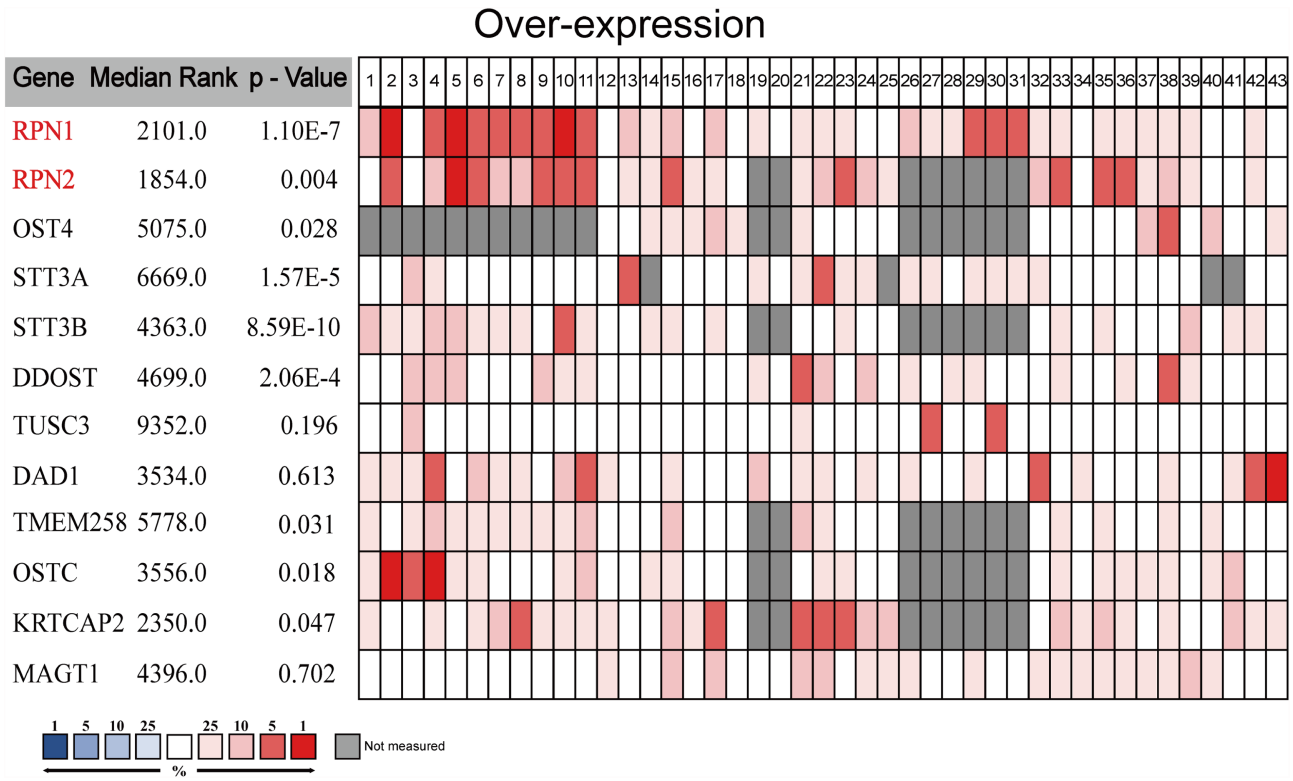


FIGURE 3 | The meta-analysis on the mRNA expression of OST subunits (Oncomine). The comparison of OST subunit mRNA expression in different datasets. Only the upregulated subunits were listed. The rank for a gene was the median rank for that gene across each of the analyses. The p-value for a gene was its p-value for the median-ranked analysis, and p<0.01 was considered statistically significant.

As the BLBC has the worst prognosis, we especially analyzed the five genes' mRNA expression between the BLBC and the non-BLBC patients. The expression levels of RPN1 (P<0.0001), STT3A (P<0.0001), and DDOST (P<0.0001) increased significantly in the BLBC compared with non-BLBC patients (**Table 1**, **Supplementary Figure S3**).

In BC, the Scarff Bloom & Richardson (SBR) grade is an important prognostic factor associated with the gland formation, the nuclear features, and the mitotic activity. The SBR is also correlated with poor clinical outcome (28, 29). As shown in **Table 1** and **Supplementary Figure S4**, higher mRNA expression levels of all five genes were associated with a higher SBR grade, while only RPN1, RPN2, STT3B, and DDOST were statistically significant (p<0.01) in all pairwise comparisons. The Nottingham Prognostic Index (NPI) is another system to evaluate the prognosis of BC after surgery, referring to the size of lesion, the number of lymph nodes involved, and the pathologic grade (30, 31). We found that higher expression levels of RPN1, RPN2, and STT3B were associated with higher NPI grade. The expression level of RPN1 was higher in NPI2 and NPI3 patients than in NPI1 patients, but there was no significant difference between NPI2 and NPI3 patients. The expression levels of RPN2 and STT3B increased only in NPI2 patients compared with NPI1 patients (**Supplementary Figure S5**). In summary, the high expression levels of RPN1, RPN2, and STT3B were associated with poor prognosis, suggesting their potential roles in BC.

Prognostic Values of OST Subunits Expression in BC

The prognostic values of all OST subunits in BC were listed in **Supplementary Table S6**, and **Supplementary Figure S6** showed the relapse-free survival (RFS) curves. As for the five selected genes (**Figure 4**), high expression of RPN1 (HR: 1.51, 95% CI: 1.35–1.69, P=1.20E-13), RPN2 (HR: 1.26, 95% CI: 1.13–1.40, P=3.60E-5), and STT3A (HR: 1.15, 95% CI: 1.03–1.28, P=0.013) were associated with worse RFS, while the expression of STT3B (P=0.11) and DDOST (P=0.69) showed no relationship with RFS. We also analyzed the correlation between mRNA expression level of all OST members and other prognostic indexes including overall survival (OS), distant metastasis-free survival (DMFS), and post-progression survival (PPS) (**Supplementary Table S6**, **Figure 5**). High expression level of RPN1 (HR: 1.35, 95% CI: 1.09–1.68, P=0.006) and RPN2 (HR: 1.49, 95% CI: 1.2–1.85, P=0.00031) indicated worse OS, while DDOST (HR: 0.8, 95% CI: 0.65–1.0, P=0045) was in the opposite. High expression level of DDOST (HR: 0.75, 95% CI: 0.61–0.91, P=0.0031) was associated with better DMFS.

We then analyzed the correlation between OST members and prognosis in different subtypes of BC (**Supplementary Table S7**

TABLE 1 | The relationship between the OST subunits and the clinicopathologic parameters of BC (bc-GenExMiner v4.5).

		RPN1		RPN2		STT3A		STT3B		DDOST	
		Comp.	P	Comp.	P	Comp.	P	Comp.	P	Comp.	P
Age	≤51		0.8184	↑	**0.0019**		0.1950		0.0578		0.4482
	>51										
ER (IHC)	Negative	↑	**<0.0001**	↑	**0.0020**	↑	**<0.0001**		0.2391	↑	**<0.0001**
	Positive										
PR (IHC)	Negative	↑	**<0.0001**	↑	**0.0003**		0.1242		0.3620	↑	**<0.0001**
	Positive										
HER2 (IHC)	Negative		0.0661	↑	**<0.0001**		0.1040	↑	**<0.0001**		0.0631
	Positive										
Nodal status	Negative		0.9624		0.1551	↑	**0.0289**		**<0.0001**		0.5209
	Positive							↑			
Intrinsic molecular subtypes	Total		**<0.0001**		**<0.0001**		**<0.0001**		**<0.0001**		**<0.0001**
	Basal-like *vs* Luminal A	>	**<0.0001**	>	**<0.0001**	>	**<0.0001**	>	**<0.0001**	>	**<0.0001**
	Basal-like *vs* Luminal B	>	**<0.0001**	<	**<0.0001**	>	**<0.0001**	<	**<0.0001**	>	**<0.0001**
	Basal-like *vs* HER2-E	>	**<0.05**	<	**<0.0001**	=	>0.01	<	**<0.0001**	>	**<0.0001**
	Luminal B *vs* Luminal A	>	**<0.0001**	>	**<0.0001**	<	**<0.001**	>	**<0.0001**	>	**<0.0001**
	Luminal B *vs* HER2-E	=	>0.01	<	**<0.0001**	<	**<0.0001**	<	**<0.001**	<	**<0.0001**
	HER2-E *vs* Luminal A	>	**<0.0001**	>	**<0.0001**	>	**<0.0001**	>	**<0.0001**	>	**<0.0001**
Basal-like status	Basal	↑	**<0.0001**		0.4492	↑	**<0.0001**		0.2047	↑	**<0.0001**
	None										
SBR	Total		**<0.0001**		**<0.0001**		**0.0012**		**<0.0001**		**<0.0001**
	SBR2 *vs* SBR1	>	**<0.0001**	>	**<0.0001**	=	>0.01	>	**<0.01**	>	**<0.01**
	SBR3 *vs* SBR1	>	**<0.0001**	>	**<0.0001**	>	**<0.01**	>	**<0.0001**	>	**<0.0001**
	SBR3 *vs* SBR2	>	**<0.0001**	>	**<0.0001**	>	**<0.01**	>	**<0.01**	>	**<0.0001**
NPI	Total		**<0.0001**		**0.0015**		0.2014		**0.0024**		0.0715
	NPI2 *vs* NPI1	>	**<0.001**	>	**<0.01**	=	>0.01	>	**<0.01**	=	>0.01
	NPI3 *vs* NPI1	>	**<0.001**	=	>0.01	=	>0.01	=	>0.01	=	>0.01
	NPI3 *vs* NPI2	=	>0.01	=	>0.01	=	>0.01	=	>0.01	=	>0.01

The data with statistical significance (P<0.01) were marked in bold text.

Comp, comparison; IHC, immunohistochemical; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; HER2-E, HER2-enriched; SBR, Scarff Bloom & Richardson grade; NPI, Nottingham Prognostic Index.

and **Table 2**). In Luminal A patients, high expression of RPN1 (P= 0.00024) and RPN2 (P=9.1E-7) indicated worse RFS. In Luminal B patients, high expression of RPN1 (P=0.025) and STT3A (P=0.028) predicted worse RFS. In HER2-enriched patients, high expression of RPN1(P=0.0093) indicated worse RFS while DDOST (P=0.049) indicated the opposite. In basal-like patients, high expression of RPN1 (P=0.038) and STT3A (P=0.0063) were significantly associated with worse RFS. In a word, these results implied that higher expression of most OST members, especially RPN1 and RPN2, were significantly correlated with poor prognostic outcome and might play a pro-tumor function.

RPN1 Is a Novel Prognostic Gene for BC

According to the above analyses, it could be concluded that RPN1 and RPN2 were the most influential subunits in BC progression due to their significant relationship between their expression level and clinical prognosis. However, the function of RPN2 in BC has been reported by several studies before (32–34). We focused on the function of RPN1 in BC. Therefore, we analyzed the expression status and prognostic value of RPN1 in BC deeply.

According to TCGA, GEO (GSE 42568), the HPA, and UALCAN database, both mRNA expression level and protein expression level of RPN1 could be found higher in BC tissues compared to the normal ones (**Figures 6A, C, D**). The details of the IHC figures of RPN1 in the HPA are listed in **Supplementary Table S8**. As mentioned above, BLBC has the worst prognosis. Our analyses showed that the mRNA expression level of RPN1 was the highest in BLBC tissues in TCGA database and the second highest in BLBC tissues in GEO database (GSE47561) (**Figure 6B**). Besides, the promoter methylation level of RPN1 in BC tissues was lower in TCGA samples according to UALCAN database (**Figure 6E**).

In addition, previous result by Kaplan–Meier Plotter analysis has shown that higher mRNA expression of RPN1 indicated worse RFS, and the same tendency could be found in different subtypes of BC (**Figures 6F–I**).

RPN1 Knockdown Inhibits the Proliferation and Invasion of BLBC Cells

To explore the function of RPN1 in BLBC, we established shRNA-mediated RPN1 knockdown cell lines in SUM159 and SUM149, the two BLBC cell lines (**Figure 7A**). RPN1 knockdown induced significant proliferation inhibition (**Figures 7B, C**), which might be due to the cell cycle arrest, because of the remarkably increased percentage of cells in G2/M phase (**Figure 7D**). In addition, migration and invasion abilities of SUM159 cells were significantly inhibited by the knockdown of RPN1 (**Figures 7E, F**).

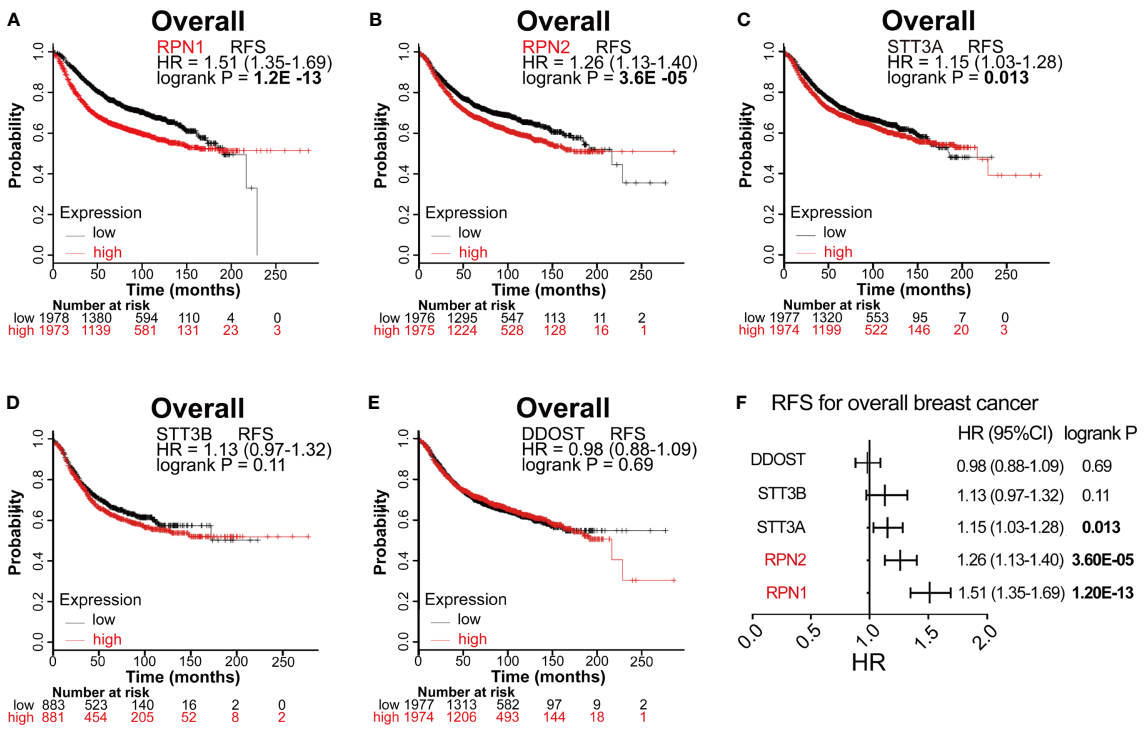


FIGURE 4 | Survival analyses of the five subunits in BC (RFS in Kaplan–Meier Plotter). **(A–E)** RFS for RPN1, RPN2, STT3A, STT3B, and DDOST in all BC. P<0.05 was considered statistically significant. **(F)** Prognostic HR of RFS for the five subunits. The data with statistical significance (P<0.05) were marked in bold text. RFS, relapse-free survival; HR, hazard ratio; CI, confidence interval.

RPN1 Knockdown Induces ERS-Dependent Cell apoptosis in BLBC

To explore the possible mechanisms of RPN1 in regulating the proliferation and invasion of BC cells, a total of 46 positively co-expressed genes of RPN1 with a Pearson correlation no less than 0.40 were obtained from the RNA-seq data in BLBC by bc-GenExMiner v4.5 database (**Supplementary Table S9**). And the GO enrichment analyses revealed that the biological process of "response to ERS", "endoplasmic reticulum unfolded protein response", "cellular response to unfolded protein", "cellular response to topologically incorrect protein", and "IRE1α-mediated unfolded protein response" were enriched for these genes (**Figure 8A**), indicating the possible important role of ERS in the knockdown of RPN1.

RPN1 plays a critical role in N-linked glycosylation, and previous studies have shown that the abnormality of the N-linked glycosylation may induce ERS in cells. Though the effect of ERS on tumor growth and metastasis was complex and dynamic, it has been proven that ERS could inhibit the growth and metastasis of tumors (35, 36). Inositol-requiring protein 1 α (IRE1α), protein kinase RNA -like endoplasmic reticulum kinase (PERK), and activating transcription factor 6 (ATF6) are endoplasmic reticulum transmembrane proteins, and each of them mediates an arm of the UPR. Normally, they are in a silent state combining with the endoplasmic reticulum chaperone immunoglobulin-binding protein (BiP). When under the ERS, they dissociate from BiP and activate their signaling functions respectively (14, 37). The results of both western blot and qRT-PCR showed the upregulation of PERK, IRE1α, ATF6, and BiP in RPN1-knockdown SUM159 cells (**Figures 8B, C**), suggesting that the ERS was induced after knockdown of RPN1. The ERS inhibitor 4-PBA could interact with unfolded or misfolded proteins to alleviate ERS (38). Treated with 4-PBA, the ERS could be significantly reduced in RPN1-knockdown SUM159 cells (**Figure 8C**). Several studies have demonstrated that the persistent ERS could play a pro-death role and trigger apoptosis (14, 35). Here, we found the knockdown of RPN1 decreased Bcl-2/Bax ratio at both protein and mRNA levels, which meant an increased apoptosis (**Figures 8D–F**), while treatment of 4-PBA increased the Bcl-2/Bax ratio (**Figure 8F**). We also found that the knockdown of RPN1 induced significant increase of early apoptosis in SUM159 cells, while treatment of 4-PBA rescued it (**Figures 8G, H**). These results demonstrated that inhibition of RPN1 could suppress BLBC cell proliferation and invasion *via* triggering the ERS.

DISCUSSION

N-glycosylation, one of important ways of post-translational modification, plays an important role in maintaining the stability of proteins. Most secreted proteins require

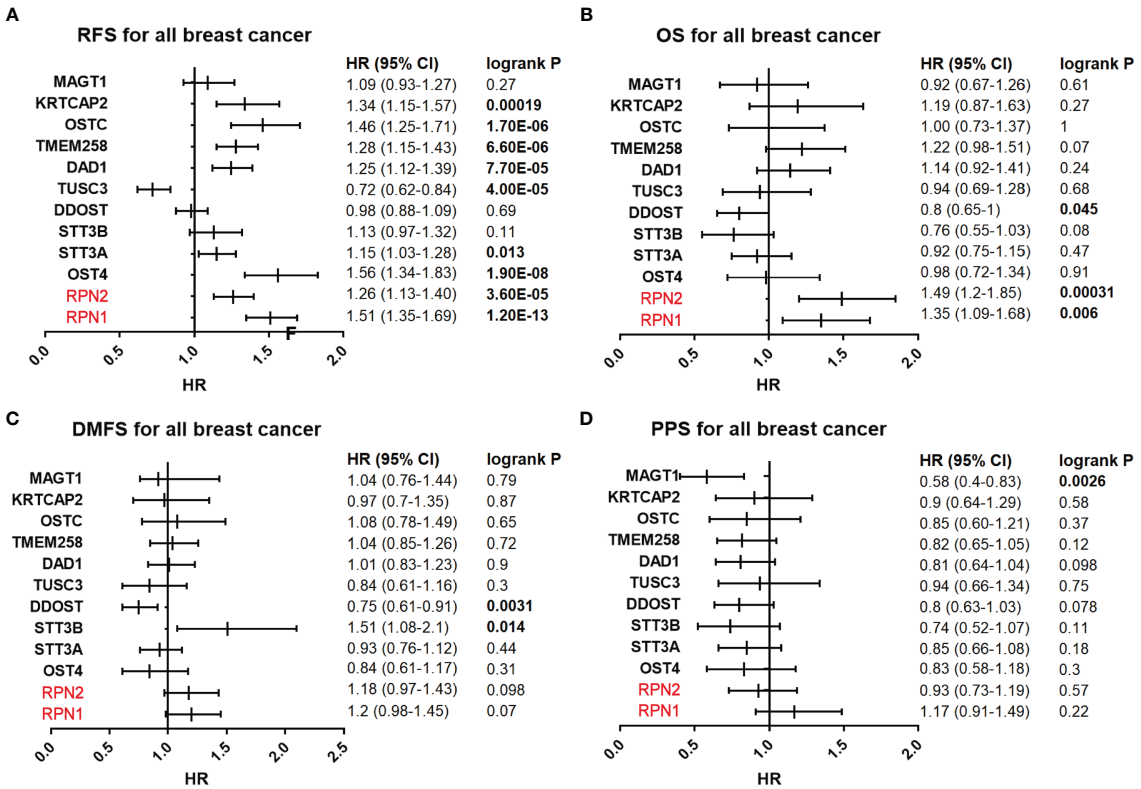


FIGURE 5 | Survival analyses of the OST subunits in breast cancer (RFS, PS, DMFS, PPS in Kaplan–Meier Plotter). **(A–D)** Prognostic HR of RFS, OS, DMFS, and PPS of individual OST subunits in all breast cancers. The data with statistical significance (P<0.05) were marked in bold text. RFS, relapse-free survival; OS, overall survival; DMFS, distant metastasis-free survival; PPS, post-progression survival; HR, hazard ratio; CI, confidence interval.

TABLE 2 | RFS of the RPN1, RPN2, STT3A, STT3B, and DDOST with different molecular subtypes in breast cancer.

BC subtypes	Gene	Affymetrix ID	Num of patients	HR (95%CI)	logrank P
Basal-like	RPN1	201011_at	618	1.31 (1.01–1.68)	**0.038**
	RPN2	213491_x_at	618	0.87 (0.67–1.12)	0.27
	STT3A	202223_at	618	1.42 (1.1–1.83)	**0.0063**
	STT3B	224700_at	360	1.24 (0.9–1.72)	0.19
	DDOST	208675_s_at	618	0.94 (0.73–1.21)	0.65
Luminal A	RPN1	201011_at	1,933	1.38 (1.16–1.64)	**0.00024**
	RPN2	213491_x_at	1,933	1.54 (1.29–1.83)	**9.1E-07**
	STT3A	202223_at	1,933	0.88 (0.74–1.04)	0.14
	STT3B	224700_at	831	1.12 (0.87–1.43)	0.38
	DDOST	208675_s_at	1,933	0.93 (0.79–1.11)	0.42
Luminal B	RPN1	201011_at	1,149	1.24 (1.03–1.51)	**0.025**
	RPN2	213491_x_at	1,149	1.12 (0.92–1.35)	0.26
	STT3A	202223_at	1,149	1.24 (1.02–1.5)	**0.028**
	STT3B	224700_at	407	1.23 (0.91–1.68)	0.18
	DDOST	208675_s_at	1,149	1 (0.82–1.21)	0.98
HER2-enriched	RPN1	201011_at	251	1.67 (1.13–2.47)	**0.0093**
	RPN2	213491_x_at	251	1.2 (0.82–1.77)	0.35
	STT3A	202223_at	251	1.21 (0.82–1.78)	0.33
	STT3B	224700_at	156	0.75 (0.48–1.18)	0.21
	DDOST	208675_s_at	251	0.68 (0.46–1)	**0.049**

The molecular subtypes were based on the 2013 St Gallen criteria. All of the data above were obtained from the Kaplan-Meier Plotter database.
The data with statistical significance (P<0.05) were marked in bold text.
HR, hazard ratio; CI, confidence interval.

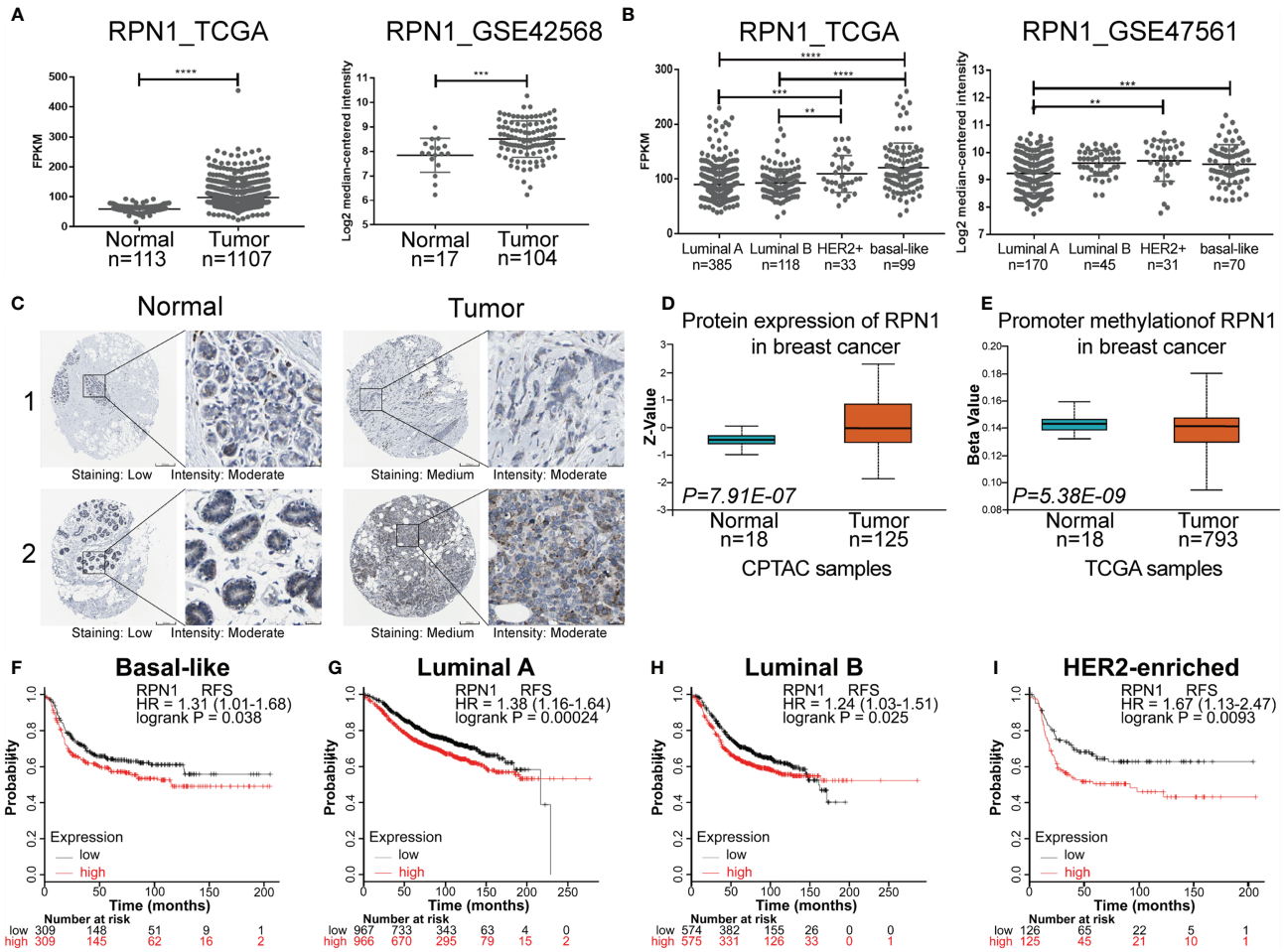


FIGURE 6 | The expression status and prognostic value of RPN1 in BC. **(A)** Gene expression analysis of RPN1 in BC tissues and normal breast tissues according to TCGA database and GEO database (GSE42568). **(B)** Gene expression analysis of RPN1 in different subtypes of BC according to TCGA database and GEO database (GSE47561). **(C)** Representative IHC staining of RPN1 expression in BC tissues and normal tissues according to the HPA database. **(D)** The protein expression of RPN1 in BC tissues and normal tissues in CPTAC samples according to the UALCAN database. Z-values represented standard deviations from the median across samples. **(E)** The promoter methylation level of RPN1 in BC tissues and normal tissues in TCGA samples according to the UALCAN database. The Beta value indicated the level of DNA methylation ranging from 0 (unmethylated) to 1 (fully methylated). **(F–I)** RFS for RPN1 in BLBC, Luminal A, Luminal B, and HER2-enriched BC. P<0.05 was considered statistically significant. **P < 0.01, ***P < 0.001 and ****P < 0.0001.

glycosylation to maintain stability and solubility, and N-glycosylation could assist proteins forming a proper folded structure by increasing the hydrophilicity of them or determining the chaperone bound to them (39). The OST complex is important for N-glycosylation, the abnormality of which is involved in tumors. Liu et al. found that the N-glycan profiles of membrane proteins in BC tissues significantly changed compared to the adjacent normal ones (40). Furthermore, previous reports have demonstrated the N-glycan alterations were essential for tumorigenesis, proliferation, and metastasis *via* modifying critical proteins or triggering mechanisms involved in the maintenance of cell homeostasis, such as ERS (41–45).

The 12 known subunits of OST complex play different roles in N-glycosylation. Some of the subunits have been reported to be associated with tumor. Takahashi et al. found that RPN2 could stabilize mutant p53 by inactivation of glycogen synthase kinase-3b, and the overexpression of RPN2 promoted the growth of BC (32). Burgermeister et al. revealed that the silence of TUSC3 by methylation was associated with the tumorigenesis of colorectal cancer, and epidermal growth factor receptor could be one of the target proteins (46).

In our study, we found that the mRNA expression levels of RPN1, RPN2, STT3A STT3B, and DDOST were significantly upregulated in BC tissues, and the expression levels of RPN1 and DDOST were significantly higher in the BLBC tissues compared to the non-BLBC. As for SBR and NPI, with the increasing of the grade of both SBR and NPI, the expression levels of RPN1, RPN2, and STT3B increased. As for the survival, the high expression of RPN1, RPN2, and STT3A were associated with

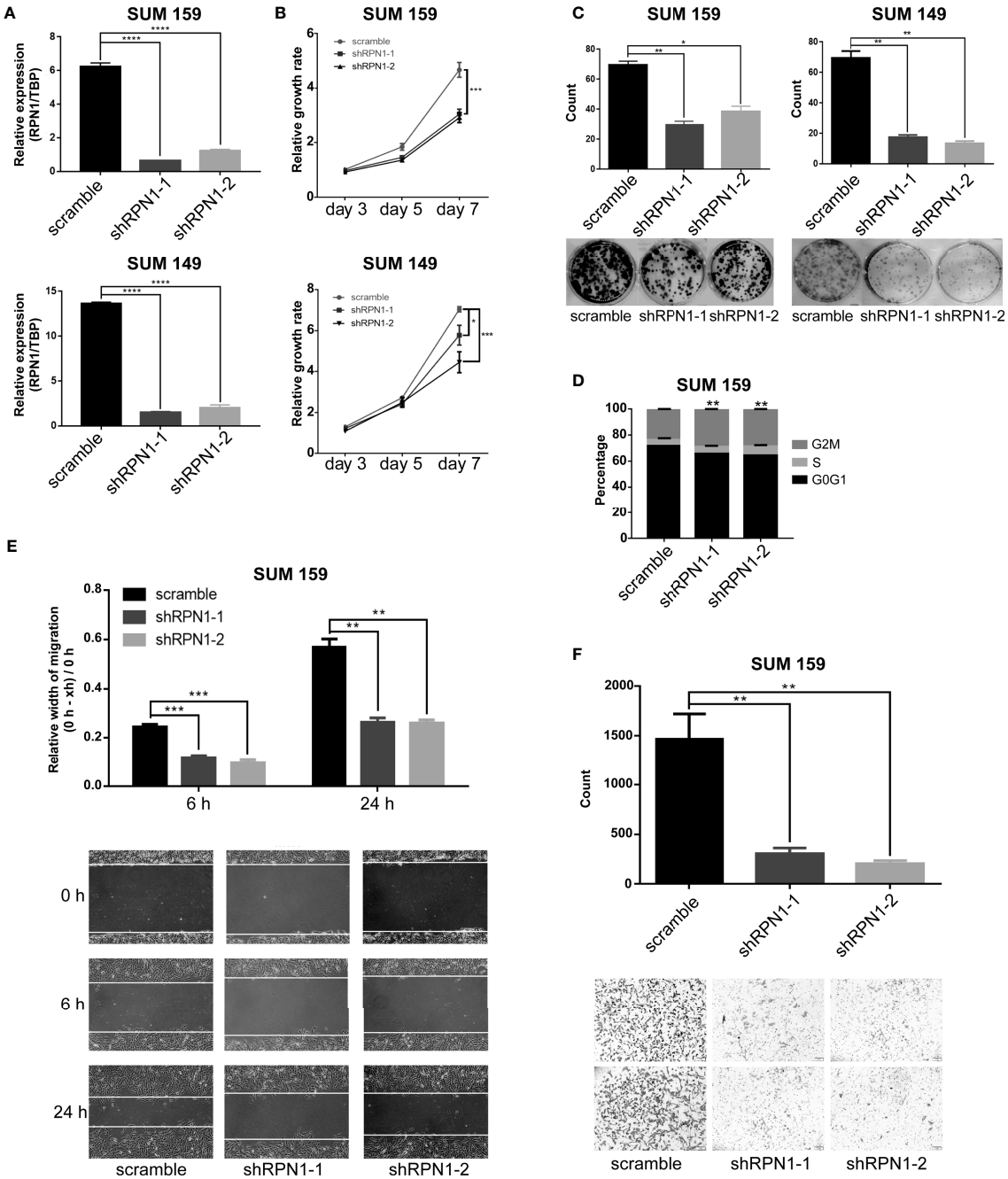


FIGURE 7 | RPN1 knockdown inhibited the growth and invasion of BLBC cells. **(A)** RPN1 was knocked down (scramble was the control). The expression of RPN1 was detected by qRT-PCR in SUM159 and SUM149. **(B, C)** MTT assay and Colony formation assay were used to measure the cell proliferation ability. **(D)** Cell cycle distribution analyzed by flow cytometry in SUM159 cells. **(E, F)** Wound healing assay and transwell assay were used to measure the cell migration and invasion ability. *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001.

worse RFS. Considering about both expression level and survival value, RPN1 and RPN2 could be the most effective biomarker and the most potential therapeutic target of OST subunits in BC. However, some studies have revealed that RPN2 plays a critical role in different cancers (32–34, 47, 48), while there was almost no study reporting the effect of RPN1.

RPN1 has been confirmed to be a type I transmembrane protein located on the endoplasmic reticulum, regulating N-glycosylation by interaction with the ribosomes and facilitating the specific precursors to the catalytic STT3A and STT3B subunits as a chaperone (15, 49). We conducted *in-vitro* experiments after knockdown of RPN1 in cells. And it turned

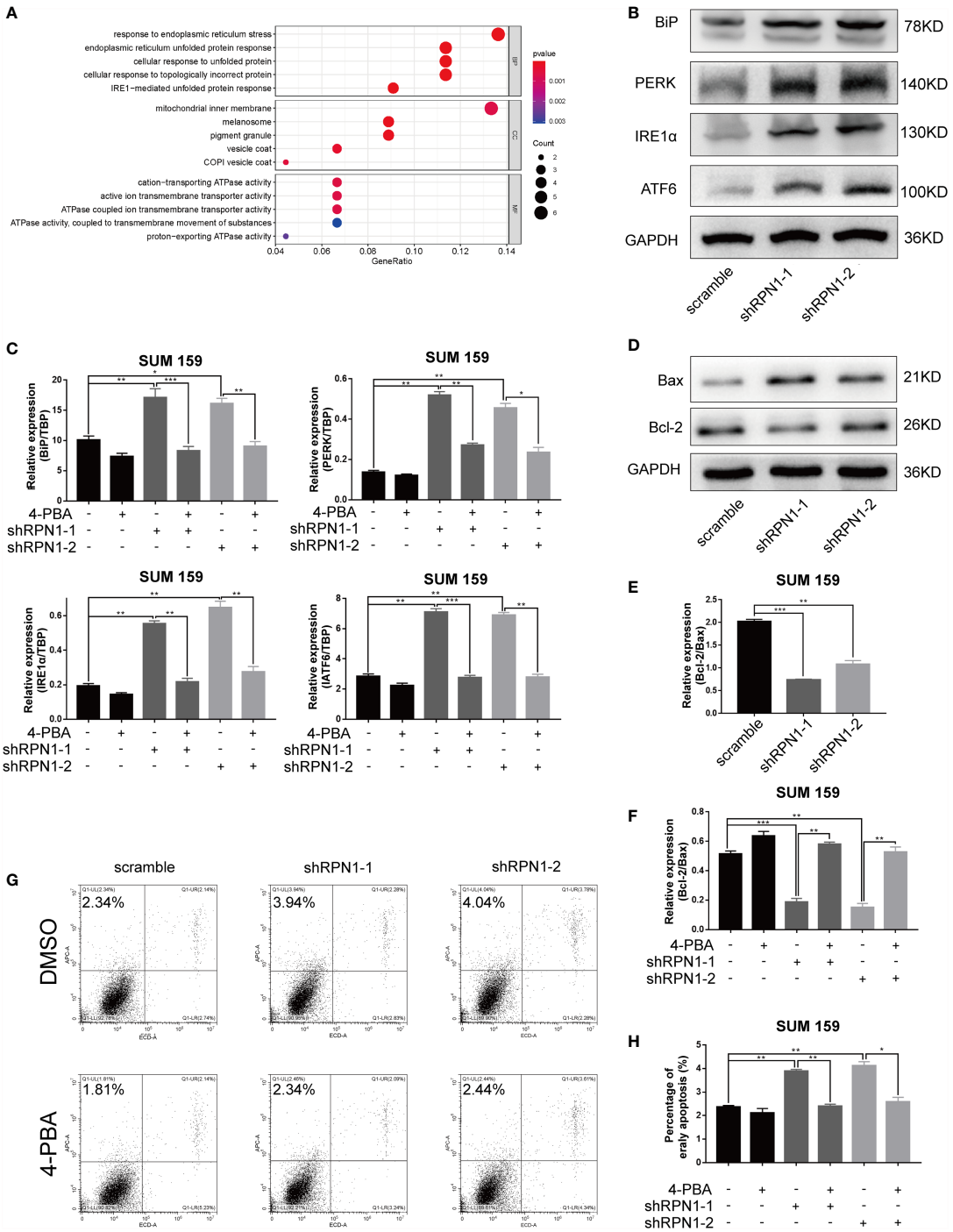


FIGURE 8 | ERS-dependent apoptosis was triggered by the knockdown of RPN1. **(A)** Bubble plot of the GO function enrichment analysis of the genes positively correlated with RPN1 RNA expression level in BLBC. Y-axis represents the name of the function, and X-axis represents the ratio of the number of the genes assigned to a term to the total number of the genes. The Bubble size represents the number of the genes annotated to the function. The color of the bubble represents the enriched P-value, while the red indicates a greater significance level. **(B)** The expression of ERS-related proteins was detected by western blot in SUM159 cells. **(C)** Scramble and shRPN1-infuected SUM159 cells were treated with 4-PBA (2 uM) or same volume of DMSO for 48 h, and the ERS markers were determined by qRT-PCR. **(D, E)** The protein expression of Bax and Bcl-2 in SUM159 detected by western blot and the ratio of the protein expression of Bcl-2 and Bax were also shown. **(F)** The ratio of the mRNA expression level of Bcl-2 and Bax was determined by qRT-PCR in 4-PBA- or DMSO-treated SUM159 cells. **(G, H)** Apoptosis analyzed by flow cytometry in 4-PBA- or DMSO-treated SUM159 cells. *P < 0.05, **P < 0.01 and ***P < 0.001. BP, Biological process; CC, cellular component; MF, molecular function; 4-PBA, Sodium phenylbutyrate.

out that the knockdown of RPN1 by shRNA led to poorer proliferation rate and less migration as well as invasion.

ERS is a mechanism to maintain the homeostasis of cell. And the aberrant glycosylation of proteins can lead to ERS and activate a set of signaling pathways (6). As mentioned previously, PERK, IRE1α, and ATF6 mediate three arms of UPR independently, and the signal pathways initiated by them could induce cell apoptosis. PERK, as a Ser/Thr kinase, mediates phosphorylation of eukaryotic initiation factor 2 (eIF2α) and then leads to the translation of transcription factor ATF4 (50). IRE1α can act not only as a protein kinase but also as an endoribonuclease. On the one hand, IRE1α can activate a pathway leading to c-Jun N-terminal kinase phosphorylation, which can promote apoptosis in several pathways (51). On the other hand, IRE1α is able to splice the mRNA of the transcription factor X-Box Binding Protein 1 (XBP1), producing XBP1s (52). ATF6 could not only cleave itself as a downstream signal molecule but also induce the modification of XBP1 (53). C/EBP homologous protein (CHOP), as an important pro-apoptotic transcription factor, can be the shared target of the three branches of UPR. It can be upregulated by the increased of ATF4, XBP1s, and cleaved ATF6. CHOP could induce the upregulation of various essential genes including Bcl-2 family members (54), thereby increase cell apoptosis directly.

The role ERS plays in tumorigenesis, proliferation, invasion, and apoptosis has been extensively reported (35, 55). In our study, we found that the knockdown of RPN1 inhibited the proliferation, migration, and invasion of BC. And the knockdown of RPN1 induced the upregulation of BiP, PERK, IRE1α, and ATF6 and the increase of cell apoptosis, while the treatment of ERS inhibitor could rescue them. These phenomena indicated that the RPN1 played a pro-tumor role by maintaining the endoplasmic reticulum homoeostasis in BLBC cells. However, the main target of RPN1 and the specific downstream pathway of ERS need further exploration.

In conclusion, clinically, the high expression level of RPN1 not only predicts a worse prognosis but is also related to a variety of recognized indicators of poor prognosis like negative ER status, negative PR status, BLBC subtype, higher SBR, and higher NPI. Biologically, our *in vitro* experiments clearly confirm that the proliferation, migration, and invasion of BC cells are significantly inhibited after interfering the expression of RPN1. Mechanismly, RPN1 inhibition leads to the activation of ERS and subsequent cell apoptosis. Although the detailed molecular mechanism is still not clear, it can be apparent that RPN1 plays an important part in BC and may be a novel biomarker as well as a potential therapeutic target.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

JD analyzed the databases and focused this study on RPN1 and was the major contributor in writing the manuscript. JX finished the most *in-vitro* experiments and contributed a lot in the manuscript. QD and XH contributed a lot in the experiments. RZ verified the accuracy of all the results and collated all the pictures. LX and SL were responsible for final approval of the version to be submitted and are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the manuscript are appropriately investigated and resolved. And all authors commented on previous versions of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The work was supported by funds from the National Key Research and Development Program of China (2018YFA0507501, 2020YFA0112300), NSFC Grant (81530075, 82073067, 81773155, 81772799), Program for Outstanding Medical Academic Leader 2019LJ04, Fudan University Research Foundation (IDH1340042), and Research Foundation of the Fudan University Shanghai Cancer Center (YJRC1603).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.722624/full#supplementary-material

Supplementary Figure S1 | Protein-protein interaction network of OST family members (data from GeneMANIA). Protein-protein interaction network among OST subunits. Each node indicates a gene, and the node size represents the strength of interactions. The internode connection lines represent the types of gene-gene interactions, and the line color represents the types of interactions, while the color of node represents the possible functions of these genes.

Supplementary Figure S2 | The relationships between the OST subunits and intrinsic molecular subtypes of breast cancer (data from bc-GenExMiner v4.7). **(A–L)** Box plots of individual OST subunit's expression according to the intrinsic molecular subtype of breast cancer (including basal-like, luminal A, luminal B, and HER2-enriched). Significant differences between groups were assessed by Welch's test, and Dunnett–Tukey–Kramer's test computed for each pairwise comparison. $P<0.05$ was considered statistically significant. In addition, the data in this figure was obtained from bc-GenExMiner v4.7 due to the update of the website. HER2-E, human epidermal growth factor receptor 2 enriched.

Supplementary Figure S3 | The relationships between the OST subunits and basal-like status of breast cancer (data from bc-GenExMiner v4.5). **(A–L)** Box plots of individual OST subunit's expression according to the basal-like status of breast cancer (basal-like or not). Significant differences between groups were assessed by Welch's test, and $P<0.05$ was considered statistically significant.

Supplementary Figure S4 | The relationships between the OST subunits and the SBR criteria (data from bc-GenExMiner v4.5). **(A–L)** Box plots of individual OST subunit's expression according to SBR. Global significant differences between groups were assessed by Welch's test, and Dunnett–Tukey–Kramer's test

computed for each pairwise comparison. P<0.05 was considered statistically significant. SBR, Scarff Bloom & Richardson grade.

Supplementary Figure S5 | The relationships between the OST subunits and the NPI criteria (data from bc-GenExMiner v4.5). **(A–L)** Box plots of individual OST subunit's expression according to NPI. Global significant differences between groups were assessed by Welch's test, and Dunnett–Tukey–Kramer's test computed for each pairwise comparison. P<0.05 was considered statistically significant. NPI, Nottingham Prognostic Index.

Supplementary Figure S6 | Survival analyses of the OST subunits in breast cancer (RFS in Kaplan–Meier Plotter). **(A–L)** RFS for individual OST subunits in all breast cancers. P<0.05 was considered statistically significant. RFS, relapse-free survival; HR, hazard ratio.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68 (6):394–424. doi: 10.3322/caac.21492
2. Fredslund SO, Gravholt CH, Laursen BE, Jensen AB. Key Metabolic Parameters Change Significantly in Early Breast Cancer Survivors: An Explorative PILOT Study. *J Transl Med* (2019) 17(1):105. doi: 10.1186/s12967-019-1850-2
3. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, et al. ONCOMINE: A Cancer Microarray Database and Integrated Data-Mining Platform1. *Neoplasia (New York NY)* (2004) 6(1):1–6. doi: 10.1016/S1476-5586(04)80047-2
4. Jezequel P, Campone M, Gouraud W, Guerin-Charbonnel C, Leux C, Ricolleau G, et al. Bc-GenExMiner: An Easy-to-Use Online Platform for Gene Prognostic Analyses in Breast Cancer. *Breast Cancer Res Treat* (2012) 131(3):765–75. doi: 10.1007/s10549-011-1457-7
5. Bai L, Wang T, Zhao G, Kovach A, Li H. The Atomic Structure of a Eukaryotic Oligosaccharyltransferase Complex. *Nature* (2018) 555 (7696):328–33. doi: 10.1038/nature25755
6. Cherepanova N, Shrimal S. Gilmore R. N-Linked Glycosylation and Homeostasis of the Endoplasmic Reticulum. *Curr Opin Cell Biol* (2016) 41:57–65. doi: 10.1016/j.ceb.2016.03.021
7. Pfeffer S, Dudek J, Gogala M, Schorr S, Linxweiler J, Lang S, et al. Structure of the Mammalian Oligosaccharyl-Transferase Complex in the Native ER Protein Translocon. *Nat Commun* (2014) 5(1):4072. doi: 10.1038/ncomms4072
8. Kelleher DJ, Gilmore R. An Evolving View of the Eukaryotic Oligosaccharyltransferase. *Glycobiology* (2006) 16(4):47R–62R. doi: 10.1093/glycob/cwj066
9. Oakes SA, Papa FR. The Role of Endoplasmic Reticulum Stress in Human Pathology. *Annu Rev Pathol* (2015) 10:173–94. doi: 10.1146/annurev-pathol-012513-104649
10. Gardner BM, Pincus D, Gotthardt K, Gallagher CM, Walter P. Endoplasmic Reticulum Stress Sensing in the Unfolded Protein Response. *Csh Perspect Biol* (2013) 5(3):a13169. doi: 10.1101/cshperspect.a013169
11. Clarke R, Cook KL, Hu R, Facey COB, Tavassoly I, Schwartz JL, et al. Endoplasmic Reticulum Stress, the Unfolded Protein Response, Autophagy, and the Integrated Regulation of Breast Cancer Cell Fate. *Cancer Res* (2012) 72 (6):1321–31. doi: 10.1158/0008-5472.CAN-11-3213
12. Shen K, Johnson DW, Vesey DA, McGuckin MA, Gobe GC. Role of the Unfolded Protein Response in Determining the Fate of Tumor Cells and the Promise of Multi-Targeted Therapies. *Cell Stress Chaperones* (2018) 23 (3):317–34. doi: 10.1007/s12192-017-0844-3
13. Li C, Fan Q, Quan H, Nie M, Luo Y, Wang L. The Three Branches of the Unfolded Protein Response Exhibit Differential Significance in Breast Cancer Growth and Stemness. *Exp Cell Res* (2018) 367(2):170–85. doi: 10.1016/j.yexcr.2018.03.033
14. Verfaillie T, Garg AD, Agostinis P. Targeting ER Stress Induced Apoptosis and Inflammation in Cancer. *Cancer Lett* (2013) 332(2SI):249–64. doi: 10.1016/j.canlet.2010.07.016
15. Wilson CM, High S. Ribophorin I Acts as a Substrate-Specific Facilitator of N-Glycosylation. *J Cell Sci* (2007) 120(4):648–57. doi: 10.1242/jcs.000729
16. Győrffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An Online Survival Analysis Tool to Rapidly Assess the Effect of 22,277 Genes on Breast Cancer Prognosis Using Microarray Data of 1,809 Patients. *Breast Cancer Res Tr* (2010) 123(3):725–31. doi: 10.1007/s10549-009-0674-9
17. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA Prediction Server: Biological Network Integration for Gene Prioritization and Predicting Gene Function. *Nucleic Acids Res* (2010) 38 (suppl_2):W214–20. doi: 10.1093/nar/gkq537
18. Thul PJ, Lindskog C. The Human Protein Atlas: A Spatial Map of the Human Proteome. *Protein Sci* (2018) 27(1):233–44. doi: 10.1002/pro.3307
19. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, et al. A Pathology Atlas of the Human Cancer Transcriptome. *Science* (2017) 357 (6352):n2507. doi: 10.1126/science.aan2507
20. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-Based Map of the Human Proteome. *Science* (2015) 347(6220):1260419. doi: 10.1126/science.1260419
21. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* (2017) 19 (8):649–58. doi: 10.1016/j.neo.2017.05.002
22. Curtis C, Shah SP, Chin S, Turashvili G, Rueda OM, Dunning MJ, et al. The Genomic and Transcriptomic Architecture of 2,000 Breast Tumours Reveals Novel Subgroups. *Nature* (2012) 486(7403):346–52. doi: 10.1038/nature10983
23. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated Observation of Breast Tumor Subtypes in Independent Gene Expression Data Sets. *Proc Natl Acad Sci USA* (2003) 100(14):8418–23. doi: 10.1073/pnas.0932692100
24. Zhao H, Langerød A, Ji Y. Different Gene Expression Patterns in Invasive Lobular and Ductal Carcinomas of the Breast. *Mol Biol Cell* (2004) 6 (15):2523–36. doi: 10.1091/mbc.E03
25. Ma X, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene Expression Profiling of the Tumor Microenvironment During Breast Cancer Progression. *Breast Cancer Res: BCR* (2009) 11(1):R7. doi: 10.1186/bcr2222
26. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal Gene Expression Predicts Clinical Outcome in Breast Cancer. *Nat Med* (2008) 14(5):518–27. doi: 10.1038/nm1764
27. Karnoub AE, Dash AB, Vo AP, Sullivan A, Brooks MW, Bell GW, et al. Mesenchymal Stem Cells Within Tumour Stroma Promote Breast Cancer Metastasis. *Nature* (2007) 449(7162):557–63. doi: 10.1038/nature06188
28. BLOOM H. Histologocal Grading and Prognosis in Breast Cancer; a Study of 1409 Cases of Which 359 Have Been Followed for 15 Years. *Br J Cancer* (1957) 3(11):359–77. doi: 10.1038/bjc.1957.43
29. Amat S, Penault-Llorca F, Cure H, Le Bouedëc G, Achard J, Van Praagh I, et al. Scarff-Bloom-Richardson (SBR) Grading: A Pleiotropic Marker of Chemosensitivity in Invasive Ductal Breast Carcinomas Treated by Neoadjuvant Chemotherapy. *Int J Oncol* (2002) 20(4):791. doi: 10.3892/ijo.20.4.791
30. Haybittle JL, Blamey RW, Elston CW, Johnson J, Doyle PJ, Campbell FC, et al. A Prognostic Index in Primary Breast Cancer. *Br J Cancer* (1982) 45(3):361–6. doi: 10.1038/bjc.1982.62
31. Winzer KJ, Buchholz A, Schumacher M, Sauerbrei W. Improving the Prognostic Ability Through Better Use of Standard Clinical Data - The Nottingham Prognostic Index as an Example. *PloS One* (2016) 11(3): e149977. doi: 10.1371/journal.pone.0149977
32. Takahashi R, Takeshita F, Honma K, Ono M, Kato K, Ochiya T. Ribophorin II Regulates Breast Tumor Initiation and Metastasis Through the Functional Suppression of GSK3β. *Sci Rep-Uk* (2013) 3(1):2474–86. doi: 10.1038/srep02474
33. Fujita Y, Yagishita S, Takeshita F, Yamamoto Y, Kuwano K, Ochiya T. Prognostic and Therapeutic Impact of RPN2-Mediated Tumor Malignancy in non-Small-Cell Lung Cancer. *Oncotarget* (2015) 6(5):3335–45. doi: 10.18632/oncotarget.2793
34. Ono M, Tsuda H, Kobayashi T, Takeshita F, Takahashi R, Tamura K, et al. The Expression and Clinical Significance of Ribophorin II (RPN2) in Human Breast Cancer. *Pathol Int* (2015) 65(6):301–8. doi: 10.1111/pin.12297

35. Luo B, Lee AS. The Critical Roles of Endoplasmic Reticulum Chaperones and Unfolded Protein Response in Tumorigenesis and Anticancer Therapies. *Oncogene* (2013) 32(7):805–18. doi: 10.1038/onc.2012.130
36. Clarke HJ, Chambers JE, Liniker E, Marciniak SJ. Endoplasmic Reticulum Stress in Malignancy. *Cancer Cell* (2014) 25(5):563–73. doi: 10.1016/j.ccr.2014.03.015
37. Ron D, Walter P. Signal Integration in the Endoplasmic Reticulum Unfolded Protein Response. *Nat Rev Mol Cell Bio* (2007) 8(7):519–29. doi: 10.1038/nrm2199
38. Deng L, Gao X, Liu B, He X, Xu J, Qiang J, et al. NMT1 Inhibition Modulates Breast Cancer Progression Through Stress-Triggered JNK Pathway. *Cell Death Dis* (2018) 9(12):1143–56. doi: 10.1038/s41419-018-1201-x
39. Shental-Bechor D, Levy Y. Folding of Glycoproteins: Toward Understanding the Biophysics of the Glycosylation Code. *Curr Opin Struct Biol* (2009) 19(5):524–33. doi: 10.1016/j.sbi.2009.07.002
40. Liu X, Nie H, Zhang Y, Yao Y, Maitikabili A, Qu Y, et al. Cell Surface-Specific N-Glycan Profiling in Breast Cancer. *PloS One* (2013) 8(8):e72704. doi: 10.1371/journal.pone.0072704
41. Liu C, Qiu H, Lin D. C-Jun-Dependent β 3gnt8 Promotes Tumorigenesis and Metastasis of Hepatocellular Carcinoma by Inducing CD147 Glycosylation and Altering N-Glycan Patterns. *Oncotarget* (2018) 26(9):18327–40. doi: 10.18632/oncotarget.24192
42. Boon L, Ugarte-Berzal E, Vandooren J, Opdenakker G. Glycosylation of Matrix Metalloproteases and Tissue Inhibitors: Present State, Challenges and Opportunities. *Biochem J* (2016) 473(11):1471–82. doi: 10.1042/BJ20151154
43. Guo H, Johnson H, Randolph M, Pierce M. Regulation of Homotypic Cell-Cell Adhesion by Branched N-Glycosylation of N-Cadherin Extracellular EC2 and EC3 Domains. *J Biol Chem* (2009) 284(50):34986–97. doi: 10.1074/jbc.M109.060806
44. Zhou F, Cui C, Ge Y, Chen H, Li Q, Yang Z, et al. 2,3-Sialylation Regulates the Stability of Stem Cell Marker CD133. *J Biochem* (2010) 148(3):273–80. doi: 10.1093/jb/mvq062
45. Wu J, Chen S, Liu H, Zhang Z, Ni Z, Chen J, et al. Tunicamycin Specifically Aggravates ER Stress and Overcomes Chemoresistance in Multidrug-Resistant Gastric Cancer Cells by Inhibiting N-Glycosylation. *J Exp Clin Canc Res* (2018) 37(1):272–83. doi: 10.1186/s13046-018-0935-8
46. Burgermeister E, Höde P, Betge J, Gutting T, Merkel A, Wu W, et al. Epigenetic Silencing of Tumor Suppressor Candidate 3 Confers Adverse Prognosis in Early Colorectal Cancer. *Oncotarget* (2017) 8(49):84714. doi: 10.18632/oncotarget.20950
47. Tominaga N, Hagiwara K, Kosaka N, Honma K, Nakagama H, Ochiya T. RPN2-Mediated Glycosylation of Tetraspanin CD63 Regulates Breast Cancer Cell Malignancy. *Mol Cancer* (2014) 13(1):134. doi: 10.1186/1476-4598-13-134
48. Zhang H, Jiang H, Zhang H, Liu J, Hu X, Chen L. Ribophorin II Potentiates P-Glycoprotein- and ABCG2-Mediated Multidrug Resistance *via* Activating ERK Pathway in Gastric Cancer. *Int J Biol Macromol* (2019) 128:574–82. doi: 10.1016/j.ijbiomac.2019.01.195
49. Mohorko E, Glockshuber R, Aebi M. Oligosaccharyltransferase: The Central Enzyme of N-Linked Protein Glycosylation. *J Inherit Metab Dis* (2011) 34(4):869–78. doi: 10.1007/s10545-011-9337-1
50. Cullinan SB, Diehl JA. Coordination of ER and Oxidative Stress Signaling: The PERK/Nrf2 Signaling Pathway. *Int J Biochem Cell Biol* (2006) 38(3):317–32. doi: 10.1016/j.biocel.2005.09.018
51. Dhanasekaran DN, Reddy EP. JNK Signaling in Apoptosis. *Oncogene* (2008) 27(48):6245–51. doi: 10.1038/onc.2008.301
52. Lee AH, Iwakoshi NN, Glimcher L. XBP-1 Regulates a Subset of Endoplasmic Reticulum Resident Chaperone Genes in the Unfolded Protein Response. *Mol Cell Biol* (2003) 21(23):7448–59. doi: 10.1128/MCB.23.21.7448
53. Bommiasamy H, Back SH, Fagone P, Lee K, Meshinchi S, Vink E, et al. Atf6α Induces XBP1-Independent Expansion of the Endoplasmic Reticulum. *J Cell Sci* (2009) 122(10):1626–36. doi: 10.1242/jcs.045625
54. Oyadomari S, Mori M. Roles of CHOP GADD153 in Endoplasmic Reticulum Stress. *Cell Death Differ* (2004) 11(4):381–9. doi: 10.1038/sj.cdd.4401373
55. Ghosh C, Nandi A, Basu S. Supramolecular Self-Assembly of Triazine-Based Small Molecules: Targeting the Endoplasmic Reticulum in Cancer Cells. *Nanoscale* (2019) 11(7):3326–35. doi: 10.1039/c8nr08682f

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

REVIEW
published: 28 October 2021
doi: 10.3389/fonc.2021.750315

KDM4 Involvement in Breast Cancer and Possible Therapeutic Approaches

Benluvankar Varghese[1†], *Nunzio Del Gaudio*[1†], *Gilda Cobellis*[1†], *Lucia Altucci*[1,2] *and Angela Nebbioso*[1,3*]

[1] Department of Precision Medicine, University of Campania Luigi Vanvitelli, Napoli, Italy, [2] Biogem Institute of Molecular Biology and Genetics, Ariano Irpino, Italy, [3] Saint Camillus International University of Health and Medical Sciences, Rome, Italy

Breast cancer (BC) is the second leading cause of cancer death in women, although recent scientific and technological achievements have led to significant improvements in progression-free disease and overall survival of patients. Genetic mutations and epigenetic modifications play a critical role in deregulating gene expression, leading to uncontrolled cell proliferation and cancer progression. Aberrant histone modifications are one of the most frequent epigenetic mechanisms occurring in cancer. In particular, methylation and demethylation of specific lysine residues alter gene accessibility *via* histone lysine methyltransferases (KMTs) and histone lysine demethylases (KDMs). The KDM family includes more than 30 members, grouped into six subfamilies and two classes based on their sequency homology and catalytic mechanisms, respectively. Specifically, the *KDM4* gene family comprises six members, *KDM4A-F*, which are associated with oncogene activation, tumor suppressor silencing, alteration of hormone receptor downstream signaling, and chromosomal instability. Blocking the activity of KDM4 enzymes renders them "druggable" targets with therapeutic effects. Several KDM4 inhibitors have already been identified as anticancer drugs *in vitro* in BC cells. However, no KDM4 inhibitors have as yet entered clinical trials due to a number of issues, including structural similarities between KDM4 members and conservation of the active domain, which makes the discovery of selective inhibitors challenging. Here, we summarize our current knowledge of the molecular functions of KDM4 members in BC, describe currently available KDM4 inhibitors, and discuss their potential use in BC therapy.

Keywords: epigenetics, histone demethylation, KDM4 inhibitors, JMJD2, KDM4

OPEN ACCESS

Edited by:
Bruno M. Simões, The University of Manchester, United Kingdom

Reviewed by:
Rakesh Kumar, Rajiv Gandhi Centre for Biotechnology, India
Cyril Ribeyre, Centre National de la Recherche Scientifique (CNRS), France

***Correspondence:**
Angela Nebbioso
angela.nebbioso@unicampania.it

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: *30 July 2021*
Accepted: *13 October 2021*
Published: *28 October 2021*

Citation:
Varghese B, Del Gaudio N, Cobellis G, Altucci L and Nebbioso A (2021) KDM4 Involvement in Breast Cancer and Possible Therapeutic Approaches. Front. Oncol. 11:750315. doi: 10.3389/fonc.2021.750315

INTRODUCTION

Breast cancer (BC) is the second leading cause of cancer death in women worldwide with a 0.5% increase in incidence rate per year. Advances in diagnosis and treatment in 64% of BC cases at earlier stages has increased 5-year survival to 99% (National Breast Cancer Foundation).

Much is known about oncogenes, tumor suppressors, and DNA repair genes, which play a role in breast tumorigenesis, promoting aberrant cell growth and/or mismatch error repair (1, 2). Research on molecular hallmarks of BC has identified several diagnostic markers including:

i) immunohistochemical markers, such as estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2); ii) genetic markers, such as *BRCA1*, *BRCA2*, and *PIK3CA* mutations; iii) immunomarkers, such as programmed death-ligand 1 (PD-L1) and tumor infiltrating lymphocytes; iv) proliferation markers, such as Ki-67. All of these have significantly changed the prediction of prognosis and therapy decisions (3).

The Cancer Genome Atlas classifies BC into five different subtypes: normal-like, luminal A, luminal B, HER2-positive ($HER2^+$), and basal-like. Luminal A and B tumors are ER- and PR-positive (ER^+PR^+), while the $HER2^+$ and basal-like subtypes are hormone-independent (ER^-PR^-) and positive for high levels of Ki-67, showing the worst prognosis (4).

Current therapeutic strategies for BC are based on tumor heterogeneity associated with different histotypes and specific molecular profiles: ER^+ and PR^+ patients are treated with hormonal therapy, $HER2^+$ patients with anti-HER2 therapy, and *BRCA* mutation carriers with poly (ADP-ribose) polymerase (PARP) inhibitors plus adjuvant therapies (chemotherapy, immunotherapy, and radiation therapy) (5).

Despite advancements in our knowledge of BC biology as well as intense disease prevention programs and therapies able to block tumor progression, the incidence of BC continues to rise. High-throughput analysis reveals a massive transcriptional deregulation in BC, for which a tight interplay between genetic and epigenetic factors has been hypothesized. Progressive dedifferentiation of cell identity to a progenitor-like state due to increased cell plasticity is observed in the early phase of cancer formation, whereas epigenetic modifications support oncogenic progression (6).

Epigenetic alterations such as DNA methylation and reversible histone modifications (methylation, acetylation, ubiquitination) alter gene accessibility, resulting in aberrant gene expression.

A promising opportunity to rewind cell fate comes from epigenetic-based therapies, which make use of small molecule drugs (epidrugs) able to interfere with the activity of epigenetic regulators and thus correct cancer-associated chromatin states (7). Following confirmation of the efficacy of epidrug-based therapies in oncology by several *in vitro* studies, many epidrugs have moved to clinical trials for different cancer types (8), and some have been clinically approved by the US Food and Drug Administration (9).

Eukaryotic chromatin is organized in active euchromatin and inactive heterochromatin and the histone methylations define these two interchangeable functional states. Histone lysine methylation is regulated by methyltransferases (KMTs) and demethylases (KDMs) (10). KDMs are classified into two groups: i) the KDM1 or LSD1 family, dependent on flavin adenine dinucleotide (FAD) and ii) the JmjC family, dependent on 2-oxoglutarate (2-OG) for their demethylase activity. JmjC domain-containing KDMs form the larger KDM class with 20 members grouped into five subfamilies (KDM2/7, KDM3, KDM4, KDM5, and KDM6), and their deregulation is associated with cancer, including BC (11).

Based on their catalytic activity, KDM4 subfamily members catalyze N-methyl-lysine demethylation by removing mono-, di-, and trimethyl marks *via* an oxidative mechanism. KDM4 uses 2-OG and O_2 as cosubstrates, Fe(II) as a cofactor for the enzymatic oxygenase reaction (**Figure 1A**). This activity contributes to the control of gene expression in a context-dependent manner, either by influencing the compaction of chromatin or through regulation of signaling pathways and recruitment of other protein complexes. The most frequent modifications occur on H3K4, H3K36 and H3K79 associated with gene activation, whereas H3K9, H3K27, H4K20 and H3K56 associated with gene silencing (11–13).

The KDM4 (JHDM3/JMJD2) subfamily is highly conserved (14, 15). In humans, this subfamily comprises *KDM4A*, *KDM4B*, *KDM4C*, and *KDM4D* genes, with *KDM4E* and *KDM4F* considered as pseudogenes, although a partial catalytic activity is reported (16) (**Figure 1B**).

KDM4A-C enzymes have five different domains: JmjN, JmjC, tandem PHD, Tudor, and F-box, whereas KDM4D-F lack PHD and Tudor domains (**Figure 1B**) (17). The stability and catalytic activity of KDM4s depends on the interaction between JmjN and JmjC domains, and their structural integrity maintains overall protein stability (18). The crystal structure of the KDM4A Tudor domain revealed it as histone reader, identifying methylated lysine residues at histone H3 and H4 tails; the function of the PHD domain is still unclear, although in other PHD-containing proteins this domain is able to bind modified and unmodified histone residues (19).

Concerning KDM4 mRNA levels, they are tightly regulated to guarantee proper biological processes (20). Next-generation sequencing in normal tissues revealed that KDM4A/B/C are broadly expressed in most tissues, although at different levels. They share more than 50% protein sequence identity, however the variations in expression levels suggest that these proteins have not-overlapping functions, as evidenced also by single/double knockout mouse models, that were viable and showed no evident abnormalities. Cell-specificity is thus guaranteed by specific interaction with regulatory factors. For instance, the control of KDM4A expression rely on ubiquitin-proteasome pathway through FBX022, a key regulator of histone methylation. This evidence suggests that posttranslational modifications of KDM4A regulates its abundance, conferring it the cell/tissue-specificity (21). Other studies suggested that KDM4s have peculiar cell-type functions. Heart-specific KDM4A conditional knockout showed cardiac hypertrophy and no compensatory effect has been observed (22).

KDM4 subfamily members control different biological functions, to ensure proliferation, differentiation, migration and adhesion (23), as well as regulation of transcription (24) and genome stability (25) (**Figure 2**). In embryonic stem cells (ESCs), KDM4 proteins (4B and 4C) control stem cell identity by interacting with the pluripotency factors such as Sox2, Oct4, c-Myc, and Klf4, but also by modulating, alone or in combination, gene expression during the differentiation program (26). In addition, depletion of KDM4C in ESCs causes downregulation

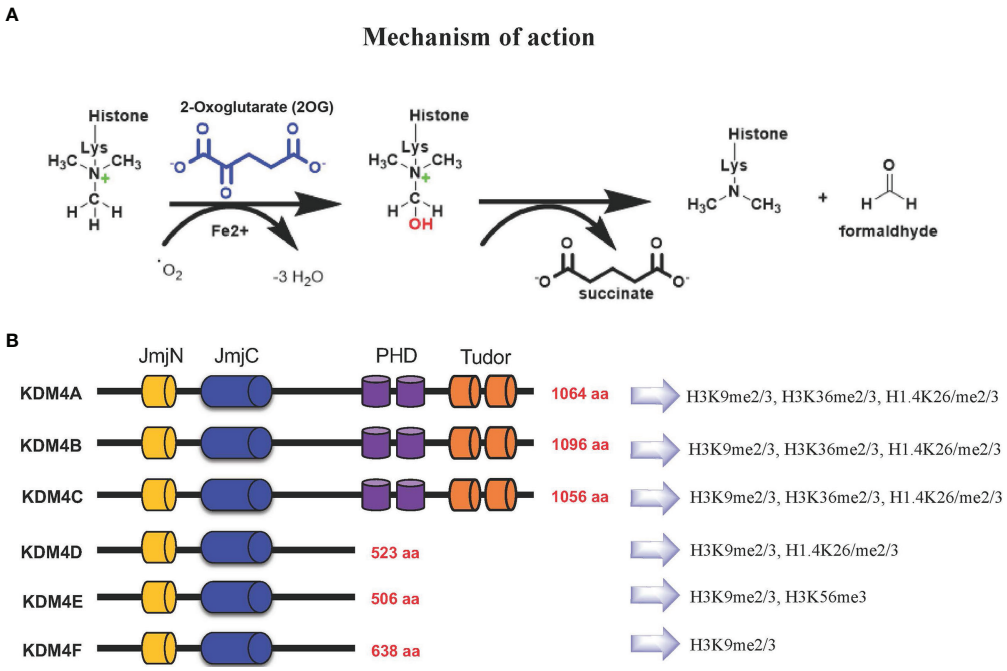


FIGURE 1 | **(A)** KDM4s: mechanism of demethylation; **(B)** KDM4s have conserved JmjN and JmjC domains, while substrate recognition domains such as PHD and Tudor are present only in KDM4A-C. Histone targets of KDM4 family members are shown on the right.

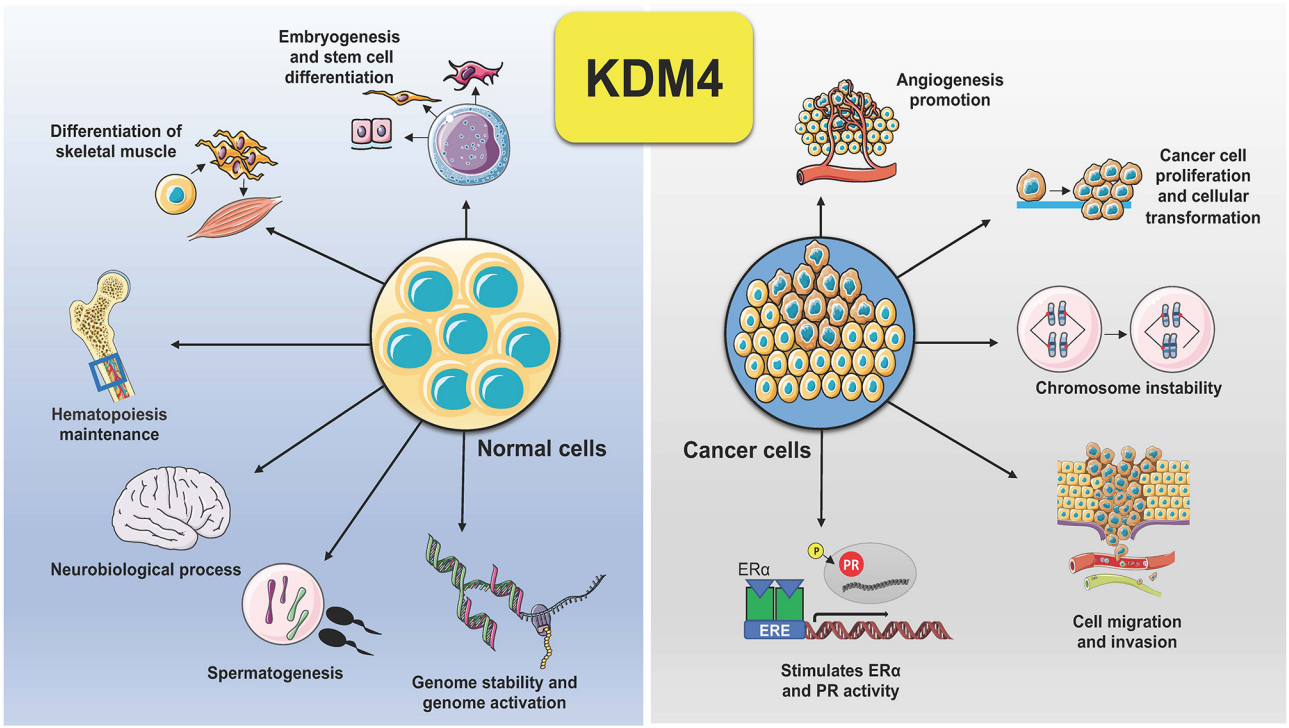


FIGURE 2 | Schematic representation of KDM4 functions in normal and cancer cells. Deregulation of KDM4 promotes cancer cell proliferation, migration, and invasion, angiogenesis, chromosome instability, and stimulation of ER and PR activity.

of *Myc* and *Klf4* genes associated with cell proliferation during early embryo stage, leading to developmental defects (27).

KDM4s are also involved in cell differentiation: KDM4B promotes osteogenic differentiation of human mesenchymal stem cells, activating expression of *DLX* genes by removing trimethyl groups from H3K9me3 marks (28, 29), whereas depletion of KDM4B reduces osteogenic differentiation *via DLX* gene suppression (28).

Downregulation of *Taf1b* and *Nom1* in hematopoietic stem cells was observed due to accumulation of H3K9me3 on their transcription start sites (30), following knockout of KDM4A/B/C *in vivo* resulting in aberrant differentiation.

KDM4A was found to play an important role in skeletal muscle differentiation (31), while KDM4B deletion in mice leads to neurodevelopmental disorders and defects in spinal maturation (32). KDM4A knockdown or inhibition decreases leukocyte adhesion and transmigration in inflammatory response, by modulating expression of vascular adhesion proteins (ICAM1 and VCAM1) in cerebral microvessels (33). KDM4A/D play a role in female fertility (34) and in spermatogenesis, respectively (35). KDM4D was recently reported to maintain genome stability by facilitating double-strand DNA damage repair mechanisms in a PARP1-dependent manner; specifically, the interaction between KDM4D and RNA seems to be essential for chromatin localization and efficient demethylation of trimethyl H3K9 (36).

Concerning breast tissue, KDM4B is important for transcriptional regulation and development of mammary gland. Deletion of KDM4B in mammary epithelium produces immature mammary gland development in female mice (37). KDM4B is also involved in ER signaling cascade and is required for ER-mediated gene transcription, essential for normal development of ovarian follicles, luteal function, and ovulation (38–40). In summary, these findings revealed that KDM4B plays a critical role in regulation of transcriptional program in the mammary gland.

Dysregulation of KDM4s is behind several hallmarks of cancer (**Figure 2**). Tumorigenesis is a complex adaptive process that involves alterations in different cellular functions, as proliferation, differentiation, adaptation to altered microenvironment, many of them controlled by KDM4s, found overexpressed in various human cancers, sustaining tumor progression and acting as oncoproteins (11, 41).

Thus, KDM4s have emerged as a druggable targets in cancer to restore cell homeostasis by erasing inappropriate histone modifications distributed across the genome that are responsible of cell transformation. Although the drug discovery rationale is straightforward, the efficacy of KDM4 inhibitors identified to date is limited, mainly due to their lack of selectivity and/or specificity to the different KDM4 isoforms (42).

High expression levels of KDM4A were observed in squamous cell carcinoma as well as in ovarian and prostate cancer, where it is highly associated with chromosomal instability (43). KDM4A/C/D bind androgen receptor (AR) *in vitro* and *in vivo*, resulting in tumor cell proliferation through demethylation of H3K9me3 in AR target genes, stimulating AR-dependent transcription in combination with KDM1A (44, 45).

KDM4A is also directly involved in upregulation of the lung cancer-associated genes *CXCL5*, *ADAM12*, and *JAG1*, involved in angiogenesis promotion, tumor cell growth, and cell proliferation (46–50). Overexpression of KDM4C was found in non-small cell lung carcinoma (51) and osteosarcoma (52), where upregulation of fibroblast growth factor 2, promoted by KDM4B/C modulates cell migration, invasion, and proliferation in osteosarcoma metastasis (52).

Demethylation of H3K9 marks by KDM4D is involved in tumor necrosis factor α activation, associated with tumorigenesis and inflammatory response (53). KDM4D stimulates p53-dependent gene expression and acts as a pro-oncogenic factor, specifically on AR target genes in prostate and colon cancer cell growth (54). Further, KDM4A reduces activity of p53 pathway through inhibition of Ras-mediated chromodomain-helicase-DNA-binding protein 5 (CHD5) induction, blocking senescence and thereby promoting cell transformation (55).

In ovarian cancer, reduced levels of KDM4B led to an increase in H3K9me3 in the promoter regions of genes such as *PDGFB*, *LCN2*, and *LOXL2*, suppressing cell migration, invasion, and formation of spheroids *in vitro* (56). In gastrointestinal tumors, KDM4D promotes cancer progression by directly interacting with hypoxia-inducible factor (HIF) 1β gene and activating its expression *via* H3K9me3 demethylation of the vascular endothelial growth factor A promoter region (57). KDM4B expression was found to be activated by *HIF* genes, promoting cancer cell survival in a hypoxic setting (58–60).

In conclusion, KDM4s exert their effect mainly by altering the chromatin state and therefore the expression of genes involved in physiological functions that, when disrupted, cancer occurs.

FUNCTIONAL ROLE OF KDM4s IN BREAST CANCER

KDM4s are responsible in controlling development and proliferation of mammary gland (61), and their altered expression (mainly gene amplification) can promote cell transformation, migration, and invasion, all hallmarks of tumorigenesis in BC (47) (**Figure 3**). A recent meta-analysis of KDM4 gene expression in BC subtypes identified overexpression of KDM4A/D in basal-like BC, whereas KDM4B was predominantly expressed in ER^+ luminal-type BC (61).

The development of potential KDM4 inhibitors with high selectivity in different BC subtypes therefore remains a major challenge. To address this issue, a better understanding of the molecular mechanisms of KDM4s as well as their specific target sites is urgently required to develop new treatments targeting molecular pathways crucial for BC progression.

KDM4A in BC

KDM4A mainly demethylates H3K9me2/me3 and, at a lower rate, H3K36me2/me3 *in vivo* and H1.4K26me3 *in vitro* (17),

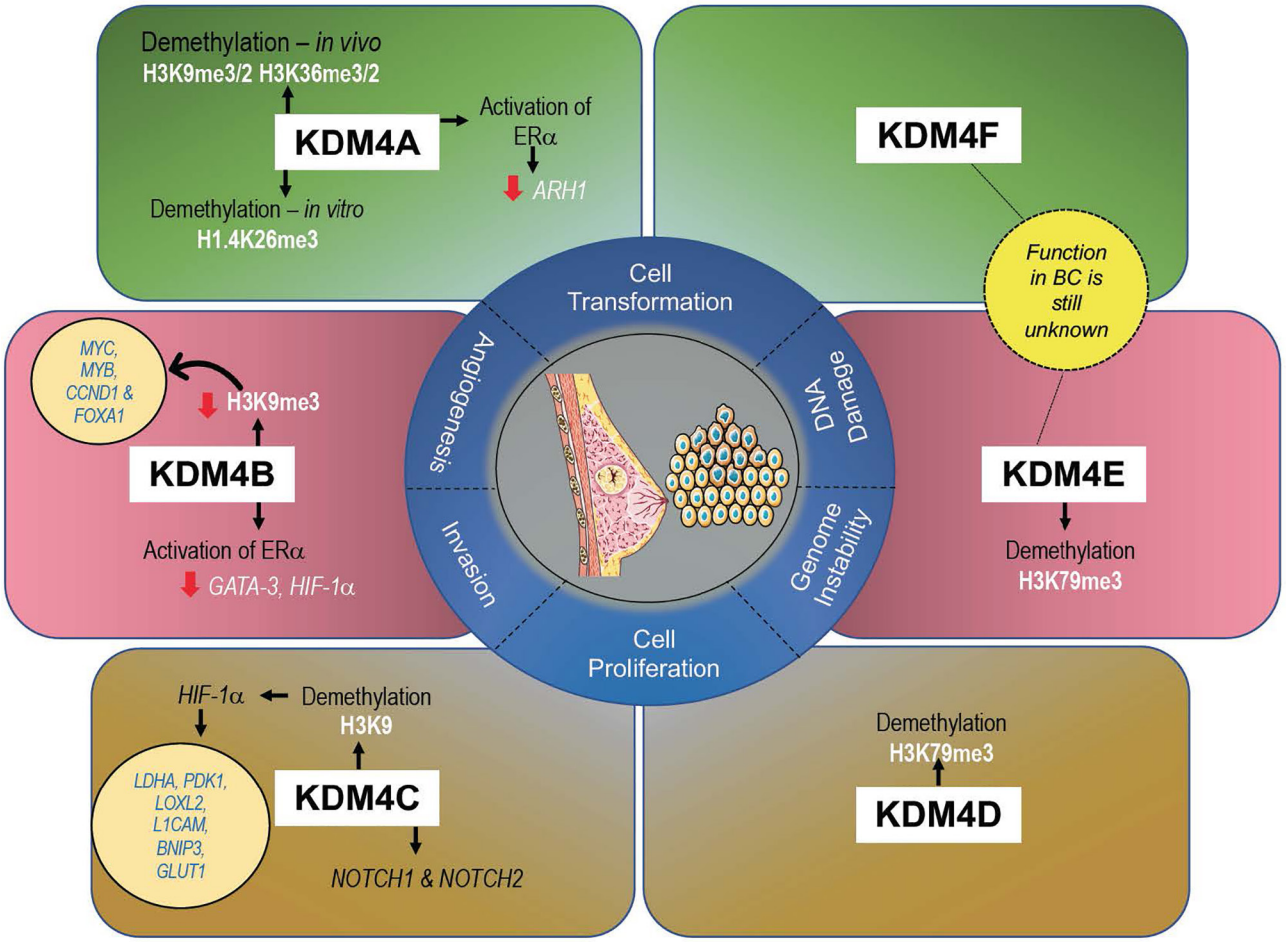


FIGURE 3 | Expression levels of each KDM4 and its histone marks associated with BC progression.

promoting chromatin decompaction. *Via* histone deacetylase (HDAC) and p53 association, KDM4A may repress gene expression (62, 63).

Overexpression of KDM4A was observed in 60% of BC tissue at both mRNA and protein level (64). KDM4A overexpression leads to upregulation of estrogen-dependent genes, whereas depletion of KDM4A decreases transcription of ERα target genes, such as *JUN* and *CCND1*, promoting cell growth arrest. Taken together, these interconnections suggest that KDM4A promotes BC growth *via* hormone-dependent and -independent mechanisms (65). KDM4A overexpression was also found to contribute to BC growth through downregulation of the tumor suppressor gene ADP-ribosylarginine hydrolase 1 (ARH1), highly expressed in normal breast tissue (66). Furthermore, downregulation of the ubiquitous transcription factor Sp1 was reported in highly invasive and in advanced stages of BC, showing a clear correlation with the TNM staging system, confirmed by KDM4A overexpression (67). *In vitro*, knockdown of KDM4A in MCF-7 cells blocks JUN expression, inhibiting invasion, migration, and tumor formation (68–70). In these cells, expression levels of KDM4A were also found to be modulated by hsa-mir-23a-3p, hsa-mir-23b-3p, and hsa-mir-137. Inhibition of these microRNAs enhances KDM4A levels, with a consequent increase in some drug-resistant genes such as CDC28 protein kinase regulatory subunit 1B (*CKS1B*) (71), contributing to the outgrowth of chemoresistant cells (68).

CHD5, a tumor suppressor gene, is under the control of KDM4A, whose silencing restores *CHD5* expression by decreasing H3K36me2/me3 histone marks in its locus (72).

KDM4B in BC

KDM4B is similar to KDM4A in structure and enzymatic activity, demethylating both H3K9 and H3K36. Unlike KDM4A/C, KDM4B acts as a monomer and not as a homodimer or heterodimer (73).

KDM4B is a key regulator of estrogen signaling cascade, and its depletion attenuates BC growth both *in vitro* and *in vivo* (37, 40). Noteworthy, *KDM4B* is itself an ER-responsive gene (58). Taken together, these findings suggest a positive feedback mechanism between KDM4 and ER whereby estrogen-induced KDM4 expression in turn coregulates and, unexpectedly, upregulates ER-target genes, sustaining BC growth. KDM4B is

required in ER-mediated gene transcription essential not only in mammary gland, but also in ovarian follicles, suggesting a possible correlation of KDM4B between these gynecological cancers.

H3K4 methylation and H3K9 demethylation are coordinated by binding of KDM4B/mixed-lineage leukemia 2 (MLL2 or KMT2D) complex, with ERα driving ERα-dependent transcription (74). Some studies report the interaction of KDM4B/ERα with SWI/SNF-B chromatin complex, regulating numerous genes involved in resistance and invasiveness of BC (37). Decreased levels of H3K9me3, corresponding to overexpression of KDM4B, facilitate transcription of ER-responsive genes such as *MYC*, *MYB*, *CCND1* (37), and *FOXA1* (40). GATA-3 is a transcription factor highly expressed in luminal A-type BC and is associated with ER expression. The demethylation process mediated by KDM4B is fundamental for activation of ER by GATA-3, whereas downregulation of KDM4B levels induces H3K9 methylation and a reduction in GATA-3 binding on ER promoter, suppressing ER targets (40). Moreover, ERα regulates expression of KDM4B through HIF-1α, promoting its expression in a feed-forward regulatory circuit (58).

Several ER coregulated genes are primed to activate gene expression upon histone modifications induced by KDM4 proteins. One example is the KDM3A/KDM4B/FOXA1 complex, which leads to an increase in pro-proliferative and ERα-dependent gene expression and dual knockdown of KDM4A and KDM4B, strongly inhibiting ERα activity and blocking cell proliferation (75).

Additionally, high KDM4B-mediated demethylation levels of H3K9 were found on the promoter of long interspersed nuclear element-1, increasing its expression and improving the effectiveness of retrotransposition (76). A direct correlation was found between KDM4B expression and the absence of H3K9me3 in pericentromeric regions, suggesting the involvement of this enzyme in chromosomal instability and aneuploidy cell formation (77).

Interestingly, KDM4B also plays a role at cytoplasmic level, where it regulates the unfolded protein response (UPR) pathway through direct interaction with eukaryotic initiation factor 2α (eIF2). UPR is commonly hyperactivated as result of severe and prolonged cellular stress, triggering cell death. Inhibiting the association between KDM4B and eIF2 also allows activation of UPR cell death pathway in triple-negative breast cancer (TNBC), deficient in PTEN, and therefore increases responsiveness to therapy with PI3K-AKT inhibitors (78).

Selective estrogen receptor modulators (SERMs) are beneficial in treating premenopausal ER-positive BC resistant to tamoxifen. However, no effect was obtained in tumors where *Fbxo22* gene is low expressed as Fbxo22 ubiquitinates tamoxifen-bound KDM4B (79), resulting in KDM4B overexpression and poor prognosis.

KDM4C in BC

KDM4C (also known as GASC1) is amplified in many cancers including BC, mainly in basal-like and in ER^- and PR^- subtypes (80), making this enzyme a negative prognostic marker (81, 82). KDM4C overexpression is mediated by gene amplification of 9p24 chromosomal region, which contains several candidate tumor genes, including *KDM4C*.

KDM4C regulates expression of genes involved in stem cell self-renewal and induces phenotypic changes in cancer cells. However, despite its involvement in tumor development, proliferation, and aggression, very little is known about this enzyme compared to KDM4A/B. In MCF-10A cells, the expression of KDM4C induces a transformed phenotype (80). KDM4C upregulates many genes responsible for cell growth, migration, and metastasis and interacts with HIF-1α, mediating KDM4C recruitment on hypoxia-inducible genes and demethylation of H3K9 on metabolic genes, such as *LDHA*, *PDK1*, *LOXL2*, *L1CAM*, *BNIP3*, and *GLUT1*. The physical interaction of these two proteins is a critical epigenetic mechanism, given that HIF-1α involvement in BC is responsible for an aggressive phenotype, characterized by metastasis progression and resistance to drug therapy (83).

A D396N polymorphism found in the caspase-3 cleavage site of *KDM4C* in BC cells and contributes to drug resistance, indicating the involvement of KDM4C in BC-resistant progression (81).

Unlike KDM4A, KDM4C is recruited to mitotic chromosomes, modulating correct chromosomal stability and gene expression. This suggests that total inhibition of the enzyme in TNBC should induce a reduction in cell multiplication (84) and an increase in γ-H2AX, a marker of DNA damage (81). Through modulation of steroid receptor co-activator 1 (SRC-1), KDM4C also regulates CD24 and the apoptotic protein PAWR. In endocrine-resistant BC cell lines, SRC-1/KDM4C complex together with JUN mediates transcriptional repression of these two oncogenic proteins (85).

Despite evidence that KDM4C silencing or inhibition may represent an effective epigenetic therapy in BC treatment, a study conducted on 355 patients with invasive BC found that KDM4C was negatively associated with the development of a more aggressive BC histopathological type (grade II/III, ductal-type, PR^-, and ER^-). Women with KDM4C-positive tumors responded better to radiation therapy and hormone treatment (82).

KDM4D/E/F in BC

Unlike other subfamily members, KDM4D has JmjN and JmjC domains encoding only a small peptide protein. KDM4D potentially regulates H3K79me3, suggesting its involvement in DNA repair, telomeric silencing regulation, cellular development, transcriptional regulation, and cell cycle checkpoints (86). The role of KDM4D in cancer is relatively less studied than that of other KDM4s. A recent study reported that KDM4D was significantly overexpressed in basal-like BC, with an amplification frequency of 3.6%, and was found ubiquitously expressed in ER^+, MCF-10A, and basal-like cell lines (61).

The catalytic domain of KDM4E was found to demethylate H3K9me3/me2 regulated by the availability of O_2 in an *in vitro* assay (87).

The expression of *KDM4E/F* in BC is still unknown. Further studies on these genes may unveil their potential role in BC and in other cancers.

KDM4 Inhibitors

Depending on their mechanism of action, KDM4 inhibitors are divided into different classes: 2-OG cofactor mimics, metal cofactor disruptors, histone substrate-competitive inhibitors, and natural and peptide inhibitors (**Figure 4**) (88).

Cofactor mimics are metal-chelated inhibitors that competitively bind Fe(II) molecules of the catalytic site of KDM4 members, blocking their enzymatic activity. Cancer cells are able to reprogram their metabolism to support the increased energy demand required for cell survival and rapid proliferation. Metabolic disruption can alter KDM4 activity by modifying the availability of the required cofactor, 2-OG. Therefore, the intermediates of the tricarboxylic acid (TCA) cycle can inhibit KDM activity. The first identified KDM4 inhibitors were the natural molecules fumarate and succinate, which act as competitive antagonists for 2-OG (89).

Among 2-OG analogs, the oxalic acid-derivative N-oxalylglycine (NOG: IC_{50} = 78 μM), pyridine dicarboxylic acid (PCA: IC_{50} = 1.4 μM), and 8-hydroxyquinoline (8-HQ: IC_{50} = 0.6 μM) showed antiproliferative activity (42).

The hydroxamate-based 2-OG analog NCDM-32B was identified as a good inhibitor of KDM4 subfamily members, and its therapeutic potential was investigated in basal-like BC (61). In enzymatic assays, NCDM-32B displayed IC_{50} values of 3.0 μM for KDM4A and 1.0 μM for KDM4C. Treatment with NCDM-32B in BC cell lines induced a global increase in H3K9me3/me2 marks, and microarray Gene Ontology analysis of differentially expressed genes revealed pathways that control cell proliferation, growth, DNA replication, and DNA repair. Of note, the compound suppressed the expression of oncogenes, such as the *MET* proto-oncogene, as well as genes involved in cell cycle regulation including *CDC26* and *CDK6*. These data suggest that NCDM-32B may be a regulator of different cell growth and transformation pathways activated in BC (61).

The orally available KDM4 inhibitor QC6352 has an IC_{50} value of 0.104 μM for KDM4A, 0.056 μM for KDM4B, and 0.035 μM for KDM4C (90). This molecule showed a strong capability to inhibit proliferation, sphere formation, and xenograft tumor formation of BC stem-like cells derived from tissue of TNBC patients after neoadjuvant chemotherapy. *Via* H3K9me3 induction, QC6352 inhibited expression of epidermal growth factor receptor, a pivotal gene in therapy resistance mechanisms in TNBC (91).

A very recent study characterized TACH101 as a first-in-class pan inhibitor of KDM4s, with promising pharmacological applicability. Surprisingly, the compound displayed potent inhibitory activity on four KDM4 isoforms (A-D) with IC_{50} values below 0.100 μM. Furthermore, it increased H3K36me3 levels and induced apoptosis in human esophageal cancer, TNBC, and colorectal cancer cell lines. *In vivo*, TACH101 showed 100% tumor growth inhibition in BC xenograft models, reducing tumor-initiating cell frequencies by 4.4-fold, and exhibited good oral availability. However, further preclinical studies are required to drive progression of the compound to clinical trials (92).

Another interesting study reported that KDM4 inhibitors such as A1 (CGC00247751), B3 (NCGC00244536), and I9 (NCGC00247743) repress the transcriptional activity of AR

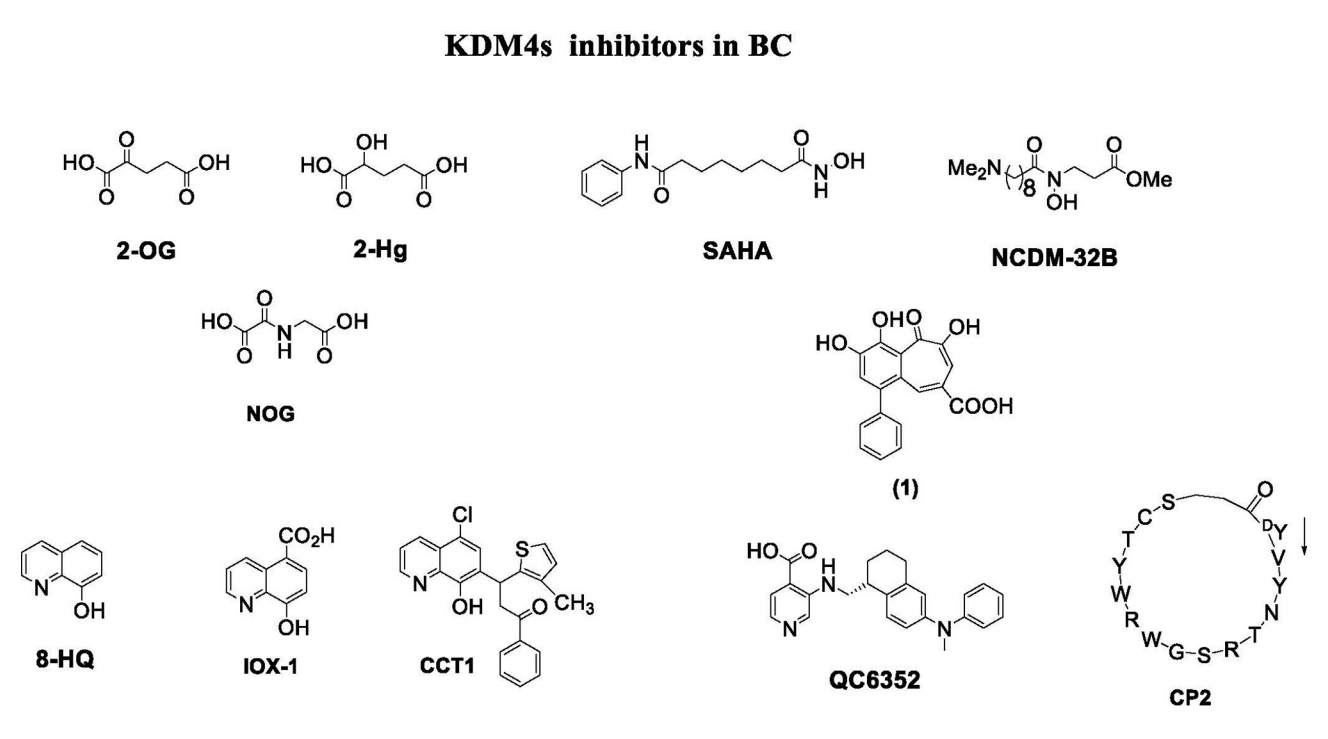


FIGURE 4 | KDM4 inhibitors in BC.

and B-MYB, regulating genes such as *PLK1*, involved in cell cycle progression. Interestingly, the compound B3 showed antiproliferative effects in BC cell lines. Findings from this study suggest that the inhibitor specifically targets KDM4B in late S-phase due to activation of *PLK1* transcription *via* B-MYB, justifying the development of this KDM4B inhibitor for AR^+ prostate cancer and opening up the possibility for new treatments in the AR^+ subgroup of BC (93).

Other selective KDM family inhibitors are also described as anticancer agents in BC. Many reports indicate that KDM5 maintains tumor-initiating cells and promotes the development of drug tolerance (94). A selective inhibitor of KDM5B, KDOAM-21, significantly increased global levels of H3K4me3 in MCF-7 and TNBC cells. The compound also inhibited the growth of MCF-7 cells at 5 μM in colony-formation experiments (95). In another study, YUKA1, a small molecule inhibitor of KDM5A, displayed the ability to prevent drug tolerance in $HER2^+$ BC cells treated with trastuzumab (96).

Natural Inhibitors

Quercetin (WO2007104314) is a natural flavonoid that was found to inhibit KDM4C in demethylation assays and to modify H3K9me3 demethylation status in esophageal carcinoma and bone osteosarcoma cells. A hydroxamate analog (JP2011168581) showed selective inhibition of KDM4A/C (88). Methylstat (US20130137720) is a methyl ester analog inhibiting KDM4C that increases hypermethylation levels of H3K9me3 and H3K36me3 in a concentration-dependent manner, blocking the growth of MCF-7 cells (97).

Curcumin derivatives show good inhibition of KMD4s at cellular level. For example, efficient histone demethylation was observed by FLLL compounds. Notably, FLLL-8 and FLLL-24 displayed inhibitory activity against KDM4C, while FLLL-60 showed inhibition of KMD4A/D (98). Recently, a new compound synthesized from the natural product purpurogallin was reported to be a KDM4 inhibitor. This compound, called 9bf, exhibited a potent inhibitory activity on KDM4A and antiproliferative activity in many solid cancer cells (99).

Peptide Inhibitors

In 2014, the first peptide-based KDM4 inhibitors displaying major selectivity and minor off-target effects were described (100). Two cyclic peptides were identified and both were active against KDM4C. Interestingly, this study proposed a novel approach to developing selective KDM4 inhibitors, regardless of the substrate and cofactor used (100).

An *in vitro* screening of a cyclic peptide library identified selective substrate-competitive inhibitors of KDM4s, showing alteration of H3K9me3 levels and inhibition of cell proliferation. The cyclic peptide CP2 showed potent IC_{50} values (IC_{50} = 0.42/0.33/0.39 μM against KDM4A/B/C, respectively) and exceptional intra-subfamily selectivity. The compound displayed high potency against KDM4A/B/C but was much less active against KDM4D (IC_{50} = 6.2 μM) and KDM4E (IC_{50} = 9.2 μM) (101). Although further studies are needed to evaluate stability, cell permeability, and subcellular localization, this approach may lead to the discovery and characterization of potent peptide inhibitors of KDM4 for the treatment of BC and other cancers. Because the functions of non-catalytic domains of KDM4 subfamily members such as PHD and Tudor domains are still unknown, the development of KDM4 inhibitors against non-catalytic domains remains challenging.

Dual and Other Inhibitors

Since epigenetic machinery such as DNA methylation and histone modifications often work in parallel, the use of single agents in combination has recently drastically increased as this approach enhances their efficacy (102). Such a drug combination approach has also been exploited toward non-epigenetic targets (103). For example, combinations of HDAC-HSP90 inhibitors (104), HDAC-DNMT inhibitors (105), HDAC-KDM1 inhibitors (106), HDAC-BET protein inhibitors (107, 108), HDAC-EZH2 inhibitors (109), and HDAC-PI3K inhibitors (110) showed good efficacy in different cancer cells.

By way of an example, the dual KDM inhibitor MC3324 showed inhibition of KDM1 and KDM6A with a consequent increase in H3K4me2 and H3K24me3 levels and induction of apoptosis in hormone-responsive MCF-7 cells. Downregulation of ERα was observed at both transcriptional and translational level, indicating that the compound affects the transcription of genes regulating cell proliferation, hormonal response, and apoptosis. Interestingly, MC3324 reduced cell proliferation in *ex vivo* BC models and showed absence of toxicity and good oral efficacy in chicken embryo and mouse xenograft models. Thus, the simultaneous inhibition of multiple targets could be beneficial in BC (111).

Combining different drugs could be a feasible strategy to target multiple oncogenic pathways (112–114). Currently, many two-in-one drug approaches are being investigated in clinical trials for various cancers. The well-known HDAC inhibitor vorinostat (SAHA) in combination with tamoxifen (NCT00365599), and carboplatin and nab-paclitaxel (NCT00616967) is at different stages of clinical trials. In another trial, entinostat (MS-275) in combination with immunotherapy and monoclonal antibodies (nivolumab, ipilimumab) is under evaluation in patients with metastatic BC and $HER2^-$ BC (NCT02453620). The synergy between HDAC inhibitors and anti-HER2 therapy with trastuzumab showed promising results (NCT00258349), but the adverse effects of trastuzumab resistance need to be further evaluated. In sum, in order to develop and optimize the effective use of epidrugs alone or in combination, there is an urgent need to identify new epigenetic targets that will pave the way for new cancer treatments.

In recent years, epigenetic studies combined with advanced computational methods have brought substantial advancements in drug discovery. A recent cutting-edge technology known as "epi-informatics" has been exploited to create a plethora of targeted compounds that may eventually lead to the discovery of new drugs. Computer-aided drug design could be used to explore and identify much needed selective KDM4 inhibitors for BC (115).

DISCUSSION

The role of KDM4s in cancer has been extensively studied, and promising targets for BC therapy have been proposed. Through demethylation of H3K9 and H3K36, KDM4s regulate chromatin structure and gene expression in numerous cancer types. Notably, overexpression of KDM4 subfamily members promotes cancer cell proliferation, invasion, and migration, DNA damage, tumor angiogenesis, and metastasis. Although some epigenetic mechanisms and functions of KDM4 proteins associated with carcinogenesis remain unclear, a growing body of evidence indicates that KDM4 inhibitors are good candidates as anticancer drugs for various malignancies, including BC. To date, however, reported inhibitors do not have a sufficient level of enzyme specificity and are not commercially available for the treatment of any cancer types. The specific role of KDM4s and their mechanism of action in BC is less well known. Another challenging task is to explore new compounds against KDM4 activity through computational screening, which may identify more specific KDM4 inhibitors. Further studies could drive the future development of potent and selective targets for specific KDM4s in BC.

To improve overall healthcare outcomes in BC, a substantial endeavor aimed at reducing mortality and increasing survival in patients is needed. Several studies have investigated the crucial role of KDM4 subfamily members in different cancers, and KDM4 targeting has been revealed as a promising strategy to inhibit BC development. However, no KDM4 modulators have as yet been approved for clinical use. Targeting these molecules has thus been attracting considerable interest among the scientific community (62, 88). The development of KDM4 inhibitors is still in its premature stage, with a limited number of scientific publications and patents. Although the development of potent and selective KDM4 inhibitors for BC is a complicated process, efforts in a number of different directions might be of help: i) The functions of KDM4E/F are still unclear, and more extensive investigations into these two enzymes may open up new avenues in cancer research. In addition, KDM4A/B/C share the same substrates, further complicating the development of selective inhibitors for KDM4 subtypes. Structural studies could help better define the catalytic pockets of these enzymes for more precise targeting (116, 117); ii) Specific gene expression patterns/programs controlling KDM4 activity are poorly studied and need to be further explored; iii) Findings related to KDM4 inhibitors usually derive from *in vitro* or cell-based assays, with a lower amount of *in vivo* data being reported. Characterizing their *in vivo* activity might provide greater insights useful for more potent drug development; iv) The activity of KDM4s in regulating DNA damage, non-histone proteins, and other posttranslational modifications is still unclear. New research directed at understanding these mechanisms may lead to the identification of novel molecules with higher selectivity: v) Due to their structural similarities and the presence of a JmjC domain in all isoforms of KDM4, engineering KDM4 inhibitors with isoform specificity is challenging. However, elucidating the distinct physiological function of each KDM4 enzymes in cancer is necessary.

Moving from a single- to a multi-KDM4s target therapeutic approach may be a useful strategy to improve BC treatment. Specifically, hybrid scaffolds coupling two individually well-known KDM4i compounds in a single unit (dual compound) could be a valid option to simultaneously target different KDM4 isoforms. As reported, the molecule MC3324 was more effective in blocking cell proliferation, targeting ER, and inducing cell death of BC cells, compared to its constituent moieties and other known inhibitors used alone or in combination. Alternatively, dual compounds could also be used to directly target KDM4 isoforms and their co-regulators in a highly specific manner. Hybrid scaffolds bridging binders of KDM4 isoform domains could be used to target a dual compound to KDM4 isoforms, thus overcoming the lack of specificity towards isoforms. However, the complexity and vulnerability of epigenetic regulation limits the use of epigenetic molecules to specific treatment contexts, which may contribute to poor therapeutic outcome. Studies into combinatorial epigenetic therapy have recently paved the way toward exploring new effective therapeutic strategy in cancer. For instance, polyclonal tumors are characterized by the presence of multiple coactive deregulated pathways, and in these tumors epigenetic alterations are favoring, permissive, or secondary events. In this scenario, testing novel targeted treatments in a single-agent approach may thus be problematic and may underestimate their effectiveness. Combining epigenetic drugs with conventional protocols, both targeted and immune therapies, may therefore represent a successful anticancer approach.

Exploiting single-cell omics approaches could capture cancer cell heterogeneity and provide a better understanding of the involvement of different KDM4s in the sequential stages of breast transformation at both bulk and single-cell level. This approach may ensure a more accurate patient stratification and unravel the role of each KDM4 in BC transformation, allowing evaluation of the efficacy of targeted selective modulators and opening the way toward personalized medicine in BC driven by specific KDM4 aberrations.

Unlike genetic events, epigenetic changes are reversible and because of this inherent plasticity, epigenome-targeted therapy has emerged as a potential strategy for the treatment of cancer. The results of investigational and approved epigenetic therapies in other clinical contexts have proven that this approach can be effective. KDM4s have been found to control many aspects of BC, including cancer initiation and progression. Additionally, traditional BC treatments fail in targeting therapy-resistant cancer stem cells strongly characterized by alteration of epi-regulators. Thus, considering that KDM4s are epigenetic regulators with overlapping functions in controlling gene expression of crucial signaling pathways, KDM4s inhibition reflects their target potential for BC therapy. Targeting these histone demethylases will pave the way toward improving the treatment of BC patients.

AUTHOR CONTRIBUTIONS

Conceptualization and writing BV, NDG, and GC. Supervision AN and LA. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Italian Association for Cancer Research (AIRC-17217), the Campania Regional Government "Lotta alle Patologie Oncologiche" iCURE (CUP B21C17000030007), MIUR Proof Of Concept (POC01_00043), VALERE: Vanvitelli per la Ricerca Program: AdipCare (ID263) and EPInhibitDRUGre (CUP B66J20000680005).

ACKNOWLEDGMENTS

We thank C. Fisher for English language editing.

REFERENCES

1. Lee EY, Muller WJ. Oncogenes and Tumor Suppressor Genes. *Cold Spring Harb Perspect Biol* (2010) 2(10):a003236. doi: 10.1101/cshperspect.a003236
2. Torgovnick A, Schumacher B. DNA Repair Mechanisms in Cancer Development and Therapy. *Front Genet* (2015) 6:157. doi: 10.3389/fgene.2015.00157
3. Loibl S, Poortmans P, Morrow M, Denkert C, Curigliano G. Breast Cancer. *Lancet* (2021) 397(10286):1750–69. doi: 10.1016/s0140-6736(20)32381-3
4. Tomczak K, Czerwińska P, Wiznerowicz M. Review The Cancer Genome Atlas (TCGA): An Immeasurable Source of Knowledge. *Współczesna Onkol* (2015) 1A:68–77. doi: 10.5114/wo.2014.47136
5. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast Cancer. *Nat Rev Dis Primers* (2019) 5(1):66. doi: 10.1038/s41572-019-0111-2
6. Smigiel JM, Taylor SE, Bryson BL, Tamagno I, Polak K, Jackson MW. Cellular Plasticity and Metastasis in Breast Cancer: A Pre- and Post-Malignant Problem. *J Cancer Metastasis Treat* (2019) 5:47. doi: 10.20517/2394-4722.2019.26
7. Garcia-Martinez L, Zhang Y, Nakata Y, Chan HL, Morey L. Epigenetic Mechanisms in Breast Cancer Therapy and Resistance. *Nat Commun* (2021) 12(1):1786. doi: 10.1038/s41467-021-22024-3
8. Diesch J, Zwick A, Garz AK, Palau A, Buschbeck M, Gotze KS. A Clinical-Molecular Update on Azanucleoside-Based Therapy for the Treatment of Hematologic Cancers. *Clin Epigenet* (2016) 8:71. doi: 10.1186/s13148-016-0237-y
9. Ganesan A, Arimondo PB, Rots MG, Jeronimo C, Berdasco M. The Timeline of Epigenetic Drug Discovery: From Reality to Dreams. *Clin Epigenet* (2019) 11(1):174. doi: 10.1186/s13148-019-0776-0
10. Mosammaparast N, Shi Y. Reversal of Histone Methylation: Biochemical and Molecular Mechanisms of Histone Demethylases. *Annu Rev Biochem* (2010) 79:155–79. doi: 10.1146/annurev.biochem.78.070907.103946
11. Labbe RM, Holowatyj A, Yang ZQ. Histone Lysine Demethylase (KDM) Subfamily 4: Structures, Functions and Therapeutic Potential. *Am J Trans Res* (2013) 6(1):1–15.
12. Klose RJ, Kallin EM, Zhang Y. JmjC-Domain-Containing Proteins and Histone Demethylation. *Nat Rev Genet* (2006) 7(9):715–27. doi: 10.1038/nrg1945
13. Slama P. Identification of Family Determining Residues in Jumonji-C Lysine Demethylases: A Sequence-Based, Family Wide Classification. *Proteins* (2016) 84(3):397–407. doi: 10.1002/prot.24986
14. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, et al. Endogenous Human microRNAs That Suppress Breast Cancer Metastasis. *Nature* (2008) 451(7175):147–52. doi: 10.1038/nature06487
15. Paolicchi E, Crea F, Farrar WL, Green JE, Danesi R. Histone Lysine Demethylases in Breast Cancer. *Crit Rev Oncol Hematol* (2013) 86(2):97–103. doi: 10.1016/j.critrevonc.2012.11.008
16. Liu X, Wang Y, Gao Y, Su J, Zhang J, Xing X, et al. H3K9 Demethylase KDM4E Is an Epigenetic Regulator for Bovine Embryonic Development and a Defective Factor for Nuclear Reprogramming. *Development* (2018) 145(4):dev158261. doi: 10.1242/dev.158261
17. Berry WL, Janknecht R. KDM4/JMJD2 Histone Demethylases: Epigenetic Regulators in Cancer Cells. *Cancer Res* (2013) 73(10):2936–42. doi: 10.1158/0008-5472.CAN-12-4300
18. Huang F, Chandrasekharan MB, Chen YC, Bhaskara S, Hiebert SW, Sun ZW. The JmjN Domain of Jhd2 Is Important for its Protein Stability, and the Plant Homeodomain (PHD) Finger Mediates its Chromatin Association Independent of H3K4 Methylation. *J Biol Chem* (2010) 285(32):24548–61. doi: 10.1074/jbc.M110.117333
19. Ozboyaci M, Gursoy A, Erman B, Keskin O. Molecular Recognition of H3/H4 Histone Tails by the Tudor Domains of JMJD2A: A Comparative Molecular Dynamics Simulations Study. *PloS One* (2011) 6(3):e14765. doi: 10.1371/journal.pone.0014765
20. Black JC, Van Rechem C, Whetstine JR. Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. *Mol Cell* (2012) 48(4):491–507. doi: 10.1016/j.molcel.2012.11.006
21. Tan MK, Lim HJ, Harper JW. SCF(FBXO22) Regulates Histone H3 Lysine 9 and 36 Methylation Levels by Targeting Histone Demethylase KDM4A for Ubiquitin-Mediated Proteasomal Degradation. *Mol Cell Biol* (2011) 31(18):3687–99. doi: 10.1128/MCB.05746-11
22. Skarnes WC, Rosen B, West AP, Koutsourakis M, Bushell W, Iyer V, et al. A Conditional Knockout Resource for the Genome-Wide Study of Mouse Gene Function. *Nature* (2011) 474(7351):337–42. doi: 10.1038/nature10163
23. Wan M, Liang J, Xiong Y, Shi F, Zhang Y, Lu W, et al. The Trithorax Group Protein Ash2l Is Essential for Pluripotency and Maintaining Open Chromatin in Embryonic Stem Cells. *J Biol Chem* (2013) 288(7):5039–48. doi: 10.1074/jbc.M112.424515
24. Tsurumi A, Xue S, Zhang L, Li J, Li WX. Genome-Wide Kdm4 Histone Demethylase Transcriptional Regulation in Drosophila. *Mol Genet Genomics* (2019) 294(5):1107–21. doi: 10.1007/s00438-019-01561-z
25. Ferrand J, Rondinelli B, Polo SE. Histone Variants: Guardians of Genome Integrity. *Cells* (2020) 9(11):2424. doi: 10.3390/cells9112424
26. Das PP, Shao Z, Beyaz S, Apostolou E, Pinello L, De Los Angeles A, et al. Distinct and Combinatorial Functions of Jmjd2b/Kdm4b and Jmjd2c/Kdm4c in Mouse Embryonic Stem Cell Identity. *Mol Cell* (2014) 53(1):32–48. doi: 10.1016/j.molcel.2013.11.011
27. Wang J, Zhang M, Zhang Y, Kou Z, Han Z, Chen DY, et al. The Histone Demethylase JMJD2C Is Stage-Specifically Expressed in Preimplantation Mouse Embryos and Is Required for Embryonic Development. *Biol Reprod* (2010) 82(1):105–11. doi: 10.1095/biolreprod.109.078055
28. Ye L, Fan Z, Yu B, Chang J, Al Hezaimi K, Zhou X, et al. Histone Demethylases KDM4B and KDM6B Promotes Osteogenic Differentiation of Human MSCs. *Cell Stem Cell* (2012) 11(1):50–61. doi: 10.1016/j.stem.2012.04.009
29. Scarborough GA, Hennessey JP Jr. Identification of the Major Cytoplasmic Regions of the Neurospora Crassa Plasma Membrane H(+)-ATPase Using Protein Chemical Techniques. *J Biol Chem* (1990) 265(27):16145–9. doi: 10.1016/S0021-9258(17)46200-4
30. Agger K, Nishimura K, Miyagi S, Messling JE, Rasmussen KD, Helin K. The KDM4/JMJD2 Histone Demethylases Are Required for Hematopoietic Stem Cell Maintenance. *Blood* (2019) 134(14):1154–8. doi: 10.1182/blood.2019000855
31. Verrier L, Escaffit F, Chailleux C, Trouche D, Vandromme M. A New Isoform of the Histone Demethylase JMJD2A/KDM4A Is Required for Skeletal Muscle Differentiation. *PloS Genet* (2011) 7(6):e1001390. doi: 10.1371/journal.pgen.1001390
32. Fujiwara K, Fujita Y, Kasai A, Onaka Y, Hashimoto H, Okada H, et al. Deletion of JMJD2B in Neurons Leads to Defective Spine Maturation,

Hyperactive Behavior and Memory Deficits in Mouse. *Trans Psychiatry* (2016) 6:e766. doi: 10.1038/tp.2016.31
33. Choi JY, Yoon SS, Kim SE, Ahn Jo S. KDM4B Histone Demethylase and G9a Regulate Expression of Vascular Adhesion Proteins in Cerebral Microvessels. *Sci Rep* (2017) 7:45005. doi: 10.1038/srep45005
34. Sankar A, Kooistra SM, Gonzalez JM, Ohlsson C, Poutanen M, Helin K. Maternal Expression of the Histone Demethylase Kdm4a Is Crucial for Pre-Implantation Development. *Development* (2017) 144(18):3264–77. doi: 10.1242/dev.155473
35. Iwamori N, Zhao M, Meistrich ML, Matzuk MM. The Testis-Enriched Histone Demethylase, KDM4D, Regulates Methylation of Histone H3 Lysine 9 During Spermatogenesis in the Mouse But Is Dispensable for Fertility. *Biol Reprod* (2011) 84(6):1225–34. doi: 10.1095/biolreprod.110.088955
36. Khoury-Haddad H, Nadar-Ponniah PT, Awwad S, Ayoub N. The Emerging Role of Lysine Demethylases in DNA Damage Response: Dissecting the Recruitment Mode of KDM4D/JMJD2D to DNA Damage Sites. *Cell Cycle* (2015) 14(7):950–8. doi: 10.1080/15384101.2015.1014147
37. Kawazu M, Saso K, Tong KI, McQuire T, Goto K, Son DO, et al. Histone Demethylase JMJD2B Functions as a Co-Factor of Estrogen Receptor in Breast Cancer Proliferation and Mammary Gland Development. *PloS One* (2011) 6(3):e17830. doi: 10.1371/journal.pone.0017830
38. Couse JF, Yates MM, Deroo BJ, Korach KS. Estrogen Receptor-Beta Is Critical to Granulosa Cell Differentiation and the Ovulatory Response to Gonadotropins. *Endocrinology* (2005) 146(8):3247–62. doi: 10.1210/en.2005-0213
39. Emmen JM, Couse JF, Elmore SA, Yates MM, Kissling GE, Korach KS. *In Vitro* Growth and Ovulation of Follicles From Ovaries of Estrogen Receptor (ER){alpha} and ER{beta} Null Mice Indicate a Role for ER{beta} in Follicular Maturation. *Endocrinology* (2005) 146(6):2817–26. doi: 10.1210/en.2004-1108
40. Gaughan L, Stockley J, Coffey K, O'Neill D, Jones DL, Wade M, et al. KDM4B Is a Master Regulator of the Estrogen Receptor Signalling Cascade. *Nucleic Acids Res* (2013) 41(14):6892–904. doi: 10.1093/nar/gkt469
41. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-Cancer Patterns of Somatic Copy Number Alteration. *Nat Genet* (2013) 45(10):1134–40. doi: 10.1038/ng.2760
42. Lin H, Li Q, Li Q, Zhu J, Gu K, Jiang X, et al. Small Molecule KDM4s Inhibitors as Anti-Cancer Agents. *J Enzyme Inhib Med Chem* (2018) 33 (1):777–93. doi: 10.1080/14756366.2018.1455676
43. Black JC, Manning AL, Van Rechem C, Kim J, Ladd B, Cho J, et al. KDM4A Lysine Demethylase Induces Site-Specific Copy Gain and Rereplication of Regions Amplified in Tumors. *Cell* (2013) 154(3):541–55. doi: 10.1016/j.cell.2013.06.051
44. Wissmann M, Yin N, Muller JM, Greschik H, Fodor BD, Jenuwein T, et al. Cooperative Demethylation by JMJD2C and LSD1 Promotes Androgen Receptor-Dependent Gene Expression. *Nat Cell Biol* (2007) 9(3):347–53. doi: 10.1038/ncb1546
45. Shin S, Janknecht R. Activation of Androgen Receptor by Histone Demethylases JMJD2A and JMJD2D. *Biochem Biophys Res Commun* (2007) 359(3):742–6. doi: 10.1016/j.bbrc.2007.05.179
46. Xu W, Jiang K, Shen M, Chen Y, Huang HY. Jumonji Domain Containing 2A Predicts Prognosis and Regulates Cell Growth in Lung Cancer Depending on miR-150. *Oncol Rep* (2016) 35(1):352–8. doi: 10.3892/or.2015.4349
47. Kveiborg M, Albrechtsen R, Couchman JR, Wewer UM. Cellular Roles of ADAM12 in Health and Disease. *Int J Biochem Cell Biol* (2008) 40(9):1685–702. doi: 10.1016/j.biocel.2008.01.025
48. Narayan S, Moyes B, Wolff S. Family Characteristics of Autistic Children: A Further Report. *J Autism Dev Disord* (1990) 20(4):523–35. doi: 10.1007/BF02216057
49. Li A, King J, Moro A, Sugi MD, Dawson DW, Kaplan J, et al. Overexpression of CXCL5 Is Associated With Poor Survival in Patients With Pancreatic Cancer. *Am J Pathol* (2011) 178(3):1340–9. doi: 10.1016/j.ajpath.2010.11.058
50. Hofmann JJ, Luisa Iruela-Arispe M. Notch Expression Patterns in the Retina: An Eye on Receptor-Ligand Distribution During Angiogenesis. *Gene Expr Patterns* (2007) 7(4):461–70. doi: 10.1016/j.modgep.2006.11.002
51. Soini Y, Kosma VM, Pirinen R. KDM4A, KDM4B and KDM4C in non-Small Cell Lung Cancer. *Int J Clin Exp Pathol* (2015) 8(10):12922–8.
52. Li X, Dong S. Histone Demethylase JMJD2B and JMJD2C Induce Fibroblast Growth Factor 2: Mediated Tumorigenesis of Osteosarcoma. *Med Oncol* (2015) 32(3):53. doi: 10.1007/s12032-015-0503-4
53. Zhu Y, van Essen D, Saccani S. Cell-Type-Specific Control of Enhancer Activity by H3K9 Trimethylation. *Mol Cell* (2012) 46(4):408–23. doi: 10.1016/j.molcel.2012.05.011
54. Kim TD, Oh S, Shin S, Janknecht R. Regulation of Tumor Suppressor P53 and HCT116 Cell Physiology by Histone Demethylase JMJD2D/KDM4D. *PloS One* (2012) 7(4):e34618. doi: 10.1371/journal.pone.0034618
55. Mallette FA, Richard S. JMJD2A Promotes Cellular Transformation by Blocking Cellular Senescence Through Transcriptional Repression of the Tumor Suppressor CHD5. *Cell Rep* (2012) 2(5):1233–43. doi: 10.1016/j.celrep.2012.09.033
56. Wilson C, Qiu L, Hong Y, Karnik T, Tadros G, Mau B, et al. The Histone Demethylase KDM4B Regulates Peritoneal Seeding of Ovarian Cancer. *Oncogene* (2017) 36(18):2565–76. doi: 10.1038/onc.2016.412
57. Hu F, Li H, Liu L, Xu F, Lai S, Luo X, et al. Histone Demethylase KDM4D Promotes Gastrointestinal Stromal Tumor Progression Through HIF1beta/VEGFA Signalling. *Mol Cancer* (2018) 17(1):107. doi: 10.1186/s12943-018-0861-6
58. Yang J, Jubb AM, Pike L, Buffa FM, Turley H, Baban D, et al. The Histone Demethylase JMJD2B Is Regulated by Estrogen Receptor Alpha and Hypoxia, and Is a Key Mediator of Estrogen Induced Growth. *Cancer Res* (2010) 70(16):6456–66. doi: 10.1158/0008-5472.CAN-10-0413
59. Fu L, Chen L, Yang J, Ye T, Chen Y, Fang J. HIF-1alpha-Induced Histone Demethylase JMJD2B Contributes to the Malignant Phenotype of Colorectal Cancer Cells *via* an Epigenetic Mechanism. *Carcinogenesis* (2012) 33 (9):1664–73. doi: 10.1093/carcin/bgs217
60. Beyer S, Kristensen MM, Jensen KS, Johansen JV, Staller P. The Histone Demethylases JMJD1A and JMJD2B Are Transcriptional Targets of Hypoxia-Inducible Factor HIF. *J Biol Chem* (2008) 283(52):36542–52. doi: 10.1074/jbc.M804578200
61. Ye Q, Holowatyj A, Wu J, Liu H, Zhang L, Suzuki T, et al. Genetic Alterations of KDM4 Subfamily and Therapeutic Effect of Novel Demethylase Inhibitor in Breast Cancer. *Am J Cancer Res* (2015) 5 (4):1519–30.
62. Kim TD, Shin S, Berry WL, Oh S, Janknecht R. The JMJD2A Demethylase Regulates Apoptosis and Proliferation in Colon Cancer Cells. *J Cell Biochem* (2012) 113(4):1368–76. doi: 10.1002/jcb.24009
63. Zhang D, Yoon HG, Wong J. JMJD2A Is a Novel N-CoR-Interacting Protein and Is Involved in Repression of the Human Transcription Factor Achaete Scute-Like Homologue 2 (ASCL2/Hash2). *Mol Cell Biol* (2005) 25(15):6404–14. doi: 10.1128/MCB.25.15.6404-6414.2005
64. Patani N, Jiang WG, Newbold RF, Mokbel K. Histone-Modifier Gene Expression Profiles Are Associated With Pathological and Clinical Outcomes in Human Breast Cancer. *Anticancer Res* (2011) 31(12):4115–25.
65. Berry WL, Shin S, Lightfoot SA, Janknecht R. Oncogenic Features of the JMJD2A Histone Demethylase in Breast Cancer. *Int J Oncol* (2012) 41 (5):1701–6. doi: 10.3892/ijo.2012.1618
66. Li LL, Xue AM, Li BX, Shen YW, Li YH, Luo CL, et al. JMJD2A Contributes to Breast Cancer Progression Through Transcriptional Repression of the Tumor Suppressor ARHI. *Breast Cancer Res* (2014) 16(3):R56. doi: 10.1186/bcr3667
67. Li L, Gao P, Li Y, Shen Y, Xie J, Sun D, et al. JMJD2A-Dependent Silencing of Sp1 in Advanced Breast Cancer Promotes Metastasis by Downregulation of DIRAS3. *Breast Cancer Res Treat* (2014) 147(3):487–500. doi: 10.1007/s10549-014-3083-7
68. Li BX, Luo CL, Li H, Yang P, Zhang MC, Xu HM, et al. Effects of siRNA-Mediated Knockdown of Jumonji Domain Containing 2A on Proliferation, Migration and Invasion of the Human Breast Cancer Cell Line MCF-7. *Exp Ther Med* (2012) 4(4):755–61. doi: 10.3892/etm.2012.662
69. Li BX, Zhang MC, Luo CL, Yang P, Li H, Xu HM, et al. Effects of RNA Interference-Mediated Gene Silencing of JMJD2A on Human Breast Cancer Cell Line MDA-MB-231 *In Vitro*. *J Exp Clin Cancer Res* (2011) 30:90. doi: 10.1186/1756-9966-30-90
70. Wang J, Wang H, Wang LY, Cai D, Duan Z, Zhang Y, et al. Silencing the Epigenetic Silencer KDM4A for TRAIL and DR5 Simultaneous Induction

and Antitumor Therapy. *Cell Death Differ* (2016) 23(11):1886–96. doi: 10.1038/cdd.2016.92
71. Black JC, Zhang H, Kim J, Getz G, Whetstine JR. Regulation of Transient Site-Specific Copy Gain by MicroRNA. *J Biol Chem* (2016) 291(10):4862–71. doi: 10.1074/jbc.M115.711648
72. Guerra-Calderas L, Gonzalez-Barrios R, Patino CC, Alcaraz N, Salgado-Albarran M, de Leon DC, et al. CTCF-KDM4A Complex Correlates With Histone Modifications That Negatively Regulate CHD5 Gene Expression in Cancer Cell Lines. *Oncotarget* (2018) 9(24):17028–42. doi: 10.18632/oncotarget.24798
73. Levin M, Stark M, Assaraf YG. The JmjN Domain as a Dimerization Interface and a Targeted Inhibitor of KDM4 Demethylase Activity. *Oncotarget* (2018) 9(24):16861–82. doi: 10.18632/oncotarget.24717
74. Shi L, Sun L, Li Q, Liang J, Yu W, Yi X, et al. Histone Demethylase JMJD2B Coordinates H3K4/H3K9 Methylation and Promotes Hormonally Responsive Breast Carcinogenesis. *Proc Natl Acad Sci USA* (2011) 108 (18):7541–6. doi: 10.1073/pnas.1017374108
75. Jones D, Wilson L, Thomas H, Gaughan L, Wade MA. The Histone Demethylase Enzymes KDM3A and KDM4B Co-Operatively Regulate Chromatin Transactions of the Estrogen Receptor in Breast Cancer. *Cancers* (2019) 11(8):1122. doi: 10.3390/cancers11081122
76. Xiang Y, Yan K, Zheng Q, Ke H, Cheng J, Xiong W, et al. Histone Demethylase KDM4B Promotes DNA Damage by Activating Long Interspersed Nuclear Element-1. *Cancer Res* (2019) 79(1):86–98. doi: 10.1158/0008-5472.CAN-18-1310
77. Slee RB, Steiner CM, Herbert BS, Vance GH, Hickey RJ, Schwarz T, et al. Cancer-Associated Alteration of Pericentromeric Heterochromatin may Contribute to Chromosome Instability. *Oncogene* (2012) 31(27):3244–53. doi: 10.1038/onc.2011.502
78. Wang W, Oguz G, Lee PL, Bao Y, Wang P, Terp MG, et al. KDM4B-Regulated Unfolded Protein Response as a Therapeutic Vulnerability in PTEN-Deficient Breast Cancer. *J Exp Med* (2018) 215(11):2833–49. doi: 10.1084/jem.20180439
79. Johmura Y, Maeda I, Suzuki N, Wu W, Goda A, Morita M, et al. Fbxo22-Mediated KDM4B Degradation Determines Selective Estrogen Receptor Modulator Activity in Breast Cancer. *J Clin Invest* (2018) 128(12):5603–19. doi: 10.1172/JCI121679
80. Liu G, Bollig-Fischer A, Kreike B, van de Vijver MJ, Abrams J, Ethier SP, et al. Genomic Amplification and Oncogenic Properties of the GASC1 Histone Demethylase Gene in Breast Cancer. *Oncogene* (2009) 28 (50):4491–500. doi: 10.1038/onc.2009.297
81. Hong Q, Yu S, Yang Y, Liu G, Shao Z. A Polymorphism in JMJD2C Alters the Cleavage by Caspase-3 and the Prognosis of Human Breast Cancer. *Oncotarget* (2014) 5(13):4779–87. doi: 10.18632/oncotarget.2029
82. Berdel B, Nieminen K, Soini Y, Tengstrom M, Malinen M, Kosma VM, et al. Histone Demethylase GASC1–A Potential Prognostic and Predictive Marker in Invasive Breast Cancer. *BMC Cancer* (2012) 12:516. doi: 10.1186/1471-2407-12-516
83. Luo W, Chang R, Zhong J, Pandey A, Semenza GL. Histone Demethylase JMJD2C Is a Coactivator for Hypoxia-Inducible Factor 1 That Is Required for Breast Cancer Progression. *Proc Natl Acad Sci USA* (2012) 109(49): E3367–76. doi: 10.1073/pnas.1217394109
84. Garcia J, Lizcano F. KDM4C Activity Modulates Cell Proliferation and Chromosome Segregation in Triple-Negative Breast Cancer. *Breast Cancer* (2016) 10:169–75. doi: 10.4137/BCBCR.S40182
85. Walsh CA, Bolger JC, Byrne C, Cocchiglia S, Hao Y, Fagan A, et al. Global Gene Repression by the Steroid Receptor Coactivator SRC-1 Promotes Oncogenesis. *Cancer Res* (2014) 74(9):2533–44. doi: 10.1158/0008-5472.CAN-13-2133
86. Hillringhaus L, Yue WW, Rose NR, Ng SS, Gileadi C, Loenarz C, et al. Structural and Evolutionary Basis for the Dual Substrate Selectivity of Human KDM4 Histone Demethylase Family. *J Biol Chem* (2011) 286 (48):41616–25. doi: 10.1074/jbc.M111.283689
87. Sanchez-Fernandez EM, Tarhonskaya H, Al-Qahtani K, Hopkinson RJ, McCullagh JS, Schofield CJ, et al. Investigations on the Oxygen Dependence of a 2-Oxoglutarate Histone Demethylase. *Biochem J* (2013) 449(2):491–6. doi: 10.1042/BJ20121155
88. Chin YW, Han SY. KDM4 Histone Demethylase Inhibitors for Anti-Cancer Agents: A Patent Review. *Expert Opin Ther Pat* (2015) 25(2):135–44. doi: 10.1517/13543776.2014.991310
89. Smith EH, Janknecht R, Maher LJ3rd. Succinate Inhibition of Alpha-Ketoglutarate-Dependent Enzymes in a Yeast Model of Paraganglioma. *Hum Mol Genet* (2007) 16(24):3136–48. doi: 10.1093/hmg/ddm275
90. Chen YK, Bonaldi T, Cuomo A, Del Rosario JR, Hosfield DJ, Kanouni T, et al. Design of KDM4 Inhibitors With Antiproliferative Effects in Cancer Models. *ACS Med Chem Lett* (2017) 8(8):869–74. doi: 10.1021/acsmedchemlett.7b00220
91. Metzger E, Stepputtis SS, Strietz J, Preca BT, Urban S, Willmann D, et al. KDM4 Inhibition Targets Breast Cancer Stem-Like Cells. *Cancer Res* (2017) 77(21):5900–12. doi: 10.1158/0008-5472.CAN-17-1754
92. Yoo S, Chandhasin C, Rosario JRD, Chen YK, Stafford J, Quake S, et al. Abstract 2128: TACH101, A First-in-Class Pan Inhibitor of KDM4 Histone Lysine Demethylases. *Cancer Res* (2021) 81(13 Supplement):2128. doi: 10.1158/1538-7445.am2021-2128
93. Duan L, Rai G, Roggero C, Zhang Q-J, Wei Q, Ma Shi H, et al. KDM4/JMJD2 Histone Demethylase Inhibitors Block Prostate Tumor Growth by Suppressing the Expression of AR and BMYB-Regulated Genes. *Chem Biol* (2015) 22(9):1185–96. doi: 10.1016/j.chembiol.2015.08.007
94. Rotili D, Mattevi A. At Long Last Potent and Selective KDM5 Inhibitors. *Cell Chem Biol* (2016) 23(7):749–51. doi: 10.1016/j.chembiol.2016.07.003
95. Hatch SB, Yapp C, Montenegro RC, Savitsky P, Gamble V, Tumber A, et al. Assessing Histone Demethylase Inhibitors in Cells: Lessons Learned. *Epigenet Chromatin* (2017) 10:9. doi: 10.1186/s13072-017-0116-6
96. Gale M, Sayegh J, Cao J, Norcia M, Gareiss P, Hoyer D, et al. Screen-Identified Selective Inhibitor of Lysine Demethylase 5A Blocks Cancer Cell Growth and Drug Resistance. *Oncotarget* (2016) 7(26):39931–44. doi: 10.18632/oncotarget.9539
97. Luo X, Liu Y, Kubicek S, Myllyharju J, Tumber A, Ng S, et al. A Selective Inhibitor and Probe of the Cellular Functions of Jumonji C Domain-Containing Histone Demethylases. *J Am Chem Soc* (2011) 133(24):9451–6. doi: 10.1021/ja201597b
98. Kim TD, Fuchs JR, Schwartz E, Abdelhamid D, Etter J, Berry WL, et al. Pro-Growth Role of the JMJD2C Histone Demethylase in HCT-116 Colon Cancer Cells and Identification of Curcuminoids as JMJD2 Inhibitors. *Am J Trans Res* (2014) 6(3):236–47.
99. Souto JA, Sarno F, Nebbioso A, Papulino C, Alvarez R, Lombino J, et al. A New Family of Jumonji C Domain-Containing KDM Inhibitors Inspired by Natural Product Purpurogallin. *Front Chem* (2020) 8:312. doi: 10.3389/fchem.2020.00312
100. Leurs U, Lohse B, Rand KD, Ming S, Riise ES, Cole PA, et al. Substrate- and Cofactor-Independent Inhibition of Histone Demethylase KDM4C. *ACS Chem Biol* (2014) 9(9):2131–8. doi: 10.1021/cb500374f
101. Kawamura A, Munzel M, Kojima T, Yapp C, Bhushan B, Goto Y, et al. Highly Selective Inhibition of Histone Demethylases by *De Novo* Macrocyclic Peptides. *Nat Commun* (2017) 8:14773. doi: 10.1038/ncomms14773
102. Majchrzak-Celinska A, Warych A, Szoszkiewicz M. Novel Approaches to Epigenetic Therapies: From Drug Combinations to Epigenetic Editing. *Genes* (2021) 12(2):208. doi: 10.3390/genes12020208
103. Nepali K, Liou JP. Recent Developments in Epigenetic Cancer Therapeutics: Clinical Advancement and Emerging Trends. *J Biomed Sci* (2021) 28(1):27. doi: 10.1186/s12929-021-00721-x
104. George P, Bali P, Annavarapu S, Scuto A, Fiskus W, Guo F, et al. Combination of the Histone Deacetylase Inhibitor LBH589 and the Hsp90 Inhibitor 17-AAG Is Highly Active Against Human CML-BC Cells and AML Cells With Activating Mutation of FLT-3. *Blood* (2005) 105(4):1768–76. doi: 10.1182/blood-2004-09-3413
105. Pathania R, Ramachandran S, Mariappan G, Thakur P, Shi H, Choi JH, et al. Combined Inhibition of DNMT and HDAC Blocks the Tumorigenicity of Cancer Stem-Like Cells and Attenuates Mammary Tumor Growth. *Cancer Res* (2016) 76(11):3224–35. doi: 10.1158/0008-5472.CAN-15-2249
106. Kalin JH, Wu M, Gomez AV, Song Y, Das J, Hayward D, et al. Targeting the CoREST Complex With Dual Histone Deacetylase and Demethylase Inhibitors. *Nat Commun* (2018) 9(1):53. doi: 10.1038/s41467-017-02242-4
107. Zhang X, Zegar T, Weiser T, Hamdan FH, Berger BT, Lucas R, et al. Characterization of a Dual BET/HDAC Inhibitor for Treatment of Pancreatic Ductal Adenocarcinoma. *Int J Cancer* (2020) 147(10):2847–61. doi: 10.1002/ijc.33137

108. Romanelli A, Stazi G, Fioravanti R, Zwergel C, Di Bello E, Pomella S, et al. Design of First-In-Class Dual EZH2/HDAC Inhibitor: Biochemical Activity and Biological Evaluation in Cancer Cells. *ACS Med Chem Lett* (2020) 11 (5):977–83. doi: 10.1021/acsmedchemlett.0c00014
109. Lue JK, Prabhu SA, Liu Y, Gonzalez Y, Verma A, Mundi PS, et al. Precision Targeting With EZH2 and HDAC Inhibitors in Epigenetically Dysregulated Lymphomas. *Clin Cancer Res* (2019) 25(17):5271–83. doi: 10.1158/1078-0432.CCR-18-3989
110. Thakur A, Tawa GJ, Henderson MJ, Danchik C, Liu S, Shah P, et al. Design, Synthesis, and Biological Evaluation of Quinazolin-4-One-Based Hydroxamic Acids as Dual PI3K/HDAC Inhibitors. *J Med Chem* (2020) 63 (8):4256–92. doi: 10.1021/acs.jmedchem.0c00193
111. Benedetti R, Dell'Aversana C, De Marchi T, Rotili D, Liu NQ, Novakovic B, et al. Inhibition of Histone Demethylases LSD1 and UTX Regulates ERalpha Signaling in Breast Cancer. *Cancers* (2019) 11(12):2027. doi: 10.3390/cancers11122027
112. Damaskos C, Garmpis N, Valsami S, Kontos M, Spartalis E, Kalampokas T, et al. Histone Deacetylase Inhibitors: An Attractive Therapeutic Strategy Against Breast Cancer. *Anticancer Res* (2017) 37(1):35–46. doi: 10.21873/anticanres.11286
113. Meng J, Dai B, Fang B, Bekele BN, Bornmann WG, Sun D, et al. Combination Treatment With MEK and AKT Inhibitors Is More Effective Than Each Drug Alone in Human non-Small Cell Lung Cancer *In Vitro* and *In Vivo*. *PloS One* (2010) 5(11):e14124. doi: 10.1371/journal.pone.0014124
114. Doroshow JH, Simon RM. On the Design of Combination Cancer Therapy. *Cell* (2017) 171(7):1476–8. doi: 10.1016/j.cell.2017.11.035
115. Sessions Z, Sanchez-Cruz N, Prieto-Martinez FD, Alves VM, Santos HP Jr, Muratov E, et al. Recent Progress on Cheminformatics Approaches to Epigenetic Drug Discovery. *Drug Discov Today* (2020) 25(12):2268–76. doi: 10.1016/j.drudis.2020.09.021
116. Whetstine JR, Nottke A, Lan F, Huarte M, Smolikov S, Chen Z, et al. Reversal of Histone Lysine Trimethylation by the JMJD2 Family of Histone Demethylases. *Cell* (2006) 125(3):467–81. doi: 10.1016/j.cell.2006.03.028
117. Trojer P, Zhang J, Yonezawa M, Schmidt A, Zheng H, Jenuwein T, et al. Dynamic Histone H1 Isotype 4 Methylation and Demethylation by Histone Lysine Methyltransferase G9a/KMT1C and the Jumonji Domain-Containing JMJD2/KDM4 Proteins. *J Biol Chem* (2009) 284(13):8395–405. doi: 10.1074/jbc.M807818200

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 10 November 2021
doi: 10.3389/fonc.2021.752651

Plasma-Derived Extracellular Vesicles Circular RNAs Serve as Biomarkers for Breast Cancer Diagnosis

Li Lin[1†], Geng-Xi Cai[2,3†], Xiang-Ming Zhai[1†], Xue-Xi Yang[1], Min Li[1], Kun Li[4], Chun-Lian Zhou[1], Tian-Cai Liu[1], Bo-Wei Han[1], Zi-Jia Liu[1], Mei-Qi Chen[1], Guo-Lin Ye[2], Ying-Song Wu[1*] and Zhi-Wei Guo[1*]*

[1] Key Laboratory of Antibody Engineering of Guangdong Higher Education Institutes, School of Laboratory Medical and Biotechnology, Southern Medical University, Guangzhou, China, [2] Department of Breast Surgery, The First People's Hospital of Foshan, Foshan, China, [3] Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China, [4] Department of Cancer Biology, Guangzhou XGene Co., Ltd., Guangzhou, China

Breast cancer is the second cause of cancer-associated death among women and seriously endangers women's health. Therefore, early identification of breast cancer would be beneficial to women's health. At present, circular RNA (circRNA) not only exists in the extracellular vesicles (EVs) in plasma, but also presents distinct patterns under different physiological and pathological conditions. Therefore, we assume that circRNA could be used for early diagnosis of breast cancer. Here, we developed classifiers for breast cancer diagnosis that relied on 259 samples, including 144 breast cancer patients and 115 controls. In the discovery stage, we compared the genome-wide long RNA profiles of EVs in patients with breast cancer (n=14) and benign breast (n=6). To further verify its potential in early diagnosis of breast cancer, we prospectively collected plasma samples from 259 individuals before treatment, including 144 breast cancer patients and 115 controls. Finally, we developed and verified the predictive classifies based on their circRNA expression profiles of plasma EVs by using multiple machine learning models. By comparing their circRNA profiles, we found 439 circRNAs with significantly different levels between cancer patients and controls. Considering the cost and practicability of the test, we selected 20 candidate circRNAs with elevated levels and detected their levels by quantitative real-time polymerase chain reaction. In the training cohort, we found that BC_{ExoC}, a nine-circRNA combined classifier with SVM model, achieved the largest AUC of 0.83 [95% CI 0.77-0.88]. In the validation cohort, the predictive efficacy of the classifier achieved 0.80 [0.71-0.89]. Our work reveals the application prospect of circRNAs in plasma EVs as non-invasive liquid biopsies in the diagnosis and management of breast cancer.

Keywords: breast cancer, cancer diagnosis, extracellular vesicles, circular RNA, predictive classifier

OPEN ACCESS

Edited by:
Biao Huang,
Jiangsu Institute of Nuclear Medicine, China

Reviewed by:
Zhaohui Huang,
Affiliated Hospital of Jiangnan University, China
Lei Zheng,
Southern Medical University, China
Cheng Yang,
Nankai University, China

****Correspondence:***
Guo-Lin Ye
yglin@fsyyy.com
Ying-Song Wu
wg@smu.edu.cn
Zhi-Wei Guo
gzw188@126.com

†These authors share first authorship

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: *03 August 2021*
Accepted: *15 October 2021*
Published: *10 November 2021*

Citation:
Lin L, Cai G-X, Zhai X-M, Yang X-X, Li M, Li K, Zhou C-L, Liu T-C, Han B-W, Liu Z-J, Chen M-Q, Ye G-L, Wu Y-S and Guo Z-W (2021) Plasma-Derived Extracellular Vesicles Circular RNAs Serve as Biomarkers for Breast Cancer Diagnosis. Front. Oncol. 11:752651. doi: 10.3389/fonc.2021.752651

INTRODUCTION

Breast cancer is a major kind of malignant tumor that seriously endangers women's health. According to cancer statistics in 2018, breast cancer accounts for more than 10% of all new diagnoses and causes about 600,000 deaths every year (1). Although the overall prognosis of breast cancer is good, the five-year relative survival rate of stage IV patients is still lower than 30% (2). Early diagnosis of cancer could effectively improve their therapeutic effects. Therefore, it is necessary to develop an early diagnosis method for breast cancer identification.

Plasma extracellular vesicles (EVs), such as exosomes and microvesicles, are mainly derived from cancer and hematopoietic cells in cancer patients, which host cell-information of their original tissues (3, 4). Since the contents of EVs could reflect the characteristics of cancer cells, they have been used to develop a variety of non-invasive methods for cancer-related applications, such as early diagnosis and prognosis prediction of cancer (5–8). For example, microRNAs and proteins derived from EVs have been used for the early diagnosis of various cancers (5, 9, 10). However, instability of microRNA and low abundance of proteins may limit their clinical applications. Therefore, a stable biomarker with appropriate concentrations may be more suitable for the early diagnosis of breast cancer.

Circular RNA (circRNA) is a new type of RNA, which shows the remarkable feature of being covalently closed continuous loops without 5' to 3' polar structure (11). CircRNA is stable, temporospatial (often exhibit type-specific, tissue-specific, and stage-specific manner), and conserved (12, 13). Functional studies have shown that they may play important roles in tumorigenesis by becoming microRNA sponge or translating into proteins (14, 15). Recently, a variety of RNAs, especially circRNAs, were discovered in EVs of different types of cancer (16, 17). Since circRNA is stable and type-specific, we assume that the circRNA in the EVs can be used for early diagnosis of breast cancer.

In this study, we first implemented genome-wide long RNA sequencing to determine the difference of RNA profiles in EVs of plasma between breast cancer patients and controls. In the training stage, the circRNAs with significantly different levels were selected and their relative levels were evaluated among 182 participants by quantitative real-time polymerase chain reaction (qPCR). Based on the relative levels, we then constructed the diagnosis classifier with multiple machine learning models, including SVM, LR, and LDA. According to the results of their cross-validation, the classifier with the largest AUC was selected, and its performance was further studied in the validation cohort.

METHODS

Participants and Research Design

In total, we collected 259 plasma samples from two groups of individuals: breast cancer patients and controls (including healthy individuals and benign breast patients including fibroadenoma and benign epithelial proliferation, **Figure 1**). Participants were enrolled from the First People's Hospital of Foshan and plasma samples of all cancer patients were collected prospectively before cancer therapy. The samples used in the discovery stage were collected prospectively from January 2018 to July 2018. Plasma samples used in the training and validation stage were collected prospectively from August 2018 to May 2019. All plasma samples were obtained under institutional review board of the First People's Hospital of Foshan approved protocols with written informed consent from all participants for research use [ID: L(2021)-7]. More details about the clinical information of all participants involved in this study were shown in **Supplemental Table 1**.

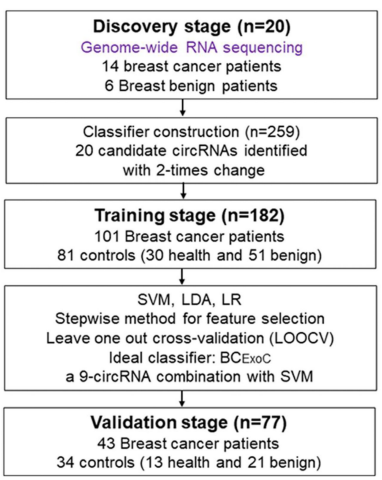


FIGURE 1 | Study design. To develop classifiers for the early diagnosis of breast cancer, the workflow of our study consists of three stages, including the discovery stage, training, and validation stage. In the discovery stage, we used whole-genome sequencing to identify circRNAs with significantly different levels. In the training stage, we developed classifiers with three regression models by using the circRNAs levels detected by qPCR. In the validation stage, the predictive efficacy of the classifiers was validated. qPCR, quantitative real-time polymerase chain reaction. circRNA, circular RNA.

RNA Extraction and Sequencing

RNAs of EVs were extracted from about 5 mL of plasma with exoRNeasy Serum/Plasma Maxi Kit following the manufacturer's instructions (QIAGEN, Germany). In brief, the plasma was prefiltered, then was mixed with 2x binding buffer. The mixture is added to the exoEasy membrane affinity to bind the EVs to the membrane. After centrifugation, the wash buffer was added to wash off non-specific material in the column. After enriching EVs, QIAZOL was added to the column to lyse the vesicles and chloroform was added to the lysate collected after centrifugation. After the aqueous phase is recovered and mixed with ethanol, and the sample-ethanol mixture is added to the RNeasy MinElute spin column and centrifuged. Washing the column with buffer RWT, then wash twice with buffer RPE. And finally elute RNA in water. The rRNAs in total RNAs were first removed using Ribo-Zero rRNA Removal Kits (Illumina, USA) and the libraries of RNA-sequencing were constructed with TruSeq Stranded Total RNA Library Prep Kit (Illumina, USA). Subsequently, quality and quantification of libraries were assessed using the BioAnalyzer 2100 system (Agilent Technologies, USA). Finally, 10 pM libraries were denatured as single-stranded DNA molecules, captured on Illumina flow cells, amplified *in situ* as clusters, and finally sequenced for 150 cycles on Illumina HiSeq 4000 sequencer following the manufacturer's instructions.

Process of High Throughput RNA-Sequencing Data

The 3' adapter of the raw read was trimmed, and the low-quality read was removed by using cutadapt software (v1.9.3). At first, the reads were aligned to the reference genome (hg19) and transcriptome with STAR software (v2.5.1b) (18). Then, the circRNAs were detected and identified by DCC software (v0.4.4) (19). According to their genomic localization of known genes, the circRNAs were separated into five different types, including exon, intronic, intergenic, antisense, and sense overlapping circRNAs. In addition, the identified circRNAs were annotated with circBase and some previous studies (20–22). Normalized expression values of circRNAs were calculated by using edgeR software (v3.16.5) (23). For LncRNA and mRNA, the reads were aligned to the human reference genome with hisat2 software (v2.0.4) (24).

The dysregulated circRNAs were determined by the edgeR package of R software with a cutoff threshold of $|\log_2 \text{fold change}| \geq 2$ and P-value < 0.05 (**Supplemental Table 2**). The principal component analysis (PCA) and result visualization were realized by rgl package (v0.1). The enrichment of GO function and KEGG pathway were implemented and visualized by using Metascape (25) and OmicShare tools (www.omicshare.com/tools).

Detection of qPCR

TaKaRa PrimeScript™ RT reagent was done with equal quality of input RNA. The qPCR for human circRNAs was done on an Applied Biosystems 7500 Real-Time PCR System using the TaKaRa TB Green™ Premix Ex Taq™ II. The value of the cycle threshold (Ct) was processed and exported by the software of Applied Biosystems SDS (v2.3.0). CircRNAs from the training and testing cohort were detected by qPCR with a human endogenous mRNA, U6, as a reference. Relative quantification was used and the levels of circRNAs were normalized against the level of reference by $2^{-\Delta ct}$, where $\Delta Ct = Ct_{target} - Ct_{reference}$. Their primer sequence was shown in **Supplemental Table 3**.

Construction of Classifiers for Early Diagnosis of Breast Cancer

The workflow of classifier construction was shown in **Figure 1**. Firstly, the circRNAs with significantly elevated levels in breast cancer patients were selected. In the training stage, the relative level of 20 candidate circRNAs in 182 participants, including 101 breast cancer patients and 81 controls (30 healthy individuals and 51 breast benign patients [42 fibroadenoma and 9 benign epithelial proliferation]), was assessed using qPCR (**Figure 1**). To construct circRNA classifiers that could distinguish breast cancer patients from controls, the qPCR was used to develop classifiers with three regression modes, including support vector machine (SVM), logistic regression (LR), and linear discriminate analysis (LDA). The SVM classifier was constructed with the linear kernel in e1071 package using the default setting. The glm and lda function in base package of R software was used to develop the LR and LDA classifier with default setting, respectively.

Since quite a number of studies have reported that discrete data may improve classifier performance (26), before classifier construction, the continuous variable was first discretized according to the optimal cut-off point. The optimal cut-off point of each variable was defined as the maximum value of (sensitivity + specificity)/2 in the training cohort. Then the continuous value set to one when it was larger than the corresponding optimal cut-off in each subject; Otherwise, it was set to zero (**Supplemental Table 4**). The stepwise method was used to select the optimal classifier with the largest AUC. To estimate the robustness and prediction error of the selected classifiers, we applied the leave one out cross-validation (LOOCV) method. Briefly, each subject in the training cohort was withheld in turn, and the rest of subjects were submitted to train the model. As there were 182 samples in the training cohort, this procedure was repeated 182 times. In the validation cohort, the relative levels of the circRNAs in the selected classifiers were detected, which included 77 participants, including 43 breast cancer patients and 34 controls [13 healthy individuals and 21 breast benign patients (17 fibroadenoma and 4 benign epithelial proliferation)]. Finally, the predictive efficacy of the optimal classifier in the validation cohort was calculated.

Statistical Analysis

The fisher exact test and the χ^2 test were used for comparison of categorical variables. P-value < 0.05 for two-sided tests was considered to be statistically significant. Hierarchical clustering was applied to the circRNAs with significantly different levels, using the average-linkage clustering algorithms in Cluster (ver. 3.0). Heat maps were plotted using the pheatmap package of R

(version 3.0.1). The receiver operating characteristic curve (ROC) was drawn and the difference of the area under the curve (AUC) was calculated by using the pROC package (27).

RESULTS

Genome-Wide Long RNA Profiles of Plasma EVs

Previous studies have shown that a variety of RNA was found in EVs of plasma. Therefore, we first analyzed the genome-wide long RNA profiles of breast cancer patients and controls. We found that there existed different types of RNAs in EV, such as circRNA, lncRNA, and mRNA, and each type of RNA showed many entities [circRNA (n=34,749), lncRNA (n=68,298) and mRNA (n=20,324); **Figure 2A**]. The amount of circRNAs derived from breast cancer patients was significantly higher than that of benign patients (**Figure 2A**). However, this phenomenon has not been observed in lncRNA and mRNA. In addition, the circular structure of circRNAs was normally more stable than the linear RNA, so they may be suitable to be disease biomarkers.

By using public databases and literatures to annotate the circRNAs, we found that approximately 71.30% of circRNAs were novel circRNAs (**Figure 2B**). According to previous classification criteria, we characterized the circRNAs into five types, including exonic, intronic, intergenic, sense overlapping and antisense circRNAs. We found that the exonic and intronic circRNAs took up the largest proportion of circRNAs (55.1%, **Figure 2C**). By further analyzing their length distribution, we found that the majority of circRNAs in EVs were less than 2,500 nucleotides (nt), which took up over 85.92% (**Figure 2D**).

Distinct CircRNA Profiles Between Breast Cancer Patients and Controls

The workflow of classifier construction was shown in **Figure 1**. In the discovery stage, by comparing the circRNA profiles of 14 breast cancer and 6 benign patients, we identified 439 circRNAs of EVs with significantly different levels, including 162 increased and 277 decreased circRNAs, and the cut-off threshold ($|\log_2$ fold change$|>2$, P-value<0.05) was calculated by edgR (**Figure 3A**, **Supplemental Table 2**). The PCA results showed that the expression profiles between cancer patients and controls showed different patterns (**Figure 3B**). There is approximately 64.69% of circRNAs were novel circRNAs and the exonic and

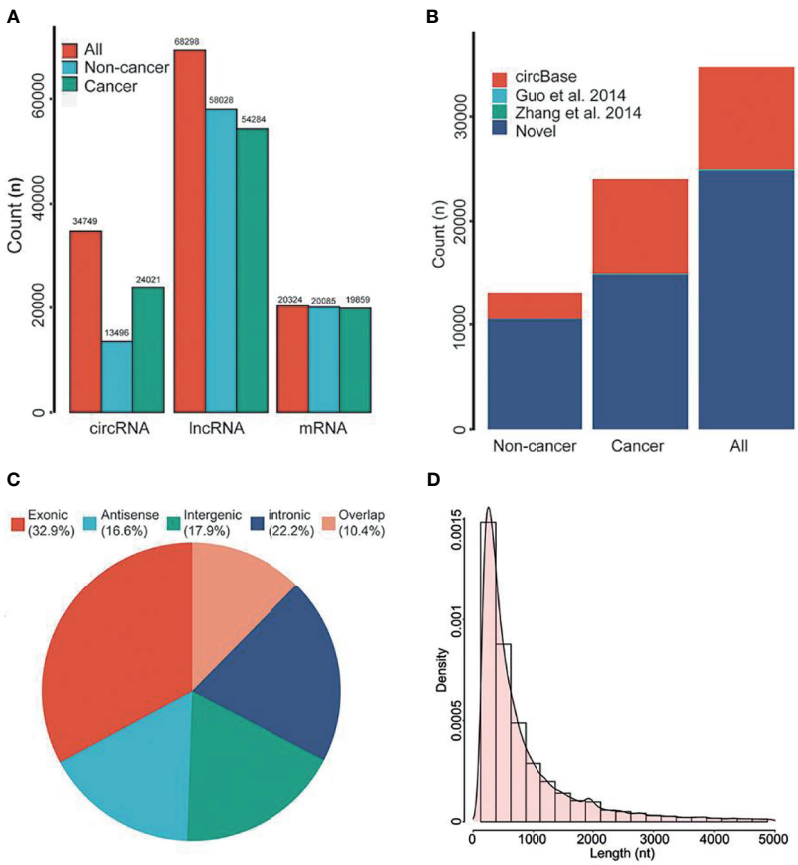


FIGURE 2 | RNA composition in EVs. **(A)** The types of RNAs in EVs. **(B)** Annotation of circRNAs. **(C)** Source of circRNAs. **(D)** Length distribution of circRNAs. nt, Nucleotide. All, all of individuals. Overlap, sense overlapping circRNAs.

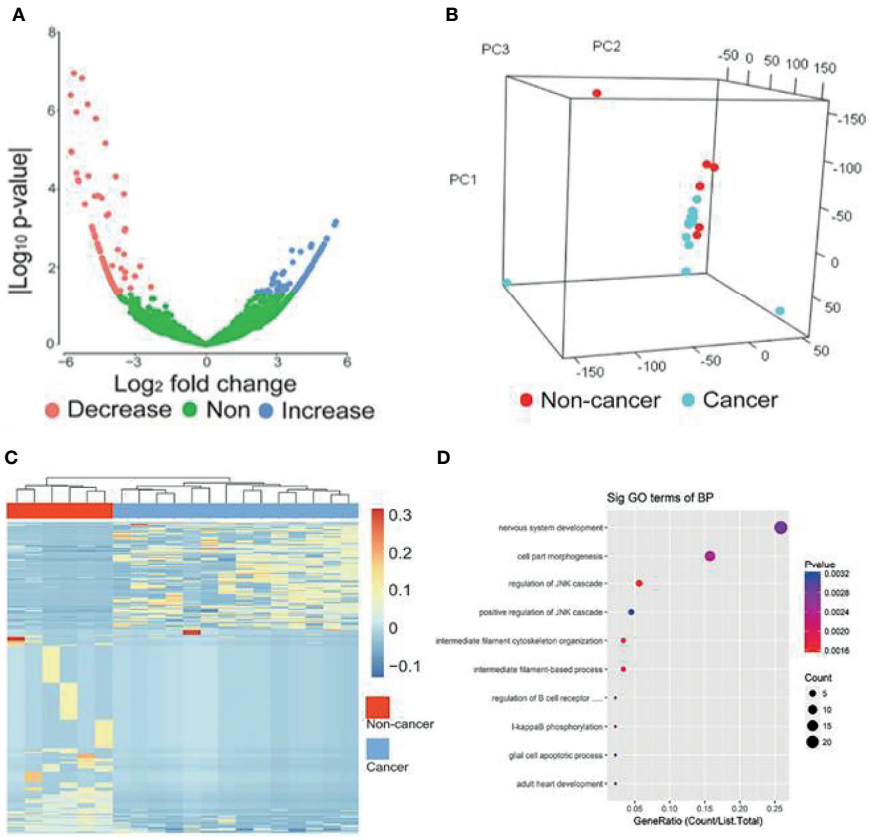


FIGURE 3 | circRNAs with significantly different levels. **(A)** Volcano plots of circRNAs with significant different levels (|log2 fold change| ≥ 2 and *P*-value < 0.05 produced by edgR package of R software) between cancer and control groups. **(B)** PCA analysis of genome-wide RNA sequencing data derived from 14 breast cancer patients and 6 benign patients. **(C)** Heat map of the z-scores of circRNAs with significantly different levels. **(D)** Gene function enrichment analysis of the host genes of the circRNAs with significantly different levels. Decrease, circRNAs with decreased levels. Non, circRNAs with non-significant changes. Increase, circRNAs increased levels.

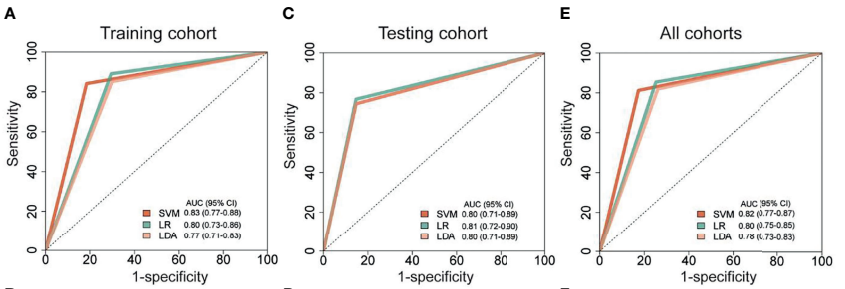
intronic type took up 54.44% among the dysregulated circRNAs. Next, we implemented unsupervised cluster analysis, and found that there was a distinct pattern between the cancer patients and controls (**Figure 3C**), which indicated that circRNAs in EVs may be used for the diagnosis of breast cancer.

To further reveal the relationships between the dysregulated circRNAs and breast cancer, we implemented gene function enrichment analysis on the host genes of circRNAs. The results showed that these terms were enriched in multiple processes, such as cell part morphogenesis, regulation of the JNK pathway and I-kappaB phosphorylation (**Figure 3D**). Previous studies have reported that the enriched pathways were related to the tumorigenesis of breast cancer. For example, the JNK pathway influences proliferation, differentiation, survival and migration in different cancers (28).

Classifiers for Early Diagnosis of Breast Cancer

Since the up-regulated features were more practical in clinical detection, we focused on the 20 circRNAs, which were increased in breast cancer patients compared to controls (**Supplemental Figure 1**). In the training cohort, we evaluated the relative levels of 20 increased circRNAs in 182 plasma samples, including 101 breast cancer patients and 81 controls. Three regression models, including SVM, LDA and LR, were used to construct circRNA classifiers which could distinguish breast cancer patients from controls. The AUC, accuracy, sensitivity and specificity of the classifiers were cross-verified by LOOCV cross-validation method (**Figure 4**). Among all combinations with three different regression models, a nine-circRNA combination, named BC_{ExoC}, achieved high performance [AUC=0.83 (95% confidence interval 0.77-0.88) and accuracy=0.83] in the training cohort after LOOCV, displaying the maximum AUC of SVM model (**Figures 4A,B**). The circRNAs in BC_{ExoC} were hsa_circ_0002190, hsa_circ_0007177, hsa_circ_0000642, hsa_circ_0001439, hsa_circ_0001417, hsa_circ_0005552, hsa_circ_0001073, hsa_circ_0000267, and hsa_circ_0006404 (**Supplemental Table 4**).

Then, the performance of BC_{ExoC} was investigated and verified in the validation cohort. In the validation cohort, the AUC of BC_{ExoC} was 0.80 (95% CI 0.71-0.89, **Figures 4C, D**).



B Training cohort

Model	Acc	Sen	Spe	P
SVM	83.0	84.2	81.5	
LR	80.8	89.1	70.4	0.2628
LDA	78.5	85.1	70.0	0.037

D Testing cohort

Model	Acc	Sen	Spe	P
SVM	79.2	74.4	85.3	
LR	80.5	76.7	85.3	0.317
LDA	79.2	74.4	85.3	1

F All cohorts

Model	Acc	Sen	Spe	P
SVM	81.9	81.2	82.6	
LR	80.7	85.4	74.8	0.355
LDA	78.4	81.9	73.9	0.038

FIGURE 4 | Performance of classifiers for breast cancer prediction. The performance of training cohort **(A, B)**, testing cohort **(C, D)** and all sample **(E, F)** with three models was showed. The AUC in the training cohort was cross-validated using leave one out cross-validation (LOOCV). Receiver operating characteristic, ROC; Acc, accuracy; Sen, sensitivity; SVM, support vector machine; LDA, linear discriminate analysis; LR, logistic regression; Spe, specificity; P, the *P*-value of DeLong's test.

By contrast with the training cohort, the AUC of the validation cohort was similar to those in the training cohort (*P*-value= 0.582; DeLong's test).

DISCUSSION

By analyzing genome-wide long RNA sequencing data of EVs, we identified a large number of novel circRNAs in human blood. By comparing the levels between breast cancer patients and those of controls, we found 439 circRNAs with significantly different levels. Based on their levels in EVs, we developed classifiers with three regression models in the training cohort. The optimal classifier (BC_{ExoC}) composed of nine circRNAs with the highest AUC was selected, and then it was verified in the validation cohort. The AUC of BC_{ExoC} was 0.82 [0.77-0.87] in all cohorts (**Figures 4E, F**). These findings highlight the potential of BC_{ExoC} as a non-invasive assessment for breast cancer in preclinical stages.

Compared with other studies, our method has several strengths: The circRNAs showed temporospatial characteristics (exhibit patient-specific and stage-specific manner), and the heat map results showed distinct patterns between breast cancer patients and controls, indicating their potentials as early diagnostic biomarkers of breast cancer. In this study, we used circRNAs in plasma EVs, so our method is non-invasive, which could reduce the harm of biopsy to patients and avoid the heterogeneity of cancer. In addition, the circular structure of circRNAs was normally more stable than the linear RNA, therefore, they may be more suitable taken as disease biomarkers. However, our study also has some limitations: as all samples were merely collected from one center, the performance of our classifier needs to be validated with more independent cohorts prior to their clinical applications.

By literature search, we found that three of the top 20 up-regulated circRNAs have been studied in breast cancer cells, and their functions were closely related to tumorigenesis (29, 30). For example, silencing of has_circ_0004771 inhibits proliferation and induces apoptosis in breast cancer through activation of miR-653 by targeting ZEB2 signaling pathway (29). Results of gene function enrichment analysis showed that some of these circRNAs were related to the tumorigenesis of breast cancer. Searching for the functions of these genes is expected to be biomarkers or therapeutic targets for breast cancer.

In summary, our data showed that BC_{ExoC} is a promising noninvasive method for the early diagnosis of breast cancer. Our techniques required for circRNA detection, such as plasma collection, RNA extraction, and qPCR, are routinely used in clinics. What's more, the cost of reagents and consumables is relatively low and the result of BC_{ExoC} is easy to be explained. Therefore, it is feasible to analyze BC_{ExoC} in clinical practice.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional review board of The First People's Hospital of Foshan. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

G-lY, Y-sW designed and supervised the study. LL and Z-wG analyzed and interpreted the data and prepared the manuscript.

X-mZ, LL, B-w H, Z-j L and M-qC designed the study, provided samples and interpreted clinical data. X-xY, ML, KL, C-lZ and T-cL analyzed and interpreted the data. All authors contributed to the article and approved the submitted version.

FUNDING

The work was supported by National Natural Science Foundation of China (81872416, 82173001, 81802435, 81900191), Medical Scientific Research Foundation of Guangdong Province of China (B2017006), and China Postdoctoral Science Foundation funded project (2019M662998) and Special fund of Foshan Summit plan (2020G010).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.752651/full#supplementary-material

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2019. *CA Cancer J Clin* (2019) 69:7–34. doi: 10.3322/caac.21551
3. Valadi H, Ekstrom K, Bossios A, Sjostrand M, Lee JJ, Lotvall JO. Exosome-Mediated Transfer of mRNAs and microRNAs Is a Novel Mechanism of Genetic Exchange Between Cells. *Nat Cell Biol* (2007) 9:654–9. doi: 10.1038/ncb1596
4. Ratajczak J, Wysoczynski M, Hayek F, Janowska-Wieczorek A, Ratajczak MZ. Membrane-Derived Microvesicles: Important and Underappreciated Mediators of Cell-to-Cell Communication. *Leukemia* (2006) 20:1487–95. doi: 10.1038/sj.leu.2404296
5. Sun Z, Shi K, Yang S, Liu J, Zhou Q, Wang G, et al. Effect of Exosomal miRNA on Cancer Biology and Clinical Applications. *Mol Cancer* (2018) 17:147. doi: 10.1186/s12943-018-0897-7
6. Shah R, Patel T, Freedman JE. Circulating Extracellular Vesicles in Human Disease. *N Engl J Med* (2018) 379:958–66. doi: 10.1056/NEJMra1704286
7. Soung YH, Ford S, Zhang V, Chung J. Exosomes in Cancer Diagnostics. *Cancers (Basel)* (2017) 9:8. doi: 10.3390/cancers9010008
8. Lener T, Gimona M, Aigner L, Borger V, Buzas E, Camussi G, et al. Applying Extracellular Vesicles Based Therapeutics in Clinical Trials - an ISEV Position Paper. *J Extracell Vesicles* (2015) 4:30087. doi: 10.3402/jev.v4.30087
9. Shen S, Song Y, Zhao B, Xu Y, Ren X, Zhou Y, et al. Cancer-Derived Exosomal miR-7641 Promotes Breast Cancer Progression and Metastasis. *Cell Commun Signal* (2021) 19:20. doi: 10.1186/s12964-020-00700-z
10. Thery C. Cancer: Diagnosis by Extracellular Vesicles. *Nat* (2015) 523:161–2. doi: 10.1038/nature14626
11. Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs are the Predominant Transcript Isoform From Hundreds of Human Genes in Diverse Cell Types. *PloS One* (2012) 7:e30733. doi: 10.1371/journal.pone.0030733
12. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a Large Class of Animal RNAs With Regulatory Potency. *Nat* (2013) 495:333–8. doi: 10.1038/nature11928
13. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, et al. Circular RNAs are Abundant, Conserved, and Associated With ALU Repeats. *RnaBioinf* (2013) 19:141–57. doi: 10.1261/rna.035667.112
14. Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, et al. Translation of CircRNAs. *Mol Cell* (2017) 66:9–21. doi: 10.1016/j.molcel.2017.02.021
15. Han D, Li J, Wang H, Su X, Hou J, Gu Y, et al. Circular RNA Circmto1 Acts as the Sponge of microRNA-9 to Suppress Hepatocellular Carcinoma Progression. *Hepatol* (2017) 66:1151–64. doi: 10.1002/hep.29270
16. Wang Y, Liu J, Ma J, Sun T, Zhou Q, Wang W, et al. Exosomal circRNAs: Biogenesis, Effect and Application in Human Diseases. *Mol Cancer* (2019) 18:116. doi: 10.1186/s12943-019-1041-z
17. Li S, Li Y, Chen B, Zhao J, Yu S, Tang Y, et al. ExoRBase: A Database of circRNA, lncRNA and mRNA in Human Blood Exosomes. *Nucleic Acids Res* (2018) 46:D106–12. doi: 10.1093/nar/gkx891
18. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinf* (2013) 29:15–21. doi: 10.1093/bioinformatics/bts635
19. Cheng J, Metge F, Dieterich C. Specific Identification and Quantification of Circular RNAs From Sequencing Data. *Bioinf* (2016) 32:1094–6. doi: 10.1093/bioinformatics/btv656
20. Zhang Z, Qi S, Tang N, Zhang X, Chen S, Zhu P, et al. Discovery of Replicating Circular RNAs by RNA-Seq and Computational Algorithms. *PloS Pathog* (2014) 10:e1004553. doi: 10.1371/journal.ppat.1004553
21. Guo JU, Agarwal V, Guo H, Bartel DP. Expanded Identification and Characterization of Mammalian Circular RNAs. *Genome Biol* (2014) 15:409. doi: 10.1186/s13059-014-0409-z
22. Glazar P, Papavasileiou P, Rajewsky N. CircBase: A Database for Circular RNAs. *RnaBioinf* (2014) 20:1666–70. doi: 10.1261/rna.043687.113
23. Robinson MD, McCarthy DJ, Smyth GK. EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinf* (2010) 26:139–40. doi: 10.1093/bioinformatics/btp616
24. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-Based Genome Alignment and Genotyping With HISAT2 and HISAT-Genotype. *Nat Biotechnol* (2019) 37:907–15. doi: 10.1038/s41587-019-0201-4
25. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10:1523. doi: 10.1038/s41467-019-09234-6
26. Lin XJ, Chong Y, Guo ZW, Xie C, Yang XJ, Zhang Q, et al. A Serum microRNA Classifier for Early Detection of Hepatocellular Carcinoma: A Multicentre, Retrospective, Longitudinal Biomarker Identification Study With a Nested Case-Control Study. *Lancet Oncol* (2015) 16:804–15. doi: 10.1016/S1470-2045(15)00048-0
27. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. PROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinf* (2011) 12:77. doi: 10.1186/1471-2105-12-77
28. Wagner EF, Nebreda AR. Signal Integration by JNK and P38 MAPK Pathways in Cancer Development. *Nat Rev Cancer* (2009) 9:537–49. doi: 10.1038/nrc2694
29. Xie R, Tang J, Zhu X, Jiang H. Silencing of Hsa_Circ_0004771 Inhibits Proliferation and Induces Apoptosis in Breast Cancer Through Activation of miR-653 by Targeting ZEB2 Signaling Pathway. *Biosci Rep* (2019) 39: BSR20181919. doi: 10.1042/BSR20181919
30. Lu WY. Roles of the Circular RNA Circ-Foxo3 in Breast Cancer Progression. *Cell Cycle* (2017) 16:589–90. doi: 10.1080/15384101.2017.1278935

Conflict of Interest: Author KL was employed by Guangzhou XGene Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer LZ declared a shared affiliation, with no collaboration, with the authors to the handling editor at the time of the review

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 12 November 2021
doi: 10.3389/fonc.2021.736769

Circulating Tumor DNA as a Predictive Marker of Recurrence for Patients With Stage II-III Breast Cancer Treated With Neoadjuvant Therapy

*Po-Han Lin[1,2], Ming-Yang Wang[3], Chiao Lo[3], Li-Wei Tsai[3], Tzu-Chun Yen[1], Thomas Yoyan Huang[1], Wei-Chih Huang[1], Karen Yang[4], Chih-Kai Chen[1], Sheng-Chih Fan[1], Sung-Hsin Kuo[5] and Chiun-Sheng Huang[3,6]**

[1] Department of Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan, [2] Institute of Medical Genomics and Proteomics, College of Medicine, National Taiwan University, Taipei, Taiwan, [3] Department of Surgery, National Taiwan University Hospital, Taipei, Taiwan, [4] Department of Molecular Biology, Princeton University, Princeton, NJ, United States, [5] Department of Medical Oncology, National Taiwan University Hospital, Taipei, Taiwan, [6] Department of Surgery, College of Medicine, National Taiwan University, Taipei, Taiwan

OPEN ACCESS

Edited by:
Raquel Nunes,
Johns Hopkins University,
United States

Reviewed by:
Islam M. Miligy,
University of Nottingham,
United Kingdom
Francesco Pepe,
University of Naples Federico II, Italy
Jenna Canzoniero,
Johns Hopkins University,
United States

***Correspondence:**
Chiun-Sheng Huang
huangcs@ntu.edu.tw

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: 05 July 2021
Accepted: 25 October 2021
Published: 12 November 2021

Citation:
Lin P-H, Wang M-Y, Lo C, Tsai L-W, Yen T-C, Huang TY, Huang W-C, Yang K, Chen C-K, Fan S-C, Kuo S-H and Huang C-S (2021) Circulating Tumor DNA as a Predictive Marker of Recurrence for Patients With Stage II-III Breast Cancer Treated With Neoadjuvant Therapy. Front. Oncol. 11:736769. doi: 10.3389/fonc.2021.736769

Background: Patients with stage II to III breast cancer have a high recurrence rate. The early detection of recurrent breast cancer remains a major unmet need. Circulating tumor DNA (ctDNA) has been proven to be a marker of disease progression in metastatic breast cancer. We aimed to evaluate the prognostic value of ctDNA in the setting of neoadjuvant therapy (NAT).

Methods: Plasma was sampled at the initial diagnosis (defined as before NAT) and after breast surgery and neoadjuvant therapy(defined as after NAT). We extracted ctDNA from the plasma and performed deep sequencing of a target gene panel. ctDNA positivity was marked by the detection of alterations, such as mutations and copy number variations.

Results: A total of 95 patients were enrolled in this study; 60 patients exhibited ctDNA positivity before NAT, and 31 patients exhibited ctDNA positivity after NAT. A pathologic complete response (pCR) was observed in 13 patients, including one ER(+)Her2(-) patient, six Her2(+) patients and six triple-negative breast cancer (TNBC) patients. Among the entire cohort, multivariate analysis showed that N3 classification and ctDNA positivity after NAT were independent risk factors that predicted recurrence (N3, hazard ratio (HR) 3.34, 95% confidence interval (CI) 1.26 – 8.87, p = 0.016; ctDNA, HR 4.29, 95% CI 2.06 – 8.92, p < 0.0001). The presence of ctDNA before NAT did not affect the rate of recurrence-free survival. For patients with Her2(+) or TNBC, patients who did not achieve pCR were associated with a trend of higher recurrence (p = 0.105). Advanced nodal status and ctDNA positivity after NAT were significant risk factors for recurrence (N2 – 3, HR 3.753,

95% CI 1.146 – 12.297, p = 0.029; ctDNA, HR 3.123, 95% CI 1.139 – 8.564, p = 0.027). Two patients who achieved pCR had ctDNA positivity after NAT; one TNBC patient had hepatic metastases six months after surgery, and one Her2(+) breast cancer patient had brain metastasis 13 months after surgery.

Conclusions: This study suggested that the presence of ctDNA after NAT is a robust marker for predicting relapse in stage II to III breast cancer patients.

Keywords: circulating tumor DNA, neoadjuvant therapy, breast cancer, recurrence, next-generation sequencing

INTRODUCTION

Although breast cancer prognosis has improved during the past two decades, breast cancer-related death remains a major cause of cancer-related mortality in women (1, 2). The main reason is that a significant proportion of breast cancer patients develop recurrence and distant metastases (3, 4). Once metastases occur, breast cancer is treatable but no longer curable (5).

For breast cancer patients, early detection of recurrence remains a major unmet need. In the neoadjuvant setting, pathological complete response (pCR) is a favorable prognostic marker in patients with Her2 (+) and triple-negative breast cancer (TNBC) (6). However, some patients with pCR may still experience recurrence or metastasis; on the other hand, the absence of pCR does not necessarily correlate with recurrence (6, 7). Recent studies have shown circulating tumor DNA (ctDNA), which are circulating DNA fragments that carry tumor-specific sequence alterations found in the cell-free fraction of blood, to be a promising and sensitive tool for targeted monitoring (8–12). The detection of resistance mutations using ctDNA can also occur significantly earlier than radiographic progression (13). In previous reports of metastatic cancer patients, serial quantification of ctDNA allowed for noninvasive assessment of therapeutic response and understanding of resistance mechanisms (8, 11, 14, 15). For patients with early-stage breast, lung and colon cancer, studies reported that ctDNA in the plasma can be used to detect minimal residual disease (16–18). Serial detection of ctDNA after surgery and adjuvant chemotherapy of breast cancer could identify recurrent disease earlier than clinical overt tumor presenting in the radiologic images (19, 20). However, for breast cancer patients receiving neoadjuvant therapy (NAT), the prognostic value of ctDNA before and after NAT is uncertain. It is unknown whether ctDNA or pCR has a more prognostic value for breast cancer patients, either. To determine the prognostic value of ctDNA in the context of NAT, we collected the patients' plasma before and after NAT and used next-generation sequencing (NGS)-based deep sequencing to detect ctDNA and evaluated the impact of ctDNA on disease recurrence.

Abbreviations: ctDNA, Circulating tumor DNA; CHIP, Clonal hematopoiesis of indeterminate potential; ER, Estrogen receptor; Her2, Human epidermal growth factor receptor; NAT, Neoadjuvant therapy; pCR, Pathologic complete response; RFS, Recurrence-free survival; TNBC, Triple-negative breast cancer.

METHODS

Patients and Sample Collection

Stage II or III breast cancer patients who received NAT were enrolled in this study. The clinical and pathologic characteristics were reviewed retrospectively from medical records. The presence of estrogen receptors (ER), progesterone receptors (PR), and Her2 were determined by immunohistochemical staining. The ER or PR status was considered negative when less than 1% of the tumor cells showed positive staining. For Her2 staining, a score of 0 or 1+ was considered negative; specimens with a score of 2+ were further tested with fluorescence *in situ* hybridization analysis. The tumor histological grade was defined using the Nottingham combined histological grading system. This study was approved by the institutional review board (IRB number: 201704009RINC).

At the initial diagnosis (defined as before NAT), a 10-mL sample of blood was collected and stored in an EDTA-containing tube. Then, all patients were treated with NAT and received breast surgery. After NAT and breast surgery (defined as after NAT), another 10 mL of blood was sampled. Within three hours of blood sampling, the plasma was extracted after centrifugation at 1000× G for 10 minutes then stored at -80°C (21). Cell-free DNA was extracted using a QIAamp Circulating Nucleic Acid Kit (Qiagen, Germantown, MD, USA) according to the manufacturer's protocol.

Library Preparation and Next-Generation Sequencing

The library was constructed using a QIAseq Targeted DNA Panel with a customized gene list. The customized panel was designed to amplify the coding regions of the following genes: *TP53*, *PIK3CA*, *Her2*, *GATA3*, *CDH1*, *PTEN*, *AKT1*, *ESR1*, *S100A7-9*, *ZNF703*, *B2M*, *CCND1*, *GATA3* and *c-MYC*. According to the manufacturer's protocol, 10 ng of DNA was digested briefly into small fragments by a fragmentation enzyme at 32°C and 72°C. The DNA fragments were added to the QIAseq IL-N7 adapters, followed by target enrichment polymerase chain reaction (PCR) using the QIAGEN IL-Forward primer and the targeted DNA Panel primers. Finally, the library was amplified with universal PCR. The DNA library was then checked by using an Agilent Chip High Sensitivity DNA kit. KAPA library quantification kits were used to quantify the final concentration. The final DNA library was sequenced with the following

Illumina platforms: Illumina MiSeq Reagent Kit v2, 2 x 150 bp reads or Illumina NextSeq 550 system Mid-Output Kit, 2 x 150 bp reads.

Post-Sequencing Analysis

Previously, we have constructed an analytic pipeline of post-NGS bioinformatics (22). First, BWA software (version 0.5.9) was used to align the raw sequencing data to the reference human genome [Feb. 2009, GRCh37/hg19; SAMtools (version 0.1.18)]. Picard (version 1.54) was used to perform the necessary data conversion, sorting, and indexing. GATK was used for variant calling with the Mutect2 and VariantFiltration parameters. Finally, ANNOVAR was used to annotate the genetic variants. Pathogenic and likely pathogenic variants were defined according to the American College of Medical Genomics and Genetics (ACMG) guidelines (23). The presence of ctDNA was determined by the presence of pathogenic and likely pathogenic variants, which are also considered tumor mutations. For variants of uncertain significance, if the prevalence of the variants in the normal population was less than 0.01 in a genomic database (1000 Genomics, ESP6500 and ExAC) and predicted to be deleterious by computer software (SIFT, PolyPhen2, and CADD), then they were classified as "highly suspected deleterious". The above filtering analyses removes germline variants as much as possible (24); these variants are highly suspected to originate from tumors, so the detection of these variants could be considered indicative of ctDNA.

Analysis of Copy Number Changes

Since the *Her2*, *c-Myc*, *CCND1* and *S100A* genes can be amplified in some breast cancer tumors, we decided to use copy number variations (CNV) to indicate the presence of ctDNA (25–27). Copy number variations were analyzed by OncoCNV (https://github.com/BoevaLab/ONCOCNV) according to the authors' instructions. The baseline control consisted of the ctDNA BAM files of 14 healthy people. The ctDNA BAM files from the breast cancer patients were compared to the BAM files from the control population by using OncoCNV's default *cghseg* segmentation algorithm (28). The sequencing region of each targeted gene was divided into several segments. When the mean of all segments of each gene was significantly different from the baseline, such as when the copy number predicted was greater than three copies or fewer than one copy from the baseline, we considered that to indicate a CNV alteration, which indicated the presence of ctDNA.

Statistics

The chi-squared test and Fisher's exact test were used to calculate the significance of the variance between each group. Survival was estimated by Kaplan-Meier analysis. Cox proportional hazards regression analysis was used to estimate the hazards ratios of RFS with a corresponding 95% confidence interval (CI) for various factors. All *p* values are two-sided, and *p*-values less than 0.05 were considered statistically significant.

RESULTS

Evaluation of Assay Performance

First, to confirm the accuracy of the NSG-based deep sequencing, we checked whether this method could distinguish the existence of low-abundance mutants from background errors arising from the polymerase chain reaction (PCR) or sequencing process. We constructed a *TP53* mutant (NM_000546.6: c.844C>A) as a reference sample; then we utilized this *TP53* mutant with serial concentrations of 100%, 10%, 1%, and 0.1% to test whether the experimental method could detect these mutants at these concentrations (**Supplementary Methods**). The results demonstrated that the signal from the 0.1% mutant was significantly higher than background errors (**Supplementary Figure S1A**), suggesting that NGS testing accurately detected mutants present at 0.1%. In addition, the mutation level could be measured with a linear fashion (R^2 = 0.9997, **Supplementary Figure S1B**).

Second, in deep cell-free analyses, another source of variants that makes it hard to distinguish cancer mutations is clonal hematopoiesis of indeterminate potential (CHIP) (29–31). The CHIP mutations mostly occur in the *DNMT3A*, *TET2*, *PPM1D*, *ASXL1* and *TP53* genes (29), whereas pathogenic variants of breast cancer were most prevalent in *TP53*, *PIK3CA*, *MAP3KA1*, *CDH1*, and *PTEN* (32). Variants most likely to be indistinguishable from CHIP were located in *TP53*. Twenty-two tumors from the pre-neoadjuvant core biopsy tumors were available for DNA extraction and sequencing (**Supplementary Table S1**). Among them, 6 patients had *TP53* variants, and their *TP53* variants co-existed in the ctDNA and DNA from tumor biopsies (**Supplementary Table S1** and **Supplementary Figure S2**), suggesting the *TP53* variants origin from breast cancer, not CHIP mutations.

Patients

A total of 95 patients were enrolled in this study. The median age was 50.0 years old. Forty-one patients had ER(+) Her2(-) breast cancer, 29 patients had Her2(+) breast cancer, and 25 patients had triple-negative breast cancer (TNBC). Before NAT, tumors with T1, T2 and T3-4 size classifications were found in three, 54 and 38 patients of each population, respectively. Eighty-two patients had positive axillary lymph nodes. According to standard clinical practice, ER(+) Her2(-) breast cancer patients with large tumors were treated with NAT. Out of the 95 patients, 77 patients received anthracycline while 80 patients received taxane in their NAT regimens. All Her2(+) patients received adjuvant anti-Her2 target therapy (27 patients receiving trastuzumab, one another receiving trastuzumab/pertuzumab and the other receiving trastuzumab-DM1). After NAT, 13 patients achieved a pCR of their primary breast tumors; 82 patients did not have pCR. Among the 13 pCR patients, there was one ER(+) Her2(-), six Her2(+) and six TNBC patients. The frequency of pCR was significantly higher in patients with Her2(+) breast cancer or TNBC than ER(+)Her2(-) patients (p = 0.002). CtDNA was detected in 60 patients before NAT and 31 patients after NAT. All of the clinical and pathologic characteristics are shown in **Table 1**.

TABLE 1 | Clinical and pathologic characteristics of enrolled patients stratified by immunophenotypes.

	All	ER(+) Her2 (-)	ER(±) Her2 (+)	TNBC
Number	95	41	29	25
Age (mean ± SD)	50.0 ± 8.8	49.2 ± 7.8	49.3 ± 8.7	52.0 ± 10.2
T classification (before NAT)				
T1	3	0	1	2
T2	54	19	16	19
T3-4	38	22	12	4
N classification (before NAT)				
N-negative	13	4	6	3
N-positive	82	37	23	22
T classification (after NAT)				
no tumor	13	1	6	6
T1	32	10	13	9
T2	29	16	6	7
T3-4	21	14	4	3
N classification (after NAT)				
N0	34	7	16	11
N1	29	9	10	10
N2	22	17	2	3
N3	10	8	1	1
Response				
pCR	13	1	6	6
absence of pCR	82	40	23	19
NAT regimen				
Anthracycline	77	33	24	20
Taxane	80	29	29	22
Trastuzumab/pertuzumab	29	0	29	0
Presence of ctDNA				
before NAT	60	33	15	12
after NAT	31	11	10	10
Adjuvant chemotherapy	30	18	3	9
anthracycline	15	8	3	4
taxane	15	10	0	5
Adjuvant anti-Her2 target therapy*	29	0	29	0

anti-Her2 target therapy: 27 patients receiving trastuzumab, one another receiving trastuzumab/pertuzumab and the other receiving trastuzumab-DM1.*

Genetic Alterations in Tumor ctDNA

Among the 95 patients, 19 patients were found to have ctDNA before and after NAT; 41 patients had ctDNA only before NAT, 12 patients had ctDNA only after NAT, and 23 patients had ctDNA neither before nor after NAT (**Supplementary Table S2**). The most common genetic variants were in the *TP53* (n = 28), followed by *PIK3CA* (n = 16), *CDH1* (n = 15), and *Her2* (n = 7) genes. Eighteen patients had altered CNVs in their ctDNA, including of *AKT1*, *CCND1*, *CDH1*, *c-MYC*, *Her2*, *PIK3CA*, *S100A*, and *ZNF703*, either before or after NAT (**Supplementary Table S2** and **Figure 1**). Before NAT, Patient #73 (**Figure 1A**) and Patient #24 (**Figure 1B**) exhibited copy number gains of the *S100A* and *Her2* genes in ctDNA, respectively; after NAT, the copy numbers of these genes in ctDNA returned to normal levels. Patient #3 (**Figure 1C**) had a new copy loss of the *PTEN* gene after NAT. We observed gains of *Her2* and *c-MYC* in patient #27 (**Figure 1D**) before NAT that were only partially resolved after NAT.

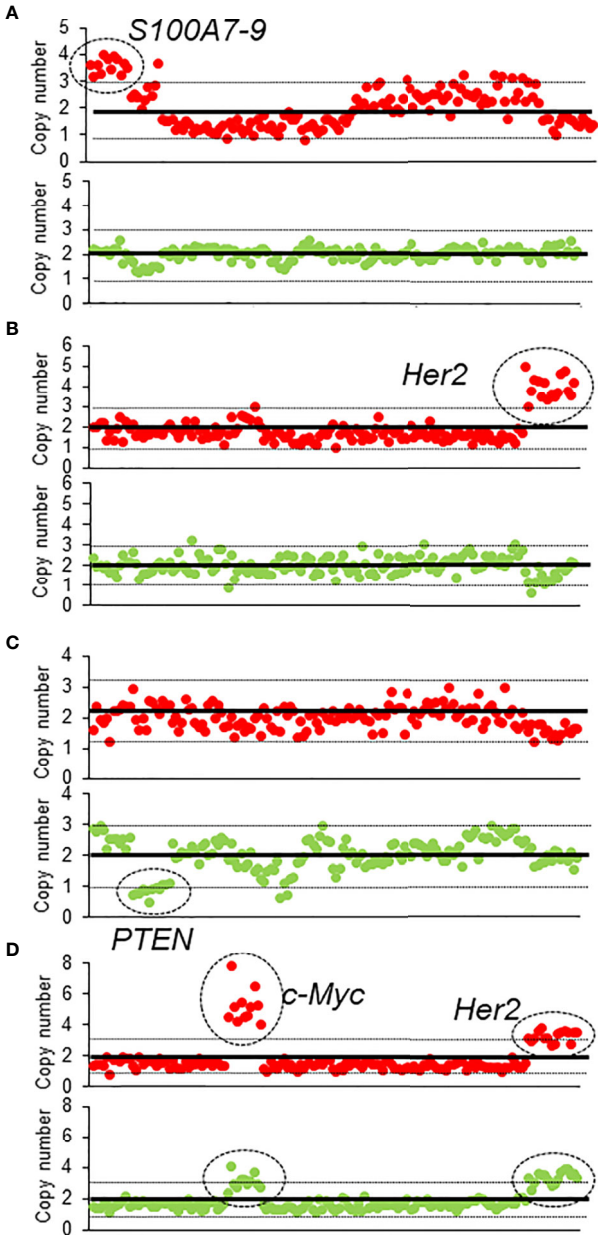


FIGURE 1 | The CNV of four patients before and after NAT **(A–D)**. The red dots represent the CNV before NAT, and green dots represent the CNV after NAT.

Association Between ctDNA and Clinical Characteristics

Patients who had ctDNA before NAT tended to have a larger tumor size than those who did not have ctDNA before NAT (mean 5.0 cm *vs.* 4.3 cm, p = 0.104). However, the presence of ctDNA after NAT did not correlate with the tumor size or LN numbers after NAT. Although the difference was not statistically significant, patients with pCR had a lower detection of ctDNA after NAT than patients with no pCR (patients with pCR *vs.* absence of pCR:

15.4% *vs.* 35.4%, p = 0.132). Additionally, the presence of ctDNA was not correlated with the immunophenotype of breast cancer.

Impact of Clinical Factors and ctDNA on RFS

The median follow-up time of the entire cohort was 5.1 years, and the 5-year recurrence-free survival (RFS) was 58% (95% CI 48.0 – 68.0%). For clinical factors, Kaplan-Meier analysis showed that the residual tumor size after NAT and N classification after NAT were prognostic factors for RFS; patients who achieved pCR tended to have a better RFS than patients who did not achieve pCR (**Figures 2A–C** and **Table 2**). On the other hand, patients with ctDNA after NAT had a significantly inferior RFS (p < 0.001, **Figure 2D**). Other factors, such as age, ctDNA detection before NAT, immunophenotype, initial tumor size before NAT and N classification before NAT and adjuvant chemotherapy did not influence RFS. RFS was similar between patients with and without *TP53*, *PIK3CA* and *CDH1* mutations (**Table 2** and **Supplementary Table S3**).

We then analyzed the clinical and pathologic characteristics of patients with and without ctDNA after NAT, and no difference was found between the two patient groups (**Supplementary Table S4**). After incorporating the residual tumor size, N classification after NAT, pCR and ctDNA after NAT, multivariate analysis showed that an N3 classification and ctDNA positivity after NAT were independent risk factors that predicted tumor recurrence (N3, hazard ratio (HR) 3.352, 95% CI 1.267 – 8.870, p = 0.015; ctDNA, HR 4.135, 95% CI 2.014 – 8.491, p < 0.0001). Other factors did not significantly impact RFS (**Table 2**).

Next, we analyzed the 72 patients with detected ctDNA, either before or after NAT. Patients with ctDNA positivity after NAT had a significantly inferior RFS compared to those without detectable ctDNA (**Supplementary Figure S3**, p<0.001). After adjusting for tumor size (after NAT), N classification (after NAT) and pCR, multivariate analysis with the Cox model revealed that ctDNA positivity after NAT was the most significant risk factor that predicted tumor recurrence (HR

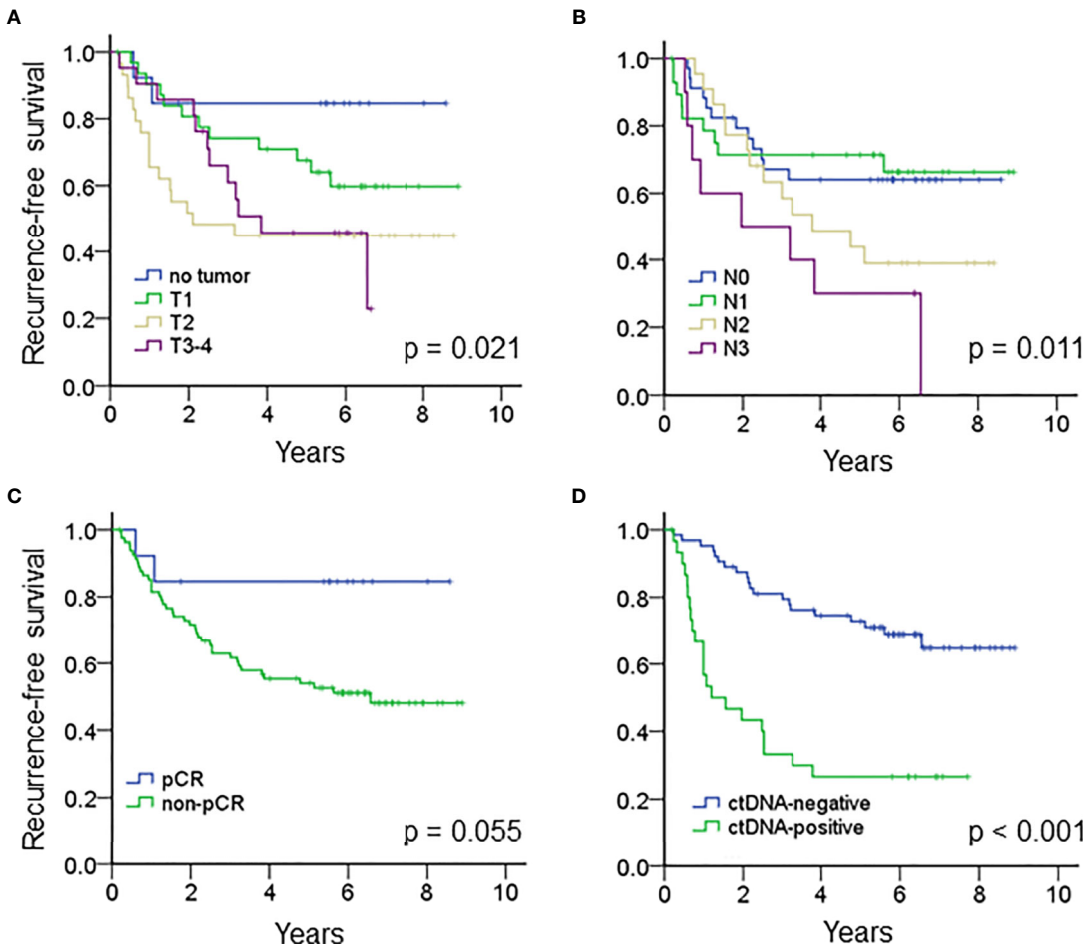


FIGURE 2 | Kaplan-Meier analysis estimated the recurrence-free survival of the entire cohort according to **(A)** the tumor size after NAT (p = 0.021), **(B)** N classification after NAT (p = 0.011), **(C)** pCR (p = 0.055) and **(D)** ctDNA after NAT (p < 0.001).

TABLE 2 | Univariate and multivariate analysis of recurrence-free survival of the entire cohort.

variables	univariate				multivariate			
	HR	lower	upper	P value	HR	lower	upper	P value
Age (>50 *vs.* <50)	0.962	00.525	1.763	.899				
T classification (before NAT)								
T1-2	1							
T3-4	1.026	.553	1.903	0.936				
N classification (before NAT)								
N-negative	1							
N-positive	2.266	0.700	7.336	0.172				
T classification (after NAT)								
no tumor	1				1			
T1	2.536	0.568	11.333	0.223	1.963	0.333	11.575	0.456
T2	4.842	1.112	21.083	0.036	2.435	0.450	13.186	0.302
T3-4	4.158	0.929	18.604	0.062	2.338	0.488	11.202	0.288
N classification (after NAT)								
N0	1				1			
N1	0.953	0.401	2.263	0.914	1.378	.526	3.606	0.514
N2	1.750	0.798	3.838	0.163	1.418	.611	3.293	0.416
N3	3.055	1.246	7.487	0.015	3.352	1.267	8.870	0.015
Response								
pCR	1				1			
absence of pCR	3.656	0.883	15.134	0.074	2.230	0.468	10.623	0.314
Immunophenotype								
ER/PR(+)Her2(-)	1							
ER/PR(+)Her2(+)	0.611	0.284	1.314	0.207				
TNBC	1.294	0.639	2.622	0.474				
ctDNA								
before NAT*	0.700	0.378	1.298	0.257				
after NAT*	3.894	2.113	7.177	<0.001	4.135	2.014	8.491	<0.001
Adjuvant chemotherapy								
No	1							
Yes	1.141	0.601	2.169	0.686				
Genes								
TP53#	1.156	0.609	2.197	0.657				
CDH1#	0.669	0.263	1.704	0.399				
PIK3CA#	1.313	0.607	2.837	0.489				

**The presence of ctDNA vs. nonpresence of ctDNA; #gene mutation vs. nonmutation.*

8.02, 95% CI 3.24 – 19.86, p < 0.0001) (**Supplementary Table S5**).

The Impact of ctDNA on Disease Recurrence in Different Immunophenotypes of Breast Cancer

The median RFS of all the patients with ctDNA positivity after NAT was 1.19 years. When stratified by the immunophenotypes, ctDNA positivity after NAT was associated with a significantly inferior RFS for ER(+) breast cancer or TNBC patients and a trend of higher recurrence rates for patients with the Her2 subtype (**Figures 3A–C**). The median RFS of ER(+) breast cancer, Her2 (+) breast cancer and TNBC patients with ctDNA positivity after NAT were 0.90, 2.52 and 0.74 years, respectively.

The Impact of ctDNA on Disease Recurrence in Patients With and Without a pCR

For the entire cohort, the presence of ctDNA after NAT was a significant risk factor associated with recurrence in both patients who achieved and did not achieve pCR (**Figures 3D, E**, all p < 0.001). Because pCR was previously reported as a surrogate marker for survival in patients with Her2(+) and TNBC (6), we analyzed these patient subgroups. Between the two patient populations, pCR was related to a trend of improved survival compared to absence of pCR (HR 3.328, 95% CI 0.777 – 14.243, p = 0.105, **Supplementary Table S6**). Multivariate analysis showed that advanced nodal status and ctDNA after NAT were independently correlated with high risk (N2-3, HR 3.753, 95% CI 1.146–12.297, p = 0.029; ctDNA, HR 3.123, 95% CI. 1.139 – 8.564, p = 0.027), and pCR status did show a not significant correlation with recurrence (**Table 3**). A potential reason for this phenomenon is that pCR only represents the therapeutic efficacy of local breast tumor and the ctDNA may indicate that an occult lesion is present that is not effectively treated with NAT. In our study, 13 patients achieved pCR after NAT, and among those patients, two exhibited ctDNA positivity after NAT. One TNBC patient (case #50) received neoadjuvant docetaxel/epirubicin (four cycles) and achieved pCR for her primary breast and axillary tumors. However, she had hepatic metastases at 6 months after mastectomy (**Supplementary Figure S4**). The other patient (case #5) had Her2-positive

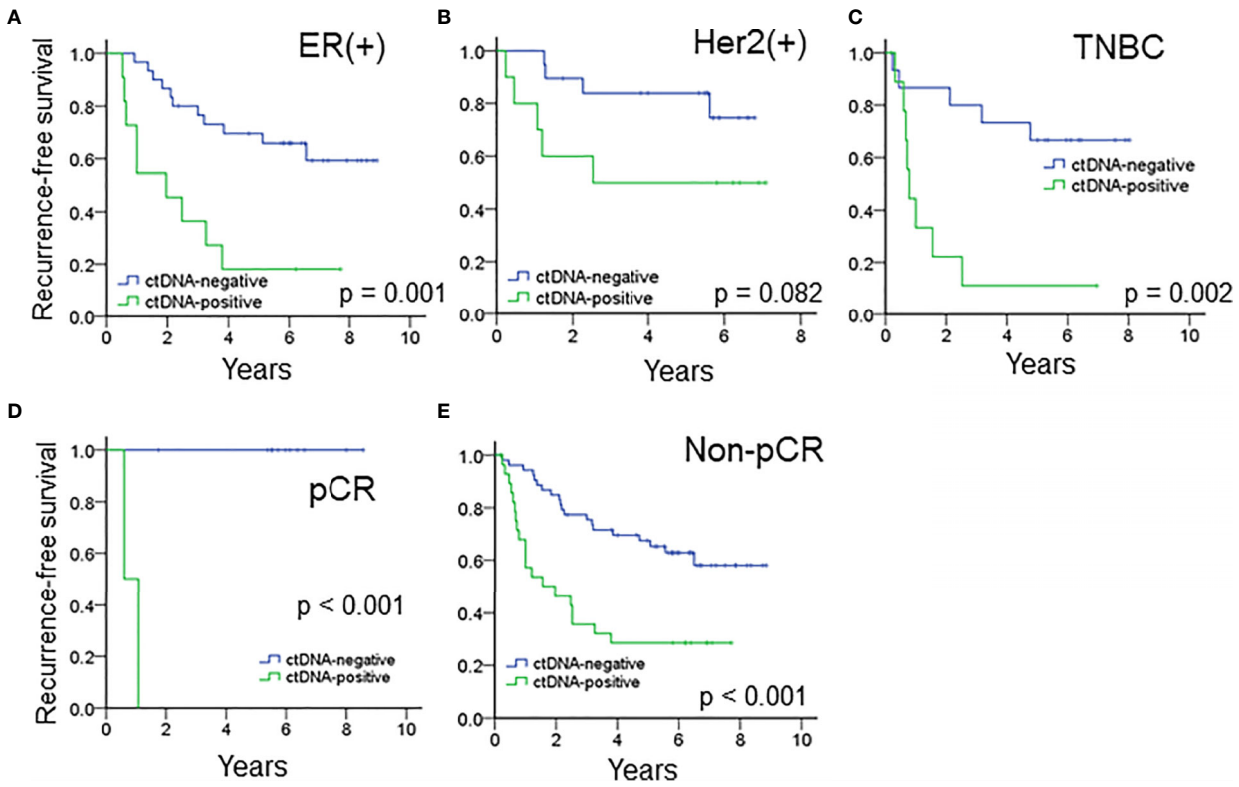


FIGURE 3 | The prognostic impact of ctDNA after NAT in patients with **(A)** ER(+) breast cancer, **(B)** Her2(+) breast cancer and **(C)** TNBC. ctDNA after NAT predicted RFS in **(D)** pCR and **(E)** patients who did not achieve pCR.

breast cancer and received neoadjuvant docetaxel/trastuzumab (four cycles) and epirubicin/cyclophosphamide (four cycles). The pathology showed no residual tumors. Trastuzumab was continuously maintained for one year. At the end of trastuzumab treatment (13 months after mastectomy), a cerebellar metastasis was found. The other 11 patients who achieved a pCR did not have ctDNA after NAT nor did they experience recurrence or metastasis.

TABLE 3 | Multivariate analysis of recurrence-free survival in patients with Her2(+) breast cancer and TNBC.

Variables	HR	lower	upper	P value
T classification (after NAT)				
no tumor	1			
T1	0.909	0.167	4.952	0.912
T2	2.461	0.435	13.917	0.308
T3-4	4.082	0.756	22.038	0.102
N classification (after NAT)				
N0	1			
N1	1.845	.633	5.378	0.262
N2-3	3.753	1.146	12.297	0.029
Response				
pCR	1			
absence of pCR	4.082	0.756	22.038	0.102
Adjuvant chemotherapy				
No	1			
Yes	1.137	0.419	3.084	0.801
ctDNA after NAT				
undetected	1			
detected	3.123	1.139	8.564	0.027

DISCUSSION

Our data suggested that the presence of ctDNA after NAT is a prognostic factor that predicts breast cancer recurrence after mastectomy. Traditionally, the therapeutic response to NAT was considered a marker for predicting prognosis (6). In our study, multivariate analysis showed a greater predictive value for ctDNA than the response of the primary breast tumor to NAT treatment. Therefore, ctDNA seems more representative of the therapeutic efficacy of primary and potential micrometastatic tumors treated with NAT.

During the median 5.1-year follow-up, the overall positive predictive value of ctDNA positivity after NAT for disease relapse was 70.9%, which was higher than the predictive value of 48.8% for relapse in patients who did not achieve pCR. After stratifying patients into pCR and absence of pCR, ctDNA positivity after NAT remained a significant risk factor for RFS among the two patient groups (**Figures 3D, E**). Although patients who did not achieve pCR usually had a significantly inferior RFS than pCR patients, ctDNA negativity after NAT in patients who did not achieve pCR was associated with a better RFS (**Figure 3E**), compatible with previous findings that ctDNA clearance associated with the improved survival in patients who did not achieve pCR (33). In contrast, pCR after NAT was a surrogate marker for predicting disease-free Her2(+) and TNBC patients. However, in our cohort, two patients (one Her2(+) and one TNBC) who achieved a pCR and exhibited ctDNA positivity after NAT developed distal metastasis at six months and one year, respectively. A possible reason is that the pCR was assessed using only primary breast tumor detection without evaluating systemic micrometastatic tumor cells. The patient who had Her2-positive breast cancer and achieved a pCR after NAT developed brain metastasis after trastuzumab maintenance therapy. This was compatible with previous report that trastuzumab was difficult to penetrate the blood-brain barrier to treat brain micrometastatic tumor cells (34). However, ctDNA positivity suggested that ctDNA could cross the blood–brain barrier to be detected in the plasma (35). Thus, ctDNA is more suitable than pCR for representing the overall disease state and could be a robust marker for predicting the survival rate.

Although patients with ctDNA positivity after NAT had inferior RFS, the length of RFS varied among patients with different immunophenotypes. Among patients with ctDNA positivity after NAT, patients with Her2- positive breast cancer had a significantly longer RFS than patients with TNBC and luminal breast cancers. The maintenance of anti-Her2 antibody therapy and the potential long-term preservation of antibody-dependent cellular cytotoxicity may explain the risk attenuation and delayed relapse of Her2-positive breast cancer patients (36). In this study, some patients received adjuvant chemotherapy according to physician decision. However, adjuvant chemotherapy did not influence the RFS in the overall cohort (**Table 2**) or in each subtype of breast cancer (**Supplementary Table S3**). For patients with detected ctDNA after NAT, all twelve Her2-positive breast cancer patients received postmastectomy adjuvant anti-Her2 therapy; one received trastuzumab emtansine, another received trastuzumab plus pertuzumab, and the remaining patients received trastuzumab for one year. For the eight TNBC patients, only one received adjuvant chemotherapy. Out of the eleven patients with ER(+) breast cancer, six received adjuvant chemotherapy, and all of them received hormone therapy. Notably, the median RFS of TNBC and ER(+) breast cancer patients was less than one year. This result might suggest that current standard chemotherapy and hormone therapy treatments were not effective for these patients. CtDNA has the potential to identify actionable genetic variants that provide sensitivity or resistance mechanisms for chemotherapy and/or targeted therapy (37); this information can be used to guide personalized therapy in the future (38). Alternative adjuvant therapy options can be explored for these patients.

The concordance between pCR and the clearance of ctDNA was moderate. The ctDNA concentration usually decreases after NAT (17, 39). In a previous report, the decrease in ctDNA levels in patients who achieved a pCR was greater than that in those who did not achieve a pCR (39). Similarly, our data revealed that a lower proportion of patients who achieved a pCR exhibited ctDNA positivity after NAT than that in patients who did not achieve pCR (pCR *vs.* absence of pCR: 15.4% *vs.* 35.4%, p = 0.132). Among the 72 patients with ctDNA positivity (before and after NAT), 81.0% of responders had a decrease in ctDNA (defined as a tumor size reduction of more than 30% of the original size) (40), whereas 58.9% of nonresponders had a decrease in ctDNA concentrations (Pearson's chi-squared, p = 0.088, **Figures 4A, B**).

One limitation to this study is the possibility that some ctDNA mutations may have originated from CHIP mutations (41). Although we observed a good concordance of genetic variants between ctDNA and available pre-neoadjuvant biopsy tumors, the possibility that some ctDNA mutations originated from CHIP mutations could not be ruled out because we did not have all of the biopsy tumors for sequencing. To reduce the possibility of detecting CHIP mutations, first, we designed a sequencing panel by selecting genes that are often mutated in breast cancer, not in hematologic cells (32). This strategy decreases the possibility of mixing the CHIP mutations into breast cancer mutations. Second, we only considered pathogenic/likely pathogenic or highly-suspicious deleterious variants as proof of ctDNA positivity. These variants may have biological implications for breast cancer. For example, PIK3CA H1047R is a driver mutation in breast cancer (42), suggesting that it could be a ctDNA specific to breast cancer. Third, we not only analyzed the genetic variants but also the CNV. The amplification of *Her2*, *S100A* and *CCND1* have biological significance in breast cancer pathology (25, 43), and amplification of *c-MYC* is related to high-grade malignancy (44). These CNVs are considered to be derived from breast cancer. Thus, we can reduce the possibility to contaminate CHIP mutations in the ctDNA.

The second limitation was that we only examined ctDNA before and after NAT and did not perform longitudinal monitoring; as a result, we were not able to detect late recurrence. In our cohort, 42 patients had disease recurrence. Out of those 42 patients, 22 exhibited ctDNA positivity after NAT. The 22 patients with ctDNA positivity had a significantly

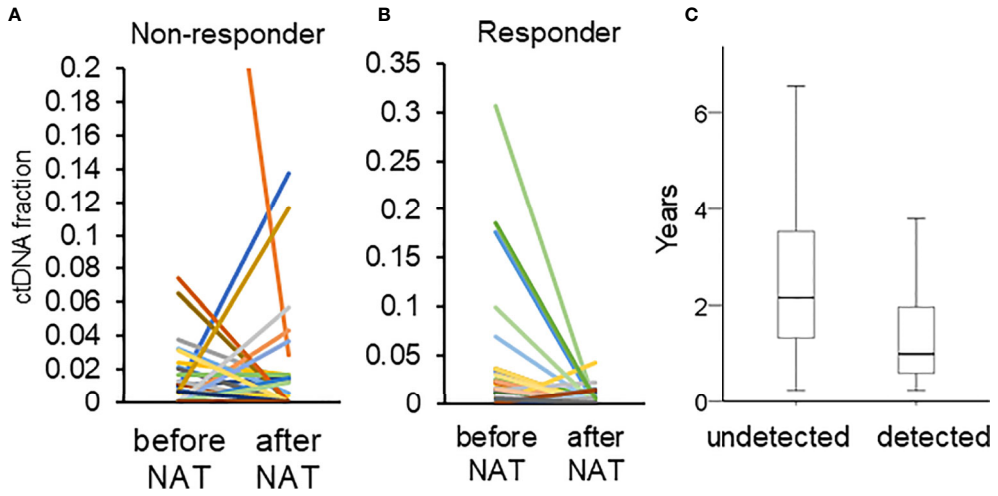


FIGURE 4 | **(A, B)** Changes in the fraction of ctDNA in patients who did and did not respond. The different color represented different mutations. **(C)** The duration of RFS in patients with (detected) and without (undetected) ctDNA after NAT.

shorter time to recurrence than those with ctDNA negativity (with ctDNA *vs.* without ctDNA: 1.31 *vs.* 2.64 years, p = 0.004, **Figure 4C**). A single time point sample of ctDNA after NAT was a significant predictor of only early recurrence. Longitudinally tracking ctDNA may improve the predictive value for both early and late recurrence (19, 20, 39).

CONCLUSIONS

We showed that ctDNA detection after NAT has great clinical utility potential as a prognostic marker in patients with breast cancer. CtDNA detection can identify and define a subset of high-risk patients. The next step is to determine the type of adjuvant therapy strategies that can effectively reduce recurrence. Since actionable genetic variants can be detected by ctDNA, further prospective trials should focus on incorporating ctDNA detection and exploring how to guide patient treatment, which could maximize the utility of ctDNA detection.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

Ethical approval was obtained from the ethical committees of National Taiwan University Hospital (IRB number: 201704009RINC). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

C-SH had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: P-HL and C-SH. Patient collection: P-HL, M-YW, LW-T, CL, S-HK, and C-SH. Performing experiments and bioinformatics: T-CY, TH, C-KC, KY, W-CH, and S-CF. Acquisition, analysis, or interpretation of data: all authors. Drafting of the manuscript: P-HL and C-SH. Critical revision of the manuscript for important intellectual content: all authors.

FUNDING

This work was supported, in part, by research grants from the National Taiwan University Hospital (NTUH. 107-004068 and 108-004128, to P-HL) and the Ministry of Science and Technology (MOST 104-2314-B-002-106-MY3 to C-SH, and MOST 109WFA0111726 to P-HL).

ACKNOWLEDGMENTS

We also thank the National Applied Research Laboratories for providing access to their high-performance computer to analyze the post-NGS data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.736769/full#supplementary-material

REFERENCES

1. DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A. Breast Cancer Statistics, 2015: Convergence of Incidence Rates Between Black and White Women. *CA Cancer J Clin* (2016) 66(1):31–42. doi: 10.3322/caac.21320
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2016. *CA Cancer J Clin* (2016) 66(1):7–30. doi: 10.3322/caac.21332
3. Redig AJ, McAllister SS. Breast Cancer as a Systemic Disease: A View of Metastasis. *J Internal Med* (2013) 274(2):113–26. doi: 10.1111/joim.12084
4. Hudis CA, Gianni L. Triple-Negative Breast Cancer: An Unmet Medical Need. *Oncologist* (2011) 16 Suppl 1:1–11. doi: 10.1634/theoncologist.2011-S1-01
5. Anders CK, Zagar TM, Carey LA. The Management of Early-Stage and Metastatic Triple-Negative Breast Cancer: A Review. *Hematol Oncol Clinics North Am* (2013) 27(4):737–49, viii. doi: 10.1016/j.hoc.2013.05.003
6. von Minckwitz G, Untch M, Blohmer JU, Costa SD, Eidtmann H, Fasching PA, et al. Definition and Impact of Pathologic Complete Response on Prognosis After Neoadjuvant Chemotherapy in Various Intrinsic Breast Cancer Subtypes. *J Clin Oncol* (2012) 30(15):1796–804. doi: 10.1200/JCO.2011.38.8595
7. van Hagen P, Wijnhoven BP, Nafteux P, Moons J, Haustermans K, De Hertogh G, et al. Recurrence Pattern in Patients With a Pathologically Complete Response After Neoadjuvant Chemoradiotherapy and Surgery for Oesophageal Cancer. *Br J Surg* (2013) 100(2):267–73. doi: 10.1002/bjs.8968
8. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, et al. Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. *N Engl J Med* (2013) 368(13):1199–209. doi: 10.1056/NEJMoa1213261
9. Figg WD2nd, Reid J. Monitor Tumor Burden With Circulating Tumor DNA. *Cancer Biol Ther* (2013) 14(8):697–8. doi: 10.4161/cbt.25361
10. Rothe F, Laes JF, Lambrechts D, Smeets D, Vincent D, Maetens M, et al. Plasma Circulating Tumor DNA as an Alternative to Metastatic Biopsies for Mutational Analysis in Breast Cancer. *Ann Oncol* (2014) 25(10):1959–65. doi: 10.1093/annonc/mdu288
11. Lau E, McCoy P, Reeves F, Chow K, Clarkson M, Kwan EM, et al. Detection of ctDNA in Plasma of Patients With Clinically Localised Prostate Cancer Is Associated With Rapid Disease Progression. *Genome Med* (2020) 12(1):72. doi: 10.1186/s13073-020-00770-1
12. Lee B, Lipton L, Cohen J, Tie J, Javed AA, Li L, et al. Circulating Tumor DNA as a Potential Marker of Adjuvant Chemotherapy Benefit Following Surgery for Localized Pancreatic Cancer. *Ann Oncol* (2019) 30(9):1472–8. doi: 10.1093/annonc/mdz200
13. Oxnard GR, Paweletz CP, Kuang Y, Mach SL, O'Connell A, Messineo MM, et al. Noninvasive Detection of Response and Resistance in EGFR-Mutant Lung Cancer Using Quantitative Next-Generation Genotyping of Cell-Free Plasma DNA. *Clin Cancer Res* (2014) 20(6):1698–705. doi: 10.1158/1078-0432.CCR-13-2482
14. Bachet JB, Blons H, Hammel P, Hariry IE, Portales F, Mineur L, et al. Circulating Tumor DNA Is Prognostic and Potentially Predictive of Eryaspase Efficacy in Second-Line in Patients With Advanced Pancreatic Adenocarcinoma. *Clin Cancer Res* (2020) 26(19):5208–16. doi: 10.1158/1078-0432.CCR-20-0950
15. Darrigues L, Pierga JY, Bernard-Tessier A, Bieche I, Silveira AB, Michel M, et al. Circulating Tumor DNA as a Dynamic Biomarker of Response to Palbociclib and Fulvestrant in Metastatic Breast Cancer Patients. *Breast Cancer Res* (2021) 23(1):31. doi: 10.1186/s13058-021-01411-0
16. Tie J, Wang Y, Tomasetti C, Li L, Springer S, Kinde I, et al. Circulating Tumor DNA Analysis Detects Minimal Residual Disease and Predicts Recurrence in Patients With Stage II Colon Cancer. *Sci Trans Med* (2016) 8(346):346ra92. doi: 10.1126/scitranslmed.aaf6219
17. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, et al. Mutation Tracking in Circulating Tumor DNA Predicts Relapse in Early Breast Cancer. *Sci Trans Med* (2015) 7(302):302ra133. doi: 10.1126/scitranslmed.aab0021
18. Chaudhuri AA, Chabon JJ, Lovejoy AF, Newman AM, Stehr H, Azad TD, et al. Early Detection of Molecular Residual Disease in Localized Lung Cancer by Circulating Tumor DNA Profiling. *Cancer Discov* (2017) 7(12):1394–403. doi: 10.1158/2159-8290.CD-17-0716
19. Coombes RC, Page K, Salari R, Hastings RK, Armstrong A, Ahmed S, et al. Personalized Detection of Circulating Tumor DNA Antedates Breast Cancer Metastatic Recurrence. *Clin Cancer Res* (2019) 25(14):4255–63. doi: 10.1158/1078-0432.CCR-18-3663
20. Garcia-Murillas I, Chopra N, Comino-Mendez I, Beaney M, Tovey H, Cutts RJ, et al. Assessment of Molecular Relapse Detection in Early-Stage Breast Cancer. *JAMA Oncol* (2019) 5(10):1473–8. doi: 10.1001/jamaoncol.2019.1838
21. Narayan A, Carriero NJ, Gettinger SN, Kluytenaar J, Kozak KR, Yock TI, et al. Ultrasensitive Measurement of Hotspot Mutations in Tumor DNA in Blood Using Error-Suppressed Multiplexed Deep Sequencing. *Cancer Res* (2012) 72 (14):3492–8. doi: 10.1158/0008-5472.CAN-11-4037
22. Lin PH, Chen M, Tsai LW, Lo C, Yen TC, Huang TY, et al. Using Next-Generation Sequencing to Redefine BRCAness in Triple-Negative Breast Cancer. *Cancer Sci* (2020) 111(4):1375–84. doi: 10.1111/cas.14313
23. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* (2015) 17(5):405–24. doi: 10.1038/gim.2015.30
24. He MM, Li Q, Yan M, Cao H, Hu Y, He KY, et al. Variant Interpretation for Cancer (VIC): A Computational Tool for Assessing Clinical Impacts of Somatic Variants. *Genome Med* (2019) 11(1):53. doi: 10.1186/s13073-019-0664-4
25. Goh JY, Feng M, Wang W, Oguz G, Yatim S, Lee PL, et al. Chromosome 1q21.3 Amplification Is a Trackable Biomarker and Actionable Target for Breast Cancer Recurrence. *Nat Med* (2017) 23(11):1319–30. doi: 10.1038/nm.4405
26. Naab TJ, Gautam A, Ricks-Santi L, Esnakula AK, Kanaan YM, DeWitty RL, et al. MYC Amplification in Subtypes of Breast Cancers in African American Women. *BMC Cancer* (2018) 18(1):274. doi: 10.1186/s12885-018-4171-6
27. Lundberg A, Lindstrom LS, Li J, Harrell JC, Darai-Ramqvist E, Sifakis EG, et al. The Long-Term Prognostic and Predictive Capacity of Cyclin D1 Gene Amplification in 2305 Breast Tumours. *Breast Cancer Res* (2019) 21(1):34. doi: 10.1186/s13058-019-1121-4
28. Page K, Guttery DS, Fernandez-Garcia D, Hills A, Hastings RK, Luo J, et al. Next Generation Sequencing of Circulating Cell-Free DNA for Evaluating Mutations and Gene Amplification in Metastatic Breast Cancer. *Clin Chem* (2017) 63(2):532–41. doi: 10.1373/clinchem.2016.261834
29. Ptashkin RN, Mandelker DL, Coombs CC, Bolton K, Yelskaya Z, Hyman DM, et al. Prevalence of Clonal Hematopoiesis Mutations in Tumor-Only Clinical Genomic Profiling of Solid Tumors. *JAMA Oncol* (2018) 4(11):1589–93. doi: 10.1001/jamaoncol.2018.2297
30. Coombs CC, Gillis NK, Tan X, Berg JS, Ball M, Balasis ME, et al. Identification of Clonal Hematopoiesis Mutations in Solid Tumor Patients Undergoing Unpaired Next-Generation Sequencing Assays. *Clin Cancer Res* (2018) 24 (23):5918–24. doi: 10.1158/1078-0432.CCR-18-1201
31. Heuser M, Thol F, Ganser A. Clonal Hematopoiesis of Indeterminate Potential. *Dtsch Arztebl Int* (2016) 113(18):317–22. doi: 10.3238/arztebl.2016.0317
32. Cancer Genome Atlas N. Comprehensive Molecular Portraits of Human Breast Tumours. *Nature* (2012) 490(7418):61–70. doi: 10.1038/nature11412
33. Magbanua MJM, Swigart LB, Wu HT, Hirst GL, Yau C, Wolf DM, et al. Circulating Tumor DNA in Neoadjuvant-Treated Breast Cancer Reflects Response and Survival. *Ann Oncol* (2021) 32(2):229–39. doi: 10.1016/j.annonc.2020.11.007
34. Gabos Z, Sinha R, Hanson J, Chauhan N, Hugh J, Mackey JR, et al. Prognostic Significance of Human Epidermal Growth Factor Receptor Positivity for the Development of Brain Metastasis After Newly Diagnosed Breast Cancer. *J Clin Oncol* (2006) 24(36):5658–63. doi: 10.1200/JCO.2006.07.0250
35. Piccioni DE, Achrol AS, Kiedrowski LA, Banks KC, Boucher N, Barkhoudarian G, et al. Analysis of Cell-Free Circulating Tumor DNA in 419 Patients With Glioblastoma and Other Primary Brain Tumors. *CNS Oncol* (2019) 8(2):CNS34. doi: 10.2217/cns-2018-0015
36. Nami B, Maadi H, Wang Z. Mechanisms Underlying the Action and Synergism of Trastuzumab and Pertuzumab in Targeting HER2-Positive Breast Cancer. *Cancers (Basel)* (2018) 10(10):342. doi: 10.3390/cancers10100342

37. Zhang X, Zhao W, Wei W, You Z, Ou X, Sun M, et al. Parallel Analyses of Somatic Mutations in Plasma Circulating Tumor DNA (ctDNA) and Matched Tumor Tissues in Early-Stage Breast Cancer. *Clin Cancer Res* (2019) 25 (21):6546–53. doi: 10.1158/1078-0432.CCR-18-4055
38. Zhou Q, Perakis SO, Ulz P, Mohan S, Riedl JM, Talakic E, et al. Cell-Free DNA Analysis Reveals POLR1D-Mediated Resistance to Bevacizumab in Colorectal Cancer. *Genome Med* (2020) 12(1):20. doi: 10.1186/s13073-020-0719-6
39. McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, et al. Personalized Circulating Tumor DNA Analysis to Detect Residual Disease After Neoadjuvant Therapy in Breast Cancer. *Sci Trans Med* (2019) 11(504):eaax7392. doi: 10.1126/scitranslmed.aax7392
40. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026
41. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, et al. Direct Detection of Early-Stage Cancers Using Circulating Tumor DNA. *Sci Trans Med* (2017) 9(403):eaan2415. doi: 10.1126/scitranslmed.aan2415
42. Martinez-Saez O, Chic N, Pascual T, Adamo B, Vidal M, Gonzalez-Farre B, et al. Frequency and Spectrum of PIK3CA Somatic Mutations in Breast Cancer. *Breast Cancer Res* (2020) 22(1):45. doi: 10.1186/s13058-020-01284-9
43. Roy PG, Pratt N, Purdie CA, Baker L, Ashfield A, Quinlan P, et al. High CCND1 Amplification Identifies a Group of Poor Prognosis Women With Estrogen Receptor Positive Breast Cancer. *Int J Cancer* (2010) 127(2):355–60. doi: 10.1002/ijc.25034
44. Deming SL, Nass SJ, Dickson RB, Trock BJ. C-Myc Amplification in Breast Cancer: A Meta-Analysis of its Occurrence and Prognostic Relevance. *Br J Cancer* (2000) 83(12):1688–95. doi: 10.1054/bjoc.2000.1522

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 23 November 2021
doi: 10.3389/fonc.2021.779185

Impact of Preoperative *vs* Postoperative Radiotherapy on Overall Survival of Locally Advanced Breast Cancer Patients

Yujiao Deng[1,2†], Hongtao Li[3†], Yi Zheng[1,2], Zhen Zhai[1,2], Meng Wang[2], Shuai Lin[2], Yizhen Li[1,2], Bajin Wei[1], Peng Xu[2], Ying Wu[2], Xinyue Deng[2], Si Yang[1,2], Jun Lyu[4], Jingjing Hu[5], Huaying Dong[6] and Zhijun Dai[1,2*]*

[1] Department of Breast Surgery, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou, China, [2] Department of Oncology, The 2nd Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [3] Department of Breast Head and Neck Surgery, The 3rd Affiliated Teaching Hospital of Xinjiang Medical University (Affiliated Tumor Hospital), Urumqi, China, [4] Department of Clinical Research, The First Affiliated Hospital of Jinan University, Guangzhou, China, [5] Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, United States, [6] Department of General Surgery, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, China

OPEN ACCESS

Edited by:
Alberto Farolfi,
Istituto Scientifico Romagnolo per lo Studio e il Trattamento dei Tumori (IRCCS), Italy

Reviewed by:
Yiqun Han,
Chinese Academy of Medical Sciences and Peking Union Medical College, China
Qi Tian,
The First Affiliated Hospital of Xi'an Jiaotong University, China

***Correspondence:**
Zhijun Dai
dzj0911@126.com
Huaying Dong
dr_dhy@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: 18 September 2021
Accepted: 08 November 2021
Published: 23 November 2021

Citation:
Deng Y, Li H, Zheng Y, Zhai Z, Wang M, Lin S, Li Y, Wei B, Xu P, Wu Y, Deng X, Yang S, Lyu J, Hu J, Dong H and Dai Z (2021) Impact of Preoperative vs Postoperative Radiotherapy on Overall Survival of Locally Advanced Breast Cancer Patients. Front. Oncol. 11:779185. doi: 10.3389/fonc.2021.779185

Background: The treatment for locally advanced breast cancer (LABC) is a severe clinical problem. The postoperative radiotherapy is a conventional treatment method for patients with LABC, whereas the effect of preoperative radiotherapy on outcome of LABC remains controversial. This study aimed to examine and compare the overall survival (OS) in patients with LABC who underwent preoperative radiotherapy or postoperative radiotherapy.

Methods: This retrospective cohort study included 41,618 patients with LABC from the National Cancer Database (NCDB) between 2010 and 2014. We collected patients' demographic, clinicopathologic, treatment and survival information. Propensity score was used to match patients underwent pre-operative radiotherapy with those who underwent post-operative radiotherapy. Cox proportional hazard regression model was performed to access the association between variables and OS. Log-rank test was conducted to evaluate the difference in OS between groups.

Results: The estimated median follow-up of all included participants was 69.6 months (IQR: 42.84-60.22); 70.1 months (IQR: 46.85-79.97) for postoperative radiotherapy, 68.5 (IQR: 41.13-78.23) for preoperative radiotherapy, and 67.5 (IQR: 25.92-70.99) for no radiotherapy. The 5-year survival rate was 80.01% (79.56-80.47) for LABC patients who received postoperative radiotherapy, 64.08% (57.55-71.34) for preoperative radiotherapy, and 59.67% (58.60-60.77) for no radiotherapy. Compared with no radiation, patients receiving postoperative radiotherapy had a 38% lower risk of mortality (HR=0.62, 95%CI: 0.60-0.65, p<0.001), whereas those who received preoperative radiotherapy had no significant survival benefit (HR=0.88, 95%CI: 0.70-1.11, p=0.282). Propensity score matched analysis indicated that patients treated with

preoperative radiotherapy had similar outcomes as those treated with postoperative radiotherapy (AHR=1.23, 95%CI: 0.88-1.72, p=0.218). Further analysis showed that in C0 (HR=1.45, 95%CI: 1.01-2.07, p=0.044) and G1-2 (AHR=1.74, 95%CI: 1.59-5.96, p=0.001) subgroup, patients receiving preoperative radiotherapy showed a worse OS than those who received postoperative radiotherapy.

Conclusions: Patients with LABC underwent postoperative radiotherapy had improved overall survival, whereas no significant survival benefit was observed in patients receiving preoperative radiotherapy. Preoperative radiotherapy did not present a better survival than postoperative radiotherapy for LABC patients.

Keywords: locally advanced breast cancer, National Cancer Database, preoperative radiotherapy, postoperative radiotherapy, overall survival

INTRODUCTION

Breast cancer has become the most common cancer worldwide. Early breast cancer accounts for an increasing proportion of new breast cancer cases, and the disease burden continues to increase over time (1). Locally advanced breast cancer (LABC) encompasses stage III of the disease and a subset of patients with stage II (2), with a maximum lesion diameter of more than 5cm or lesion involving the surrounding skin or muscle, with or without axillary lymph node fusion and intramammary node, or ipsilateral supraconavicular node involvement.

The treatment of LABC is still a major challenge in patients with breast cancer because of the large space occupied by the primary lesions and serious local adhesions (3). Due to its low rate of overall survival (OS), high rate of recurrence and distant metastasis, LABC affects the overall survival of breast cancer largely (4). Currently, common adjuvant treatments for breast cancer are postoperative chemotherapy and radiotherapy (5). Radiotherapy is an effective treatment to reduce metastasis and improve the survival rate of breast cancer (6).

Recently, with the development of radiotherapy techniques, the value of preoperative radiotherapy has been reevaluated (7–10). Preoperative radiotherapy has been proven to prolong the prognosis of many cancers, such as rectal cancer (11), cervical cancer (12), et al. Some studies stated that preoperative radiotherapy could reduce the stage of tumor, increase the rate of surgical resection, alleviate symptoms and pain in patients, and improve the life quality of patients (9, 13). At present, there are few clinical studies on preoperative radiotherapy, and its effect for LABC patients is controversial (14–16). Early studies were mainly single-center, uncontrolled retrospective studies with small sample sizes, and the results were limited (17). In terms of long-term survival, the comparison between preoperative radiotherapy and postoperative radiotherapy lacks high-grade evidence-based data, and further investigation is needed.

Abbreviations: LABC, locally advanced breast cancer; OS, overall survival; NCDB, National Cancer Database; Her-2, human epidermal growth factor receptor 2; BCS, breast-conserving or -preserving surgery; IQR inter-quartile range; PSM, propensity score matching; AHR, adjusted hazard ratio.

The Nationally recognized National Cancer Database (NCDB), co-sponsored by the American College of Surgeons and the American Cancer Society, is a clinical oncology database derived from hospital registries collected by more than 1,500 Cancer Council accredited institutions. NCDB data were used to analyze and track patients with malignant cancer, their treatment and outcomes. The data represent more than 70 percent of newly diagnosed cancer cases and more than 34 million historical records nationwide (18). Based on the NCDB, we conducted this study to determine whether preoperative radiotherapy is superior to postoperative radiotherapy for the prognosis of patients with LABC. In this study, we analyzed the radiotherapy status of LABC patients who underwent surgery, and discussed the status and role of preoperative radiotherapy and postoperative radiotherapy in the treatment of LABC, as well as their prognostic value.

MATERIALS AND METHODS

Study Design and Data Sources

We performed a retrospective review of the NCDB data of LABC patients diagnosed between January 1, 2010, and December 31, 2014. All adult women with LABC were selected by the ICD-O-3 (histological code <8800), and were assigned according to the 7th AJCC TNM edition. Cases with LABC were defined as patients with stage III ($T_{0\text{-}2}N_2M_0$, $T_{3\text{-}4}N_{0\text{-}2}M_0$, $T_{0\text{-}4}N_3M_0$) and part of stage II B ($T_3N_0M_0$).

The inclusion criteria were as follows: (1) patients diagnosed with LABC in 2010-2014, microscopically confirmed, and only one malignant or *in situ* primary tumor in the patient's lifetime; (2) patients who underwent breast surgery with a specific surgical procedure; (3) patients with no distant metastasis; (4) cases were females and aged ≥18.

We excluded cases for any of the following reasons: (1) lack of data on estrogen receptor, progesterone receptor, or human epidermal growth factor receptor 2 (Her-2); (2) unknown tumor grade or stage; (3) unknown status of chemotherapy, hormone therapy, or immunotherapy treatment; (3) lack of data on insurance, income, home location, vital status, or follow-up time; (4) if the patient received radiation therapy both before and

after surgery or if they received intraoperative radiation with or without another therapy, in an unknow sequence except for postoperative radiotherapy, preoperative radiotherapy, and no radiation.

Data Extraction and Outcomes

All included LABC patients were confirmed by cytology, histopathology, or microscopy and had only one lifetime history of malignancy or *in situ* recurrence, with no distant metastasis. We used the Charlson-Deyo Comorbidity Index (CCI) to quantify comorbid conditions. In total, eighteen factors were extracted: age at diagnosis, race, insurance provider (Medicaid, Medicare, or Private insurance/managed Care), median household income (high, high-middle, low-middle, or low), home location (rural, urban, or metro); CCI, grade (G1, well differentiated; G2, moderately differentiated; G3, poorly differentiated; G4, undifferentiated); tumor stage (T stage), nodal stage (N stage), molecular subtype (luminal, Her-2 positive, and triple-negative breast cancer); clinical stage, chemotherapy, hormone therapy, immunotherapy, surgery method; sequence of patients receiving radiotherapy and surgery, vital status, and follow-up time. The surgical procedure included total (simple) mastectomy, breast-conserving or -preserving surgery (BCS), and radical mastectomy. The race of the patients was divided into white, black, Asian/other. The pathological results of patients were classified into three categories based on ER, PR, and ERBB2 status. Luminal subtype was ER or PR positive, with or without ERBB2 positive. Her-2 positive subtype meant that both ER and PR are negative and ERBB2 is positive. Triple-negative subtype was defined as negative for estrogen receptor (ER), progesterone receptor (PR) and ERBB2 or Her-2. ER and PR were considered negative if less than 1% of cells stain positive. If the immunohistochemistry score was 0 to 1+ or fluorescence *in situ* hybridization and color *in situ* hybridization do not amplify, ERBB2 status was considered negative. The primary outcome was the rate of overall survival after breast surgery and radiotherapy. The endpoint was defined as the vital status of patients at last contact (alive or deceased). And the number of months to last contact were recorded. The diagram outlining all the selection criteria is presented in **Figure 1**.

Statistical Analysis

We used frequency (percentage) to express categorical variables data and reported quantitative variables in quartile range (IQR). χ^2 test or Fisher's exact test was used for qualitative variables, and unpaired Kruskal-Wallis test was applied in quantitative variables. The Bonferroni test was conducted to compare sociodemographic, therapeutic, and tumor characteristics between the three treatment groups. In addition, from diagnosis to the last contact or death, the OS rate was calculated on a monthly basis. Univariate and multivariate Cox proportional hazard models were used to investigate the factors affecting OS in the unmatched and matched cohort. To solve the imbalance between patients receiving postoperative and preoperative radiotherapy, we conducted propensity score matching (PSM) analysis (19). We matched the conditional probability propensity scores for adjuvant radiotherapy before and after surgery. The variables included in the PSM analysis were age, race, insurance, income, home location, CCI, grade, T stage, N stage, molecular subtype, clinical stage, chemotherapy, hormone therapy, immunotherapy, surgery method. These variables are potential factors affecting the probability of receiving radiotherapy treatment. To avoid over-fitness, items (radiation and surgery sequence) entered into the PSM were excluded from the multivariate Cox regression analysis. The Kaplan-Meier curve was fit to calculate cumulative survival in unmatched and propensity matched cohorts. A log-rank test was performed to test the differences in the cumulative proportions across different treatment groups (20). Our study was reported followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline (**eTable 1**). All statistical tests were two-sided, the significance level of the Bonferroni test was 0.0167, and the significance level of other tests was 0.05. All statistical analyses were conducted using R software for Windows, version 4.0.5 (R Project for Statistical Computing).

RESULTS

Patient Characteristics

A total of 41,618 cases met the inclusion criteria outlined above and were enrolled in our initial non-matched analysis (**Figure 1**). Among these patients, 32,625 (78.39%) experienced postoperative radiotherapy, 8,787 (21.11%) received no adjuvant radiation, and 206 (0.49%) endured preoperative radiotherapy. Compared with patients experienced preoperative radiotherapy, the postoperative radiotherapy cohort was younger (mean age, 59.24 *vs* 59.27, $p<0.001$), more Asians ($p<0.001$), more private insurance payers ($p=0.002$), more luminal tumors ($p<0.001$); and had better differentiation levels ($p<0.001$), lower tumor stage ($p<0.001$), higher nodal stage ($p=0.005$), better prognosis ($p<0.001$); more patients received hormone therapy ($p<0.001$) and BCS ($p<0.001$). There were no significant differences of distribution between preoperative and postoperative groups in income, home location, CCI grade, clinical stage, chemotherapy, and immunotherapy (**Table 1**).

Univariate and Multivariate Analysis

The estimated median follow-up time was 70.1 months (IQR: 46.85-79.97, range: 2.92-112.95, 95%CI: 69.7-70.5) for postoperative radiotherapy, 68.5 (IQR: 41.13-78.23, range: 4.99-111.57, 95%CI: 65.2- 74.8) for preoperative radiotherapy. The 5-year survival rate was 80.01% (79.56-80.47) for LABC patients receiving postoperative radiotherapy, 64.08% (57.55-71.34) for preoperative radiotherapy. In the survival analysis of the unmatched cohort, postoperative radiotherapy was related associated with improved OS compared to no radiation ($p<0.0001$) (**Figure 2A**). Similarly, in the multivariable Cox analysis adjusted for confounders, patients who received

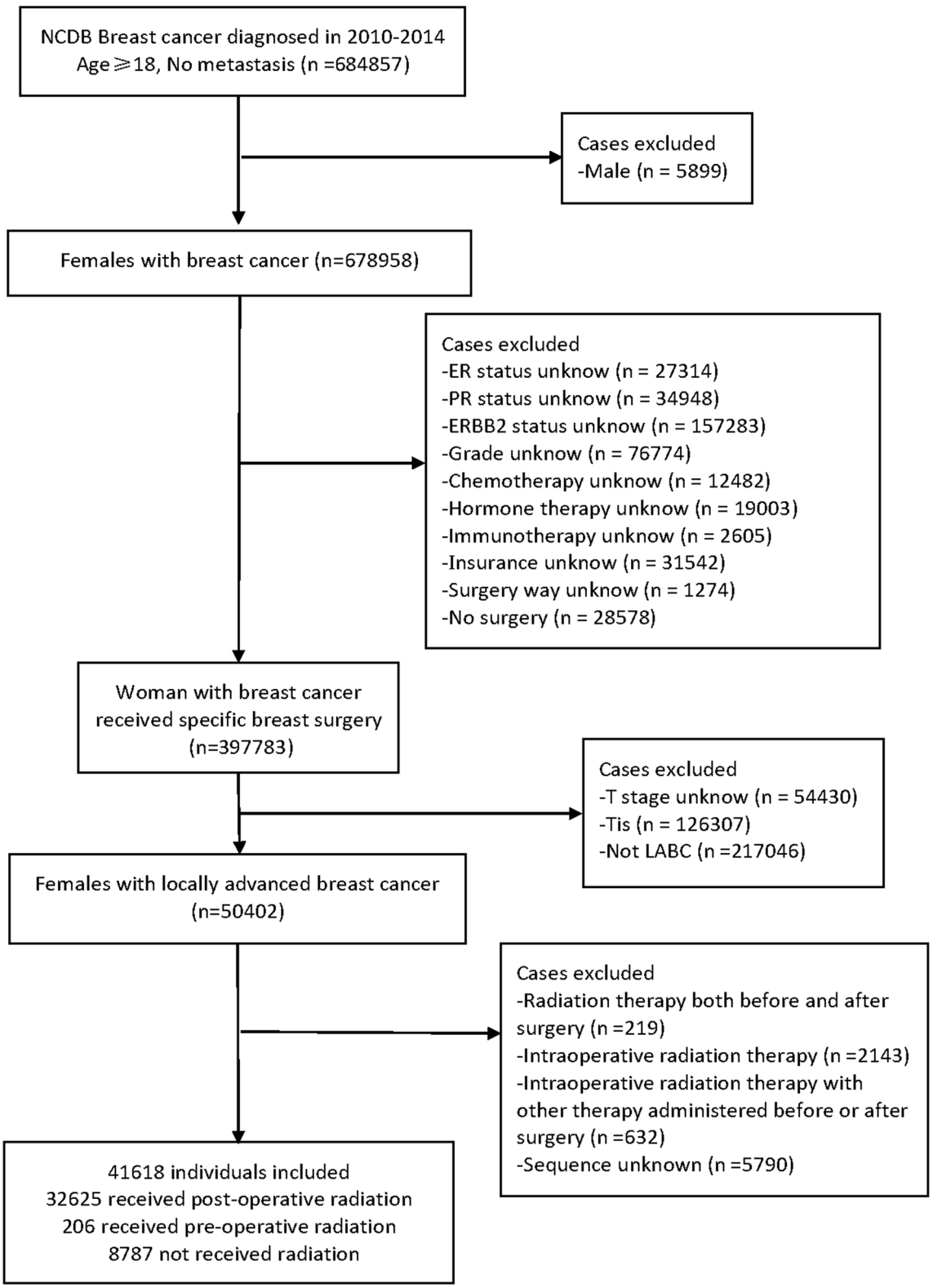


FIGURE 1 | Flow Chart of participants Selection in National Cancer Database.

TABLE 1 | Patient demographic, disease, and treatment characteristics of locally advanced breast cancer grouped by radiation status.

Variable	Total population (No.)	No radiation		Postoperative radiotherapy		Preoperative radiotherapy		P value		
		(No.)	%	(No.)	%	(No.)	%	No radiation *vs* Postoperative radiotherapy	No radiation *vs* Preoperative radiotherapy	Postoperative radiotherapy *vs* Preoperative radiotherapy
Age, mean (SD), years	59.82(12.91)	60.98 (14.64)		59.24 (11.89)		56.27 (13.00)		<0.001[a]	<0.001[a]	<0.001[a]
Age distribution (years)										
<35	1415	187	13.22	1212	85.65	16	1.13	<0.001[a]	<0.001[a]	0.015[a]
35-50	10860	1575	14.50	9236	85.05	49	0.45			
50-70	20345	3434	16.88	16804	82.60	107	0.53			
≥70	8998	3591	39.91	5373	59.71	34	0.38			
Race										
White	33144	6927	20.90	26073	78.67	144	0.43	<0.001[a]	0.001a	<0.001[a]
Asia/ other	2140	401	18.74	1731	80.89	8	0.37			
Black	6334	1459	23.03	4821	76.11	54	0.85			
Insurance										
Not insured	1436	267	18.59	1153	80.29	16	1.11	<0.001[a]	<0.001[a]	0.002[a]
Medicaid	4800	941	19.60	3830	79.79	29	0.60			
Medicare	12995	4258	32.77	8676	66.76	61	0.47			
Private Insurance/ Managed Care	22387	3321	14.83	18966	84.72	100	0.45			
Income										
Low	7839	1937	24.71	5856	74.70	46	0.59	<0.001[a]	<0.001[a]	0.381
High	15333	2878	18.77	12384	80.77	71	0.46			
High-middle	9612	1980	20.60	7583	78.89	49	0.51			
Low-middle	8834	1992	22.55	6802	77.00	40	0.45			
Home location										
Rural/urban	5720	1207	21.10	4491	78.51	22	0.38	0.957	0.246	0.238
Metro	35898	7580	21.12	28134	78.37	184	0.51			
Charlson Comorbidity Index										
C0	34199	6684	19.54	27344	79.96	171	0.50	<0.001[a]	0.053	0.11
C1	5921	1574	26.58	4323	73.01	24	0.41			
C2-3	1498	529	35.31	958	63.95	11	0.73			
Grade										
G1-2	22435	4344	19.36	18008	80.27	83	0.37	<0.001[a]	0.0116[a]	<0.001[a]
G3-4	19183	4443	23.16	14617	76.20	123	0.64			
Tumor stage										
T0-1	7016	767	10.93	6224	88.71	25	0.36	<0.001[a]	<0.001[a]	<0.001[a]
T2	13610	2007	14.75	11551	84.87	52	0.38			
T3	17616	4828	27.41	12718	72.20	70	0.40			
T4	3376	1185	35.10	2132	63.15	59	1.75			
Nodal stage										
N0	8255	3153	38.20	5053	61.21	49	0.59	<0.001[a]	<0.001[a]	0.005[a]
N1	5787	1558	26.92	4198	72.54	31	0.54			
N2	19298	2771	14.36	16437	85.17	90	0.47			
N3	8278	1305	15.76	6937	83.80	36	0.43			
Stage										
S0-2	8050	2883	35.81	5134	63.78	33	0.41	<0.001[a]	<0.001[a]	0.988
S3-4	33568	5904	17.59	27491	81.90	173	0.52			
Chemotherapy										
No	8054	4207	52.23	3827	47.52	20	0.25	<0.001[a]	<0.001[a]	0.429
Yes	33564	4580	13.65	28798	85.80	186	0.55			
Hormone therapy										
No	12563	4408	35.09	8055	64.12	100	0.80	<0.001[a]	0.697	<0.001[a]
Yes	29055	4379	15.07	24570	84.56	106	0.36			
Immunotherapy										
No	39252	8457	21.55	30597	77.95	198	0.50	<0.001[a]	0.99	0.215
Yes	2366	330	13.95	2028	85.71	8	0.34			

(Continued)

TABLE 1 | Continued

Variable	Total population (No.)	No radiation		Postoperative radiotherapy		Preoperative radiotherapy		P value		
		(No.)	%	(No.)	%	(No.)	%	No radiation *vs* Postoperative radiotherapy	No radiation *vs* Preoperative radiotherapy	Postoperative radiotherapy *vs* Preoperative radiotherapy
Subtype										
Luminal	31257	5895	18.86	25247	80.77	115	0.37	<0.001[a]	<0.001[a]	<0.001[a]
Triple negative	7677	2209	28.77	5391	70.22	77	1.00			
Her-2	2684	683	25.45	1987	74.03	14	0.52			
Surgery										
Simple mastectomy	13582	3586	26.40	9938	73.17	58	0.43	<0.001[a]	<0.001[a]	<0.001[a]
BCS/other	9374	1200	12.80	8150	86.94	24	0.26			
Radical mastectomy	18662	4001	21.44	14537	77.90	124	0.66			
Vital status										
Alive	30352	5108	16.83	25116	82.75	128	0.42	<0.001[a]	0.28	<0.001[a]
Deceased	11266	3679	32.66	7509	66.65	78	0.69			

BCS, breast-conserving surgery; SD, standard deviation.

[a]*Significance was evaluated using Bonferroni test. The statistical tests were two-sided, the significance level was 0.0167.*

postoperative radiotherapy had a 38% lower risk of mortality [Adjusted HR (AHR) =0.62, 95%CI: 0.60-0.65, p<0.001]. However, there was no significant difference in prognosis between patients who received preoperative radiotherapy and those who did not (HR=0.85, 95%CI: 0.68-1.06, p=0.148; AHR=0.88, 95%CI: 0.70-1.11, p=0.282, **Table 2**).

Multivariable Cox analysis revealed that some factors were independently associated with improved or worse OS in LABC patients. Among these, the highest HR was for high nodal stage of LABC (N1/N2/N3 *vs*. N0), with N3 patients having an AHR of 3.49 (95%CI: 3.14-3.89, p<0.001, **Table 2**) and tumor stage ≥T2, and those with T4 having an AHR of 2.18 (95%CI: 2.02-2.36, p<0.001). Compared with well or moderately differentiated LABC, patients with poorly differentiated or undifferentiated histology had a 53% higher mortality risk (AHR= 1.53, 95%CI: 1.46-1.60, p<0.001). Compared with patients aged 35-50 years, patients aged <35 years, 50-70 years, and ≥70 years had a 23% (p<0.001), 11% (p<0.001), and 63% (p<0.001) higher mortality risk, respectively. Black patients had a 16% higher mortality risk (AHR= 1.16, 95%CI: 1.10-1.22, p<0.001) than white patients. Patients classified as C1 and C2-3 on the CCI had higher mortality risk values compared to C0 patients (C1: AHR=1.27, 95%CI: 1.21-1.33, p<0.001; C2-3: AHR=1.67, 95%CI: 1.55-1.80, p<0.001). Other factors associated with poor survival included clinical stage (stage 0-2 *vs* 3-4: AHR=1.14, 95%CI: 1.03-1.26, p=0.012), triple-negative subtype (triple-negative *vs*. luminal: AHR=1.94, 95%CI: 1.81-2.09, p<0.001), and the receipt of radical mastectomy (radical *vs*. simple: AHR=1.12, 95%CI: 1.07-1.17, p<0.001). In addition, some factors were associated with improved survival of patients with LABC. Asian and other races had a 23% lower mortality risk than white patients (AHR= 0.77, 95%CI: 0.69-0.85, p<0.001). Compared with patients who were not insured, private insurance payers had a 22% lower mortality risk (AHR= 0.78, 95%CI: 0.70-0.86, p<0.001). In addition, compared with low-income patients, those who carried a high median household income had a 6% lower mortality risk (AHR= 0.94, 95%CI: 0.89-0.99, p=0.041). Patients who lived in metro had a 7% lower mortality risk (AHR= 0.93, 95%CI: 0.88-0.99, p=0.014) than those who lived in rural or urban areas. As presented in **eTable 2**, patients who received preoperative radiotherapy combined with chemotherapy (HR= 0.34, 95%CI: 0.19-0.62, p<0.001) or hormone therapy (HR= 0.56, 95%CI: 0.36-0.88, p=0.012) showed better outcomes compared with their counterparts without corresponding treatments. The univariate Cox analysis results of patients who received postoperative radiotherapy were shown in **eTable 3**.

Propensity Score–Matched Analysis and Outcomes

The estimated median follow-up time was 71.4 months (IQR: 34.37-75.22, range: 4.50-107.04, 95%CI: 67.40-75.20) for patients who received postoperative radiotherapy and 68.5 months (IQR: 65.20-74.80, range: 4.99-111.57, 95%CI: 65.2-74.8) for those who experienced preoperative radiotherapy. The 5-year survival rate was 66.29% (59.82-73.47) for those who received postoperative radiotherapy and 64.08% (57.55-71.34) for those who endured preoperative radiotherapy. In the multivariable analysis of the matched cohort (**Table 3**), patients aged ≥70 years had a three times higher risk of mortality (AHR= 3.83, 95%CI: 1.81-8.11, p<0.001) compared to those aged 35-50 years. Black patients had a 59% worse OS (AHR= 1.59, 95%CI: 1.07-2.37, p<0.001) than white patients. In addition, factors associated with poor OS in the matched cohort

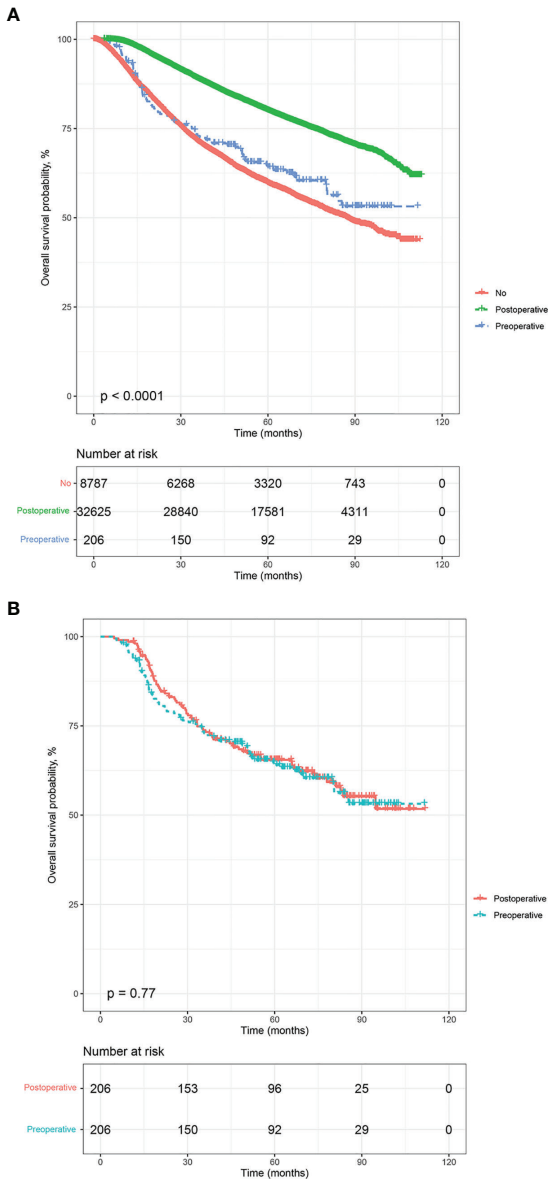


FIGURE 2 | Kaplan-Meier overall survival analysis, before and after propensity score matching. **(A)** all participants, **(B)** matched population.

included tumor stages T3 (T3 *vs* T0-1: AHR= 2.09, 95%CI: 1.09-4.02, p=0.027) and T4 (T4 *vs* T0-1:AHR= 3.45, 95%CI: 1.82-6.54, p<0.001), nodal stages N1 (N1 *vs* N0: AHR= 3.37, 95%CI: 1.48-7.68, p=0.004), N2 (N2 *vs* N0: AHR= 10.01, 95% CI: 4.59-21.83, p<0.001), and N3 (N3 *vs* N0: AHR= 10.26, 95% CI: 4.62-22.78, p<0.001), triple-negative subtype (Triple negative *vs* Luminal: AHR= 9.02, 95%CI: 3.90-20.86, p<0.001), Her-2 positive subtype (Her-2 positive *vs* Luminal: AHR= 4.17, 95%CI: 1.48-11.72, p=0.007), and patients underwent radical mastectomy (AHR= 1.71, 95%CI: 1.10-2.66, p=0.017). Finally, patients who endured preoperative radiotherapy had a statistically similar prognosis to those who received postoperative radiotherapy (AHR=1.23, 95%CI: 0.88-1.72, p=0.218). Survival analysis indicated no difference existed in the OS of LABC patients between preoperative radiotherapy and postoperative radiotherapy (p=0.77, **Figure 2B**). In addition, patients in C0 (HR=1.45, 95%CI: 1.01-2.07, p=0.044) and G1-2 subgroup (AHR=1.74, 95%CI: 1.59-5.96, p=0.001) experienced preoperative radiotherapy showed a worse OS than those who received postoperative radiotherapy (**Figure 3**).

DISCUSSION

In this hospital-based registry analysis, postoperative radiotherapy presented a significant benefit for improved OS of LABC patients compared to no radiation, which appears to be consistent with a previous study (21). The benefit was also observed in patients who endured preoperative radiotherapy. However, the benefit was not statistically significant. PSM matched analysis indicated that, compared with postoperative radiotherapy, no survival improvement was observed in LABC patients who experienced preoperative radiotherapy. The effect of postoperative radiotherapy for LABC patients had been confirmed by several large clinical trials, which could significantly increase the local control rates and improve their OS rates (22).

In recent years, the value of preoperative radiotherapy in the treatment of LABC patients has been reassessed. Studies showed that new adjuvant chemotherapy improved the pathological complete response of tumors (23). A Previous study reported on the benefits of preoperative radiotherapy or chemotherapy on tumor treatment (24) and the impact of breast reconstruction surgery, as well as its value in tumor biology and translational medicine research. Our analysis illustrated that patients receiving preoperative radiotherapy combined with chemotherapy or hormone therapy showed prognosis benefit, which is consistent with published studies (25, 26). Through the combined use of preoperative radiotherapy and drugs, clinicians can obtain a clinical effect evaluation in a relatively short period of time and guide follow-up treatment by observing lesion changes (27, 28). However, approximately 1/3 of LABC patients are resistant to neoadjuvant chemotherapy, and there is still no manual resection opportunity for the tumor after chemotherapy. In this case, preoperative radiotherapy (21) or preoperative concurrent chemoradiotherapy is an important salvage treatment measure which could reduce the tumor load in some patients and provide the opportunity for surgical resection (2). Preoperative radiotherapy could increase the sensitivity of radiotherapy (29), cause tumor tissue fibrosis, reduce the risk of intraoperative implantation and metastasis, change the tumor microenvironment, transform the tumor immune escape state into a tumor immune attack state, and activate the immune system to produce long-distance effects (9, 30). However, the high incidence of acute toxic reactions is

TABLE 2 | Univariable and multivariable Cox analysis of overall survival for patients with locally advanced breast cancer.

Variable	Total population	Alive	Deceased		Univariable Analysis		Multivariable Analysis	
		No.	No.	%	HR (95% CI)	P value	HR (95% CI)	P value
Age distribution (years)								
35-50	10860	8731	2129	19.60	1 (Ref.)		1 (Ref.)	
<35	1415	1055	360	25.44	1.36 (1.22-1.53)	<0.001	1.23 (1.10-1.37)	<0.001[a]
50-70	20345	15582	4763	23.41	1.22 (1.16-1.28)	<0.001	1.11 (1.05-1.17)	<0.001[a]
≥70	8998	4984	4014	44.61	2.80 (2.65-2.95)	<0.001	1.63 (1.52-1.76)	<0.001[a]
Race								
White	33144	24421	8723	26.32	1 (Ref.)			
Asia/other	2140	1758	382	17.85	0.68 (0.61-0.75)	<0.001	0.77 (0.69-0.85)	<0.001[a]
Black	6334	4173	2161	34.12	1.42 (1.36-1.49)	<0.001	1.16 (1.10-1.22)	<0.001[a]
Insurance								
Not insured	1436	1049	387	26.95	1 (Ref.)		1 (Ref.)	
Medicaid	4800	3416	1384	28.83	1.07 (0.95-1.19)	0.264	0.99 (0.89-1.11)	0.898
Medicare	12995	7975	5020	38.63	1.50 (1.35-1.66)	<0.001	0.97 (0.87-1.09)	0.646
Private Insurance/Managed Care	22387	17912	4475	19.99	0.66 (0.60-0.74)	<0.001	0.78 (0.70-0.86)	<0.001[a]
Income								
Low	7839	5336	2503	31.93	1 (Ref.)		1 (Ref.)	
High	15333	11767	3566	23.26	0.67 (0.63-0.70)	<0.001	0.94 (0.89-0.99)	0.041[a]
High-middle	9612	7020	2592	26.97	0.81 (0.76-0.85)	<0.001	0.97 (0.92-1.03)	0.344
Low-middle	8834	6229	2605	29.49	0.90 (0.85-0.95)	<0.001	1.03 (0.98-1.092)	0.27
Home location								
Rural/urban	5720	4041	1679	29.35	1 (Ref.)		1 (Ref.)	
Metro	35898	26311	9587	26.71	0.88 (0.84-0.93)	<0.001	0.93 (0.88-0.99)	0.014[a]
Charlson Comorbidity Index								
C0	34199	25771	8428	24.64	1 (Ref.)		1 (Ref.)	
C1	5921	3843	2078	35.10	1.54 (1.46-1.61)	<0.001	1.27 (1.21-1.33)	<0.001[a]
C2-3	1498	738	760	50.73	2.59 (2.41-2.79)	<0.001	1.67 (1.55-1.80)	<0.001[a]
Grade								
G1-2	22435	18100	4335	19.32	1 (Ref.)		1 (Ref.)	
G3-4	19183	12252	6931	36.13	2.18 (2.09-2.26)	<0.001	1.53 (1.46-1.60)	<0.001[a]
Tumor stage								
T0-1	7016	5454	1562	22.26	1 (Ref.)		1 (Ref.)	
T2	13610	9939	3671	26.97	1.25 (1.18-1.33)	<0.001	1.12 (1.05-1.19)	<0.001[a]
T3	17616	13198	4418	25.08	1.19 (1.12-1.26)	<0.001	1.57 (1.47-1.68)	<0.001[a]
T4	3376	1761	1615	47.84	2.75 (2.56-2.95)	<0.001	2.18 (2.02-2.36)	<0.001[a]
Nodal stage								
N0	8255	6478	1777	21.53	1 (Ref.)		1 (Ref.)	
N1	5787	4329	1458	25.19	1.18 (1.10-1.26)	<0.001	1.39 (1.24-1.54)	<0.001[a]
N2	19298	14438	4860	25.18	1.14 (1.08-1.21)	<0.001	2.35 (2.12-2.62)	<0.001[a]
N3	8278	5107	3171	38.31	1.90 (1.79-2.01)	<0.001	3.49 (3.14-3.89)	<0.001[a]
Stage								
S0-2	8050	6400	1650	20.50	1 (Ref.)		1 (Ref.)	
S3-4	33568	23952	9616	28.65	1.44 (1.37-1.52)	<0.001	1.14 (1.03-1.26)	0.012[a]
Chemotherapy								
No	8054	4794	3260	40.48	1 (Ref.)	<0.001	1 (Ref.)	
Yes	33564	25558	8006	23.85	0.49 (0.47-0.51)	<0.001	0.57 (0.54-0.60)	<0.001[a]
Hormone therapy								
No	12563	7024	5539	44.09	1 (Ref.)		1 (Ref.)	
Yes	29055	23328	5727	19.71	0.34 (0.33-0.35)	<0.001	0.62 (0.58-0.66)	<0.001[a]
Immunotherapy								
No	39252	28387	10865	27.68	1 (Ref.)			
Yes	2366	1965	401	16.95	0.69 (0.63-0.77)	<0.001	0.85 (0.77-0.95)	0.003[a]
Subtype								
Luminal	31257	24683	6574	21.03	1 (Ref.)		1 (Ref.)	
Triple negative	7677	3822	3855	50.21	3.34 (3.21-3.48)	<0.001	1.94 (1.81-2.09)	<0.001[a]
Her-2	2684	1847	837	31.18	1.60 (1.49-1.72)	<0.001	0.93 (0.84-1.02)	0.125
Surgery								
Simple mastectomy	13582	10482	3100	22.82	1 (Ref.)		1 (Ref.)	
BCS/other	9374	7279	2095	22.35	0.96 (0.91-1.01)	0.137	0.95 (0.90-1.01)	0.085
Radical mastectomy	18662	12591	6071	32.53	1.48 (1.42-1.55)	<0.001	1.12 (1.07-1.17)	<0.001[a]
Radiotherapy								

(Continued)

TABLE 2 | Continued

Variable	Total population	Alive	Deceased		Univariable Analysis		Multivariable Analysis	
		No.	No.	%	HR (95% CI)	P value	HR (95% CI)	P value
No	8787	5108	3679	41.87	1 (Ref.)		1 (Ref.)	
Postoperative radiotherapy	32625	25116	7509	23.02	0.43 (0.42-0.45)	<0.001	0.62 (0.60-0.65)	<0.001[a]
Preoperative radiotherapy	206	128	78	37.86	0.85 (0.68-1.06)	0.148	0.88 (0.70-1.11)	0.282

BCS, breast-conserving surgery; HR, hazard ratio.
[a]*The statistical tests were two-sided, the significance level was 0.05.*

TABLE 3 | Propensity-adjusted multivariable Cox regression analysis of overall survival for locally advanced breast cancer.

Variable	HR (95% CI)	P value
Age distribution (years)		
35-50	1 (Ref.)	
<35	1.09 (0.54-2.21)	0.814
50-70	1.33 (0.83-2.14)	0.233
≥70	3.83 (1.81- 8.11)	<0.001[a]
Race		
White	1 (Ref.)	
Asia/other	1.31 (0.45-3.81)	0.625
Black	1.59 (1.072-2.37)	0.021[a]
Insurance		
Not insured	1 (Ref.)	
Medicaid	1.27 (0.48-3.35)	0.63
Medicare	0.67 (0.26-1.74)	0.406
Private Insurance/Managed Care	1.09 (0.44-2.67)	0.858
Income	1 (Ref.)	
Low	1.35 (0.80-2.27)	0.259
High	1.09 (0.64-1.88)	0.741
High-middle	1.70 (0.99-2.90)	0.054
Low-middle		
Home location		
Rural/urban	1 (Ref.)	
Metro	1.13 (0.59-2.19)	0.71
Charlson Comorbidity Index		
C0	1 (Ref.)	
C1	1.12 (0.62-2.02)	0.706
C2-3	1.48 (0.74-2.97)	0.266
Grade		
G1-2	1 (Ref.)	
G3-4	1.07 (0.72-1.60)	0.736
Tumor stage		
T0-1	1 (Ref.)	
T2	1.76 (0.96-3.25)	0.07
T3	2.09 (1.09-4.02)	0.027[a]
T4	3.45 (1.82-6.54)	<0.001[a]
Nodal stage		
N0	1 (Ref.)	
N1	3.37 (1.48-7.68)	0.004[a]
N2	10.01 (4.59-21.83)	<0.001[a]
N3	10.26 (4.62-22.78)	<0.001[a]
Stage		
S0-2	1 (Ref.)	
S3-4	0.71 (0.30-1.68)	0.436
Chemotherapy		
No	1 (Ref.)	
Yes	1.06 (0.58-1.94)	0.855
Hormone therapy		
No	1 (Ref.)	
Yes	1.60 (0.71-3.58)	0.254

(Continued)

TABLE 3 | Continued

Variable	HR (95% CI)	P value
Immunotherapy		
No	1 (Ref.)	
Yes	1.46 (0.52-4.13)	0.472
Subtype		
Luminal	1 (Ref.)	
Triple negative	9.02 (3.90-20.86)	<0.001[a]
Her-2	4.17 (1.48-11.72)	0.007[a]
Surgery		
Simple mastectomy	1 (Ref.)	
BCS/other	1.80 (0.89-3.62)	0.1
Radical mastectomy	1.71 (1.10-2.66)	0.017[a]
Radiotherapy		
Postoperative radiotherapy	1 (Ref.)	
Preoperative radiotherapy	1.23 (0.88-1.72)	0.218

BCS, breast-conserving surgery; HR, hazard ratio.

[a]*The statistical tests were two sided, the significance level was 0.05.*

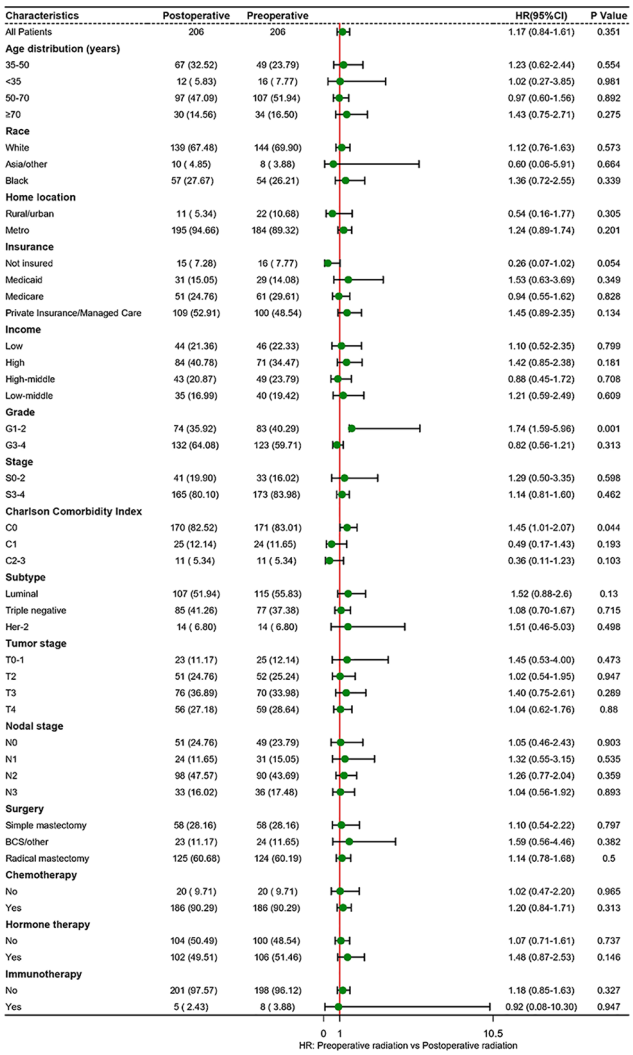


FIGURE 3 | Subgroup analyses of radiotherapy treatment based on matched population. BCS, breast-conserving surgery; HR, hazard ratio.

attributed to the lack of therapeutic experience and/or technical limitations due to factors such as concurrent chemotherapy, a high total dose of radiotherapy, and the limit of radiation techniques. Severe toxic reactions are the most important reason for the limited clinical application of preoperative radiotherapy or preoperative neoadjuvant concurrent chemoradiotherapy (31). Several studies have demonstrated the favorable effect of preoperative radiotherapy on tumor treatment and breast reconstruction surgery, as well as its value in tumor biology and translational medicine research.

Radiotherapy is important for the treatment of breast cancer, improving the local control rate and OS of patients at a high risk of recurrence. For advanced breast cancer (16), preoperative radiotherapy can reduce tumor stage, increase the resection rate, and alleviate the symptoms of patients.

Clinically, the selection of neoadjuvant radiotherapy for patients is limited to a certain extent, and there is currently no unified standard. Most clinical decisions depend on the clinical experience of doctors, so there may be the possibility of overtreatment. In our analysis, black patients with LABC were more inclined to endure preoperative radiotherapy, especially for patients with T4 stage tumors, aged 50-70 years, uninsured, triple negative subtype, poorly or undifferentiated. As for surgery method, the proportion of patients undergoing radical breast cancer resection undergoing preoperative radiotherapy was higher than that of patients undergoing other surgical procedures. Besides, neoadjuvant radiotherapy or chemoradiotherapy may lead to vascular injury and microcirculation disturbance, resulting in tissue cell degeneration and necrosis, breast fibrosis and skin injury. However, the fibrotic and damaged skin of the breast increases the difficulty of operation and prolongs the operation time, making radiotherapy as a neoadjuvant therapy method not widely employed for breast cancer (7). Suitable and safe treatment plans timelines, and treatment modalities with long survival rates, short and convenient reconstruction processes, and good appearance should be determined for LABC patients. In addition, biomarkers that are sensitive to radiation and chemotherapy should be ascertained.

A study based on 129,692 patients supported that breast-conserving surgery with radiation therapy improved the survival of breast cancer patients (26). Patients with stage IIB-IIIA breast cancer are generally considered having "operable breast cancer". In contrast, those receiving postoperative radiotherapy or with stage IIIB and IIIC cancer are likely to be classified as inoperable cases; this is due to the presence of inflammation and/or extensive skin involvement, immobilization, or very large axillary lymph node disease, and/or the involvement of supraclavicular or internal breast lymph nodes (32). However, preoperative radiotherapy provides LABC patients with no chance of surgery with the opportunity of surgical treatment, as well as the opportunity of breast-conserving surgery for patients who cannot initially undergo breast-conserving surgery (24), thus improving their quality of life (33). By comparing the tumor tissues before and after radiotherapy and analyzing the various differences at the molecular level, biological information related to the radio sensitivity of tumor cells can be obtained, which helps to understand the changes in the immune microenvironment (34).

There are some limitations inevitable in this study. A small percentage of patients with LABC received preoperative radiation. Due to the limited data, we could not perform further subgroup analysis on the radiotherapy duration and dose of patients. The study population included patients who were diagnosed with LABC and underwent breast surgery. It should be emphasized that the application of these results cannot be expanded to general breast cancer. Although we used a retrospective paired study to select the control group, there is still an unavoidable selection bias, and there are some unknown influencing factors that will affect the final study conclusion. Besides, due to the limited data of preoperative radiotherapy, we had not been able to do a preoperative and postoperative analysis of other treatments (e.g., chemotherapy, endocrine therapy, immunotherapy) in LABC patients. However, we had enrolled those potential factors into the PSM analysis, and the effect of the variables has been largely balanced. In addition, we analyzed the relationship between the two types of radiotherapy combined with other treatments independently. We recommend that patients with LABC be treated in combination with chemotherapy or hormone therapy, regardless of preoperative or postoperative radiotherapy. Nevertheless, the role and value of preoperative radiotherapy or concurrent radio-chemotherapy for the treatment of LABC under the application of novel radiotherapy technologies and medicines requires confirmation and investigation by prospective, multi-center, randomized controlled clinical studies with large sample sizes.

In this study, patients with LABC who received postoperative radiotherapy were associated with improved OS, while those who received preoperative radiotherapy had no significant benefit. In the matched analysis, there was no significant difference in survival between patients receiving postoperative radiotherapy and those who receiving preoperative radiotherapy. The conclusions still need to be confirmed in large prospective clinical trials.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The First Affiliated Hospital of Zhejiang University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

All authors read, critically reviewed and approved the final manuscript. ZD and HD designed the research. YD, HL, JL, and ZZ collected the data. MW and SL verified the accuracy of the data. YD, SY, YL, and BW performed the statistical analysis. PX, YW, JH, and XD conducted the visualization. YD wrote the manuscript. YD and HL contributed equally to this work.

ACKNOWLEDGMENTS

We thank all members of our study team for their wonderful cooperation and the National Cancer Database for their works.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.779185/full#supplementary-material

Supplementary Table 1 | Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Reporting Guideline Checklist.

Supplementary Table 2 | Univariate cox analysis of overall survival in patients with locally advanced breast cancer received preoperative radiation.

Supplementary Table 3 | Univariate cox analysis of overall survival in patients with locally advanced breast cancer received postoperative radiation.

REFERENCES

1. Li N, Deng Y, Zhou L, Tian T, Yang S, Wu Y, et al. Global Burden of Breast Cancer and Attributable Risk Factors in 195 Countries and Territories, From 1990 to 2017: Results From the Global Burden of Disease Study 2017. *J Hematol Oncol* (2019) 12 (1):140. doi: 10.1186/s13045-019-0828-0
2. Tryfonidis K, Senkus E, Cardoso MJ, Cardoso F. Management of Locally Advanced Breast Cancer-Perspectives and Future Directions. *Nat Rev Clin Oncol* (2015) 12(3):147–62. doi: 10.1038/nrclinonc.2015.13
3. Klein J, Tran W, Watkins E, Vesprini D, Wright FC, Look Hong NJ, et al. Locally Advanced Breast Cancer Treated With Neoadjuvant Chemotherapy and Adjuvant Radiotherapy: A Retrospective Cohort Analysis. *BMC Cancer* (2019) 19(1):306. doi: 10.1186/s12885-019-5499-2
4. Sikov WM. Locally Advanced Breast Cancer. *Curr Treat Options Oncol* (2000) 1(3):228–38. doi: 10.1007/s11864-000-0034-9
5. Veronesi U, Boyle P, Goldhirsch A, Orecchia R, Viale G. Breast Cancer. *Lancet* (2005) 365(9472):1727–41. doi: 10.1016/s0140-6736(05)66546-4
6. Rutqvist LE, Pettersson D, Johansson H. Adjuvant Radiation Therapy *Versus* Surgery Alone in Operable Breast Cancer: Long-Term Follow-Up of a Randomized Clinical Trial. *Radiother Oncol* (1993) 26(2):104–10. doi: 10.1016/0167-8140(93)90090-u
7. Riet FG, Fayard F, Arriagada R, Santos MA, Bourgier C, Ferchiou M, et al. Preoperative Radiotherapy in Breast Cancer Patients: 32 Years of Follow-Up. *Eur J Cancer* (2017) 76:45–51. doi: 10.1016/j.ejca.2017.01.022
8. Bollet MA, Belin L, Reyal F, Campana F, Dendale R, Kirova YM, et al. Preoperative Radio-Chemotherapy in Early Breast Cancer Patients: Long-Term Results of a Phase II Trial. *Radiother Oncol* (2012) 102(1):82–8. doi: 10.1016/j.radonc.2011.08.017
9. Adams S, Chakravarthy AB, Donach M, Spicer D, Lymberis S, Singh B, et al. Preoperative Concurrent Paclitaxel-Radiation in Locally Advanced Breast Cancer: Pathologic Response Correlates With Five-Year Overall Survival. *Breast Cancer Res Treat* (2010) 124(3):723–32. doi: 10.1007/s10549-010-1181-8
10. Corradini S, Krug D, Meattini I, Matuschek C, Bölke E, Francolini G, et al. Preoperative Radiotherapy: A Paradigm Shift in the Treatment of Breast Cancer? A Review of Literature. *Crit Rev Oncol Hematol* (2019) 141:102–111. doi: 10.1016/j.critrevonc.2019.06.003
11. Marijnen CA, Kapiteijn E, van de Velde CJ, Martijn H, Steup WH, Wiggers T, et al. Acute Side Effects and Complications After Short-Term Preoperative Radiotherapy Combined With Total Mesorectal Excision in Primary Rectal Cancer: Report of a Multicenter Randomized Trial. *J Clin Oncol* (2002) 20 (3):817–25. doi: 10.1200/jco.2002.20.3.817
12. Beskow C, Agren-Cronqvist AK, Granath F, Frankendal B, Lewensohn R. Pathologic Complete Remission After Preoperative Intracavitary Radiotherapy of Cervical Cancer Stage Ib and IIa Is a Strong Prognostic Factor for Long-Term Survival: Analysis of the Radiumhemmet Data 1989-1991. *Int J Gynecol Cancer* (2002) 12(2):158–70. doi: 10.1046/j.1525-1438.2002.01089.x
13. Paillocher N, Florczak AS, Richard M, Classe JM, Oger AS, Raro P, et al. Evaluation of Mastectomy With Immediate Autologous Latissimus Dorsi Breast Reconstruction Following Neoadjuvant Chemotherapy and Radiation Therapy: A Single Institution Study of 111 Cases of Invasive Breast Carcinoma. *Eur J Surg Oncol* (2016) 42(7):949–55. doi: 10.1016/j.ejso.2016.03.024
14. Mladenovic J, Susnjar S, Tanic M, Jankovic R, Karadzic K, Gavrilovic D, et al. Tumor Response and Patient Outcome After Preoperative Radiotherapy in Locally Advanced Non-Inflammatory Breast Cancer Patients. *J buon* (2017) 22(2):325–33.
15. Lerouge D, Touboul E, Lefranc JP, Genestie C, Moureau-Zabotto L, Blondon J. [Locally Advanced non Inflammatory Breast Cancer Treated by Combined Chemotherapy and Preoperative Irradiation: Updated Results in a Series of 120 Patients]. *Cancer Radiother* (2004) 8(3):155–67. doi: 10.1016/j.canrad.2004.01.001 Cancer du sein localement évolué non inflammatoire traité par association de chimiothérapie et de radiothérapie à dose préopératoire: réactualisation des résultats d'une série de 120 patientes. fre.
16. Lerouge D, Touboul E, Lefranc JP, Genestie C, Moureau-Zabotto L, Blondon J. Combined Chemotherapy and Preoperative Irradiation for Locally Advanced Noninflammatory Breast Cancer: Updated Results in a Series of 120 Patients. *Int J Radiat Oncol Biol Phys* (2004) 59(4):1062–73. doi: 10.1016/j.ijrobp.2003.12.034
17. Gerlach B, Audretsch W, Gogolin F, Königshausen T, Rohn R, Schmitt G, et al. Remission Rates in Breast Cancer Treated With Preoperative Chemotherapy and Radiotherapy. *Strahlenther Onkol* (2003) 179(5):306–11. doi: 10.1007/s00066-003-1019-y
18. Boffa DJ, Rosen JE, Mallin K, Loomis A, Gay G, Palis B, et al. Using the National Cancer Database for Outcomes Research: A Review. *JAMA Oncol* (2017) 13(12):1722–8. doi: 10.1001/jamaoncol.2016.6905
19. Duhamel A, Labreuche J, Gronnier C, Mariette C. Statistical Tools for Propensity Score Matching. *Ann Surg* (2017) 265(6):E79–e80. doi: 10.1097/sla.0000000000001312
20. Zhai Z, Zheng Y, Yao J, Liu Y, Ruan J, Deng Y, et al. Evaluation of Adjuvant Treatments for T1 N0 M0 Triple-Negative Breast Cancer. *JAMA Netw Open* (2020) 3(11):e2021881. doi: 10.1001/jamanetworkopen.2020.21881
21. Wang J, Shi M, Ling R, Xia Y, Luo S, Fu X, et al. Adjuvant Chemotherapy and Radiotherapy in Triple-Negative Breast Carcinoma: A Prospective Randomized Controlled Multi-Center Trial. *Radiother Oncol* (2011) 100 (2):200–4. doi: 10.1016/j.radonc.2011.07.007
22. Pondé NF, Zardavas D, Piccart M. Progress in Adjuvant Systemic Therapy for Breast Cancer. *Nat Rev Clin Oncol* (2019) 16(1):27–44. doi: 10.1038/s41571-018-0089-9
23. Kuerer HM, Newman LA, Smith TL, Ames FC, Hunt KK, Dhingra K, et al. Clinical Course of Breast Cancer Patients With Complete Pathologic Primary Tumor and Axillary Lymph Node Response to Doxorubicin-Based Neoadjuvant Chemotherapy. *J Clin Oncol* (1999) 17(2):460–9. doi: 10.1200/jco.1999.17.2.460
24. Matuschek C, Nestle-Kraemling C, Haussmann J, Bölke E, Wollandt S, Speer V, et al. Long-Term Cosmetic Outcome After Preoperative Radio-/Chemotherapy in Locally Advanced Breast Cancer Patients. *Strahlenther Onkol* (2019) 195(7):615–28. doi: 10.1007/s00066-019-01473-2 Langfristiges

kosmetisches Ergebnis nach neoadjuvanter Radio−/Chemotherapie bei lokal fortgeschrittenen Brustkrebspatientinnen.
25. Kunkler IH, Williams LJ, Jack WJ, Cameron DA, Dixon JM. Breast-Conserving Surgery With or Without Irradiation in Women Aged 65 Years or Older With Early Breast Cancer (PRIME II): A Randomised Controlled Trial. *Lancet Oncol* (2015) 16(3):266–73. doi: 10.1016/s1470-2045(14)71221-5
26. Lagendijk M, van Maaren MC, Saadatmand S, Strobbe LJA, Poortmans PMP, Koppert LB, et al. Breast Conserving Therapy and Mastectomy Revisited: Breast Cancer-Specific Survival and the Influence of Prognostic Factors in 129,692 Patients. *Int J Cancer* (2018) 142(1):165–75. doi: 10.1002/ijc.31034
27. Sharma RA, Plummer R, Stock JK, Greenhalgh TA, Ataman O, Kelly S, et al. Clinical Development of New Drug-Radiotherapy Combinations. *Nat Rev Clin Oncol* (2016) 13(10):627–42. doi: 10.1038/nrclinonc.2016.79
28. Ataman OU, Sambrook SJ, Wilks C, Lloyd A, Taylor AE, Wedge SR. The Clinical Development of Molecularly Targeted Agents in Combination With Radiation Therapy: A Pharmaceutical Perspective. *Int J Radiat Oncol Biol Phys* (2012) 84(4):e447–54. doi: 10.1016/j.ijrobp.2012.05.019
29. Zaidi M, Fu F, Cojocari D, McKee TD, Wouters BG. Quantitative Visualization of Hypoxia and Proliferation Gradients Within Histological Tissue Sections. *Front Bioeng Biotechnol* (2019) 7:397. doi: 10.3389/fbioe.2019.00397
30. Bauer JA, Chakravarthy AB, Rosenbluth JM, Mi D, Seeley EH, De Matos Granja-Ingram N, et al. Identification of Markers of Taxane Sensitivity Using Proteomic and Genomic Analyses of Breast Tumors From Patients Receiving Neoadjuvant Paclitaxel and Radiation. *Clin Cancer Res* (2010) 16(2):681–90. doi: 10.1158/1078-0432.Ccr-09-1091
31. Brunt AM, Wheatley D, Yarnold J, Somaiah N, Kelly S, Harnett A, et al. Acute Skin Toxicity Associated With a 1-Week Schedule of Whole Breast Radiotherapy Compared With a Standard 3-Week Regimen Delivered in the UK FAST-Forward Trial. *Radiother Oncol* (2016) 120(1):114–8. doi: 10.1016/j.radonc.2016.02.027
32. Matsen CB, Neumayer LA. Breast Cancer: A Review for the General Surgeon. *JAMA Surg* (2013) 148(10):971–9. doi: 10.1001/jamasurg.2013.3393
33. Engel J, Kerr J, Schlesinger-Raab A, Sauer H, Hölzel D. Quality of Life Following Breast-Conserving Therapy or Mastectomy: Results of a 5-Year Prospective Study. *Breast J* (2004) 10(3):223–31. doi: 10.1111/j.1075-122X.2004.21323.x
34. Chakravarthy AB, Kelley MC, McLaren B, Truica CI, Billheimer D, Mayer IA, et al. Neoadjuvant Concurrent Paclitaxel and Radiation in Stage II/III Breast Cancer. *Clin Cancer Res* (2006) 12(5):1570–6. doi: 10.1158/1078-0432.Ccr-05-2304

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer QT has declared a shared parent affiliation with the authors MW, SL, PX, and YW at the time of review.

frontiers in Oncology

ORIGINAL RESEARCH
published: 13 January 2022
doi: 10.3389/fonc.2021.755267

Knockdown of NAA25 Suppresses Breast Cancer Progression by Regulating Apoptosis and Cell Cycle

Jingkai Xu[1], *Zhi Li*[2], *Xianbo Zuo*[1], *Guozheng Li*[3,4], *Xuejun Zhang*[5], *Bo Zhang*[3,4*] *and Yong Cui*[1*]

[1] Department of Dermatology, China-Japan Friendship Hospital, Beijing, China, [2] Department of Dermatology, Jiangsu Province Hospital, Nanjing, China, [3] School of Life Sciences, Anhui Medical University, Hefei, China, [4] Department of Oncology, No. 2 Hospital, Anhui Medical University, Hefei, China, [5] Department of Dermatology, The First Affiliated Hospital of Anhui Medical University, Hefei, China

NAA25 gene variants were reported as risk factors for type 1 diabetes, rheumatoid arthritis and acute arterial stroke. But it's unknown whether it could contribute to breast cancer. We identified rs11066150 in *lncHSAT164*, which contributes to breast cancer, in our earlier genome-wide long non-coding RNA association study on Han Chinese women. However, rs11066150 A/G variant is also located in *NAA25* intron. Based on the public database, such as TCGA and Curtis dataset, *NAA25* gene is highly expressed in breast cancer tissues and this result has also been proved in our samples and cell lines through RT-qPCR and western blot analysis. To better understand the function of *NAA25* in breast cancer, we knocked down the expression of *NAA25* in breast cancer cell lines, FACS was used to detect cell apoptosis and cell cycle and colony formation assay was used to detect cell proliferation. We found that *NAA25*-deficient cells could increase cell apoptosis, delay G2/M phase cell and decrease cell clone formation. RNA sequencing was then applied to analyze the molecular profiles of *NAA25*–deficient cells, and compared to the control group, *NAA25* knockdown could activate apoptosis-related pathways, reduce the activation of tumor-associated signaling pathways and decrease immune response-associated pathways. Additionally, RT-qPCR was employed to validate these results. Taken together, our results revealed that NAA25 was highly expressed in breast cancer, and NAA25 knockdown might serve as a therapeutic target in breast cancer.

Keywords: breast cancer, NAA25, cell cycle, apoptosis, RNA sequencing

OPEN ACCESS

Edited by:
Dirk Geerts,
University of Amsterdam, Netherlands

Reviewed by:
Jianjun Chen,
Tongji University, China
Hou-Feng Zheng,
Westlake Institute for Advanced Study (WIAS), China
Yanhua Liang,
Southern Medical University, China

***Correspondence:**
Yong Cui
wuhucuiyong@vip.163.com
Bo Zhang
alvinbo@163.com

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: *07 September 2021*
Accepted: *21 December 2021*
Published: *13 January 2022*

Citation:
Xu J, Li Z, Zuo X, Li G, Zhang X, Zhang B and Cui Y (2022) Knockdown of NAA25 Suppresses Breast Cancer Progression by Regulating Apoptosis and Cell Cycle. Front. Oncol. 11:755267. doi: 10.3389/fonc.2021.755267

INTRODUCTION

Breast cancer is the most common and a leading cause of cancer-related deaths of women worldwide (1). And China is undergoing the cancer transition stage, with the occurrence of female breast cancer increasing rapidly (2, 3). With the development of sequencing technologies, a lot of breast cancer associated genes have been validated (4, 5). Our previous case-control genome wide lncRNA association study on Han Chinese women identified that SNP rs11066150 was

associated with breast cancer and *lncHSAT164* gene could contribute to breast cancer (6). And rs11066150 A/G was an intron variant in N-alpha-acetyltransferase 25 *(NAA25*) gene (also known as *MDM20*, *C12orf30* and *NAP1*). *NAA25* gene variants were reported to be associated with type 1 diabetes (T1D), rheumatoid arthritis, acute arterial stroke and dyslipidemia (7–10). However, the relationship between *NAA25* and breast cancer is still unknown.

NAA25 encodes the auxiliary subunit, which could then affect posttranslational modifications by forming N-terminal acetyltransferase B complex with catalytic subunit *NAA20* (11). In yeast, it can regulate actin remodeling, and stabilize actin cytoskeleton and mitochondrial targeting (12, 13). And *NAA25* knockdown can disrupt cell cycle and reduce cell growth (14). However, the physiological function and mechanism of *NAA25* in breast cancer remain unknown.

To explore the relationship between *NAA25* gene and breast cancer, we compared *NAA25* gene expression between normal tissues and breast cancer tissues in public databases, such as TCGA and Curtis dataset, and analyzed the relationship between *NAA25* gene expression and overall survival (OS) of patients. In addition, we tested *NAA25* gene expression in breast cancer tissues, para-carcinoma tissues, breast cancer cell lines and normal breast epithelial cell lines. Furthermore, we specifically knocked down *NAA25* gene expression in breast cancer cells and explored its influence on tumor cell proliferation, apoptosis and cell cycle. Finally, RNA-seq analysis was used to clarify the molecular profiles of *NAA25*-deficient cells.

MATERIALS AND METHODS

Subjects

In this study, four-pairs of breast cancer tissues and para-carcinoma tissues (all from Han Chinese women) were collected at the No.2 Hospital, Anhui Medical University. All cases were diagnosed with breast cancer by at least two pathologists. Para-carcinoma specimens were adipose/skin tissues, which were collected from breast cancer patients who underwent radical mastectomy. All tissue samples were stored in liquid nitrogen immediately after surgical resection. The information of breast cancer patients was provided in **Supplementary Table 1**.

Cell Culture

MCF10A, MCF7, T47D, and HEK293T cell lines were purchased from the Institute of Basic Medical Sciences of the Chinese Academy of Medical Sciences. MCF10A, a kind of normal human breast epithelial cell, was grown in DMEM/F12 (Gibco, Life, China) medium supplemented with 10% fetal bovine serum (FBS) (Gibco, Australia), 10 μg/ml insulin (Macklin, China), 20 ng/ml EGF (Peprotech, China), and 0.5 μg/ml hydrocortisone (Macklin, China). MCF7 and HEK293T cells were maintained in DMEM (Gibco, Life, USA) supplemented with 10% FBS (Gibco, Australia). T47D cells were maintained in RPMI-1640 medium (Gibco, Life, USA) supplemented with 10% FBS (Gibco, Australia). All medium were supplemented with 100 U/ml penicillin–streptomycin (Gibco, Life, China), and all cells were maintained at 37°C in a humidified atmosphere containing 5% CO_2 and confirmed to be mycoplasma free.

RNA Extraction and RT-qPCR Analysis

The total RNA from the cell lines, human breast cancer tissues and para-cancerous tissues used in this study was extracted with TRIzol reagent, and DNase I (Thermo Fisher, USA) was used to remove genomic DNA. First-strand cDNA was synthesized by using the SuperScript III Reverse Transcriptase Kit (Thermo Fisher, USA). Relative RNA levels determined by RT-qPCR were measured on a Rotor-Gene Q real-time PCR machine (Qiagen, Germany). GAPDH was employed as an internal control. The relative expression of RNAs was calculated using the $2^{-\Delta\Delta Ct}$ method. All primer sequences for RT-qPCR are listed in **Supplementary Table 2**.

Plasmid Construction, Transfection and Lentivirus Infection

Short hairpin RNAs (shRNAs) against NAA25 sh1 and sh2 were designed and synthesized by Taihe Biotechnology (Beijing, China) and cloned into the EGFP-Puro-pll3.7 plasmid. Based on the PSPAX2-PMD2G lentiviral system, a lentivirus was constructed according to the manufacturer's instructions. After lentivirus infection, 1 μg/ml puromycin (*In vivo*Gen, USA) was added for selection, and 48-72 hours later, the cells were harvested for further experiments. shRNA sequences are listed in **Supplementary Table 2**.

Western Blot Analysis

Tissues and cells were lysed in RIPA buffer (Beyotime, China). 40 μg of protein was used for SDS-PAGE gel electrophoresis (Bio-Rad) and transferred onto PVDF membranes (Millipore, China). Blocking was performed with 5% milk, and then the membranes were incubated with primary antibodies. Anti-NAA25 (1:1, 000 HPA039322, Sigma-Aldrich) or anti-actin (1:5000, A1978, Sigma-Aldrich) was added and incubated overnight at 4°C. After being washed, the membranes were incubated with secondary antibodies (peroxidase conjugated, suitable for each primary antibody) for 2 hours at room temperature. The signal was detected with a Bio-Rad ChemiDoc XRS + System after adding Super Signal West Pico chemiluminescence.

Colony Formation Assay

To analyze cell growth, colony formation assays were performed. 1×10^3 cells of T47D- and MCF7- Ctr, -sh1, -sh2 were seeded in a 6-well plate and incubated for 10 to 15 days at 37°C. Then, the cells were washed twice in PBS, fixed with 90% ethanol for 15 minutes and stained with 0.1% crystal violet for 20 minutes. Images of colonies were taken with a digital camera, and the number of colonies was analyzed by ImageJ v1.8.0 software.

Apoptosis Assay

For apoptosis analysis, target cells were transferred to a 15 ml centrifuge tube, and annexin V binding buffer was added. After

being centrifugated at 1,000 rpm for 5 min at 4°C, the cells were washed 3 times in PBS. Then, the cells were treated with 100 µl of binding buffer, 5 µl of Annexin V-APC and 1 µl of 100 µg/ml propidium iodide (PI) stain (Thermo Fisher, USA), and incubated in the dark for 25 min. Cell apoptosis was analyzed by flow cytometry (BD Biosciences).

Cell Cycle Assay

For cell cycle analysis, target cells were fixed with 75% ice-cold ethanol at 4°C overnight. Then, the cells were suspended in PBS supplemented with 100 mg/ml RNase A for 30 min at 37°C and then stained with 50 µg/ml PI (Thermo Fisher, USA) in the dark at room temperature for 15 min. Finally, a total of 20,000 cells were analyzed on a FACS Calibur flow cytometer equipped with Cell Quest software (BD Biosciences).

RNA-Seq Analysis

After *NAA25* knockdown in T47D cells, cells from the Ctr, sh1 and sh2 groups were harvested for RNA-seq analysis at Shanghai Majorbio Biopharm Technology Co. mRNAs were isolated from total RNA with the oligo (dT) method. The mRNAs were fragmented, and then first-strand cDNA and second-strand cDNA were synthesized. After being purified, cDNA fragments were linked to adapters. Then, cDNA fragments of suitable size were selected for PCR amplification. The sequencing platform used in this study was Illumina HiSeq, and the paired-end reads were 2×150 bp. TPM (Transcripts Per Million reads) was used to evaluate genes expression, transcript abundance was assessed with the DESeq2, and the significantly affected genes were determined by setting a fold change of ≥ 2. The differentially expressed gene (DEG), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, gene ontology (GO), GO term and gene set enrichment analysis (GSEA) described in this paper were performed on the free online platform Majorbio Cloud Platform (www.majorbio.com).

Statistical Analysis

All statistical analyses were performed using Graphpad Prism 8.0 statistical software (California, US). Experiment data are shown as the means ± SEM, and all experiments were conducted for at least three times. Significance was determined using the Student's *t*-test: N.S. $p > 0.05$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$ and $****p < 0.0001$.

RESULTS

rs11066150 Associated Gene NAA25 Highly Expressed in Breast Cancer

rs11066150 was reported in *lncHSAT164* (6), and it is also located within the fifth intron of *NAA25* gene (**Figure 1A**). Based on the eQTLGen database (https://www.eqtlgen.org/), we identified 4 cis-eQTL effects genes, *TMEM116*, *HECTD4*, *MAPKAPK5* and *NAA25*, to be associated with rs11066150 (**Supplementary Table 3**). And *TMEM116*, *HECTD4*, *MAPKAPK5* was reported to be associated with renal cell carcinoma, prostate cancer and colorectal cancer (15–17). However, *NAA25* gene has never been reported to be associated with cancers.

To explore the role of *NAA25* gene in breast cancer, we analyzed its expression in different public databases. According to TCGA and the Curtis, Finak breast and Richardson breast datasets (18–21), we found that *NAA25* was greatly up regulated in breast cancer tissues in comparison with normal breast tissues (**Figures 1B, C** and **Supplementary Figures 1A, B**). Furthermore, high mRNA levels of *NAA25* showed marginal associations with poor OS in the Curtis database ($p = 0.013$) (**Figure 1D**). Additionally, we explored the expression of *NAA25* gene in breast cancer tissues and para-cancerous tissues. RT-qPCR and western blot analyses were performed in four-pairs of tissues, and results revealed that *NAA25* was highly expressed in cancer tissues (**Figure 1E**). We also monitored *NAA25* expression in normal breast epithelial cell line MCF10A, and breast cancer cell lines MCF7 and T47D. Compared to MCF10A, *NAA25* was highly expressed in T47D cells (**Figure 2A**). Together, our analyses reveal a previously unknown role of *NAA25* in breast cancer, and highly expressed *NAA25* might influence the progress of breast cancer.

NAA25 Knockdown Inducing Apoptosis, G2/M Arrest and Suppressing Cell Proliferation

To investigate the physiological roles of *NAA25* gene in breast cancer, two shRNA targets were designed to knockdown *NAA25* gene in breast cancer cell lines, and the mRNA expression and protein expression of *NAA25* were both significantly diminished (**Figure 2B** and **Supplementary Figure 1C**). Apoptosis is a key cellular process in breast cancer. We measured the effect of *NAA25* on apoptosis and cell cycle by FACE analysis. Compared to the Ctr group, the number of apoptotic cells was relatively larger in the shRNA groups ($p < 0.01$), as shown in (**Figure 2C**), and more cells were arrested in the G2/M phase ($p < 0.05$), as shown in (**Figure 2D**).

To further investigate whether NAA25 knockdown could influence tumor growth, colony formation assays were applied in this study, which illustrated that clonogenic survival significantly decreased following *NAA25* knockdown in T47D cell line (**Figures 2E, F**). And similar results were investigated in NAA25-deficient MCF7 cells. Hence, based on these results, we conclude that *NAA25* is highly expressed in breast cancer and may lead to poor OS in patients by regulating tumor cell apoptosis and cell cycle.

RNA Sequencing Characterizing the Molecular Profile of NAA25-Deficient Breast Cancer Cells

To investigate the importance of *NAA25* gene in breast cancer, RNA-seq analysis was applied after *NAA25* knockdown in the T47D cell line. Pearson's correlation analysis (PCA) was performed to cluster all samples (**Supplementary Figure 2A**). Based on the gene expression matrix, the Venn diagram was used to analyze the co-expressed and specifically expressed genes or transcripts among the Ctr group and shRNA groups

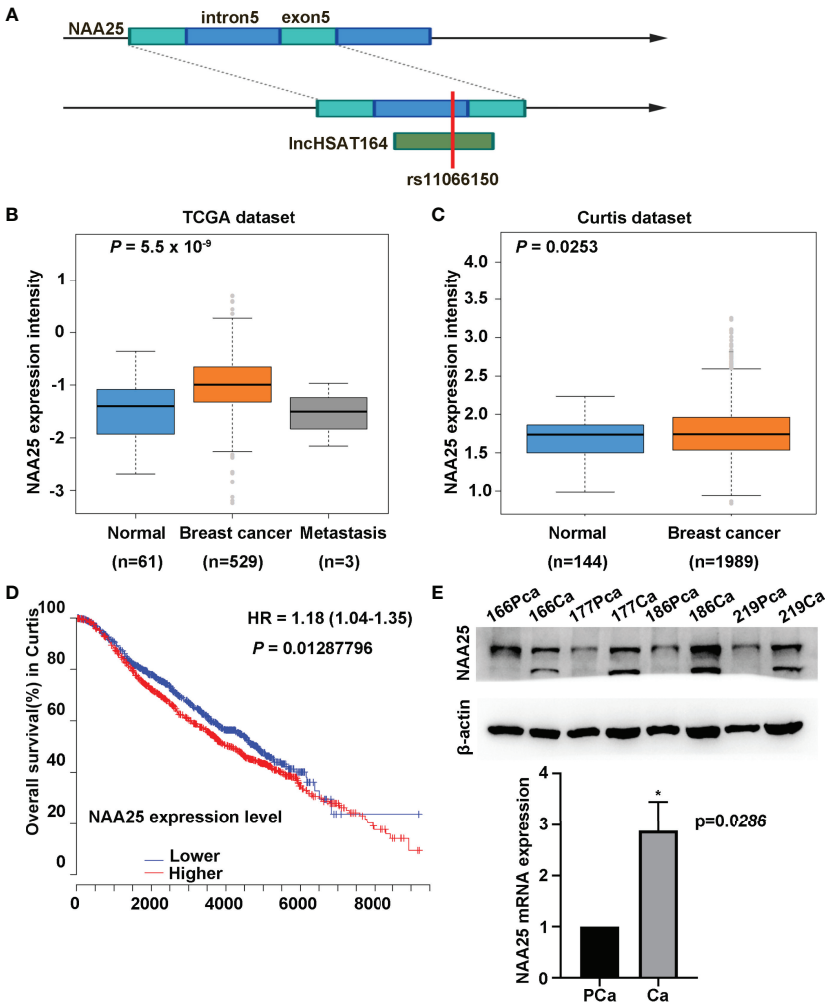


FIGURE 1 | Characterization of NAA25 gene in breast cancer. **(A)** rs11066150 variant schematic diagram in lncHSAT164 and NAA25. **(B, C)** NAA25 gene was highly expressed in breast cancer tissues compared to the controls in TCGA dataset and Curtis dataset. **(D)** OS analysis of patients with high and low NAA25 expression. The *p* value was calculated using Mann-Whitney *U* tests. **(E)** Western-blot and RT-qPCR to analyze NAA25 expression in breast cancer tissues and para-carcinoma tissues, NAA25 was highly expressed in breast cancer tissues. The two bands are all *NAA25*. *$p < 0.05$.

(**Figure 3A**). Furthermore, differentially expressed gene (DEG) analysis was conducted to compare the Ctr group and the sh1 and sh2 groups respectively, and 119 DEGs were identified (**Figures 3B, C**, **Supplementary Figure 2B**). All DEGs were presented in **Supplementary Table 4**.

Furthermore, KEGG enrichment analyses were performed among the 119 DEGs, and most of them were related to infections, immune responses, cancers and immune diseases (**Supplementary Figure 2C**). GO term analysis was performed to *NAA25*-deficient cells, and the results showed that many genes were associated with infection and immunity (**Figure 3D** and **Supplementary Figure 2D**).

To assess the molecular pathways involved in *NAA25*-deficient T47D cells, we performed gene set enrichment analysis (GSEA). And *NAA25* knockdown could increase apoptosis associated pathways, and reduce tumor associated pathways, like MYC, HIF1A, ERB2, MEK and TNF (**Figure 3E** and **Supplementary Figure 2E**). In addition, immune response associated pathways like IL4, TNF and LTE2 were reduced. Finally, RT-qPCR analysis was used to verify RNA-seq data (**Figure 3F**). *IFIT2*, *IFIT3*, *IFIT27*, *IFITM1*, *NDRG1*, *PFKFB4*, *ZNF395*, *IFI6*, *FUT11* and *OAS2* mRNA expression was upregulated after *NAA25* knockdown, and *HSPH1* gene expression was down regulated, consistent with the RNA-seq results.

DISCUSSION

A large number of breast cancer associated susceptibility SNPs and genes were identified and reported as a molecular marker in tumor incidence, metastasis, prognosis and treatment. Previously, we performed a genome-wide lncRNA association study in Han Chinese women and identified two new susceptibility SNPs, rs11066150 and rs12537 (6) (22).

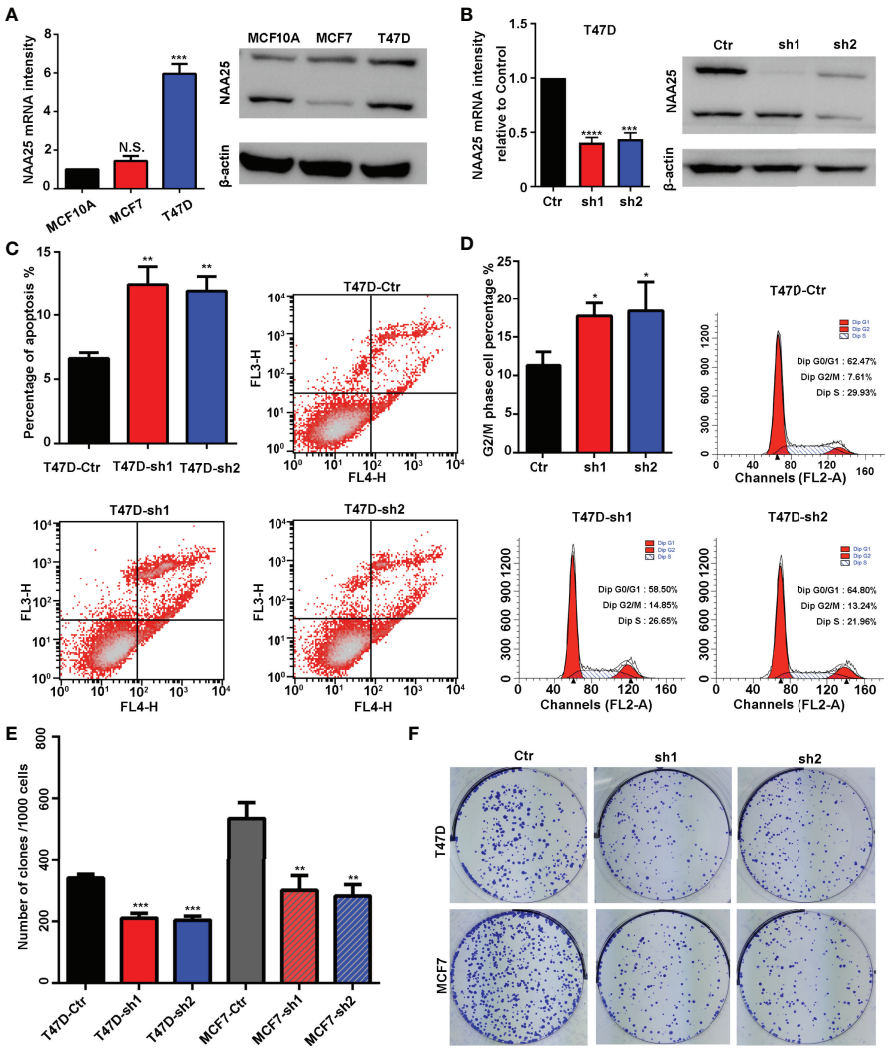


FIGURE 2 | NAA25 gene influences cell apoptosis and the cell cycle in breast cancer. **(A)** RT-qPCR and western blot analysis of NAA25 gene expression in breast cancer cell lines (The two bands are all *NAA25*.). **(B)** RT-qPCR and western blot analysis in the NAA25-deficient T47D cell line. **(C)** Cell apoptosis in the NAA25-deficient T47D cells. Compared to the Ctr group, NAA25 knockdown could increase cell apoptosis. **(D)** Cell cycle analysis of the NAA25-deficient T47D cells. Compared to the Ctr group, NAA25 knockdown induced G2/M cell cycle arrest. **(E, F)** Downregulated NAA25 reduced the clonogenic potential of breast cancer cells. N.S. $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$ and ****$p < 0.0001$.

rs11066150 variant had no relationship with the clinical characteristics of breast cancer like family history, menopausal status, and molecular subtypes (22). However, rs11066150 associated lncRNA, *lncHSAT164*, was highly expressed in breast cancer, and overexpressed *lncHSAT164* could promote colony formation and down-expressed *lncHSAT164* could promote cell apoptosis and regulate cell cycle (6). In this study, we reported rs11066150 as an intron variant SNP in *NAA25* gene. And *NAA25* gene is highly expressed in breast cancer tissues relative to normal tissues, while high *NAA25* expression is correlated with poor OS. And *NAA25* knockdown could induce cell apoptosis, delay G2/M phase cell and decrease cell clone formation. *NAA25* was reported to be associated with T1D (7, 23), arthritis (8, 24) and virus infection (25). However, *NAA25* gene was reported as a proto-oncogene in breast cancer for the first time, and more research is needed in the future to characterize the impact of rs11066150 A/G variant on breast cancer, and the relationship between *lncHSAT164* and *NAA25* gene also needs further study.

RNA-seq is a ubiquitous tool in molecular biology that is shaping nearly every aspect of our understanding of genomic function (26). The molecular features of *NAA25*-deficient T47D cell lines were analyzed by RNA-seq in this work, and analysis results indicated that many infection and immune associated genes were highly expressed, which suggests that immune therapy may be an effective approach in treating *NAA25*-overexpressed breast cancer.

Highly expressed *IFIT2* and *NDRG1* could reduce tumor migration and metastasis (27–30). And *HSPH1* was highly

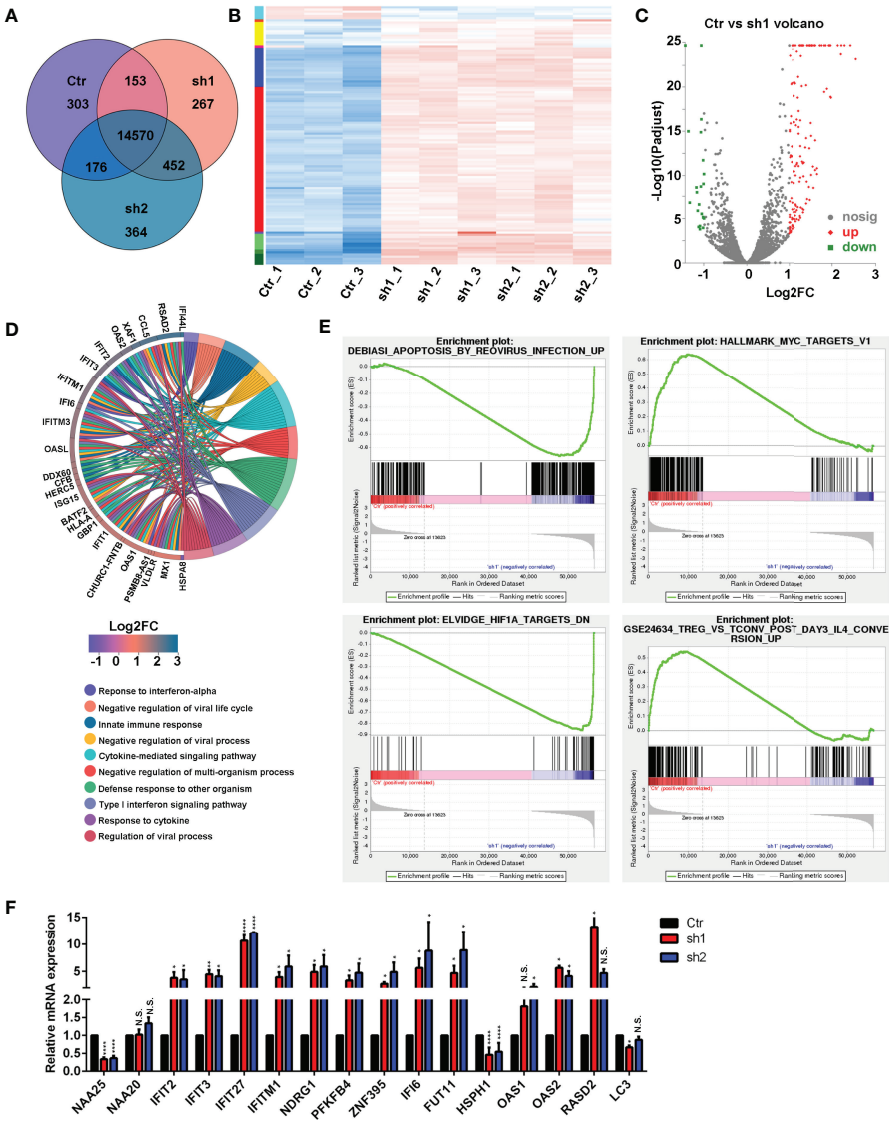


FIGURE 3 | RNA-seq analysis in the NAA25-deficient T47D cells. **(A)** Venn diagram analysis of gene or transcript expression among the Ctr group and shRNA groups. **(B)** Differentially expressed gene (DEG) heatmap analysis. Blue indicates downregulated genes. Red indicates upregulated genes. **(C)** Volcano plot showing the DEG in the Ctr group and the sh1 group. **(D)** GO term analysis between the Ctr group and the sh1 group. **(E)** Gene set enrichment analysis (GSEA) to analyze DEG between the Ctr group and the sh1 group. **(F)** RT-qPCR analysis to validate DEGs after NAA25 knockdown in the T47D cells. The data shown here are representative of at least 3 independent experiments. N.S. $p > 0.05$; *$p < 0.05$; **$p < 0.01$ and ****$p < 0.0001$.

expressed in different tumors, such as colorectal cancer, B-cell lymphoma, melanoma and esophageal squamous cell carcinoma (31–34), while *NAA25* knockdown could upregulate *IFIT2* and *NDRG1* expression and downregulate *HSPH1* expression (**Figure 3F**). These findings suggest that *NAA25* knockdown may also play a positive role in treating other cancers. As an important accessory subunit of the NatB enzymatic complex, *NAA25* could work with the *NAA20* catalytic subunit to promote enzymatic activity (26, 35), and *NAA25* knockdown did not reduce *NAA20* expression (12). It's also verified in the current study.

In conclusion, in this study, we reported *NAA25* as a candidate gene of rs11066150, which was highly expressed in breast cancer, and highly expressed *NAA25* could reduce patient's OS. In addition, *NAA25* knockdown could induce cell apoptosis, delay G2/M phase cell and decrease cell clone formation. RNA-seq analysis was also applied to clarify the molecular profiling of *NAA25*-deficient cells, and *NAA25* knockdown repressed tumor- and immune response-associated pathways. This study is among the first attempts to clarify the function of *NAA25* in breast cancer, and these results have elucidated the mechanism of *NAA25* in breast cancer and

suggests that *NAA25* may serve as a potential therapeutic target of breast cancer.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. RNA-seq data presented in the study are deposited in the SRA repository, accession number PRJNA752396. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

This study was approved by the Ethics Committee of Anhui Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JX, YC, and BZ conceived of the idea. JX, ZL, and GL performed the experiments. JX, XBZ, and XJZ analyzed the data. JX drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the scientific and technological innovation leading talents of "Ten Thousand Talents Program" (2018-WRJH-1), the National Natural Science Foundation of China (81872516), the discipline construction project of Peking Union Medical College (xhxk201903) and 2020 medical service and support capability upgrade project (2020-QTL-008).

ACKNOWLEDGMENTS

The authors would like to thank all the participating patients and healthy controls, as well as all the doctors and nurses who have contributed to this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.755267/full#supplementary-material

Supplementary Figure 1 | NAA25 gene expression in public databases. **(A, B)** Compared to normal tissues NAA25 gene was highly expressed in breast cancer tissues in Finak breast and Richardson breast databases. **(C)** RT-qPCR and western blot analysis in the NAA25-deficient MCF7 cell line. The two bands are all *NAA25*. $p < 0.05$; $^{**}p < 0.001$.

Supplementary Figure 2 | RNA-seq analysis in the NAA25-deficient T47D cells. **(A)** Pearson's correlation analysis (PCA) clarified the similarity between RNA-seq samples. **(B)** Volcano plot showing the DEG in the Ctr group and the sh2 group. **(C)** GO term analysis between the Ctr group and the sh1 group. **(D)** GO analysis between the Ctr group and the sh1 group. **(E)** Gene set enrichment analysis (GSEA) to analyze the DEG between the Ctr group and the sh2 group.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, 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(3):209–49. doi: 10.3322/caac.21660
2. Feng RM, Zong YN, Cao SM, Xu RH. Current Cancer Situation in China: Good or Bad News From the 2018 Global Cancer Statistics? *Cancer Commun (Lond Engl)* (2019) 39(1):22. doi: 10.1186/s40880-019-0368-6
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer Statistics in China, 2015. *CA Cancer J Clin* (2016) 66(2):115–32. doi: 10.3322/caac.21338
4. Zhang B, Chen MY, Shen YJ, Zhuo XB, Gao P, Zhou FS, et al. A Large-Scale, Exome-Wide Association Study of Han Chinese Women Identifies Three Novel Loci Predisposing to Breast Cancer. *Cancer Res* (2018) 78(11):3087–97. doi: 10.1158/0008-5472.CAN-17-1721
5. Milne RL, Kuchenbaecker KB, Michailidou K, Beesley J, Kar S, Lindstrom S, et al. Identification of Ten Variants Associated With Risk of Estrogen-Receptor-Negative Breast Cancer. *Nat Genet* (2017) 49(12):1767–78. doi: 10.1038/ng.3785
6. Xu JK, Li GZ, Li Z, Li WJ, Chen RS, Zhang B, et al. Genome-Wide Long non-Coding RNA Association Study on Han Chinese Women Identifies Lnchsat164 as a Novel Susceptibility Gene for Breast Cancer. *Chin Med J (Engl)* (2021) 134(10):1138–45. doi: 10.1097/CM9.0000000000001429
7. Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, et al. Robust Associations of Four New Chromosome Regions From Genome-Wide Analyses of Type 1 Diabetes. *Nat Genet* (2007) 39(7):857–64. doi: 10.1038/ng2068
8. Prahalad S, Hansen S, Whiting A, Guthery SL, Clifford B, McNally B, et al. Variants in TNFAIP3, STAT4, and C12orf30 Loci Associated With Multiple Autoimmune Diseases Are Also Associated With Juvenile Idiopathic Arthritis. *Arthritis Rheum* (2009) 60(7):2124–30. doi: 10.1002/art.24618
9. Bozpolat A, Unal E, Topaloglu T, Taheri S, Bayram AK, Ozcan A, et al. The Relationship Between the Prognosis of Children With Acute Arterial Stroke and Polymorphisms of CDKN2B, HDAC9, NINJ2, NAA25 Genes. *J Thromb Thrombolysis* (2019) 47(4):578–84. doi: 10.1007/s11239-018-01802-9
10. Yasukochi Y, Sakuma J, Takeuchi I, Kato K, Oguri M, Fujimaki T, et al. Identification of Six Novel Susceptibility Loci for Dyslipidemia Using Longitudinal Exome-Wide Association Studies in a Japanese Population. *Genomics* (2019) 111(4):520–33. doi: 10.1016/j.ygeno.2018.05.015
11. Van Damme P, Lasa M, Polevoda B, Gazquez C, Elosegui-Artola A, Kim DS, et al. N-Terminal Acetylome Analyses and Functional Insights of the N-Terminal Acetyltransferase NatB. *Proc Natl Acad Sci USA* (2012) 109(31):12449–54. doi: 10.1073/pnas.1210303109
12. Yasuda K, Takahashi M, Mori N. Mdm20 Modulates Actin Remodeling Through the Mtorc2 Pathway *via* Its Effect on Rictor Expression. *PloS One* (2015) 10(11):e0142943. doi: 10.1371/journal.pone.0142943
13. Alves S, Neiri L, Chaves SR, Vieira S, Trindade D, Manon S, et al. N-Terminal Acetylation Modulates Bax Targeting to Mitochondria. *Int J Biochem Cell Biol* (2018) 95:35–42. doi: 10.1016/j.biocel.2017.12.004
14. Starheim KK, Arnesen T, Gromyko D, Ryningen A, Varhaug JE, Lillehaug JR. Identification of the Human N(alpha)-Acetyltransferase Complex B (Hnatb): A Complex Important for Cell-Cycle Progression. *Biochem J* (2008) 415(2):325–31. doi: 10.1042/BJ20080658
15. Wrzesinski T, Szelag M, Cieslikowski WA, Ida A, Giles R, Zodro E, et al. Expression of Pre-Selected TMEMs With Predicted ER Localization as

Potential Classifiers of ccRCC Tumors. *BMC Cancer* (2015) 15:518. doi: 10.1186/s12885-015-1530-4
16. Vatapalli R, Sagar V, Rodriguez Y, Zhao JC, Unno K, Pamarthy S, et al. Histone Methyltransferase DOT1L Coordinates AR and MYC Stability in Prostate Cancer. *Nat Commun* (2020) 11(1):4153. doi: 10.1038/s41467-020-18013-7
17. Ji H, Hui B, Wang J, Zhu Y, Tang L, Peng P, et al. Long Noncoding RNA MAPKAPK5-AS1 Promotes Colorectal Cancer Proliferation by Partly Silencing P21 Expression. *Cancer Sci* (2019) 110(1):72–85. doi: 10.1111/cas.13838
18. Cancer Genome Atlas Network. Comprehensive Molecular Portraits of Human Breast Tumours. *Nature* (2012) 490(7418):61–70. doi: 10.1038/nature11412
19. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The Genomic and Transcriptomic Architecture of 2,000 Breast Tumours Reveals Novel Subgroups. *Nature* (2012) 486(7403):346–52. doi: 10.1038/nature10983
20. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal Gene Expression Predicts Clinical Outcome in Breast Cancer. *Nat Med* (2008) 14(5):518–27. doi: 10.1038/nm1764
21. Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene Expression Profiling of the Tumor Microenvironment During Breast Cancer Progression. *Breast Cancer Res* (2009) 11(1):R7. doi: 10.1186/bcr2222
22. Xu J, Li G, Chen M, Li W, Wu Y, Zhang X, et al. Rs12537 Is a Novel Susceptibility SNP Associated With Estrogen Receptor Positive Breast Cancer in Chinese Han Population. *Front Med* (2021) 8:708644. doi: 10.3389/fmed.2021.708644
23. Douroudis K, Kisand K, Nemvalts V, Rajasalu T, Uibo R. Allelic Variants in the PHTF1-PTPN22, C12orf30 and CD226 Regions as Candidate Susceptibility Factors for the Type 1 Diabetes in the Estonian Population. *BMC Med Genet* (2010) 11:11. doi: 10.1186/1471-2350-11-11
24. Lopez Herraez D, Martinez-Bueno M, Riba L, Garcia de la Torre I, Sacnun M, Goni M, et al. Rheumatoid Arthritis in Latin Americans Enriched for Amerindian Ancestry Is Associated With Loci in Chromosomes 1, 12, and 13, and the HLA Class II Region. *Arthritis Rheum* (2013) 65(6):1457–67. doi: 10.1002/art.37923
25. Oishi K, Yamayoshi S, Kozuka-Hata H, Oyama M, Kawaoka Y. N-Terminal Acetylation by NatB Is Required for the Shutoff Activity of Influenza A Virus PA-X. *Cell Rep* (2018) 24(4):851–60. doi: 10.1016/j.celrep.2018.06.078
26. Stark R, Grzelak M, Hadfield J. RNA Sequencing: The Teenage Years. *Nat Rev Genet* (2019) 20(11):631–56. doi: 10.1038/s41576-019-0150-2
27. Lai KC, Liu CJ, Chang KW, Lee TC. Depleting IFIT2 Mediates Atypical PKC Signaling to Enhance the Migration and Metastatic Activity of Oral Squamous Cell Carcinoma Cells. *Oncogene* (2013) 32(32):3686–97. doi: 10.1038/onc.2012.384
28. Koh SY, Moon JY, Unno T, Cho SK. Baicalein Suppresses Stem Cell-Like Characteristics in Radio- and Chemoresistant MDA-MB-231 Human Breast Cancer Cells Through Up-Regulation of IFIT2. *Nutrients* (2019) 11(3):624. doi: 10.3390/nu11030624
29. Godbole M, Togar T, Patel K, Dharavath B, Yadav N, Janjuha S, et al. Up-Regulation of the Kinase Gene SGK1 by Progesterone Activates the AP-1-NDRG1 Axis in Both PR-Positive and -Negative Breast Cancer Cells. *J Biol Chem* (2018) 293(50):19263–76. doi: 10.1074/jbc.RA118.002894
30. Yang Y, Liu Y, Guo R, Fu Y, Zhang Z, Zhang P, et al. The Novel Dithiocarbamate, DpdtC Suppresses HER2-Overexpressed Cancer Cells by Up-Regulating NDRG1 *via* Inactivation of HER2-ERK 1/2 Signaling. *Sci Rep* (2018) 8(1):3398. doi: 10.1038/s41598-018-21768-1
31. Buhard O, Lagrange A, Guilloux A, Colas C, Chouchene M, Wanherdrick K, et al. HSP110 T17 Simplifies and Improves the Microsatellite Instability Testing in Patients With Colorectal Cancer. *J Med Genet* (2016) 53(6):377–84. doi: 10.1136/jmedgenet-2015-103518
32. Boudesco C, Verhoeyen E, Martin L, Chassagne-Clement C, Salmi L, Mhaidly R, et al. HSP110 Sustains Chronic NF-kappaB Signaling in Activated B-Cell Diffuse Large B-Cell Lymphoma Through MyD88 Stabilization. *Blood* (2018) 132(5):510–20. doi: 10.1182/blood-2017-12-819706
33. Park HS, Park CH, Choi BR, Lim MS, Heo SH, Kim CH, et al. Expression of Heat Shock Protein 105 and 70 in Malignant Melanoma and Benign Melanocytic Nevi. *J Cutan Pathol* (2009) 36(5):511–6. doi: 10.1111/j.1600-0560.2008.01085.x
34. Gao H, Zheng Z, Mao Y, Wang W, Qiao Y, Zhou L, et al. Identification of Tumor Antigens That Elicit a Humoral Immune Response in the Sera of Chinese Esophageal Squamous Cell Carcinoma Patients by Modified Serological Proteome Analysis. *Cancer Lett* (2014) 344(1):54–61. doi: 10.1016/j.canlet.2013.10.007
35. Polevoda B, Sherman F. N-Terminal Acetyltransferases and Sequence Requirements for N-Terminal Acetylation of Eukaryotic Proteins. *J Mol Biol* (2003) 325(4):595–622. doi: 10.1016/s0022-2836(02)01269-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer, H-FZ, declared a past co-authorship with one of the authors, XJZ, to the handling editor.

frontiers in Oncology

ORIGINAL RESEARCH
published: 10 February 2022
doi: 10.3389/fonc.2022.725318

Image Analysis of Circulating Tumor Cells and Leukocytes Predicts Survival and Metastatic Pattern in Breast Cancer Patients

Giacomo Da Col[1†], Fabio Del Ben[2†], Michela Bulfoni[3], Matteo Turetta[4], Lorenzo Gerratana[2,5], Serena Bertozzi[6], Antonio Paolo Beltrami[2] and Daniela Cesselli[2,3]*

[1] Scuola Internazionale Superiore di Studi Avanzati, Trieste, Italy, [2] Department of Medicine, University of Udine, Udine, Italy, [3] Institute of Pathology, University Hospital of Udine (ASUFC), Udine, Italy, [4] Immunopathology and Cancer Biomarkers, Department of Translational Research, Centro di Riferimento Oncologico di Aviano, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Aviano, Italy, [5] Department of Medical Oncology, Centro di Riferimento Oncologico di Aviano, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Aviano, Italy, [6] Department of Surgery, AOU "S. Maria della Misericordia", Udine, Italy

OPEN ACCESS

Edited by:
Bruno M. Simões, The University of Manchester, United Kingdom

Reviewed by:
Sandra Casimiro, Universidade de Lisboa, Portugal
Hans Neubauer, University of Dusseldorf Medical School, Germany

***Correspondence:**
Fabio Del Ben
delben.fabio@spes.uniud.it;
fabio.delben@cro.it

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: *15 June 2021*
Accepted: *14 January 2022*
Published: *10 February 2022*

Citation:
Da Col G, Del Ben F, Bulfoni M, Turetta M, Gerratana L, Bertozzi S, Beltrami AP and Cesselli D (2022) Image Analysis of Circulating Tumor Cells and Leukocytes Predicts Survival and Metastatic Pattern in Breast Cancer Patients. Front. Oncol. 12:725318. doi: 10.3389/fonc.2022.725318

Background: The purpose of the present work was to test whether quantitative image analysis of circulating cells can provide useful clinical information targeting bone metastasis (BM) and overall survival (OS >30 months) in metastatic breast cancer (MBC).

Methods: Starting from cell images of epithelial circulating tumor cells (eCTC) and leukocytes (CD45pos) obtained with DEPArray, we identified the most significant features and applied single-variable and multi-variable methods, screening all combinations of four machine-learning approaches (Naïve Bayes, Logistic regression, Decision Trees, Random Forest).

Results: Best predictive features were circularity (OS) and diameter (BM), in both eCTC and CD45pos. Median difference in OS was 15 vs. 43 (months), p = 0.03 for eCTC and 19 vs. 36, p = 0.16 for CD45pos. Prediction for BM showed low accuracy (64%, 53%) but strong positive predictive value PPV (79%, 91%) for eCTC and CD45, respectively. Best machine learning model was Naïve Bayes, showing 46 vs 11 (months), p <0.0001 for eCTC; 12.5 vs. 45, p = 0.0004 for CD45pos and 11 vs. 45, p = 0.0003 for eCTC + CD45pos. BM prediction reached 91% accuracy with eCTC, 84% with CD45pos and 91% with combined model.

Conclusions: Quantitative image analysis and machine learning models were effective methods to predict survival and metastatic pattern, with both eCTC and CD45pos containing significant and complementary information.

Keywords: liquid biopsy, circulating tumor cells, image analysis, machine learning, data science

Abbreviations: BM, bone metastasis; CD45pos, leukocytes; ctDNA, circulating tumor DNA; CTC, circulating tumor cells; eCTC, epithelial circulating tumor cells; ECOG PS, Eastern Cooperative Oncology Group Peformance Status; EM-CTC, epithelial–mesenchymal CTC; MBC, metastatic breast cancer; MES, mesenchymal phenotype; NEG, CD45 negative; OS, overall survival; PPV, predictive value.

BACKGROUND

Breast cancer remains the most diagnosed tumor in the female population worldwide (1, 2). Cancer-related deaths are associated with the metastatic spread to various organs, mainly liver, bones, lungs and brain; along cancer evolution, the metastatic disease expresses the most complex picture of genetic modifications, often expressed by therapy resistance (3–8). Current methods for the detection of tumor progression are suffering from limited sensitivity, thus the development of accurate, sensitive and minimally invasive diagnostic tests is a hot topic in the clinical management of patients (9). Liquid biopsy, by the analysis of circulating tumor cells (CTC), tumor DNA (ctDNA) and exosomes, represents one of the most promising approaches to provide a complete and real-time overview of tumor evolution (10–12). In particular, the identification and characterization of CTC provide researchers with a goldmine of information that goes beyond mere DNA mutations. Epigenetics, transcriptomics, and phenotypical aspects of cancer can be probed exclusively on CTC. We focused on image analysis of immunostained whole cells, thus providing morphological and phenotypical information.

In our laboratory, we optimized a workflow to identify, count and sort viable CTC, immune-stained by an antibody cocktail recognizing CD45, epithelial and mesenchymal markers and analyzed by the DEPArray system (Menarini-Silicon Biosystems) (13). In metastatic breast cancer (MBC) patients, 4 classes of circulating cells have been described: epithelial CTC (eCTC), epithelial–mesenchymal CTC (EM-CTC), circulating cells with mesenchymal phenotype (MES), and circulating cells negative for epithelial, mesenchymal and for the CD45 pan-leukocyte markers (NEG) (13, 14). We limited the study to eCTC since their prognostic role has been widely demonstrated in breast cancer, while it is much less explored for mesenchymal CTC (15–19). Additionally, our preliminary data on the genomic profile of single CTC showed that while eCTC are a homogeneous population containing high fraction of tumor cells, mesenchymal cells represent a mix of cancer cells and normal stromal cells, constituting a significant risk of spurious results (13). Previous studies have shown that the number and phenotype of CTC represents a prognostic factor in patients with MBC (13, 18, 20, 21). However, these studies were based on image qualitative data only (presence/absence of known markers), which are used to classify cells phenotypically. No quantitative data from cell images were extracted or analyzed. The Kelley group obtained semi-quantitative information on the expression of known markers by means of magnetic gradients, and demonstrated that semi-quantitative information are valuable (22). However, to the best of our knowledge, there is no prior work considering quantitative data that can be obtained by CTC images, either morphological or fluorescence intensity of known markers, and correlating them to clinical outcomes.

The aim of the study was to evaluate whether quantitative analysis of images of CTC can provide useful information in terms of both overall survival (OS) and presence of bone metastases (BM).

Machine learning is a branch of artificial intelligence that aims to extrapolate relevant information from available data thus creating a model able to infer conclusions on future data. Machine learning has a long history of successful applications in all sorts of fields, but only recently has it received a lot of attention, mainly thanks to the neural network algorithm. Albeit the notoriety, neural networks need huge amounts of data (in the order of tens of thousands) to perform effectively, while having significant risk of losing generalization by overfitting training set when working with smaller datasets. In this study, we concentrate on algorithms with demonstrated capability of effectiveness even with small datasets; those algorithms have the advantage of being transparent with respect to the analyzed features, allowing insights into the model (23, 24).

As an additional aim, we evaluated whether the images of white blood cells contained information on OS and BM. It is in fact increasingly recognized that the immune system represents a central player in tumor occurrence, development and progression (25, 26). Recent studies illustrated that the "immunome" is generally dysfunctional in MBC patients. In particular, peripheral blood lymphocyte count is generally decreased and lymphocyte subpopulations are altered (27). Also, the cytokine signaling responsiveness of T cells is dysregulated (28). The immune status of cancer patients seems to predict response to therapy and prognosis in both localized and metastatic settings and correlates with clinical-pathological features (29–31). For these reasons, tumor-induced systemic immune changes are used as relevant biomarkers to better understand cancer evolution in women with MBC, and we hypothesized that white blood cells collected were worth to be investigated.

Thus, we focused on both the eCTC and leukocytes, to test the hypothesis whether the images of these cells can provide clinical information in MBC.

METHODS

Patients' Recruitment

The clinical study, approved by the Regional Ethics Committee (Ceur, N.152/2011/Sper and N.178/2014 Em), is a prospective observational study, carried out in collaboration between the Pathology Institute and the Oncology Department of Udine (University of Udine, Udine Academic Hospital). The criteria used for the recruitment and selection of patients were: age ≥18 years; measurable metastatic breast tumor; start of a new line of systemic therapy; Eastern Cooperative Oncology Group Peformance Status (ECOG PS) between 0 and 2; Availability of a histological sample of the primary tumor. In particular, 45 of 100 patients recruited in the period between November 2013 and December 2019 were eligible, for this study, since the others had no eCTC or were collected at a different timepoint.

Sample Processing and Staining

Approximately 7.5 ml of peripheral blood samples of the patients were processed for the isolation and characterization of CTC by

DEPArray technology. After a hypotonic red blood cell lysis (Miltenyi Biotec), the sample was enriched by an immunomagnetic depletion of the $CD45^+$ and $CD325a^+$ (Miltenyi Biotec) fraction of the blood, according to manufacturer's instructions. After incubation for 20 min at 4°C, the sample was depleted into an LD column (Miltenyi Biotec), lodged in the appropriate MidiMACS (Miltenyi Biotec) separator. The $CD45^-$ fraction, including CTC, was collected, loaded in a cartridge, and analyzed by DEPArray®. CTC were characterized alive by an antibody cocktail recognizing epithelial biomarkers in the FITC channel (EpCAM, E-Cadherin), mesenchymal markers in the PE channel (CD44, CD146, N-Cadherin) and the pan-leukocytes marker CD45 in the APC one. Nuclei were stained with HOECHST 33342 (Thermofisher Scientific). Immunostaining procedure is described in detail in the following article (13).

DEPArray Analysis and Data Selection

Circulating cell subgroups created during the DEPArray analysis were: Epithelial cells (E) characterized by nuclear positivity in blue (HOECHST 33342+) and a green signal (FITC+) specific for epithelial markers; Mesenchymal cells (M) characterized by nuclear positivity in blue (HOECHST 33342+) and by a red signal (PE+) specific for mesenchymal markers; Epithelial–Mesenchymal Cells (EM) characterized by blue nuclear positivity (HOECHST 33342+) and the simultaneous presence of a red signal (PE+) for mesenchymal markers and a green one (FITC+) for the epithelial ones; Lymphocytes (L) characterized by nuclear positivity (HOECHST 33342+) in blue and a blue signal (APC+) specific for CD45, sometimes by a mesenchymal red signal (PE+) and Negative cells (N) characterized by only the nuclear positivity in blue.

Cells of interest were selected using the CellBrowser Software (Menarini Silicon Biosystems), and sorted individually. Parameters provided by CellBrowser were morphological features such as: such as diameter, circularity, OV circularity, perimeter and fluorescence intensities for each channel (mean fluorescence intensity, max intensity, mean intensity without background) of each single cell found. All raw data were exported from the instrument and elaborated through bioinformatic tools.

Experimental Setup

All cellular parameters were analyzed first with single-variable analysis and then by means of machine-learning algorithms considering multiple variables (**Figure 1**).

The single variable analysis was conducted using a combination of GraphPad Prism 6.01 for the statistical analysis and Microsoft Excel 2016 for data handling. All the software used for the machine learning tests was written in Python. The version of the interpreter is Python 3.7. The software library used for the machine learning classifiers is scikit-learn 0.21.3, which is the de-facto standard library for data science with Python. Since scikit only provided a limited selection of naïve Bayes algorithms that did not fit our needs (in particular Gaussian and a Bernoulli naïve Bayes algorithm, which are targeted towards data following normal distributions and binary data respectively), we implemented a naïve Bayes algorithm able to deal with categorical data (a similar tool is now available directly from the scikit-learn library, from version 0.22.2 onwards). The system used for the analysis is a 64 bit processor Intel(R) Core I i7-7700HQ at 2.8 GHz equipped with 16 GB of RAM.

RESULTS

Overall Design, Patients' Selection and Cells Included in the Study

The study included 45 MBC patients. Each of these patients had a variable number of CTC and CD45pos cells, and each cell had several parameters provided by CellBrowser software. It was not possible to directly use the dataset, because single cells among patients were not comparable. Thus, we aggregated data of single cells in the form of descriptive statistics (average, st. dev, 25th percentile, etc.) to obtain a list of comparable features describing the cell population for each patient (**Supplementary Figure S1**).

A total of 2,598 cells belonging to the 45 MBC patients were processed, extracting 846 CD45pos cells and 344 eCTCs. Specifically, for each cell, DEPArray obtained a brightfield image and also 4 fluorescence images corresponding to the

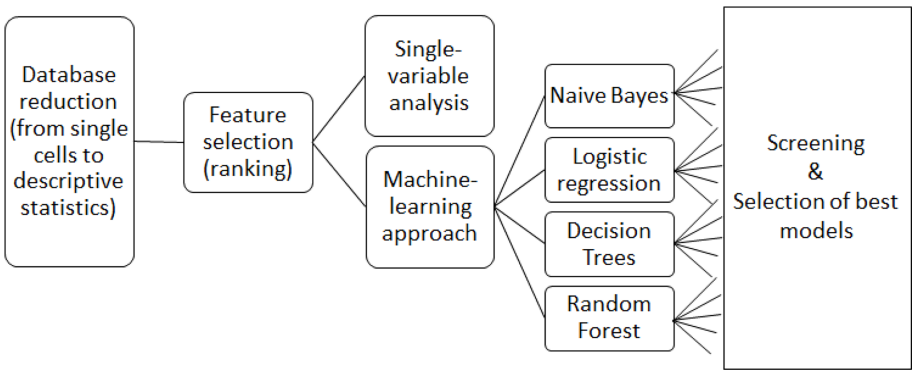


FIGURE 1 | Overview of data analysis workflow.

expression of epithelial (FITC), mesenchymal (PE), leukocyte (APC), and nuclear (DAPI) markers. From each cell image the following parameters were provided by CellBrowser software of DEPArray: circularity (using 2 algorithms, named circularity and circularityOV, the second being more effective on cells with irregular membranes), diameter, perimeter, average, and maximum intensity for each channel (both corrected and not corrected for background value). **Table 1** summarizes the clinical and pathological data of patients, while **Table 2** reports the number and type of cells for each patient.

Feature Selection and Data Preprocessing

Descriptive statistics of cell population data for each patient was performed using mean, standard deviation, 25th percentile, median and 75th percentile, resulting in 34 parameters for each patient, corresponding to the 34 features of cell images. Percentiles were included since the Shapiro–Wilk test revealed that most features did not follow a normal distribution (data not shown). In addition to data derived from image analysis, we considered the total number of cells per patient, and the absolute and relative number of eCTC and circulating CD45 positive cells.

To reduce the dimensionality of data, parameters were ranked by information gain with respect to the target variable (OS and BM). Information gain is the amount of information gained about a random variable or signal from observing another random variable; it is a method of feature selection widely used in machine-learning applications. OS was transformed into a dichotomic variable (survival ≤30 or >30 months), considering the median as threshold, so that the population could be divided in two groups equally represented. BM was transformed into a dichotomic variable as well (presence or absence of bone metastasis). Feature selection process was performed independently for eCTC and CD45pos cell populations. The ten most relevant features obtained for each of these two cell populations are listed in **Table 3**. Each selected feature for eCTC and CD45pos is visualized as box plot with respect to OS and BM in **Supplementary Figures S2**-**S5**. Since OS was originally a continuous variable, regression plot is also displayed in **Supplementary Figures S6, S7** for completeness.

With respect to OS, both morphological and phenotypic variables were selected among the most relevant, with a predominance of morphological variables. Interestingly, the number of cells was not included among this set by ranking, while known to be a good predictor of OS. With respect to BM, variables describing morphology, phenotype and the number of eCTC were included among the most relevant variables.

Most of the classification algorithms we adopted (see section *Experimental Setup*) did not need additional pre-processing to utilize the features. The only exception was naïve Bayes, which expected the features to be categorical instead of continuous. Therefore, we maintained the data in their original form when using all approaches, except for naïve Bayes, where features were discretized in four equal-frequency classes.

TABLE 1 | Demographic and clinicopathological features of the 45 MBC patients analyzed.

AGE AT THE DIAGNOSIS	
- MEDIAN (*range*)	54 (31–78)
HISTOTYPE	
Ductal	86.6%
Lobular	11.2%
Ductal and Lobular	2.2%
MOLECULAR CLASSIFICATION	
Luminal	44.4%
HER2+	31.1%
Triple negative	20.0%
N.A.	4.4%
NO. OF METASTATIC SITES	
1	31.1%
2	17.8%
>2	51.1%
METASTATIC SITES*	
Bone	66.7%
Liver	44.4%
Lymphonodes	33.3%
SNC	11.1%
Skin	20.0%
Lung	35.5%

**Patients may have more than one site involved.*

TABLE 2 | Distribution of cells in patients.

Patient id	no. of cells	CD45pos	eCTC	Patient id	no. of cells	CD45pos	eCTC
1	11	1	1				
2	125	13	1	**24**	25	12	9
3	53	6	31	**25**	77	12	4
4	80	18	51	**26**	79	41	1
5	48	9	1	**27**	87	64	2
6	73	30	11	**28**	60	12	7
7	31	6	2	**29**	21	6	4
8	21	7	5	**30**	84	35	5
9	40	13	6	**31**	7	1	2
10	21	8	7	**32**	38	26	3
11	46	0	16	**33**	24	3	0
12	52	33	2	**34**	15	0	2
13	12	0	9	**35**	98	9	8
14	94	14	3	**36**	67	51	1
15	47	7	0	**37**	127	66	3
16	98	39	2	**38**	101	27	18
17	11	5	0	**39**	72	26	16
18	56	23	4	**40**	35	24	0
19	32	23	1	**41**	15	0	11
20	144	25	62	**42**	57	11	3
21	63	32	1	**43**	62	17	6
22	72	30	8	**44**	49	15	0
23	111	25	5	**45**	57	21	10
				TOT =	**2598**	**846**	**344**

CD45pos, CD45-positive cells; eCTC, epithelial circulating tumor cells.
Bold is the total (sum) of each column.

Single Variable Analysis Demonstrated That Morphology of Both eCTC and CD45pos Predict Prognosis and Bone Metastasis

For both eCTC and CD45pos, we selected the best feature, used ROC curve analysis to detect the best cutoff for the variable with respect to the target (either OS or BM) using the Youden index (calculated as SN + SP − 1, where SN is the sensitivity and SP is the specificity), and represented Kaplan–Meier curve for OS and

TABLE 3 | Best features ranked by information gain, with respect to overall survival and bone metastasis.

OVERALL SURVIVAL			
eCTC		**CD45 positive cells**	
FEATURES	**SCORE**	**FEATURES**	**SCORE**
circularityOV_brightfield_25th	0.237*	circularityOV_brightfield_SD	0.203*
perimeter_fitc_25th	0.215*	circularityOV_fitc_25th	0.178*
circularity_brightfield_25th	0.189*	circularity_brightfield_25th	0.169*
mean_intensity_bgsub_apc_SD	0.184*	circularity_fitc_25th	0.163*
circularity_brightfield_mean	0.174*	mean_intensity_bgsub_pe_25th	0.154*
circularity_apc_mean	0.146	perimeter_fitc_75th	0.146
circularityOV_pe_75th	0.146	circularityOV_fitc_mean	0.146
max_intensity_brightfield_median	0.142	diameter_brightfield_25th	0.133
diameter_apc_median	0.138	circularity_fitc_SD	0.130
circularity_dapi_25th	0.133	circularityOV_brightfield_median	0.121
BONE METASTASIS			
eCTC		**CD45 positive cells**	
FEATURES	**SCORE**	**FEATURES**	**SCORE**
diameter_fitc_median	0.211*	diameter_pe_SD	0.203*
% of eCTC	0.189*	circularity_fitc_SD	0.203*
perimeter_apc_25 th	0.189*	perimeter_pe_SD	0.203*
circularity_fitc_SD	0.177*	perimeter_fitc_SD	0.163
circularityOV_fitc_SD	0.177*	perimeter_brightfield	0.155
max_intensity_apc_SD	0.177*	circularity_apc_75th	0.153
circularityOV_brightfield_SD	0.177*	circularityOV_brightfield_75th	0.139
mean_intensity_bgsub_apc_25th	0.170	circularity_apc_25th	0.139
diameter_apc_mean	0.167	circularity_pe_median	0.134
diameter_apc_75th	0.167	circularityOV_pe_75th	0.134

*Each feature is described by the parameter, the channel of collection (brightfield, fitc, pe or apc) and descriptive statistics feature (mean, standard deviation, median, 25th or 75th percentile). SD, standard deviation; mean_intensity_bgsub, mean intensity after background subtraction; fitc, epithelial marker expression; pe, mesenchymal marker expression; apc, CD45 expression; DAPI, nuclear staining. *features subsequently selected for the combined approach (see Experimental Setup).*

contingency tables for BM. Survival curves and contingency tables were obtained using the leave-one-out method: cut-off was assessed on all patients except for one, on which prediction for survival and bone metastasis were performed according to the established cut-off. This was iterated for all patients, so that each prediction was made on a patient who was not used for cut-off assessment. Interestingly the best variable was morphological in all cases.

Considering OS, circularity, measured in brightfield images, resulted to be the most predictive feature for both eCTC and CD45pos, although two different aspects were considered for the two types of cell: the 25th percentile for eCTC (i.e., circularity degree) and standard deviation for CD45pos (i.e., variability in circularity). The median survival of MBC patients, stratified as predicted to survive <= or > 30 months months, resulted to be 15 months vs. 43 months for eCTC ($p = 0.03$, Log-Rank) and 19 months vs. 36 months for CD45pos ($p = 0.16$, Log-Rank) (**Figure 2**).

Considering the presence of bone metastases, the best predictors resulted to be the diameter for either eCTC (increased median value) or CD45pos (increased standard deviation), measured in different fluorescence channels. Using the same iterative cut-off method to predict MBC patients as having or not BM. eCTC could predict BM with a positive predictive value (PPV) of 79% and a negative predictive value (NPV) of 48%, while CD45pos presented a PPV of 91% and an NPV of 41%. The accuracy was 64% for eCTC and 53% for CD45pos (**Table 4**).

The prediction showed strong PPV, but high number of false negatives. In the attempt of improving this results, we explored different machine learning approaches.

Machine Learning Approaches Improved the Accuracy in Predicting Overall Survival and Bone Metastasis

The machine learning approaches selected for our tests are the following:

Logistic regression: A statistical model commonly used in medicine to classify binary target variables (32–37).

Decision trees: this algorithm is considered a weak classifier, but able to organize features based on their importance and find the best cut-off value for discriminating subgroups. It is a white-box approach, therefore it offers an explanation of every choice the algorithm made, making it well suited for medical applications (24, 32, 38).

Random forest: An approach that represents an evolution of the previous: by combining several decision trees in a voting system, this algorithm is able to mitigate the error that a single decision tree might have. It is less transparent than a single decision tree, but it typically performs better in terms of classification (32, 39).

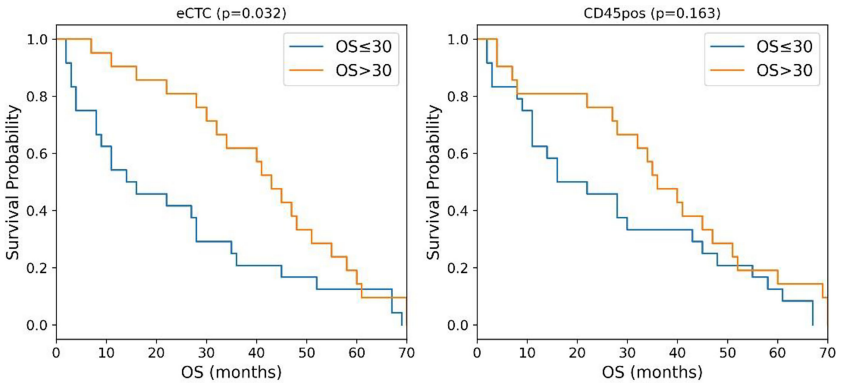


FIGURE 2 | Kaplan–Meier curves of MBC patients stratified according to the circularity of eCTC (left) and CD45 positive cells (right). P-values were calculated by Log Rank test.

TABLE 4 | Contingency tables of prediction of bone metastasis based on a single variable derived from either eCTC (left) or CD45-positive cells (right).

		eCTC-based prediction Actual				CD45pos-based prediction Actual		
		BM+	**BM–**			**BM+**	**BM–**	
Predicted	BM+	19	5	**PPV** 0.79	BM+	10	1	**PPV** 0.91
	BM–	11	10	**NPV** 0.48	BM–	20	14	**NPV** 0.41
		Sensitivity 0.63	**Specificity** 0.67	**Accuracy** 0.64		**Sensitivity** 0.33	**Specificity** 0.93	**Accuracy** 0.53

Columns indicate the actual positive and negative patients, while the rows indicate the predicted positives and negatives patients. BM+, presence of bone metastasis; BM–, absence of bone metastasis, PPV, positive predictive value; NPV, negative predictive value.

Naive Bayes: It is a probabilistic machine learning method which assumes strong independence between the features. While this assumption is typically too "naive" for non-synthetic data, where there are often hidden dependences between variables, this approach has been applied successfully in many real-world scenarios (23, 32).

As in the case of single-variable analysis, image-based features of eCTC and CD45pos cells were used as inputs and OS (≤30 vs. >30 months) or BM (absence vs. presence) as output.

For each model, we evaluated the "power set" of the best ten features identified during feature selection. The "power set" includes all possible subsets of a given set (e.g., if our set is [1, 2, 3], the power set is [1, 2], [2, 3], [1, 3], [1], [2], [3], [], [1, 2, 3]). Thus, for each model, we tested 1023 possible subsets of features with size ranging from 1 to 10 features (**Supplementary Table S1**). Thus, we screened all models with all combinations of features, to identify the best one. Each model was cross-validated with leave-one-out strategy, that is, training of the model on all patients except for one, which is in turn used as test set, doing this iteratively for all patients. The performance of the model is thus the average of all "leave-one-out" models created.

Models were trained independently for eCTC and CD45pos, then we evaluated models taking into account both cell populations combined.

Naïve Bayes resulted to be the best classifier in all cases: considering all three possible inputs (eCTC, CD45pos, eCTC & CD45pos) and all possible target variables (OS or BM) (**Supplementary Table S1**). Details on the results obtained by the Naïve Bayes approach are reported below.

Both eCTC and CD45pos Features Could Predict Overall Survival

Table 5 shows the features considered by the best models for eCTC, CD45pos and eCTC & CD45pos. The power set of 10 features was evaluated, but the best performing subset of features only contained 6 features for eCTC, 3 features for CD45pos and 4 features for eCTC & CD45pos. This underlines that addition of a feature is not always beneficial and can actually lead to worst performance, increasing noise. Regarding the parameters selected, they were mainly morphological in the case of eCTC (circularity of cell and nucleus and perimeter), while, for the CD45pos, both circularity and expression of mesenchymal markers (PE) were chosen by the Naïve Bayes model.

As shown in **Figure 3**, the Naïve Bayes model significantly stratifies patients according to prognosis using image features of either eCTC and CD45pos alone or in combination.

The median OS difference was similarly significant in all three cell subsets: eCTC (46 months versus 11 months; p <0.0001), CD45pos (12.5 vs. 45 months; p = 0.0004) and eCTC+CD45pos

TABLE 5 | Features identified by the naïve Bayes approach as the most informative to predict overall survival and bone metastasis considering eCTC features alone (left), CD45pos alone (center) or both (right).

eCTC	OVERALL SURVIVAL CD45pos	eCTC & CD45pos
circularityOV_brightfield_25th	circularityOV_brightfield_SD	eCTC: perimeter_fitc_25th
perimeter_fitc_25th	circularity_fitc_25th	eCTC: circularity_brightfield_mean
circularity_apc_mean	mean_intensity_bgsub_pe_25th	CD45pos cells: circularityOV_brightfield_SD
circularityOV_pe_75th		CD45pos cells: mean_intensity_bgsub_pe_25th
max_intensity_brightfield_median		
circularity_dapi_25th		
eCTC	**BONE METASTASIS CD45pos**	**eCTC & CD45pos**
perimeter_apc_25th	circularity_fitc_SD	eCTC: perimeter_apc_25th
percentage of eCTC	circularity_apc _75th	eCTC: percentage of eCTC
circularityOV_brightfield_SD		eCTC: circularityOV_brightfield_SD
max_intensity_apc_SD		eCTC: max_intensity_apc_SD

SD, standard deviation; mean_intensity_bgsub, mean intensity after background subtraction; fitc, epithelial marker expression; pe, mesenchymal marker expression; apc, CD45 expression.

(11 vs. 45 months; $p = 0.0003$). The combined approach was slightly more accurate in predicting OS (89%) with respect to eCTC or CD45pos considered alone (82 and 84%, respectively). Thus, the combination of the information obtained from eCTC and CD45pos worked better than considering these cell populations separately.

Altogether these data showed that, with respect to the single variable analysis (**Table 5**), adopting a machine learning approach significantly increased accuracy in stratification of patients by survival. The improvement in accuracy was significant in the case of eCTC (from 73.3 to 82%), and even higher in CD45pos (from 66.7 to 84%). Moreover, the combination of image data obtained from eCTC and CD45pos further boosted the classification accuracy to 89%, confirming the benefit of associating information from both cell types.

eCTC Predicted the Presence of Bone Metastases With Greater Accuracy Than CD45 Positive Cells

Naïve Bayes was the best performing model also concerning the BM prediction (**Supplementary Table S1**). In **Table 5** are summarized the subsets of features selected for eCTC, CD45pos and eCTC & CD45pos.

In the case of eCTC, beside features strictly related to image analysis (perimeter, circularity and aberrant expression of CD45), the percentage of eCTC was selected as an informative feature, that is the fraction of eCTC on total CTC detected in that patient, suggesting a role for the number of CTC in prediction of bone metastasis. In the case of CD45pos, circularity and expression of mesenchymal markers resulted to be informative. Interestingly, in the combined approach the features selected were all derived from eCTC, indicating no improvement derived by combining the analysis with CD45pos.

Considering the contingency tables (**Table 6**), it is apparent that, with respect to the single-variable analysis, the accuracy was strongly increased either considering eCTC (from 67 to 91%) or CD45pos alone (from 58 to 84%).

In particular, the eCTC model performed better than the CD45pos one. Indeed, specificity and PPV were both 100% for eCTC and 80 and 84% for CD45pos.

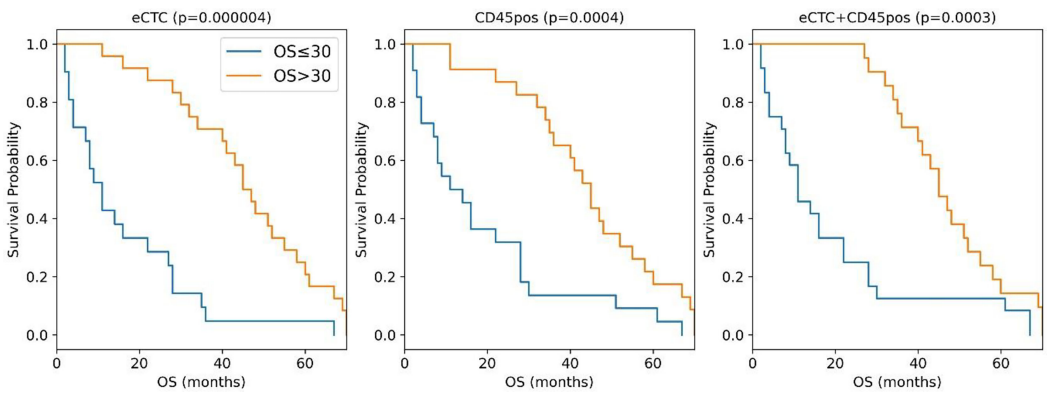


FIGURE 3 | Kaplan–Meier curves of the MBC patients stratified in OS <= 30 months (blue curve) or >30 months (orange curves) according to the naïve Bayes analysis conducted taking into consideration eCTC (left panel), CD45pos (central panel) or eCTC+C45pos (right panel).

TABLE 6 | Contingency tables of the prediction of bone metastases adopting a machine learning approach taking into consideration only eCTC (top), only CD45-positive cells (middle) or both (bottom).

eCTC				
		Actual		
		Pos	**Neg**	
Predicted	Pos	26	0	PPV = 1
	Neg	4	15	NPV = 0.79
		Sensitivity	Specificity	Accuracy
		0.87	1	0.91
CD45 Positive Cells				
		Actual		
		Pos	**Neg**	
Predicted	Pos	26	3	PPV = 0.9
	Neg	4	12	NPV = 0.75
		Sensitivity	Specificity	Accuracy
		0.87	0.80	0.84
eCTC and CD45pos				
		Actual		
		Pos	**Neg**	
Predicted	Pos	26	0	PPV = 1
	Neg	4	15	NPV = 0.79
		Sensitivity	Specificity	Accuracy
		0.87	1	0.91

PPV, positive predictive value; NPV, negative predictive value.

Differently from OS, considering eCTC & CD45pos did not improve the accuracy in predicting bone metastases. As additional evidence, the combined approach used the same features of the model set on eCTC only.

DISCUSSION

Systematic and quantitative image analysis of cells and machine-learning have been employed in CTC detection methods (40–42). Moreover, a software application named ACCEPT intended to segment images of cells and extract multiple parameters was recently published (43). Applications of ACCEPT found in literature were however limited to accurate and reproducible assessment of particular features [e.g., treatment target expression levels (43) or size (44)], or cell classification (45). To our knowledge, quantitative features extracted from images of isolated CTC have never been employed as prognostic biomarkers for clinical outcomes either alone or integrated in complex modeling. This paper offers evidence that useful information can be extracted from quantitative analysis of images of isolated CTC. Moreover and surprisingly, information about overall survival could also be extracted from images of leukocytes. We conducted both a single variable analysis and a multi-variable analysis with machine-learning approaches. In general, features that when taken alone showed poor performance in discriminating between target variables (OS and bone metastasis), were instead capable of generating effective models when integrated in a multi-features model.

Some biological insights might be gained by a closer look to features selected by ranking and model optimization. With respect to eCTC and OS, features ranking indicated predominantly morphological properties, and some protein expression data. The most represented morphological aspect was circularity, which is the most prevalent feature, in various channels and statistical variables, and it is defined as:

$$4\pi \times \frac{Area}{Perimeter^2}$$

Circularity is thus inversely proportional to the square of perimeter, meaning that membranes with higher complexity (frequency and extent of indentations) have lower levels of circularity.

Higher circularity values (simpler membranes) are linked to poor survival. In patients with lower overall survival, both nucleus and membrane of eCTC have higher circularity. In a purely speculative way, in the attempt to attribute a meaning to this information, the ideal representation of a cell with a highly circular membrane and nucleus is a small basal-like or stem-like cell with low differentiation, which might be more be responsible of cancer progression (46). Thus, the increased average circularity of CTC population might indicate an increased proportion of such highly aggressive cells.

Protein expression in patients with lower overall survival showed higher variation (SD) in CD45 expression in eCTC (higher mean_intensity_bgsub_apc_SD). Considering that eCTC do not show CD45 expression, we cannot give a biological interpretation to this feature. From a data analysis point of view, it is very interesting that a feature typically used as categorical (presence/absence of CD45 expression) seems to have instead some information when considered quantitatively, even inside the same category of "negative" CD45 expression.

Considering CD45-positive cells and OS, cells also showed significantly increased circularity (and decreased standard deviation) in lower OS, indicating a more circular and homogenous cell population in patients with lower OS. Interpretation of this variable is not easy as we do not know whether CD45pos are neutrophils, monocytes or lymphocytes.

With respect to bone metastasis, eCTC showed morphological, protein expression, and % composition features. The eCTC population associated with bone metastasis can grossly be described as bigger, more circular, and with higher fraction of epithelial cells over total CTC. This provides an interesting insight in morphological properties which could be worth investigating with deeper molecular analysis, in order to understand why these cells display such preferential trophism for bone.

Considering bone metastasis and CD45pos, cells show substantially a lower circularity when bone metastasis are present.

The majority of these variables are selected also in the independent process of model screening and optimization. With respect to the machine learning analysis, we provided an exhaustive benchmark of the available algorithms. In the totality of cases, naïve Bayes proved to be the best classifier. In the analysis for the OS prediction, there was a significant improvement compared with the single-variable analysis, in terms of both accuracy and Kaplan–Meier curve, particularly in CD45pos cells. In the single-variable analysis, CD45pos cells failed to stratify patients according to survival. By exclusively using this approach, one would conclude that no information related to survival is contained in CD45pos. The use of a more complex approach instead, able to highlight more

subtle relationships hidden in data, showed that CD45pos do actually contain information about survival, apparently comparable to eCTC, as effective stratification of patients was possible. Moreover, the combined approach boosted the performance of the model from 0.84 to 0.89 of accuracy, suggesting that information coming from CD45pos is different and complementary to eCTC.

The naïve Bayes classifier proved to be a good predictor of BM, especially in terms of specificity and positive predictive value. Contrarily to OS prediction, combining the information from CD45pos does not improve the performance of the classifier.

Thus, both CD45pos and eCTC cells are informative with respect to OS, and their information is different and complementary, because combining information coming from the two populations showed better performance than considering either CD45pos or eCTC alone. Moreover, combined model showed top-ranked features of both cell subpopulations.

In BM prediction instead, information was found mainly in eCTC population. CD45pos is informative, but information is overshadowed by eCTC. Combining information from eCTC and CD45pos did not improve performance, with the combined model showing only eCTC features.

A possible explanation of these facts is that eCTC and CD45pos contain information regarding two different aspects of patient-tumor interaction: eCTC contain information about biological features of cancer, while CD45pos offer an insight into the host immune system status. For this reason, considering both these aspects by combining information offer better prediction on survival than taken singularly. Bone metastasis instead are mainly dependent on the trophism of cancer cells, and are thus mainly predicted by eCTC features.

Conclusions

The study suggests that quantitative image analysis can reveal undiscovered meaningful information. Thanks to modern machine learning approach, the massive amount of data yielded by quantitative image analysis can be linked to clinical outcomes effectively. In our specific case, images of epithelial CTC and leukocytes revealed information predicting overall survival and metastatic pattern of MBC patients. The method uses standardized outputs (cell images and data obtained by DEPArray) and relatively simple models (e.g., Naïve Bayes), and can thus be easily scaled-up and standardized for further validation.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ceur fvg, N.152/2011/Sper and N.178/2014 Em. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conceptualization, GDC, FDB, and MT. Methodology, GDC, FDB, and MB. Software, GDC. Validation, GDC and FDB. Formal analysis, GDC. Investigation, APB and DC. Resources, APB and DC. Data curation, MB. Writing—Original draft preparation, GDC and MB. Writing—Review and editing, FDB, MT, and DC. Visualization, GDC and FDB. Supervision, FDB and DC. Project administration, APB and DC. Funding acquisition, APB and DC. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

The study was supported by the AIRC IG 2017-20443: "Dissecting the heterogeneity of circulating tumor cells in metastatic breast cancer patients to predict clinical outcome", CUP G23C17000800007, and the Project "HEaD Higher Education and Development" SISSA Operazione 2 FP1619889003, funding channel 1420AFPLO2, Region FVG, co-funded by Fondo Sociale Europeo POR 2014/2020. The TITAN X used for this research was donated by the NVIDIA Corporation.

ACKNOWLEDGMENTS

We thank Prof. A. Laio, who contributed to the project funding.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.725318/full#supplementary-material

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Pineros M, et al. Estimating the Global Cancer Incidence and Mortality in 2018: GLOBOCAN Sources and Methods. *Int J Cancer* (2019) 144:1941–53. doi: 10.1002/ijc.31937
2. Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and Mortality and Epidemiology of Breast Cancer in the World. *Asian Pac J Cancer Prev* (2016) 17:43–6. doi: 10.7314/APJCP.2016.17.S3.43
3. Byler S, Goldgar S, Heerboth S, Leary M, Housman G, Moulton K, et al. Genetic and Epigenetic Aspects of Breast Cancer Progression and Therapy. *Anticancer Res* (2014) 34:1071–7.

4. Fidler IJ. Critical Determinants of Metastasis. *Semin Cancer Biol* (2002) 12:89–96. doi: 10.1006/scbi.2001.0416
5. Fidler IJ. The Pathogenesis of Cancer Metastasis: The "Seed and Soil" Hypothesis Revisited. *Nat Rev Cancer* (2003) 3:453–8. doi: 10.1038/nrc1098
6. Hsiao YH, Chou MC, Fowler C, Mason JT, Man YG. Breast Cancer Heterogeneity: Mechanisms, Proofs, and Implications. *J Cancer* (2010) 1:6–13. doi: 10.7150/jca.1.6
7. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, et al. Metastatic Behavior of Breast Cancer Subtypes. *J Clin Oncol* (2010) 28:3271–7. doi: 10.1200/JCO.2009.25.9820
8. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The Landscape of Cancer Genes and Mutational Processes in Breast Cancer. *Nature* (2012) 486:400–4. doi: 10.1038/nature11017
9. Fittall MW, Van Loo P. Translating Insights Into Tumor Evolution to Clinical Practice: Promises and Challenges. *Genome Med* (2019) 11:20. doi: 10.1186/s13073-019-0632-z
10. Alix-Panabieres C, Pantel K. Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy. *Cancer Discov* (2016) 6:479–91. doi: 10.1158/2159-8290.CD-15-1483
11. Appierto V, Di Cosimo S, Reduzzi C, Pala V, Cappelletti V, Daidone MG. How to Study and Overcome Tumor Heterogeneity With Circulating Biomarkers: The Breast Cancer Case. *Semin Cancer Biol* (2017) 44:106–16. doi: 10.1016/j.semcancer.2017.04.007
12. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating Liquid Biopsies Into the Management of Cancer. *Nat Rev Clin Oncol* (2017). doi: 10.1038/nrclinonc.2017.14
13. Bulfoni M, Gerratana L, Del Ben F, Marzinotto S, Sorrentino M, Turetta M, et al. In Patients With Metastatic Breast Cancer the Identification of Circulating Tumor Cells in Epithelial-to-Mesenchymal Transition Is Associated With a Poor Prognosis. *Breast Cancer Res* (2016) 18:30. doi: 10.1186/s13058-016-0687-3
14. Bulfoni M, Turetta M, Del Ben F, Di Loreto C, Beltrami A, Cesselli D. Dissecting the Heterogeneity of Circulating Tumor Cells in Metastatic Breast Cancer: Going Far Beyond the Needle in the Haystack. *Int J Mol Sci* (2016) 17:1775. doi: 10.3390/ijms17101775
15. Bidard F-C, Peeters DJ, Fehm T, Nolé F, Gisbert-Criado R, Mavroudis D, et al. Clinical Validity of Circulating Tumour Cells in Patients With Metastatic Breast Cancer: A Pooled Analysis of Individual Patient Data. *Lancet Oncol* (2014) 15:406–14. doi: 10.1016/S1470-2045(14)70069-5
16. Zhang L, Riethdorf S, Wu G, Wang T, Yang K, Peng G, et al. Meta-Analysis of the Prognostic Value of Circulating Tumor Cells in Breast Cancer. *Clin Cancer Res* (2012) 18:5701–10. doi: 10.1158/1078-0432.CCR-12-1587
17. Janni WJ, Rack B, Terstappen LWMM, Pierga J-Y, Taran F-A, Fehm T, et al. Pooled Analysis of the Prognostic Relevance of Circulating Tumor Cells in Primary Breast Cancer. *Clin Cancer Res* (2016) 22:2583–93. doi: 10.1158/1078-0432.CCR-15-1603
18. Yu M, Bardia A, Wittner BS, Stott SL, Smas ME, Ting DT, et al. Circulating Breast Tumor Cells Exhibit Dynamic Changes in Epithelial and Mesenchymal Composition. *Science* (2013) 339:580–4. doi: 10.1126/science.1228522
19. Zhou J, Zhu X, Wu S, Guo J, Zhang K, Xu C, et al. Epithelial-Mesenchymal Transition Status of Circulating Tumor Cells in Breast Cancer and its Clinical Relevance. *Cancer Biol Med* (2020) 17:169–80. doi: 10.20892/j.issn.2095-3941.2019.0118
20. Satelli A, Mitra A, Brownlee Z, Xia X, Bellister S, Overman MJ, et al. Epithelial-Mesenchymal Transitioned Circulating Tumor Cells Capture for Detecting Tumor Progression. *Clin Cancer Res* (2015) 21:899–906. doi: 10.1158/1078-0432.CCR-14-0894
21. Vansant G, Wang Y, Hom B, Jendrisak A, Schonhoft J, Graf RP, et al. Analysis of Circulating Tumor Cells (CTCs) in Patients Across Multiple Metastatic Breast Cancer (mBCa) Cohorts Identifies Marked Inter- and Intra-Patient Heterogeneity in CTC Size, Shape, and Overall Morphology. *J Clin Oncol* (2018) 36:1084–4. doi: 10.1200/JCO.2018.36.15_suppl.1084
22. Kermanshah L, Poudineh M, Ahmed S, Nguyen LNM, Srikant S, Makonnen R, et al. Dynamic CTC Phenotypes in Metastatic Prostate Cancer Models Visualized Using Magnetic Ranking Cytometry. *Lab Chip* (2018) 18:2055–64. doi: 10.1039/C8LC00310F
23. Zhang H. *The Optimality of Naïve Bayes*. (2004). Available at: https://citeseerx.ist.psu.edu/viewdoc/citations;jsessionid=371ED8E43C8A9F019C2A77F0AA8E67C0?doi=10.1.1.483.2183 (Accessed January 12, 2022).
24. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software (1984).
25. Verronese E, Delgado A, Valladeau-Guilemond J, Garin G, Guillemaut S, Tredan O, et al. Immune Cell Dysfunctions in Breast Cancer Patients Detected Through Whole Blood Multi-Parametric Flow Cytometry Assay. *Oncoimmunology* (2016) 5:e1100791. doi: 10.1080/2162402X.2015.1100791
26. Law AM, Lim E, Ormandy CJ, Gallego-Ortega D. The Innate and Adaptive Infiltrating Immune Systems as Targets for Breast Cancer Immunotherapy. *Endocr Relat Cancer* (2017) 24:123–44. doi: 10.1530/ERC-16-0404
27. Holl EK, Frazier VN, Landa K, Beasley GM, Hwang ES, Nair SK. Examining Peripheral and Tumor Cellular Immunome in Patients With Cancer. *Front Immunol* (2019) 10:1767. doi: 10.3389/fimmu.2019.01767
28. Wang L, Simons DL, Lu X, Tu TY, Avalos C, Chang AY, et al. Breast Cancer Induces Systemic Immune Changes on Cytokine Signaling in Peripheral Blood Monocytes and Lymphocytes. *EBioMedicine* (2020) 52:102631. doi: 10.1016/j.ebiom.2020.102631
29. Muraro E, Comaro E, Talamini R, Turchet E, Miolo G, Scalone S, et al. Improved Natural Killer Cell Activity and Retained Anti-Tumor CD8(+) T Cell Responses Contribute to the Induction of a Pathological Complete Response in HER2-Positive Breast Cancer Patients Undergoing Neoadjuvant Chemotherapy. *J Transl Med* (2015) 13:204. doi: 10.1186/s12967-015-0567-0
30. Papatestas AE, Lesnick GJ, Genkins G, Aufses AH. The Prognostic Significance of Peripheral Lymphocyte Counts in Patients With Breast Carcinoma. *Cancer* (1976) 37:164–8. doi: 10.1002/1097-0142(197601)37:1<164::AID-CNCR2820370123>3.0.CO;2-H
31. Vicente Conesa MA, Garcia-Martinez E, Gonzalez Billalabeitia E, Chaves Benito A, Garcia Garcia T, Vicente Garcia V, et al. Predictive Value of Peripheral Blood Lymphocyte Count in Breast Cancer Patients Treated With Primary Chemotherapy. *Breast* (2012) 21:468–74. doi: 10.1016/j.breast.2011.11.002
32. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-Learn: Machine Learning in Python. *J Mach Learn Res* (2011) 12:2825–30.
33. Boyd CR, Tolson MA, Copes WS. Evaluating Trauma Care: The TRISS Method. *J Trauma Acute Care Surg* (1987) 27:370–8. doi: 10.1097/00005373-198704000-00005
34. Kologlu M, Elker D, Altun H, Sayek I. Validation of MPI and PIA II in Two Different Groups of Patients With Secondary Peritonitis. *Hepatogastroenterology* (2001) 48:147–51.
35. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple Organ Dysfunction Score: A Reliable Descriptor of a Complex Clinical Outcome. *Crit Care Med* (1995) 23:1638–52. doi: 10.1097/00003246-199510000-00007
36. Le Gall J-R, Lemeshow S, Saulnier F. A New Simplified Acute Physiology Score (SAPS II) Based on a European/North American Multicenter Study. *JAMA* (1993) 270:2957–63. doi: 10.1001/jama.270.24.2957
37. Biondo S, Ramos E, Deiros M, Ragué JM, De Oca J, Moreno P, et al. Prognostic Factors for Mortality in Left Colonic Peritonitis: A New Scoring System. *J Am Coll Surg* (2000) 191:635–42. doi: 10.1016/S1072-7515(00)00758-4
38. Cesselli D, Ius T, Isola M, Del Ben F, Da Col G, Bulfoni M, et al. Application of an Artificial Intelligence Algorithm to Prognostically Stratify Grade II Gliomas. *Cancers* (2019) 12. doi: 10.3390/cancers12010050
39. Ho TK. Random Decision Forests, in: *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Vol. 1. IEEE Computer Society (1995). pp. 278–82.
40. Soldati G, Del Ben F, Brisotto G, Biscontin E, Bulfoni M, Piruska A, et al. Microfluidic Droplets Content Classification and Analysis Through Convolutional Neural Networks in a Liquid Biopsy Workflow. *Am J Transl Res* (2018) 10:4004–16.
41. Wang S, Zhou Y, Qin X, Nair S, Huang X, Liu Y. Label-Free Detection of Rare Circulating Tumor Cells by Image Analysis and Machine Learning. *Sci Rep* (2020) 10:12226. doi: 10.1038/s41598-020-69056-1
42. Lannin TB, Thege FI, Kirby BJ. Comparison and Optimization of Machine Learning Methods for Automated Classification of Circulating Tumor Cells. *Cytometry A* (2016) 89:922–31. doi: 10.1002/cyto.a.22993
43. Zeune L, van Dalum G, Decraene C, Proudhon C, Fehm T, Neubauer H, et al. Quantifying HER-2 Expression on Circulating Tumor Cells by ACCEPT. *PloS One* (2017) 12:e0186562. doi: 10.1371/journal.pone.0186562
44. Mendelaar PAJ, Kraan J, Van M, Zeune LL, Terstappen LWMM, Oomen-de Hoop E, et al. Defining the Dimensions of Circulating Tumor Cells in a Large Series of Breast, Prostate, Colon, and Bladder Cancer Patients. *Mol Oncol* (2021) 15:116–25. doi: 10.1002/1878-0261.12802

45. de Wit S, Zeune LL, Hiltermann TJN, Groen HJM, van DG, Terstappen LWMM. Classification of Cells in CTC-Enriched Samples by Advanced Image Analysis. *Cancers (Basel)* (2018) 10:E377. doi: 10.3390/cancers10100377
46. Elston CW, Ellis IO. Pathological Prognostic Factors in Breast Cancer. I. The Value of Histological Grade in Breast Cancer: Experience From a Large Study With Long-Term Follow-Up. *Histopathology* (1991) 19:403–10. doi: 10.1111/j.1365-2559.1991.tb00229.x

Conflict of Interest: FDB and MT co-founded a start-up company focused on liquid biopsy and circulating tumor cells detection (Lighthouse Biotech srl).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 28 February 2022
doi: 10.3389/fonc.2022.811347

MRI-Based Radiomics Nomogram: Prediction of Axillary Non-Sentinel Lymph Node Metastasis in Patients With Sentinel Lymph Node-Positive Breast Cancer

Ya Qiu[1,2,3†], Xiang Zhang[1,2†], Zhiyuan Wu[4], Shiji Wu[2,5,6], Zehong Yang[1,2], Dongye Wang[1,2], Hongbo Le[1,2], Jiaji Mao[1,2], Guochao Dai[3], Xuwei Tian[3], Renbing Zhou[3], Jiayi Huang[1,2], Lanxin Hu[1,2] and Jun Shen[1,2]*

[1] Department of Radiology, Sun Yat-Sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China, [2] Guangdong Provincial Key Laboratory of Epigenetics and Gene Regulation of Malignant Tumors, Sun Yat-sen Memorial Hospital, Guangzhou, China, [3] Department of Radiology, the First People's Hospital of Kashi Prefecture, Kashi, China, [4] School of Public Health, Capital Medical University, Beijing, China, [5] Department of Ultrasound, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China, [6] Department of Ultrasound, the First People's Hospital of Kashi Prefecture, Kashi, China

Background: Overtreatment of axillary lymph node dissection (ALND) may occur in patients with axillary positive sentinel lymph node (SLN) but negative non-SLN (NSLN). Developing a magnetic resonance imaging (MRI)-based radiomics nomogram to predict axillary NSLN metastasis in patients with SLN-positive breast cancer could effectively decrease the probability of overtreatment and optimize a personalized axillary surgical strategy.

Methods: This retrospective study included 285 patients with positive SLN breast cancer. Fifty five of them had metastatic NSLNs and 230 had non-metastatic NSLNs. MRI-based radiomic features of primary tumors were extracted and MRI morphologic findings of the primary tumor and axillary lymph nodes were assessed. Four models, namely, a radiomics signature, an MRI-clinical nomogram, and two MRI-clinical-radiomics nomograms were established based on MRI morphologic findings, clinicopathologic characteristics, and MRI-based radiomic features to predict the NSLN status. The optimal predictors in each model were selected using the 5-fold cross-validation (CV) method. Their predictive performances were determined by the receiver operating characteristic (ROC) curves analysis. The area under the curves (AUCs) of different models was compared by the Delong test. Their discrimination capability, calibration curve, and clinical usefulness were also assessed.

Results: The 5-fold CV analysis showed that the AUCs ranged from 0.770 to 0.847 for the radiomics signature, from 0.720 to 0.824 for the MRI-clinical nomogram, from 0.843 to 0.932 for the MRI-clinical-radiomics nomogram. The optimal predictive factors in the radiomics signature, MRI-clinical nomogram, and MRI-clinical-radiomics nomogram were

OPEN ACCESS

Edited by:
Mangesh A. Thorat, Guy's and St. Thomas' NHS Foundation Trust, United Kingdom

Reviewed by:
Amar Ahmad, Cancer Research UK (CRUK), United Kingdom
Isaac Daimiel Naranjo, Guy's and St. Thomas' NHS Foundation Trust, United Kingdom

***Correspondence:**
Jun Shen
shenjun@mail.sysu.edu.cn

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: *08 November 2021*
Accepted: *01 February 2022*
Published: *28 February 2022*

Citation:
Qiu Y, Zhang X, Wu Z, Wu S, Yang Z, Wang D, Le H, Mao J, Dai G, Tian X, Zhou R, Huang J, Hu L and Shen J (2022) MRI-Based Radiomics Nomogram: Prediction of Axillary Non-Sentinel Lymph Node Metastasis in Patients With Sentinel Lymph Node-Positive Breast Cancer. Front. Oncol. 12:811347. doi: 10.3389/fonc.2022.811347

one texture feature of diffusion-weighted imaging (DWI), two clinicopathologic features together with one MRI morphologic finding, and the DWI-based texture feature together with the two clinicopathologic features plus the one MRI morphologic finding, respectively. The MRI-clinical-radiomics nomogram with CA 15-3 included achieved the highest AUC compared with the radiomics signature (0.868 *vs*. 0.806, P <0.001) and MRI-clinical nomogram (0.868 *vs*. 0.761; P <0.001). In addition, the MRI-clinical-radiomics nomogram without CA 15-3 showed a higher performance than that of the radiomics signature (AUC, 0.852 *vs*. 0.806, P = 0.016) and the MRI-clinical nomogram (AUC, 0.852 *vs*. 0.761, P = 0.007). The MRI-clinical-radiomics nomograms showed good discrimination and good calibration. Decision curve analysis demonstrated that the MRI-clinical-radiomics nomograms were clinically useful.

Conclusion: The MRI-clinical-radiomics nomograms developed in our study showed high predictive performance, which can be used to predict the axillary NSLN status in SLN-positive breast cancer patients before surgery.

Keywords: multiparametric magnetic resonance imaging, nomograms, sentinel lymph node, lymph node excision, breast neoplasms

INTRODUCTION

Breast cancer is the first high incidence of malignant tumor and the leading cause of death by cancer among female patients (1). Axillary lymph node (ALN) status assessment is of great significance to stage breast cancer and guides the treatment decision-making (2). Nowadays, sentinel lymph node biopsy (SLNB) has substituted for the ALN dissection (ALND) to assess the ALN metastasis in early-stage breast cancer patients (3). Despite a high risk that non-sentinel lymph nodes (NSLNs) metastasis may occur in patients with metastatic sentinel lymph nodes (SLNs) (4, 5), not all patients with a positive SLN would necessarily have a positive NSLN. Indeed, the Z0011 randomized clinical trial showed that only approximately 27.3% of patients with 1 or 2 positive SLNs had NSLN metastasis (6). Other studies showed that 32.1–63% of patients with positive SLNs had NSLNs metastasis, as confirmed by ALND following SLNB (4, 5). These results demonstrate that a considerable number of patients with positive SLN might have negative NSLN; these patients may suffer from overtreatment of ALND (7). Therefore, to avoid unnecessary ALND in a patient with positive SLN but negative NSLN, developing a method to predict the absence or presence of NSLN metastasis is desperately needed.

Previously, several clinicopathologic nomograms (Memorial Sloan Kettering Cancer Center, Mayo, Cambridge, Stanford, and Ljubljana) and scoring systems (Tenon, MD Anderson Cancer Center, and Saidi) have been established to predict the NSLN status (7–14). However, all these models were developed based on pathologic features of the SLN, which could only be obtained from invasive axillary procedures. In addition, except for the Ljubljana nomograms in which preoperative axillary US examination was used as the predictors (7), none of these models have used radiologic features from diagnostic imaging. To date, noninvasive magnetic resonance imaging (MRI) has been recommended as a sufficient tool to comprehensively evaluate ALN status before treatment (15). However, MRI mainly relies on the morphologic criteria to assess the status of the ALN, which showed high specificity but low sensitivity in identifying the ALN metastasis (16). Radiomics could quantify heterogeneity of inter-tumor and intra-tumor by extracting high-throughput data from MR images (17, 18). Previously, MRI-based radiomics of the primary breast cancer has been used to predict the ALN metastasis with an area under the curve (AUC) ranging from 0.81 to 0.92 in training and 0.74 to 0.90 in the validation datasets (19–22), and the SLN burden with a reported AUC of 0.82, 0.81, and 0.81 in the training, validation, and test dataset, respectively (23). However, whether MRI-based radiomics could be applied to predict the NSLN metastasis in breast cancer patients with positive SLNs remains to be determined.

In this study, a large cohort of patients with SLN-positive breast cancer was retrospectively included. Radiomic features of the primary breast tumor on pretreatment multiparametric MRI were extracted, and the MRI-based radiomics signature was constructed to predict the NSLN metastasis. In addition, predictive clinicopathologic features and MRI morphologic findings of breast tumors before treatment were identified to develop an integrative predictive MRI-clinical-radiomics nomogram. The purpose of this study was to develop an MRI-based radiomics model to predict the NSLN metastasis in breast cancer patients with positive SLNs.

MATERIALS AND METHODS

Patients and Study Design

This study was approved by the Institutional Review Board of Sun Yat-sen Memorial Hospital, Sun Yat-sen University, and the informed consent was waived because of the nature of the retrospective study. A total of 306 consecutive women with

pathologically confirmed primary breast carcinoma were collected from the hospital medical record system between April 2016 and September 2018. The patient enrollment workflow is shown in **Figure 1**. Patients were included if they (i) underwent multiparametric breast MRI examination before breast and axillary surgery; (ii) underwent SLNB and ALND with at least one pathologically positive SLN. The exclusion criteria were as follows: (i) chemotherapy, endocrine therapy, targeted therapy, or radiotherapy before surgery; (ii) recurrent breast malignant tumor; (iii) a history of ipsilateral breast lesion excision; (iv) distant metastasis; and (v) bilateral, multicentric, multifocal, or non-mass-type breast cancer. A total of 285 patients were included. According to the pathologic results of ALND, 285 patients were divided into two groups: the metastatic NSLN group in which at least one NSLN was metastasis (micrometastasis or macrometastasis) pathologically (n = 55) and the non-metastatic NSLN group (n = 230) in which none of NSLN was metastasis pathologically.

Clinicopathologic Characteristics

All patients were treated by surgery, namely, breast tumor resection, SLNB, and ALND. SLNB was performed by using the methylene blue technique, as previously described (24). The status of NSLN was identified by ALND and subsequent pathologic examination. The clinicopathologic data, namely, age, family history of breast cancer, palpable breast mass, clinical tumor staging, carcinoembryonic antigen (CEA) level, carbohydrate antigen 15-3 (CA 15-3) level, cytokeratin-19-fragment level, pathologic type of breast cancer, lymphovascular invasion, estrogen receptor (ER) status, progesterone receptor (PR) status, human epidermal growth factor receptor-2 (HER-2) status, Ki-67 status, the number of pathologically proved metastatic SLNs, and the number of pathologically proved metastatic ALNs were collected from the electronic medical record system and pathologic system. Clinical tumor staging was evaluated following the guidelines of the TNM staging system proposed by the American Joint Committee on Cancer (25). In addition, the ALN status determined by preoperative axillary ultrasound (US) examination or US-guided fine-needle aspiration biopsy (FNAB) was collected from the electronic medical record system. The presence of ALN metastasis on US was assessed according to the following abnormal morphologic features: lobulated or eccentric cortex, dislocated and/or absence of fatty hilum, eccentric or concentric thickening ≥2 mm, a cortex-to-hilum ratio ≥1, or a longitudinal axis-to-transverse axis ratio ≤2 (26). During US evaluation, the typical location of the SLN (i.e., axillary tail area) was paid special attention. A biopsy sample was obtained from the most suspicious ALN that showed the above abnormal morphologic characteristics (26).

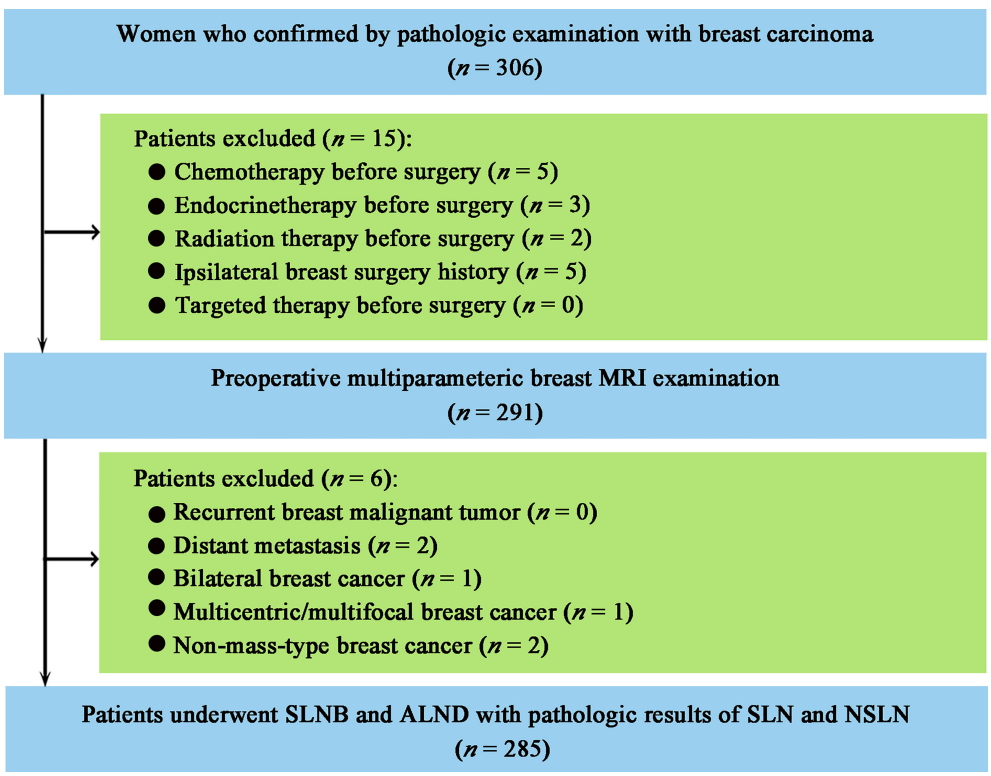


FIGURE 1 | Patient enrollment workflow. MRI, magnetic resonance imaging; SLNB, sentinel lymph node biopsy; ALND, axillary lymph node dissection; SLN, sentinel lymph node; NSLN, non-sentinel lymph node.

Multiparametric MRI Acquisition

MRI was performed on a 1.5 T MR scanner (Magnetom Avanto, Siemens Medical Solutions) with an 8-channel phased-array breast coil (Siemens Medical Solutions). The patients were placed in the prone position with a body parallel to the shoulders, and both breasts were naturally suspended in the coil. The sequences included axial T2-weighted imaging (T2WI), axial T1-weighted imaging (T1WI), axial diffusion-weighted imaging (DWI) with readout segmented echo planar imaging, followed by axial dynamic contrast-enhanced imaging (DCE), axial and coronal delayed contrast-enhanced T1WI (T1 + C). Two dynamic phases of DCE acquisition (40 phases with a temporal resolution of 8 s) were initially performed. And then, all patients underwent intravenous bolus injection of Gd-DTPA-BMA (Omniscan, GE Healthcare; dose = 0.1 mmol/kg body weight; flow rate = 3.5 ml/s) through a high-pressure contrast agent injector (Spectris, Medrad). The T1 + C images were obtained immediately after the DCE imaging was finished. The detailed acquisition parameters are shown in **Table 1**.

MRI Morphologic Analysis

Morphologic findings of MRI were assessed by two radiologists (ZY and YQ, with 12 and 7 years of clinical experience in breast MRI diagnosis, respectively) who knew breast cancer diagnosis but were blinded to other clinicopathologic information. All MRI sequences of each patient were available during the morphologic assessment. Any disagreement between the two radiologists was resolved by consultation of another senior radiologist (JS with 20 years of clinical experience in breast MRI diagnosis), and a final diagnosis was made by this senior radiologist. For morphologic analysis, MRI findings, namely, the quadrant of breast cancer, long diameter of breast cancer, presence of ALN metastasis, number of metastatic ALN, and short diameter of the largest ALN, were evaluated. The quadrant of breast cancer and the long diameter of breast cancer were measured on axial or coronal T1 + C image in which the primary tumor showed the largest section. All lymph nodes in the axilla were evaluated on axial and coronal T1 + C images. The ALN metastasis was assessed according to previously morphologic criteria as follows: the disappearance of hilum structure (27), lymphatic hilum displacement, eccentric cortical thickening, short diameter >1 cm, or the ratio of long to a short diameter less than 2 (28). The number of metastatic ALN was recorded. The short diameter of the largest ALN was measured on the axial T1 + C image.

Radiomic Feature Extraction

The flowchart and radiomics analysis workflow are shown in **Figure 2**. First, the primary breast cancer was segmented manually by investigator 1 (XZ, with 10 years of clinical experience in breast MRI diagnosis) to separately create a volume of interest (VOI) on DWI images, apparent diffusion coefficient (ADC) maps, T2WI images, and T1 + C images using the ITK-SNAP (version 3.6.0). Investigator 1 repeated the tumor segmentation in a randomized selecting dataset (n = 60) after 2 weeks, and investigator 2 (JH, with 3 years of clinical experience in breast MRI diagnosis) independently performed the segmentation in these 60 patients using the same method as that of investigator 1. Second, radiomic feature extraction was performed using the PyRadiomics toolkit (version 3.0.1) written in Python (version 3.8.3). All the segmented images were interpolated to normalize the spatial resolution in X, Y, and Z directions. For each patient, 1,595 radiomic features were extracted from the initial VOIs and the wavelet filtered, and intensity transformed DWI, ADC, T2WI, and T1 + C images. A total of 6,380 radiomic features were extracted from the primary breast tumors of these four sequences. Details of radiomic features are shown in **Supplementary Table 1**. Third, the intraclass correlation coefficients (ICCs) for the extraction of NSLN metastasis-related radiomic features were assessed by the reproducibility of intra-investigator (first segmentation of investigator 1 vs. second segmentation of investigator 1) and inter-investigator (first segmentation of investigator 1 vs. segmentation of investigator 2), respectively. A good agreement was considered when an ICC was greater than 0.75.

Development of Predictive Models

MRI morphologic findings, clinicopathologic characteristics, and MRI-based radiomic features were selected to develop three kinds of predictive models, namely, a radiomics signature and two integrative models. For the two integrative predictive models, one was the MRI-clinical nomogram where the independent predictors of MRI morphologic findings and predictive clinicopathologic characteristics were included; the other was the MRI-clinical-radiomics nomogram where the independent predictors of MRI morphologic findings,

TABLE 1 | Multiparametric MRI and acquisition parameters.

Sequence	TR/TE (ms)	FOV (mm)	Matrix	Acquisition time (s)	Slice gap (mm)	Fat suppression	Flip angle	Slice thickness (mm)	b value (s/mm^2)
T2WI	2,500/107	350 × 50	384 × 256	174	1	yes	111°	4	–
T1WI	6.86/2.39	350 × 350	384 × 256	117	1	yes	111°	4	–
DWI	5,400/119	350 × 350	128 × 128	165	1	yes	90°	4	0/800
DCE	4.95/2.28	360 × 360	384 × 224	332	0.8	yes	15°	1.6	–
T1 + C (Axial)	4.85/2.34	360 × 360	320 × 320	65	0.2	yes	5°	1.4	–
T1 + C (Coronal)	6.88/62.39	360 × 360	384 × 384	81	0.4	no	111°	2	–

TR, repetition time; TE, echo time; FOV, field of view; T2WI, T2-weighted imaging; T1WI, T1-weighted imaging; DWI, diffusion-weighted imaging; DCE, dynamic contrast-enhanced imaging; T1+C, contrast-enhanced T1-weighted imaging.

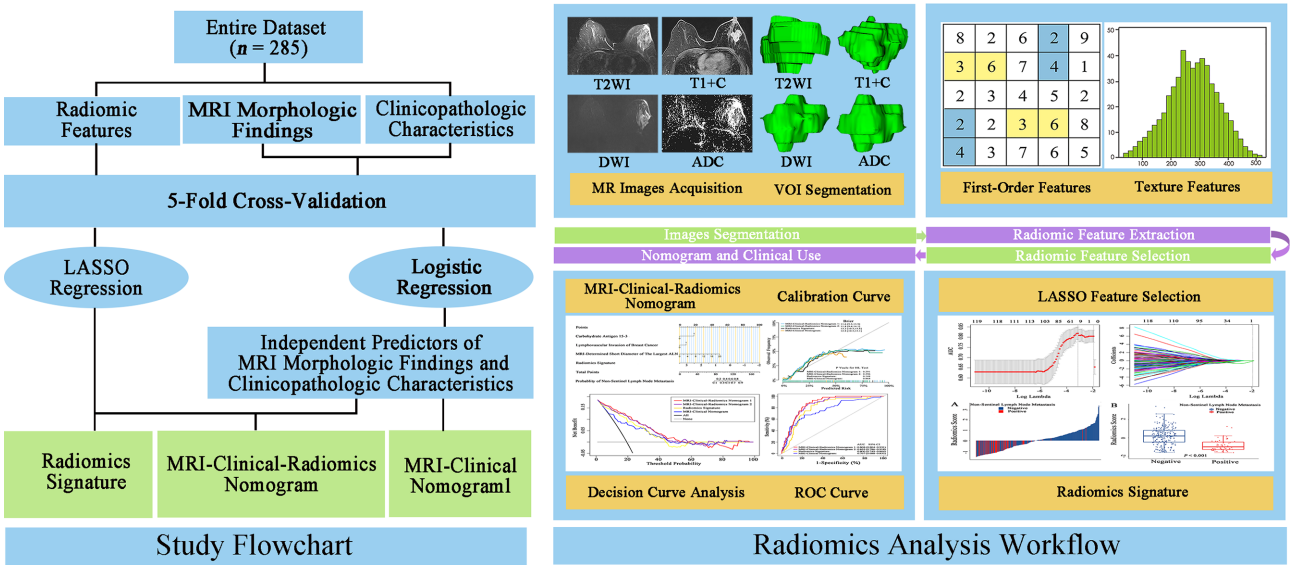


FIGURE 2 | Study flowchart and radiomics analysis workflow. The green rectangular boxes in the study flowchart represent three different non-sentinel lymph node predictive models, namely, radiomics signature, MRI-clinical-radiomics nomogram, and MRI-clinical nomogram. MRI, magnetic resonance imaging; LASSO, shrinkage and selection shrinkage and selection operator; T2WI, T2-weighted imaging; T1 + C, delayed contrast-enhanced T1-weighted imaging; DWI, diffusion-weighted imaging; ADC, apparent diffusion coefficient; VOI, volume of interest.

predictive clinicopathologic characteristics, and radiomics signature were included. To construct integrative predictive models, the Mann–Whitney *U* test was used to compare the MRI morphologic findings and clinicopathologic characteristics between the metastatic NSLN group and the non-metastatic NSLN group. Multivariable logistic regression was then applied to select independent predictors of NSLN metastasis from the MRI morphologic findings and clinicopathologic characteristics. For radiomics analysis, the Mann–Whitney *U* test was performed to select the statistically significant radiomic features between metastatic NSLN group and non-metastatic NSLN group, followed by the least absolute shrinkage and selection operator (LASSO) regression to identify the NSLN metastasis-related radiomic features. The radiomics signature was presented as a radiomics score and constructed by combining the NSLN metastasis-related radiomic features, weighted by the corresponding coefficients of LASSO regression. To determine the optimal independent predictors in each model, a 5-fold cross-validation (CV) analysis was performed by training and testing five separate models to select the most robust predictors (29). For the 5-fold CV analysis, the entire dataset was randomly divided into five subsets, four subsets used for training and another one subset used for testing. This process was repeated five times and five training CV folds and five internal validation CV folds were obtained. The receiver operating characteristic (ROC) curve analysis was used to assess the predictive performance of each model. The area under the curve (AUC) was calculated and compared among different models by the DeLong test (30).

Performance and Usefulness of Predictive Models

The most robust predictors in the radiomics signature, the MRI-clinical nomogram, and the MRI-clinical-radiomics nomogram selected by the 5-fold CV analysis were used to construct the final predictive models. The performances of the final models of the radiomics signature, MRI-clinical nomogram, and MRI-clinical-radiomics nomogram were determined by the ROC curves analysis in the entire dataset. Their AUCs were compared by the Delong test. The calibration of the final radiomics signature, MRI-clinical nomogram, and MRI-clinical-radiomics nomograms was evaluated using the calibration curves with the Hosmer–Lemeshow test. In addition, the decision curve analysis (DCA) was conducted respectively to assess the clinical use of the final predictive models presenting as the net benefit at different threshold probabilities (31).

Statistical Analysis

Descriptive statistics were summarized as median (quartile range) for continuous variables or as frequencies with percentages for categorical variables. The continuous variables were compared between different groups by using the *t*-test. The categorical variables were compared between different groups using Pearson's χ^2 or Fisher exact test. The comparison of continuous and categorical variables and ICCs for the feature extraction of intra- and inter-investigator was conducted on SPSS 25. The Mann–Whitney *U* test, LASSO regression, multivariable logistic regression, 5-fold CV, ROC analysis with AUC values calculating, calibration curves, and DCA were

performed using the R software (version 4.0.1). P <0.05 was considered statistically significant.

RESULTS

Clinicopathologic Characteristics and MRI Morphologic Findings

The clinicopathologic characteristics and MRI morphologic findings of 55 patients with metastatic NSLN and 230 patients without metastatic NSLN are summarized in **Table 2**. The time between the breast MRI and surgery ranged from 1 to 12 days, with a median of 5 days. There were significant differences in CA 15-3 status (P <0.001), pathologic types of breast cancer (P = 0.005), lymphovascular invasion (P = 0.001), MRI-determined presence of ALN metastasis (P = 0.018), and MRI-determined short diameter of the largest ALN (P <0.001) between metastatic and non-metastatic NSLN groups. Most of the patients (272 of 285, 95.4%) had preoperative US results of ALN status. Among these 272 patients, 246 patients had negative results on axillary US examination, and 26 patients had positive results on axillary US examination but negative results US-guided FNAB. Based on the entire dataset of 285 patients, multivariable logistic regression showed that one MR-determined finding (MRI-determined short diameter of the largest ALN), and two clinicopathologic characteristics (CA 15-3 and lymphovascular invasion of breast cancer) were the independent predictors of the NSLN metastasis (**Table 3**). Based on the dataset of 272 patients having preoperative axillary US results, US-reported ALN status was an independent predictor of the NSLN metastasis (**Table 3**). Other MRI morphologic findings and clinicopathologic characteristics were not selected as the independent predictors of the NSLN metastasis (**Supplementary Table 2**).

Radiomic Feature Extraction

A total of 6,380 radiomic features were extracted from DWI, ADC, T2WI, and T1 + C images of the primary breast tumors for each patient. The ICCs of these radiomic features ranged from 0.797 to 0.981 and 0.773 to 0.976 for intra- and inter-investigator segmentation, respectively, indicating a good reproducibility for radiomic feature extraction.

Development of Different Predictive Models

For the radiomics signature, the MRI-clinical nomogram and the MRI-clinical-radiomics nomogram, the selected independent predictors and their AUCs in each training and internal validation CV fold of the 5-fold CV analysis are shown in **Table 4**. The AUCs ranged from 0.774 (95% CI, 0.675–0.873) to 0.847 (95% CI, 0.757–0.937) in the training CV fold and from 0.770 (95% CI, 0.654–0.886) to 0.820 (95% CI, 0.749–0.891) in the internal validation CV fold for the radiomics signature, from 0.758 (95% CI, 0.662–0.854) to 0.824 (95% CI, 0.729–0.919) in the training CV fold and from 0.720 (95% CI, 0.598–0.843) to 0.762 (95% CI, 0.685–0.840) in the internal validation CV fold for the MRI-clinical nomogram, and from 0.850 (95% CI, 0.764–0.936) to 0.932 (95% CI, 0.871–0.993) in the training CV fold and from 0.843 (95% CI, 0.745–0.943) to 0.904 (95% CI, 0.849–0.959) in the validation CV fold for the MRI-clinical-radiomics nomogram. The comparisons of the performances among different predictive models in each training CV fold and internal validation CV fold are shown in **Table 5**. The AUCs of the MRI-clinical-radiomics nomogram were higher than those of the radiomics signature (P ≤0.001–0.059) and the MRI-clinical nomogram (P = 0.003–0.050). Although Fold 1 model of MRI-clinical-radiomics nomogram appeared to perform the best in training and also validation and in comparison with other models, the most robust variables selected by each CV fold were four features, namely, an MRI morphologic finding (short diameter of the largest ALN), two clinicopathologic features (CA 15-3 and lymphovascular invasion of breast cancer), and a texture feature of DWI (DWI_*original_GLDM_Small_Dependence_High_GrayLevel_Emphasis*), which were considered as the optimal independent predictors and used for final model construction.

Performance and Clinical Usefulness of Different Predictive Models

The final model of the MRI-clinical-radiomics nomogram is shown in **Figure 3A**. ROC analysis showed that the final model of the MRI-clinical-radiomics nomogram had an AUC of 0.868, which was significantly higher than that of radiomics signature (0.868 vs. 0.806, P <0.001) and MRI-clinical nomogram (0.868 vs. 0.761, P <0.001) (**Figure 3B**). As the CA 15-3 is not a standard of care for prediction of NSLN metastasis, the MRI-clinical-radiomics nomogram, namely, an MRI morphologic finding (short diameter of the largest ALN), a clinicopathologic features (lymphovascular invasion of breast cancer), and a texture feature of DWI (DWI_*original_GLDM_Small_Dependence_High_GrayLevel_Emphasis*) but without CA 15-3 were also constructed. This MRI-clinical-radiomics nomogram had an AUC of 0.852, which was significantly higher than those of radiomics signature (0.852 vs. 0.806, P = 0.016) and MRI-clinical nomogram (0.852 vs. 0.761, P = 0.007) in predicting NSLN metastasis in the entire dataset (**Figure 3C**). The calibration curves (**Figure 3D**) indicated an excellent calibration capability of the MRI-clinical-radiomics nomogram with or without CA 15-3, and the Hosmer–Lemeshow test showed a P-value of 0.291 and 0.296, respectively, suggesting a favorable calibration in terms of the agreement between the predicted risk and actual probability for NSLN metastasis. The decision curve analysis showed that if the threshold probability is between 0.1 and 0.6, using the MRI-clinical-radiomics nomograms with or without CA 15-3 to predict NSLN metastasis adds more benefit than either treating-all or treating-no patients (**Figure 4**). Additionally, the radiomics score of each patient is shown in **Figure 5A**. The radiomics scores in the non-metastatic NSLN group were higher than those in the metastatic NSLN group (0.210 [−0.471, 0.822] vs. −0.980 [−1.270, −0.401], P <0.001). The comparison of radiomics scores between the two groups is shown in **Figure 5B**.

TABLE 2 | Clinicopathologic characteristics and MRI morphologic findings of patients with and without metastatic NSLN.

Characteristic	Non-metastatic NSLN (*n* = 230)	Metastatic NSLN (*n* = 55)	*P*-value
Age (median, quartile range), years	49 (44, 58)	50 (45, 59)	0.337•
Family history of breast cancer			0.578Δ
No	227 (98.7)	54 (98.2)	
Yes	3 (1.3)	1 (1.8)	
Palpable breast mass			0.028◊
No	215 (93.5)	46 (83.6)	
Yes	15 (6.5)	9 (16.4)	
Clinical tumor staging			0.100◊
T1	117 (50.9)	21 (38.2)	
T2	113 (49.1)	34 (61.8)	
CEA#			0.738Δ
Negative	219 (95.2)	52 (94.5)	
Positive	11 (4.8)	3 (5.5)	
CA 15-3#			<0.001◊*
Negative	218 (94.8)	42 (76.4)	
Positive	12 (5.2)	13 (23.6)	
CYFR 21-1#			0.063◊
Negative	171 (74.3)	34 (61.8)	
Positive	59 (25.7)	21 (38.2)	
Pathologic type of breast cancer			0.005Δ*
IDC	189 (82.2)	44 (80.0)	
ILC	3 (1.3)	5 (9.1)	
Others†	38 (16.5)	6 (10.9)	
Lymphovascular invasion			0.001◊*
No	181 (78.7)	31 (56.4)	
Yes	49 (21.3)	24 (43.6)	
ER status			0.453
Negative	48 (20. 9)	9 (16.4)	
Positive	182 (79.1)	46 (83.6)	
PR status			0.546◊
Negative	81 (35.2)	17 (30.9)	
Positive	149 (64.8)	38 (69.1)	
HER-2 status			0.248Δ
Negative	3 (1.3)	2 (3.6)	
Positive	227 (98.7)	53(96.4)	
Ki-67 status			0.354◊
Negative (<14%)	46 (20)	8 (14.5)	
Positive (≥14%)	184 (80.0)	47 (85.5)	
MRI-determined quadrant of breast cancer			0.154Δ
Central quadrant	10 (4.3)	1 (1.8)	
Outer-upper quadrant	83 (36.1)	26 (47.8)	
Outer-lower quadrant	42 (18.3)	14 (25.5)	
Upper-inner quadrant	64 (27.8)	8 (14.5)	
Lower-inner quadrant	31 (13.5)	6 (10.9)	
MRI-determined long diameter of breast cancer (median, quartile range), mm	19.75 (15.1, 25.7)	22.2 (16.6, 29)	0.074•
MRI-determined presence of ALN metastasis			0.018◊
No	218 (94.8)	46 (83.6)	
Yes	12 (5.2)	9 (16.4)	

(Continued)

TABLE 2 | Continued

Characteristic	Non-metastatic NSLN (*n* = 230)	Metastatic NSLN (*n* = 55)	*P*-value
MRI-determined number of metastatic ALN			0.077[Δ]
<1	218 (94.8)	46 (83.6)	
<2	8 (3.5)	6 (10.9)	
≤3	4 (1.7)	3 (5.5)	
MRI-determined short diameter of the largest ALN (median, quartile range), mm	3.60 (2.7,5.3)	5.7 (3.8,8.9)	< 0.001[•]*
US-reported ALN status[‡]			0.041[◊]
Negative	202 (92.2)	44 (83)	
Positive	17 (7.8)	9 (17)	

Numbers in the parentheses were presented as percentages. NSLN, non-sentinel lymph node; CEA, carcinoembryonic antigen; CA 15-3, carbohydrate antigen 15-3, CYFR 21-1, cytokeratin-19-fragment; IDC, invasive ductal carcinoma, ILC, invasive lobular carcinoma; ER, estrogen receptor, PR, progesterone receptor; HER-2, human epidermal growth factor receptor-2; MRI, magnetic resonance imaging; mm, millimeter; ALN, axillary lymph node; US, ultrasound.

[†]Others include intraductal papillary carcinoma, ductal carcinoma in situ, lobular carcinoma in situ, neuroendocrine carcinoma, mucinous carcinoma.

[‡]Data was based on 272 patients who underwent US examination in Sun Yat-sen Memorial Hospital, Sun Yat-sen University.

[#]Laboratory analysis of CEA, CA 15-3, and CYFR 21-1 were performed through blood tests within 1 week before surgery. CEA level ≤5 ng/ml, CA 15-3 level ≤25 U/ml, and CYFR 21-1 level <3.3 ng/ml were set as the normal ranges.

[•]Continuous variables were compared by using the Nonparametric test.

[Δ]Categorical variables were compared by using the Fisher exact test.

[◊]Categorical variables were compared by using Pearson's χ^2 test.

**P-value <0.05.*

TABLE 3 | Multivariate logistic regression analysis of predictors of NSLN metastasis prediction in patients with breast cancer based on entire dataset.

Variables	β	Odds ratio (95% CI)[Δ]	*P*-value
MRI-determined short diameter of the largest ALN	0.342	1.408 (1.195–1.658)	<0.001*
US-reported ALN status[‡]	1.829	6.227 (1.871–20.727)	0.003*
CA 15-3	2.006	7.436 (2.237–24.719)	0.001*
Lymphovascular invasion of breast cancer	1.612	5.012 (2.213–11.355)	<0.001*

CI, confidence interval; MRI, magnetic resonance imaging; ALN, axillary lymph node; CA 15-3, carbohydrate antigen 15-3; US, ultrasound.

[Δ]Data in parentheses are 95% confidence intervals.

[‡] Data was based on 272 patients who had preoperative axillary US results.

**P-value < 0.05.*

TABLE 4 | Five-fold cross-validation analysis of different predictive models.

Predictive Model	Fold Sequence	Selected Variable	AUC (95% CI) in training CV fold	AUC (95% CI) in internal validation CV fold
Radiomics signature	Fold 1	DWI_*Original GLDM Small Dependence High Gray Level Emphasis* ADC_*Wavelet LLH First order 10 Percentile* ADC_*Wavelet HHH NGTDM Contrast* ADC_*Wavelet HHL GLDM Small Dependence Low Gray Level Emphasis*	0.837 (0.755–0.922)	0.820 (0.749–0.891)
	Fold 2	DWI_*Original GLDM Small Dependence High Gray Level Emphasis*	0.774 (0.675–0.873)	0.794 (0.673–0.915)
	Fold 3	DWI_*Original GLDM Small Dependence High Gray Level Emphasis*	0.806 (0.707–0.906)	0.787 (0.676–0.899)
	Fold 4	DWI_*Original GLDM Small Dependence High Gray Level Emphasis* ADC_*Wavelet LLH First order 10 Percentile* ADC_*Wavelet HHH NGTDM Contrast* ADC_*Wavelet HHL GLDM Small Dependence Low Gray Level Emphasis*	0.847 (0.757–0.937)	0.770 (0.654–0.886)
	Fold 5	DWI_*Original GLDM Small Dependence High Gray Level Emphasis*	0.821 (0.729–0.912)	0.787 (0.676–0.899)
MRI-clinical nomogram	Fold 1	CA 15-3 Lymphovascular invasion MRI-determined short diameter of the largest ALN	0.758 (0.662–0.854)	0.762 (0.685–0.840)
	Fold 2	CA 15-3 Lymphovascular invasion MRI-determined short diameter of the largest ALN	0.772 (0.673–0.872)	0.745 (0.734–0.950)
	Fold 3	CA 15-3 Lymphovascular invasion MRI-determined short diameter of the largest ALN	0.779 (0.675–0.883)	0.745 (0.628–0.863)
	Fold 4	CA 15-3 CYFR 21-1 Lymphovascular invasion Pathologic type of breast cancer MRI-determined short diameter of the largest ALN	0.824 (0.729–0.919)	0.720 (0.598–0.843)
	Fold 5	CA 15-3 CYFR 21-1 Lymphovascular invasion Pathologic type of breast cancer MRI-determined short diameter of the largest ALN MRI BI-RADS	0.787 (0.690–0.884)	0.745 (0.628–0.863)
MRI-clinical-radiomics nomogram	Fold 1	CA 15-3 Lymphovascular invasion MRI-determined short diameter of the largest ALN DWI_*Original GLDM Small Dependence High Gray Level Emphasis* ADC_*Wavelet LLH First order 10 Percentile* ADC_*Wavelet HHH NGTDM Contrast* ADC_*Wavelet HHL GLDM Small Dependence Low Gray Level Emphasis*	0.906 (0.839–0.973)	0.904 (0.849–0.959)
	Fold 2	CA 15-3 Lymphovascular invasion MRI-determined short diameter of the largest ALN DWI_*Original GLDM Small Dependence High Gray Level Emphasis*	0.850 (0.764–0.936)	0.898 (0.808–0.987)
	Fold 3	CA 15-3 Lymphovascular invasion MRI-determined short diameter of the largest ALN DWI_*Original GLDM Small Dependence High Gray Level Emphasis*	0.875 (0.790–0.959)	0.843 (0.745–0.943)
	Fold 4	CA 15-3 CYFR 21-1 Lymphovascular invasion Pathologic type of breast cancer MRI-determined short diameter of the largest ALN	0.929 (0.864–0.994)	0.886 (0.778–0.974)

(Continued)

TABLE 4 | Continued

Predictive Model	Fold Sequence	Selected Variable	AUC (95% CI) in training CV fold	AUC (95% CI) in internal validation CV fold
		DWI_*Original GLDM Small Dependence High Gray Level Emphasis*		
		ADC_*Wavelet LLH First order 10 Percentile*		
		ADC_*Wavelet HHH NGTDM Contrast*		
		ADC_*Wavelet HHL GLDM Small Dependence Low Gray Level Emphasis*		
	Fold 5	CA 15-3	0.932 (0.871–0.993)	0.843 (0.745–0.943)
		CYFR 21-1		
		Lymphovascular invasion		
		Pathologic type of breast cancer		
		MRI-determined short diameter of the largest ALN		
		MRI BI-RADS		
		DWI_*Original GLDM Small Dependence High Gray Level Emphasis*		

AUC, area under the curve; CI, confidence interval; CV, cross-validation; DWI, diffusion-weighted imaging; GLDM, Gray Level Dependence Matrix; ADC, apparent diffusion coefficient; NGTDM, Neighbouring Gray Tone Difference Matrix; MRI, magnetic resonance imaging; ALN, axillary lymph node; CA 15-3, carbohydrate antigen 15-3; CYFR 21-1, Cytokeratin-19-fragment; BI-RADS, Breast imaging-reporting and data system.

TABLE 5 | Comparisons of predictive performances of different predictive models in 5-fold cross-validation analysis.

Fold Sequence	*P*-Values for Comparison of AUCs in Training CV Fold		*P*-Values for Comparison of AUCs in Internal Validation CV Fold	
	MRI-Clinical-Radiomics Nomogram vs. MRI-Clinical Nomogram	MRI-Clinical-Radiomics Nomogram vs. Radiomics Signature	MRI-Clinical-Radiomics Nomogram vs. MRI-Clinical Nomogram	MRI-Clinical-Radiomics Nomogram vs. Radiomics Signature
Fold 1	0.017*	0.001*	0.007*	0.001*
Fold 2	0.006*	0.059	0.050	0.006*
Fold 3	0.015*	0.044*	0.042*	0.037*
Fold 4	0.004*	0.007*	0.007*	0.007*
Fold 5	0.003*	0.001*	0.042*	0.037*

MRI, magnetic resonance imaging; AUC, area under the curve; MRI, magnetic resonance imaging.
**P-value < 0.05.*

Additionally, since axillary US is the most robust axillary assessment tool, the 5-fold cross-validation analysis, where the US-reported ALN status was also included as a variable, was performed in 272 patients with negative axillary US examination (with or without FNAB). The results showed that the US-reported ALN status was not a strong clinical predictor (**Supplementary Table 3**). Based on these 272 patients, the MRI-clinical-radiomics nomograms with CA 15-3 and without CA 15-3 showed an AUC of 0.861 and 0.844 in predicting NSLN metastasis, respectively (**Figure 6**). After the US-reported ALN status was added, the MRI-clinical-radiomics nomograms with CA 15-3 and without CA 15-3 had an AUC of 0.862 and 0.824 in predicting NSLN metastasis in this subcohort (**Figure 7**).

DISCUSSION

In this study, we developed two MRI-clinical-radiomics nomograms that incorporate one MR-determined finding (short diameter of the largest ALN), one or two clinicopathologic characteristics (i.e. lymphovascular invasion of breast cancer or CA 15-3 plus lymphovascular invasion of breast cancer), and the radiomics signature consisting of one DWI radiomic feature based on the entire dataset of 285 patients. These two MRI-clinical-radiomics nomograms demonstrated robust and high predictive performance (AUC = 0.868 and 0.852), which were both better than the radiomics signature alone and MRI-clinical nomogram. The developed MRI-clinical-radiomics nomograms can serve as novel and easy-to-popularized tools to predict axillary NSLN metastasis in breast cancer patients with positive SLNs.

Invasive ALND is associated with potential postoperative morbidities such as pain, numbness, lymphedema, restricted arm movements, and high risk of infection (32, 33), which can be omitted for those patients at extremely low risk of NSLN metastasis (2). Previously, various clinicopathologic models, such as Memorial Sloan Kettering Cancer Center, Mayo, Cambridge, Stanford, and Ljubljana nomograms, were constructed to predict the NSLN metastasis with reported AUCs range from 0.74 to 0.84 (8–12). It is noted that these predictive models required the pathologic results both from the primary tumor and from the SLN, i.e., the SLN size, the number of positive SLN, and the proportion of positive SLN to all dissected SLN. This information is available only after the invasive SLNB (8–12). In our study, only the preoperative imaging data, clinical details, and pathologic information of

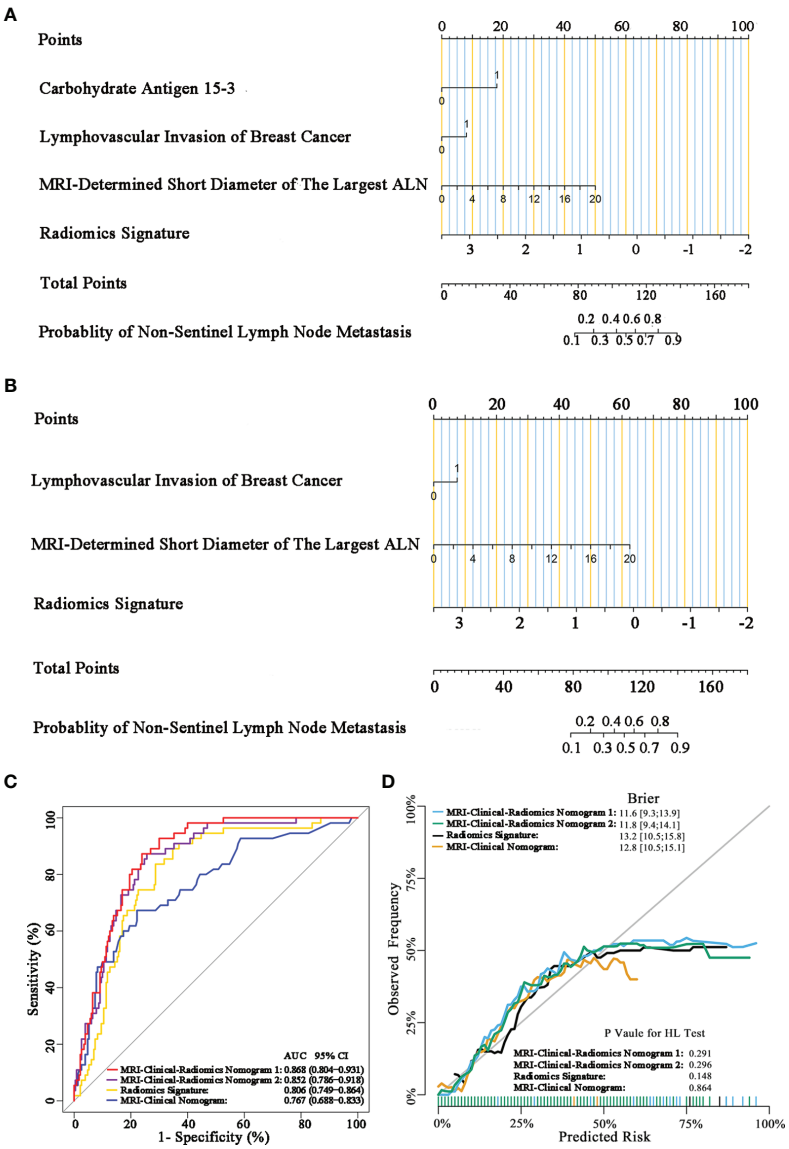


FIGURE 3 | MRI-clinical-radiomics nomograms, receiver operating characteristic (ROC) curves, and calibration curves of predictive models. MRI-clinical-radiomics nomogram **(A)** developed in the entire dataset incorporates one MRI-determined morphologic finding, two clinicopathologic characteristics (lymphovascular invasion of breast cancer plus CA 15-3), and radiomics signature. MRI-clinical-radiomics nomogram **(B)** developed in the entire dataset incorporates one MRI-determined morphologic finding, one clinicopathologic characteristics (lymphovascular invasion of breast cancer alone), and radiomics signature. ROC curves of the radiomics signature, MRI-clinical nomogram, and MRI-clinical-radiomics nomograms with CA 15-3 (MRI-Clinical-Radiomics Nomogram 1) and without CA 15-3 (MRI-Clinical-Radiomics Nomogram 2) in the entire dataset **(C)**. Calibration curves of the radiomics signature, MRI-clinical nomogram, and MRI-clinical-radiomics nomograms in the entire dataset **(D)**. ALN, axillary lymph node; AUC, area under the curve; CI, confidence interval; HL, Hosmer–Lemeshow.

the primary breast tumor obtained from biopsy were applied to develop a predictive model. Comparatively, our predictive model may be preferable in clinical practice as it can predict NSLN status without the trauma of the axilla resulting from the SLNB.

To date, a few MRI-based radiomics nomograms have been established for predicting the presence of ALN metastasis, disease-free survival, neoadjuvant chemotherapy efficacy, and tumor microenvironment status in breast cancer patients (19, 34–36). Previously, a Ljubljana nomogram was constructed using the preoperative axillary US features and clinicopathologic information to predict the likelihood of NSLN metastases, with the reported AUCs ranging from 0.75 to 0.79 (7). MRI-based radiomics nomogram to predict the axillary NSLNs metastasis in breast cancer patients with positive SLNs remains a scarcity of data. Dong et al. reported that breast cancer-specific radiomics features extracted from T2WI and DWI images could improve the performance in predicting SLN metastasis, with an AUC of 0.863 in the training set and 0.805 in the validation set (21). In addition, a

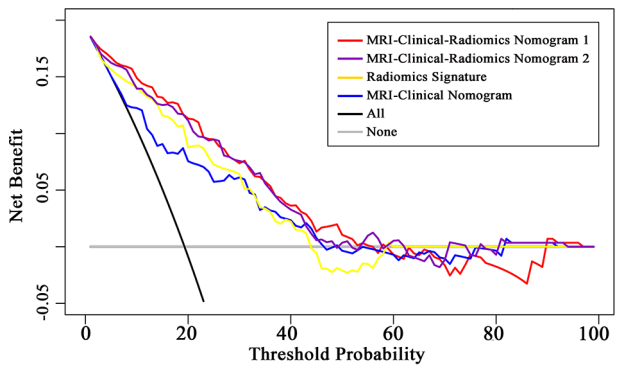


FIGURE 4 | Decision curve analysis (DCA) of the radiomics signature, MRI-clinical nomogram, and MRI-clinical-radiomics nomograms with CA 15-3 (MRI-Clinical-Radiomics Nomogram 1) and without CA 15-3 (MRI-Clinical-Radiomics Nomogram 2). The x-axis and y-axis represent the threshold probability and net benefit, respectively. The gray line and black line represent the hypothesis that all patients and no patient had NSLN metastasis, respectively. The threshold probability is where the expected benefit of treatment is equal to the expected benefit of avoiding treatment. The decision curves in the validation dataset showed that if the threshold probability is between 0.1 and 0.6, using the MRI-clinical-radiomics nomograms to predict non-sentinel lymph node metastasis add more benefit than treating all or treating no patients.

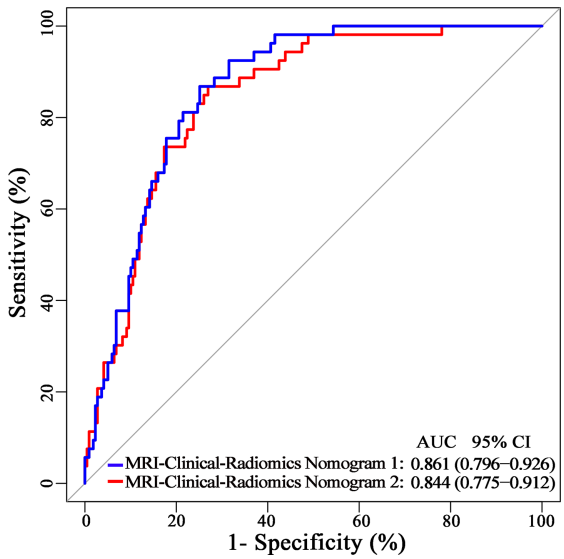


FIGURE 6 | Receiver operating characteristic curves of the MRI-clinical-radiomics nomograms with CA 15-3 (MRI-Clinical-Radiomics Nomogram 1) and without CA 15-3 (MRI-Clinical-Radiomics Nomogram 2) in predicting non-sentinel lymph node metastasis based on 272 patients with negative axillary US examination.

T2WI and DWI images-based radiomics predictive model could be utilized for preoperative stratification of the SLN low- and heavy-burden in breast cancer patients, yielding an AUC of 0.82, 0.81, and 0.81 in the training, validation, and test dataset, respectively (23). These studies indicated the potential of T2WI- and DWI-based radiomics in predicting the NSLN metastasis. In our study, radiomic features of multiparametric MRI, namely, T2WI, DWI, ADC, and T1 + C were extracted. The 5-fold CV analysis showed that one radiomic feature from DWI (DWI_*original_GLDM_Small_Dependence_High_GrayLevel_Emphasis*) ranged from 0.774 to 0.847 in the training CV fold and from 0.770 to 0.820 in the internal validation CV cohort. Moreover, DWI_*original_GLDM_Small_Dependence_High_GrayLevel_Emphasis* was a consistently selected variable during the 5-fold CV analysis, suggesting that this radiomic feature from DWI was a robust variable. As such, it was selected as the optimal predictor incorporated into the final predictive models. The final model of the one DWI feature-based radiomics signature had a favorable AUC of 0.806 in the entire cohort. This result suggested that the predictive capacity of radiomics features from DWI may be better than the radiomics features extracted from other sequences for predicting the NSLN metastasis. Moreover, this one feature-based radiomics signature might be more convenient for clinical use since fewer reproducible radiomic features imply better reproducibility (37).

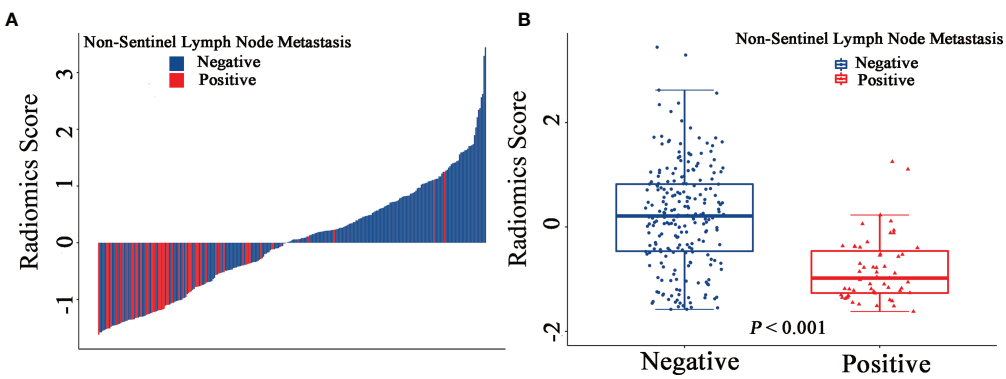


FIGURE 5 | Waterfall plots show the distribution of radiomic feature and non-sentinel lymph node metastasis status for each patient in the entire dataset **(A)**. Boxplots of the radiomic score in the entire dataset **(B)**.

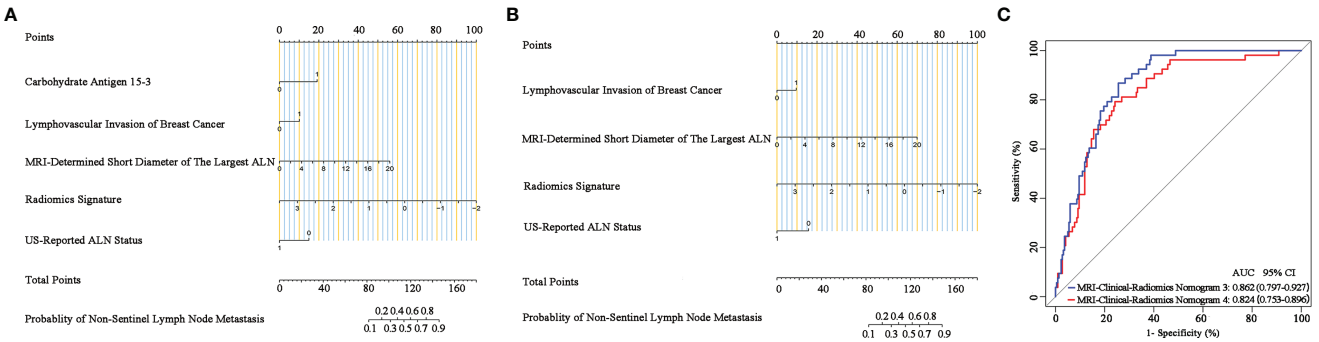


FIGURE 7 | Nomograms, receiver operating characteristic (ROC) curves of the US-reported ALN status-incorporated MRI-clinical-radiomics predictive models with CA 15-3 (MRI-Clinical-Radiomics Nomogram 3) and without CA 15-3 (MRI-Clinical-Radiomics Nomogram 4) in predicting non-sentinel lymph node metastasis based on 272 patients with negative axillary US examination. MRI-clinical-radiomics nomogram 3 **(A)** incorporates one MRI-determined morphologic finding, three clinicopathologic characteristics (lymphovascular invasion of breast cancer, CA 15-3 plus US-reported ALN status), and radiomics signature. MRI-clinical-radiomics nomogram 4 **(B)** incorporates one MRI-determined morphologic finding, two clinicopathologic characteristics (lymphovascular invasion of breast cancer plus US-reported ALN status), and radiomics signature. ROC curves **(C)** of the MRI-Clinical-Radiomics Nomogram 3 and MRI-Clinical-Radiomics Nomogram 4 in predicting non-sentinel lymph node metastasis based on 272 patients with negative axillary US examination. ALN, axillary lymph node; AUC, area under the curve; CI, confidence interval.

To further improve the predictive performance of radiomics signature, clinicopathologic information and MRI-determined morphologic findings were also assessed and incorporated to build an integrative radiomics-based predictive model in our study. Besides the radiomics signature, one MRI morphologic finding (short diameter of the largest ALN), and two clinicopathologic characteristics, including CA 15-3, lymphovascular invasion of breast cancer, were identified as the independent predictors by multivariable logistic regression for NSLN metastasis. The final model of the MRI-clinical-radiomics nomogram incorporating these predictors showed a higher performance than that of the radiomics signature (AUC, 0.868 vs. 0.806, P <0.001) and the MRI-clinical nomogram (0.868 vs. 0.761, P <0.001) in the entire dataset. In addition, the MRI-clinical-radiomics nomogram without CA 15-3 incorporated also showed a higher performance than those of the radiomics signature (AUC, 0.852 vs. 0.806, P = 0.016) and the MRI-clinical nomogram (AUC, 0.852 vs. 0.761, P = 0.007) in the entire dataset. It is seemingly that the MRI-clinical-radiomics nomograms developed in our study may serve as a preferable approach to predicting NSLN status in patients with SLN metastasis but without NSLN metastasis. Notably, the MRI-clinical-radiomics nomograms developed in our study also did not need pathologic features that should be obtained from invasive SLNB.

Our study had several limitations. First, the dataset used in our study was retrospectively collected from one center, and no independent external dataset was available for validation, which may limit the generalizability of the radiomics-based nomogram. Further multicenter studies with a larger sample size are needed to acquire high-level evidence for the clinical application of our predictive nomogram. Second, 272 patients (95.4%) underwent preoperative US scan of ALN. Unfortunately, the results of axillary US examination in the remaining 13 patients were not available in our hospital database. This might result in slightly higher than expected SLN involvement in the entire cohort. The accuracy of NSLN prediction could be affected for the constructed predictive models. Third, the proportion of the patients with metastatic NSLN enrolled in our study was relatively small. In our study, the 5-fold CV analysis was used to select the optimal variables for the development of predictive models, as previously reported (29). Forth, manual segmentation of tumors in our study was time- and labor-consuming, which could be improved by a more automatic segmentation approach with the assistance of artificial intelligence in the future. Fifth, the radiomics signature was built based on the radiomic features extracted from primary tumors but not the ALNs. However, it is ambiguous to identify the target ALN for radiomics feature extraction because it has a great challenge to match the ALNs on pathologic examination with the lymph nodes shown on preoperative axillary MRI. Sixth, non-mass-like, multicentric, and multifocal tumors were excluded, which may limit the generalizability of our results. However, it was a great challenge to delineate the boundary of non-mass-like lesions precisely on MR images. In addition, a potential possibility that a heavy burden of axillary NSLN metastasis in patients with multicentric and multifocal tumors may lead to a bias for the patient selection.

In conclusion, two MRI-clinical-radiomics nomograms were developed in our study. The proposed integrative MRI-clinical-radiomics nomograms was one feature-based radiomics signature with one MRI-determined morphologic finding, and one or two clinicopathologic characteristic incorporated, which showed high performance in predicting the axillary NSLN metastasis in patients with SLN positive breast cancer. These MRI-clinical-radiomics nomograms can serve as novel tools to predict axillary NSLN status, which may help avoid unnecessary invasive procedures on the axilla, i.e., ALND, in breast cancer patients with positive SLN but negative NSLN.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, and the informed consent was waived because of the nature of the retrospective study. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

All authors conceived and execution of this study or analysis of the study data. YQ, XZ, and JS designed the study. YQ, XZ, ZW, SW, DW, HL, JM, GD, XT, RZ, JH, and LH participated in the collection of the clinical information and data analysis. ZW and ZY did the statistical analysis. XZ and JS provided critical comments and suggestions and revised the manuscript. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This work was supported by the Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2017), the National Natural Science Foundation of China (82102130, 8210071257), the Key Areas Research and Development Program of Guangdong (2019B020235001), the Natural Science Foundation of Guangdong Province (2021A1515010385), the Medical Artificial Intelligence Project of Sun Yat-Sen Memorial Hospital (YXRGZN201905), and the Tianshan Youth Project of Xinjiang Uyghur Autonomous Region (2019Q144).

ACKNOWLEDGMENTS

We thank SY Xu and MJ Fang for their kind help for consultation of statistical method.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.811347/full#supplementary-material

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, 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. doi: 10.3322/caac.21660
2. Chang JM, Leung JWT, Moy L, Ha SM, Moon WK. Axillary Nodal Evaluation in Breast Cancer: State of the Art. *Radiology* (2020) 295:500–15. doi: 10.1148/radiol.2020192534
3. Lyman GH, Somerfield MR, Bosserman LD, Perkins CL, Weaver DL, Giuliano AE. Sentinel Lymph Node Biopsy for Patients With Early-Stage Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline Update. *J Clin Oncol* (2017) 35:561–4. doi: 10.1200/JCO.2016.71.0947
4. Turner RR, Chu KU, Qi K, Botnick LE, Hansen NM, Glass EC, et al. Pathologic Features Associated With Nonsentinel Lymph Node Metastases in Patients With Metastatic Breast Carcinoma in a Sentinel Lymph Node. *Cancer* (2000) 89:574–81. doi: 10.1002/1097-0142(20000801)
5. Maimaitiaili A, Wu D, Liu Z, Liu H, Muyiduli X, Fan Z. Analysis of Factors Related to non-Sentinel Lymph Node Metastasis in 296 Sentinel Lymph Node-Positive Chinese Breast Cancer Patients. *Cancer Biol Med* (2018) 15:282–9. doi: 10.20892/j.issn.2095-3941.2018.0023
6. Giuliano AE, Ballman KV, McCall L, Beitsch PD, Brennan MB, Kelemen PR, et al. Effect of Axillary Dissection vs No Axillary Dissection on 10-Year Overall Survival Among Women With Invasive Breast Cancer and Sentinel Node Metastasis: The ACOSOG Z0011 (Alliance) Randomized Clinical Trial. *JAMA* (2017) 318:918–26. doi: 10.1001/jama.2017.11470
7. Perhavec A, Perme MP, Hocevar M, Besić N, Zgajnar J. Ljubljana Nomograms for Predicting the Likelihood of Non-Sentinel Lymph Node Metastases in Breast Cancer Patients With a Positive Sentinel Lymph Node. *Breast Cancer Res Treat* (2010) 119:357–66. doi: 10.1007/s10549-009-0561-4
8. Van Zee KJ, Manasseh DM, Bevilacqua JL, Boolbol SK, Fey JV, Tan LK, et al. A Nomogram for Predicting the Likelihood of Additional Nodal Metastases in Breast Cancer Patients With a Positive Sentinel Node Biopsy. *Ann Surg Oncol* (2003) 10:1140–51. doi: 10.1245/aso.2003.03.015
9. Degnim AC, Reynolds C, Pantvaidya G, Zakaria S, Hoskin T, Barnes S, et al. Nonsentinel Node Metastasis in Breast Cancer Patients: Assessment of an Existing and a New Predictive Nomogram. *Am J Surg* (2005) 190:543–50. doi: 10.1016/j.amjsurg.2005.06.008
10. Pal A, Provenzano E, Duffy SW, Pinder SE, Purushotham AD. A Model for Predicting Non-Sentinel Lymph Node Metastatic Disease When the Sentinel Lymph Node is Positive. *Br J Surg* (2008) 95:302–9. doi: 10.1002/bjs.5943
11. Kohrt HE, Olshen RA, Bermas HR, Goodson WH, Wood DJ, Henry S. Et Al; New Models and Online Calculator for Predicting Non-Sentinel Lymph Node Status in Sentinel Lymph Node Positive Breast Cancer Patients. *BMC Cancer* (2008) 8:66. doi: 10.1186/1471-2407-8-66
12. Barranger E, Coutant C, Flahault A, Delpech Y, Darai E, Uzan S. An Axilla Scoring System to Predict Non-Sentinel Lymph Node Status in Breast Cancer Patients With Sentinel Lymph Node Involvement. *Breast Cancer Res Treat* (2005) 91:113–9. doi: 10.1007/s10549-004-5781-z
13. Hwang RF, Krishnamurthy S, Hunt KK, Mirza N, Ames FC, Feig B, et al. Clinicopathologic Factors Predicting Involvement of Nonsentinel Axillary Nodes in Women With Breast Cancer. *Ann Surg Oncol* (2003) 10:248–54. doi: 10.1245/aso.2003.05.020
14. Saidi RF, Dudrick PS, Remine SG, Mittal VK. Nonsentinel Lymph Node Status After Positive Sentinel Lymph Node Biopsy in Early Breast Cancer. *Am Surg* (2004) 70:101–5.
15. Byon JH, Park YV, Yoon JH, Moon HJ, Kim EK, Kim MJ, et al. Added Value of MRI for Invasive Breast Cancer Including the Entire Axilla for Evaluation of High-Level or Advanced Axillary Lymph Node Metastasis in the Post-ACOSOG Z0011 Trial Era. *Radiology* (2021) 300:46–54. doi: 10.1148/radiol.2021202683
16. Zhang X, Liu Y, Luo H, Zhang J. PET/CT and MRI for Identifying Axillary Lymph Node Metastases in Breast Cancer Patients: Systematic Review and Meta-Analysis. *J Magn Reson Imaging* (2020) 52:1840–51. doi: 10.1002/jmri.27246

17. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278:563–77. doi: 10.1148/radiol.2015151169
18. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: The Bridge Between Medical Imaging and Personalized Medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141
19. Mao N, Dai Y, Lin F, Ma H, Duan S, Xie H, et al. Radiomics Nomogram of DCE-MRI for the Prediction of Axillary Lymph Node Metastasis in Breast Cancer. *Front Oncol* (2020) 10:541849. doi: 10.3389/fonc.2020.541849
20. Liu M, Mao N, Ma H, Dong J, Zhang K, Che K, et al. Pharmacokinetic Parameters and Radiomics Model Based on Dynamic Contrast Enhanced MRI for the Preoperative Prediction of Sentinel Lymph Node Metastasis in Breast Cancer. *Cancer Imaging* (2020) 20:65. doi: 10.1186/s40644-020-00342-x
21. Dong Y, Feng Q, Yang W, Lu Z, Deng C, Zhang L, et al. Preoperative Prediction of Sentinel Lymph Node Metastasis in Breast Cancer Based on Radiomics of T2-Weighted Fat-Suppression and Diffusion-Weighted MRI. *Eur Radiol* (2018) 28:582–91. doi: 10.1007/s00330-017-5005-7
22. Santucci D, Faiella E, Cordelli E, Sicilia R, de Felice C, Zobel BB, et al. 3t MRI-Radiomic Approach to Predict for Lymph Node Status in Breast Cancer Patients. *Cancers (Basel)* (2021) 13:2228. doi: 10.3390/cancers13092228
23. Zhang X, Yang Z, Cui W, Zheng C, Li H, Shen J, et al. Preoperative Prediction of Axillary Sentinel Lymph Node Burden With Multiparametric MRI-Based Radiomics Nomogram in Early-Stage Breast Cancer. *Eur Radiol* (2021) 31 (8):5924–39. doi: 10.1007/s00330-020-07674-z
24. Varghese P, Abdel-Rahman AT, Akberali S, Mostafa A, Gattuso JM, Carpenter R. Methylene Blue Dye–a Safe and Effective Alternative for Sentinel Lymph Node Localization. *Breast J* (2008) 14:61–7. doi: 10.1111/j.1524-4741
25. Kalli S, Semine A, Cohen S, Naber SP, Makim SS, Bahl M. American Joint Committee on Cancer's Staging System for Breast Cancer, Eighth Edition: What the Radiologist Needs to Know. *Radiographics* (2018) 38:1921–33. doi: 10.1148/rg.2018180056
26. Rautiainen S, Masarwah A, Sudah M, Sutela A, Pelkonen O, Joukainen S, et al. Axillary Lymph Node Biopsy in Newly Diagnosed Invasive Breast Cancer: Comparative Accuracy of Fine-Needle Aspiration Biopsy Versus Core-Needle Biopsy. *Radiology* (2013) 269:54–60. doi: 10.1148/radiol.13122637
27. Mortellaro VE, Marshall J, Singer L, Hochwald SN, Chang M, Copeland EM, et al. Magnetic Resonance Imaging for Axillary Staging in Patients With Breast Cancer. *J Magn Reson Imaging* (2009) 30:309–12. doi: 10.1002/jmri.21802
28. Javid S, Segara D, Lotfi P, Raza S, Golshan M. Can Breast MRI Predict Axillary Lymph Node Metastasis in Women Undergoing Neoadjuvant Chemotherapy. *Ann Surg Oncol* (2010) 17:1841–6. doi: 10.1245/s10434-010-0934-2
29. Tran D, Cooke S, Illingworth PJ, Gardner DK. Deep Learning as a Predictive Tool for Fetal Heart Pregnancy Following Time-Lapse Incubation and Blastocyst Transfer. *Hum Reprod* (2019) 34:1011–8. doi: 10.1093/humrep/dez064
30. Demler OV, Pencina MJ, D'Agostino RBSr. Misuse of DeLong Test to Compare AUCs for Nested Models. *Stat Med* (2012) 31:2577–87. doi: 10.1002/sim.5328
31. Wu S, Zheng J, Li Y, Yu H, Shi S, Xie W, et al. A Radiomics Nomogram for the Preoperative Prediction of Lymph Node Metastasis in Bladder Cancer. *Clin Cancer Res* (2017) 23:6904–11. doi: 10.1158/1078-0432.CCR-17-1510
32. Kootstra JJ, Hoekstra-Weebers JE, Rietman JS, de Vries J, Baas PC, Geertzen JH, et al. A Longitudinal Comparison of Arm Morbidity in Stage I-II Breast Cancer Patients Treated With Sentinel Lymph Node Biopsy, Sentinel Lymph Node Biopsy Followed by Completion Lymph Node Dissection, or Axillary Lymph Node Dissection. *Ann Surg Oncol* (2010) 17:2384–94. doi: 10.1245/s10434-010-0981-8
33. Caudle AS, Cupp JA, Kuerer HM. Management of Axillary Disease. *Surg Oncol Clin N Am* (2014) 23:473–86. doi: 10.1016/j.soc.2014.03.007
34. Yu Y, Tan Y, Xie C, Hu Q, Ouyang J, Chen Y, et al. Development and Validation of a Preoperative Magnetic Resonance Imaging Radiomics-Based Signature to Predict Axillary Lymph Node Metastasis and Disease-Free Survival in Patients With Early-Stage Breast Cancer. *JAMA Netw Open* (2020) 3:e2028086. doi: 10.1001/jamanetworkopen.2020.28086
35. Chen S, Shu Z, Li Y, Chen B, Tang L, Mo W, et al. Machine Learning-Based Radiomics Nomogram Using Magnetic Resonance Images for Prediction of Neoadjuvant Chemotherapy Efficacy in Breast Cancer Patients. *Front Oncol* (2020) 10:1410. doi: 10.3389/fonc.2020.01410
36. Yu Y, He Z, Ouyang J, Tan Y, Chen Y, Gu Y, et al. Magnetic Resonance Imaging Radiomics Predicts Preoperative Axillary Lymph Node Metastasis to Support Surgical Decisions and is Associated With Tumor Microenvironment in Invasive Breast Cancer: A Machine Learning, Multicenter Study. *EBioMedicine* (2021) 69:103460. doi: 10.1016/j.ebiom.2021.103460
37. Sosna J. Fewer Reproducible Radiomic Features Mean Better Reproducibility Within the Same Patient. *Radiology* (2019) 293:592–3. doi: 10.1148/radiol.2019191958

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

CLINICAL TRIAL
published: 07 March 2022
doi: 10.3389/fonc.2022.775081

Dalpiciclib Combined With Pyrotinib and Letrozole in Women With HER2-Positive, Hormone Receptor-Positive Metastatic Breast Cancer (LORDSHIPS): A Phase Ib Study

Jian Zhang[1,2†], Yanchun Meng[1,2†], Biyun Wang[1,2], Leiping Wang[1,2], Jun Cao[1,2], Zhonghua Tao[1,2], Ting Li[1,2], Wenqing Yao[3] and Xichun Hu[1,2]*

[1] Department of Medical Oncology, Fudan University Shanghai Cancer Center, Shanghai, China, [2] Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China, [3] Department of Clinical Research & Development, Jiangsu Hengrui Pharmaceuticals Co., Ltd., Shanghai, China

OPEN ACCESS

Edited by:
Sercan Aksoy,
Hacettepe University, Turkey

Reviewed by:
Jürgen Geisler,
University of Oslo, Norway
Adam Brufsky,
University of Pittsburgh Medical Center, United States

***Correspondence:**
Xichun Hu
huxichun2017@163.com

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: *13 September 2021*
Accepted: *10 February 2022*
Published: *07 March 2022*

Citation:
Zhang J, Meng Y, Wang B, Wang L, Cao J, Tao Z, Li T, Yao W and Hu X (2022) Dalpiciclib Combined With Pyrotinib and Letrozole in Women With HER2-Positive, Hormone Receptor-Positive Metastatic Breast Cancer (LORDSHIPS): A Phase Ib Study. Front. Oncol. 12:775081. doi: 10.3389/fonc.2022.775081

Purpose: The LORDSHIPS study aimed to explore the safety and efficacy of a novel fully oral triplet combination of dalpiciclib (a potent cyclin-dependent kinase 4/6 inhibitor), pyrotinib (a HER2 tyrosine kinase inhibitor) and endocrine therapy letrozole in patients with HER2-positive, hormone receptor (HR)-positive metastatic breast cancer (MBC) in the front-line setting.

Patients and Methods: Postmenopausal women with HER2-positive, HR-positive MBC were recruited in the dose-finding phase Ib trial. A standard 3 + 3 design was used to determine safety, tolerability, and recommended phase II dose (RP2D) for the combination.

Results: A total of 15 patients were enrolled to three dose combination cohorts (letrozole/pyrotinib/dalpiciclib, level/I: 2.5/400/125 mg, n=5; level/L1: 2.5/400/100 mg, n=6; level/L2: 2.5/320/125 mg, n=4). Three patients experienced dose-limiting toxicities (level/I, n=2; level/L1, n=1) and level/L2 was identified as RP2D. The most frequent grade 3-4 adverse events were neutropenia (46.7%), leukopenia (40.0%), oral mucositis (26.7%) and diarrhea (20.0%). The confirmed objective response rate (ORR) was 66.7% (95% CI: 38.4% to 88.2%). The confirmed ORR of study treatment as first line (1L) and second line (2L) HER2-targeted therapy was 85.7% (6/7) and 50.0% (4/8), respectively. Median progression-free survival (PFS) was 11.3 months (95% CI: 5.3 months to not reached). PFS in 1L setting was not reached yet, while PFS in 2L setting was 10.9 months (95% CI: 1.8 to 13.7 months).

Conclusions: The fully oral combination of dalpiciclib, pyrotinib and letrozole is a promising chemotherapy-sparing treatment option for HER2-positive, HR-positive MBC patients. The planned dose-expansion phase II study is ongoing.

Clinical Trial Registration: ClinicalTrials.gov, identifier NCT03772353.

Keywords: metastatic breast cancer, HER2-positive, hormone receptor-positive, pyrotinib, CDK4/6 inhibitor, endocrine therapy

INTRODUCTION

Breast cancer (BC) is the most common cancer globally (1), with 15%-20% of BCs classified as human epidermal growth factor receptor (HER2)-positive (2). Despite successful HER2 targeted therapies, a substantial proportion of patients with HER2-positive advanced breast cancer will eventually acquire treatment resistance and succumb to their disease. The co-expression of hormone receptors (HR) is an important resistance mechanism, affecting around 50% of HER2-positive BC (2–4). Given that patients with HER2-positive, HR-positive breast cancer are less likely to respond to standard combination of anti-HER2 and chemotherapy (5–8), several studies have valuated the possibility of combined treatment with anti-HER2 and endocrine therapy. However, such regiments merely led to a modest improvement in progression-free survival (PFS) (9–11). Therefore, alternative strategies are much-needed to overcome the treatment resistance in patients with HER2-positive, HR-positive breast cancer.

Cyclin-dependent kinase 4/6 (CDK4/6) has now become a promising strategy for HER2-positive, HR-positive breast cancer treatment as it is the downstream of the estrogen receptor (ER) and HER2 pathways, as well as many other cellular pathways inducing resistance to HER2-targeted therapies (12). Preclinical studies have reported that increased levels of cyclin D1 and CDK4 confer resistance to HER2-inhibitors in tumor cells, and CDK4/6 inhibitor can regain the sensitivity to HER2-directed agents (13). Results from the MonarcHER study demonstrated that the combination of CDK4/6 inhibitor abemaciclib plus trastuzumab and fulvestrant were effective and tolerable in heavily pretreated HER2-positive, HR-positive metastatic breast cancer (MBC) patients (14). Moreover, a similar study of tucatinib, palbociclib and letrozole showed promising activity in patients with two lines of prior therapy for HER2-positive, HR-positive MBC, even in brain metastases (15, 16). Previous findings bring a glimmer of light to prevent or conquer either endocrine or anti-HER2 therapy resistance in HER2-positive, HR-positive MBC patients. However, the efficacy of the addition of CDK 4/6 inhibitors to hormonal and anti-HER2 therapies in the front-line setting remains unknown.

Dalpiciclib (SHR6390) is an oral, novel, efficient, and highly selective small-molecule CDK4/6 inhibitor (17). Phase III trial (DAWNA-1) has demonstrated improved PFS with dalpiciclib plus fulvestrant versus placebo plus fulvestrant (15.7 vs 7.2 months; hazard ratio, 0.42; p<0.0001] in pretreated HR-positive, HER2-negative advanced breast cancer (18) Pyrotinib, an irreversible pan-HER receptor tyrosine kinase inhibitor (TKI) targeting epidermal growth factor receptor/HER1, HER2, and HER4 (19), is approved for the treatment of HER2-positive metastatic breast cancer in China. In a randomized, controlled, phase III trial (PHOEBE), pyrotinib plus capecitabine yielded significantly improved PFS compared with lapatinib plus capecitabine (12.5 vs 6.8 months; hazard ratio, 0.39; one-sided p<0.0001) in HER2-positive metastatic breast cancer patients who previously received trastuzumab and taxanes (20). Notably, preclinical studies demonstrated that dalpiciclib can overcome resistance to endocrine therapy and HER2-targeted antibody in ER-positive, HER2-positive breast cancer cells (21). Additionally, dalpiciclib sensitizes pyrotinib in pyrotinib-refractory HER2-positive gastric cancer models, which has been preliminary validated in five HER2-positive gastric cancer patients (22).

Previous findings suggested that the combination of anti-HER2 agent, CDK4/6 inhibitor and endocrine therapy could be synergistic in HER2-positive, HR-positive breast cancer. To test the hypothesis, we conducted the LORDSHIPS study to investigate the safety and efficacy of a fully oral therapy that adding CDK 4/6 inhibitor dalpiciclib to the combination of pyrotinib and letrozole as front-line treatment in patients with HER2-positive, HR-positive relapsed or metastatic breast cancer.

PATIENTS AND METHODS

Study Design and Treatments

The LORDSHIPS study (ClinicalTrial.gov Identifier: NCT03772353) was a single-center, open-label, dose-finding phase Ib study. In this trial, a traditional 3 + 3 design was implemented for dose finding (**Figure 1**). The treatment consisted of letrozole (fixed dose at 2.5 mg) and pyrotinib (initial dose at 400 mg) orally once daily in 28-day cycles

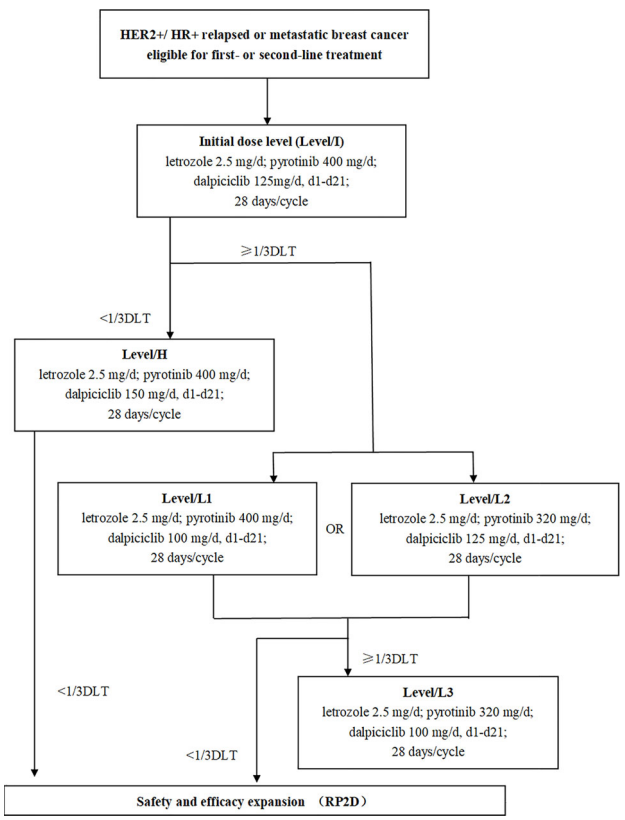


FIGURE 1 | Study design. HR, hormone receptor; HER2, human epidermal growth factor receptor 2; DLT, dose-limiting toxicity; RP2D, recommended phase II dose.

combined with dalpiciclib (initial dose at 125 mg) orally once daily for 21 days followed by 7 days off. The combination dose finding of dalpiciclib and pyrotinib followed the "3+3" principle, with a subsequent dose escalation or de-escalation based on the incidence of specified dose-limiting toxicities (DLTs) in the initial dose group. If the initial dose level (Level/I) with pyrotinib 400 mg/d and dalpiciclib 125 mg/d could be tolerated with zero out of three patients or one out of six patients experienced a DLT, subsequent participants were assigned to the higher level (Level/H) with pyrotinib 400 mg/d and dalpiciclib 150 mg/d; otherwise, to de-escalation of dalpiciclib that Level/L1 with pyrotinib 400 mg/d, and dalpiciclib 100 mg/d, or de-escalation of pyrotinib that Level/L2 with pyrotinib 320 mg/d, and dalpiciclib 125 mg/d. If two or more patients experienced a DLT in a cohort of three or six patients at the dose of Level/L1 and Level/L2, patients would be assigned to the next dose de-escalation group in Level/L3 with pyrotinib 320 mg/d and dalpiciclib 100 mg/d according to the dose adjustment principle (**Figure 1**).

Patients

Postmenopausal female patients aged 18-75 years, diagnosed with HER2-positive, HR-positive unresectable, relapsed or metastatic breast cancer confirmed by histopathology (local laboratory assessment) were recruited. Patients must have received ≤1 line of systemic chemotherapy for metastatic stage, ≤1 line of HER2 targeted therapy and ≤1 line of endocrine therapy. Other key inclusion criteria included at least one extracranial measurable lesion per Response Evaluation Criteria in Solid Tumors (RECIST) criteria version 1.1, an Eastern Cooperative Oncology Group (ECOG) performance status of 0-1, and adequate bone marrow and organ function. Key exclusion criteria were untreated central nervous system metastases, any prior treatment with CDK4/6 inhibitor, or proven primary resistance to letrozole or anastrozole. Primary resistance was defined as relapse during the first 2 years of adjuvant endocrine therapy or progression of disease within the first 6 months of first-line endocrine therapy for metastatic breast cancer.

The study was approved by the institutional ethics committee of the Fudan University Shanghai Cancer Center, following the principles of Declaration of Helsinki and Good Clinical Practice guidelines of the National Medical Products Administration of China. Informed consent was obtained from each participant. The potential risks and benefits of the protocol had been explained by the investigators before any study procedures were initiated. The protocol was designed and conducted in accordance with all applicable regulations, guidance, and local policies.

Study Objectives and Assessments

The aim of the study was to determine the safety and tolerability of dalpiciclib in combination with letrozole and pyrotinib, and the recommended dose to be used in the phase II extension study.

The primary endpoints were DLTs, maximum tolerated dose (MTD), RP2D and safety of dalpiciclib in combination with letrozole and pyrotinib. DLTs were defined as the following adverse events (AEs) definitely or possibly related to study drugs in the first cycle: grade 4 neutropenia ≥5 days; grade 4 thrombocytopenia or grade 3 thrombocytopenia with significant clinical bleeding; grade ≥3 neutropenia with fever (≥38.0 degrees Celsius for 1 hour or >38.3 degrees Celsius on single oral measurement); grade ≥4 anemia; and any grade ≥3 non-hematological toxicity (excluding grade 3-4 nausea/vomiting/diarrhea/electrolyte disturbance in patients who recovered to ≤ grade 2 within 72 hours with best supportive care, and grade 3-4 increased alkaline phosphatase or glutamyl transpeptidase related to cancer instead of drugs). MTD was defined as the dose below which ≥1 of 3 or ≥2 of 6 patients experienced DLTs in the first cycle. AE severity was classified according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) (version 4.0.3).

The secondary endpoints included investigator-assessed PFS, objective response rate (ORR), disease control rate (DCR), clinical benefit rate (CBR, the proportion of subjects with complete response (CR), partial response (PR) or stable disease (SD) ≥ 24 weeks during the study), and duration of response (DOR). CR and PR must be confirmed within 4-6 weeks after the criteria for response were first met. Enrolled patients underwent imaging evaluations at baseline and at the end of every 2 cycles (every 8 weeks ±7 days) until disease progression or the initiation of new anticancer therapy. The tumor response was evaluated according to RECIST 1.1 criteria. Following disease progression or initiation of new anticancer therapy, survival was followed up every 12 weeks until death. This study also collected samples for the analysis of the pharmacokinetic (PK) profile. Blood samples for PK analyses of dalpiciclib in combination with letrozole and pyrotinib were collected on day 21 of the first cycle at 1 hour, 3 hours, and 24 hours after administration.

Statistical Analyses

All statistical analyses, except pharmacokinetic analysis (Phoenix WinNonlin 8.1), were conducted using SAS 9.2 or above (North Carolina, USA). Continuous data were presented as mean and standard deviation, or median with maximum and minimum value. Categorical data were listed as the frequency and percentage. The adverse events and serious adverse events were assessed as the indicators of safety in each dose group. Point estimates of efficacy endpoints such as ORR, DCR, and CBR were provided with 95% confidence interval (CI) calculated by Clopper-Pearson method. The Kaplan-Meier method was used to evaluate median PFS and Brookmeyer-Crowley method was used to construct 95% CI.

RESULTS

Patient Characteristics

Between February 2019 and June 2020, a total of 15 eligible MBC patients from Fudan University Shanghai Cancer Center were enrolled in the phase Ib study. As of the January 1, 2022 data cutoff, the median follow-up was 11.4 months (range, 1.8-24.3 months). Four patients (26.7%) remained on study treatment,

whereas 11 patients (73.3%) discontinued treatment because of disease progression (9 [60.0%]) or AEs (2 [13.3%]).

Baseline patient demographics, disease characteristics, and previous systemic therapies for breast cancer are summarized in **Table 1**. The median age was 53 years old (range, 38 to 72 years old). 14 patients (93.3%) had visceral metastases with six patients (40.0%) had more than three metastatic lesions. 10 patients (66.7%) had been previously treated with trastuzumab and 11 patients (73.3%) had received prior hormonal therapy. Seven patients (46.7%) and eight patients (53.3%) received the study treatment as first-line (1L) and second-line (2L) HER2-targeted treatment, respectively.

DLTs and RP2D

Five patients were enrolled in Level/I with pyrotinib 400 mg/d, dalpiciclib 125 mg/d, and letrozole 2.5 mg/d, and two patients experienced a DLT with grade 3 oral mucositis. Subsequent participants were assigned to Level/L1 or Level/L2 with de-escalation of dalpiciclib or pyrotinib followed a 3 + 3 design. Six patients were enrolled in Level/L1 and one patient experienced a DLT with grade 3 oral mucositis, while four patients were enrolled in Level/L2 and no DLT occurred. Two different MTDs were determined as Level/L1 and Level/L2. Accordingly, Level/L2 with pyrotinib 320 mg/d, dalpiciclib 125 mg/d, and letrozole 2.5 mg/d was declared as RP2D as no DLT occurred in this cohort.

Safety

Patients who received at least one dose of protocol therapy were evaluable for safety. All patients experienced treatment-related adverse events (TRAEs) with grade 3-4 TRAEs being reported in 80.0% of patients (**Table 2**). The most common TRAEs were neutropenia (100.0%), leukopenia (100.0%), anemia (100.0%), oral mucositis (93.3%) and diarrhea (86.7%). Other common TRAEs (≥50% of patients) included increased creatinine (73.3%), ECG T wave abnormal (60.0%) and hypertriglyceridemia (53.3%). The most frequent grade 3-4 TRAEs included neutropenia (46.7%), leukopenia (40.0%), oral mucositis (26.7%) and diarrhea (20.0%). Serious adverse events (SAE) occurred in only one patient with intracranial hemorrhage, which was attributed to cerebral arteriovenous fistula instead of study drugs. TRAEs led to dose reduction in five patients (33.3%) and treatment discontinuation in 2 patients (13.3%), respectively.

TABLE 1 | Patient characteristics.

Characteristics	Dose Cohorts			
	Level/I (n=5)	Level/L1 (n=6)	Level/L2 (n=4)	Total (N=15)
Age, median (range), years	59 (38-65)	56 (42-72)	50 (44-55)	53 (38-72)
<65 years	4 (80.0)	5 (83.3)	4 (100.0)	13 (86.7)
≥65 years	1 (20.0)	1 (16.7)	0 (0)	2 (13.3)
ECOG performance status, n (%)				
0	0 (0)	0 (0)	0 (0)	0 (0)
1	5 (100.0)	6 (100.0)	4 (100.0)	15 (100.0)
ER status, n (%)				
ER <50%	1 (20.0)	2 (33.3)	0 (0)	3(20.0)
ER ≥50%	4 (80.0)	4 (66.7)	4 (100.0)	12(80.0)
No. of metastatic sites, n (%)				
<3	3 (60.0)	4 (66.7)	2 (50.0)	9 (60.0)
≥3	2 (40.0)	2 (33.3)	2 (50.0)	6 (40.0)
Metastatic sites, n (%)				
Visceral	4 (80.0)	6 (100.0)	4 (100.0)	14 (93.3)
Non-visceral	1 (20.0)	0 (0)	0 (0)	1 (6.7)
Previous lines of HER2-targeted treatment [a], n (%)				
0	1 (20.0)	4 (66.7)	2 (50.0)	7 (46.7)
1	4 (80.0)	2 (33.3)	2 (50.0)	8 (53.3)
Previous trastuzumab therapy, n (%)				
Neoadjuvant/Adjuvant only	2 (40.0)	1 (16.7)	1 (25.0)	4 (26.7)
Advanced setting	3 (60.0)	1 (16.7)	2 (50.0)	6 (40.0)
Overall	5 (100.0)	2 (33.3)	3 (75.0)	10 (66.7)
Previous endocrine therapy, n (%)				
Neoadjuvant/Adjuvant setting only	3 (60.0)	2 (33.3)	1 (25.0)	6 (40.0)
Advanced setting	2 (40.0)	1 (16.7)	2 (50.0)	5 (33.3)
Tamoxifen	2 (40.0)	3 (50.0)	3 (75.0)	8 (53.3)
Aromatase inhibitors	3 (60.0)	2 (33.3)	2 (50.0)	7 (46.7)
Overall	5 (100.0)	3 (50.0)	3 (75.0)	11(73.3)
Previous lines of chemotherapy for advanced setting, n (%)				
0	2 (40.0)	5 (83.3)	3 (75.0)	10 (66.7)
1	3 (60.0)	1 (16.7)	1 (25.0)	5 (33.3)

[a]0 line anti-HER2 treatment was defined as with no history of trastuzumab treatment or relapse more than 1 year after the end of trastuzumab-based adjuvant therapy. 1 line anti-HER2 treatment was defined as relapse during or within 1 year after the end of the adjuvant trastuzumab treatment, or progression on first line trastuzumab treatment for advanced disease.

TABLE 2 | All grade AEs related to treatment with at least two patients.

TRAEs, n (%)	All grades	Grade 3–4
Total patients with any AE	15 (100.0)	12 (80.0)
Hematologic		
Neutropenia	15 (100.0)	7 (46.7)
Leukopenia	15 (100.0)	6 (40.0)
Anemia	15 (100.0)	1 (6.7)
Thrombocytopenia	6 (40.0)	0 (0)
Gastrointestinal		
Oral mucositis	14 (93.3)	4 (26.7)
Diarrhea	13 (86.7)	3 (20.0)
Anorexia	4 (26.7)	0 (0)
Nausea	2 (13.3)	0 (0)
Laboratory		
Increased creatinine	11 (73.3)	0 (0)
Hypertriglyceridemia	8 (53.3)	0 (0)
Hyperglycemia	7 (46.7)	0 (0)
Hypophosphatemia	7 (46.7)	1 (6.7)
Hyperuricemia	6 (40.0)	0 (0)
Increased ALT	6 (40.0)	0 (0)
Haematuria	6 (40.0)	0 (0)
Hypokalemia	5 (33.3)	1 (6.7)
Increased AST	4 (26.7)	0 (0)
Hypoproteinemia	4 (26.7)	0 (0)
Hypomagnesemia	4 (26.7)	0 (0)
Hypocalcemia	4 (26.7)	0 (0)
Positive urine leukocyte	4 (26.7)	0 (0)
Increased ALP	3 (20.0)	0 (0)
Increased GGT	2 (13.3)	0 (0)
Hypercholesterolemia	2 (13.3)	0 (0)
Hyponatremia	2 (13.3)	0 (0)
Constitutional		
ECG T wave abnormal	9 (60.0)	0 (0)
Weight loss	7 (46.7)	0 (0)
Rash	3 (20.0)	0 (0)
Fatigue	3 (20.0)	0 (0)
Dermatitis acneiform	2 (13.3)	0 (0)
Palmar-plantar erythrodysaesthesia syndrome	2 (13.3)	0 (0)
Periodontal disease	2 (13.3)	0 (0)

ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, γ-Glutamyl transpeptidase; ALP, alkaline phosphatase; ECG, electrocardiogram.
Note: no patients died from treatment-emergent adverse events.

Efficacy

As of 1 January 2022, 15 patients were considered evaluable for efficacy. Majority of patients (93.3%, 14/15) showed tumor shrinkage (**Figure 2A**). 10 of 15 (66.7%; 95% CI: 38.4% to 88.2%) patients achieved confirmed partial response (PR) as assessed by the investigator (n=3 [60.0%], Level/I; n=3 [50.0%], Level/L1; n=4 [100.0%], Level/L2) (**Figures 2A, B; Table 3**). Responses were still ongoing in 4 of the 10 responders, and the median DOR was 15.6 months (95% CI: 3.7 months to not reached). The DCR was 93.3% (95% CI: 68.1% to 99.8%) and the CBR was 80.0% (95% CI: 51.9% to 95.7%) for all 15 patients (**Table 3**). With 9 (60.0%) disease progression events, the median PFS was 11.3 months (95% CI: 5.3 months to not reached) (**Figure 3A**).

Preliminary subgroup analysis by the number of HER2-targeted treatment lines for advanced breast cancer showed that ORR of study treatment as 1L and 2L HER2-targeted therapy was 85.7% (6/7) and 50.0% (4/8), respectively. PFS in 2L setting was 10.9 months (95% CI: 1.8 to 13.7 months), while PFS in 1L setting was not reached yet. (**Figure 3B**). In addition, patients with ER≥50% had better ORR (83.3%, 10/12) compared with those with ER<50% (0/3).

PK Analysis

Plasma samples for PK analysis were available from 8 patients. The PK parameters are summarized in **Supplementary 1**. The means of C_{max} for dalpiciclib (125 mg) were 130.38 ng/mL and 139.20 ng/mL, and AUC_{last} of dalpiciclib (125 mg) were 2.65 μg·h/mL and 2.52 μg·h/mL, with pyrotinib doses of 320 mg and 400 mg, respectively. Based on the preliminary data, pyrotinib did not alter the PK profile of dalpiciclib in each cohort. The exposures of pyrotinib were different when combined with dalpiciclib, indicating more data would be needed to identify the drug-drug interaction between pyrotinib and dalpiciclib in the phase II trial.

DISCUSSION

To the best of our knowledge, this was the first study to establish a fully oral therapy of the novel CDK4/6 inhibitor dalpiciclib combined with HER2-targeted tyrosine kinase inhibitor pyrotinib and aromatase inhibitor letrozole as first- or second-line treatment in patients with HR-positive, HER2-positive relapsed or metastatic breast cancer. This approach showed promising anticancer activities and tolerable toxicities. The TRAEs of the combination of pyrotinib, dalpiciclib, and letrozole observed in this study were as expected for each drug toxicity profile, with mild or moderate neutropenia (100%), leukopenia (100%), anemia (100%), oral mucositis (93.3%) and diarrhea (86.7%) as the most common TRAEs (17, 19, 23). Similar to dalpiciclib combined with fulvestrant in DAWNA-1 study, the incidence of hematological toxicities with dalpiciclib, pyrotinib plus letrozole was high, whereas grade≥3 neutropenia and leukopenia were reported less frequently (grade ≥3 neutropenia: 84.2% vs 46.7%; grade ≥ 3 leukopenia: 62.1% vs 33.3%) (24). Diarrhea occurred in 86.7% of patients, by only 20.0% with grade ≥ 3 diarrhea, which compared favorably with pyrotinib plus capecitabine (all grade: 95%; grade ≥ 3: 31%) (20). Diarrhea was generally reversible with anti-diarrhea treatment, treatment interruption, or dose reduction, and it did not lead to treatment termination. Three cases of oral mucositis were identified as DLTs: two cases in Level/I and one case in Level/L1 cohort, respectively. Despite the low incidence of oral mucositis with single agent [dalpiciclib, all grade: <10% and grade ≥ 3:<3% (18); pyrotinib, all grade: 9.9% and grade ≥ 3:1.4% (25)], the events were considered as possibly related to both dalpiciclib and pyrotinib. In this study, dalpiciclib 125 mg/d, pyrotinib 320 mg/d, and letrozole 2.5 mg/d was defined as the recommended phase II dose.

Regardless of the HR status, patients with HER2 overexpression/amplification should receive a combination of HER2-targeted therapy and chemotherapy as the standard 1L treatment (26). However, data from clinical trials showed that the subgroups of HER2-positive, HR-positive tumor are less likely to respond to standard chemotherapy combined with trastuzumab

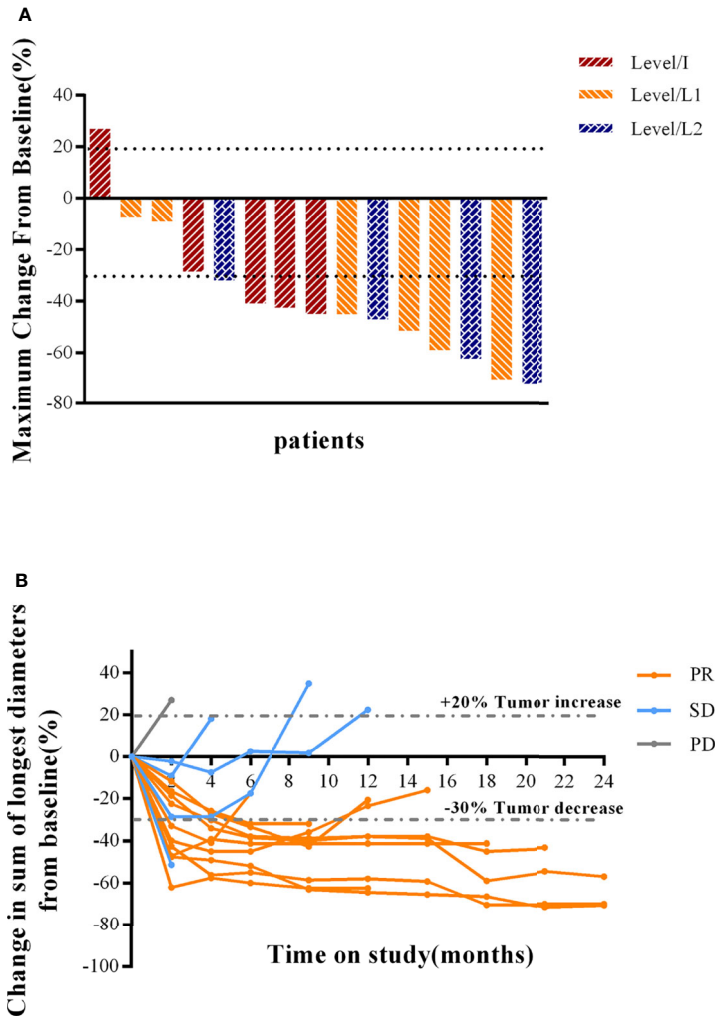


FIGURE 2 | Clinical response to combination therapy in patients. **(A)** Maximum reduction of target lesions from baseline for patients in the Level/I, Level/L1, and Level/L2 dose cohorts. The best response for target lesions per patient was determined on the basis of RECIST 1.1 criteria. **(B)** Change in tumor burden over time, measured as the sum of longest diameters (SLD), in patients with MBC. PR was confirmed by investigator-assessed RECIST 1.1 criteria. PR, partial response; SD, stable disease; PD, progressive disease; HER2, human epidermal growth factor receptor 2.

and pertuzumab, or with T-DM1 (5–8). In CLEOPATRA study, 1L treatment of dual HER2-targeted pertuzumab and trastuzumab plus docetaxel yielded inferior PFS and OS benefits in HR-positive/HER2-positive subsets compared to HR-negative/HER2-positive subsets (hazard ratio for PFS: 0.73 vs 0.64; hazard ratio for OS: 0.71 vs 0.61) (5). In the Chinese bridging study PUFFIN, subgroup analysis suggested that 1L treatment with dual HER2-targeted pertuzumab and trastuzumab plus docetaxel failed to prolong PFS compared to trastuzumab plus docetaxel significantly (14.5 months vs 12.5 months; hazard ratio: 0.80; 95%CI 0.50 to 1.29) in HR-positive, HER2-positive MBC patients (7). The present study showed that the median PFS in the 1L setting was not reached and the ORR was 85.7%, which was equivalent to dual-targeted HER2 agents combined with chemotherapy in CLEOPATRA (ORR in HER2-positive patients: 80.2%) (27) or PUFFIN trial (ORR in HR-positive/HER2-positive patients: 81.7%) (7). T-DM1 is the standard second-line treatment for HER2-positive metastatic breast cancer patients based on the results of EMILIA study, with an ORR of 43.6% and a median PFS of 9.6 months regardless of HR status (6). In China, pyrotinib plus capecitabine has become an alternative 2L treatment option with better PFS and OS compared to lapatinib plus capecitabine (PFS: 12.5 months vs 6.8 months; hazard ratio, 0.39; p<0.0001; ORR 67% vs 52%) (20). In our study, patients in the 2L setting had an ORR of 50.0% and a median PFS of 10.9 months (95% CI: 1.8 to 13.7 months), which was similar to standard treatment of T-DM1 or pyrotinib plus capecitabine. Although cross-trial comparisons should be made with caution, our results indicate a promising treatment option for HR-positive, HER2-positive breast cancer and support further investigations.

TABLE 3 | Response in the evaluable population.

Parameter	Dose Cohorts			
	Level/l (n=5)	Level/L1 (n=6)	Level/L2 (n=4)	Total (N=15)
CR, n (%)	0 (0)	0 (0)	0 (0)	0 (0)
PR, n (%)	3 (60.0)	3 (50.0)	4 (100.0)	10 (66.7)
SD, n (%)	1[a] (20.0)	3[b] (50.0)	0 (0)	4 (26.7)
PD, n (%)	1 (20.0)	0 (0)	0 (0)	1 (6.7)
ORR, n (%)	3 (60.0)	3 (50.0)	4 (100.0)	10 (66.7)
95% CI				38.4- 88.2
DCR, n (%)	4 (80.0)	6 (100.0)	4 (100.0)	14 (93.3)
95% CI				68.1-99.8
CBR, n (%)	4 (80.0)	4 (66.7)	4 (100.0)	12 (80.0)
95% CI				51.9-95.7

[a]*1 patient with SD ≥ 24 weeks.*

[b]*Among 3 patients, 1 patient with SD ≥ 24 weeks.*

CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; ORR, objective response rate (CR + PR); DCR, disease control rate (CR + PR + SD); CBR, clinical benefit rate (CR + PR + SD ≥ 24 weeks); CI, confidence interval.

It's speculated that the inferior response of anti-HER2 and chemotherapy in HR-positive subgroups compared to HR-negative subgroups was in part attributed to the bidirectional crosstalk between the HER2 and ER-α pathways (28, 29). As a result, a growing number of clinical studies have explored the combination of HER2 targeted and endocrine therapy in the subsets of breast cancer patients (9–11, 30). In TAnDEM trial (11), 1L treatment of trastuzumab plus anastrozole achieved PFS benefits compared to anastrozole (4.8 months vs 2.4 months; hazard ratio, 0.63; p=0.0016). In addition, subgroup results of phase III EGF 30008 trial showed that lapatinib plus letrozole achieved ORR 28% and PFS 8.2 months in first line patients (9). These data showed a promising but still modest PFS benefits in HR-positive, HER2-positive patients, indicating that intervention of the crosstalk between HER2 and ER-α might be insufficient and additional treatment are of value to be explored.

Preclinical models showed that CDK4/6 inhibitors could sensitize HER2-targeted therapy and delay tumor recurrence in

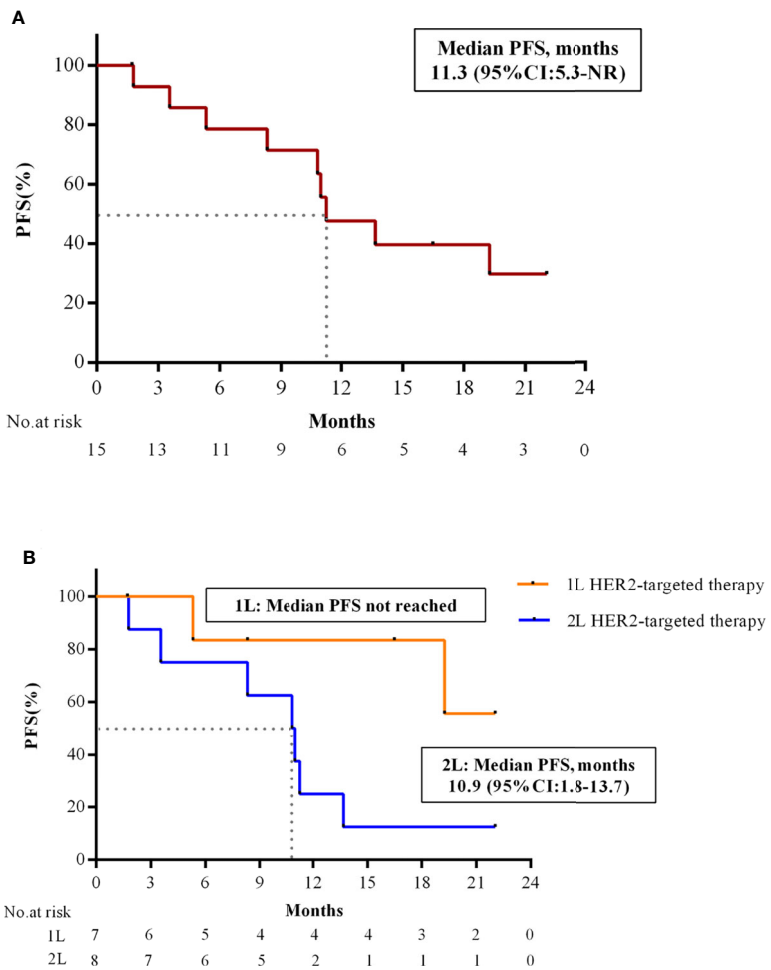


FIGURE 3 | PFS of combination therapy in patients with HER2+/HR+ MBC. **(A)** Kaplan-Meier estimates of PFS in all patients (N = 15). **(B)** Kaplan-Meier estimates of PFS in patients with 1L and 2L HER2-targeted therapy. PFS, progression-free survival; HR, hormone receptor; HER2, human epidermal growth factor receptor 2; 1L, first-line; 2L, second-line; CI, confidence interval; NR, not reached.

HER2+ breast cancer (13). The monarcHER trial built on these preclinical findings and reported that CDK4/6 inhibitor abemaciclib combined with trastuzumab and fulvestrant significantly improved PFS compared to trastuzumab plus standard-of-care chemotherapy (8.3 months vs 5.7 months; hazard ratio, 0.67; p=0.051) as third-line or later treatment in HR-positive/HER2-positive MBC patients (14). Moreover, in heavily pretreated patients with HR-positive/HER2-positive MBC, the combination of tucatinib with letrozole and palbociclib showed a considerable anti-tumor activity with median PFS of 8.7 months (10.1 months for patients without brain metastasis and 6.0 months for those with brain metastasis) (15), and the central nervous system metastases PFS was 8 months in patients with untreated asymptomatic or treated stable patients with brain metastases (16). As 60.5% of HR-positive/HER2-positive MBC patients chose first-line hormonal therapy over chemotherapy in real-world (39.5%) (31), whether patients could obtain benefits from this new kind of chemo-free combination in the 1L or 2L setting would be worthy of investigation. In our study, the combination of pyrotinib, dalpiciclib, and letrozole achieving an ORR of 66.7% (95% CI: 38.4% to 88.2%) with a median PFS of 11.3 months (95% CI: 5.3 months to not reached), shows potential to be a promising chemo-sparing regimen for patients with HR-positive/HER2-positive MBC in the front-line setting. Furthermore, identifying patients who are likely to gain the most benefits from the combination of HER2-targeted and endocrine therapy with CDK4/6 inhibitor is important. Given our results, HR-positive/HER2-positive MBC patients with higher ER expression seemed to be associated with greater benefits from the combination. However, it should be noted that subgroup analysis is inconclusive due to the limited sample size.

Limitations of this early-phase study included its nonrandomized, single-arm design, small sample size and lack of direct comparator with pyrotinib plus letrozole or chemotherapy. In addition, the preliminary pharmacokinetic analyses had not yielded conclusive results because of large variation within individuals and limited blood sampling points. More patients and samples are planned to be included in further study to verify efficacy and pharmacokinetics of the combination. Meanwhile, our study excluded patients with brain metastases in phase Ib. Based on the clinical efficacy of tucatinib with letrozole and palbociclib in heavily treated patients with brain metastases, our study would further investigate the efficacy of the triplet regimen in this population in the front-line setting. Recently, T-DXd was recommended as the new standard second-line therapy by guidelines based on DESTINY-Breast03 trial with highly clinically meaningful and statistically significant improvement in PFS compared with T-DM1 in patients with HER2-positive MBC (PFS HR of 0.28 ($P = 7.8\times10^{-22}$)), similarly in HR-positive subgroup (22.4 months vs 6.9 months; hazard ratio: 0.3191) (32, 33). Since T-DXd has made a breakthrough in HER2-positive breast cancer, anti-HER2 ADC combined with CDK4/6 inhibitor and endocrine therapy may be the future exploring direction of HR-positive/HER2-positive breast cancer.

In conclusion, this is the first study to evaluate a fully oral treatment with the CDK4/6 inhibitor dalpiciclib plus HER2 TKI pyrotinib and letrozole in front-line HR-positive, HER2-positive MBC patients. The triplet combination of dalpiciclib, pyrotinib and letrozole has been proven to be safe and effective, potentially offering a chemotherapy-sparing treatment option for patients with HER2-positive, HR-positive MBC. The dose expansion phase II trial is ongoing to further evaluate its efficacy and safety.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Fudan University Shanghai Cancer Center. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors had full access to all data in the trial and take responsibility for the integrity of the data and the accuracy of the data analysis. JZ, YM, and XH contributed to the conceptualisation and design of the trial. BW, LW, JC, ZT, TL, and WY were responsible for collection and assembly of data. JZ and YM completed the statistical analyses. All authors participated in writing the paper and approved the final version of the paper.

ACKNOWLEDGMENTS

We thank all of patients for their participation in the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.775081/full#supplementary-material

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, 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(3):209–49. doi: 10.3322/caac.21660
2. Wu VS, Kanaya N, Lo C, Mortimer J, Chen S. From Bench to Bedside: What do We Know About Hormone Receptor-Positive and Human Epidermal

Growth Factor Receptor 2-Positive Breast Cancer? *J Steroid Biochem Mol Biol* (2015) 153:45–53. doi: 10.1016/j.jsbmb.2015.05.005
3. Schettini F, Buono G, Cardalesi C, Desideri I, De Placido S, Del Mastro L. Hormone Receptor/Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer: Where We Are Now and Where We Are Going. *Cancer Treat Rev* (2016) 46:20–6. doi: 10.1016/j.ctrv.2016.03.012
4. Zhao S, Liu XY, Jin X, Ma D, Xiao Y, Shao ZM, et al. Molecular Portraits and Trastuzumab Responsiveness of Estrogen Receptor-Positive, Progesterone Receptor-Positive, and HER2-Positive Breast Cancer. *Theranostics* (2019) 9 (17):4935–45. doi: 10.7150/thno.35730
5. Swain SM, Baselga J, Kim SB, Ro J, Semiglazov V, Campone M, et al. Pertuzumab, Trastuzumab, and Docetaxel in HER2-Positive Metastatic Breast Cancer. *N Engl J Med* (2015) 372(8):724–34. doi: 10.1056/NEJMoa1413513
6. Verma S, Miles D, Gianni L, Krop IE, Welslau M, Baselga J, et al. Trastuzumab Emtansine for HER2-Positive Advanced Breast Cancer. *N Engl J Med* (2012) 367(19):1783–91. doi: 10.1056/NEJMoa1209124
7. Xu B, Li W, Zhang Q, Shao Z, Li Q, Wang X, et al. Pertuzumab, Trastuzumab, and Docetaxel for Chinese Patients With Previously Untreated HER2-Positive Locally Recurrent or Metastatic Breast Cancer (PUFFIN): A Phase III, Randomized, Double-Blind, Placebo-Controlled Study. *Breast Cancer Res Treat* (2020) 182(3):689–97. doi: 10.1007/s10549-020-05728-w
8. Saura C, Oliveira M, Feng YH, Dai MS, Chen SW, Hurvitz SA, et al. Neratinib Plus Capecitabine Versus Lapatinib Plus Capecitabine in HER2-Positive Metastatic Breast Cancer Previously Treated With >/= 2 HER2-Directed Regimens: Phase III NALA Trial. *J Clin Oncol* (2020) 38(27):3138–49. doi: 10.1200/JCO.20.00147
9. Johnston S, Pippen JJr., Pivot X, Lichinitser M, Sadeghi S, Dieras V, et al. Lapatinib Combined With Letrozole Versus Letrozole and Placebo as First-Line Therapy for Postmenopausal Hormone Receptor-Positive Metastatic Breast Cancer. *J Clin Oncol* (2009) 27(33):5538–46. doi: 10.1200/JCO.2009.23.3734
10. Johnston SRD, Hegg R, Im SA, Park IH, Burdaeva O, Kurteva G, et al. Phase III, Randomized Study of Dual Human Epidermal Growth Factor Receptor 2 (HER2) Blockade With Lapatinib Plus Trastuzumab in Combination With an Aromatase Inhibitor in Postmenopausal Women With HER2-Positive, Hormone Receptor-Positive Metastatic Breast Cancer: ALTERNATIVE. *J Clin Oncol* (2018) 36(8):741–8. doi: 10.1200/JCO.2017.74.7824
11. Kaufman B, Mackey JR, Clemens MR, Bapsy PP, Vaid A, Wardley A, et al. Trastuzumab Plus Anastrozole Versus Anastrozole Alone for the Treatment of Postmenopausal Women With Human Epidermal Growth Factor Receptor 2-Positive, Hormone Receptor-Positive Metastatic Breast Cancer: Results From the Randomized Phase III TAnDEM Study. *J Clin Oncol* (2009) 27 (33):5529–37. doi: 10.1200/JCO.2008.20.6847
12. O'Sullivan CC, Suman VJ, Goetz MP. The Emerging Role of CDK4/6i in HER2-Positive Breast Cancer. *Ther Adv Med Oncol* (2019) 11:1758835919887665. doi: 10.1177/1758835919887665
13. Goel S, Wang Q, Watt AC, Tolaney SM, Dillon DA, Li W, et al. Overcoming Therapeutic Resistance in HER2-Positive Breast Cancers With CDK4/6 Inhibitors. *Cancer Cell* (2016) 29(3):255–69. doi: 10.1016/j.ccell.2016.02.006
14. Tolaney SM, Wardley AM, Zambelli S, Hilton JF, Troso-Sandoval TA, Ricci F, et al. Abemaciclib Plus Trastuzumab With or Without Fulvestrant Versus Trastuzumab Plus Standard-of-Care Chemotherapy in Women With Hormone Receptor-Positive, HER2-Positive Advanced Breast Cancer (monarcHER): A Randomised, Open-Label, Phase 2 Trial. *Lancet Oncol* (2020) 21(6):763–75. doi: 10.1016/S1470-2045(20)30112-1
15. Shagisultanova E, Gradishar W, Brown-Glaberman U, Chalasani P, Brenner AJ, Stopeck A, et al. Abstract PS10-03: Interim Safety and Efficacy Analysis of Phase IB/II Clinical Trial of Tucatinib, Palbociclib and Letrozole in Patients With Hormone Receptor and HER2-Positive Metastatic Breast Cancer. *Cancer Res* (2021) 81(4 Supplement):PS10–03-PS10-03. doi: 10.1158/1538-7445.sabcs20-ps10-03
16. Shagisultanova E, Gradishar W, Brown-Glaberman U, Chalasani P, Brenner A, Stopeck A, et al. (2021). Abstract P1-18-26: Intracranial Efficacy of Tucatinib, Palbociclib and Letrozole Combination in Patients With HR +/HER2+ Breast Cancer and Brain Metastases, in: *Presented at: 2021 San Antonio Breast Cancer Symposium*, San Antonio, TX, December 7-10, 2021.
17. Zhang P, Xu B, Gui L, Wang W, Xiu M, Zhang X, et al. A Phase 1 Study of Dalpiciclib, A Cyclin-Dependent Kinase 4/6 Inhibitor in Chinese Patients With Advanced Breast Cancer. *Biomark Res* (2021) 9(1):24. doi: 10.1186/s40364-021-00271-2
18. Xu B, Zhang Q, Zhang P, Hu X, Li W, Tong Z, et al. Dalpiciclib or Placebo Plus Fulvestrant in Hormone Receptor-Positive and HER2-Negative Advanced Breast Cancer: A Randomized, Phase 3 Trial. *Nat Med* (2021) 27 (11):1904–9. doi: 10.1038/s41591-021-01562-9
19. Ma F, Li Q, Chen S, Zhu W, Fan Y, Wang J, et al. Phase I Study and Biomarker Analysis of Pyrotinib, a Novel Irreversible Pan-ErbB Receptor Tyrosine Kinase Inhibitor, in Patients With Human Epidermal Growth Factor Receptor 2-Positive Metastatic Breast Cancer. *J Clin Oncol* (2017) 35 (27):3105–12. doi: 10.1200/JCO.2016.69.6179
20. Xu B, Yan M, Ma F, Hu X, Feng J, Ouyang Q, et al. Pyrotinib Plus Capecitabine Versus Lapatinib Plus Capecitabine for the Treatment of HER2-Positive Metastatic Breast Cancer (PHOEBE): A Multicentre, Open-Label, Randomised, Controlled, Phase 3 Trial. *Lancet Oncol* (2021) 22(3):351–60. doi: 10.1016/S1470-2045(20)30702-6
21. Long F, He Y, Fu H, Li Y, Bao X, Wang Q, et al. Preclinical Characterization of SHR6390, a Novel CDK 4/6 Inhibitor, *In Vitro* and in Human Tumor Xenograft Models. *Cancer Sci* (2019) 110(4):1420–30. doi: 10.1111/cas.13957
22. Chen Z, Xu Y, Gong J, Kou F, Zhang M, Tian T, et al. Pyrotinib Combined With CDK4/6 Inhibitor in HER2-Positive Metastatic Gastric Cancer: A Promising Strategy From AVATAR Mouse to Patients. *Clin Transl Med* (2020) 10(4):e148. doi: 10.1002/ctm2.148
23. Dellapasqua S, Colleoni M. Letrozole. *Expert Opin Drug Metab Toxicol* (2010) 6(2):251–9. doi: 10.1517/17425250903540246
24. Xu B, Zhang Q, Zhang P, Hu X, Li W, Tong Z, et al. Dalpiciclib Versus Placebo Plus Fulvestrant in HR+/HER2- Advanced Breast Cancer That Relapsed or Progressed on Previous Endocrine Therapy (DAWNA-1): A Multicenter, Randomized, Phase 3 Study. *J Clin Oncol* (2021) 39(15_suppl):1002–2. doi: 10.1200/JCO.2021.39.15_suppl.1002
25. Yan M, Bian L, Hu X, Zhang Q, Ouyang Q, Feng J, et al. Pyrotinib Plus Capecitabine for Human Epidermal Factor Receptor 2-Positive Metastatic Breast Cancer After Trastuzumab and Taxanes (PHENIX): A Randomized, Double-Blind, Placebo-Controlled Phase 3 Study. *Trans Breast Cancer Res* (2020) 1. doi: 10.21037/tbcr-20-25
26. Montemurro F, Di Cosimo S, Arpino G. Human Epidermal Growth Factor Receptor 2 (HER2)-Positive and Hormone Receptor-Positive Breast Cancer: New Insights Into Molecular Interactions and Clinical Implications. *Ann Oncol* (2013) 24(11):2715–24. doi: 10.1093/annonc/mdt287
27. Baselga J, Cortes J, Kim SB, Im SA, Hegg R, Im YH, et al. Pertuzumab Plus Trastuzumab Plus Docetaxel for Metastatic Breast Cancer. *N Engl J Med* (2012) 366(2):109–19. doi: 10.1056/NEJMoa1113216
28. Puglisi F, Minisini AM, De Angelis C, Arpino G. Overcoming Treatment Resistance in HER2-Positive Breast Cancer: Potential Strategies. *Drugs* (2012) 72(9):1175–93. doi: 10.2165/11634000-000000000-00000
29. Giuliano M, Hu H, Wang YC, Fu X, Nardone A, Herrera S, et al. Upregulation of ER Signaling as an Adaptive Mechanism of Cell Survival in HER2-Positive Breast Tumors Treated With Anti-HER2 Therapy. *Clin Cancer Res* (2015) 21 (17):3995–4003. doi: 10.1158/1078-0432.CCR-14-2728
30. Thanopoulou E, Khader L, Caira M, Wardley A, Ettl J, Miglietta F, et al. Therapeutic Strategies for the Management of Hormone Receptor-Positive, Human Epidermal Growth Factor Receptor 2-Positive (HR+/HER2+) Breast Cancer: A Review of the Current Literature. *Cancers (Basel)* (2020) 12(11). doi: 10.3390/cancers12113317
31. Statler AB, Hobbs BP, Wei W, Gupta A, Blake CN, Nahleh ZA. Real-World Treatment Patterns and Outcomes in HR+/HER2+ Metastatic Breast Cancer Patients: A National Cancer Database Analysis. *Sci Rep* (2019) 9(1):18126. doi: 10.1038/s41598-019-54402-9
32. Cortés J, Kim S, Chung W, Im S, Park YH, Hegg R, et al. LBA1 - Trastuzumab Deruxtecan (T-DXd) vs Trastuzumab Emtansine (T-DM1) in Patients (Pts) With HER2+ Metastatic Breast Cancer (mBC): Results of the Randomized Phase III DESTINY-Breast03 Study. *Ann Oncol* (2021) 32(suppl_5):S1283–346. doi: 10.1016/annonc/annonc741
33. Gennari A, Andre F, Barrios CH, Cortes J, de Azambuja E, DeMichele A, et al. ESMO Clinical Practice Guideline for the Diagnosis, Staging and Treatment of Patients With Metastatic Breast Cancer. *Ann Oncol* (2021) 32(12):1475–95. doi: 10.1016/j.annonc.2021.09.019

Conflict of Interest: Author WY was employed by Jiangsu Hengrui Pharmaceuticals Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership