## Appendix 1

### *S1 MR image acquisition*

All patients underwent breast MRI examination using a 1.5 Tesla (T) or 3.0T MRI scanner from Siemens (Avanto; Skyra; Siemens, Erlangen, Germany). The imaging protocols for both 1.5T and 3.0T MRI scanners included T1-weighted imaging (T1WI), fat-suppressed T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced MR imaging (DCE).

The scanning parameters for 1.5T MRI scanner were as follows: (I) T1WI: TR 559 ms, TE 12 ms, reverse angle 180°, FOV 380 mm × 380 mm, layer thickness 4.0 mm, layer spacing 0.40 mm, matrix 448×448; (II) Fat-suppressed T2WI: TR 5,000 ms, TE 58 ms, flip angle 150°, FOV 380 mm × 380 mm, layer thickness 4.0 mm, layer spacing 2.0 mm, matrix 384×384. (III) DWI: b=50, 500, 1,000 s/mm$^2$, TR 6,400 ms, TE 97 ms, FOV 340 mm × 340 mm, layer thickness 4.0 mm, layer spacing 2.0 mm, excitement times 2, matrix 192×192. (IV) DCE-MRI: TR 5.16 ms, TE 2.39 ms, flip angle 10°, FOV 340 mm × 340 mm, layer thickness 1.1 mm, layer spacing 0.22 mm, matrix 384×384, scanned before and 90, 150, 210, 270, 330 s after injecting gadopentetatedimeglumine (Gd-DTPA, Magnevist; Bayer Healthcare, Berlin, Germany) at a dose of 0.2 mmol/kg at a rate of 2.0 mL/s and flushed with 20 mL saline.

The scanning parameters for the 3.0T MRI scanner were as follows: (I) T1WI: TR 6 ms, TE 2.46 ms, reverse angle 15°, FOV 360 mm × 360 mm, layer thickness 1.6 mm, layer spacing 0.32 mm, matrix 448×448; (II) fat-suppressed T2WI: TR 3,400 ms, TE 54 ms, flip angle 120°, FOV 340 mm × 340 mm, layer thickness 4.0 mm, layer spacing 0.8 mm, matrix 384×384. (III) DWI: b=50, 400, 800 s/mm$^2$, TR 5,700 ms, TE 59 ms, FOV 340 mm × 340 mm, layer thickness 4.0 mm, layer spacing 0.8 mm, excitement times 3, 4 and 5, matrix 192×192. (IV) DCE-MRI: TR 4.66 ms, TE 1.68 ms, flip angle 10°, FOV 380 mm × 380 mm, layer thickness 1.6 mm, layer spacing 0.32 mm, matrix 448×448, scanned before and 90, 150, 210, 270, 330 s after injecting gadopentetatedimeglumine (Gd-DTPA, Magnevist; Bayer Healthcare, Berlin, Germany) at a dose of 0.2 mmol/kg at a rate of 2.0 mL/s and flushed with 20 mL saline.

### *S2 MR image preprocessing*

The original MRI images were preprocessed before image feature extraction. The specific preprocessing steps are as follows:

(I) Intensity standardization: pixel value intensities were constrained within the range of 0 to 1,000, mitigating the influence of extreme values and outliers.

(II) Spatial normalization: to counteract voxel spacing disparities among various volumes of interest (VOIs), the fixed resolution resampling method was adopted for spatial normalization. This strategy achieved a consistent voxel spacing of 1 mm ×1 mm × 1 mm, enabling precise comparisons and evaluations by aligning the spatial dimensions across images.

### *S3 feature extraction*

In this study, handcrafted features were extracted from intra- and peri-tumoral regions and classified into three categories: (I) geometry, (II) intensity, and (III) texture (39). The geometry features describe the tumor's 3D shape characteristics, while the intensity features depict the first-order statistical distribution of voxel intensities. Additionally, texture features capture patterns, as well as the second and high-order spatial distributions of the intensities. The texture features are extracted using various methods, including the gray-level co-occurrence matrix (GLCM), gray-level dependence matrix (GLDM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), and neighborhood gray-tone difference matrix (NGTDM) methods.

A total of 1,834 hand-crafted features were extracted from intratumoral and peritumoral regions, respectively, comprising 14 geometry features, 360 intensity features, and 1,460 texture features. The proportion of each group of handcrafted features is illustrated in *Figure S2*.

All features were extracted using the PyRadiomics tool (version 3.0.1; https://pyradiomics.readthedocs.io), with a majority aligning with the feature definitions outlined by the Imaging Biomarker Standardization Initiative (IBSI).


https://dx.doi.org/10.21037/qims-24-558

All features were extracted using the PyRadiomics tool (version 3.0.1; https://pyradiomics.readthedocs.io), with a majority aligning with the feature definitions outlined by the Imaging Biomarker Standardization Initiative (IBSI).

### *S4 feature selection*

#### Intraclass correlation coefficient (ICC)

The test-retest and inter-rater analysis were used to assess segmentation uncertainty and ensure feature robustness. One rater performed dual tumor segmentations on 30 randomly chosen patients for test-retest, and two raters did independent tumor segmentations on another 30 randomly chosen patients for inter-rater analysis. ICC evaluated features from multiple-segmented subregions. ICC ≥0.85 indicated robustness against segmentation uncertainties, and these features were included in the subsequent analysis.

#### Normalization and *t*-test

After ICC screening, features were normalized using Z-scores for normal distribution. A t-test calculated p-values for imaging features, with only P value <0.05 retaining radiomic features.

#### Correlation

Pearson's correlation coefficient was calculated to identify highly correlated features (exceeding 0.9). A recursive deletion strategy was employed to filter the features, removing the most redundant one at each step.

#### LASSO

LASSO regression selected the final feature set for the radiomic signature, zeroing irrelevant features' coefficients based on regularization weight λ. Optimal λ was determined via 10-fold cross-validation, selecting the λ with the lowest mean standard error.

In our experiment, the Intra model identified 12 features, the Peri2mm model selected 12 features, the Peri4mm model extracted 15 features, the Peri6mm model recognized 18 features, the Peri8mm model pinpointed 15 features, and the Habitat model identified four features. These selected features were then used to construct the habitat-based radiomics signature. The results of the LASSO 10-fold cross-validation are presented in *Figure S4*.

### *S5 model construction*

In this study, we conducted a comparative assessment of diverse tumor regions in breast cancer, compassing the entire tumor region (Intra), multiple peritumoral regions, and tumor habitat analyses for ALNM prediction in breast cancer patients.

#### Intratumoral radiomics signature (Intra)

Following LASSO feature screening, the finalized features underwent machine learning methods to establish the radiomics signature. The ALNM prediction model was formulated by employing six commonly used machine learning models: the support vector machine (SVM), K-nearest neighbors (KNN), RandomForest, XGBoost, LightGBM, and ExtraTrees. We performed a 5-fold cross-validation to search for the optimal hyperparameters. During the hyperparameter search, we applied the Grid-search algorithm to find the best hyperparameters for each model. The specific hyperparameters for each model are outlined as follows:

(I) RandomForest: employing 300 estimators, a maximum depth of 4, and a minimum sample split of 2.
(II) XGBoost: utilizing 35 estimators, a binary logistic objective, a maximum depth of 3, and a minimum child weight of 0.2.
(III) LightGBM: utilizing 60 estimators with a maximum depth of 2.
(IV) ExtraTrees: employing 1,000 estimators, a maximum depth of 6, and a minimum sample split of 2.
(V) SVM and KNN retained the default parameters.

Once the modeling was completed, a best-fit model was selected from the six constructed models for the subsequent analysis.


https://dx.doi.org/10.21037/qims-24-558

**Peritumoral radiomics signature (PeriXmm)**

Here, 'X' denotes the peritumoral region of breast cancer in mm. Features extracted from the peritumoral areas of breast cancer were selected using the same feature selection process employed for the Intra Radiomics Signature, and the ultimate model was constructed using the identical machine learning algorithm. Similarly, an optimal model was selected for each specific peritumoral region.

**Habitat signature (Habitat)**

Given the reliance on unsupervised clustering algorithms for characterizing the internal tumor habitat, the mean feature values were computed. Furthermore, due to the unsupervised nature of clustering, the ICC assessment was omitted from the feature selection process of Habitat Signature. However, other configurations aligned with those of Intra and Peri models.

The development of 42 models was conducted using six algorithms to individually model clinical features, intratumoral features, peritumoral features within four specific ranges, and habitat features. Additionally, a fusion nomogram was constructed utilizing the logistic regression algorithm. Consequently, a total of 43 models were ultimately established. Ultimately, eight models were selected from the 43 constructed models for subsequent analysis:

(I) clinicopathological features (clinical model) employing the ExtraTrees algorithms.
(II) intratumoral radiomic signatures (Intra model) employing the ExtraTrees algorithm.
(III) peritumoral radiomic signatures of 2 mm dilation (Peri2mm model) employing the LightGBM algorithm.
(IV) peritumoral radiomic signatures of 4mm dilation (Peri4mm model) employing the XGBoost algorithm.
(V) peritumoral radiomic signatures of 6mm dilation (Peri6mm model) employing the RandomForest algorithm.
(VI) peritumoral radiomic signatures of 8mm dilation (Peri8mm model) employing the ExtraTrees algorithm.
(VII) habitat signatures (habitat model) employing the XGBoost algorithm.
(VIII) clinicopathological features, peritumoral radiomic signatures (choose the best from 2–8 mm), and habitat signatures (fusion nomogram model) employing the logistic regression algorithm.

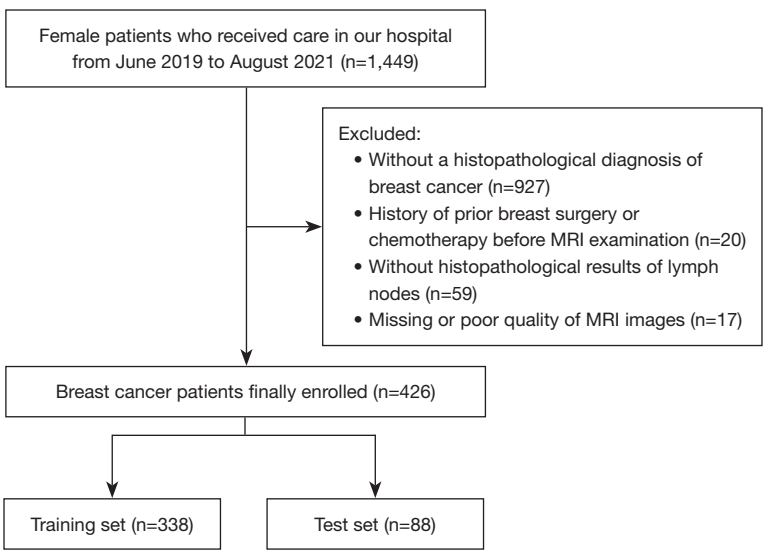


**Figure S1** Flowchart of the patient recruitment process. MRI, magnetic resonance imaging.


https://dx.doi.org/10.21037/qims-24-558

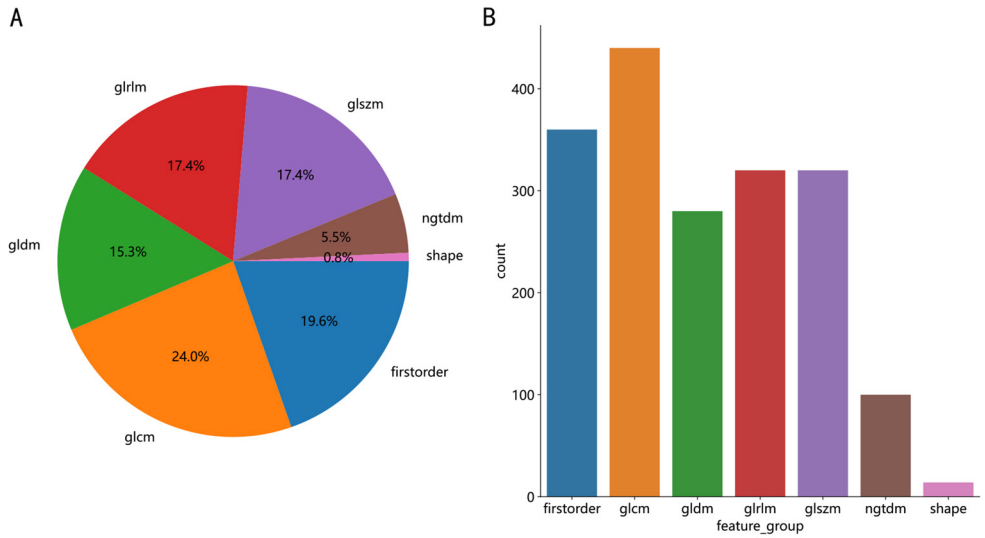


**Figure S2** Ratio (A) and number (B) of handcrafted features extracted from the segmented images of DCE-MRI. DCE-MRI, dynamic contrast-enhanced magnetic resonance imaging.

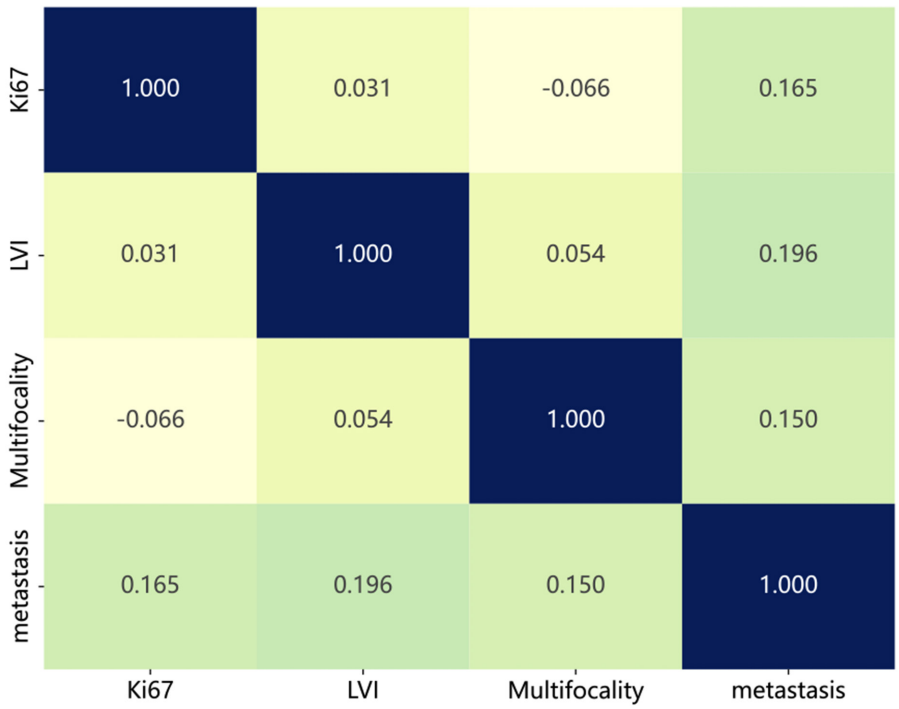


**Figure S3** Spearman correlation coefficients of each clinicopathological feature of breast cancer whose P value <0.05 in univariable regression analysis. LVI, lymphovascular invasion; Ki67, proliferation marker protein Ki67.


https://dx.doi.org/10.21037/qims-24-558

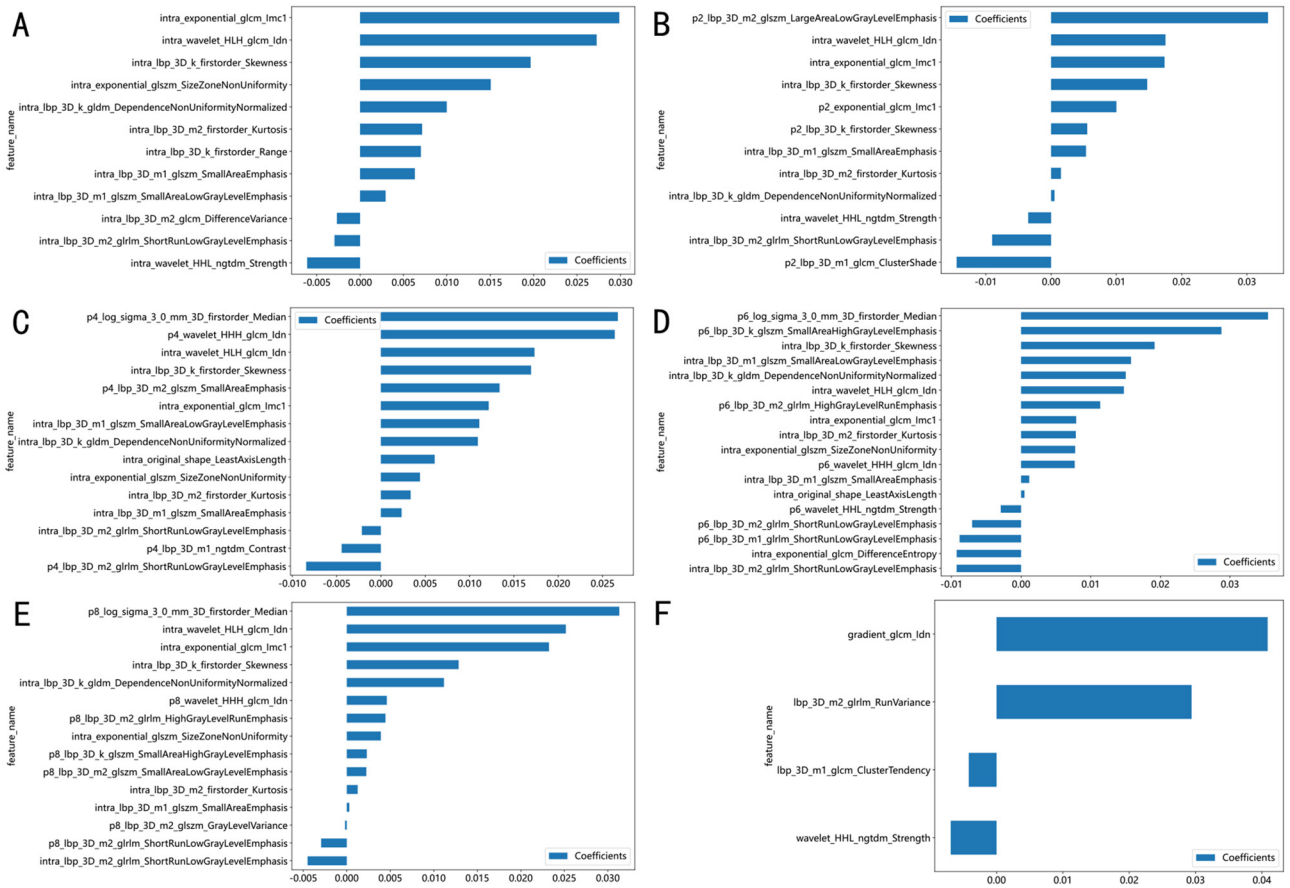


**Figure S4** The feature weights obtained using the LASSO feature selection algorithm for the following models: Intra (A), Peri2mm (B), Peri4mm (C), Peri6mm (D), Peri8mm (E), and Habitat (F). LASSO, least absolute shrinkage and selection operator.


https://dx.doi.org/10.21037/qims-24-558

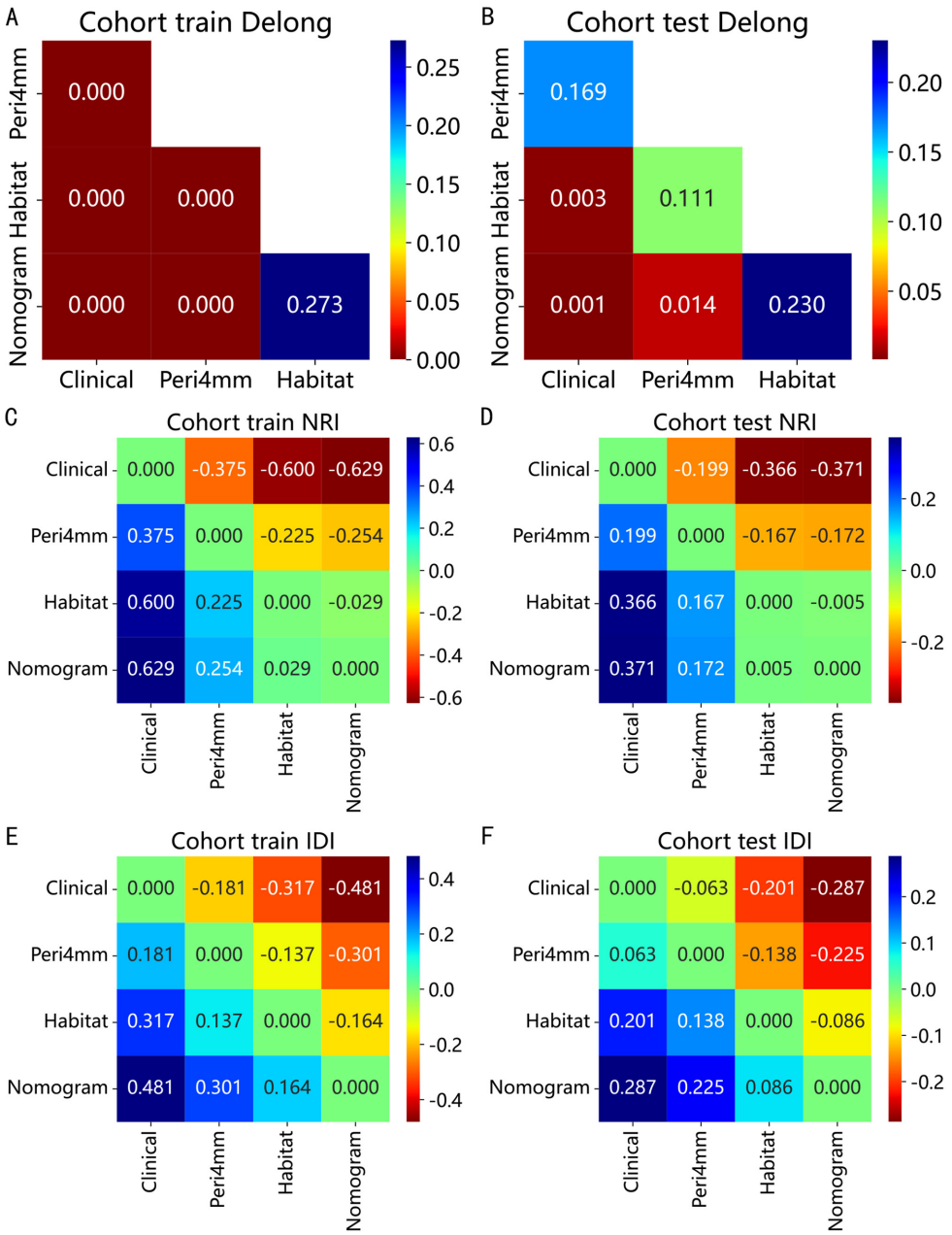


**Figure S5** The Delong, NRI, and IDI metrics of the Clinical, Peri4mm, Habitat, and Nomogram models in the training and test sets. Delong metrics of the four models in the training (A) and test sets (B). NRI metrics of the four models in the training (C) and test sets (D). IDI metrics of the four models in the training (E) and test sets (F). NRI, net reclassification improvement; IDI, integrated discrimination improvement.


https://dx.doi.org/10.21037/qims-24-558

**Table S1** Univariable and multivariable analysis of associations between clinicopathological features and ALNM in breast cancer patients

| Feature name | Univariable analysis | | | | Multivariable analysis | | | |
|---|---|---|---|---|---|---|---|---|
| | OR | 95% CI | | P value | OR | 95% CI | | P value |
| | | Lower | Upper | | | Lower | Upper | |
| Age | 0.995 | 0.991 | 0.999 | 0.050 | | | | |
| ER | 0.999 | 0.900 | 1.108 | 0.987 | | | | |
| Histological grade | 1.004 | 0.921 | 1.094 | 0.935 | | | | |
| PR | 1.045 | 0.930 | 1.175 | 0.537 | | | | |
| HER-2 | 1.151 | 1.012 | 1.310 | 0.073 | | | | |
| Ki67 | 1.153 | 1.055 | 1.260 | 0.008 | 1.176 | 1.078 | 1.281 | 0.002 |
| Multifocality | 1.256 | 1.110 | 1.420 | 0.003 | 1.263 | 1.119 | 1.426 | 0.001 |
| LVI | 1.274 | 1.131 | 1.436 | 0.001 | 1.259 | 1.120 | 1.416 | 0.002 |

ALNM, axillary lymph node metastasis; OR_UNI, odds ratio of univariable analysis; CI, confidence interval; OR_Mul, odds ratio of multivariable analysis; ER, estrogen receptor; PR, progesterone receptor; HER-2, human epidermal growth factor receptor 2; Ki67, proliferation marker protein Ki67; LVI, lymphovascular invasion.

**Table S2** Performance of the intratumoral and peritumoral radiomic signature models in the training and test sets

| Signature | AUC (95% CI) | Accuracy | Sensitivity | Specificity | PPV | NPV | Threshold |
|---|---|---|---|---|---|---|---|
| Training set | | | | | | | |
| Intra | 0.75 (0.699–0.802) | 0.695 | 0.604 | 0.763 | 0.654 | 0.722 | 0.455 |
| Peri2mm | 0.851 (0.812–0.891) | 0.760 | 0.812 | 0.722 | 0.684 | 0.838 | 0.398 |
| Peri4mm | 0.871 (0.834–0.908) | 0.796 | 0.757 | 0.825 | 0.762 | 0.821 | 0.457 |
| Peri6mm | 0.85 (0.811–0.890) | 0.766 | 0.743 | 0.784 | 0.718 | 0.804 | 0.421 |
| Peri8mm | 0.746 (0.694–0.798) | 0.683 | 0.688 | 0.680 | 0.615 | 0.746 | 0.422 |
| Test set | | | | | | | |
| Intra | 0.745 (0.643–0.847) | 0.739 | 0.405 | 0.980 | 0.937 | 0.694 | 0.506 |
| Peri2mm | 0.769 (0.671–0.866) | 0.693 | 0.811 | 0.608 | 0.600 | 0.816 | 0.358 |
| Peri4mm | 0.773 (0.672–0.873) | 0.716 | 0.838 | 0.627 | 0.620 | 0.842 | 0.378 |
| Peri6mm | 0.725 (0.620–0.830) | 0.636 | 0.919 | 0.431 | 0.540 | 0.880 | 0.311 |
| Peri8mm | 0.719 (0.611–0.827) | 0.693 | 0.568 | 0.784 | 0.656 | 0.714 | 0.445 |

AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.


https://dx.doi.org/10.21037/qims-24-558

**Table S3** Performance of the Clinical, Peri4mm, Habitat, and Nomogram models in the training and test sets

| Signature | AUC (95% CI) | Accuracy | Sensitivity | Specificity | PPV | NPV | Threshold |
|---|---|---|---|---|---|---|---|
| Training set | | | | | | | |
| Clinical | 0.655 (0.600–0.711) | 0.577 | 0.819 | 0.399 | 0.502 | 0.748 | 0.438 |
| Peri4mm | 0.871 (0.834–0.908) | 0.796 | 0.757 | 0.825 | 0.762 | 0.821 | 0.457 |
| Habitat | 0.973 (0.959–0.987) | 0.911 | 0.924 | 0.902 | 0.875 | 0.941 | 0.436 |
| Nomogram | 0.977 (0.965–0.989) | 0.923 | 0.875 | 0.959 | 0.940 | 0.912 | 0.529 |
| Test set | | | | | | | |
| Clinical | 0.667 (0.558–0.777) | 0.614 | 0.703 | 0.549 | 0.531 | 0.718 | 0.438 |
| Peri4mm | 0.773 (0.672–0.873) | 0.716 | 0.838 | 0.627 | 0.620 | 0.842 | 0.378 |
| Habitat | 0.854 (0.778–0.931) | 0.807 | 0.811 | 0.804 | 0.750 | 0.854 | 0.403 |
| Nomogram | 0.873 (0.802–0.945) | 0.818 | 0.730 | 0.882 | 0.818 | 0.818 | 0.500 |

AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

**Table S4** HL test statics in the training and test sets for the Clinical, Peri4mm, Habitat, and Nomogram models

| Model | Training set | Test set |
|---|---|---|
| Clinical | 0.933 | 0.868 |
| Peri4mm | 0.293 | 0.736 |
| Habitat | 0.238 | 0.301 |
| Nomogram | 0.828 | 0.406 |

HL, Hosmer-Lemeshow.


https://dx.doi.org/10.21037/qims-24-558