

RESEARCH

Open Access



MRI radiomics and biological correlations for predicting axillary lymph node burden in early-stage breast cancer

Minping Hong^{1†}, Sijia Fan^{2,9†}, Zeyan Xu^{6†}, Zhen Fang^{2†}, Keng Ling⁷, Penghao Lai^{2,9}, Chaokang Han^{2,9}, Zhonghua Chen⁴, Jie Hou⁵, Yanting Liang³, Changyu Zhou², Junyan Wang¹, Xiaobo Chen^{3,8*}, Yanqi Huang^{3,8*} and Maosheng Xu^{2,9*} 

Abstract

Background and aims Preoperative prediction of axillary lymph node (ALN) burden in patients with early-stage breast cancer is pivotal for individualised treatment. This study aimed to develop a MRI radiomics model for evaluating the ALN burden in early-stage breast cancer and to provide biological interpretability to predictions by integrating radiogenomic data.

Methods This study retrospectively analyzed 1211 patients with early-stage breast cancer from four centers, supplemented by data from The Cancer Imaging Archive (TCIA) and Duke University (DUKE). MRI radiomic features were extracted from dynamic contrast-enhanced MRI images and an ALN burden-related radscore was constructed by the backpropagation neural network algorithm. Clinical and combined models were developed, integrating ALN-related clinical variables and radscore. The Kaplan–Meier curve and log-rank test were used to assess the prognostic differences between the predicted high- and low-ALN burden groups in both Center I and DUKE cohorts. Gene set enrichment and immune infiltration analyses based on transcriptomic TCIA and TCIA Breast Cancer dataset were used to investigate the biological significance of the ALN-related radscore.

Results The MRI radiomics model demonstrated an area under the curve of 0.781–0.809 in three validation cohorts. The predicted high-risk population demonstrated a poorer prognosis (log-rank $P < .05$ in both cohorts). Radiogenomic analysis revealed migration pathway upregulation and cell differentiation pathway downregulation in the high radscore groups. Immune infiltration analysis confirmed the ability of radiological features to reflect the heterogeneity of the tumor microenvironment.

[†]Minping Hong, Sijia Fan, Zeyan Xu and Zhen Fang are joint first authors.

*Correspondence:

Xiaobo Chen
chenxb@gdph.org.cn
Yanqi Huang
huangyanqi@gdph.org.cn
Maosheng Xu
xums166@zcmu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Conclusions The MRI radiomics model effectively predicted the ALN burden and prognosis of early-stage breast cancer. Moreover, radiogenomic analysis revealed key cellular and immune patterns associated with the radscore.

Keywords Breast cancer, Radiomic, Magnetic resonance imaging, Axillary lymph node, Genomics

Introduction

The status of axillary lymph node (ALN) of breast cancer playing a critical role in determining patient prognosis and guiding treatment strategies [1, 2]. Recent guidelines from the American College of Surgeons Oncology Group now advise against axillary lymph node dissection (ALND) in specific cases. Specifically, patients with early-stage (T1 or T2) breast cancer who have one or two positive sentinel lymph nodes are recommended not to undergo ALND if they are undergoing mastectomy followed by radiotherapy [3–6]. Conversely, a higher ALN burden (≥ 3 metastatic ALNs) often leads to recommendations for more aggressive interventions, such as comprehensive chemotherapy or additional surgical procedures, to effectively address the spread of cancer. [7].

Currently, the principal method for evaluating ALN burden is sentinel lymph node biopsy (SLNB) [8]. However, as an invasive axillary technique, SLNB not only contributes to additional medical expenses and anesthesia duration but also may induce complications, such as lymphedema or upper limb numbness, with rates ranging from 1.8–29.9% [9]. If a reliable method existed to predict ALN burden preoperatively existed, patients with three or more positive lymph nodes could be directly selected for ALND or receive a new adjuvant therapy without SLNB [10].

Currently, MRI and ultrasound examinations are the primary noninvasive imaging modalities used for assessing ALN burden [11, 12]. Their efficacy in this context is constrained by their comparatively high false-negative rates [12]. Radiomics, a novel approach for extracting quantitative features from medical images, can be regarded as a form of digital biopsy, facilitating comprehensive tumor phenotype and spatial heterogeneity characterization across diverse clinical scenarios [13–15]. Several previous radiomics models aimed to predict ALN metastasis, but small sample sizes and insufficient independent validation limited their efficacy [16–19]. Additionally, the exploration of the association between radiomic phenotypes and their biological foundations remains superficial [20]. Elucidating the biological underpinnings of radiomic predictions is crucial for their broader acceptance by clinicians as independent diagnostic, predictive, or prognostic tools [21]. Therefore, further research is crucial to validate radiomics' robustness in predicting ALN burden and investigate its molecular genetic basis, thereby facilitating its clinical application.

This study aims to develop a radiomic model for predicting ALN burden in early-stage breast cancer

preoperatively and elucidate the association of biological basis of MRI radiomic phenotypes with ALN burden through radiomic and genomic data integration.

Materials and methods

Patient recruitment and clinicopathological data collection

The ethics committees of four institutions granted ethical approval for this study. Informed consent was waived because of the retrospective study design. Patients with early-stage invasive breast cancer (T1 or T2, tumor size ≤ 5 cm) were recruited from four institutions. These individuals had undergone breast-enhanced MRI examinations within two weeks before SLNB or ALN surgery. This study enrolled 1211 patients, with Center I enrolling 532 cases (May 2014 to December 2020) for model training, and Centers II, III, and IV enrolling 113, 185, and 381 cases, respectively, for validation over different periods. Furthermore, The Cancer Genome Atlas-Breast Cancer (TCGA-BRCA) and the Cancer Imaging Archive-Breast Cancer (TCIA-BRCA) data, including dynamic contrast-enhanced MRI (DCE-MRI) and RNA data, serve as radiogenomics cohorts ($n=99$). Data from the Duke Breast Cancer DCE-MRI set, also from TCIA, served as a prognostic cohort of 226 patients. Figure 1 provides detailed inclusion and exclusion criteria for the study.

The baseline clinical characteristics and pathological data were obtained from the patient's medical records, including age, menopausal status, clinical T stage, estrogen receptor (ER) status, progesterone receptor (PR) status, human epidermal growth factor receptor-2 (HER-2) status, histological grade, histological type, molecular subtype [22], Ki-67 proliferation index [23], ALN burden, and ALN status. Patients were categorized into specific groups based on the pathological ALN burden. These categories included the low-ALN burden group (< 3 metastatic ALNs) and the high-ALN burden group (≥ 3 metastatic ALNs).

MRI information acquisition

The parameters of breast MRI collection and scanning of the four centers, Duke University (DUKE), and TCIA are detailed in the supplementary materials (Table S1). MRI enhancement pattern of the tumor region was recorded as mass or nonmass patterns. MRI-ALN status is assessed by MRI [24, 25], where the presence of one or more suspicious metastatic lymph nodes was considered positive, and the absence was deemed negative. MRI-ALN burden is categorized as low for 0–2 and high for ≥ 3 suspicious nodes. Additionally, MRI images were

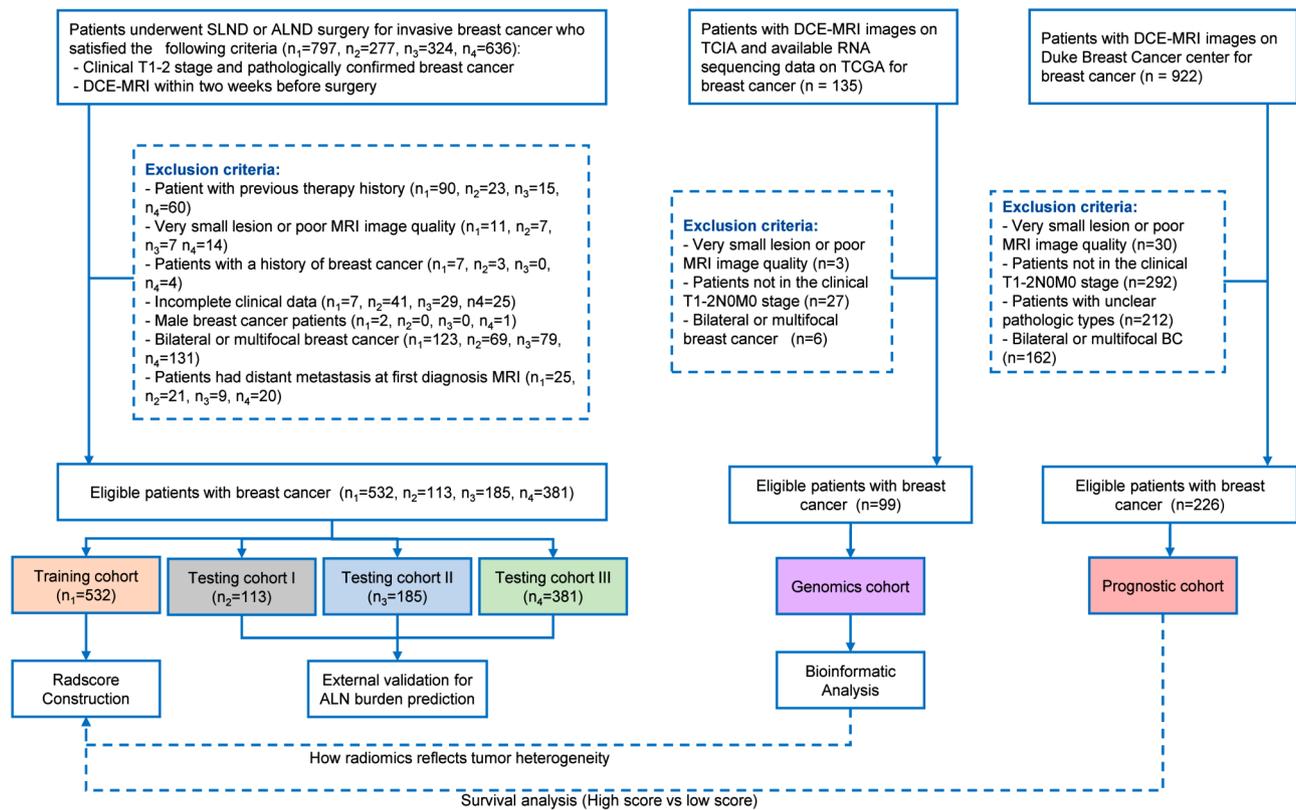


Fig. 1 Flowchart of patient selection and study design. SLNB: sentinel lymph node biopsy; ALND: axillary lymph node dissection; ALN: axillary lymph node burden

independently evaluated by two experienced radiologists, referred to here as Physician A (M.P.H., 8 years of experience) and Physician B (Z.F, 7 years of experience). To ensure consistency in evaluation and to resolve any discrepancies between the two initial assessments, a third, senior radiologist (M.S.X, 31 years of experience) was consulted. This senior radiologist reviewed any conflicting cases and facilitated a consensus decision.

Radiomics analysis

The U2Net network architecture deep learning model was used for automatic volumn of interest segmentation of the peak phase images of the DCE-MRI scan. Two experienced radiologists, M.P.H. and J.Y.W. (22 years of experience) corrected the segmented volumn of interest. The MRI image preprocessing strategy was applied to minimize the effect of device differences between different scanners before feature extraction. The images were first resampled to a standardized voxel spacing of $1 \times 1 \times 1 \text{ mm}^3$ (x, y, z) using a linear interpolation algorithm. The tumor and the surrounding 15-pixel three-dimensional area were segmented, and histogram equalization was conducted based on the field strength intensity. The PyRadiomics package (version 3.0.1) was used to extract radiomic features, such as shape, first-order, texture, wavelet, exponential, and square transform. The same

image segmentation and feature extraction processes were reapplied to 50 cases in the cohort after 2 months to assess the intraclass correlation coefficient (ICC).

Construction of the ALN-burden prediction radiomic model

ComBat Tool and z-scores were first applied to standardize radiomics data. [26, 27] A series of feature selection strategies were used for feature dimensionality reduction. Initially, features were selected for further analysis based on both inter- and intra-observer ICC values > 0.75 . Subsequently, rank-sum tests were performed to determine features with significant differences between low and high-ALN burden groups. The Spearman correlation coefficients (r) were then computed between each pair of features. Feature pairs with $|r| > 0.6$ were selected, following which the feature with the higher mean absolute correlation in each pair was eliminated. The upsampling method was conducted to mitigate model bias favoring the majority class. Elastic-logistic analysis was applied to identify the most predictive features. Finally, a backpropagation neural method was used to develop a radiomic model, and the predicted probability was regarded as the radscore.

Prognostic stratification analysis of ALN status-related Radscore

To evaluate and validate the potential prognostic stratification value of ALN status-related radscore, patients were divided into predicted high- and low-ALN burden groups based on the radscore cutoff value calculated from the minimum log-rank P value method of the training cohort to evaluate and validate the potential prognostic stratification value of ALN status-related radscore. The Kaplan–Meier curve and log-rank test were then performed to assess the prognostic differences between the predicted high- and low-ALN burden groups in both Center I and DUKE cohorts.

External validation of ALN-burden prediction

Cohorts from Centers II–IV were used for the external validation of the radscore to further confirm the generalization performance of the radscore in assessing the ALN burden. The evaluation metrics used included receiver operating characteristic curves and area under the curve (AUC), sensitivity, accuracy, and specificity.

Development and validation of the clinical model

Univariate and multivariate logistic regression analyses were used for selecting the ALN-related clinical variables. The candidate clinical factors included age, menopausal status, histological grade, histological type, MRI ALN status, MRI ALN burden, Ki67, ER, PR, HER-2, molecular subtype, and clinical T stage (Table S2). Subsequently, a combined model integrating all significant clinical predictive factors and radscore was developed, whereas a separate clinical model was established exclusively from relevant clinical factors.

Analysis of potential biological significance underlying the ALN-related Radscore

RNA transcriptomics data was collected from the TCGA-BRCA cohort and matched with the MRI data from TCIA-BRCA to reflect the potential biological significance underlying the ALN-related radscore to improve the biological interpretability of the radiomics model. A total of 99 patients were categorized into the predicted high- and low-ALN burden groups based on the same image preprocessing and radiomic modeling analysis.

For differential expression analysis, we utilized count type gene data, which is ideal for analyzing raw RNA-seq outputs. This analysis was performed on the TCGA-BRCA dataset using the edgeR R package, applying stringent selection criteria of a false discovery rate (FDR) less than 0.01 and an absolute log₂ fold change ($|\log_2\text{FC}|$) greater than 2. The results from this differential expression analysis served as the basis for subsequent Gene Ontology (GO) Biological Process enrichment analysis. This enrichment analysis was conducted using the

clusterProfiler R package, which facilitated the identification of significantly enriched molecular functions and pathways through hierarchical clustering analysis of the differential gene set.

For Gene Set Enrichment Analysis (GSEA) and immune infiltration analysis, we utilized log₂-transformed FPKM type gene data. This transformation standardizes expression levels, enabling more accurate comparisons across samples. The GSEA was employed to determine pathways differentially enriched between patients with varying ALN burdens, thereby providing insights into the biological underpinnings of radiomic features.

Association analysis between radiomics and tumor immune infiltration

Scores for 64 types of immune cells were calculated using the XCell package from fpkm data in RNA-seq, obtained through count normalization, to validate the hypothesis that ALN-related radscore can reflect the heterogeneity of the tumor microenvironment. Rank-sum tests were used to identify differences in the scores of 64 types of immune cells between the high- and low-ALN burden groups. Finally, the Spearman method was used to evaluate the correlation between radiomic features and most immune cell scores.

Statistical analysis

Chi-squared, Fisher's exact, or Mann–Whitney *U* tests were used to assess differences in clinical features. AUC values were used to assess model performance. DeLong test was used to compare differences in AUC values. McNemar's test was used to assess performance differences between the radiomic model and expert reader review. The cutoff value of prediction model and survival analysis was calculated by the maximum Youden index and minimum log-rank P value method of the training cohort, respectively. Kaplan–Meier curve and log-rank test were performed to assess the prognostic differences between the predicted high- and low-ALN burden groups. R (version 4.3.1) and Python (version 3.7.3) were used for all analyses. *P*-values < 0.05 were considered statistically significant.

Result

Characteristics of the study population

Table 1 shows the clinicopathologic factors of 1536 patients with early-stage breast cancer from four Chinese centers, TCIA, and DUKE. The average age of the enrolled patients was 52.0 ± 10.7 years (range: 21–85 years), with 540 (44.6%) and 671 (55.4%) having positive and negative ALN status, respectively. Of the 540 patients with positive ALN status, 215 (39.8%) were identified with a high-ALN burden. Patients from Center

Table 1 Demographic and clinicopathological characteristics

Characteristics	levels	Centre I (n=532)	Centre II (n=113)	Centre III (n=185)	Centre IV (n=381)	TCIA (n=99)	Duke (n=226)
ALN burden	Low	448 (84.2%)	90 (79.6%)	149 (80.5%)	309 (81.1%)	NA	NA
	High	84 (15.8%)	23 (20.4%)	36 (19.5%)	72 (18.9%)	NA	NA
ALN status	Negative	272 (51.1%)	62 (54.9%)	119 (64.3%)	218 (57.2%)	NA	NA
	Positive	260 (48.9%)	51 (45.1%)	66 (35.7%)	163 (42.8%)	NA	NA
Menopausal Status	Postmenopausal	241 (45.3%)	52 (46%)	124 (67%)	297 (78%)	NA	NA
	Premenopausal	291 (54.7%)	61 (54%)	61 (33%)	84 (22%)	NA	NA
Histological grade	Low	34 (6.4%)	9 (7.9%)	14 (7.6%)	45 (11.8%)	NA	NA
	Intermediate	313 (58.8%)	35 (31.0%)	88 (47.6%)	130 (34.1%)	NA	NA
	High	185 (35.8%)	69 (61.1%)	83 (44.9%)	206 (54.1%)	NA	NA
Histological type	Ductal	494 (92.9%)	100 (88.5%)	165 (89.2%)	341 (89.5%)	NA	207 (91.6%)
	Lobular	12 (2.3%)	3 (2.7%)	4 (2.2%)	14 (3.7%)	NA	16 (7.1%)
	Other	26 (4.9%)	10 (8.8%)	16 (8.6%)	26 (6.8%)	NA	3 (1.2%)
Enhanced pattern	Mass	418 (78.6%)	91 (80.5%)	140 (75.7%)	305 (80.1%)	NA	NA
	No-mass	114 (21.4%)	22 (19.5%)	45 (24.3%)	76 (19.9%)	NA	NA
MRI ALN status	Negative	405 (76.1%)	88 (77.9%)	164 (88.6%)	296 (77.7%)	NA	NA
	Positive	127 (23.9%)	25 (22.1%)	21 (11.4%)	85 (22.3%)	NA	NA
MRI ALN burden	Low	505 (94.9%)	104 (92%)	174 (94.1%)	358 (94%)	NA	NA
	High	27 (5.1%)	9 (8%)	11 (5.9%)	23 (6%)	NA	NA
Ki67 (%)	<14	135 (25.4%)	29 (25.7%)	49 (26.5%)	139 (36.5%)	NA	NA
	≥ 14	397 (74.6%)	84 (74.3%)	136 (73.5%)	242 (63.5%)	NA	NA
ER	Negative	85 (16%)	35 (31%)	57 (30.8%)	112 (29.4%)	NA	56 (24.8%)
	Positive	447 (84%)	78 (69%)	128 (69.2%)	269 (70.6%)	NA	170 (75.2%)
PR	Negative	120 (22.6%)	34 (30.1%)	69 (37.3%)	121 (31.8%)	NA	74 (32.7%)
	Positive	412 (77.4%)	79 (69.9%)	116 (62.7%)	260 (68.2%)	NA	152 (67.3%)
HER2	Negative	417 (78.4%)	63 (55.8%)	107 (57.8%)	278 (73%)	NA	190 (84.1%)
	Positive	97 (18.2%)	30 (26.5%)	34 (18.4%)	70 (18.4%)	NA	36 (15.9%)
	Uncertainty	18 (3.4%)	20 (17.7%)	44 (23.8%)	33 (8.7%)	NA	NA
Molecular subtype	Luminal	456 (85.9%)	80 (71.1%)	134 (72.4%)	277 (72.7%)	NA	NA
	HER2-positive	35 (6.6%)	17 (15.1%)	18 (9.7%)	32 (8.4%)	NA	NA
	Triple-negative	40 (7.5%)	11 (10.2%)	18 (9.9%)	53 (13.9%)	NA	NA
Clinical Tumor stage	T1	206 (38.7%)	51 (45.1%)	75 (40.5%)	150 (39.4%)	43 (43.4%)	128 (56.6%)
	T2	326 (61.3%)	62 (54.9%)	110 (59.5%)	231 (60.6%)	56 (55.6%)	98 (43.4%)
Age	Median (IQR)	50.00	51.00	56.00	53.00	53.00	NA
		[44.00, 59.00]	[44.00, 60.00]	[49.00, 68.00]	[46.00, 62.00]	[45.00, 62.00]	

Note ALN, axillary lymph node burden; ER, estrogen receptor; PR, progesterone receptor. TCIA, The Cancer Imaging Archive

I and DUKE underwent follow-up for overall survival (OS), consequently demonstrating a median [interquartile range] OS of 19.8 [9.42–41.2] months and 46.4 [28.8–63.1] months, respectively.

Feature selection and radscore calculation

A total of 944 features were extracted from volumn of interest. Figures S1a–1b show that data distributions from different centers were relatively scattered before using ComBat, but these datasets converged after the center effect was eliminated using ComBat. First, 736 features with ICC>0.75 were selected. Second, 388 features were selected using independent sample *t*-tests or Mann–Whitney *U* tests ($p<.05$). Third, 37 features were selected based on correlation analysis. Fourth, elastic-logistic regression analysis was performed on

the remaining features after upsampling high-burden patients four times (Figures S1c–1d), reducing inter-sequence redundancy, ultimately yielding nine features (Table S3). Finally, a neural network method (BPNN model) was used to develop an ALN burden prediction model using the nine most predictive features, with their predictive probability constituting the radscore.

Performance of the prediction models

Univariate and multivariate logistic analyses identified menopausal status, MRI-ALN status, MRI-ALN burden, and radscore as independent predictive factors for ALN burden (Table 2). A clinical model and a combined model were constructed based on these factors (Table 2 and Table S2).

Table 2 Univariate and multivariate logistic regression analysis to assess the association of clinical characteristics and radscore with ALN burden

Characteristics	OR (univariable)	OR (multivariable)
Clinical T stage		
T1		
T2	3.17 (2.33–4.32, $p < .001$)	NA
Menopausal Status		
Postmenopausal		
Premenopausal	1.60 (1.22–2.10, $p < .001$)	1.68 (1.21–2.34, $p = .002$)
MRI ALN status		
Negative		
Positive	7.09 (5.18–9.69, $p < .001$)	3.14 (2.19–4.49, $p < .001$)
MRI ALN burden		
Low		
High	59.33 (18.69–188.32, $p < .001$)	19.68 (5.99–64.62, $p < .001$)
Radscore	7.47 (5.12–10.88, $p < .001$)	3.90 (2.63–5.78, $p < .001$)

Note OR, Odds Ratio. ALN, axillary lymph node burden

Table 3 Model performance evaluation and comparison

	Model	AUC (95% CI)	Accuracy	Sensitivity	Specificity	P value
Training Cohort (Centre I)	Clinical model	0.771 (0.741–0.802)	0.715	0.595	0.828	0.000
	Rad ANN model	0.856 (0.830–0.880)	0.791	0.964	0.629	0.000
	Combined model	0.899 (0.878–0.920)	0.831	0.929	0.739	-
Validation Cohort I (Centre II)	Clinical model	0.689 (0.566–0.805)	0.752	0.435	0.833	0.005
	Rad ANN model	0.781 (0.669–0.870)	0.681	0.826	0.644	0.173
	Combined model	0.826 (0.732–0.910)	0.717	0.739	0.711	-
Validation Cohort II (Centre III)	Clinical model	0.620 (0.514–0.718)	0.789	0.250	0.919	0.000
	Rad ANN model	0.809 (0.733–0.875)	0.686	0.806	0.658	0.850
	Combined model	0.812 (0.735–0.881)	0.762	0.694	0.779	-
Validation Cohort III (Centre IV)	Clinical model	0.643 (0.572–0.716)	0.756	0.444	0.828	0.000
	Rad ANN model	0.783 (0.722–0.835)	0.706	0.764	0.693	0.112
	Combined model	0.803 (0.748–0.852)	0.738	0.667	0.754	-

Note CI, confidence interval; AUC, Receiver Operating Characteristic curves and Area Under the Curve. BPNN, the back propagation neural network algorithm. The P value was calculated by DeLong test

The BPNN radiomics model (with AUCs of 0.856, 0.781, 0.809, and 0.783 in training and three external validation cohorts) performed comparable to the combined model (AUCs of 0.899, 0.826, 0.812, and 0.803, DeLong's test, $P = .112–0.850$), significantly outperforming the clinical model (AUCs of 0.771, 0.689, 0.620, and 0.643, $P < .01$). Table 3; Fig. 2a and d show the model performance details. The performance of BPNN radiomics model performed better than the MRI-ALN burden in four cohorts (McNemar's test, $p < .001$). Figure 3 presents three typical cases demonstrating the clinical application of the radiomics model. Figure 3 illustrates that patient 1 was pathologically ALN-negative but misclassified as ALN-positive by MRI. In contrast, patient 2 was initially diagnosed as ALN-negative but was later found to have a low-ALN burden pathologically. Similarly, patient 3 was initially deemed ALN-negative but was later revealed to have a high-ALN burden.

Prognostic stratification analysis of ALN status-related radscore

A cutoff value of radscore with 0.542 was calculated, and patients were categorized into predicted high- and low-ALN burden groups. Kaplan–Meier survival curve revealed that the predicted low-ALN burden group had significantly better OS than the predicted high-ALN burden group in the Center I (hazard ratio [HR] = 31.52, $P = .034$) and DUKE cohort (HR = 20.72, $P = .031$) (Fig. 2e and f).

Biological significance of the ALN-related Radscore

This study identified 231 differentially expressed genes (DEGs) between the predicted high- and low-ALN burden groups, comprising 120 upregulated and 111 downregulated DEGs in the predicted high-ALN burden group (Fig. 4a). Hierarchical clustering analysis revealed that the DEGs were mainly distributed in five functional modules, including response chemotaxis to adhesion; ERK1, ERK2, and MAPK cascade; blood activation body

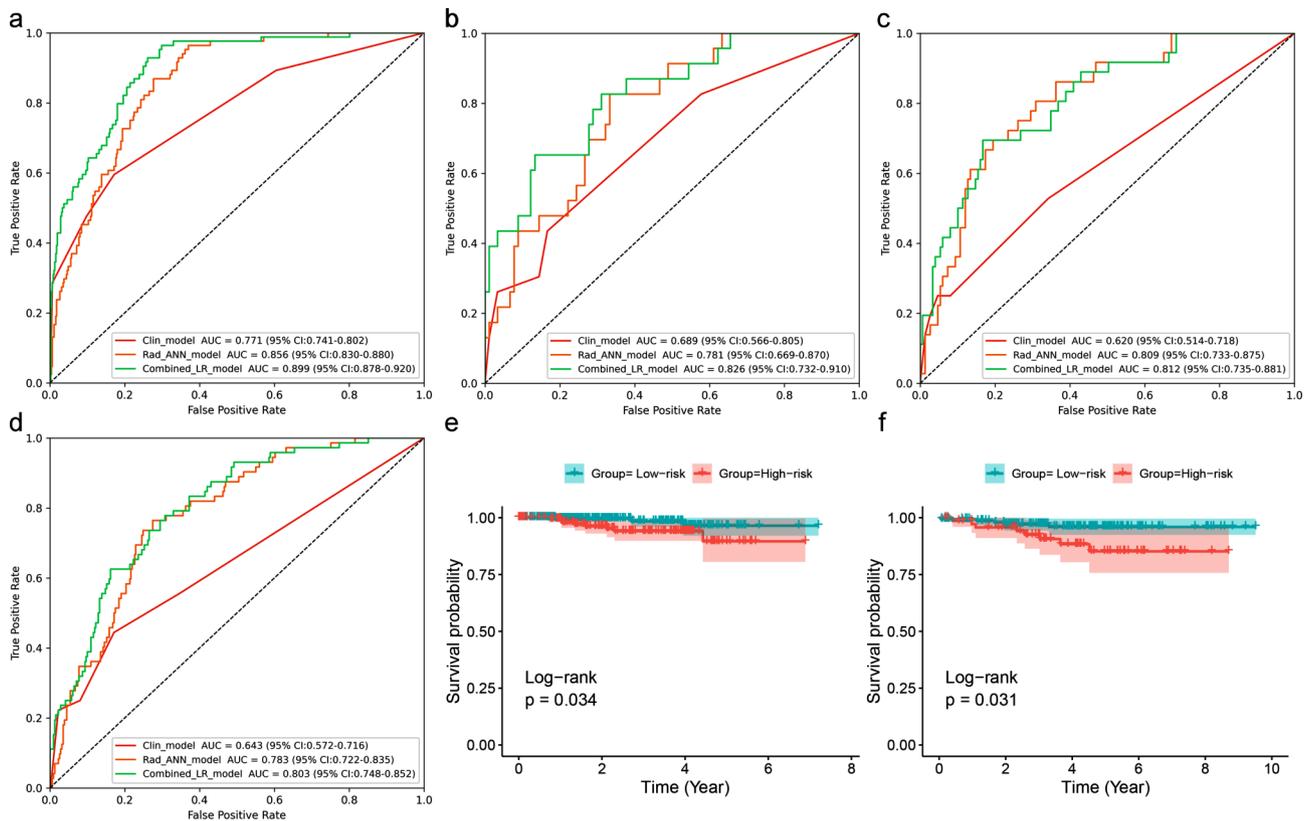


Fig. 2 Model performance analysis. Receiver operating characteristic (ROC) curves for various models are presented for centers I (a), II (b), III (c), and IV (d), along with the AUC and 95% confidence intervals shown in the bottom right. Kaplan–Meier curves illustrate the survival probabilities of patients in the predicted high- and low-ALN burden groups for in Center I (e) and DUKE (f)

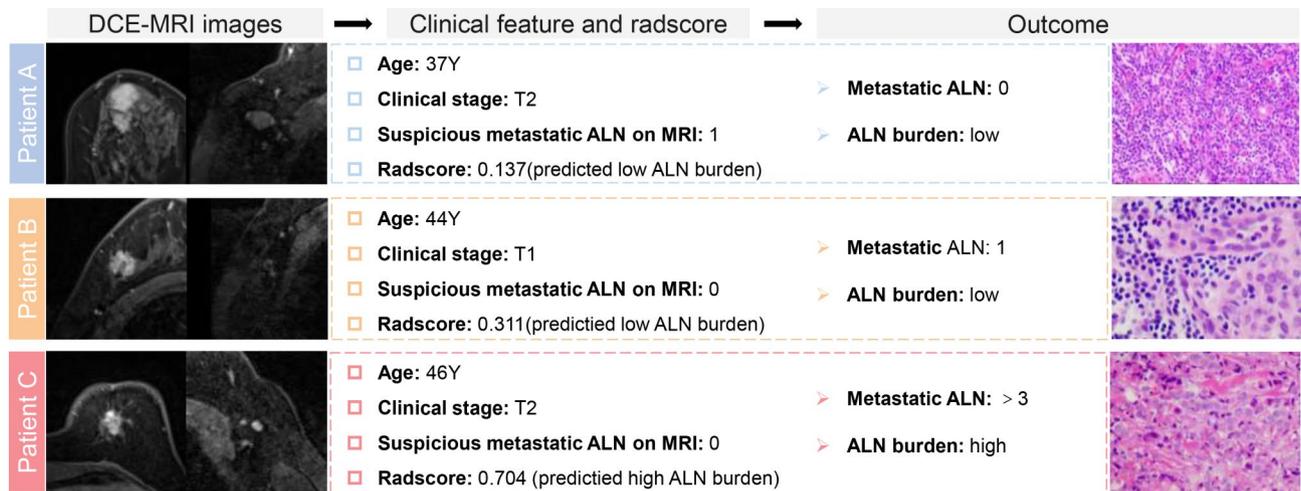


Fig. 3 Three typical cases demonstrated the clinical application of radiomics models. The ALN burden was correctly classified according to the radscore, and MRI image-based decision support was provided. Pathological images are 20×Microscope field

coagulation; receptor surface signaling pathway; and epidermal epidermis development differentiation (Fig. 4b). GO functional analysis revealed that epidermal cell differentiation-related, keratinocyte differentiation, and epidermis pathways were downregulated in the predicted high-ALN burden group (Figures S2a-c). Conversely,

migration/invasion pathways, such as cell chemotaxis, regulation of chemotaxis, and cell-substrate adhesion were upregulated in the predicted high-ALN burden group (Figures S2d-f).

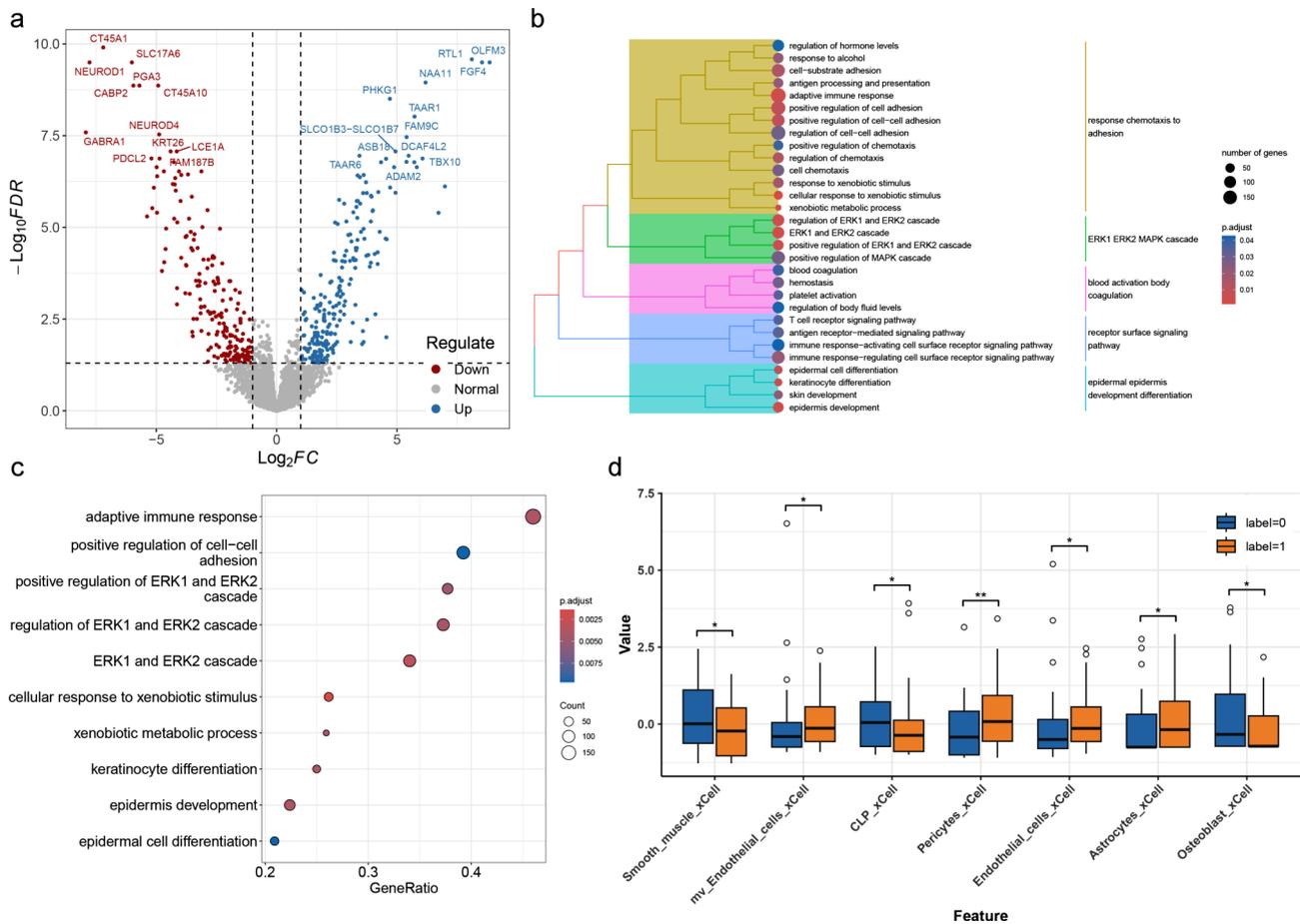


Fig. 4 Transcriptomic and immunological analysis related to radscore. **(a)** Volcano plot of differentially expressed genes, with upregulation indicated in red, downregulation in blue, and nonsignificant genes in gray. **(b)** Clustered heatmap of pathway enrichment analysis. **(c)** Bubble chart of pathway enrichment based on gene ratio, with bubble size representing gene count. **(d)** Box plots of immune cell scores across two radscore groups

Associations between radiomic features and tumor immune infiltration

Significant differences in eight types of RNA-based immune markers between the low and high radscore groups were observed (Fig. 4d). Specifically, Mv Endothelial, Pericytes cells, and others were more abundant in the predicted high-ALN burden group, whereas common lymphocyte precursors, smooth muscle were more prevalent in the predicted low-ALN burden group. These results indicate differences in tumor immunity and the tumor microenvironment between the high- and low-score groups. Correlation analysis between radiomics features and immune cell scores revealed a strong negative correlation between radiomic features and immune scores (Fig. 5).

Discussion

This study developed a radiomic model for assessing the ALN burden in patients with early-stage breast cancer based on the BPNN algorithm and successfully validated its performance across three external cohorts.

Furthermore, we observed a correlation between higher radscores and poorer breast cancer prognosis. The heterogeneity of tumor cells and the tumor microenvironment across different radscore groups was revealed by integrating genomic data.

The performance of our radiomics signature was moderate compared to previous models although judged by AUC values (AUCs of 0.781–0.856 versus 0.810–0.865, respectively) [16–18], and our model exhibited robust generalizability across validation datasets comprising a large number of patients ($n=1211$) scanned under six different scanners. The broad applicability of the model to diverse patient samples may be attributed to the use of the ComBat technique, which effectively mitigates potential center effects [28]. Additionally, this study used the deep learning method to automatically sketch volumn of interest, further improving the reproducibility of radiological features.

The status of ALN metastasis is well-established as a critical factor influencing the prognosis of patients with breast cancer [29, 30]. This study revealed a significant

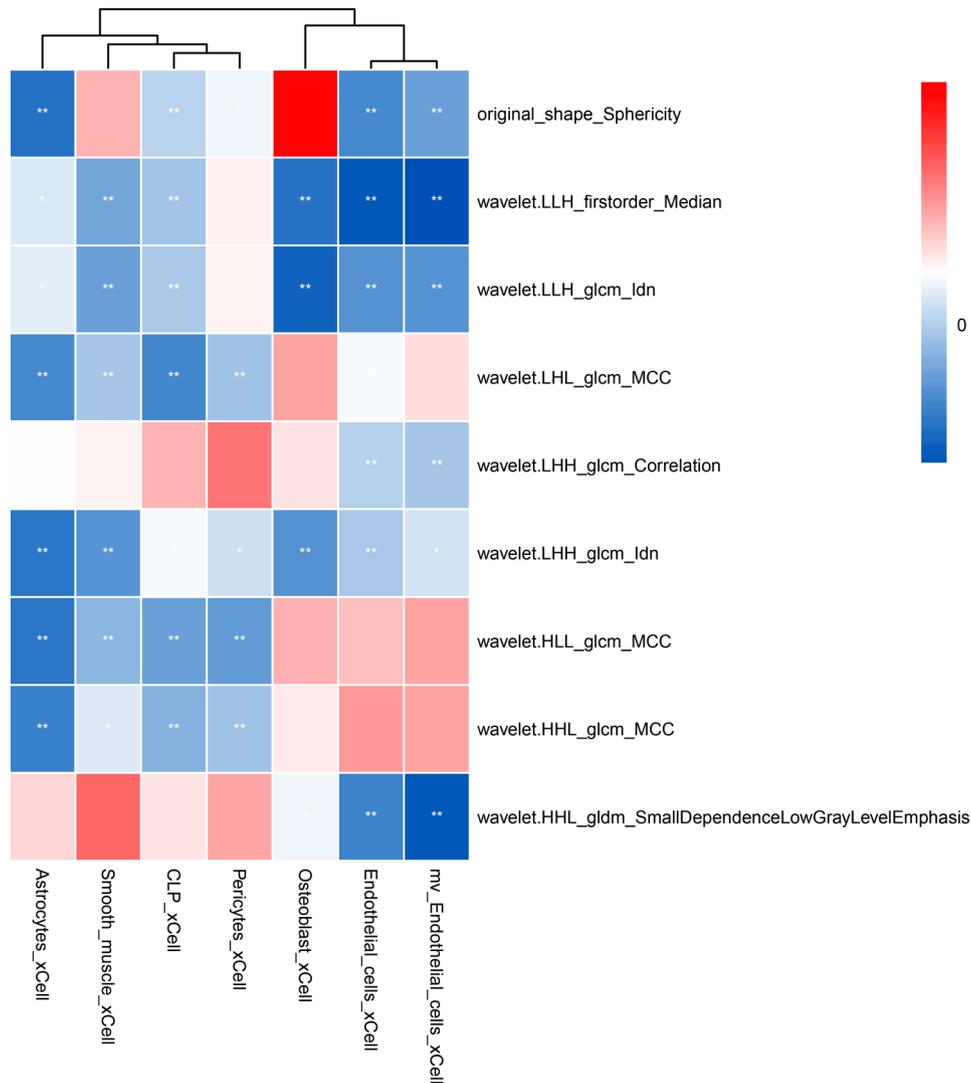


Fig. 5 Spearman correlation heatmap between radiomic features and immune marks. Blue indicates a negative correlation, while red signifies a positive correlation. The deeper the color, the stronger the correlation. CLP, common lymphocyte precursors

correlation between the radscore and breast cancer prognosis, indicating that the lower the radscore, the better the prognosis. This result indicates that radscore can be used as a valuable tool to tailor personalized treatment plans for patients.

Further exploration of RNA-seq data provided insight into the molecular underpinnings of the radiomic predictions in validating the radiomic score. GSEA revealed that the upregulation of cell adhesion pathways was particularly pronounced in individuals with a higher radiomic score. Noteworthy, the pathway from cell chemotaxis to adhesion plays a critical role in the mechanisms of tumor invasion and metastasis, possessing the capacity to facilitate tumor cell migration through the extracellular matrix and develop connections with distant tissues [31]. This process likely represents a potential mechanism associated with the observed increase

in ALN burden. Conversely, pathways associated with epidermal cell differentiation and keratinocyte differentiation were downregulated in the high radiomic score group. The suppression of differentiation pathways indicates a disruption in normal cell function and an inclination toward a more undifferentiated, aggressive tumor phenotype capable of enhanced metastatic dissemination [32]. These results not only improve the biological interpretability of the model at the genetic level but also provide potential therapeutic strategies to target these pathways for curbing tumor dissemination.

Notably, significant differences in immune scores were observed between high and low radscore groups, indicating that radiomics predictions reflect variations in tumor immune status and tumor microenvironment. Higher scores were noted for common lymphocyte precursors within the low radscore group. Common lymphocyte

precursors, closely related to new blood vessel formation, can disrupt the local microenvironment balance through active metabolism, potentially promoting tumor growth [33, 34]. Further analysis revealed a strong negative correlation between radiomic features and immune scores, indicating a potential for stronger immune escape capabilities in tumors within the high radscore group. This observation indicates that tumors may suppress effective immune cell infiltration through specific mechanisms, thereby diminishing the body's immune response to the tumor and resulting in a poorer prognosis [35].

Additionally, the study determined that menopausal status also serves as an independent clinical predictor of ALN burden in early-stage breast cancer along with MRI-ALN status and MRI-ALN burden. This could be attributed to hormone receptor-positive breast cancer being more prevalent in premenopausal women, with hormonal level fluctuations potentially affecting tumor aggressiveness and metastatic potential [36]. However, the addition of the clinical features did not cause a significant incremental increase in predictive value for ALN burden (AUCs of 0.826, 0.812, and 0.823, Delong test, $P > .05$). This phenomenon stemmed from the interplay of coexisting features and the superior predictive performance of radiomic features compared to clinical features.

This study has several limitations. First, being a retrospective analysis, and future research will necessitate larger datasets in multicenter and prospective studies despite the application of an independent external validation cohort. Second, the result that radiomic features may reflect biological characteristics is based on indirect evidence. Plans include conducting targeted modulation of immune cells in breast cancer metastasis mouse models and performing MRI examinations to monitor dynamic changes in radiomic features.

In summary, our study successfully developed and validated an MRI radiomics model for predicting ALN burden in patients with early-stage breast cancer. The model demonstrated robust performance across different imaging centers, along with prognostic significance. Furthermore, our study provided valuable insights into the biological implications of the radiomic predictions, uncovering associations with tumor microenvironment variations, key signaling pathways, and cellular processes.

Abbreviations

ALN	axillary lymph node
TCIA	The Cancer Imaging Archive
BRCA	Breast Cancer
ICC	intraclass correlation coefficient
ER	estrogen receptor
PR	progesterone receptor
HER-2	human epidermal growth factor receptor 2
SLNB	sentinel lymph node biopsy
ALND	axillary lymph node dissection
AUC	area under the curve
ROC	receiver operating characteristic

GO	Gene Ontology
OS	overall survival
DEGs	differentially expressed genes

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-024-05619-4>.

Supplementary Material 1

Acknowledgements

Not applicable.

Author contributions

Experimental design: HMP, FSJ, and FZ. Collection and curation of radiological and clinical data: XZY, FSJ, HMP, ZCY and ZCH. Data analysis and interpretation: LK, LPH, CXB, and HCK. Project supervision: XMS, WJY, HJ and HYQ. Manuscript writing: HMP, CXB, HYQ. All authors contributed to the article and approved the final manuscript.

Funding

The study was supported by the grants of Medical Science and Technology Project of Zhejiang Province (Grant No. 2023KY338, 2024KY454, 2024KY1200), Public Welfare Research Project of Jiaxing (Grant No. 2024AY10026, 2023AD31065), Zhejiang Provincial Outstanding Youth Science Foundation (Grant No. LTGY24H180007), Zhejiang Traditional Chinese Medicine Administration (Grant No. 2024ZL1058, 2024ZL475), and the Natural Science Foundation of Guangdong Province (Grant No. 2023A1515030251).

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The code used in this work can be available on <https://github.com/idcast-Tracy/Image-genomics>.

Declarations

Ethics approval and consent to participate

This study was approved by the institutional review boards of each institution involved and informed consent was waived because of the retrospective study design.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Radiology, Zhejiang Chinese Medical University Affiliated Jiaxing TCM Hospital, Jiaxing, China

²Department of Radiology, The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), Hangzhou, China

³Department of Radiology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China

⁴Department of Radiology, Haining Hospital, The First Affiliated Hospital of Zhejiang University, Haining, China

⁵Department of Radiology, Zhejiang Provincial People's Hospital, Hangzhou, China

⁶Department of Radiology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan, China

⁷Department of Clinical Laboratory, Jiaxing Maternity and Children Health Care Hospital, Jiaxing, China

⁸Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China

⁹The First School of Clinical Medicine, Zhejiang Chinese Medical University, Hangzhou, China

Received: 14 April 2024 / Accepted: 15 August 2024

Published online: 06 September 2024

References

1. Danko ME, Bennett KM, Zhai J, et al. Improved staging in node-positive breast cancer patients using lymph node ratio: results in 1,788 patients with long-term follow-up. *J Am Coll Surg*. 2010;210(1):797–805. <https://doi.org/10.1016/j.jamcollsurg.2010.02.045>.
2. Arnold M, Morgan E, Rungby H, et al. Current and future burden of breast cancer: global statistics for 2020 and 2040. *Breast Edinb Scotl*. 2022;66:15–23. <https://doi.org/10.1016/j.breast.2022.08.010>.
3. Brackstone M, Baldassarre FG, Perera FE, et al. Management of the Axilla in early-stage breast Cancer: Ontario Health (Cancer Care Ontario) and ASCO Guideline. *J Clin Oncol off J Am Soc Clin Oncol*. 2021;39:3056–82. <https://doi.org/10.1200/JCO.21.00934>.
4. Tinterri C, Gentile D, Gatzemeier W, et al. Preservation of Axillary Lymph Nodes Compared with complete dissection in T1-2 breast Cancer patients presenting one or two metastatic Sentinel Lymph nodes: the SINODAR-ONE Multicenter Randomized Clinical Trial. *Ann Surg Oncol*. 2022;29:5732–44. <https://doi.org/10.1245/s10434-022-11866-w>.
5. Giuliano AE, Hunt KK, Ballman KV, et al. Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial. *JAMA*. 2011;305:569–75. <https://doi.org/10.1001/jama.2011.90>.
6. Giuliano AE, Ballman KV, McCall L, 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. <https://doi.org/10.1001/jama.2017.11470>.
7. Rao R, Euhus D, Mayo HG, Balch C. Axillary node interventions in breast cancer: a systematic review. *JAMA*. 2013;310:1385–94. <https://doi.org/10.1001/jama.2013.277804>.
8. Giaquinto AN, Sung H, Miller KD, et al. Breast Cancer statistics, 2022. *CA Cancer J Clin*. 2022;72:524–41. <https://doi.org/10.3322/caac.21754>.
9. Langer I, Guller U, Berclaz G, et al. Morbidity of sentinel lymph node biopsy (SLN) alone versus SLN and completion axillary lymph node dissection after breast cancer surgery: a prospective Swiss multicenter study on 659 patients. *Ann Surg*. 2007;245:452–61. <https://doi.org/10.1097/01.sla.0000245472.47748.ec>.
10. Houvenaeghel G, Cohen M, Raro P, et al. Sentinel node involvement with or without completion axillary lymph node dissection: treatment and pathologic results of randomized SERC trial. *NPJ Breast Cancer*. 2021;7:133. <https://doi.org/10.1038/s41523-021-00336-3>.
11. Di Paola V, Mazzotta G, Pignatelli V, et al. Beyond N staging in breast Cancer: importance of MRI and ultrasound-based imaging. *Cancers*. 2022;14:4270. <https://doi.org/10.3390/cancers14174270>.
12. Li Z, Ma Q, Gao Y, et al. Diagnostic performance of MRI for assessing axillary lymph node status after neoadjuvant chemotherapy in breast cancer: a systematic review and meta-analysis. *Eur Radiol*. 2023. <https://doi.org/10.1007/s00330-023-10155-8>.
13. Pesapane F, De Marco P, Rapino A, et al. How Radiomics can improve breast Cancer diagnosis and treatment. *J Clin Med*. 2023;12:1372. <https://doi.org/10.3390/jcm12041372>.
14. Shi Z, Huang X, Cheng Z, et al. MRI-based quantification of Intratumoral Heterogeneity for Predicting Treatment response to neoadjuvant chemotherapy in breast Cancer. *Radiology*. 2023;308:e222830. <https://doi.org/10.1148/radiol.222830>.
15. Wang Z, Fang M, Zhang J, et al. Radiomics and Deep Learning in Nasopharyngeal Carcinoma: a review. *IEEE Rev Biomed Eng*. 2024;17:118–35. <https://doi.org/10.1109/RBME.2023.3269776>.
16. Zhang X, Yang Z, Cui W, et al. Preoperative prediction of axillary sentinel lymph node burden with multiparametric MRI-based radiomics nomogram in early-stage breast cancer. *Eur Radiol*. 2021;31:5924–39. <https://doi.org/10.1007/s00330-020-07674-z>.
17. Jiang M, Li C-L, Luo X-M, et al. Radiomics model based on shear-wave elastography in the assessment of axillary lymph node status in early-stage breast cancer. *Eur Radiol*. 2022;32:2313–25. <https://doi.org/10.1007/s00330-021-08330-w>.
18. Chen Y, Wang L, Dong X, et al. Deep learning radiomics of preoperative breast MRI for prediction of Axillary Lymph Node Metastasis in breast Cancer. *J Digit Imaging*. 2023;36:1323–31. <https://doi.org/10.1007/s10278-023-00818-9>.
19. Chen W, Lin G, Kong C, et al. Non-invasive prediction model of axillary lymph node status in patients with early-stage breast cancer: a feasibility study based on dynamic contrast-enhanced-MRI radiomics. *Br J Radiol*. 2024;97:439–50. <https://doi.org/10.1093/bjr/tqad034>.
20. Tomaszewski MR, Gillies RJ. The Biological meaning of Radiomic features. *Radiology*. 2021;299:E256. <https://doi.org/10.1148/radiol.2021219005>.
21. Stanzione A, Cuocolo R, Ugga L, et al. Oncologic imaging and Radiomics: a Walkthrough Review of Methodological challenges. *Cancers*. 2022;14:4871. <https://doi.org/10.3390/cancers14194871>.
22. Goldhirsch A, Winer EP, Coates AS, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast Cancer 2013. *Ann Oncol off J Eur Soc Med Oncol*. 2013;24:2206–23. <https://doi.org/10.1093/annonc/mdt303>.
23. Cheang MCU, Chia SK, Voduc D, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst*. 2009;101:736–50. <https://doi.org/10.1093/jnci/djp082>.
24. Abel MK, Greenwood H, Kellil T, et al. Accuracy of breast MRI in evaluating nodal status after neoadjuvant therapy in invasive lobular carcinoma. *NPJ Breast Cancer*. 2021;7:25. <https://doi.org/10.1038/s41523-021-00233-9>.
25. Xu Z, Ding Y, Zhao K, et al. MRI characteristics of breast edema for assessing axillary lymph node burden in early-stage breast cancer: a retrospective bicentric study. *Eur Radiol*. 2022;32:8213–25. <https://doi.org/10.1007/s00330-022-08896-z>.
26. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic phenotype. *Cancer Res*. 2017;77:e104–7. <https://doi.org/10.1158/0008-5472.CAN-17-0339>.
27. Orlhac F, Lecler A, Savatovski J, et al. How can we combat multicenter variability in MR radiomics? Validation of a correction procedure. *Eur Radiol*. 2021;31:2272–80. <https://doi.org/10.1007/s00330-020-07284-9>.
28. Ibrahim A, Refaee T, Primakov S, et al. The effects of In-Plane spatial resolution on CT-Based Radiomic features' Stability with and without ComBat Harmonization. *Cancers*. 2021;13:1848. <https://doi.org/10.3390/cancers13081848>.
29. Bhushan A, Gonsalves A, Menon JU. Current state of breast Cancer diagnosis, treatment, and Theranostics. *Pharmaceutics*. 2021;13:723. <https://doi.org/10.3390/pharmaceutics13050723>.
30. Allison KH. Prognostic and predictive parameters in breast pathology: a pathologist's primer. *Mod Pathol off J U S Can Acad Pathol Inc*. 2021;34:94–106. <https://doi.org/10.1038/s41379-020-00704-7>.
31. B P P B. Current methods for studying metastatic potential of tumor cells. *Cancer Cell Int*. 2022;22. <https://doi.org/10.1186/s12935-022-02801-w>.
32. Moltrasio C, Romagnuolo M, Marzano AV. Epigenetic mechanisms of epidermal differentiation. *Int J Mol Sci*. 2022;23:4874. <https://doi.org/10.3390/ijms23094874>.
33. Ghaedi M, Steer CA, Martinez-Gonzalez I, et al. Common-lymphoid-progenitor-independent pathways of Innate and T Lymphocyte Development. *Cell Rep*. 2016;15:471–80. <https://doi.org/10.1016/j.celrep.2016.03.039>.
34. Salemme V, Centonze G, Cavallo F, et al. The Crosstalk between Tumor Cells and the Immune Microenvironment in breast Cancer: implications for Immunotherapy. *Front Oncol*. 2021;11:610303. <https://doi.org/10.3389/fonc.2021.610303>.
35. Costa A, Kieffer Y, Scholer-Dahirel A, et al. Fibroblast heterogeneity and immunosuppressive environment in human breast Cancer. *Cancer Cell*. 2018;33:463–e47910. <https://doi.org/10.1016/j.ccell.2018.01.011>.
36. Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol*. 2012;13:1141–51. [https://doi.org/10.1016/S1470-2045\(12\)70425-4](https://doi.org/10.1016/S1470-2045(12)70425-4).

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.