



OPEN ACCESS

EDITED BY

Habibollah Dadgar,
Independent researcher, Mashad, Iran

REVIEWED BY

Lingyun Wang,
Shanghai Jiao Tong University, China
Yu Feng,
The First Affiliated Hospital of Soochow
University, China

*CORRESPONDENCE

Xiaoyuan Sun
✉ 1340565393@qq.com

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

RECEIVED 05 July 2025

ACCEPTED 03 October 2025

PUBLISHED 15 October 2025

CITATION

Dong N, Yan Y, Li Y, Li G, Wang P, Li L,
Zhang H, Sheng H and Sun X (2025)
CT-based habitat radiomics for preoperative
differentiation of adenocarcinoma in situ/
minimally invasive adenocarcinoma from
invasive adenocarcinoma manifesting as
ground-glass nodules: a multicenter study.
Front. Oncol. 15:1660071.
doi: 10.3389/fonc.2025.1660071

COPYRIGHT

© 2025 Dong, Yan, Li, Li, Wang, Li, Zhang,
Sheng and Sun. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

CT-based habitat radiomics for preoperative differentiation of adenocarcinoma in situ/ minimally invasive adenocarcinoma from invasive adenocarcinoma manifesting as ground-glass nodules: a multicenter study

Ning Dong^{1†}, Yumao Yan^{2†}, Yunxin Li¹, Guochao Li¹,
Ping Wang³, Lin Li¹, Hu Zhang⁴, Hui Sheng¹ and Xiaoyuan Sun^{1*}

¹Department of Radiology, Yantaishan Hospital, Yantai, Shandong, China, ²Department of Radiology, Yantai Qishan Hospital, Yantai, Shandong, China, ³Department of Radiology, Yantai Yuhuangding Hospital, Yantai, Shandong, China, ⁴Department of Radiology, Affiliated Hospital of Binzhou Medical University, Binzhou, Shandong, China

Objectives: To develop a CT-based habitat radiomics model for preoperative differentiation of adenocarcinoma in situ/minimally invasive adenocarcinoma (AIS/MIA) from invasive adenocarcinoma (IAC) manifesting as ground-glass nodules (GGNs), and to construct a combined model integrating clinical risk factors for optimizing individualized treatment decisions.

Methods: We retrospectively collected imaging and clinical data from 630 patients with pathologically confirmed ground-glass nodules (GGNs) who underwent surgical resection at two medical centers between January 2020 and December 2024. Patients from Center 1 were randomly divided into training and internal validation sets at a 7:3 ratio, while patients from Center 2 served as the external validation set. Tumor habitats were generated using K-means clustering, and radiomics features were extracted from intratumoral, peritumoral 1mm, peritumoral 2mm, and habitat regions. Feature selection was performed using Least Absolute Shrinkage and Selection Operator (LASSO) regression, and predictive models were constructed using multiple machine learning algorithms. A combined nomogram was developed by integrating the Habitat model, Intratumoral model, and Clinic model. Model performance was evaluated using receiver operating characteristic (ROC) curves, calibration curves, and decision curve analysis (DCA).

Results: In the training set, the Combined model demonstrated optimal performance (AUC = 0.928), followed by the Habitat model (AUC = 0.924), both significantly outperforming the Intratumoral model (AUC = 0.879), Peritumoral 1mm model (AUC = 0.874), Peritumoral 2mm model (AUC = 0.868), and Clinic model (AUC = 0.807) ($P < 0.05$). In the external validation set, the Combined model maintained superior performance (AUC = 0.897), significantly exceeding all other models ($P < 0.05$). The Habitat model showed

the second-best performance in external validation (AUC = 0.840). Hosmer-Lemeshow test and calibration curves demonstrated good calibration for both the Combined and Habitat models across all cohorts. DCA indicated high net benefit for both models in clinical applications.

Conclusion: CT-based habitat radiomics effectively quantifies intratumoral heterogeneity, significantly improving the differentiation between AIS/MIA and IAC. The combined nomogram integrating habitat features, intratumoral features, and clinical factors demonstrates excellent diagnostic performance and generalizability, providing a reliable preoperative assessment tool for individualized treatment decision-making in ground-glass nodular lung adenocarcinoma.

KEYWORDS

computed tomography, habitat, lung adenocarcinoma, radiomics, ground-glass nodules

1 Introduction

Lung cancer remains the most prevalent cancer type globally and the leading cause of cancer-related mortality (1). Lung adenocarcinoma represents the most common histological subtype of lung cancer (2). With the widespread implementation of low-dose computed tomography (CT) in lung cancer screening, the detection rate of ground-glass nodules (GGNs) has increased substantially (3), with GGNs being a common manifestation of lung adenocarcinoma (4). The 2021 World Health Organization Classification of Thoracic Tumors categorizes lung adenocarcinoma into precursor glandular lesions (including atypical adenomatous hyperplasia and adenocarcinoma *in situ* [AIS]), minimally invasive adenocarcinoma (MIA), and invasive adenocarcinoma (IAC) (5). AIS/MIA demonstrates excellent prognosis with a 5-year disease-free survival (DFS) rate of 100% after surgery (6), whereas IAC shows poorer outcomes with 5-year DFS rates ranging from 38% to 86% (7, 8). Surgical approaches also differ significantly: lobectomy remains the standard treatment for IAC, while sublobar resection is preferred for AIS/MIA (9). Therefore, accurate preoperative differentiation between AIS/MIA and IAC is crucial for developing individualized treatment strategies and avoiding overtreatment or undertreatment.

Conventional imaging examinations have limitations in differentiating the invasiveness of GGNs. Although nodule size, morphological features, and density correlate with invasiveness, these qualitative or semi-quantitative assessment methods are subjective and demonstrate limited accuracy in distinguishing AIS/MIA from IAC (10, 11). Radiomics, an emerging artificial intelligence-based imaging analysis approach, efficiently extracts high-throughput feature information from massive medical images, encompassing shape, texture, signal intensity, and numerous other aspects. These rich and detailed features have been widely applied in

disease diagnosis, prognosis assessment, and treatment response monitoring, demonstrating significant clinical value and development potential (12–14). Recent years have witnessed substantial progress in CT radiomics-based differentiation of AIS/MIA from IAC. These studies provide important evidence for early diagnosis and treatment decision-making of IAC by extracting and analyzing high-dimensional features from CT images (8, 15–17). Despite their innovation and promising predictive performance, these studies treat the entire tumor as a single region of interest (ROI) for feature extraction, overlooking the significant heterogeneity characteristic of ground-glass nodular lung adenocarcinoma (18).

The tumor microenvironment plays a pivotal role in shaping tumor heterogeneity. The diversity of stromal cell types and functional heterogeneity directly sculpts the complex environmental landscape within tumors (19). Stromal components including cancer-associated fibroblasts, tumor-associated macrophages, and vascular endothelial cells create spatially heterogeneous microenvironmental gradients through secretion of different growth factors and cytokines. This spatial microenvironmental heterogeneity further promotes adaptive evolution of tumor cells under selective pressure, leading to the emergence of tumor cell subpopulations with different phenotypic and functional characteristics, ultimately forming complex patterns of intratumoral heterogeneity (20). Given this inherent spatial complexity within tumors, traditional radiomics approaches that analyze tumors as single homogeneous entities may inadequately capture the full spectrum of biological diversity present in these heterogeneous tissues (21, 22). To address this limitation and better reflect the spatial complexity of tumor biology, habitat radiomics quantifies intratumoral heterogeneity by segmenting complex tumors into distinct subregions (called habitats) (23). This approach overcomes the limitation of traditional radiomics that treats tumors as

homogeneous entities, enabling deeper analysis of biological differences between tumor regions and providing more reliable imaging evidence for personalized treatment strategies (24). Multiple studies have demonstrated the promising application value of habitat radiomics in predicting glioma molecular markers, Human Epidermal Growth Factor Receptor 2 expression status in breast cancer, and lymphovascular space invasion in cervical cancer (25–27). The peritumoral region, as an integral component of the tumor microenvironment, contains information related to tumor molecular subtypes, invasiveness, and lymph node metastasis, holding significant value in tumor molecular subtyping, prognosis assessment, and metastasis prediction (28–30).

This study aims to develop a CT-based habitat radiomics model for differentiating AIS/MIA from IAC manifesting as GGNs. Furthermore, we integrate the habitat model with intratumoral (or peritumoral) features and clinical risk factors to construct a combined nomogram model, providing clinicians with more comprehensive and accurate diagnostic evidence to optimize individualized treatment decision-making.

2 Materials and methods

2.1 Patients

This multicenter study was approved by the ethics committees of Yantaishan Hospital and Affiliated Hospital of Binzhou Medical University. Given the retrospective nature of this study, the requirement for informed consent was waived. Figure 1 illustrates the specific workflow of this study.

We retrospectively collected imaging and clinical data from patients with GGNs who underwent surgical resection at Center 1 (Yantaishan Hospital) and Center 2 (Affiliated Hospital of Binzhou Medical University) between January 2020 and December 2024. Inclusion criteria were as follows (1): pathologically confirmed AIS, MIA, or IAC after surgery (2); nodule long diameter <3 cm measured on lung window (window width: 1200 Hounsfield Units [HU]; window level: -600 HU) (3); thin-slice chest CT examination within two weeks before surgical resection, with slice thickness less than 2 mm. Exclusion criteria were (1): poor CT

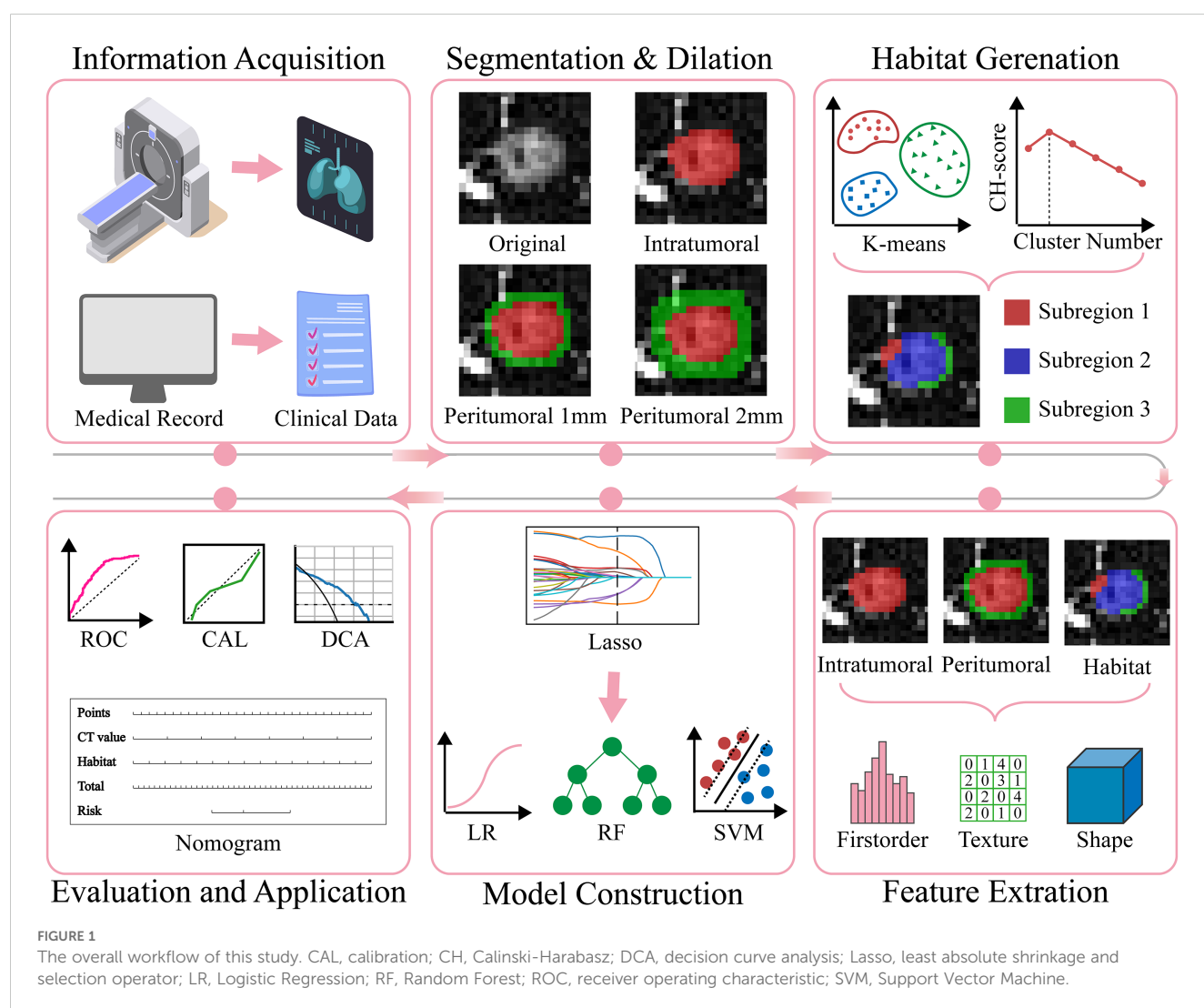
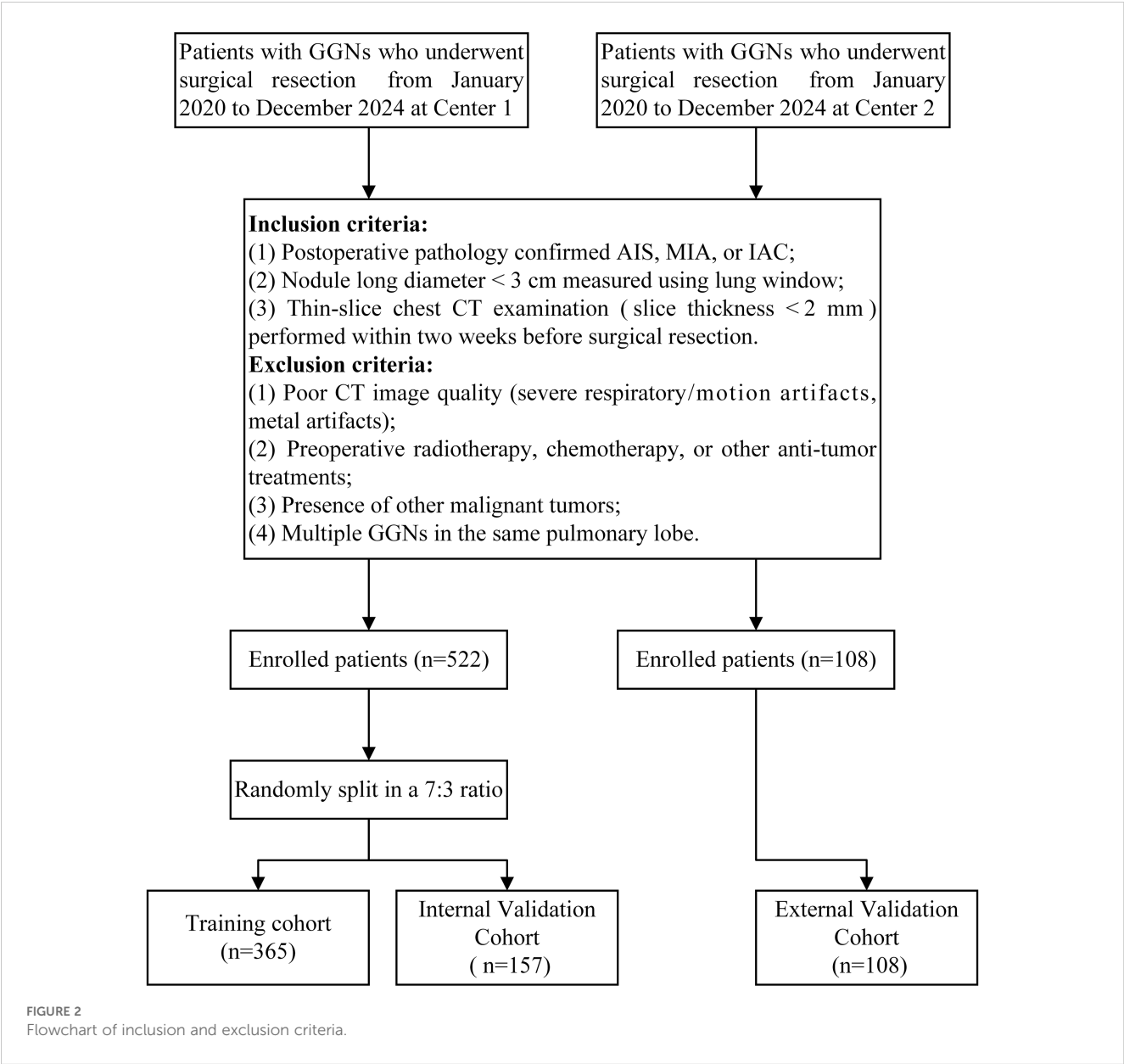


image quality (severe respiratory artifacts, metal artifacts, etc.) (2); previous radiotherapy, chemotherapy, or other antitumor treatment (3); concomitant other malignancies (4); multiple GGNs in the same lobe. Ultimately, 630 GGNs from 630 eligible patients were included in this study. The 522 GGNs from Center 1 were randomly divided into training and internal validation sets at a 7:3 ratio, while the 108 GGNs from Center 2 served as the external validation set (Figure 2).

2.2 Image acquisition and preprocessing

This study employed a multicenter imaging acquisition protocol. Both centers were equipped with CT scanners from Philips Medical Systems (Cleveland, USA), including Brilliance 64, Brilliance 128, and

Incisive 64. All patients received standardized breathing training before scanning and were positioned supine (head first, arms raised and placed beside the head). Scanning was performed at maximum inspiratory breath-hold. For the pulmonary nodule region, a targeted scanning protocol was used to obtain non-contrast high-resolution images with the following parameters: tube voltage 120 kV, tube current 300 mA, pitch 0.6, collimation 0.625 mm × 64, matrix size 1024 × 1024, field of view 200 mm, reconstruction slice thickness 0.670 mm, reconstruction slice interval 0.340 mm, and sharp reconstruction algorithm. To reduce inter-equipment variability and improve the comparability and reproducibility of radiomics features, thereby enhancing model robustness and generalizability, voxel spacing was first resampled to 1 mm × 1 mm × 1 mm using nearest neighbor interpolation, followed by histogram standardization of intensity values.



2.3 ROI segmentation and peritumoral region generation

A junior radiologist A (5 years of experience in chest imaging diagnosis) used ITK-SNAP software (version 3.8.0; <http://www.itksnap.org>) to manually delineate ROI along nodule edges layer by layer under lung window settings (window width: 1200 HU; window level: -600 HU) until the entire nodule was covered, obtaining three-dimensional volume of interest (VOI). Large vessels and bronchi within nodules were carefully excluded during delineation. Subsequently, a senior radiologist B (20 years of experience in chest imaging diagnosis) reviewed the delineation results. Disagreements between the two radiologists were resolved through consensus. Both radiologists were blinded to pathological results throughout the process to ensure objectivity. Finally, using the VOI outer surface as a reference, morphological dilation algorithms were applied to generate peritumoral regions extending 1 mm and 2 mm outward. Non-lung tissues such as chest wall, ribs, and heart covered during the dilation process were manually excluded.

2.4 Habitat generation

To generate tumor habitats, 12 local features were extracted from each voxel within the three-dimensional VOI (Figure 3 shows feature visualization), followed by K-means clustering to delineate

habitat regions. Cluster numbers from 2 to 9 were evaluated, with the optimal number selected based on Calinski-Harabasz scores (31). Specific details regarding habitat generation are provided in the Supplementary Materials.

2.5 Feature extraction and selection

Multi-regional radiomics feature extraction was performed using the PyRadiomics platform (version 3.0.1), including (1): intratumoral region (2); peritumoral 1mm region (3); peritumoral 2mm region (4); tumor habitat regions. Feature extraction strictly followed the Imaging Biomarker Standardization Initiative guidelines (32), encompassing three major categories (1): first-order statistics features, characterizing signal intensity distribution (2); shape features, quantifying spatial geometric attributes of lesions (3); higher-order texture features, analyzing inter-pixel correlation patterns through Gray Level Co-occurrence Matrix, Gray Level Dependence Matrix, Gray Level Run Length Matrix, Gray Level Size Zone Matrix, and Neighboring Gray Tone Difference Matrix to characterize microscopic heterogeneity.

To assess feature extraction consistency and reproducibility, 30 GGNs were randomly selected for independent ROI segmentation by radiologists A and B, with interclass correlation coefficients calculated; 2 weeks later, radiologist A repeated segmentation of the same nodules to calculate intraclass correlation coefficients. Features with both intraclass and interclass correlation coefficients

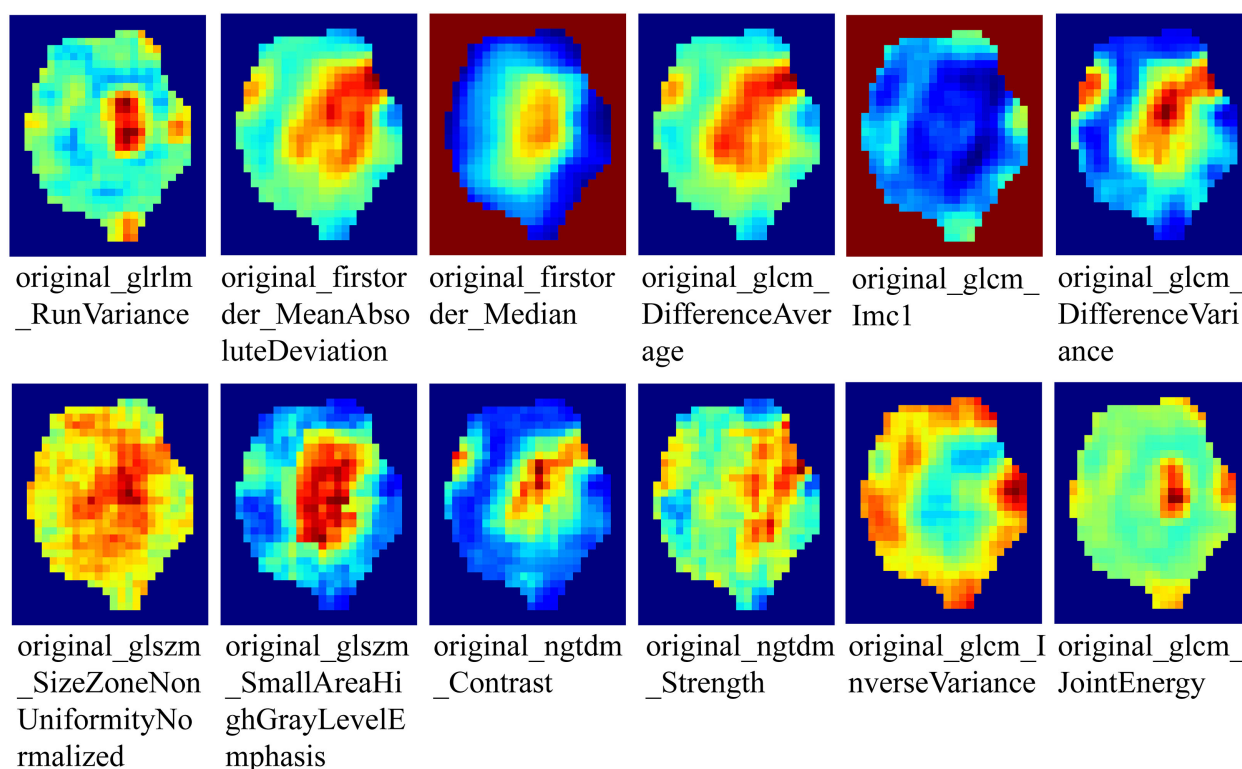


FIGURE 3
Visualization of local features of voxels within the VOI.

greater than 0.75 were retained for subsequent analysis. Due to the unsupervised nature of clustering, this process was omitted for habitat model feature selection. Feature values were standardized using Z-score normalization based on the mean and standard deviation of the training cohort to eliminate scale effects. Features with $p < 0.05$ by t-test were retained. Pearson correlation coefficients were calculated to identify highly correlated features, with a threshold of 0.9. The minimum Redundancy Maximum Relevance algorithm was used to select the top 30 features most relevant to outcomes with low mutual redundancy. To further improve model generalizability, the Least Absolute Shrinkage and Selection Operator (LASSO) regression model was constructed on the training set, with the optimal regularization parameter λ determined through 10-fold cross-validation. Features with non-zero coefficients based on the optimal λ value were selected for final predictive model construction.

2.6 Model construction

In this study, we constructed the following four radiomics models based on different regions (1): Intratumoral (Intra) model (2); Peritumoral 1mm (Peri 1mm) model (3); Peritumoral 2mm (Peri 2mm) model (4); Habitat model. For Clinic model construction, we first performed univariate logistic regression analysis on all clinical and imaging variables, selecting variables with $p < 0.05$, followed by multivariate logistic regression analysis to identify independent risk factors for IAC for modeling. For radiomics and clinic model construction, we employed various advanced machine learning algorithms, including Logistic Regression, Support Vector Machine, Random Forest, eXtreme Gradient Boosting, and Light Gradient Boosting Machine. To ensure model performance and stability, we used five-fold cross-validation and grid search algorithms to determine optimal hyperparameters for each algorithm. To construct the combined model, we performed a comprehensive evaluation of model performance, complementarity, and clinical applicability. The Habitat model demonstrated superior performance with strong generalizability (external validation AUC = 0.840), while the Intra model provided comprehensive tumor characterization (external validation AUC = 0.756). Although the peritumoral models showed predictive ability, their relatively lower performance (Peri 1mm external validation AUC = 0.747; Peri 2mm external validation AUC = 0.730) led to their exclusion from the final combined model. Additionally, including too many radiomics models could increase model complexity and risk of overfitting. Finally, the Intra model, Habitat model, and Clinic model were integrated to construct a combined model, visualized in nomogram form.

2.7 Model evaluation

Model performance was evaluated using receiver operating characteristic (ROC) curve metrics, specifically including area

under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value, and negative predictive value. Through comparison and analysis of different machine learning algorithms, the algorithm with the maximum AUC in the internal validation set was selected as the basis for constructing corresponding radiomics and Clinic models. To validate differences in predictive performance between models, pairwise comparisons were performed using the DeLong test. For model calibration assessment, calibration curves were used to visually present the consistency between predicted probabilities and actual occurrence probabilities, with the Hosmer-Lemeshow test providing quantitative assessment of calibration ability. Decision curve analysis (DCA) was employed to evaluate the clinical net benefit of models at different risk thresholds.

2.8 Statistical analysis

Statistical analysis was performed using SPSS (version 26.0) and Python (version 3.9.7). Continuous variables were expressed as mean $\pm$ standard deviation, with between-group comparisons using independent sample t-tests. Categorical variables were expressed as frequencies and percentages, with between-group differences compared using chi-square tests or Fisher's exact test. All statistical tests were two-sided, with $p < 0.05$ considered statistically significant.

3 Results

3.1 Patient characteristics

This study included 630 patients from two centers, comprising 365 patients in the training set (mean age 56.46 ± 11.59 years), 157 patients in the internal validation set (mean age 55.79 ± 11.43 years), and 108 patients in the external validation set (mean age 53.75 ± 10.75 years). Age, long diameter, short diameter, CT value, lobulation, spiculation, vessel changes, shape, and type showed statistically significant differences between AIS/MIA and IAC groups across all three cohorts. Specifically, in all cohorts, the IAC group had higher age of onset than the AIS/MIA group, with larger long and short diameters and higher CT values. Additionally, the incidence of lobulation, spiculation, vessel changes, round nodules, and mixed ground-glass nodules was significantly higher in the IAC group than in the AIS/MIA group. Detailed data are presented in [Table 1](#).

3.2 Habitat generation

When generating habitat subregions, we evaluated subregion numbers from 2 to 9. As shown in [Supplementary Figure S1](#), the Calinski-Harabasz score increased when the number of subregions increased from 2 to 3, then gradually decreased, indicating the optimal number of subregions was 3. The different subregions were named Habitat 1, Habitat 2, and Habitat 3.

TABLE 1 Comparison of clinical and imaging characteristics between the AIS/MIA group and the IAC group in the three cohorts.

Features	Training cohort (n=365)		P	Internal validation cohort (n=157)		P	External validation cohort (n=108)		P
	AIS/MIA (n=195)	IAC (n=170)		AIS/MIA (n=85)	IAC (n=72)		AIS/MIA (n=70)	IAC (n=38)	
Age (years)	52.76 ± 12.24	60.69 ± 9.14	<0.001	52.35 ± 11.77	59.85 ± 9.60	<0.001	51.03 ± 10.21	58.76 ± 10.01	<0.001
CEA (ng/ml)	1.74 ± 1.22	2.08 ± 1.17	<0.001	1.83 ± 1.46	3.22 ± 11.17	0.189	1.73 ± 0.81	2.20 ± 1.67	0.256
NSE (ng/ml)	15.53 ± 5.56	15.59 ± 4.58	0.496	14.05 ± 4.15	16.30 ± 6.48	0.040	12.79 ± 3.37	13.05 ± 3.74	0.731
Long diameter (mm)	9.12 ± 2.66	15.64 ± 5.79	<0.001	9.67 ± 3.55	15.20 ± 5.57	<0.001	8.86 ± 4.50	15.40 ± 5.57	<0.001
Short diameter (mm)	7.07 ± 2.07	11.06 ± 4.25	<0.001	7.46 ± 2.74	10.63 ± 3.87	<0.001	7.07 ± 2.15	11.76 ± 4.10	<0.001
CT value (HU)	476.56 ± 175.74	333.11 ± 181.83	<0.001	466.15 ± 180.94	329.00 ± 180.22	<0.001	484.31 ± 184.17	334.37 ± 220.86	<0.001
Gender [n (%)]			0.072			0.457			0.026
Male	71(36.41)	46(27.06)		26(30.59)	27(37.50)		14(20.00)	16(42.11)	
Female	124(63.59)	124(72.94)		59(69.41)	45(62.50)		56(80.00)	22(57.89)	
Smoking [n (%)]			0.770			0.025			0.087
No	175(89.74)	150(88.24)		78(91.76)	56(77.78)		62(88.57)	28(73.68)	
Yes	20(10.26)	20(11.76)		7(8.24)	16(22.22)		8(11.43)	10(26.32)	
Location [n (%)]			0.932			0.436			0.705
RUL	61(31.28)	53(31.18)		29(34.12)	21(29.17)		24(34.29)	14(36.84)	
RML	12(6.15)	12(7.06)		4(4.71)	7(9.72)		4(5.71)	1(2.63)	
RLL	29(14.87)	30(17.65)		16(18.82)	15(20.83)		11(15.71)	9(23.68)	
LUL	58(29.74)	48(28.24)		27(31.76)	17(23.61)		21(30.00)	8(21.05)	
LLL	35(17.95)	27(15.88)		9(10.59)	12(16.67)		10(14.29)	6(15.79)	
Lobulation [n (%)]			<0.001			<0.001			<0.001
No	71(36.41)	11(6.47)		33(38.82)	3(4.17)		36(51.43)	2(5.26)	
Yes	124(63.59)	159(93.53)		52(61.18)	69(95.83)		34(48.57)	36(94.74)	
Spiculation [n (%)]			<0.001			0.003			<0.001
No	156(80.00)	81(47.65)		60(70.59)	33(45.83)		59(84.29)	15(39.47)	
Yes	39(20.00)	89(52.35)		25(29.41)	39(54.17)		11(15.71)	23(60.53)	
Margin [n (%)]			<0.001			0.100			0.003
Clear	140(71.79)	92(54.12)		58(68.24)	39(54.17)		51(72.86)	16(42.11)	
Unclear	55(28.21)	78(45.88)		27(31.76)	33(45.83)		19(27.14)	22(57.89)	
Vessel changes [n (%)]			<0.001			<0.001			<0.001
No	102(52.31)	24(14.12)		44(51.76)	5(6.94)		36(51.43)	5(13.16)	
Yes	93(47.69)	146(85.88)		41(48.24)	67(93.06)		34(48.57)	33(86.84)	
Bubble lucency [n (%)]			0.007			0.057			0.018
No	148(75.90)	106(62.35)		71(83.53)	50(69.44)		49(70.00)	17(44.74)	
Yes	47(24.10)	64(37.65)		14(16.47)	22(30.56)		21(30.00)	21(55.26)	
Pleural retraction [n (%)]			<0.001			<0.001			0.076

(Continued)

TABLE 1 Continued

Features	Training cohort (n=365)		P	Internal validation cohort (n=157)		P	External validation cohort (n=108)		P
	AIS/MIA (n=195)	IAC (n=170)		AIS/MIA (n=85)	IAC (n=72)		AIS/MIA (n=70)	IAC (n=38)	
No	152(77.95)	65(38.24)		59(69.41)	29(40.28)		55(78.57)	23(60.53)	
Yes	43(22.05)	105(61.76)		26(30.59)	43(59.72)		15(21.43)	15(39.47)	
Shape [n (%)]			<0.001			<0.001			<0.001
Round	106(54.36)	29(17.06)		42(49.41)	11(15.28)		58(82.86)	13(34.21)	
Irregular	89(45.64)	141(82.94)		43(50.59)	61(84.72)		12(17.14)	25(65.79)	
Type [n (%)]			<0.001			<0.001			0.013
pGGN	135(69.23)	48(28.24)		53(62.35)	15(20.83)		16(22.86)	1(2.63)	
mGGN	60(30.77)	122(71.76)		32(37.65)	57(79.17)		54(77.14)	37(97.37)	

AIS, adenocarcinoma *in situ*; CEA, carcinoembryonic antigen; CT, computed tomography; HU, Hounsfield units; IAC, invasive adenocarcinoma; LLL, left lower lobe; LUL, left upper lobe; mGGN, mixed round-glass nodules; MIA, minimally invasive adenocarcinoma; NSE, neuron specific enolase; pGGN, pure ground-glass nodules; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.

3.3 Feature selection and model construction

From the intratumoral region, peritumoral 1mm region, peritumoral 2mm region, and each habitat subregion, 1834 features were extracted respectively, including first-order features, shape features, and texture features, resulting in a total of 5502 features extracted from the overall tumor habitat regions. Through LASSO screening, 15 features were retained from the intratumoral region, 11 features from the peritumoral 1mm region, 16 features from the peritumoral 2mm region, and 18 features from the habitat regions for corresponding model construction (Supplementary Figures S2-S4, Figure 4). In Clinic model construction, univariate logistic regression showed long diameter, short diameter, CT value, lobulation, spiculation, margin, vessel changes, pleural retraction, shape, and type as potential risk factors. Further multivariate logistic regression identified long diameter and CT value for subsequent modeling (Table 2). Based on AUC evaluation results in the internal validation set, the Intra model, Peri 1mm model, and Habitat model all adopted the Logistic Regression algorithm, the Peri 2mm model adopted the eXtreme Gradient Boosting algorithm, and the Clinic model adopted the Random Forest algorithm. Specific performance of each model under different algorithms is detailed in Supplementary Tables S1-S5 and Supplementary Figure S5. Finally, the Intra model, Habitat model, and Clinic model were integrated to construct the combined nomogram model (Figure 5).

3.4 Model performance and evaluation

ROC curve analysis showed that in the training cohort, the combined model exhibited optimal diagnostic performance with an AUC of 0.928 (95% CI: 0.901-0.956) and accuracy of 0.874, ranking first among all models in both AUC and accuracy. The habitat

model closely followed with an AUC of 0.924 (95% CI: 0.896-0.953) and accuracy of 0.871. The Intra model, Peri 1mm model, and Peri 2mm model showed similar diagnostic performance (AUC range: 0.868-0.879), all significantly lower than the Combined and Habitat models. The Clinic model achieved only an AUC of 0.807, indicating limited clinical diagnostic value. DeLong test further confirmed that the Combined and Habitat models significantly outperformed the Intra model, Peri 1mm model, Peri 2mm model, and Clinic model ($p < 0.05$). In the internal validation cohort, models performance generally decreased, but the Combined model (AUC: 0.871, 95% CI: 0.815-0.926) and Habitat model (AUC: 0.859, 95% CI: 0.799-0.919) continued to maintain leading advantages. In the external validation cohort, the Combined model demonstrated the most excellent generalization ability, with its AUC (0.897, 95% CI: 0.836-0.957) significantly superior to all other single models (Intra AUC: 0.756, Peri 1mm AUC: 0.747, Peri 2mm AUC: 0.730, Clinic AUC: 0.712, Habitat AUC: 0.840), with DeLong test showing p -values all less than 0.05. Meanwhile, the Habitat model also showed good robustness, being the second-best performing model in external validation. Details are shown in Table 3 and Figure 6.

DCA indicated that both the Combined and Habitat models demonstrated high net benefit across all three cohorts, particularly at lower threshold probabilities, suggesting their potential value in early diagnosis. In contrast, while other models showed some clinical benefit within specific threshold probability ranges, their overall performance was notably inferior to the Combined and Habitat models (Figure 6).

To evaluate model calibration performance, this study employed Hosmer-Lemeshow test and calibration curves for analysis. Hosmer-Lemeshow test results indicated that only the Combined and Habitat models consistently maintained good calibration ability across training, internal validation, and external validation sets ($p > 0.05$) (Table 4). Calibration curves further visualized the predictive accuracy of each model. In the calibration curves, both Habitat and Combined model curves

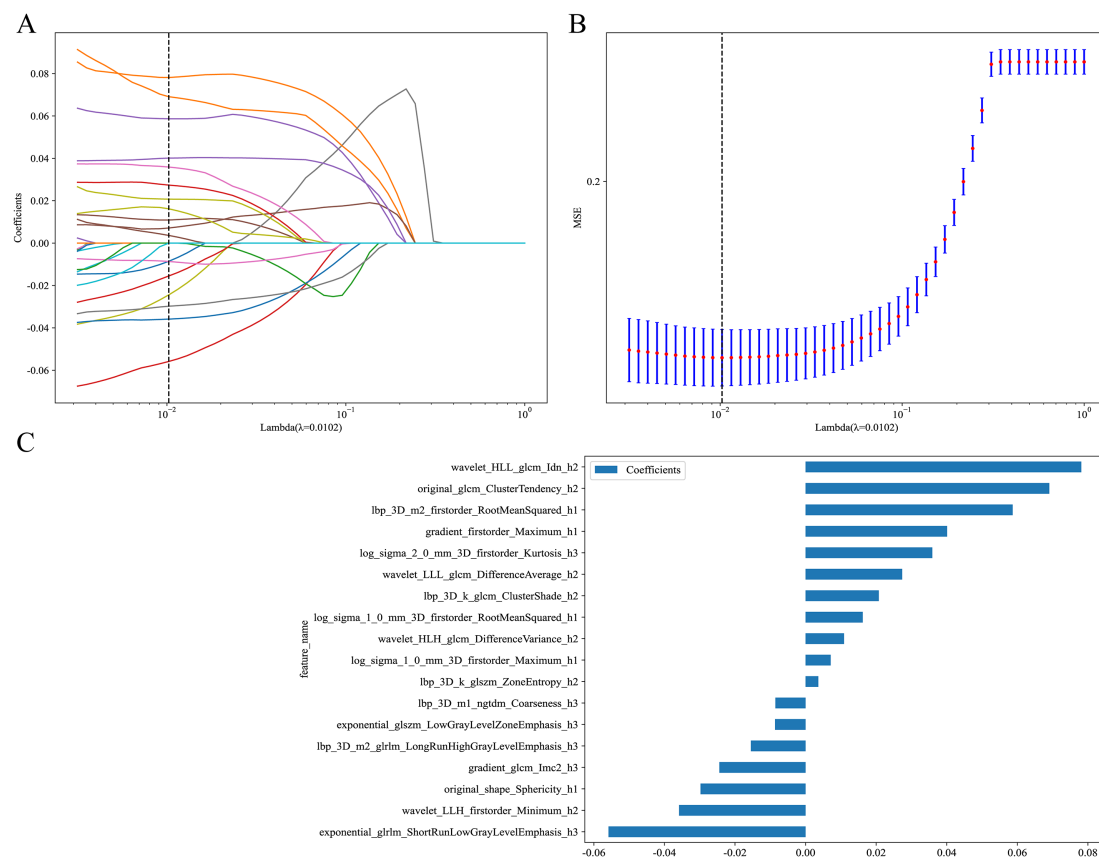


FIGURE 4

LASSO regression screening of radiomic features in the Habitat model. (A) LASSO coefficient path plot. This plot illustrates the trajectories of feature coefficients as the regularization parameter (Lambda) varies. As Lambda increases, the coefficients shrink toward zero, identifying key features at the optimal Lambda value (dashed line). (B) LASSO regression MSE curve plot. The dashed line marks the optimal Lambda value where MSE is minimized, determining the final feature subset. (C) LASSO-screened feature coefficient distribution plot. This shows the coefficients of features selected by LASSO regression. MSE, mean squared error.

remained close to the ideal calibration line (dashed line) across all three cohorts, indicating high concordance between predicted probabilities and actual observations (Figure 6).

4 Discussion

This study developed and validated a CT-based habitat radiomics model for differentiating AIS/MIA from IAC manifesting as GGNs. The results demonstrate that the habitat radiomics model possesses unique advantages in capturing intratumoral heterogeneity, with diagnostic performance significantly superior to traditional intratumoral and peritumoral radiomics models. The combined nomogram model integrating habitat features, intratumoral features, and clinical risk factors exhibited optimal diagnostic performance, achieving an AUC of 0.897 in the external validation set, providing a reliable quantitative tool for clinical precision diagnosis and treatment.

Accurate preoperative pathological grading of lung adenocarcinoma is crucial for developing individualized treatment strategies. Previous studies have shown that AIS/MIA patients can achieve a 5-year DFS rate of 100% after surgery, while IAC patients

have significantly poorer prognosis (7, 33). Therefore, accurate preoperative differentiation between these two lesion types is essential for avoiding overtreatment or undertreatment. This study found that CT value and long diameter were independent risk factors for IAC, consistent with previous research (4, 34). Lung adenocarcinoma generally progresses through four stages: AAH, AIS, MIA, and IAC. During this gradual evolution, increased nodule size often reflects enhanced tumor cell proliferative activity and invasiveness (18). Larger nodules are more likely to contain solid components, typically indicating invasive growth patterns (35). During invasive adenocarcinoma development, tumor cells grow along alveolar walls, initially maintaining alveolar structural integrity with only slight density increases (36). As invasion deepens, tumor cell density increases, fibrous tissue proliferates, and angiogenesis increases, leading to further CT value elevation (37). However, models based solely on clinical factors demonstrated insufficient diagnostic performance (external test set AUC: 0.712), indicating that traditional imaging features alone cannot meet the needs for clinical precision diagnosis.

In recent years, some studies have explored the application value of radiomics in differentiating lung adenocarcinoma pathological subtypes. Zheng et al. (38) constructed a model based on 11

TABLE 2 Univariate and multivariate logistic regression analysis of clinical and imaging variables.

Features	Univariate logistic regression	P	Multivariate logistic regression	P
	OR (95% CI)		OR (95% CI)	
Age	1.000 (0.997-1.003)	0.979		
CEA	1.015 (0.940-1.096)	0.747		
NSE	0.992 (0.982-1.003)	0.228		
Long diameter	1.028 (1.014-1.042)	0.001	1.237 (1.116-1.372)	0.001
Short diameter	1.030(1.011-1.049)	0.007	1.063 (0.932-1.213)	0.442
CT value	1.001 (1.001-1.001)	0	1.007 (1.006-1.009)	0
Gender	1.000 (0.811-1.232)	1.000		
Smoking	1.000 (0.595-1.682)	1.000		
Location	0.947 (0.882-1.016)	0.202		
Lobulation	1.282 (1.053-1.562)	0.038	0.785 (0.409-1.504)	0.540
Spiculation	2.282 (1.664-3.130)	0	1.104 (0.650-1.878)	0.759
Margin	1.418 (1.062-1.895)	0.047	0.760 (0.448-1.289)	0.392
Vessel changes	1.570 (1.262-1.952)	0.001	0.732 (0.433-1.239)	0.330
Bubble lucency	1.362 (0.993-1.868)	0.108		
Pleural retraction	2.442 (1.813-3.290)	0	1.671 (0.999-2.795)	0.101
Shape	1.584 (1.267-1.980)	0.001	1.737 (0.962-3.139)	0.124
Type	2.033 (1.568-2.635)	0	0.700 (0.418-1.172)	0.256

CEA, carcinoembryonic antigen; CI, confidence interval; CT, computed tomography; NSE, neuron specific enolase; OR, odds ratio.

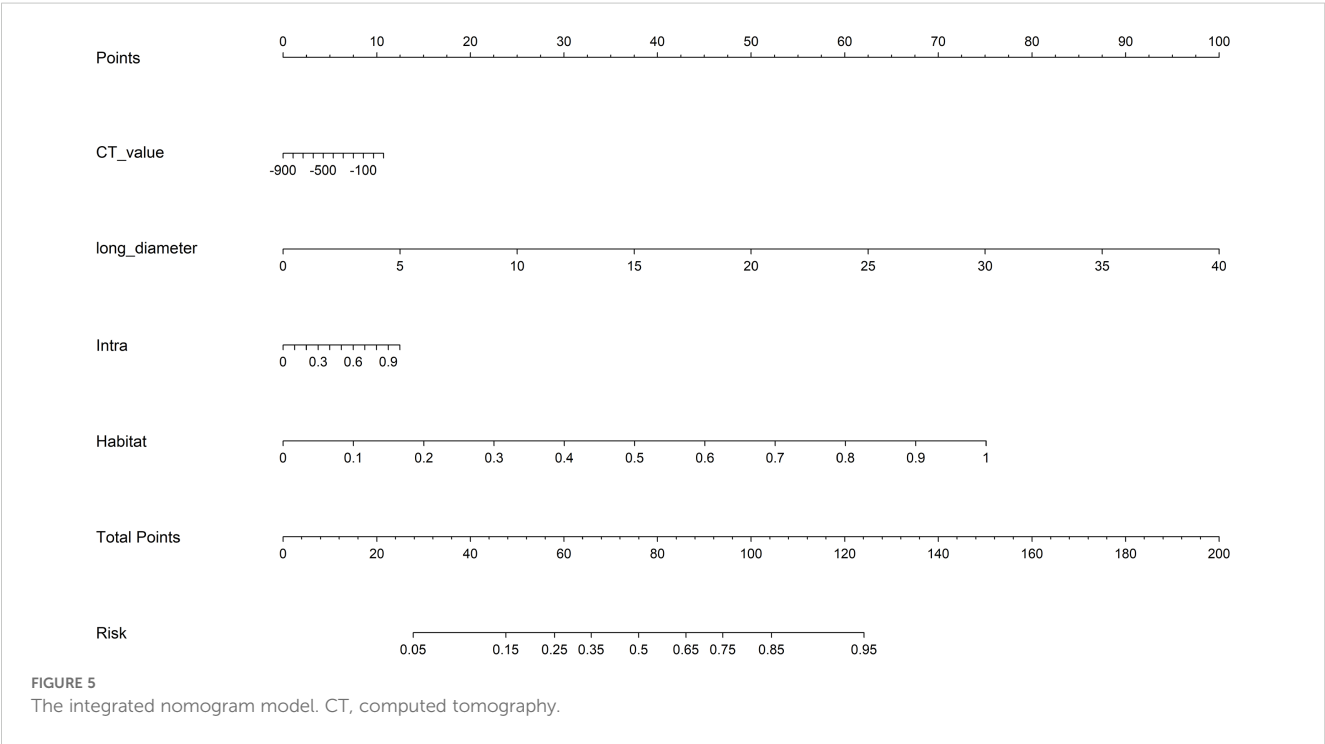


TABLE 3 Comparison of prediction performance of different models.

Model	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV
Training cohort						
Intra	0.879 (0.840 - 0.918)	0.827	0.800	0.851	0.824	0.830
Peri 1mm	0.874 (0.835 - 0.914)	0.827	0.712	0.928	0.896	0.787
Peri 2mm	0.868 (0.828 - 0.908)	0.822	0.741	0.892	0.857	0.798
Clinic	0.807 (0.758 - 0.856)	0.797	0.747	0.841	0.804	0.792
Habitat	0.924 (0.896 - 0.953)	0.871	0.794	0.938	0.918	0.839
Combined	0.928 (0.901 - 0.956)	0.874	0.771	0.964	0.949	0.828
Internal validation cohort						
Intra	0.855 (0.795 - 0.916)	0.809	0.639	0.953	0.920	0.757
Peri 1mm	0.850 (0.791 - 0.909)	0.758	0.917	0.624	0.673	0.898
Peri 2mm	0.823 (0.758 - 0.888)	0.752	0.764	0.741	0.714	0.787
Clinic	0.731 (0.650 - 0.812)	0.752	0.792	0.718	0.704	0.803
Habitat	0.859 (0.799 - 0.919)	0.803	0.778	0.824	0.789	0.814
Combined	0.871 (0.815 - 0.926)	0.809	0.667	0.929	0.889	0.767
External validation cohort						
Intra	0.756 (0.649 - 0.862)	0.750	0.842	0.700	0.604	0.891
Peri 1mm	0.747 (0.637 - 0.857)	0.769	0.737	0.786	0.651	0.846
Peri 2mm	0.730 (0.626 - 0.834)	0.750	0.553	0.857	0.677	0.779
Clinic	0.712 (0.602 - 0.822)	0.731	0.711	0.743	0.600	0.825
Habitat	0.840 (0.759 - 0.922)	0.806	0.789	0.814	0.698	0.877
Combined	0.897 (0.836 - 0.957)	0.843	0.921	0.800	0.714	0.949

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

radiomics features achieving an AUC of 0.820 in the training set. Meng et al. (4) selected 8 key features through LASSO regression to establish a Rad-score for differentiating AIS/MIA from IAC, achieving a training set AUC of 0.892. Our Intra model achieved a training set AUC of 0.879, comparable to previous studies but still inferior to the Habitat model’s predictive performance. This study systematically evaluated the diagnostic performance of Peri 1mm and Peri 2mm models (training set AUCs of 0.874 and 0.868, respectively), which, while superior to the Clinic model, were significantly inferior to the Habitat model. These results indicate that although the peritumoral region contains important information related to tumor invasiveness, both intratumoral and peritumoral models analyze these regions as wholes, failing to fully explore their internal heterogeneity information. Additionally, we found that Peri models performed worse than the Intra model, differing from some previous studies (39, 40). Possible reasons for this discrepancy include: first, this study included only GGNs, whose peritumoral microenvironmental changes may be less pronounced than solid nodules; second, the relatively small peritumoral extension distances (1-2mm) selected in this study may not have adequately captured key biological information in the peritumoral region, suggesting that future research should explore larger peritumoral ranges (such as 3-

5mm or even broader regions) to comprehensively assess the tumor microenvironment and potentially identify more valuable predictive features.

Traditional radiomics treats tumors as single homogeneous entities for feature extraction, ignoring intratumoral heterogeneity. However, increasing evidence indicates that tumors are highly heterogeneous ecosystems containing cell subpopulations with different phenotypic and functional characteristics (41). Spatial variations in the tumor microenvironment, including oxygen concentration gradients, nutrient distribution, and interstitial pressure differences, drive adaptive evolution of tumor cells, leading to coexistence of cell subpopulations with different proliferative capacities, invasiveness, and treatment sensitivities (19). This heterogeneity is particularly evident in ground-glass nodular lung adenocarcinoma: the tumor center may have already undergone invasion while peripheral regions maintain *in situ* or minimally invasive growth characteristics (42). This study employed K-means clustering based on 12 local features to divide GGNs into 3 habitat subregions, thereby more precisely capturing spatial heterogeneity information within tumors. Results showed that the Habitat model’s predictive performance significantly exceeded other single models (internal validation set AUC: 0.924

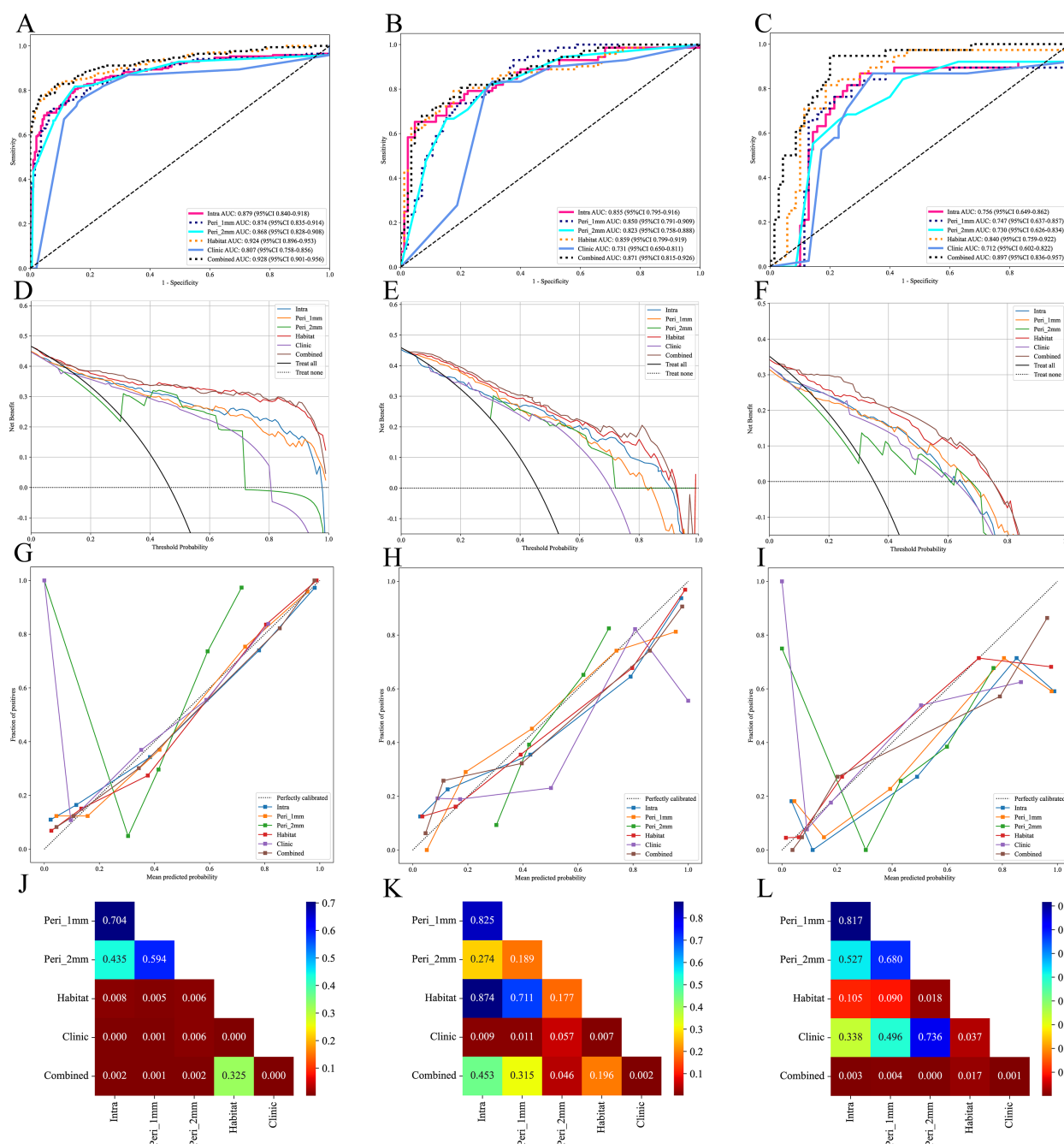


FIGURE 6

Performance evaluation of different models in the training, internal validation, and external validation cohorts. (A–C), ROC curves of different models in the training, internal validation and external validation cohorts; (D–F), DCA curves of different models in the training, internal validation and external validation cohorts; (G–I), calibration curves of different models in the training, internal validation and external validation cohorts; (J–L), Delong test of different models in the training, internal validation and external validation cohorts.

vs 0.879, 0.874, 0.868, 0.807, all $p < 0.05$), consistent with previous research conclusions. Wu et al. (43) reported that the Habitat model improved AUC by 6.5% compared to traditional radiomics models when predicting epidermal growth factor receptor mutation status in stage I non-small cell lung cancer. Bi et al. (44) reported similar results in predicting drug resistance in ovarian cancer patients. Among the 18 features selected through LASSO for the Habitat model, the top two features—wavelet_HLL_glcM_Idn_h2 and

original_glcM_ClusterTendency_h2—are both texture features derived from gray-level co-occurrence matrix (GLCM) analysis. The Inverse Difference Normalized (Idn) feature quantifies local homogeneity in image texture, with higher values indicating more homogeneous regions, which may reflect areas of uniform cellular density characteristic of less invasive tumor components (45). Cluster Tendency measures the grouping of pixels with similar gray-level values, capturing the spatial organization patterns that

TABLE 4 Hosmer-Lemeshow test results for different models in three cohorts.

Cohort	Intra	Peri 1mm	Peri 2mm	Clinic	Habitat	Combined
Training	0.080	1.977×10^{-2}	5.737×10^{-9}	4.682×10^{-7}	0.073	0.350
Internal Validation	0.309	1.272×10^{-1}	1.282×10^{-1}	3.421×10^{-1}	0.063	0.176
External Validation	0.022	1.869×10^{-4}	1.274×10^{-4}	2.775×10^{-4}	0.292	0.131

distinguish between the lepidic growth pattern of AIS/MIA and the more disorganized invasive growth pattern of IAC (46). The prominence of these texture features in our model aligns with pathological observations that IAC exhibits greater architectural complexity and cellular heterogeneity compared to AIS/MIA, manifesting as more heterogeneous texture patterns on CT imaging (47). Notably, the Habitat model demonstrated stable diagnostic performance across all cohorts, particularly in external validation, where performance decline (training set AUC 0.924 vs external validation set AUC 0.840) was smaller than other single models. This suggests that habitat features possess stronger generalizability and robustness. The reasons may include: first, habitat features reflect tumor microenvironmental heterogeneity, capturing complex spatial distribution patterns within tumors, with this heterogeneity information remaining relatively stable across different patient populations (48); second, habitat analysis segments tumors into subregions with similar phenotypic characteristics through clustering methods, and this quantification of spatial heterogeneity more accurately reflects tumor biological properties, providing better transferability between different centers (49).

The combined nomogram model integrating habitat features, intratumoral features, and clinical risk factors demonstrated optimal diagnostic performance across all cohorts. The advantages of this multi-dimensional information fusion strategy include: first, different types of features provide complementary diagnostic information, with habitat features capturing intratumoral spatial heterogeneity, intratumoral features reflecting overall attributes, and clinical features providing macroscopic morphological information (50); second, the nomogram format is intuitive and user-friendly, enabling clinicians to quickly assess individual patient IAC risk probability and provide quantitative evidence for clinical decision-making; third, the model maintained good performance in external validation (AUC = 0.897), demonstrating feasibility for cross-center application. DCA showed that the Combined model generated clinical net benefit across a wide range of threshold probabilities, particularly excelling in low threshold intervals. This holds significant importance for GGN management, as early identification of IAC can guide timely surgical intervention and prevent disease progression. Simultaneously, accurate identification of AIS/MIA can avoid overtreatment, reduce unnecessary lobectomies, and preserve more lung function.

This study has certain limitations. First, selection bias inherent to retrospective studies may affect result reliability. To further validate the clinical application value of the constructed models, future large-sample, multicenter prospective studies are necessary, employing rigorous study design and standardized data collection processes to improve evidence level and clinical credibility. Second, the sample size used for inter- and intra-observer consistency assessment (n=30)

was relatively limited, which may not fully capture the variability in feature extraction reproducibility across a broader range of cases. Future studies should employ larger sample sizes for consistency evaluation to enhance the robustness of reproducibility assessment. Third, this study analyzed only non-contrast CT images, failing to fully utilize rich information provided by multimodal functional imaging such as contrast-enhanced CT and Positron Emission Tomography-CT. These imaging techniques provide important information about tissue perfusion and metabolic activity, and integrating multimodal imaging data could significantly improve model diagnostic accuracy and clinical utility (51). Last, current habitat generation relies primarily on unsupervised clustering algorithms. While capable of identifying subregions with different imaging characteristics, it lacks direct validation against pathological gold standards. Although these habitat subregions theoretically may reflect different microenvironments within tumors, the accuracy of these correspondences requires validation through systematic imaging-pathology correlation studies to provide a more solid biological theoretical foundation for clinical translation of habitat radiomics technology.

This study successfully constructed a CT-based habitat radiomics diagnostic model achieving precise assessment of ground-glass nodular lung adenocarcinoma invasiveness. Habitat analysis significantly improved diagnostic accuracy by capturing intratumoral spatial heterogeneity information. The nomogram model combining clinical risk factors demonstrated excellent performance and clinical applicability in multicenter validation. This innovative approach provides a new tool for early precision diagnosis of lung adenocarcinoma, with potential to improve patient treatment decisions and clinical outcomes.

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 humans were approved by Ethics Committee of Yantaishan Hospital Yantaishan Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

ND: Writing – original draft, Conceptualization. YY: Writing – original draft, Methodology, Software. YL: Data curation, Writing – original draft. GL: Formal analysis, Writing – original draft. PW: Writing – original draft, Supervision. LL: Writing – original draft, Investigation. HZ: Validation, Writing – original draft. HS: Writing – original draft, Resources. XS: Writing – review & editing, Project administration.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This study has received funding by Yantai Science and Technology Innovation Development Plan (No. 2024YD018) and Medical and Health Technology Project of Shandong Province (No. 202309010633).

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.

References

- Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca: A Cancer J Clin.* (2024) 74:229–63. doi: 10.3322/caac.21834
- Luo G, Zhang Y, Rungay H, Morgan E, Langselius O, Vignat J, et al. Estimated worldwide variation and trends in incidence of lung cancer by histological subtype in 2022 and over time: a population-based study. *Lancet Respir Med.* (2025) 13:348–63. doi: 10.1016/S2213-2600(24)00428-4
- Qi H, Zuo Z, Lin S, Chen Y, Li H, Hu D, et al. Assessment of intratumor heterogeneity for preoperatively predicting the invasiveness of pulmonary adenocarcinomas manifesting as pure ground-glass nodules. *Quant Imaging Med Surg.* (2025) 15:272–86. doi: 10.21037/qims-24-734
- Meng F, Guo Y, Li M, Lu X, Wang S, Zhang L, et al. Radiomics nomogram: a noninvasive tool for preoperative evaluation of the invasiveness of pulmonary adenocarcinomas manifesting as ground-glass nodules. *Transl Oncol.* (2021) 14:100936. doi: 10.1016/j.tranon.2020.100936
- Nicholson AG, Tsao MS, Beasley MB, Borczuk AC, Brambilla E, Cooper WA, et al. The 2021 who classification of lung tumors: impact of advances since 2015. *J Thorac Oncol.* (2022) 17:362–87. doi: 10.1016/j.jtho.2021.11.003
- Yotsukura M, Asamura H, Motoi N, Kashima J, Yoshida Y, Nakagawa K, et al. Long-term prognosis of patients with resected adenocarcinoma *in situ* and minimally invasive adenocarcinoma of the lung. *J Thorac Oncol.* (2021) 16:1312–20. doi: 10.1016/j.jtho.2021.04.007
- Dziedzic R, Marjański T, Rzyman W. A narrative review of invasive diagnostics and treatment of early lung cancer. *Transl Lung Cancer Res.* (2021) 10:1110–23. doi: 10.21037/tlcr-20-728
- Fan L, Fang M, Li Z, Tu W, Wang S, Chen W, et al. Radiomics signature: a biomarker for the preoperative discrimination of lung invasive adenocarcinoma manifesting as a ground-glass nodule. *Eur Radiol.* (2019) 29:889–97. doi: 10.1007/s00330-018-5530-z
- Feng H, Shi G, Xu Q, Ren J, Wang L, Cai X. Radiomics-based analysis of ct imaging for the preoperative prediction of invasiveness in pure ground-glass nodule lung adenocarcinomas. *Insights Imaging.* (2023) 14:24. doi: 10.1186/s13244-022-01363-9

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2025.1660071/full#supplementary-material>

- Lee SM, Park CM, Goo JM, Lee H, Wi JY, Kang CH. Invasive pulmonary adenocarcinomas versus preinvasive lesions appearing as ground-glass nodules: differentiation by using ct features. *Radiology.* (2013) 268:265–73. doi: 10.1148/radiol.13120949
- Chae H, Park CM, Park SJ, Lee SM, Kim KG, Goo JM. Computerized texture analysis of persistent part-solid ground-glass nodules: differentiation of preinvasive lesions from invasive pulmonary adenocarcinomas. *Radiology.* (2014) 273:285–93. doi: 10.1148/radiol.14132187
- Li H, Sui Y, Tao Y, Cao J, Jiang X, Wang B, et al. Coupling habitat radiomic analysis with the diversification of the tumor ecosystem: illuminating new strategy in the assessment of postoperative recurrence of non-muscle invasive bladder cancer. *Acad Radiol.* (2025) 32:821–33. doi: 10.1016/j.acra.2024.09.036
- Bao D, Zhao Y, Li L, Lin M, Zhu Z, Yuan M, et al. A mri-based radiomics model predicting radiation-induced temporal lobe injury in nasopharyngeal carcinoma. *Eur Radiol.* (2022) 32:6910–21. doi: 10.1007/s00330-022-08853-w
- Shin J, Seo N, Baek S, Son N, Lim JS, Kim NK, et al. Mri radiomics model predicts pathologic complete response of rectal cancer following chemoradiotherapy. *Radiology.* (2022) 303:351–8. doi: 10.1148/radiol.211986
- Zhu M, Yang Z, Wang M, Zhao W, Zhu Q, Shi W, et al. A computerized tomography-based radiomic model for assessing the invasiveness of lung adenocarcinoma manifesting as ground-glass opacity nodules. *Respir Res.* (2022) 23:96. doi: 10.1186/s12931-022-02016-7
- Wu G, Woodruff HC, Shen J, Refaee T, Sanduleanu S, Abdalla I, et al. Diagnosis of invasive lung adenocarcinoma based on chest ct radiomic features of part-solid pulmonary nodules: a multicenter study. *Radiology.* (2020) 297:E282. doi: 10.1148/radiol.2020209019
- She Y, Zhang L, Zhu H, Dai C, Xie D, Xie H, et al. The predictive value of ct-based radiomics in differentiating indolent from invasive lung adenocarcinoma in patients with pulmonary nodules. *Eur Radiol.* (2018) 28:5121–8. doi: 10.1007/s00330-018-5509-9
- Wang Z, Li Z, Zhou K, Wang C, Jiang L, Zhang L, et al. Deciphering cell lineage specification of human lung adenocarcinoma with single-cell rna sequencing. *Nat Commun.* (2021) 12:6500. doi: 10.1038/s41467-021-26770-2

19. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature*. (2013) 501:346–54. doi: 10.1038/nature12626
20. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell*. (2012) 21:309–22. doi: 10.1016/j.ccr.2012.02.022
21. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036
22. Ye G, Wu G, Zhang C, Wang M, Liu H, Song E, et al. Ct-based quantification of intratumoral heterogeneity for predicting pathologic complete response to neoadjuvant immunochemotherapy in non-small cell lung cancer. *Front Immunol*. (2024) 15:1414954. doi: 10.3389/fimmu.2024.1414954
23. Huang H, Chen H, Zheng D, Chen C, Wang Y, Xu L, et al. Habitat-based radiomics analysis for evaluating immediate response in colorectal cancer lung metastases treated by radiofrequency ablation. *Cancer Imaging*. (2024) 24:44. doi: 10.1186/s40644-024-06092-w
24. Wu Y, Zhang W, Liang X, Zhang P, Zhang M, Jiang Y, et al. Habitat radiomics analysis for progression free survival and immune-related adverse reaction prediction in non-small cell lung cancer treated by immunotherapy. *J Transl Med*. (2025) 23:393. doi: 10.1186/s12967-024-06092-y
25. Zhang H, Ouyang Y, Zhang H, Zhang Y, Su R, Zhou B, et al. Sub-region based radiomics analysis for prediction of isocitrate dehydrogenase and telomerase reverse transcriptase promoter mutations in diffuse gliomas. *Clin Radiol*. (2024) 79:e682–91. doi: 10.1016/j.crad.2024.01.030
26. Wang S, Wang T, Guo S, Zhu S, Chen R, Zheng J, et al. Whole tumour- and subregion-based radiomics of contrast-enhanced mammography in differentiating her2 expression status of invasive breast cancers: a double-centre pilot study. *Br J Cancer*. (2024) 131:1613–22. doi: 10.1038/s41416-024-02871-9
27. Wang S, Liu X, Wu Y, Jiang C, Luo Y, Tang X, et al. Habitat-based radiomics enhances the ability to predict lymphovascular space invasion in cervical cancer: a multi-center study. *Front Oncol*. (2023) 13:1252074. doi: 10.3389/fonc.2023.1252074
28. Mo S, Luo H, Wang M, Li G, Kong Y, Tian H, et al. Machine learning radiomics based on intra and peri tumor pa/us images distinguish between luminal and non-luminal tumors in breast cancers. *Photoacoustics*. (2024) 40:100653. doi: 10.1016/j.pacs.2024.100653
29. Liu H, Wang M, Wang Q, Lu Y, Lu Y, Sheng Y, et al. Multiparametric mri-based intratumoral and peritumoral radiomics for predicting the pathological differentiation of hepatocellular carcinoma. *Insights Imaging*. (2024) 15:97. doi: 10.1186/s13244-024-01623-w
30. Wang X, Zhao X, Li Q, Xia W, Peng Z, Zhang R, et al. Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage t1 lung adenocarcinoma on ct? *Eur Radiol*. (2019) 29:6049–58. doi: 10.1007/s00330-019-06084-0
31. Liu Y, Li Z, Xiong H, Gao X, Wu J, Wu S. Understanding and enhancement of internal clustering validation measures. *IEEE Trans Cybern*. (2013) 43:982–94. doi: 10.1109/TSMCB.2012.2220543
32. Zwanenburg A, Vallières M, Abdallah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. (2020) 295:328–38. doi: 10.1148/radiol.2020191145
33. Zhang Y, Jheon S, Li H, Zhang H, Xie Y, Qian B, et al. Results of low-dose computed tomography as a regular health examination among chinese hospital employees. *J Thorac Cardiovasc Surgery*. (2020) 160:824–31. doi: 10.1016/j.jtcvs.2019.10.145
34. Heidinger BH, Anderson KR, Nemec U, Costa DB, Gangadharan SP, VanderLaan PA, et al. Lung adenocarcinoma manifesting as pure ground-glass nodules: correlating ct size, volume, density, and roundness with histopathologic invasion and size. *J Thorac Oncol*. (2017) 12:1288–98. doi: 10.1016/j.jtho.2017.05.017
35. Kakinuma R, Noguchi M, Ashizawa K, Kuriyama K, Maeshima AM, Koizumi N, et al. Natural history of pulmonary sub-solid nodules: a prospective multicenter study. *J Thorac Oncol*. (2016) 11:1012–28. doi: 10.1016/j.jtho.2016.04.006
36. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol*. (2011) 6:244–85. doi: 10.1097/JTO.0b013e318206a221
37. Suzuki K, Koike T, Asakawa T, Kusumoto M, Asamura H, Nagai K, et al. A prospective radiological study of thin-section computed tomography to predict pathological noninvasiveness in peripheral clinical ia lung cancer (Japan clinical oncology group 0201). *J Thorac Oncol*. (2011) 6:751–6. doi: 10.1097/JTO.0b013e31821038ab
38. Zheng H, Zhang H, Wang S, Xiao F, Liao M. Invasive prediction of ground glass nodule based on clinical characteristics and radiomics feature. *Front Genet*. (2021) 12:783391. doi: 10.3389/fgene.2021.783391
39. Liu K, Li K, Wu T, Liang M, Zhong Y, Yu X, et al. Improving the accuracy of prognosis for clinical stage i solid lung adenocarcinoma by radiomics models covering tumor per se and peritumoral changes on ct. *Eur Radiol*. (2022) 32:1065–77. doi: 10.1007/s00330-021-08194-0
40. Shang Y, Chen W, Li G, Huang Y, Wang Y, Kui X, et al. Computed tomography-derived intratumoral and peritumoral radiomics in predicting egfr mutation in lung adenocarcinoma. *La Radiologia Medica*. (2023) 128:1483–96. doi: 10.1007/s11547-023-01722-6
41. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell*. (2017) 168:613–28. doi: 10.1016/j.cell.2017.01.018
42. Deng Y, Xia L, Zhang J, Deng S, Wang M, Wei S, et al. Multicellular ecotypes shape progression of lung adenocarcinoma from ground-glass opacity toward advanced stages. *Cell Rep Med*. (2024) 5:101489. doi: 10.1016/j.xcrm.2024.101489
43. Wu J, Meng H, Zhou L, Wang M, Jin S, Ji H, et al. Habitat radiomics and deep learning fusion nomogram to predict egfr mutation status in stage i non-small cell lung cancer: a multicenter study. *Sci Rep*. (2024) 14:15877. doi: 10.1038/s41598-024-66751-1
44. Bi Q, Miao K, Xu N, Hu F, Yang J, Shi W, et al. Habitat radiomics based on mri for predicting platinum resistance in patients with high-grade serous ovarian carcinoma: a multicenter study. *Acad Radiol*. (2024) 31:2367–80. doi: 10.1016/j.acra.2023.11.038
45. Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. Ct texture analysis: definitions, applications, biologic correlates, and challenges. *Radiographics*. (2017) 37:1483–503. doi: 10.1148/rg.2017170056
46. Aerts HJWL, Velazquez ER, Leijenaar RTH, 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
47. Wu L, Gao C, Xiang P, Zheng S, Pang P, Xu M. Ct-imaging based analysis of invasive lung adenocarcinoma presenting as ground glass nodules using peri- and intra-nodular radiomic features. *Front Oncol*. (2020) 10:838. doi: 10.3389/fonc.2020.00838
48. 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
49. Li S, Dai Y, Chen J, Yan F, Yang Y. Mri-based habitat imaging in cancer treatment: current technology, applications, and challenges. *Cancer Imaging*. (2024) 24:107–17. doi: 10.1186/s40644-024-00758-9
50. 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
51. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (i or ii) non-small cell lung cancer. *Radiology*. (2016) 281:947–57. doi: 10.1148/radiol.2016152234