Predicting pathological staging of non-small cell lung cancer using a multi-task radiomics model integrating intratumoral and peritumoral features

RUONAN PAN[1*], XIAOQIAN LU[1*], XIN DONG[1], LIANG GUO[2], XIANG LI[1] and DIANBO CAO[1]

[1]Department of Radiology, The First Hospital of Jilin University, Changchun, Jilin 130021, P.R. China;
[2]Department of Pathology, The First Hospital of Jilin University, Changchun, Jilin 130021, P.R. China

Received January 6, 2025; Accepted June 19, 2025

DOI: 10.3892/ol.2025.15177

Abstract. Pathological staging is essential for guiding treatment decisions and determining prognosis in patients with non-small cell lung cancer (NSCLC). The present study aimed to establish a model using intratumoral and peritumoral features from computed tomography radiomics data and a multi-task learning algorithm, and evaluate its predictive performance for the pathological stage of NSCLC. Data from 198 eligible patients with NSCLC from The Cancer Imaging Archive database were retrospectively analyzed, which was used to develop four radiomics models to classify the pathological stages of NSCLC. These models combined the traditional random forest and multi-task random forest algorithms with the volumes of interest for the intratumoral region alone and with the intratumoral and peritumoral regions combined. Subsequently, the data from 90 patients from a real-world dataset were collected for use as an external test set. Diagnostic performance was evaluated using accuracy, precision, sensitivity, specificity, F1 scores and receiver operating characteristic curves. The results revealed that, in the internal test set, the area under the curve (AUC) values for model 1 (single-task model based on the intratumoral region), model 2 (single-task model combining intra- and peritumoral regions), model 3 (multi-task model based on the intratumoral region) and model 4 (multi-task model combining intra- and peritumoral regions) were 0.814, 0.900, 0.896 and 0.938, respectively. In the external test set, the AUC values were 0.821, 0.921, 0.858 and 0.939, separately. Moreover, the results of the DeLong test indicated that the AUC difference between model 1 and 4 was statistically significant ($P<0.05$). In conclusion, the multi-task radiomic model incorporating both intratumoral and peritumoral regions demonstrated favorable diagnostic efficacy in the predictive pathological staging of NSCLC.

Introduction

Lung cancer is one of the most prevalent types of malignant tumors and the leading cause of cancer-related mortality worldwide (1,2), with an estimated 2 million new cases and 1.76 million deaths per year (3). Non-small cell lung cancer (NSCLC) is the predominant type, accounting for ~85% of primary lung malignancies (3). It can be categorized into two major histological subtypes: Squamous cell carcinoma (SCC) and adenocarcinoma (ADC) (4,5).

The 5-year survival rate for patients with NSCLC is variable, as the stage of the disease markedly impacts treatment options and patients prognosis (6). Prognosis varies considerably depending on the tumor stage at diagnosis, with studies reporting that the 5-year survival rate for patients with stage IA-IIIA NSCLC ranges from 23-83% (7,8). In addition, treatment options differ according to the pathological staging, with surgical resection being the preferred approach for most patients with stage I and II NSCLC, whilst non-surgical treatments are more suitable for those patients with stage III and IV NSCLC (9,10). Therefore, accurately defining the pathological staging of NSCLC before initiating treatment is crucial for clinicians to implement optimal strategies related to actual stage and ultimately improve clinical outcomes. Currently, achieving accurate pathological staging for SCC and ADC exclusively using preoperative data remains a challenge for clinicians.

Imaging techniques such as computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET)/CT are widely used for tumor

Correspondence to: Professor Dianbo Cao, Department of Radiology, The First Hospital of Jilin University, 1 Xinmin Street, Changchun, Jilin 130021, P.R. China
E-mail: caodb@jlu.edu.cn

*Contributed equally

Abbreviations: NSCLC, non-small cell lung cancer; TCIA, The Cancer Imaging Archive; RF, random forest; MT-RF, multi-task RF; VOI, volume of interest; ROC, receiver operating characteristic; AUC, area under the curve; SCC, squamous cell carcinoma; ADC, adenocarcinoma; MRI, magnetic resonance imaging; PET/CT, positron emission tomography/computed tomography; ICC, inter- and intra-class correlation coefficient

Key words: radiomics, NSCLC, pathological staging, multi-task, intratumoral region, peritumoral region

screening (11). However, PET/CT is restricted by radiation exposure, high costs and the unavailability of equipment. MRI also has disadvantages concerning longer acquisition times and high costs (12,13). By contrast, CT, with its high resolution and short acquisition time, is utilized in clinical practice worldwide for diagnosis, staging, survival prediction and monitoring treatment response in patients with lung cancer (14-18).

Radiomics is an emerging technology that uses high-throughput quantitative imaging features for disease diagnosis and prognosis assessment. It systematically quantifies tumor phenotypes and genetic information by extracting and analyzing large volumes of image data (19). The radiomics models developed by Yu *et al* (20) were the first to suggest potential imaging biomarkers for predicting the pathological staging of NSCLC. The authors established radiomics models that accurately predicted the pathological stages of NSCLC (stage IA, IIA, IIB, IIIA, IIIB and IV). However, due to the limited sample size of each stage, the authors further analyzed binarized stages of early cancer (stage I-II) and advanced cancer (stage III-IV), which achieved an AUC of 0.82. Aside from the aforementioned study, the present study did not identify any other literature on machine learning for predicting the pathological staging of NSCLC in the PubMed database. However, the study by Yu *et al* (20) focused solely on features extracted from intratumoral tissue, neglecting the implications of peritumoral tissue. Certain studies have indicated that the peritumoral region also contains valuable information associated with predicting histological subtype, prognosis and lymph node metastasis (21,22). To date, the role of peritumoral regions in predicting the pathological staging of NSCLC has not been explored.

Most current research has focused on single-task models. Compared with single-task learning, multi-task learning models have demonstrated improved performance across ≥1 tasks, with enhanced generalization ability and robustness (23). Several studies have assessed the association between histological subtype and clinical staging. The multi-task deep learning model developed by Chen *et al* (4) and the multi-task radiomics model by Lin *et al* (24) have shown promising results in predicting histological subtypes and clinical staging, respectively. However, these studies exclusively included clinical staging as part of the analysis. Therefore, whether multi-task models can enhance the ability to predict the pathological staging of NSCLC remains a subject for further investigation.

The aim of the present study was to evaluate the performance of joint peritumoral features and multi-task learning in predicting the pathological staging of NSCLC. Single-task and multi-task learning were combined with the volumes of interest (VOIs) from the intratumoral region and the intratumoral and peritumoral regions combined, to develop four radiomics models for classifying the pathological staging of NSCLC.

Materials and methods

Patients. Patients were selected from two publicly available datasets (NSCLC-Radiogenomics and NSCLC-Radiomics-Genomics) on The Cancer Imaging Archive (TCIA) platform (25,26), supported by funding from the National Cancer Institute. Additionally, a real-world dataset comprising patients with NSCLC from the First Hospital of Jilin University (Changchun, China), collected between January 2021 and December 2022, was retrospectively analyzed. All patient data were anonymous. The inclusion criteria were as follows: i) Newly diagnosed and untreated NSCLC; ii) confirmed pathological staging (stage I-IV, according to the American Joint Committee on Cancer 8th edition, 2017) (27) and the histological subtype (SCC and ADC), as confirmed by postoperative pathology; and iii) available pre-treatment CT images. The exclusion criteria were as follows: i) Previous neoadjuvant therapy; ii) multiple primary tumors in the lungs; and iii) incomplete or poor-quality imaging data.

Data from all patients from the NSCLC-Radiogenomics and NSCLC-Radiomics-Genomics datasets were collected and the patients were randomly divided into two groups: 70% for the training set and 30% for the internal test set, including 151 cases of early stage and 47 cases of advanced stage NSCLC, and 69 cases of SCC and 129 cases of ADC. An additional 90 patients from the First Hospital of Jilin University were assigned to the external test set. The recruitment process for these patients is illustrated in Fig. 1.

The present study was approved by the Ethical Institution Review Committee of The First Hospital of Jilin University and was exempt from the requirement for informed consent. Informed consent was obtained by the TCIA database for the imaging and clinical data of the NSCLC-Radiogenomics and NSCLC-Radiomics-Genomics publicly accessible datasets.

Image acquisition. The NSCLC-Radiogenomics dataset, sourced from the Stanford University School of Medicine and the Palo Alto Veterans Affairs Healthcare System, as well as the NSCLC-Radiomics-Genomics dataset from the MAASTRO Clinic of the Netherlands, were publicly available in the TCIA database. The data for the external test set were obtained from the Picture Archiving and Communications System of the First Hospital of Jilin University. Based on the information provided by all databases, the scanning parameters were as follows: 80-140 kV; 124-699 mAs; matrix size, 512x512; slice intervals: 0.625-3 mm. Moreover, the entire lung of each patient was scanned.

Clinical information and data processing. Among eligible cases, sex, pathological staging and histological subtype classification information were collected. The tumor-node-metastasis (TNM) stages of all eligible patients were reassessed according to the 8th edition of the TNM staging (27) system by a pathologist with >5 years of experience. Based on pathological staging, patients were categorized into early stage and advanced stage groups.

Image segmentation. To eliminate differences in voxel size between different datasets, images were resampled to 1x1x1 mm^3 using RIASEG software (28,29) (https://github.com/lisherlock/RIASEG; version 1.0). In the present study, semi-automatic segmentation was employed to delineate the VOI. After the software automatically segmented the intratumoral region layer by layer, a junior radiologist (radiologist A; 3 years of experience) checked and adjusted the segmented

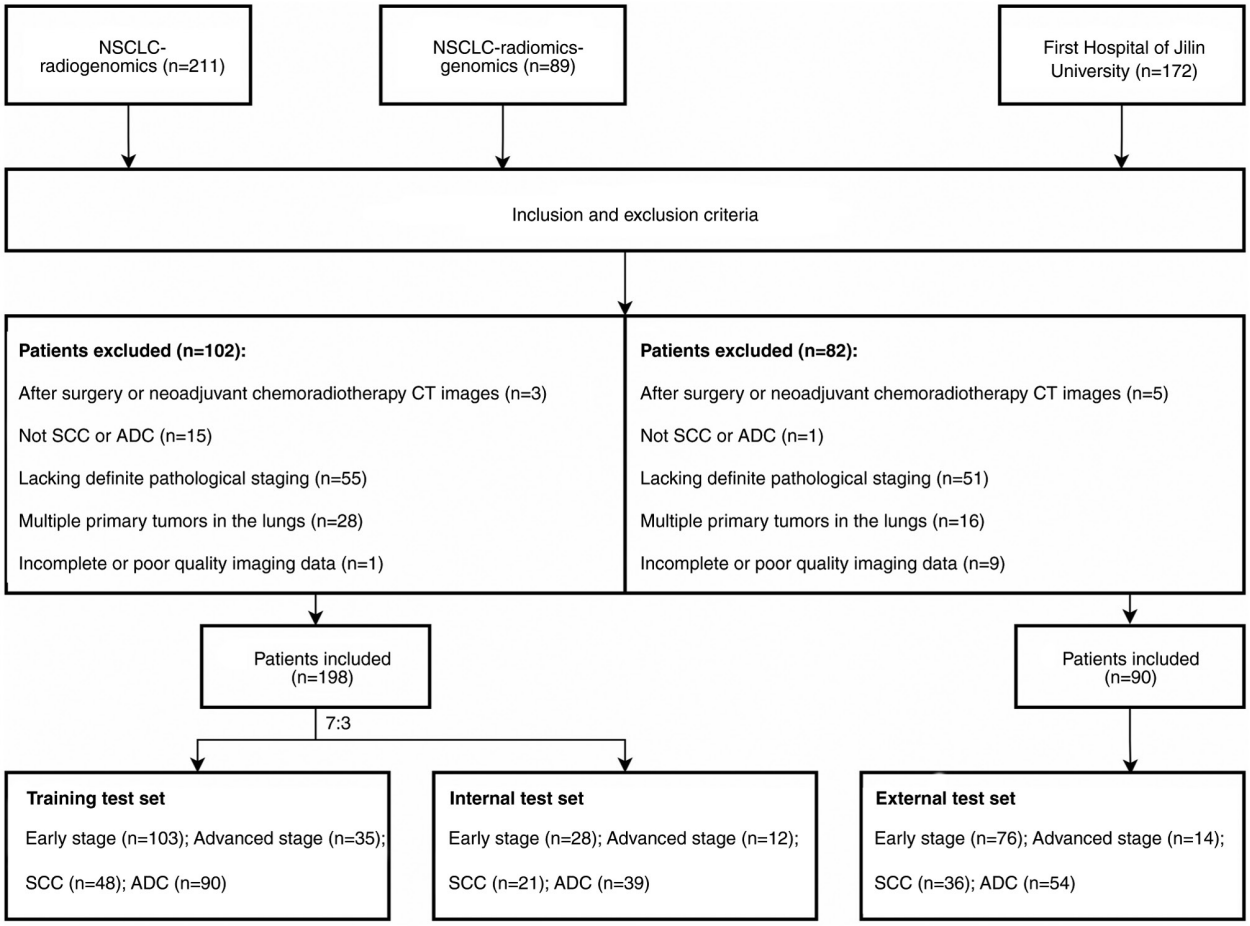

Figure 1. Flow chart for patient inclusion and exclusion. NSCLC, non-small cell lung cancer; CT, computed tomography; SCC, squamous cell carcinoma; ADC, adenocarcinoma.

areas. The final segmentation was evaluated again by a senior radiologist (radiologist B; 10 years of experience) for quality control. Following the segmentation of the intratumoral region VOI, the built-in extension function software of the was used to automatically expand the outlined VOI outward by 10 mm (21) to create a VOI that included the peritumoral region, thus generating the intratumoral and peritumoral regions VOI (Fig. 2). During the segmentation process, attention was paid to exclude large blood vessels, bronchi, soft tissues of the chest wall, ribs and mediastinum to ensure that the VOI reflected the actual lesion. Both the intratumoral and the combined intratumoral and peritumoral VOIs were mapped for each lesion.

Radiomics feature selection and radiomics model construction. The process of feature selection and model construction was divided into two steps: i) Evaluation of the reliability and reproducibility of radiomics features among observers; and ii) feature selection and model construction using two algorithms. Firstly, to assess intra-class reproducibility, 20 cases were randomly selected and segmented twice by radiologist A, with a 2-week interval. To evaluate inter-class reproducibility, the same 20 cases were segmented by radiologist C (5 years of experience) after segmentation by radiologist A. Inter- and intra-class correlation coefficients (ICCs) were calculated to evaluate the intra- and inter-observer agreement in feature extraction (30,31), with an ICC of >0.75 indicating good agreement. Thus, features with an ICC of >0.75 were retained for further feature extraction. Secondly, the retained features were normalized using standard deviation, followed by feature selection and model construction using the random forest (RF) or multi-task (MT)-RF algorithms, performed in Python 3.7.1 (https://www.python.org/downloads/release/python-371/). RF is a type of embedded feature selector which can automatically produce the relative importance of features during the model training process. MT-RF is an extension of the traditional RF algorithm that accounts for correlations and common features among related tasks, which means to identify and select features that are important for all tasks (4,23). Namely, it simultaneously considers relationships between multiple tasks and leverages feature correlations to enhance the performance and predictive accuracy of the model.

The multi-task learning model proposed in the present study consisted of two components: i) A pathological staging task branch, which classified cases into early and advanced stages; and ii) a histologic subtype task branch, which classified cases as ADC and SCC. However, the single-task model, which used the traditional RF algorithm, focused solely on pathological staging.

Statistical analysis and comparative methods. Accuracy, precision, sensitivity, specificity, F1 score and AUC were utilized as the evaluation metrics for classification results. The

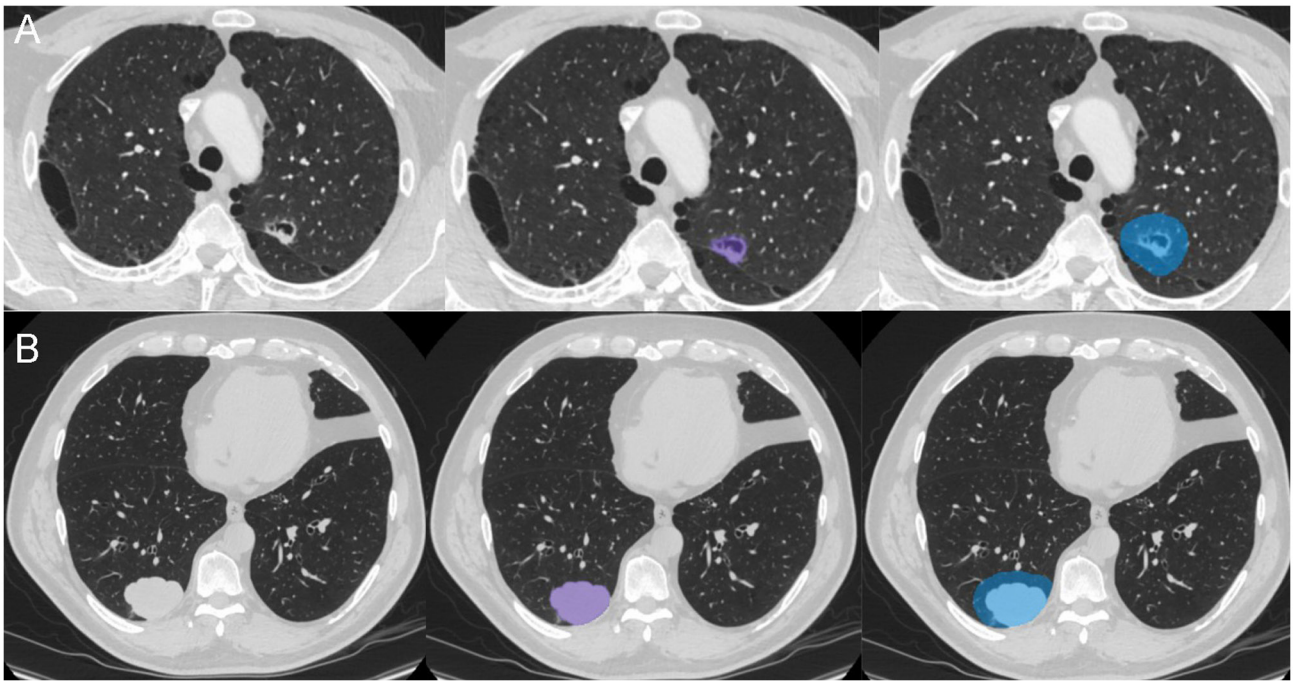

Figure 2. Image segmentation. Representative images from (A) Patient 1 and (B) patient 2 in the present study. Purple represents the intratumoral region, and blue represent the intratumoral and peritumoral regions.

AUC comprehensively reflects the sensitivity and specificity of the model, which is suitable for class imbalance problems. As the DeLong test is applicable to the non-parametric test of the paired receiver operating characteristic (ROC) curve, it can effectively evaluate the statistical significance of model performance and avoid the influence of class imbalance on indicators. Therefore, the present study used the DeLong test to compare the AUC values of different models. Demographic characteristics in the training, internal test and external test sets were compared using the χ^2 test. Statistical analyses of the demographic data were performed using SPSS software (version 26.0; IBM, Corp.). $P<0.05$ was considered to indicate a statistically significant difference and all reported statistical significance levels were two-sided.

Model comparisons were performed in two aspects: i) Comparison of different VOI models; and ii) comparison between the multi-task and single-task models. First, models that integrated features from the intratumoral and peritumoral regions were compared with those based solely on features from the intratumoral region. This comparison aimed to determine whether the inclusion of peritumoral features could enhance the classification performance of the model. Second, to assess whether the multi-task model, constructed by adding a histological subtype classification branch, could improve its performance in predicting the pathological stage of NSCLC, the multi-task model was compared with the single-task model.

Results

Patient characteristics. A total of 198 patients from NSCLC-Radiogenomics and NSCLC-Radiomics-Genomics datasets were included in the present study and subsequently assigned to either the training set (n=138) or the internal test set (n=60). Additionally, 90 patients from the First Hospital of Jilin University were assigned as an external test set (age range, 42-76 years). The demographic and tumor characteristics of all patients were presented in Table I. Statistical analysis revealed no significant differences in these characteristics among the training, internal test and external test sets.

Radiomic feature selection and model construction. Radiomic features were extracted from the delineated VOIs. Using an ICC threshold of >0.75 as a reliability standard, a total of 1,259 stable radiomic features were retained from the intratumoral VOI, whilst 1,331 stable features were retained from the combined intratumoral and peritumoral VOI. Subsequently, the MT-RF and RF algorithms were employed for feature selection and model construction. The MT-RF algorithm generated a feature importance score for each task, which was then aggregated to provide a comprehensive assessment of feature importance across all tasks. Similarly, the RF algorithm generated a feature importance score to all retained features. The features were ranked from highest to lowest importance. Figs. 3 and 4 present histograms of the top 20 most important radiomic features used in model development by the RF and MT-RF algorithms, respectively.

Performance evaluation of models

Single-task radiomics model. Table II and Fig. 5 demonstrate the performance of the single-task models on both the internal and external test sets in the present study. The results revealed that model 1 achieved an AUC of 0.814 and 0.821 for the internal and external test sets, whilst model 2 achieved an AUC of 0.900 and 0.921 for the internal and external test sets, respectively.

Multi-task radiomics model. Table III and Fig. 6 demonstrate the performance of the multi-task models on both the internal and external test sets in the present study. The results

Table I. Patient characteristics.

Characteristic	Training set (n=138)	Internal test set (n=60)	External test set (n=90)	P-value
Sex				0.211
Male	93 (67.4)	47 (78.3)	59 (65.6)	
Female	45 (32.6)	13 (21.7)	31 (34.4)	
Pathological staging				0.202
Early cancer (stage I-II)	103 (74.6)	48 (80.0)	76 (84.4)	
Advanced cancer (stage III-IV)	35 (25.4)	12 (20.0)	14 (15.6)	
Histologic subtype				0.296
SCC	44 (31.9)	25 (41.7)	36 (40.0)	
ADC	94 (68.1)	35 (58.3)	54 (60.0)	

Data are presented as n (%). Differences were assessed using the χ^2 test. SCC, squamous cell carcinoma; ADC, adenocarcinoma.

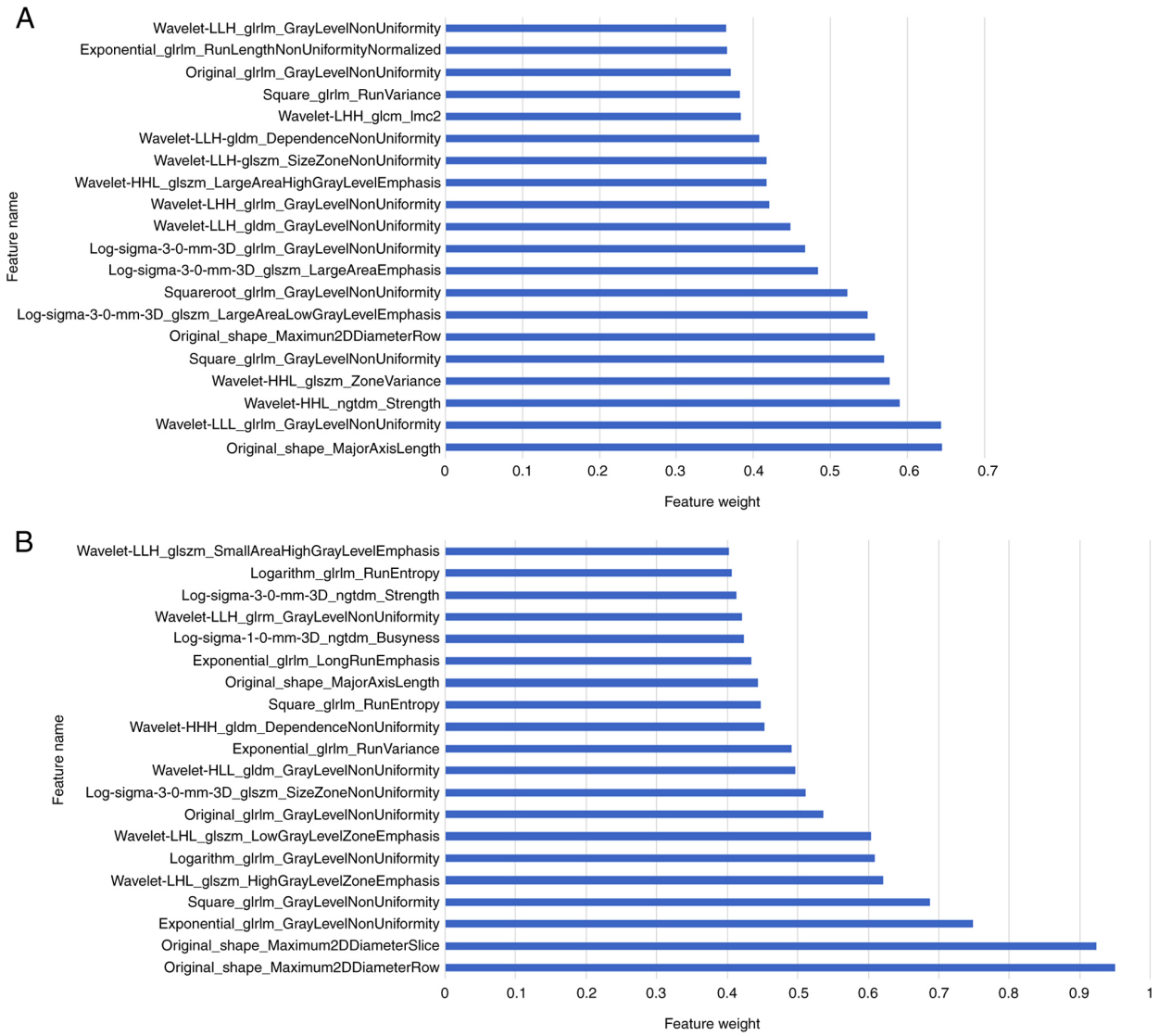

Figure 3. Weight histograms of the top 20 most important radiomic features identified using the random forest algorithms. (A) Intratumoral region. (B) Intratumoral and peritumoral regions

revealed that the AUC for the model 3 was 0.896 and 0.858 for the internal test set and external test set, respectively. By contrast, model 4 achieved AUC values of 0.938 and 0.939 for the internal and external test sets, respectively.

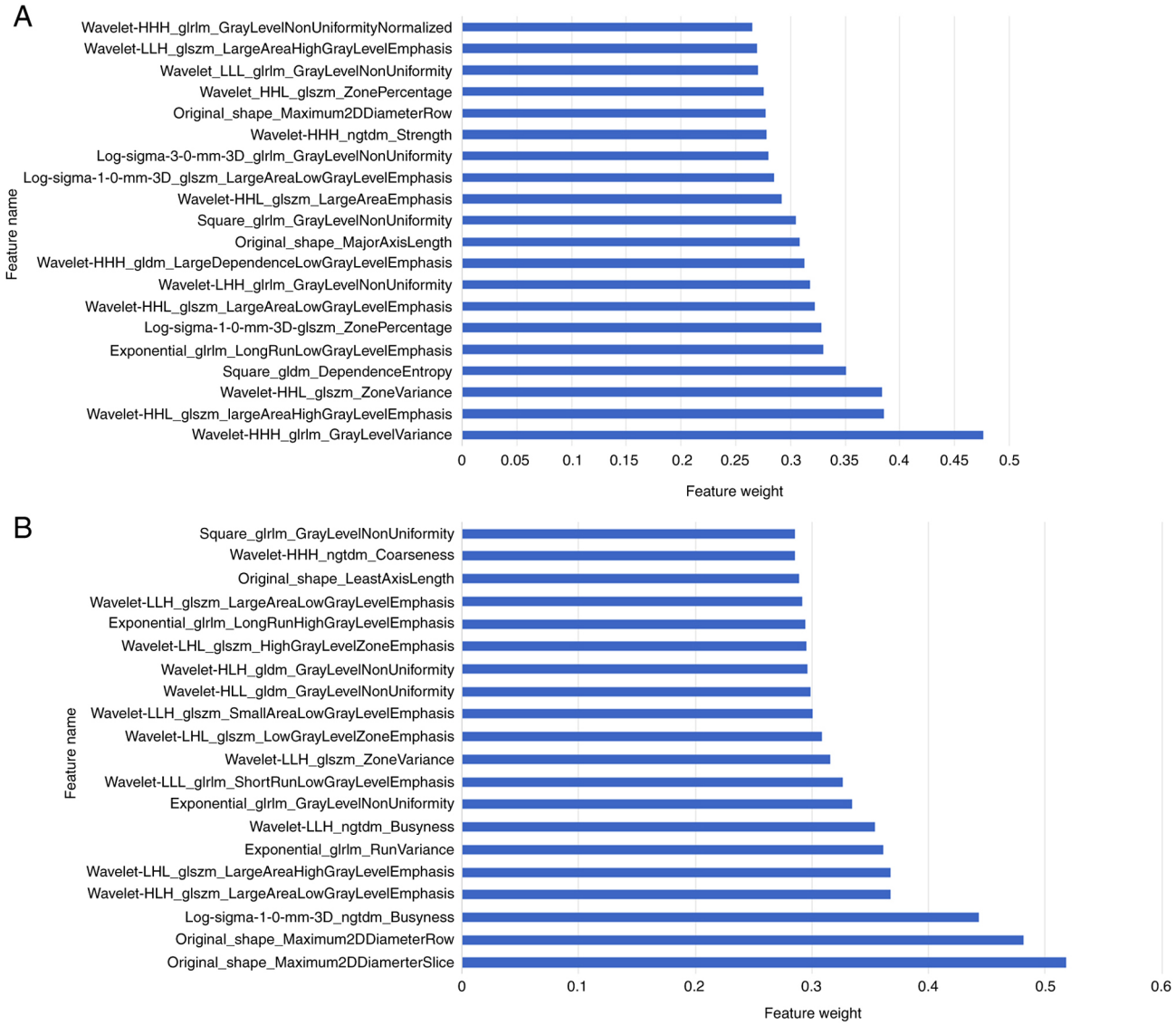

Figure 4. Weight histograms of the top 20 most important radiomic features identified using the multi-task-random forest algorithms. (A) Intratumoral region. (B) Intratumoral and peritumoral regions.

Comparison of radiomics models. Fig. 7 presents the ROC curves for all models. The results indicated that model 4 achieved the highest AUC (internal test set, 0.938; external test set, 0.939). Additionally, models incorporating both intratumoral and peritumoral regions (models 2 and 4) had markedly higher AUC values than the corresponding models based solely on the intratumoral region (models 1 and 3). The multi-task models (models 3 and 4) also exhibited notably higher AUC values compared with the corresponding single-task models (models 1 and 2).

The results of the DeLong test for the four models are presented in Table IV. The findings indicated that the AUC of models 1 and 4 were significantly different in both the internal test set (P=0.049) and the external test set (P=0.046).

Additionally, to assess the diagnostic performance of model 4 for peripheral and central types, tumor (T) stages of ≤T2 and ≥T3, and its applicability to cases with atelectasis, the present study performed subgroup analyses. The results are presented in Table SI. The findings demonstrated that the model in the present study performed well for both central and peripheral lung cancers, suggesting that tumor location does not markedly affect model performance. However, the model exhibited a slightly improved predictive performance for T2- cases compared with T3+ cases. This discrepancy may stem from the smaller sample size in the T3+ subgroup, which limited the sensitivity of the model to minority groups. Future research should incorporate more complex cases to enhance the generalization capability of the model.

Discussion

The TNM staging of tumors is crucial for evaluating patients, as it implies the extent of the disease and notably impacts therapeutic decisions and prognosis. Clinicians must assess the clinical staging of patients with NSCLC before surgery; however, discrepancies between clinical and pathological staging are commonly encountered (9). In practice, treatment plans for patients with resectable lung cancer are primarily based on the final pathological stage rather than the clinical stage (32). Therefore, accurately identifying the pathological

Table II. Performance of the single-task radiomics models in internal and external test sets.

A, Internal test set

Model	AUC	Accuracy	Precision	Sensitivity	Specificity	F1 score
1	0.814	0.883	0.692	0.750	0.916	0.720
2	0.900	0.866	0.666	0.666	0.916	0.666

B, External test set

Model	AUC	Accuracy	Precision	Sensitivity	Specificity	F1 score
1	0.821	0.866	0.562	0.642	0.807	0.600
2	0.921	0.800	0.423	0.785	0.802	0.550

AUC, area under the curve.

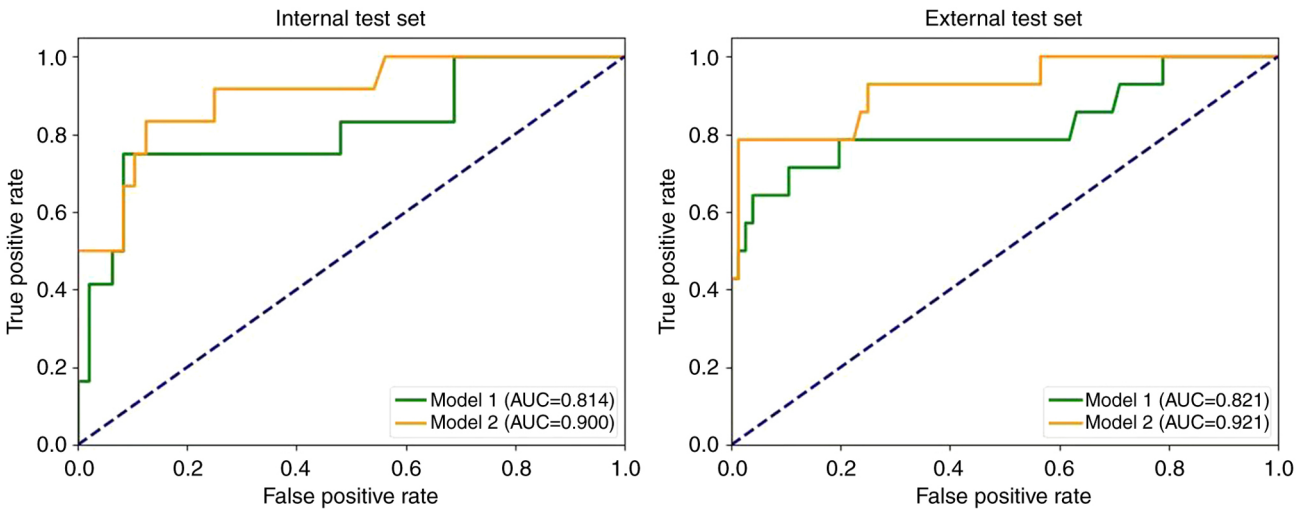

Figure 5. Receiver operating characteristic curves of the single-task radiomics models in the internal and external test sets. AUC, area under the curve.

stage of NSCLC before treatment can aid clinicians in tailoring personalized therapies and improving clinical outcomes.

The present study developed four radiomics models to predict the pathological staging of NSCLC: Model 1, a single-task model based on the intratumoral region; model 2, a single-task model combining intra- and peritumoral regions; model 3, a multi-task model based on the intratumoral region; and model 4, a multi-task model combining intra- and peritumoral regions. The results indicated that model 4 demonstrated the best performance for pathological staging, achieving an AUC of 0.938 on the internal test set and 0.939 on the external test set. Shape features and textural features served critical roles in model 4. Shape features are easy to interpret, with original_shape_Maximum2DDiameterSlice and original_shape_Maximum2DDiameterRow quantifying the maximum 2D diameters of tumors in axial planes and arbitrary slice orientations, respectively. These metrics directly reflect local tumor extension. As tumor size is a key determinant of T-staging, these features demonstrate strong associations with tumor dimensions, thereby markedly contributing to distinguishing early from advanced stage cases. Texture features were used to describe the heterogeneity within the tumor volume. Studies have reported that textural features analysis may identify adverse biological characteristics of lung tumors, including aggressive phenotypes, malignancy grading, resistance to radiotherapy/chemotherapy and prognostic outcomes (33-36). The most prominent texture feature in model 4, log-sigma-1-0-mm-3D_ngtdm_Busyness, was calculated from a neighborhood gray-tone difference matrix after edge enhancement via 3D Gaussian filtering (σ=1.0 mm). This metric reflects the frequency of local texture variations. High Busyness values indicate complex texture patterns, which may be associated with high tumor malignancy and invasive potential (37,38). Therefore, it may guide clinicians in selecting optimal treatment strategies and predicting patient outcomes.

Several retrospective studies have reported the notable potential of radiomics in predicting NSCLC staging. For instance, Yu *et al* (20) identified potential imaging markers to predict NSCLC pathological stages, though this study focused

Table III. Performance of the multi-task radiomics models in internal and external test sets.

A, Internal test set

Model	AUC	Accuracy	Precision	Sensitivity	Specificity	F1 score
3	0.896	0.866	0.666	0.666	0.916	0.666
4	0.938	0.866	0.666	0.666	0.916	0.666

B, External test set

Model	AUC	Accuracy	Precision	Sensitivity	Specificity	F1 score
3	0.858	0.877	0.600	0.642	0.921	0.620
4	0.939	0.900	0.727	0.571	0.960	0.640

AUC, area under the curve.

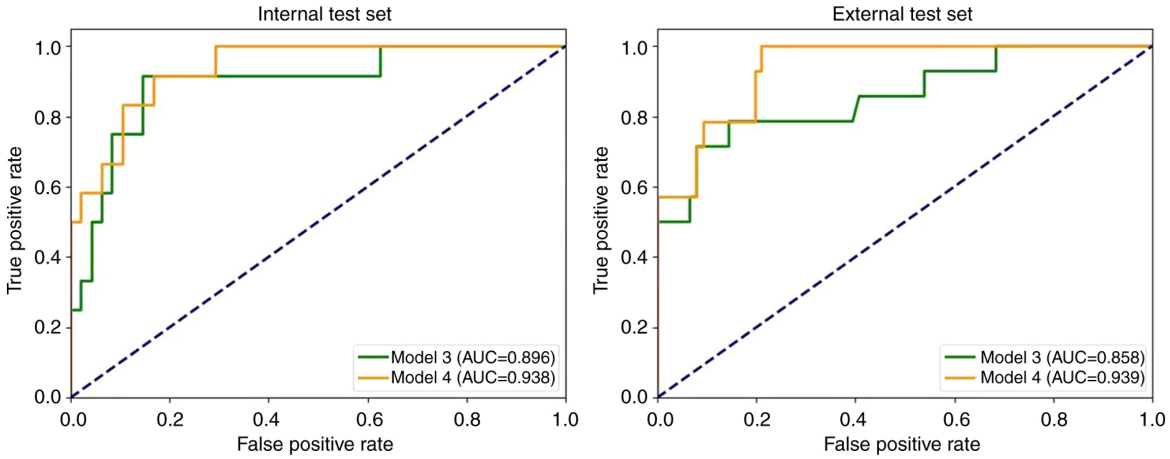

Figure 6. Receiver operating characteristic curves of the multi-task radiomics models in the internal and external test sets.

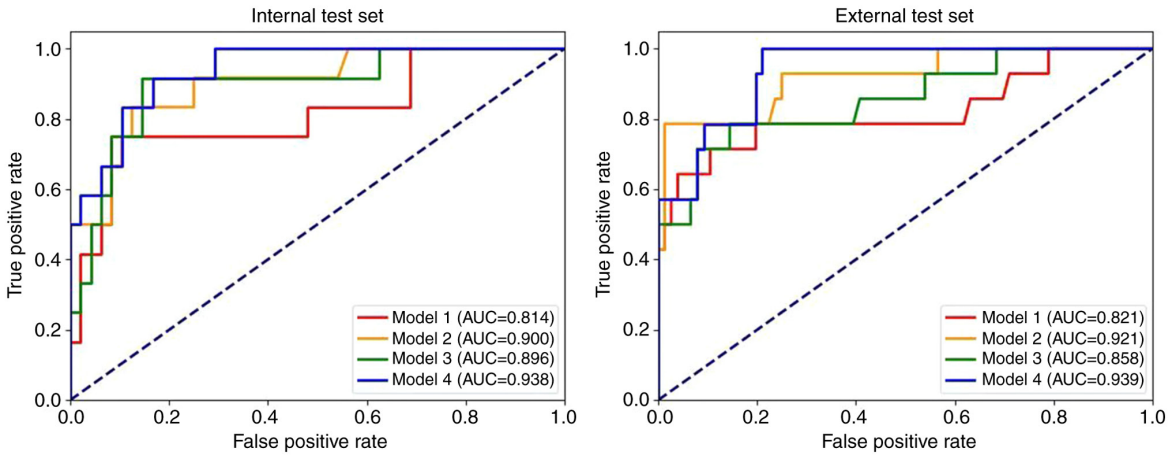

Figure 7. Receiver operating characteristic curves of all radiomics models in the internal and external test sets.

exclusively on intratumoral features and did not involve peritumoral characteristics. To further assess the role of peritumoral features in the pathological staging of NSCLC, the present study developed radiomics models that incorporated the intratumoral and peritumoral regions and compared them with models based solely on the intratumoral features. Using both multi-task and single-task models, it was demonstrated that radiomic models with the intratumoral and peritumoral

Table IV. DeLong test for the comparison of the area under the curve between the four models.

A, Internal test set

Model	AUC (95% CI)	Z	P-value
1	0.814 (0.660-0.968)	1.966	0.049
2	0.900 (0.802-0.998)	0.944	0.345
3	0.896 (0.791-1.000)	0.975	0.329
4	0.938 (0.875-1.000)	Standard	-

B, External test set

Model	AUC (95% CI)	Z	P-value
1	0.821 (0.670-0.973	1.999	0.046
2	0.921 (0.834-1.000)	0.475	0.635
3	0.858 (0.738-0.978)	1.912	0.056
4	0.939 (0.886-0.991)	Standard	-

AUC, area under the curve; CI, confidence intervwal.

regions combined achieved higher AUCs (0.938 and 0.900 in the internal test set, respectively) than models relying solely on intratumoral regions. The external test set also yielded similar results, supporting the hypothesis that integrating peritumoral features may enhance the predictive accuracy for the pathological stage of NSCLC.

Currently, most studies have focused on single-task models; however, certain research has demonstrated that multi-task models could extract more information by integrating multiple related tasks, thereby enhancing the learning outcomes for ≥1 tasks. These multi-task approaches exhibited greater generalization ability and robustness compared with single-task learning (4,23). Chen *et al* (4) developed a multi-task deep learning model comprising a histological subtype branch and clinical stage branch, elaborating the potential relationship between these factors. The study reported that the incorporation of staging tasks effectively improved the classification accuracy of histological subtypes, suggesting an association between these two elements. Similarly, Lin *et al* (24) constructed a multi-task radiomics model differentiating between histological subtypes and clinical stages in patients with NSCLC. The study reported that this multi-task model performed well in predicting clinical stages, but the lack of external validation raised concerns about the generalizability of the model. The results of the aforementioned studies underscore the association between histological subtype and clinical staging tasks, indicating that predictive performance may be enhanced through a multi-task modeling approach. However, these investigations emphasized clinical staging rather than pathological staging.

To assess the potential association between the pathological stage and histological subtype of NSCLC, and to assess whether predictive performance could be improved through a multi-task modeling approach, the present study incorporated the histological subtype into the framework for pathological staging of NSCLC. Using preoperative CT images from 198 patients, multi-task radiomics models were constructed based on the MT-RF algorithm. To evaluate the generalizability of models, an external validation was performed using data from the First Hospital of Jilin University. Furthermore, to assess the efficacy of the multi-task approach in classification, the multi-task models were compared with the single-task radiomics models built via identical VOIs. The results indicated that the AUCs of the two multi-task radiomics models (models 3 and 4) were superior to those of the single-task models (models 1 and 2). These findings suggest that integrating the histological subtype into a multi-task framework may enhance the diagnostic performance of pathological stage prediction in NSCLC.

Meanwhile, the results of the DeLong test revealed that the multi-task model, which combined the intratumoral and peritumoral features, markedly outperformed the single-task model that established solely on intratumoral features. This multi-task model achieved superior performance compared with the AUC of 0.82 reported by Yu *et al* (20). This finding suggests that the radiomics model incorporating the intratumoral and peritumoral regions, as well as multi-task learning, enhanced the diagnostic performance of pathological stage classification in NSCLC.

However, the present study has certain limitations. Firstly, the sample size was relatively small, and the data was unbalanced with regard to staging. To address this, external validation was performed to demonstrate the generalization of the model. Whilst the external test set demonstrated a good generalization performance of the model (AUC=0.939), the data were from a single source and may be limited by specific scanning protocols or population characteristics. To further validate the generality of the model, future studies should perform multi-center studies in collaboration with multiple healthcare institutions to incorporate data from different centers, devices and populations to fully evaluate the stability of the model in different clinical settings. Secondly, the lack of relevant clinical variants, such as genetic mutations and molecular biomarkers in the TCIA database, limited the current analysis. Future studies should prospectively collect multimodal data (including genomic and pathological biomarkers) to develop an integrated predictive model containing imaging-clinical-molecular features. The current research focuses on establishing a foundational framework for radiomics, which provides the cornerstone for subsequent exploration. Additionally, the present study utilized the 8th edition of TNM staging system, which is widely used in clinical practice, rather than the latest 9th edition. This choice was made as the 9th edition had not yet been broadly utilized. Therefore, further studies are necessary to advance this research.

In conclusion, the inclusion of peritumoral features significantly enhanced the AUC of predictive models for pathological staging in NSCLC. Additionally, the implementation of multi-task learning further contributed to this improvement. Finally, the multi-task radiomics model incorporating the intratumoral and peritumoral regions demonstrated good diagnostic performance in the pathological staging of NSCLC.

Acknowledgements

Not applicable.

Funding

The study was supported by grants from the Finance Department of Jilin Province (grant no. JLSWSRCZX2021-054), the Jilin Provincial Department of Science and Technology (grant no. 20220203128SF), the Jilin Provincial Development and Reform Commission (grant no. JLSDRC2023C039-2), the Jilin Provincial Key Laboratory of Medical Imaging & Big Data (grant no. 20200601003JC) and the Radiology and Technology Innovation Center of Jilin Province (grant no. YDZJ202402029CXJD).

Availability of data and materials

The data generated in the present study may be requested from the corresponding author.

Authors' contributions

RNP and XQL collected, analyzed and interpreted the patient data and were major contributors in writing the manuscript. XD and XL helped with data collection and mapping. LG assessed the pathological data and checked pathological staging. DBC conceived and designed the research, and modified the manuscript. All authors read and approved the final manuscript. RNP and XQL confirm the authenticity of all the raw data.

Ethics approval and consent to participate

The present study was approved by the Ethical Institution Review Committee of The First Hospital of Jilin University (approval no. AF-IRB-032-06) and was exempt from the requirement for informed consent.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Takenaka M, Hanagiri T, Shinohara S, Kuwata T, Chikaishi Y, Oka S, Shigematsu Y, Nagata Y, Shimokawa H, Nakagawa M, *et al*: The prognostic significance of HER2 overexpression in non-small cell lung cancer. Anticancer Res 31: 4631-4636, 2011.
2. Zhu X, Dong D, Chen Z, Fang M, Zhang L, Song J, Yu D, Zang Y, Liu Z, Shi J and Tian J: Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. Eur Radiol 28: 2772-2778, 2018.
3. Thai AA, Solomon BJ, Sequist LV, Gainor JF and Heist RS: Lung cancer. Lancet 398: 535-554, 2021.
4. Chen K, Wang M and Song Z: Multi-task learning-based histologic subtype classification of non-small cell lung cancer. Radiol Med 128: 537-543, 2023.
5. Bashir U, Kawa B, Siddique M, Mak SM, Nair A, McLean E, Bille A, Goh V and Cook G: Non-invasive classification of non-small cell lung cancer: A comparison between random forest models utilising radiomic and semantic features. Br J Radiol 92: 20190159, 2019.
6. Kutob L and Schneider F: Lung cancer staging. Surg Pathol Clin 13: 57-71, 2020.
7. Valente IR, Cortez PC, Neto EC, Soares JM, de Albuquerque VH and Tavares JM: Automatic 3D pulmonary nodule detection in CT images: A survey. Comput Methods Programs Biomed 124: 91-107, 2016.
8. León-Atance P, Moreno-Mata N, González-Aragoneses F, Cañizares-Carretero M, García-Jiménez MD, Genovés-Crespo M, Honguero-Martínez AF, Rombolá CA, Simón-Adiego CM and Peñalver-Pascual R: Multicenter analysis of survival and prognostic factors in pathologic stage I non-small-cell lung cancer according to the new 2009 TNM classification. Arch Bronconeumol 47: 441-446, 2011 (In English, Spanish).
9. Molina JR, Yang P, Cassivi SD, Schild SE and Adjei AA: Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc 83: 584-594, 2008.
10. Zarogoulidis K, Zarogoulidis P, Darwiche K, Boutsikou E, Machairiotis N, Tsakiridis K, Katsikogiannis N, Kougioumtzi I, Karapantzos I, Huang H and Spyratos D: Treatment of non-small cell lung cancer (NSCLC). J Thorac Dis 5 (Suppl 4): S389-S396, 2013.
11. Duma N, Santana-Davila R and Molina JR: Non-small cell lung cancer: Epidemiology, screening, diagnosis, and treatment. Mayo Clin Proc 94: 1623-1640, 2019.
12. Magome T, Froelich J, Takahashi Y, Arentsen L, Holtan S, Verneris MR, Brown K, Haga A, Nakagawa K, Chakrabarty JL, *et al*: Evaluation of functional marrow irradiation based on skeletal marrow composition obtained using dual-energy computed tomography. Int J Radiat Oncol Biol Phys 96: 679-687, 2016.
13. Liam CK, Andarini S, Lee P, Ho JC, Chau NQ and Tscheikuna J: Lung cancer staging now and in the future. Respirology 20: 526-534, 2015.
14. Yu W, Tang C, Hobbs BP, Li X, Koay EJ, Wistuba, II, Sepesi B, Behrens C, Canales JR, Cuentas ER, *et al*: Development and validation of a predictive radiomics model for clinical outcomes in stage I non-small cell lung cancer. Int J Radiat Oncol Biol Phys 102: 1090-1097, 2018.
15. Shim SS, Lee KS, Kim BT, Chung MJ, Lee EJ, Han J, Choi JY, Kwon OJ, Shim YM and Kim S: Non-small cell lung cancer: Prospective comparison of integrated FDG PET/CT and CT alone for preoperative staging. Radiology 236: 1011-1019, 2005.
16. van Timmeren JE, Leijenaar RTH, van Elmpt W, Reymen B, Oberije C, Monshouwer R, Bussink J, Brink C, Hansen O and Lambin P: Survival prediction of non-small cell lung cancer patients using radiomics analyses of cone-beam CT images. Radiother Oncol 123: 363-369, 2017.
17. Fass L: Imaging and cancer: A review. Mol Oncol 2: 115-152, 2008.
18. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, *et al*: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer 45: 228-247, 2009.
19. Choi ER, Lee HY, Jeong JY, Choi YL, Kim J, Bae J, Lee KS and Shim YM: Quantitative image variables reflect the intratumoral pathologic heterogeneity of lung adenocarcinoma. Oncotarget 7: 67302-67313, 2016.
20. Yu L, Tao G, Zhu L, Wang G, Li Z, Ye J and Chen Q: Prediction of pathologic stage in non-small cell lung cancer using machine learning algorithm based on CT image feature analysis. BMC Cancer 19: 464, 2019.
21. Tang X, Huang H, Du P, Wang L, Yin H and Xu X: Intratumoral and peritumoral CT-based radiomics strategy reveals distinct subtypes of non-small-cell lung cancer. J Cancer Res Clin Oncol 148: 2247-2260, 2022.
22. Zhang X, Zhang G, Qiu X, Yin J, Tan W, Yin X, Yang H, Liao L, Wang H and Zhang Y: Radiomics under 2D regions, 3D regions, and peritumoral regions reveal tumor heterogeneity in non-small cell lung cancer: A multicenter study. Radiol Med 128: 1079-1092, 2023.
23. Caruana R: Multitask learning. Machine Learning 28: 41-75, 1997.
24. Lin J, Yu Y, Zhang X, Wang Z and Li S: Classification of histological types and stages in non-small cell lung cancer using radiomic features based on CT images. J Digit Imaging 36: 1029-1037, 2023.
25. Bakr S, Gavaert O, Echegaray S, Ayers K, Zhou M, Shafiq M, Zheng H, Zhang W, Leung A, *et al*: Data for NSCLC Radiogenomics. Version 4 [Data set]. The Cancer Imaging Archive, 2017. https://www.cancerimagingarchive.net/collection/nsclc-radiogenomics/. Accessed December 13, 2024.

26. Aerts HJWL, Rios Velazques E, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, *et al*: Data From NSCLC-Radiomics-Genomics. Version 1 [Data set]. The Cancer Imaging Archive, 2015. https://www.cancerimagingarchive.net/collection/nsclc-radiomics-genomics/#citations. Accessed December 13, 2024.
27. Rami-Porta R, Bolejack V, Crowley J, Ball D, Kim J, Lyons G, Rice T, Suzuki K, Thomas CF Jr, Travis WD, *et al*: The IASLC lung cancer staging project: Proposals for the revisions of the T descriptors in the forthcoming eighth edition of the TNM classification for lung cancer. J Thorac Oncol 10: 990-1003, 2015.
28. Li M, Li X, Guo Y, Miao Z, Liu X, Guo S and Zhang H: Development and assessment of an individualized nomogram to predict colorectal cancer liver metastases. Quant Imaging Med Surg 10: 397-414, 2020.
29. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S and Aerts H: Computational radiomics system to decode the radiographic phenotype. Cancer Res 77: e104-e107, 2017.
30. Zwanenburg A, Leger S, Agolli L, Pilz K, Troost EGC, Richter C and Löck S: Assessing robustness of radiomic features by image perturbation. Sci Rep 9: 614, 2019.
31. McHugh DJ, Porta N, Little RA, Cheung S, Watson Y, Parker GJM, Jayson GC and O'Connor JPB: Image contrast, image pre-processing, and T(1) mapping affect MRI radiomic feature repeatability in patients with colorectal cancer liver metastases. Cancers (Basel) 13: 240, 2021.
32. Wu CY, Fu JY, Wu CF, Liu YH, Hsieh MJ, Wu YC, Yang CT and Tsai YH: Pathologic stage of nonsmall cell lung cancer patients presenting as resectable cases after neoadjuvant therapy did not predict the prognosis. Medicine 94: 1, 2015.
33. Song F, Song X, Feng Y, Fan G, Sun Y, Zhang P, Li J, Liu F and Zhang G: Radiomics feature analysis and model research for predicting histopathological subtypes of non-small cell lung cancer on CT images: A multi-dataset study. Med Phys 50: 4351-4365, 2023.
34. Ganeshan B, Abaleke S, Young RC, Chatwin CR and Miles KA: Texture analysis of non-small cell lung cancer on unenhanced computed tomography: Initial evidence for a relationship with tumour glucose metabolism and stage. Cancer Imaging 10: 137-143, 2010.
35. Erol M, Önner H and Küçükosmanoğlu İ: Association of fluorodeoxyglucose positron emission tomography radiomics features with clinicopathological factors and prognosis in lung squamous cell cancer. Nucl Med Mol Imaging 56: 306-312, 2022.
36. Shin SY, Hong IK and Jo YS: Quantitative computed tomography texture analysis: Can it improve diagnostic accuracy to differentiate malignant lymph nodes? Cancer Imaging 19: 25, 2019.
37. Ren D, Liu L, Sun A, Wei Y, Wu T, Wang Y, He X, Liu Z, Zhu J and Wang G: Prediction of solid pseudopapillary tumor invasiveness of the pancreas based on multiphase contrast-enhanced CT radiomics nomogram. Front Oncol 15: 1513193, 2025.
38. Miao L, Xiao G, Chen W, Yang G, Hong D, Wang Z, Zhang L and Huang W: Non-invasive assessment of programmed cell death ligand-1 expression using 18F-FDG PET-CT imaging in esophageal squamous cell carcinoma. Sci Rep 14: 26082, 2024.