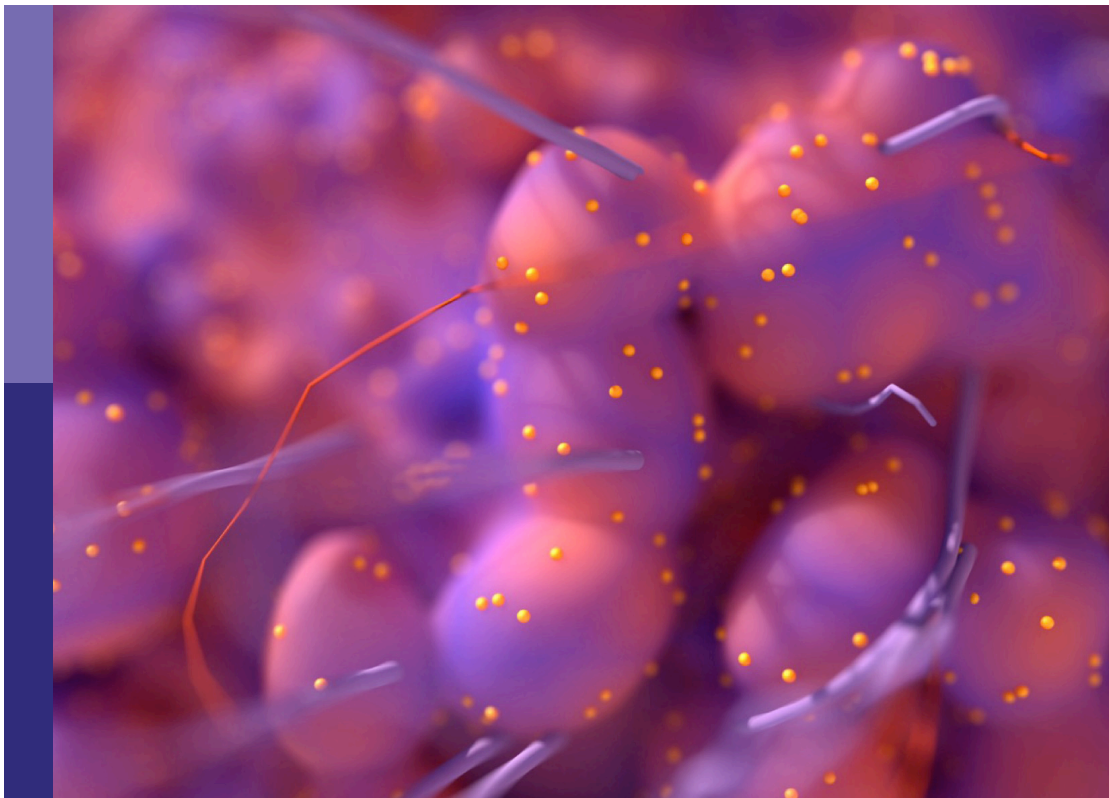
frontiers | Research Topics

Imaging in non-small cell lung cancer

Edited by
Yiyan Liu and Georgios S. Limouris

Published in
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-83251-736-9
DOI 10.3389/978-2-83251-736-9

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Imaging in non-small cell lung cancer

Topic editors

Yiyan Liu — University of Louisville, United States

Georgios S. Limouris — National and Kapodistrian University of Athens, Greece

Citation

Liu, Y., Limouris, G. S., eds. (2023). *Imaging in non-small cell lung cancer*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-736-9

Table of contents

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 07 February 2023
DOI 10.3389/fonc.2023.1143204

OPEN ACCESS

EDITED AND REVIEWED BY
Zaver Bhujwalla,
School of Medicine, Johns Hopkins University, United States

*CORRESPONDENCE
Yiyan Liu
yiyan.liu@louisville.edu

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 12 January 2023
ACCEPTED 31 January 2023
PUBLISHED 07 February 2023

CITATION
Liu Y (2023) Editorial: Imaging in non-small cell lung cancer.
Front. Oncol. 13:1143204.
doi: 10.3389/fonc.2023.1143204

Editorial: Imaging in non-small cell lung cancer

Yiyan Liu*

Department of Radiology, University of Louisville School of Medicine, Louisville, KY, United States

KEYWORDS

non-small cell lung cancer (NSCLC), radiomics, CT, PET/CT, NSCLC

Editorial on the Research Topic
Imaging in non-small cell lung cancer

Non-Small Cell Lung Cancer (NSCLC) accounts for approximately 80% to 85% of all lung cancer cases and is the leading cause of cancer death worldwide (1, 2). Beyond invasive histopathological diagnosis, imaging plays critical roles in screening, diagnosing, staging, restaging, detecting recurrence and/or metastasis, and monitoring therapeutic response of NSCLC. Significant limitations of traditional image modalities such as radiography, CT, and MRI have prompted the development of novel imaging strategies for better and more accurate tumor characterization and guidance for targeted therapy. Our Research Topic aimed to shed light on the latest imaging advances in NSCLC, which has been very attractive to our contributors and readers. We would like to express great appreciation for valuable contributions to the Research Topic by all authors.

Over the last two decades, F18-fluorodeoxyglucose PET/CT has been widely used for NSCLC, which has the unique advantage of metabolic evaluations with semi-quantitative parameters compared to the traditional images. Although there has been significant progress in FDG PET/CT applications and performance, further advancements can optimize imaging interventions as well as establish emerging approaches, such as radiomics, to improve the management of lung cancer. 8 of 17 published articles in our Research Topic are about the value of radiomics features in NSCLC.

Radiomics has been considered an “imaging biomarker”, and it is a quantitative approach with advanced mathematical analysis to medical imaging, which extracts many features from medical images using data-characterization algorithms. By providing 3-D characterization of the tumor, radiomics features based on CT, MRI and/or PET describe intra-tumor heterogeneity and have the potential to uncover disease characteristics. To date, radiomics has been most extensively studied in NSCLC compared to other tumors (3, 4). Published articles in the Research Topic have demonstrated new information about the values of radiomics in NSCLC.

Adequate identification of genotype and gene mutation is the basis for target therapy. Recent research showed correlation between the radiomics features and the epithelial growth factor receptor (EGFR) (5–7). In this topic collection, Zhang et al. reported that with correction of FDG PET/CT radiomics features to EGFR mutation analyses of tissue samples, FDG PET/CT radiomics models could help in discriminating EGFR positive from negative mutations in 173 preoperative patients with NSCLC, which could guide target therapy in patients with EGFR mutations. In another larger multicenter study including 728 patients with lung adenocarcinoma, Zhang et al. reported a two-user friendly nomogram by

calculating radiomics score to predict the EGFR mutation status. Wu et al. also demonstrated that pre-treatment CT based radiomics features could reliably predict EGFR mutation in 67 NSCLC patients, and the addition of clinical models further improved radiomics performance.

Beyond invasive approach to histopathologic diagnosis by surgery or biopsy, radiomics may help in the prediction of histology and stage of the tumors. Tang et al. analyzed FDG PET/MRI radiomics in 61 NSCLC patients and found that FDG PET/MRI radiomics features revealed different degrees of correlation with different tumors and could predict the preoperative histological classifications of the lesions: adenocarcinoma or squamous cell cancer.

Differentiation between lung metastasis and the 2nd lung primary is of incremental significance of treatment and prognosis in oncologic patients. Zhong et al. evaluated the value of CT radiomics in discriminating the second lung primary from lung metastasis in 252 oncologic patients with suspected lung lesions. 16 radiomics features and 4 clinical-radiographic features were selected to build the final model, which showed good discriminative capacity for the 2nd lung primary and lung metastasis.

Currently immunotherapy by immune checkpoint blockade is standard care in advanced NSCLC. Radiomics has been used in assessment of PD-L1 expression in NSCLC (8). Zhou et al. assessed the predictive role of FDG PET/CT based radiomics for tumor microenvironment immune types (TMIT) in 103 treatment-naïve NSCLC patients, including the expressions of programmed death ligand (PD-L1), programmed death 1 (PD-1) and CDF8+ tumor infiltrating lymphocytes. The results suggested that FDG PET/CT radiomics features had good performance in predicting the TMIT.

Zha et al. developed and validated a nomogram model based on CT radiomics features for preoperative prediction of visceral pleural invasion in 659 patients with lung adenocarcinoma. The results showed that the nomogram combining clinical and radiomics features had markedly improved accuracy, specificity, positive predictive value, and AUC for predicting visceral pleural invasion.

The micropapillary pattern is a marker of poor prognosis in NSCLC (Xu et al.). Li et al. reported a radiomics model based on nodule type stratification for preoperative prediction of micropapillary pattern in lung adenocarcinoma less than 2 cm. They found that ground glass opacity nodule type affected performance of the prediction.

In conclusion, radiomics is expected to optimize and augment image capabilities, and holds the great promise of valuable clinical applications in diagnosis, staging, prediction of treatment outcome and survival in NSCLC. Today we are on the brink of a new era in radiology artificial intelligence (AI). AI and deep learning models will facilitate faster clinical translation and implementation of radiomics in NSCLC. However, the significant variability in radiomics features used in different studies, in addition to the lack of reproducibility, suggests that the current data are still preliminary and more comprehensive studies are needed for validation of radiomics applications.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Acknowledgments

I would like to thank all authors for their contributions to the Research Topic.

Conflict of interest

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

Publisher's note

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics. *A Cancer J Clin* (2022) 72 (1):7–33. doi: 10.3322/caac.21708

2. Bade BC, Cruz C. Lung cancer 2020: Epidemiology, etiology, and prevention. *Clin Chest Med* (2020) 41(1):1–24. doi: 10.1016/j.ccm.2019.10.001

3. Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA. Nonm-small cell lung cancer: histopathologic correlates for texture parameters at CT. *Radiology* (2013) 266 (1):326–36. doi: 10.1148/radiol.12112428

4. Thawani R, McLane M, Beig N, Ghose S, Prasanna P, Velcheti V, et al. Radiomics and radiogenomics in lung cancer: a review for the clinician. *Lung Cancer* (2018) 115:34–41. doi: 10.1016/j.lungcan.2017.10.015

5. Avanzo M, Stancanello J, Pirrone G, Sartor G. Radiomics and deep learning in lung cancer. *Strahlenther Onkel* (2020) 196(10):879–87. doi: 10.1007/s00066-020-01625-9

6. Li S, Ding C, Zhang H, Song J, Wu L. Radiomics for the prediction of EGFR mutation subtypes in non-small cell lung cancer. *Myd Phys* (2019) 46(10):4545–52. doi: 10.1002/mp.13747

7. Fan L, Fang M, Tu W, Zhang D, Wang Y, Zhou X, et al. Radiomics signature: A biomarker for the preoperative distant metastatic prediction of stage I non-small cell lung cancer. *Acad Radiol* (2019) 26(9):1253–61. doi: 10.1016/j.acra.2018.11.004

8. Jiang M, Sun D, Guo Y, Guo Y, Xiao J, Wang L, et al. Asessing PD-L1 expression level by radiomics features from PET/CT in nonsmall cell lung cancer patients: an initial result. *Acad Radiol* (2020) 27(2):171–79. doi: 10.1016/j.acra.2019.04.016

frontiers in Oncology

ORIGINAL RESEARCH
published: 07 October 2020
doi: 10.3389/fonc.2020.542957

CT Radiomics in Predicting EGFR Mutation in Non-small Cell Lung Cancer: A Single Institutional Study

Shanshan Wu[1], *Guiquan Shen*[1], *Jujiang Mao*[1] *and Bo Gao*[1,2]*

[1] *Department of Radiology, The Affiliated Hospital of Guizhou Medical University, Guiyang, China,* [2] *Key Laboratory of Brain Imaging, Guizhou Medical University, Guiyang, China*

Objective: To evaluate the value of CT radiomics in predicting the epidermal growth factor receptor (EGFR) mutation of patients with non-small cell lung cancer (NSCLC), and combing with the clinical characteristic to construct the prediction model.

Methods: Sixty-seven cases of NSCLC confirmed by pathology were enrolled. The pre-treatment chest CT enhanced images were used in Radiomics analysis. Two experienced radiologists delineated the region of interest (ROI) on open source software 3D-Slicer. The feature of ROI was extracted by Pyradiomics software package and a total of 849 features were extracted. By calculating Pearson correlation coefficient between pair-wise features and LASSO method for feature screening. The prediction model was constructed by logical regression, diagnostic efficacy of the model by the area under the receiver operating characteristic (ROC) curve was calculated.

Results: Based on clinical model and the radiomics model, the AUC under the ROC was 0.8387 and 0.8815, respectively. The model combining clinical and radiomics features perfect best, the AUC under the ROC was 0.9724, the sensitivity and specificity were 85.3 and 90.9%, respectively.

Conclusions: Compared with clinical features or radiomics features alone, the model constructed by combining clinical and pre-treatment chest enhanced CT features may show more utility for improved patient stratification in EGFR mutation and EGFR wild.

Keywords: radiomics, non-small cell lung cancer, enhancement CT, EGFR, region of interest

OPEN ACCESS

Edited by:
Zhongxiang Ding,
Zhejiang University, China

Reviewed by:
Yuming Jiang,
Stanford University, United States
Long Jiang Zhang,
Nanjing General Hospital of Nanjing Military Command, China

***Correspondence:**
Bo Gao
gygb2004@gmc.edu.cn

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *15 March 2020*
Accepted: *03 September 2020*
Published: *07 October 2020*

Citation:
Wu S, Shen G, Mao J and Gao B (2020) CT Radiomics in Predicting EGFR Mutation in Non-small Cell Lung Cancer: A Single Institutional Study. Front. Oncol. 10:542957. doi: 10.3389/fonc.2020.542957

INTRODUCTION

Lung cancer is accounting for 13% of the new cancer in the world (1), which is the main cause of cancer-related death. As a lung solid tumor, lung cancer shows a wide range of molecular heterogeneity. In the past decade, the treatment of non-small cell lung cancer (NSCLC) has evolved from previous cytotoxic chemotherapy to target therapy based on molecular changes, due to significant breakthrough in molecular research for its theranostics (2, 3). Small molecule tyrosine kinase inhibitors (TKIs) targeting specific EGFR mutations are the first targeted drugs for the treatment of NSCLC. Riely et al. (4) reported the response rate of EGFR-TKIs in patients with EGFR mutations (60–80%) was significantly higher than that in patients with EGFR wild-type or unknown mutation status (10–20%). In addition, a large number of clinical trials of Schuler et al. (5) found treatments with errotene, gefitinib, or afatinibthe in NSCLC with EGFR mutation would get a longer progress-free survival and higher objective response rate, compared with standard first-line chemotherapy. However, they also found if a non-EGFR mutant lung cancer patient was

treated with targeted drugs such as gefitinib, the progression-free survival of the patient would be significantly shorter than those with first-line standard chemotherapy drugs (3). These data highlight the importance of accurately identifying the status of a patient's gene mutations in clinical to guide treatment.

For a clinical patient, the biopsy sample may be the only tumor material that can be used to detect the EGFR mutation status (6). However, tumor, it is possible to make the mutant DNA allele difficult to detect due to the sampling error and the observer difference, resulting in the occurrence of a false negative result (7). Second, living tissue examination is invasive, which increases the risk of complications. Last, genetic test is often expensive and some patients will not be able to afford it, which partly limits the clinical application of molecular test. Unfortunately, so far, there is still no stable, reliable clinical feature that can accurately predict the mutation status of EGFR. Radiogenomics has become a promising technique for identifying the gene phenotype in a tumor. As a new field of research, radiogenomics

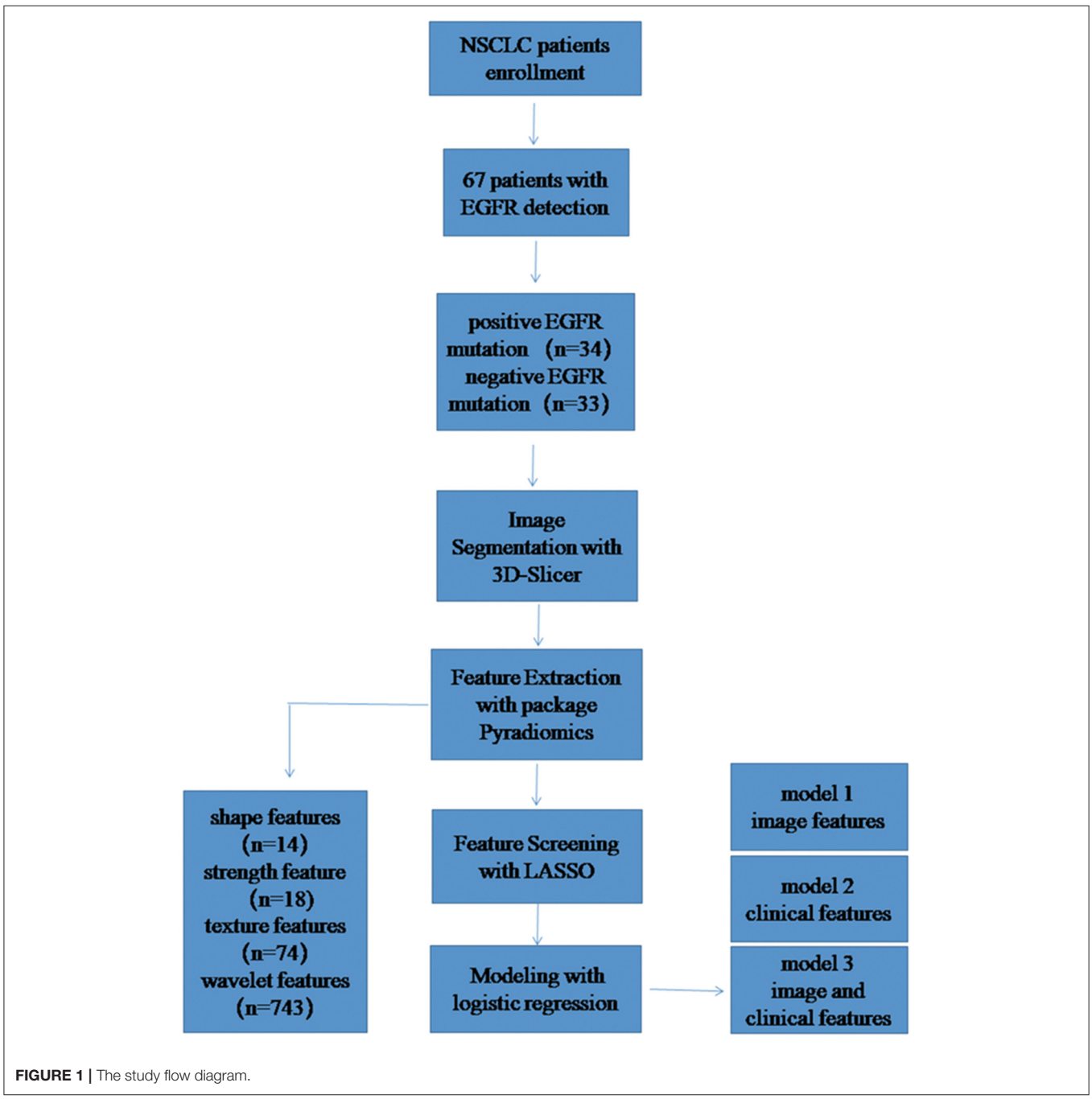


FIGURE 1 | The study flow diagram.

can integrate the image features and genome data of the disease, and to help explore the features of the image that can reflect the polymorphism or expression of the gene by digging the relation between the two. It is reported that using radiogenomics, Liu et al. (8) revealed EGFR mutation status of NSCLC was related to image features and could be predicted by five types of image feature. In this retrospective study, we tried to explore the relationships between CT imaging features and EGFR mutation in patients with NSCLC before treatment, and to construct a predictive model combined with clinical features.

TABLE 1 | General information of 67 NSCLC patients.

Subgroups	EGFR+	EGFR–	P-value
Cases	34	33	
Age, mean ± SD	53.0882 ± 8.1997	59.6061 ± 7.8020	0.023
Gender			<0.001
Males	11 (32.35%)	27 (81.82%)	
Females	23 (67.65%)	6 (18.18%)	
Smoking			<0.001
Yes	8 (23.53%)	26 (78.79%)	
No	26 (76.47%)	7 (21.21%)	
Pathological stages			0.518
IIIb, IIIc	15 (44.12%)	12 (36.36%)	
Iva, IVb	19 (55.88%)	21 (63.64%)	
Histological types			<0.001
Adenocarcinoma	32 (94.12%)	14 (42.42%)	
Squamous cell carcinoma	2 (5.88%)	19 (57.58%)	
Rad_score (mean)	4.167	−3.947	<0.001

MATERIALS AND METHODS

Patient Population

All NSCLC cases confirmed by pathology from May 2017 to November 2018 were retrospectively collected in our hospital. The inclusion criteria were: (1) NSCLC confirmed by histopathology with complete clinical data; (2) Complete contrast enhancement chest CT imaging before treatment and follow-up images; (3) Complete genetic testing information. The exclusion criteria were: (1) Small cell lung cancer confirmed by histopathology; (2) Incomplete clinical data; (3) Image quality does not meet the requirements due to serious respiratory motion artifacts. Finally, 67 patients with EGFR detection information were analyzed, including 34 patients with positive EGFR mutation and 33 patients with negative EGFR mutation.

Detection of EGFR Mutations

EGFR mutational analysis was performed on four tyrosine kinase domains (exons 18–21) that frequently mutated in lung cancer. The EGFR gene mutation was tested using human EGFR gene mutation detection kit (Beijing ACCB Biotechnology Company) using the Amplified refractory mutation system (ARMS) real-time technology.

CT Scanning Protocols

Chest CT contrast enhancement examinations were performed using TOSHIBA CT (Aquilion PRIME TSX-302A, Japan), scanning parameters: tube voltage 120 KV, tube current 350 mA, field of view 390.0 × 390.0 mm, reconstruction thickness 5 mm, reconstruction interval 5 mm. Non-ionic iodine contrast agent (Bayer Pharmaceuticals, Berlin, Germany; Ompaque, Shanghai,

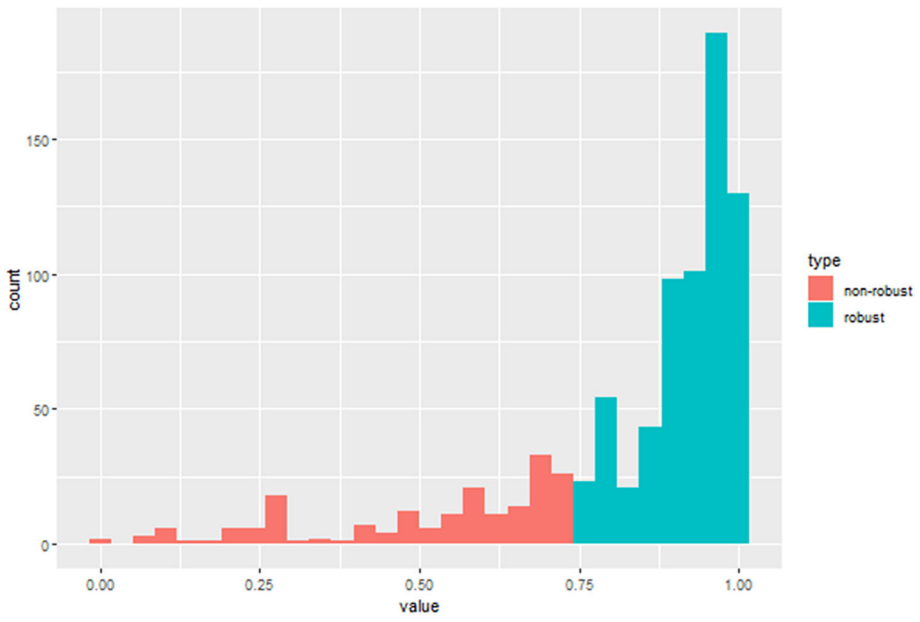


FIGURE 2 | ICC histogram of radiomics features.

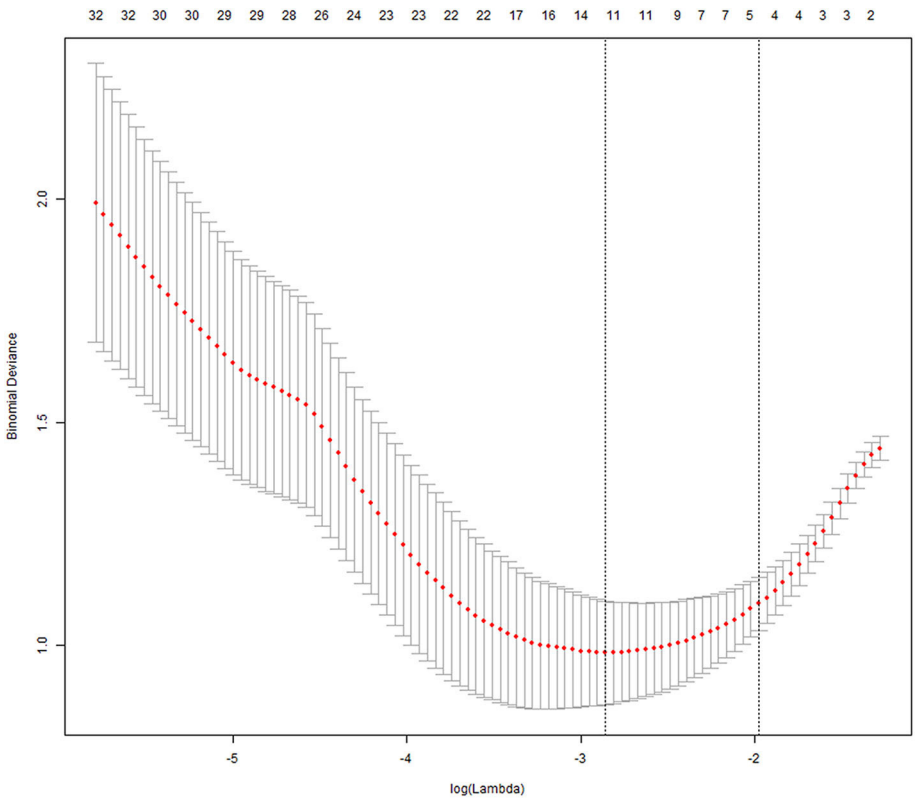


FIGURE 3 | LASSO method for screening of radiomics features.

GE) was intravenously injected at a rate of 2.5 ml/s using a high-pressure syringe at a dose of 1.3–1.5 ml/kg. The CT scanning was performed 50 s after the contrast injection.

Radiomics Analysis

Image Segmentation and Feature Extraction

The segmentation of the region of interest (ROI) was performed by two experienced radiologist, all segmented in the open source software 3D-Slicer (9) (v4.10.0, download address: https://www.slicer.org/), the image data is saved in DICOM format. According to the three-dimensional (3D) volume of the tumor after segmentation, the tumor 3D features were extracted in the open source software package Pyradiomics (10) (https://github.com/Radiomics/pyradiomics/releases), and a total of 849 features were extracted for each patient. Among them, there are 14 shape features, 18 first-order statistical features (strength features), 74 texture features, and 743 wavelet features. Then, calculating the consistency of two expert extraction features, select the features with intra-class correlation coefficient (ICC) >0.75 for subsequent feature screening.

Feature Screening

After the feature consistency test, we were calculating the Pearson correlation coefficient between all the features, one is randomly excluded from each pair of features with a correlation coefficient >0.9, the other feature is selected, and so on, the redundancy of the feature is reduced (11).The feature is then screened by the least absolute shrinkage and selection operator (LASSO), LASSO is an accepted algorithm that has been used for feature selection in high-dimensional variables. Finally, the model is constructed by logistic regression method (12). LASSO method was applied to select the features that were most distinguishable and build a logistic regression model. A radiomic score (Rad_score) was obtained for each patient using features selected and weighted by the respective coefficients. The detailed study flow is shown in **Figure 1**.

Development of the Multivariable Prediction Model

A multivariable logistic regression model was built using clinical variable and radiomics features. All of these features were included in the development of a diagnostic model to predict EGFR mutation. We also developed a radiomic nomogram based on a multivariable logistic analysis. We used radiomic signature (Rad_signature) to represent the possibility for each patient, which was obtained by the LASSO regression model developed by radiomic features.

Statistical Analysis

Statistical analysis was performed using R-3.4.4-win. In the clinical data of the patients, two independent sample *T*-tests were used to compare the age of the two samples, and

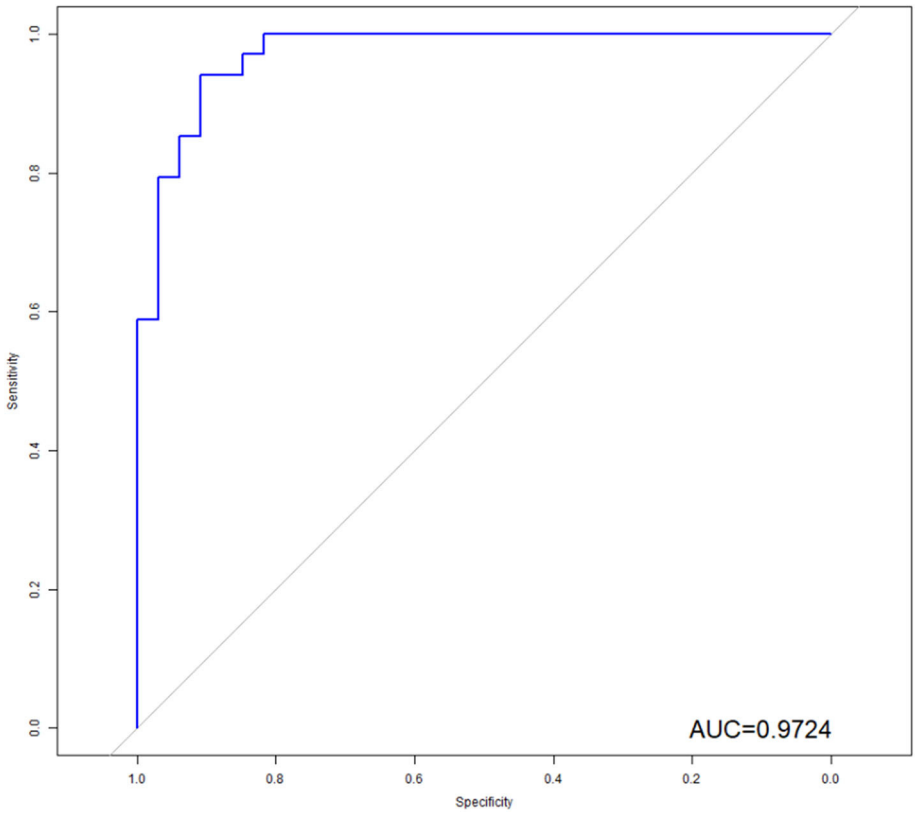


FIGURE 4 | ROC analysis of the model. As shown above, radiomics features combined with clinical variables had the potential ability to predict the EGFR mutation, the AUC was 0.9724.

TABLE 2 | Diagnostic accuracy of prediction model.

Sensibility %	Specificity %	Positive predictive value %	Negative predictive value %	Accuracy %
85.3 (29/34)	90.9 (30/33)	90.6 (29/32)	85.7 (30/35)	88.1 (59/67)

the other clinical data were compared by chi-square test, $p < 0.05$ as the difference was statistically significant. The diagnostic efficacy of the radiomics model was analyzed by the receiver operating characteristic (ROC) curve of the subjects, and the area under curve (AUC) was calculated, and the sensitivity and specificity of the model were also calculated. The "glmnet" package is used to realize LASSO. In order to fit the excellent model better, the 10-fold cross-validation method is adopted (13). The nomogram was depicted based on the results of the multivariate analysis using the "rms" package in R. The "Hmisc" package was used to investigate the performance of the nomogram in concordance with the C-index.

RESULTS

Patient

There were 34 patients with positive EGFR mutation and 33 patients with negative EGFR mutation, and there were significant differences in both cohorts detected in terms of smoking status, histological subtype, age, gender, or Rad_score. In the positive EGFR mutation, male patient comprised 32% (11/34) and female patients comprised 68% (23/34) of the total EGFR mutation cohort. The mean age was 53.1 ± 8.2 years. Adenocarcinoma was 94% (32/34) of all cases, squamous cell carcinoma was 6% (2/34). Smokers accounted for 24% (8/34) of patients, and non-smokers accounted for 76 (26/34). In the negative EGFR mutation, males were 82% (27/33) and females were 18% (6/33) of total patients. The mean age was 59.6 ± 7.8 years. Adenocarcinoma was 42% (14/33) of cases; squamous cell carcinoma was 58% (19/33). Smokers were 79% (26/33) of patients; non-smoker were 21 (7/33). More general information of patients was shown in **Table 1**.

Feature Consistency Testing and Screening

After the ICC analysis, there were 658 features ICC > 0.75 (**Figure 2**), these features combined with clinical variables will be

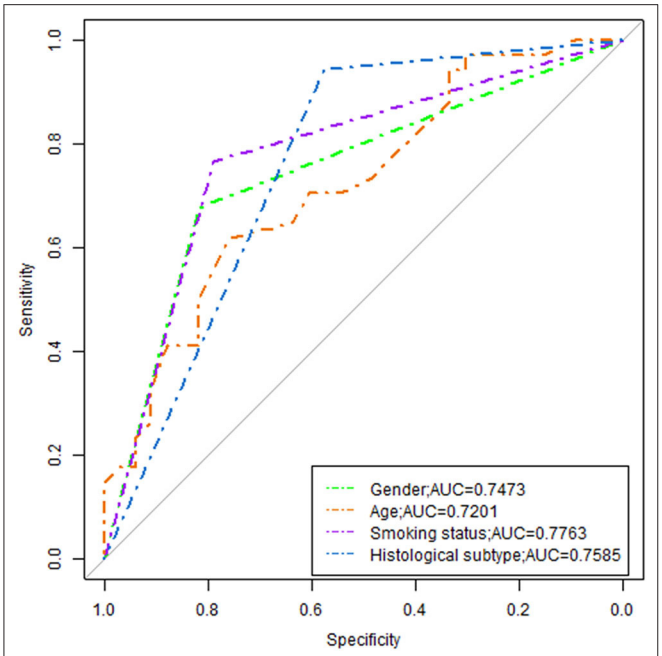


FIGURE 5 | ROC analysis of the clinical variable.

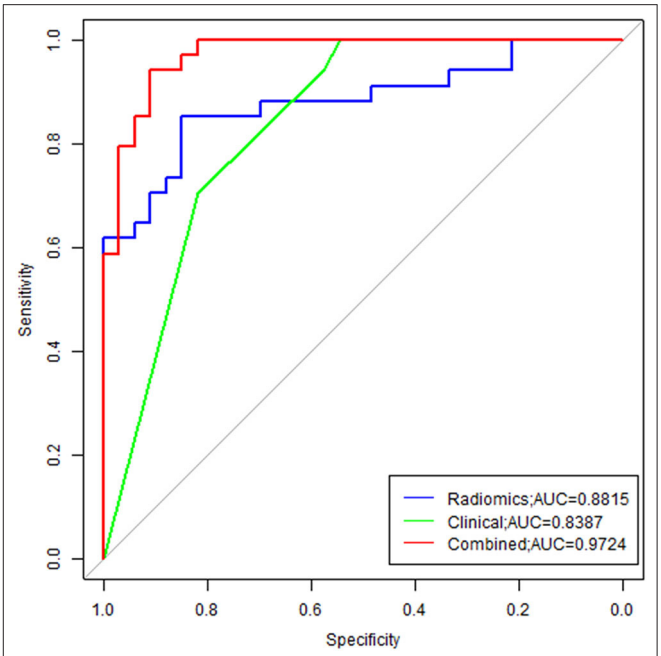


FIGURE 6 | Comparison of ROC between three models. Radiomics model was marked with blue line, and the AUC was 0.8815; Clinical model was marked with green line, and the AUC was 0.8387; The combine model was marked with red line, the AUC was 0.9724.

used for subsequent feature screening. After Pearson correlation coefficient filtering, the number of features changed from 658 to 382. Based on the LASSO dimension reduction, as **Figure 3**, when the variable is equal to 12, the error classification value is lower, and 12 features related to the patient's EGFR mutation are selected to construct the LASSO logistic regression model.

Development of the Multivariable Prediction Model

The LASSO logistic regression analysis (**Figure 4**) revealed that 10 radiomic features combined with two clinical features had the potential to build the prediction model for EGFR mutation by the training of 67 cases, which include Shape.Surfacevolume ratio; Gldm.dependencevariance; W_{HLL}.glcm.idn; w_{LHL} .ngtdm.Contrast; W_{LHH}.gldm.Large_emhpasis; W_{LHH}.gldm. Small_emphasis; W_{HHL}.gldm.variance; W_{HHL}.ngtdm.busyness; W_{LLL}.gldm.variance; W_{LLL}.glrlm.long_emphasis; smoking status and histological subtype. In addition, we calculated the sensitivity, specificity, positive predictive value, negative predictive value and accuracy to show the ability, the details are shown in **Table 2**.

ROC Curves Analysis for Radiomic Features and Clinical Predictors

The model revealed that gender, smoking status, clinical stage, and histological subtype were independent predictors of EGFR mutation. However, the model was merely developed by these features showing poor performance, the AUC ranged from 0.72 to 0.78 (**Figure 5**).

In order to illustrate the potential ability for prediction of EGFR mutation, we compared the models developed by radiomics features, clinical variables, and combination of them. As we can see from **Figure 6**, the ROC curves showed the good performance and generalization for the model built by radiomics features, AUC for radiomics model was 0.8815. When the model built by the both radiomics features and clinical variables, the AUC was 9724, which was significantly higher than the single radiomics model and the single clinical variables model, this indicates that the combine model showed best performance to predict EGFR gene mutations.

Analysis of an Individualized Prediction Model

Multivariate logistic regression analysis identified the clinical stage, age, gender, smoking status, and Rad_signature as independent predictors. The individualized EGFR mutation prediction model that consisted of the above independent predictors was visualized by the nomogram (**Figure 7**).

DISCUSSION

The malignant degree of NSCLC is high and the 5-years survival rate of patients is low, exploring effective treatment has become a widespread concern in recent years. Because of the high expression of epidermal growth factor in 40–80% of NSCLC patients, the molecule can be used as a target molecule for specific targeting therapy when the patient detects an EGFR mutation. So, the purpose of this study was to analyze the imaging features and clinical characteristics of 67 patients, to explore the

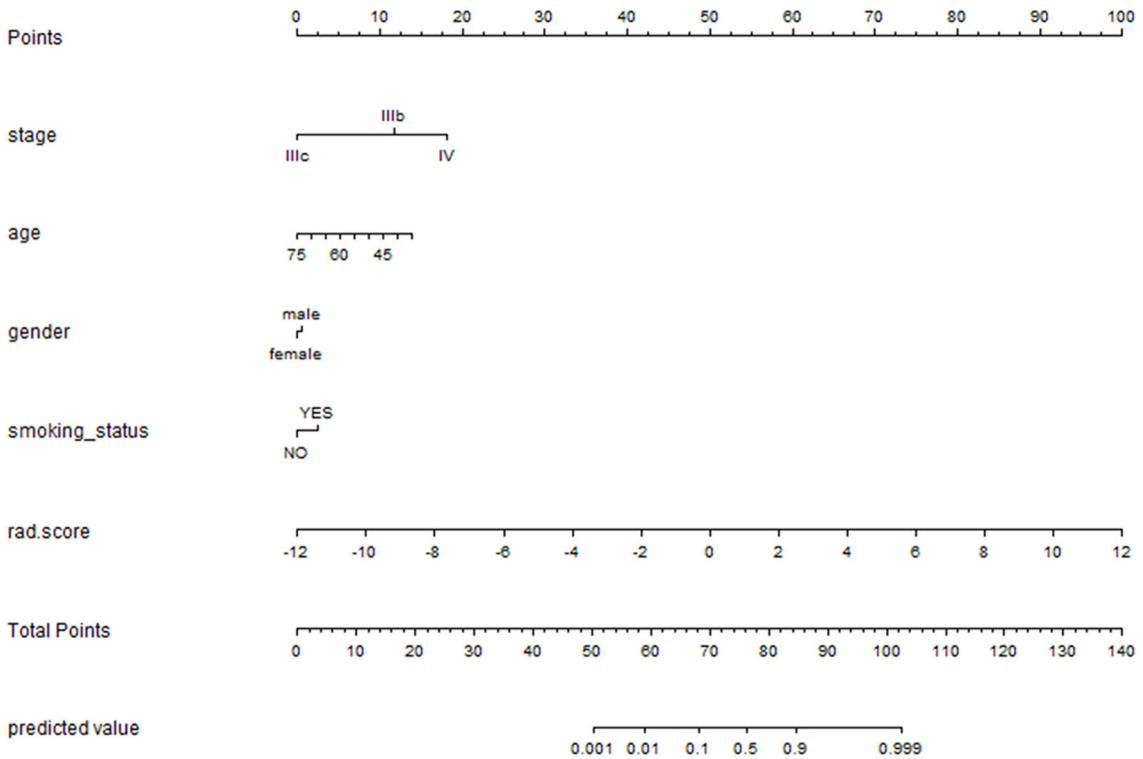


FIGURE 7 | The nomogram was depicted to present the relationship between radiomic features and clinical features and visually show the potential ability individually. The nomogram was built with the stage, age, gender, smoking status and rad_signatures.

correlation between chest enhanced CT imaging features and EGFR mutation status, to construct a prediction model combined with imaging features and clinical characteristics, and to draw a clinical predictive nomogram.

Early studies have shown inconsistent findings regarding the correlation between CT findings and EGFR mutation status in patients with NSCLC (14–16). According to a study by Zhou et al. (17), EGFR mutants and wild type were no difference in CT morphological features of lung tumors. In contrast, Rizzo et al. (18) reported that EGFR mutations are closely related to air bronchography, pleural retraction, lesion size, and presence or absence of fibrosis. In recent years, with the deepening of research, a large number of studies by Ozkan et al. (19) have demonstrated the potential of CT-based quantitative imaging features in identifying EGFR mutants and wild type in NSCLC. Mei et al. (20) also found that CT texture features are not only related to EGFR mutation status, but also can further distinguish patient mutation sequences (18–20).

In our study, a total of 67 patients with NSCLC were included in the analysis of clinical features, imaging findings and EGFR mutations. We extracted 849 features for each patient, which can provide us with more details and help us complete the evaluation of the lesion more comprehensively and accurately, it's very important. In response to EGFR mutations, we constructed three predictive models (based on patient clinical, radiomics, combined radiomics and clinical), the AUC values were 0.8387, 0.8815, and 0.9724, respectively. The results of this study are consistent with the study of Liu et al. (8), indicating the imaging features are associated with EGFR mutation status in NSCLC patients. Different from Liu et al., 849 features were extracted for each patient in this study, while Liu's study only extracted 219 features. The increase in the number of features provides us with more valuable information and greatly improved the performance of the model (0.709 vs. 0.972). we also perform a consistency test on the extracted features, only highly stable features (ICC $>$ 0.75) can be used for subsequent feature screening to ensure the stability of the study. In addition, we developed a radiomic nomogram based on a multivariable logistic analysis and radiomic signature (Rad_signature) to represent the possibility for each patient, these are not available in Liu et al. We also can find that in our study the image features used to construct the predictive model are texture, wavelet features, the results were not match Hsu et al. (21) previously reported that tumor size are correlation, which may be due to previous research based on only one level of tumor to extract tumor size characteristics (most the larger tumor areas), and our study focused on the size of the entire tumor extraction, leading to the emergence of differences. When compare the three models, we can find that the model combining imaging features and clinical has higher diagnostic efficiency than the single image model or clinical model; This also indicates that the image features and clinical

features may reflect different valuable prediction information, the combination of two can complement the information and improve the prediction ability.

Compared with recent studies by Tu et al. (22) and Zhao et al. (23), the diagnostic efficacy of this study is significantly higher. This may be because our study is based on enhanced CT images. Compared with plain images, enhanced images can reflect tumors more valuable information on blood supply, internal lesions and more. Secondly, in some CT images of patients with lung cancer and atelectasis, enhanced images can more accurately delineate the edge of the lesion, reduce the impact on imaging features, and improve the diagnostic efficacy of the model for EGFR mutations in non-small cell lung cancer. In addition, Tu et al. and Zhao et al. Studies have lots of patients, but the sample size in our study is small, and the samples from the same institution will also have an impact on the results.

Although this study has achieved high diagnostic performance, the study still has some limitations. First, according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) (24), in addition to internal cross-verification, the developed prediction model needs to evaluate the performance in external data to avoid overfitting. In this study, due to the small sample size, we have only performed internal cross-validation, and the independent model assessment could not be performed, which may have an impact on the true diagnostic performance of the prediction model. The next study requires large sample data from multi-center and independent model validation to confirm our findings. Secondly, the 5-mm slice thickness image is used in this study, due to the acquisition parameter have influence in the feature stability (25, 26), it can also lead to an inaccurate evaluation. In the future research, a prospective study is adopted, and the comparison between the models can be more facilitated by the standardized image acquisition and reconstruction algorithm, so as to improve the generalization of the model and the clinical application capability.

Our preliminary study indicates CT radiomics features are associated with EGFR mutation in patients with NSCLC, and CT imaging features of lesions on pretreatment may function as non-invasive biomarkers for improve stratification in patients with EGFR mutation and EGFR wild.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

BG and SW: project design. BG: project support. SW and GS: data collection. SW, GS, and JM: data analysis. SW: manuscript writing. All authors: final approval of manuscript.

FUNDING

This work was supported by National Natural Science Foundation of China (Grant No. 81871333); Guizhou Province Science and Technology Infrastructure Project (No. [2020]4Y-159); Guiyang science and technology project (No. 2020-10-3); Innovation group project of Guizhou Province Educational Commission; Key Laboratory Construction Project of Guizhou Medical University.

REFERENCES

1. Halpenny DF, Riely GJ, Hayes S, Yu H, Zheng J, Moskowitz CS, et al. Are there imaging characteristics associated with lung adenocarcinomas harboring ALK rearrangements? *Lung Cancer.* (2014) 86:190–4. doi: 10.1016/j.lungcan.2014.09.007
2. Jackman DM, Miller VA, Cioffredi LA, Yeap BY, Janne PA, Riely GJ, et al. Impact of epidermal growth factor receptor and KRAS mutations on clinical outcomes in previously untreated non-small cell lung cancer patients: results of an online tumor registry of clinical trials. *Clin Cancer Res.* (2009) 15:5267–73. doi: 10.1158/1078-0432.ccr-09-0888
3. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med.* (2009) 361:947–57. doi: 10.1056/NEJMoa0810699
4. Riely GJ, Pao W, Pham D, Li AR, Rizvi N, Venkatraman ES, et al. Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. *Clin Cancer Res.* (2006) 12(3 Pt 1):839–44. doi: 10.1158/1078-0432.ccr-05-1846
5. Schuler M, Yang CH, Yamamoto N, O'Byrne K, Hirsh V, Mok T, et al. LUX-lung 3: A randomized, open-label, phase III study of afatinib vs pemetrexed and cisplatin as first-line treatment for patients with advanced adenocarcinoma of the lung harboring EGFR-activating mutations. *Lung Cancer.* (2012) 77:S25–6. doi: 10.1016/j.lungcan.2012.05.043
6. Guillermo PJ, JNne PA, Lee JC, Sean T, Heidi G, Stacey G, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science.* (2004) 304:1497–1500. doi: 10.1126/science.1099314
7. Devarakonda S, Morgensztern D, Govindan, R. Genomic alterations in lung adenocarcinoma. *Lancet Oncol.* (2015) 16:e342–e351. doi: 10.1016/S1470-2045(15)00077-7
8. Liu Y, Kim J, Balagurunathan Y, Li Q, Garcia AL, Stringfield O, et al. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas. *Clin. Lung Cancer.* (2016) 17:441–8.e446. doi: 10.1016/j.cllc.2016.02.001
9. Emmanuel Rios V, Chintan P, Mohammed J, Mak RH, Angela VB, Fennessy FM, et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer. *Scient Rep.* (2013) 3:3529. doi: 10.1038/srep03529
10. Jjm VG, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 77:e104. doi: 10.1158/0008-5472.CAN-17-0339

11. Ishibashi N, Maebayashi T, Aizawa T, Sakaguchi M, Nishimaki H, Masuda S. Correlation between the Ki-67 proliferation index and response to radiation therapy in small cell lung cancer. *Radiat Oncol.* (2017) 12:16. doi: 10.1186/s13014-016-0744-1
12. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Resonance Imaging.* (2012) 30:1234–48. doi: 10.1016/j.mri.2012.06.010
13. Chen A, Lu L, Pu X, Yu T, Yang H, Schwartz LH, et al. CT-based radiomics model for predicting brain metastasis in category T1 lung adenocarcinoma. *AJR Am J Roentgenol.* (2019) 213:1–6. doi: 10.2214/ajr.18.20591
14. Lynda C, Andersen JN, Andrew, F. (2011). Cancer genomics: from discovery science to personalized medicine. *Nat. Med.* 17:297–303. doi: 10.1038/nm.2323
15. Garraway LA, Verweij J, Ballman KV. Precision oncology: an overview. *J Clin Oncol.* (2013) 31:1803–5. doi: 10.1200/jco.2013.49.4799
16. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* (2016) 278:563–77. doi: 10.1148/radiol,0.2015151169
17. Zhou JY, Zheng J, Yu ZF, Xiao WB, Zhao J, Sun K, et al. Comparative analysis of clinicoradiologic characteristics of lung adenocarcinomas with ALK rearrangements or EGFR mutations. *Eur Radiol.* (2015) 25:1257–66. doi: 10.1007/s00330-014-3516-z
18. Rizzo S, Petrella F, Buscarino V, Maria FD, Raimondi S, Barberis M, et al. CT radiogenomic characterization of EGFR, K-RAS, and ALK mutations in non-small cell lung cancer. *Eur Radiol.* (2016) 26:32–42. doi: 10.1007/s00330-015-3814-0
19. Ozkan E, West A, Dedelow JA, Chu BF, Zhao W, Yildiz VO, et al. CT gray-level texture analysis as a quantitative imaging biomarker of epidermal growth factor receptor mutation status in adenocarcinoma of the lung. *AJR Am J Roentgenol.* (2015) 205:1016–25. doi: 10.2214/ajr.14.14147
20. Mei D, Luo Y, Wang Y, Gong J. CT texture analysis of lung adenocarcinoma: can Radiomic features be surrogate biomarkers for EGFR mutation statuses. *Cancer Imaging.* (2018) 18:52. doi: 10.1186/s40644-018-0184-2
21. Hsu JS, Huang MS, Chen CY, Liu GC, Liu TC, Chong IW, et al. Correlation between EGFR mutation status and computed tomography features in patients with advanced pulmonary adenocarcinoma. *J Thorac Imaging.* (2014) 29:357–63. doi: 10.1097/rti.0000000000000116
22. Tu W, Sun G, Fan L, Wang Y, Xia Y, Guan Y, et al. Radiomics signature: a potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology. *Lung Cancer.* (2019) 132:28–35. doi: 10.1016/j.lungcan.2019.03.025
23. Zhao W, Yang J, Ni B, Bi D, Sun Y, Xu M, et al. Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. *Cancer Med.* (2019) 8:3532–43. doi: 10.1002/cam4.2233
24. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Europ Urol.* (2015) 67:1142–51. doi: 10.1016/j.eururo.2014.11.025
25. Li Y, Lu L, Xiao M, Dercle L, Huang Y, Zhang Z, et al. CT slice thickness and convolution kernel affect performance of a radiomic model for predicting EGFR status in non-small cell lung cancer: a preliminary study. *Sci Rep.* (2018) 8:17913. doi: 10.1038/s41598-018-36421-0
26. Owens CA, Peterson CB, Tang C, Koay EJ, Yu W, Mackin DS, et al. Lung tumor segmentation methods: impact on the uncertainty of radiomics features for non-small cell lung cancer. *PLoS ONE.* (2018) 13:e0205003. doi: 10.1371/journal.pone.0205003

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 08 October 2020
doi: 10.3389/fonc.2020.568857

Performance of ^{18}F-FDG PET/CT Radiomics for Predicting EGFR Mutation Status in Patients With Non-Small Cell Lung Cancer

Min Zhang[1†], Yiming Bao[2†], Weiwei Rui[3], Chengfang Shangguan[4], Jiajun Liu[1], Jianwei Xu[2], Xiaozhu Lin[1], Miao Zhang[1], Xinyun Huang[1], Yilei Zhou[1], Qian Qu[1], Hongping Meng[1], Dahong Qian[1,2] and Biao Li[1*]*

[1] Department of Nuclear Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2] Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai, China, [3] Department of Pathology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [4] Department of Oncology, Rujin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

OPEN ACCESS

Edited by:
Yiyan Liu,
The State University of New Jersey, United States

Reviewed by:
Xiaohua Zhu,
Huazhong University of Science and Technology, China
Ding Chong Yang,
The First Affiliated Hospital of Nanjing Medical University, China

***Correspondence:**
Biao Li
lb10363@rjh.com.cn
Dahong Qian
dahong.qian@sjtu.edu.cn

[†]These authors share first authorship

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 02 June 2020
Accepted: 22 September 2020
Published: 08 October 2020

Citation:
Zhang M, Bao Y, Rui W, Shangguan C, Liu J, Xu J, Lin X, Zhang M, Huang X, Zhou Y, Qu Q, Meng H, Qian D and Li B (2020) Performance of ^{18}F-FDG PET/CT Radiomics for Predicting EGFR Mutation Status in Patients With Non-Small Cell Lung Cancer. Front. Oncol. 10:568857. doi: 10.3389/fonc.2020.568857

Objective: To assess the performance of pretreatment ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) radiomics features for predicting EGFR mutation status in patients with non-small cell lung cancer (NSCLC).

Patients and Methods: We enrolled total 173 patients with histologically proven NSCLC who underwent preoperative ^{18}F-FDG PET/CT. Tumor tissues of all patients were tested for EGFR mutation status. A PET/CT radiomics prediction model was established through multi-step feature selection. The predictive performances of radiomics model, clinical features and conventional PET-derived semi-quantitative parameters were compared using receiver operating curves (ROCs) analysis.

Results: Four CT and two PET radiomics features were finally selected to build the PET/CT radiomics model. Compared with area under the ROC curve (AUC) equal to 0.664, 0.683 and 0.662 for clinical features, maximum standardized uptake values (SUV_{max}) and total lesion glycolysis (TLG), the PET/CT radiomics model showed better performance to discriminate between EGFR positive and negative mutations with the AUC of 0.769 and the accuracy of 67.06% after 10-fold cross-validation. The combined model, based on the PET/CT radiomics and clinical feature (gender) further improved the AUC to 0.827 and the accuracy to 75.29%. Only one PET radiomics feature demonstrated significant but low predictive ability (AUC = 0.661) for differentiating 19 Del from 21 L858R mutation subtypes.

Conclusions: EGFR mutations status in patients with NSCLC could be well predicted by the combined model based on ^{18}F-FDG PET/CT radiomics and clinical feature, providing an alternative useful method for the selection of targeted therapy.

Keywords: positron emission tomography/computed tomography, radiomics, lung cancer, epidermal growth factor receptor, ^{18}F-fluorodeoxyglucose

INTRODUCTION

Lung cancer is the leading cause of cancer-related death in the world (1). Non-small cell lung cancer (NSCLC) accounts for approximately 80% to 85% of all lung cancers (2). Epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor (TKI) has become a first-line drug in the treatment of NSCLC. Because the efficacy of TKI therapy is closely related to EGFR mutation status, identification of mutation status before the administration of TKI is crucial in achieving the best curative effect. Furthermore, exon 19 deletion (19 del) and exon 21 L858R point mutation (21 L858R),the most common mutation subtypes of EGFR (3), demonstrate different clinical outcomes in patients with NSCLC after TKI treatment (4, 5). *Current molecular testing for identifying EGFR mutation status is mainly based on tumor tissue from biopsies and surgical resection* (6). However, focal tissue testing may sometimes be limited by invasive procedures or tissue samples that are not readily available (7), causing patients to lose potential opportunities for EGFR-TKI treatment.

Medical imaging can reflect tumor gene-driven phenotype (8). ^{18}F-fluorodeoxyglucose (^{18}F-FDG) PET/CT, as a noninvasive molecular imaging tool, has been widely used in the evaluation of glucose metabolic phenotype of tumor (6). Previous data has suggested that several genes associated with glucose metabolism, including GLUT1 (9), GPI, G6PD, PKM2, and GAPDH (10), are down-regulated in EGFR-mutated lung cancer. Therefore, numerous studies have explored the relationship between ^{18}F-FDG PET/CT images and EGFR mutation status. Some studies suggested that there was significantly lower maximum standardized uptake values (SUV_{max}) of NSCLCs with EGFR mutations than those with wild type (11–14), but other studies reported non-significant (15) or opposite results (16). These confusing findings may be related to intra-tumoral heterogeneity of EGFR mutation (17) that the PET-derived semi-quantitative parameters cannot well reflect.

Radiomics data obtained using mathematical algorithms can quantitatively describe the spatial relationship between voxels, and become an important tool to study tumor heterogeneity *in vivo* (18). To date, most studies using radiomics for the prediction of EGFR mutation status in NSCLC are based on chest CT images (19, 20), whereas few studies about the relationship between PET or PET/CT radiomics features and EGFR mutation status in lung cancer (21–23) are conducted.

In the present study, both PET and CT radiomics features that significantly discriminated EGFR mutation status were extracted and selected for establishing a robust predictive model. Then we compared the predictive performances of the radiomics model, clinical features, and conventional PET-derived semi-quantitative parameters. Moreover, we tried to investigate the possibility of PET/CT radiomics features for distinguishing the 19 del from the 21 L858R mutation which both are two main mutation subtypes of EGFR.

MATERIALS AND METHODS

Subjects

A total of 173 patients (115 men, 58 women; mean [± SD] age 60.9 ± 10.9 years [range, 27–86 years]) with histologically proven NSCLC, who had undergone pre-treatment ^{18}F-FDG PET/CT between January 2017 and March 2018, were included in this study. This retrospective study was approved by the Ethics Committee of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine.

^{18}F-FDG PET/CT Imaging

Patients were required to fast for at least 6 h before ^{18}F-FDG PET/CT scan using GE Discovery VCT64 system, and their serum glucose levels were maintained to < 7.8 mmol/L. Whole-body imaging was performed approximately 60 min after the intravenous administration of 5.55 MBq of ^{18}F-FDG per kilogram of body weight. Emission images were acquired for 3 min per bed position using 128 × 128 matrix size, 28 subsets, 2 iterations and full-width half-maximum post-filtering. CT images were acquired using 140 kV tube voltage, 220 mA tube current, and 3.75 mm section thickness. PET images were reconstructed based on an ordered-subset expectation maximization algorithm with photon attenuation correction from CT data.

EGFR Mutation Status Analysis

Tissue samples from lung tumors were obtained through biopsy or surgical resection followed by 10% formalin fixation, paraffin embedding, and sectioning. After extracting DNA from sample sections, the nucleotide sequence encoding the kinase domain (exons 18-21) of EGFR was tested using an amplification refractory mutation system polymerase chain reaction (24) or target sequencing method based on polymerase chain reaction (25) using the X10 system (Illumina, San Diego, CA, USA).

PET/CT Image Feature Extraction, Selection and Model Establishment

All segmentation was performed by experienced nuclear medicine physicians blinded to the mutation data using an open-source ITK-SNAP software (version 3.6, https://www.itksnap.org) to manually outline the contour of the volume of interest on CT images, and automatically delineated on PET images using a fixed SUV_{max} threshold of 2.5 as previously reported (21). The extraction and selection of radiomics features were performed according to the following steps (**Figure 1**):

1. Before extraction of radiomics features, filters including Laplacian of Gaussian, wavelet, square, square root, logarithm and exponential (26) (**Supplemental Table 1**), were applied to the original PET/CT images to highlight image features for more efficient feature extraction.
2. Based on the original and filtered PET/CT images mentioned above, several types of well-designed image features were calculated using pyradiomics python package (27). These features are designed in compliance with the Image Biomarker Standardization Initiative (28) including First-order statistics, Shape, Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Dependence Matrix (GLDM), Gray Level Run Length Matrix (GLRLM) and Neighboring Gray Tone Difference

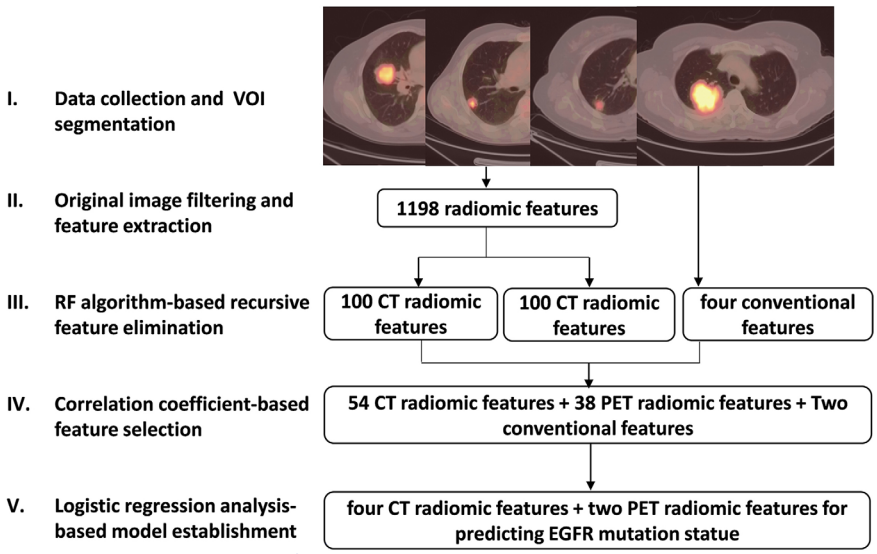


FIGURE 1 | Schematic diagram of image feature extraction and selection steps.

Matrix (NGTDM) (**Supplemental Table 2**). A total of 1198 PET and CT radiomics features were then extracted.

3. A recursive feature elimination (29) method based on random forest (RF) algorithm was developed to delete features with minimum weight coefficient. Compared to other regularization based embedded methods like Lasso and Ridge, this random forest-based wrapper feature selection method is more convenient and more intuitive for researchers to find out the most relevant features corresponding to the predication target. Among 1198 radiomics features, that with the lowest correlation with EGFR mutation status was removed during current random forest model training iteration, and the most suitable feature sub-package was reserved for next iteration. Finally, 100 CT and 100 PET radiomics features were retained (**Supplemental Table 3**).
4. The Spearman correlation coefficient (r) was used to assess the correlation between 100 PET/CT radiomics features and four conventional PET-derived semi-quantitative parameters including SUV_{max}, mean SUV (SUV_{mean}), metabolic tumor volume (MTV) and total lesion glycolysis (TLG) (illustrated in **Supplemental Figure 1**). In all pair features with $r > 0.85$ that were highly correlated and likely to provide redundance rather than complementary information about the mutation status, the one with the lower area under the curve (AUC) by receiver operating characteristic (ROC) analysis for predicting EGFR mutation status was excluded. As a result, 54 CT and 38 PET radiomics features as well as SUV_{max} and TLG were included.
5. The univariate and multivariate (**Supplemental Tables 4, 6**) logistic regression (LR) was ultimately used to screen out the CT and PET radiomics and clinical features that can be significant to establish a robust prediction model for differentiating EGFR mutation status, and the PET/CT radiomics prediction score for EGFR mutation probability of each patient was calculated based on this model.

Statistical Analysis

Data were analyzed using SPSS version 19.0 (IBM Corporation, Armonk, NY, USA). Spearman correlation analysis was performed to remove redundant radiomics features. Continuous data were compared using the independent samples t test. The χ^2 test was used to compare categorical data such as patient sex. Univariate and multivariate logistic regression was used to screen out final significant variables. 10-fold cross-validation of prediction model based on selected features using machine learning algorithm of RF, support vector machine (SVM) or traditional statistics of LR were performed to test the generalization ability of the models. ROC curves were analyzed to evaluate the performance of PET/CT radiomics model for predicting EGFR mutation status. Statistical significance was set at $p < 0.05$.

RESULTS

Patient Characteristics

As shown in **Table 1**, 173 patients with NSCLC were enrolled in the present study, among whom 71 (41%) tested positively for an EGFR mutation (EGFR+) and 102 (59%) were EGFR-negative (EGFR-). Female patients demonstrated a significantly higher EGFR mutation rate (64% [37/58]) than male patients (30% [34/115]). There was no statistical difference in age between patients with or without EGFR mutations. 39% (68/173) and 61% (105/173) of patients were stage I/II and stage III/IV, respectively. Seventy-one percent (122/173) of the pathological types of NSCLCs were adenocarcinoma. Among the 71 patients who

TABLE 1 | Patient Characteristics.

	EGFR+	EGFR-	*p* Value
No. of patients	71	102	
Sex			
Male	34	81	<0.001
Female	37	21	
Age (y)			
Mean ± SD	60.06 ± 10.93	61.54 ± 10.89	*NS*
Range	27 ~ 86	32 ~ 83	
Clinical Stage			
I	18	28	*NS*
II	8	14	
III	19	23	
IV	26	37	
Histology			
Adenocarcinoma	60	62	0.004
Squamous cell carcinoma	8	31	
Large cell neuroendocrine carcinoma	0	4	
NSCLC-NOS	3	5	
EGFR mutation subtype			
18 G719S	3	/	
19 Del	29	/	
20 T790M	1	/	
21 L858R	38	/	

NOS, not otherwise specified; NS, not significant.

were EGFR+, 38 (54%) harbored the 21 L858R mutation, 29 (41%) had the 19 del mutation, 3 (4%) had the 18 G719A substitution mutation, and 1 (1%) had the 20 T790M substitution mutation. In all clinical features, only gender was an independent and significant variable for differentiating EGFR mutation status after multivariate logistic regression analysis (**Supplemental Table 4**).

Characteristic of Selected PET/CT Radiomics Features

Eventually, four CT and two PET radiomics features were selected to build the radiomics model based on the 173 patients, including ct_original_glszm_High Gray Level Zone Emphasis (GLSZM_HGLZE), ct_wavelet_HLL_glszm_Gray Level Non-Uniformity Normalized (GLSZM_GLNN), ct_wavelet_HLL_glszm_Zone Entropy (GLSZM_ZE), ct_exponential_gldm_Dependence Variance (GLDM_DV), pet_wavelet_LHH_firstorder_Skewness (First-order_Skewness (LHH)), pet_wavelet_LLL_firstorder_Skewness (First-order_Skewness (LLL)). The definitions of these selected radiomics features were shown in **Supplemental Table 5**. The PET/CT radiomics model prediction score for EGFR mutation probability of each patient was calculated using the following formula:

$$\text{PET/CT radiomics model prediction score} = -6.142 - 2.736 \times \text{GLSZM\_HGLZE} + 5.815 \times \text{GLSZM\_GLNUN} + 5.173 \times \text{GLSZM\_ZE} + 7.737 \times \text{GLDM\_DV} - 1.734 \times \text{First-order\_Skewness (LHH)} - 6.142 \times \text{First-order\_Skewness (LLL)}.$$

The median and the interquartile range for selected PET/CT radiomics features and conventional PET parameters (SUV_{max} and TLG) was shown in **Table 2**. There was significant difference of every individual radiomics feature, SUV_{max} and TLG between the EGFR+ and EGFR- groups. Meanwhile, the tumors with EGFR+ had higher radiomics model score than those with EGFR- (0.722 vs. 0.170, $p < 0.001$). The PET/CT radiomics model prediction score for each patient was displayed in **Figure 2**.

Performance of the PET/CT Radiomics Model

The performance of PET/CT radiomics model was evaluated and compared with conventional PET-derived semi-quantitative parameters and clinical features for distinguishing EGFR+ from EGFR-. Both CT (AUC=0.792) and PET alone (AUC=0.738) radiomics model had better predictive performance than SUV_{max} (AUC=0.683), TLG (AUC=0.662) and gender (AUC=0.664). The AUC of PET/CT radiomics model further reached 0.868 with sensitivity of 92.8%, specificity of 66.3% and accuracy of 77.1%. Gender was only significant clinical predictor of EGFR mutation status (AUC=0.664), and used in the combined model in our study, whereas other clinical characteristics were excluded from the diagnostic model after multivariate regression analysis. The combined model, based on the PET/CT radiomics features and gender showed a comparable AUC (0.866) to PET/CT radiomics model. The sensitivity, specificity, and accuracy of different models and individual parameter in the training set were shown in **Table 3**. Subsequently, 10-fold cross-validation of the diagnostic model based on selected features using machine learning algorithm of SVM (**Table 3**), RF or traditional statistics of LR (**Supplemental Table 7**) were performed to further test the generalization ability of the models. The AUCs of PET radiomics, CT radiomics, PET/CT radiomics and combined models based on SVM were respectively 0.750, 0.754, 0.769 and 0.827.

In addition, we tried to investigate the possibility of radiomics features for discriminating two main mutation subtypes (**Table 4**). As previous reported, there was no difference of SUV_{max} or TLG between the 19 del and the 21 L858R mutation group. In all radiomics features, only one PET radiomics feature (pet_logarithm_glcm_Difference Variance, GLCM_DV) was significantly predictive (AUC=0.661) for differentiating these two mutation subtypes. However, it had low accuracy (43.1%) for the prediction of EGFR mutation subtypes.

DISCUSSION

EGFR-TKI is an important treatment for patients with NSCLC. When treated with TKI, patients with EGFR mutations experience significantly longer survival than those with wild-type EGFR. As such, identification of EGFR mutation status is crucial for TKI treatment to be effective; however, the molecular test for EGFR mutation status sometimes cannot be performed when a tumor sample is not available.

TABLE 2 | Characteristic of selected PET/CT radiomic features and conventional PET parameters.

Characteristic	EGFR- (N=102)	EGFR+ (N=71)	*p* Value
Conventional PET parameters			
SUV_{max}	11.500 (7.070-16.950)	6.900 (4.895-10.890)	<0.001
TLG	143.181 (25.241-358.192)	33.120 (8.854-168.031)	0.018
CT Radiomic features			
GLSZM_HGLZE	0.523 (0.353-0.659)	0.314 (0.240-0.445)	<0.001
GLDM_DV	0.390 (0.248-0.501)	0.530 (0.446-0.725)	<0.001
GLSZM_GLNUN	0.286 (0.218-0.379)	0.374 (0.283-0.483)	0.001
GLSZM_ZE	0.737 (0.610-0.849)	0.631 (0.479-0.725)	<0.001
PET Radiomic features			
First-order_Skewness (LHH)	0.561 (0.392-0.764)	0.374 (0.125-0.815)	0.019
First-order_Skewness (LLL)	1.008 (0.653-1.615)	0.773 (0.537-0.982)	<0.001
PET/CT Radiomic Score	0.170 (0.051-0.359)	0.722 (0.388-0.893)	<0.001

Data were expressed as median (interquartile range).

GLSZM, Gray Level Size Zone Matrix; GLDM, Gray Level Dependence Matrix; HGLZE, High Gray Level Zone Emphasis; DV, dependence variance; GLNUN, Gray Level Non Uniformity Normalized; ZE, zone entropy; LHH and LLL are two subtypes of wavelet filters.

Although a significant correlation between the tumor glucose metabolism level captured on PET images and EGFR mutation status has been found in multiple previous studies (11–14), namely lower SUV_{max} in NSCLCs with EGFR mutation than those with wild type EGFR, conventional PET-derived semi-quantitative parameters didn't show enough satisfactory predictive ability to be applied in clinical practice. Consistent with previous studies, SUV_{max} as a single pixel value only showed moderate AUC for distinguishing mutant EGFR from wild type in our study, whereas total lesion glycolysis (TLG) as a volumetric measurement of tumor glucose metabolism showed no higher predictive performance either. Therefore, our present study established a model based on ^{18}FDG PET/CT radiomics to improve the predictive performance for EGFR mutation status in patients with NSCLC.

In our study, four CT radiomics features and two PET features were selected to establish the predictive model with significantly higher AUC than that of SUV_{max} and TLG. Among these selected radiomics features, GLSZM_HGLZE from CT images measures the distribution of the higher gray-level values with a higher value indicating larger high-density areas proportion in tumor, which suggested that the tumors with EGFR+ had lower density than the EGFR- group in our study. In agreement with our finding, more ground-glass opacity and less solid components were observed in lung cancers with EGFR mutation (30) with lower mean CT values when compared to those with wild-type EGFR (31). The remaining 5 radiomics features, including three CT features (GLDM_DV, GLSZM_GLNN, GLSZM_ZE) and two PET features (First-order_Skewness (LHH), First-order_Skewness (LLL)), are all related to image uniformity and heterogeneity. In our study, the EGFR+ group was more heterogeneous on both PET and CT images than the EGFR- group. Our findings were similar to previous studies (21–23). They found that those image texture feature measuring the variability of gray-level intensity or the asymmetry of the distribution of gray-level values were significant predictive of EGFR mutation status. In summary, the NSCLCs with EGFR mutation had lower glucose metabolism and density, with more heterogeneity on both PET and CT images than those with wild-type EGFR. Owing to the bi-modal image features, PET/CT radiomics model in recent studies (0.79 in Zhang J's study (23); 0.80 in Li X's study (22); 0.77 in our study) has showed higher AUC than those generated by PET (0.67 in Yip, SS's study (21)) or CT (0.69 in Rios Velazquez, E's study (19); 0.56-0.75 in Sacconi, B's study (31)) radiomics features alone for predicting EGFR mutation status. However, compared with larger sample size in CT radiomics research, the current sample size in PET/CT radiomics-related studies is generally limited, and thus the generalization ability of PET/CT radiomics-based model remains to be further tested.

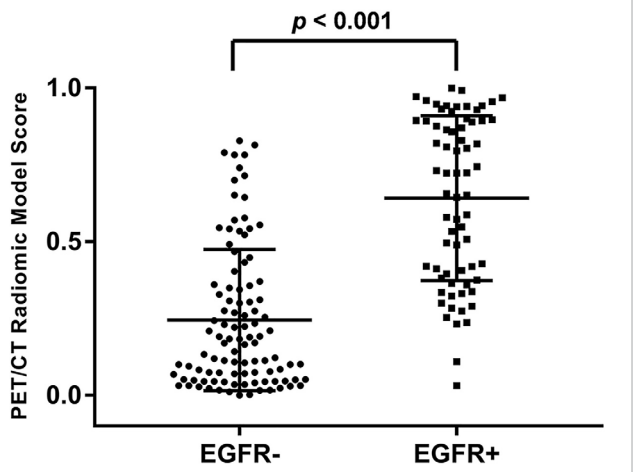


FIGURE 2 | Distribution of PET/CT radiomic model prediction score of all patients. The tumors with EGFR+ had significantly higher score than those with EGFR- ($p < 0.001$).

Clinical features in patients with NSCLC are also non-negligible variables in the evaluation of EGFR mutations, which are more likely to occur in Asians, adenocarcinomas, females, and nonsmokers (32). In our study, gender was only significant clinical predictor of EGFR mutation status. Smoking history was not included in our study due to the complexity of its definition, including the length of history, whether to quit or repeat smoking, etc. This complexity of smoking history made the simple

TABLE 3 | Predictive performance of EGFR mutation status using different models compared with conventional PET parameters and clinical feature.

Model/Parameters	Training set				10-fold cross validation using SVM algorithm			
	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)
Combined Model	0.866	82.60%	81.20%	81.80%	0.827	73.74%	76.07%	75.29%
PET/CT Radiomics Model	0.868	92.80%	66.30%	77.10%	0.769	67.11%	67.03%	67.06%
CT Radiomics Model	0.792	58.00%	87.10%	75.30%	0.754	64.22%	69.87%	67.65%
PET Radiomics Model	0.738	55.10%	82.20%	71.20%	0.750	60.29%	69.69%	67.06%
Gender	0.664	53.60%	79.20%	68.80%	/	/	/	/
SUV_{max}	0.683	84.10%	49.50%	63.50%	/	/	/	/
TLG	0.662	66.70%	64.40%	65.30%	/	/	/	/

SVM, support vector machine.

TABLE 4 | Predictive performance of EGFR mutation subtypes using PET/CT radiomic features compared with conventional PET parameters.

Parameters/Feature	21 L858R mutation(N =38)	19 Del mutation(N=29)	*p* Value	AUC	Sensitivity(%)	Specificity(%)	Accuracy(%)
SUV_{max}	7.5 (5.355-11.650)	6.695 (4.450-10.450)	0.134	/	/	/	/
TLG	37.98 (12.703-180.620)	26.014 (4.529-164.770)	0.408	/	/	/	/
GLCM_DV	1134.093 (801.011-1667.094)	808.42 (433.669-1353.409)	0.016	0.661	75.70%	57.10%	43.10%

Data were expressed as median (interquartile range).
GLCM, Gray Level Co-occurrence Matrix; DV, dependence variance.

classification of yes or no meaningless. Gender as only clinical characteristic was selected in combine model of our study. The addition of clinical characteristics to PET/CT radiomics model, to varying degrees, increase the diagnostic performance of diagnostic model in previous studies (22, 23) and our study, which finally reached 82.6% of diagnostic accuracy in Li X's study (22), 80.0% in Zhang J's study (23), and 75.3% in our study for predicting EGFR mutation status. It suggested that the combined model might be an alternative indicator of EGFR mutations when tissue samples are not available.

The 19 del and 21 L858R mutations are the main two EGFR mutation subtypes. Although both mutation subtypes are sensitive to EGFR-TKI treatment, it is now being recognized that patients with the 19 del mutation experience better clinical outcomes compared to those with the 21 L858R mutation (33, 34). Similar to a previous study investigating a large cohort of Chinese patients (14), we found that SUV_{max} or TLG had no ability to classify the 19 del and the 21 L858R mutation. We tried to investigate the possibility of PET/CT radiomics features for distinguishing these two subtypes. Only GLCM_DV from PET images, which measures the heterogeneity of different intensity-level matrix, showed significant but unsatisfactory predictive performance in our study (AUC=0.661). Liu Q, et al. recent study (35) established a predictive model for EGFR mutation subtypes using machine learning algorithm, which seemed to have better classification performance (AUC=0.77 and 0.92 for respectively predicting exons 19 del and 21 L858R mutations) than ours. However, the number of exons 19 del and 21 L858R mutations was small in Liu Q's study (only 44 and 31 cases respectively), especially when divided as the train and test cohorts, so the generalization ability of the predictive model was not clear.

The limited number of cases was one of the main factors restricting our research to obtain more reliable conclusions. Larger-scale data based on multi-center may be a solution in our next research. However, multi-center study may bring another important issue that affects the generalization ability of the model, that is, the variable PET imaging protocols among multi-centers including image acquisition and reconstruction conditions will be unable to ensure the uniformity and comparability of extracted radiomics features, thus affecting the sensitivity and specificity of radiomics model. Papp L, et al. suggested that larger matrix size/smaller voxel size, point-spread function reconstruction algorithms, and narrow Gaussian post-filtering helped minimize feature variations (36). The variability of PET radiomics is also feature-dependent. GLCM and shape features are the least sensitive to PET imaging system variations (36, 37). Although the single center study maintains the image acquisition and reconstruction methods consistent in all enrolled patients, thus avoiding the influence of the above-described factors as much as possible, the standardization of large databases from multi-centers will remain an unavoidable key step in further research.

In conclusion, EGFR mutations status in patients with NSCLC could be well predicted by the model based on ^{18}F-FDG PET/CT radiomics and clinical features, providing an alternative useful method for the selection of TKI therapy.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/ **Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ruijin Hospital Ethics Committee Shanghai Jiao Tong University School of Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conceptualization: MZ, CS, DQ, and BL. Data curation: JL. Formal analysis: YB, WR, XL, MZ, and XH. Investigation: MZ, YZ, QQ, and HM. Methodology: YB and JX. Project administration: MZh and BL. Supervision: DQ and BL. Writing—original draft: MZ and YB. Writing—review and editing: DQ and BL. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by grants from the National Key Research and Development Program of China (2018YFC0116402), National Natural Science Foundation of China (81974276), Shanghai Jiao Tong University Med-X Interdisciplinary Research Funding (YG2017MS61), Shanghai Pujiang Program (18PJD030) and Shanghai Municipal Key Clinical Specialty (shslczdzk03403).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.568857/full#supplementary-material

REFERENCES

1. Rosell R, Karachaliou N. Lung cancer in 2014: optimizing lung cancer treatment approaches. *Nat Rev Clin Oncol* (2015) 12(2):75–6. doi: 10.1038/nrclinonc.2014.225
2. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. *Adv Exp Med Biol* (2016) 893:1–19. doi: 10.1007/978-3-319-24223-1_1
3. Wu YL, Zhou C, Hu CP, Feng J, Lu S, Huang Y, et al. Afatinib versus cisplatin plus gemcitabine for first-line treatment of Asian patients with advanced non-small-cell lung cancer harbouring EGFR mutations (LUX-Lung 6): an open-label, randomised phase 3 trial. *Lancet Oncol* (2014) 15(2):213–22. doi: 10.1016/S1470-2045(13)70604-1
4. Riely GJ, Pao W, Pham D, Li AR, Rizvi N, Venkatraman ES, et al. Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. *Clin Cancer Res* (2006) 12(3 Pt 1):839–44. doi: 10.1158/1078-0432.CCR-05-1846
5. Yang W, Gao Y, Li X, Zhang J, Liu T, Feng X, et al. Postoperative survival of EGFR-TKI-targeted therapy in non-small cell lung cancer patients with EGFR 19 or 21 mutations: a retrospective study. *World J Surg Oncol* (2017) 15 (1):197. doi: 10.1186/s12957-017-1251-z
6. Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman J, Chirieac LR, et al. Non-Small Cell Lung Cancer, Version 5.2017, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* (2017) 15(4):504–35. doi: 10.6004/jnccn.2017.0050
7. Suda K, Murakami I, Sakai K, Tomizawa K, Mizuuchi H, Sato K, et al. Heterogeneity in resistance mechanisms causes shorter duration of epidermal growth factor receptor kinase inhibitor treatment in lung cancer. *Lung Cancer* (2016) 91:36–40. doi: 10.1016/j.lungcan.2015.11.016
8. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin* (2019) 69(2):127–57. doi: 10.3322/caac.21552
9. Chen L, Zhou Y, Tang X, Yang C, Tian Y, Xie R, et al. EGFR mutation decreases FDG uptake in nonsmall cell lung cancer via the NOX4/ROS/GLUT1 axis. *Int J Oncol* (2019) 54(1):370–80. doi: 10.3892/ijo.2018.4626
10. Takamochi K, Mogushi K, Kawaji H, Imashimizu K, Fukui M, Oh S, et al. Correlation of EGFR or KRAS mutation status with 18F-FDG uptake on PET-CT scan in lung adenocarcinoma. *PloS One* (2017) 12(4):e0175622. doi: 10.1371/journal.pone.0175622
11. Lee EY, Khong PL, Lee VH, Qian W, Yu X, Wong MP. Metabolic phenotype of stage IV lung adenocarcinoma: relationship with epidermal growth factor receptor mutation. *Clin Nucl Med* (2015) 40(3):e190–5. doi: 10.1097/RLU.0000000000000684
12. Cho A, Hur J, Moon YW, Hong SR, Suh YJ, Kim YJ, et al. Correlation between EGFR gene mutation, cytologic tumor markers, 18F-FDG uptake in non-small cell lung cancer. *BMC Cancer* (2016) 16:224. doi: 10.1186/s12885-016-2251-z
13. Yoshida T, Tanaka H, Kuroda H, Shimizu J, Horio Y, Sakao Y, et al. Standardized uptake value on (18)F-FDG-PET/CT is a predictor of EGFR T790M mutation status in patients with acquired resistance to EGFR-TKIs. *Lung Cancer* (2016) 100:14–9. doi: 10.1016/j.lungcan.2016.07.022
14. Lv Z, Fan J, Xu J, Wu F, Huang Q, Guo M, et al. Value of (18)F-FDG PET/CT for predicting EGFR mutations and positive ALK expression in patients with non-small cell lung cancer: a retrospective analysis of 849 Chinese patients. *Eur J Nucl Med Mol Imaging* (2018) 45(5):735–50. doi: 10.1007/s00259-017-3885-z
15. Lee SM, Bae SK, Jung SJ, Kim CK. FDG uptake in non-small cell lung cancer is not an independent predictor of EGFR or KRAS mutation status: a retrospective analysis of 206 patients. *Clin Nucl Med* (2015) 40(12):950–8. doi: 10.1097/RLU.0000000000000975
16. Ko KH, Hsu HH, Huang TW, Gao HW, Shen DH, Chang WC, et al. Value of (1)(8)F-FDG uptake on PET/CT and CEA level to predict epidermal growth factor receptor mutations in pulmonary adenocarcinoma. *Eur J Nucl Med Mol Imaging* (2014) 41(10):1889–97. doi: 10.1007/s00259-014-2802-y
17. Zhang Y, Chang L, Yang Y, Fang W, Guan Y, Wu A, et al. Intratumor heterogeneity comparison among different subtypes of non-small-cell lung cancer through multi-region tissue and matched ctDNA sequencing. *Mol Cancer* (2019) 18(1):7. doi: 10.1186/s12943-019-0939-9
18. Yip SS, Aerts HJ. Applications and limitations of radiomics. *Phys Med Biol* (2016) 61(13):R150–166. doi: 10.1088/0031-9155/61/13/R150
19. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. *Cancer Res* (2017) 77(14):3922–30. doi: 10.1158/0008-5472.CAN-17-0122
20. Wang S, Shi J, Ye Z, Dong D, Yu D, Zhou M, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J* (2019) 53(3):1800986. doi: 10.1183/13993003.00986-2018
21. Yip SS, Kim J, Coroller TP, Parmar C, Velazquez ER, Huynh E, et al. Associations Between Somatic Mutations and Metabolic Imaging Phenotypes in Non-Small Cell Lung Cancer. *J Nucl Med* (2017) 58(4):569–76. doi: 10.2967/jnumed.116.181826
22. Li X, Yin G, Zhang Y, Dai D, Liu J, Chen P, et al. Predictive Power of a Radiomic Signature Based on (18)F-FDG PET/CT Images for EGFR Mutational Status in NSCLC. *Front Oncol* (2019) 9:1062. doi: 10.3389/fonc.2019.01062
23. Zhang J, Zhao X, Zhao Y, Zhang J, Zhang Z, Wang J, et al. Value of pre-therapy (18)F-FDG PET/CT radiomics in predicting EGFR mutation status in patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging* (2019) 47(5):1137–46. doi: 10.1007/s00259-019-04592-1
24. Newton CR, Graham A, Heptinstall LE, Powell SJ, Summers C, Kalsheker N, et al. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res* (1989) 17(7):2503–16. doi: 10.1093/nar/17.7.2503

25. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* (2012) 9(4):357–9. doi: 10.1038/nmeth.1923
26. Chen JS, Huertas A, Medioni G. Fast Convolution with Laplacian-of-Gaussian Masks. *IEEE Trans Pattern Anal Mach Intell* (1987) 9(4):584–90. doi: 10.1109/TPAMI.1987.4767946
27. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77(21):e104–7. doi: 10.1158/0008-5472.CAN-17-0339
28. Zwanenburg A, Vallières M, Abdalah M, 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(2):328–38. doi: 10.1148/radiol.2020191145
29. Menze BH, Kelm BM, Masuch R, Himmelreich U, Bachert P, Petrich W, et al. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinf* (2009) 10:213. doi: 10.1186/1471-2105-10-213
30. Liu Y, Kim J, Qu F, Liu S, Wang H, Balagurunathan Y, et al. CT Features Associated with Epidermal Growth Factor Receptor Mutation Status in Patients with Lung Adenocarcinoma. *Radiology* (2016) 280(1):271–80. doi: 10.1148/radiol.2016151455
31. Sacconi B, Anzidei M, Leonardi A, Boni F, Saba L, Scipione R, et al. Analysis of CT features and quantitative texture analysis in patients with lung adenocarcinoma: a correlation with EGFR mutations and survival rates. *Clin Radiol* (2017) 72(6):443–50. doi: 10.1016/j.crad.2017.01.015
32. Shi Y, Au JS, Thongprasert S, Srinivasan S, Tsai CM, Khoa MT, et al. A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER). *J Thorac Oncol* (2014) 9(2):154–62. doi: 10.1097/JTO.0000000000000033
33. Yu JY, Yu SF, Wang SH, Bai H, Zhao J, An TT, et al. Clinical outcomes of EGFR-TKI treatment and genetic heterogeneity in lung adenocarcinoma patients with EGFR mutations on exons 19 and 21. *Chin J Cancer* (2016) 35:30. doi: 10.1186/s40880-016-0086-2
34. Zhou J, Ben S. Comparison of therapeutic effects of EGFR-tyrosine kinase inhibitors on 19Del and L858R mutations in advanced lung adenocarcinoma and effect on cellular immune function. *Thorac Cancer* (2018) 9(2):228–33. doi: 10.1111/1759-7714.12568
35. Liu Q, Sun D, Li N, Kim J, Feng D, Huang G, et al. Predicting EGFR mutation subtypes in lung adenocarcinoma using (18)F-FDG PET/CT radiomic features. *Transl Lung Cancer Res* (2020) 9(3):549–62. doi: 10.21037/tlcr.2020.04.17
36. Papp L, Rausch I, Grahovac M, Hacker M, Beyer T. Optimized Feature Extraction for Radiomics Analysis of (18)F-FDG PET Imaging. *J Nucl Med* (2019) 60(6):864–72. doi: 10.2967/jnumed.118.217612
37. Tsujikawa T, Rahman T, Yamamoto M, Yamada S, Tsuyoshi H, Kiyono Y, et al. (18)F-FDG PET radiomics approaches: comparing and clustering features in cervical cancer. *Ann Nucl Med* (2017) 31(9):678–85. doi: 10.1007/s12149-017-1199-7

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 08 January 2021
doi: 10.3389/fonc.2020.591485

Evaluating Heterogeneity of Primary Lung Tumor Using Clinical Routine Magnetic Resonance Imaging and a Tumor Heterogeneity Index

Nan Hu [1,2,3†], ShaoHan Yin [1,2,4†], Qiwen Li [1,2,3], Haoqiang He [1,2,4], Linchang Zhong [1,2,4], Nan-Jie Gong [5], Jinyu Guo [1,2,3], Peiqiang Cai [1,2,4], Chuanmiao Xie [1,2,4], Hui Liu [1,2,3] and Bo Qiu [1,2,3*]*

[1] Department of Radiation Oncology,State Key Laboratory of Oncology in South China, Guangzhou, China, [2] Department of Radiation Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China, [3] Department of Radiation Oncology, Guangdong Association Study of Thoracic Oncology, Guangzhou, China, [4] Department of Radiology, Sun Yat-sen University Cancer Center, Guangzhou, China, [5] Vector Lab for Intelligent Medical Imaging and Neural Engineering, International Innovation Center of Tsinghua University, Shanghai, China

OPEN ACCESS

Edited by:
Aditya Juloori, University of Chicago Medical Center, United States

Reviewed by:
Lin Lu, Columbia University Irving Medical Center, United States
Bilgin Kadri Aribas, Bülent Ecevit University, Turkey

***Correspondence:**
Bo Qiu
qiubo@sysucc.org.cn
Hui Liu
liuhuisysucc@sina.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *04 August 2020*
Accepted: *23 November 2020*
Published: *08 January 2021*

Citation:
Hu N, Yin S, Li Q, He H, Zhong L, Gong N-J, Guo J, Cai P, Xie C, Liu H and Qiu B (2021) Evaluating Heterogeneity of Primary Lung Tumor Using Clinical Routine Magnetic Resonance Imaging and a Tumor Heterogeneity Index. Front. Oncol. 10:591485. doi: 10.3389/fonc.2020.591485

Objective: To improve the assessment of primary tumor heterogeneity in magnetic resonance imaging (MRI) of non-small cell lung cancer (NSCLC), we proposed a method using basic measurements from T1- and T2-weighted MRI.

Methods: One hundred and four NSCLC patients with different T stages were studied. Fifty-two patients were analyzed as training group and another 52 as testing group. The ratios of standard deviation (SD)/mean signal value of primary tumor from T1-weighted (T1WI), T1-enhanced (T1C), T2-weighted (T2WI), and T2 fat suppression (T2fs) images were calculated. In the training group, correlation analyses were performed between the ratios and T stages. Then an ordinal regression model was built to generate the tumor heterogeneous index (THI) for evaluating the heterogeneity of tumor. The model was validated in the testing group.

Results: There were 11, 32, 40, and 21 patients with T1, T2, T3, and T4 disease, respectively. In the training group, the median SD/mean on T1WI, T1C, T2WI, and T2fs sequences was 0.11, 0.19, 0.16, and 0.15 respectively. The SD/mean on T1C (p=0.003), T2WI (p=0.000), and T2fs sequences (p=0.002) correlated significantly with T stages. Patients with more advanced T stage showed higher SD/mean on T2-weighted, T2fs, and T1C sequences. The median THI in the training group was 2.15. THI correlated with T stage significantly (p=0.000). In the testing group, THI was also significantly related to T stages (p=0.001). Higher THI had relevance to more advanced T stage.

Conclusions: The proposed ratio measurements and THI based on MRI can serve as functional radiomic markers that correlated with T stages for evaluating heterogeneity of lung tumors.

Keywords: non-small cell lung cancer, MRI, heterogeneity, T stage, radiation therapy

BACKGROUND

Malignant tumors usually consist of sub-clonal cells with different gene mutations, histology and morphology in a single lesion, which is termed as intratumoral heterogeneity (1–3). High level of heterogeneity has been reported to be associated with adverse survival outcomes in multiple cancer types (4, 5). Non-small cell lung cancer (NSCLC) is a highly heterogeneous disease regarding the genetic and phenotypic features (2, 3). The heterogeneity provides the fuel for drug resistance and treatment failure. The assessment of intratumoral heterogeneity helps in treatment decision and survival prediction. In patients with inoperable NSCLC, the diagnosis and treatment usually rely on a small amount of tissue by biopsy, which could not represent the chunk of tumor. Therefore, it's important to develop a noninvasive method to evaluate the full spectrum of heterogeneity for primary lesions in NSCLC.

Intratumoral genetic heterogeneity leads to regional variety in stromal architecture, vascularity, glucose uptake, and water diffusion, which can be identified and quantified by medical imaging. Heterogeneity quantification by imaging has been reported to assist in distinctions on tumor types, grading, and different survival outcomes (6). Studies for lung cancer has primarily focused on computed tomography (CT) and positron emission tomography (PET) images. For example, texture analysis of computed tomography (CT) images in NSCLC have the potential to correlate with tumor hypoxia and angiogenesis (7). Intratumoral metabolic heterogeneity on ^{18}F-FDG PET imaging has been shown to be associated with pathological type, differentiation, T stages, and recurrence in NSCLC (8, 9).

Magnetic resonance images (MRI) provides detailed anatomic information with high spatial contrast. The superior soft tissue resolution and lack of radiation make it a useful imaging modality for radiomic analysis. Advances in MR imaging technique, together with quantitative, and qualitative analysis, have expanded the role of MR imaging in lung cancer. The role of MRI in NSCLC has been investigated in multiple settings, including diagnosis, staging, response prediction and assessment, and postoperative lung function prediction (10–14). Moreover, MRI provides great insights into characterization of tumor heterogeneity (15, 16). Measuring heterogenous vascular features using dynamic contrast-enhanced (DCE) MRI and heterogenous cellular morphology using diffusion-weighted MRI (DWI) could yield important predictive biomarkers in lung cancer (16, 17). On the other hand, the assessment of intratumoral heterogeneity using routine MRI sequences has been less prominently studied in NSCLC (18), partly due to the variations in imaging protocol and acquisition signal. With the introduction of MR simulation and MR guidance into the radiation therapy workflow, the signal intensity analysis based on routine sequences become increasingly important and need further exploration.

We hypothesized that the texture features from MRI may be efficient in evaluating intratumoral heterogeneity. The aim of this study was to propose a method using basic texture measurements from T1- and T2-weighted MRI, which can improve the assessment of primary pulmonary tumor heterogeneity and provide more information on future MRI-guided radiation therapy.

MATERIALS AND METHODS

Patients

From January 2016 to December 2018, 120 histologically diagnosed NSCLC patients with stage I–III disease who underwent radiation therapy in our center were included. Clinical data were collected from each patient including age, sex, histology, and tumor stage. Patients were staged based on the 8th AJCC staging system for lung cancer. All patients had biopsy-approved pathological diagnosis of lung primary lesions. Chest MRI including unenhanced T2-weighted images (T2WI), T2WI with fat suppression, unenhanced/enhanced T1-weighted images (T1WI) had been acquired before radiation therapy. Informed consent was obtained from patients for the use of clinical and imaging data. This study was approved by institutional review board.

MRI Acquisition

All MRI examinations were performed using the same 1.5 Tesla unit (GE Signa HDx 1.5; GE Healthcare, Milwaukee, Wisconsin, USA) with a combined eight channel phased-array surface coil. The following sequences were obtained for each patient: unenhanced T2WI in the coronal and axial planes; unenhanced T2WI with fat suppression in the axial planes; three-dimensional liver acquisition with volume acceleration (3D-LAVA) enhanced-scanning in the axial, sagittal, and coronal planes. The parameters of these sequences are listed in **Table 1**. For contrast enhancement, a 0.1 mmol/kg body weight bolus injection of gadopentetate dimeglumine was administered and the enhanced image was acquired ~25 s after the injection.

Image Processing and Analysis

All MR images were viewed on a picture archiving and communication system workstation monitor (AW4.6; GE Healthcare, Milwaukee, Wisconsin, USA). The largest cross-sectional slice of primary tumor was selected. The region of

TABLE 1 | Parameters for the magnetic resonance sequences.

Sequence botained	Scanning method	TR (ms)	TE (ms)	NEX	ST/spacing (mm)	FOV (cm)	Matrix
T2WI (sagittal plane)	FSE	>1,500	80	2	5/1	25	320×224
T2WI (axial plane)	FSE	≥2,000	85	2	3/1.5	25	320×224
FS T2WI (axial plane)	FSE	≥2,000	85	2	5/1	25	320×224
Enhanced-scanning (axial and coronal planes)	3D-LAVA	3–5	1–2.5	1	4/-2	25	512×224

T2WI, T2-weighted image; TR, repetition time; TE, echo time; NEX, nember of excitations; ST, slice thickness; FOV, field of view; FSE, fast spin-echo; SE, spin-echo; FS, fat suppression; SE-EPI, spin echo planar imaging.

interest (ROI) was manually contoured on T1WI, enhance T1 (T1C), T2WI, and T2 fat suppression (T2fs) images to encompass the entire cross-sectional area of the primary tumor (**Figure 1**). The encompassment of any adjacent normal lung tissue was avoided. Contouring was performed by one radiation oncologist and reviewed by a senior radiation oncologist and a radiologist as well. The mean and standard deviation (SD) of signal intensity of ROI was read from the GE workstation. To determine the intratumoral heterogeneity, the ratios between SD and mean value (SD/mean) were calculated. The higher the ratios value was, the higher the heterogeneity of the primary tumor, and vice versa.

Statistical Analysis

We divided patients into the training and testing groups using propensity score matching (PSM) method with a caliper of 0.1 in a 1:1 ratio, with four covariates including sex, age, histology, and T stage. The SD/mean value from each sequence was presented as median and range, and compared using paired-sample t-test. In the training group, Spearman correlation analyses were performed to test the correlation between the SD/mean value and T stage. Variables with a p value <0.05 were selected into an ordinal logistic regression analysis. The ordinal logistic regression model is an extension of the binary model to the case of more than three outcomes which are naturally ordered. Then we generated a tumor heterogeneity index (THI) defined as the algebraic sum of imaging values in the model multiplied by their coefficients. The THI could be used to calculate the probability of each stage for an individual patient according to ordinal logistic regression equations. External validation of the model was performed in the testing group by calculating THI and the probability of T stage for each patient. Spearman correlation analyses were performed to test the correlation between the calculated THI and T stage. All statistical analyses were performed using SPSS ver. 24.0 software (IBM Corp., Armonk, NY), and differences were considered significant at a p-value < 0.05.

RESULTS

Baseline Characteristics

A total of 104 of the included 120 consecutive patients were divided into the training and testing groups based on the above-mentioned PSM procedure. Patients in the training and testing groups were well matched with respect to age, sex, T stage, and histology (p>0.1) (**Table 2**). The demographic and clinical characteristics of all 104 patients were listed in **Table 2**. The median age was 59 years, ranging from 30 to 82. Eighty patients were male and 24 patients were female. The histology was squamous cell carcinoma in 62 patients and non-squamous cell carcinoma in 42 patients. Eleven, 32, 40, and 21 patients had stage T1, T2, T3, and T4 disease, respectively.

TABLE 2 | The demographic and clinical characteristics of the 104 matched patients.

	Training group (n=52)	Testing group (n=52)	P value
Age (median, range)	58, 34~82	59, 30~78	0.985
Sex			0.642
Male	41	39	
Female	11	13	
Histology			0.424
Squamous	33	29	
Non-squamous	19	23	
T stage			0.779
T1	4	7	
T2	17	15	
T3	21	19	
T4	10	11	

Age, sex, histology, and T stage were well matched between training and testing group.

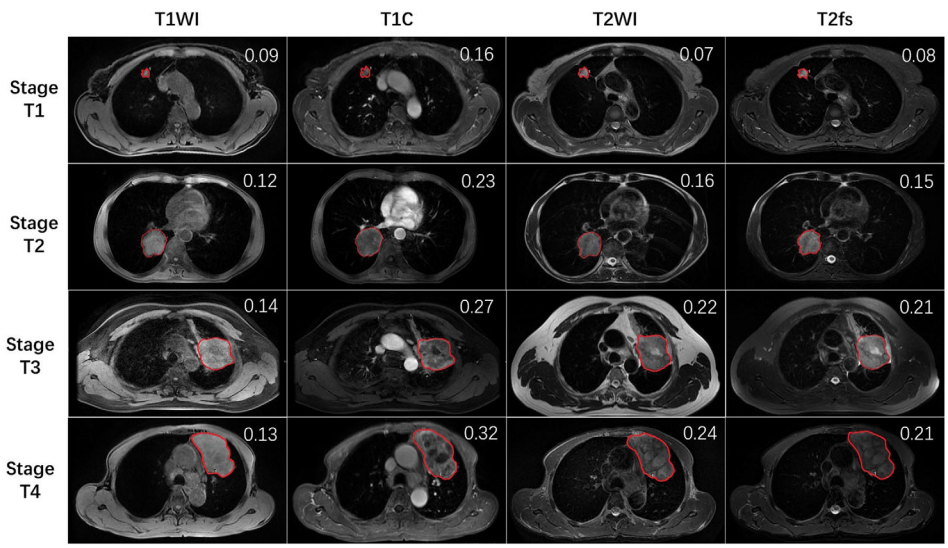


FIGURE 1 | The representative MR images of stage T1–4 patients. Red lines delineate the ROI, which was delineated on the largest cross-sectional slice of primary lung tumor. The SD/mean value was shown on the right upper top of each image.

The Heterogeneity of Primary Tumor on MRI

In the training group, the median SD/mean on T1WI, T1C, T2WI, and T2fs sequences was 0.11, 0.19, 0.16, and 0.15 respectively. The SD/mean was greatest on T1C, while smallest on T1WI. The SD/mean in T1C (p=0.003), T2WI (p=0.000), and T2fs sequences (p=0.002) correlated significantly with T stages. Patients with more advanced T stage showed higher SD/mean on T2WI, T2fs, and T1C sequences (**Figure 2**).

The Development of Regression Model in the Training Group

In the training group, the SD/mean on T1C, T2WI, and T2fs sequences were entered into an ordinal logistic regression model to predict the T stage (**Table 3**). The model fitting information indicated a significance of 0.000. The SD/mean on T2fs (p=0.017) and T1C (p=0.043) were independently predictive of T stage.

Based on the model, a tumor heterogeneous index (THI) that consists of these three variates was developed as the following equation:

$$THI = 7.748 \times SD/mean\_T1C - 11.301 \times SD/mean\_T2 + 14.906 \times SD/mean\_T2fs$$

The median THI in the training group was 2.15 (range 0.81~4.85). THI correlated with T stage significantly (p=0.000).

The probability of any T stage for each patient could be calculated by the ordinal logistic regression rule:

TABLE 3 | Ordinal logistic regression analysis.

	Coefficient	P value	OR
Threshold[T = 1]	−0.453	.604	—
Threshold[T = 2]	1.924	.022	—
Threshold[T = 3]	4.159	.000	—
SD/mean_T2WI	−11.301	.077	0.462
SD/mean_T2fs	14.906	.017	4.428
SD/mean_T1C	7.748	.043	0.744

The SD/mean on T1C, T2WI, and T2fs sequences were entered into an ordinal logistic regression model to predict the T stage. The model fitting information indicated a significance of 0.000.

$$T1 = P(Y \le 1) = 1/(1 + exp[-(-0.453 - TH1)])$$

$$T2 = P(Y \le 2) - P(Y \le 1) = 1/(1 + exp[-(1.924 - TH1)]) - 1/(1 + exp[-(0.453 - TH1)])$$

$$T3 = P(Y \le 3) - P(Y/ \le 2) = 1/(1 + exp[-(4.159 - TH1)]) - 1/(1 + exp[-(1.924 - TH1)])$$

$$T4 = 1 - P(Y \le 3) = 1 - 1/(1 + exp[-(5.140 - TH1)])$$

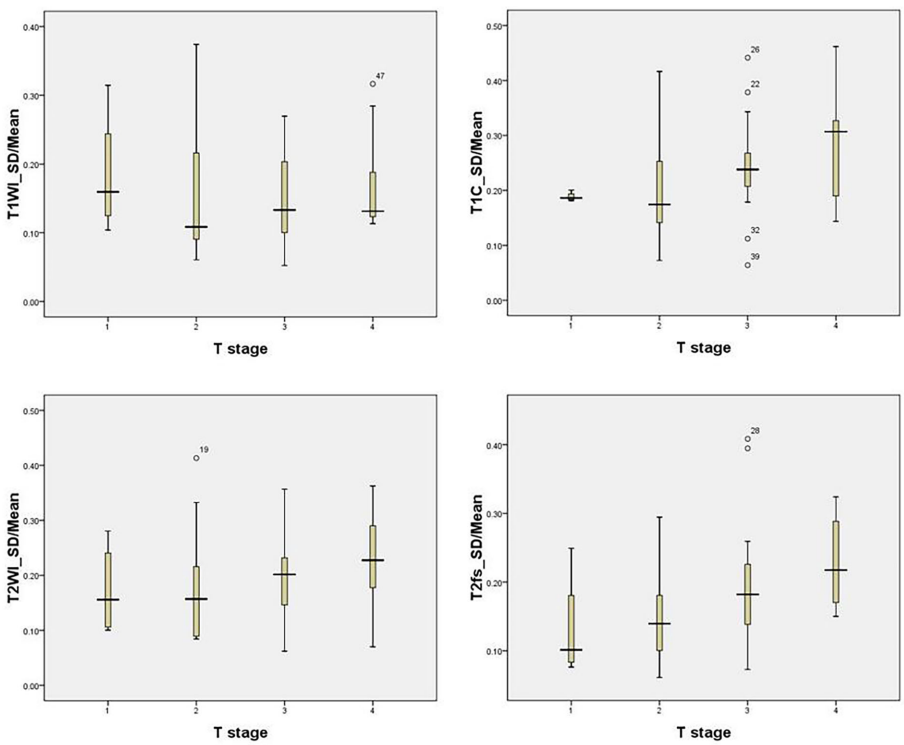


FIGURE 2 | The SD/mean on T1WI, T1C, T2WI, and T2fs grouped by T stage in the training group. Patients with more advanced T stage showed higher SD/mean in T2-weighted, T2fs, and T1C sequences.

Validation of the Model and Tumor Heterogeneity Index in the Testing Group

THI had been calculated for each patient in the testing group according to the above equation. The median THI was 2.06 (range, 0.33~5.06). THI correlated with T stage significantly (p=0.001). The proportion of more advanced stages grew gradually as the THI increased in the testing group (**Figure 3**).

The probability of each stage was calculated for individual patient according to the above equations to generated the estimated T stage. The actual and estimated T stage were listed in **Table 4**. The model predicted T stage accurately in 61.5% (32/52) of patients.

TABLE 4 | The actual and estimated T stages in the testing group.

Actual T stage	Estimated T stage				
	1	**2**	**3**	**4**	**Total**
1	**5**	2	0	0	7
2	3	**10**	2	0	15
3	0	3	**10**	6	19
4	0	1	3	**7**	11

The bold values are the number of cases in which the predicted T stage matches the actual T stage successfully.

DISCUSSION

Intratumoral heterogeneity is an important feature of malignant tumors. There has been considerable effect to use medical imaging to depicts spatial heterogeneity in tumors (6). The advantage of MRI compared with other imaging approach, such as CT and PET, resides in its potential to provide a multi-parameter sequences (T1WI, T2WI, diffusion-weighted, flow-weighted, etc.) at a high spatial resolution. In NSCLC, the signal intensity of primary lesion on MRI is often complexed by fibrous tissue signal, necrotic signal, septations, and vascular void signal. In this study, we quantified the signal heterogeneity on T1WI, T1C, T2WI, and T2fs sequences by measurement of SD/Mean. The signal heterogeneity on T1C, T2WI, and T2fs sequences correlated with T stages significantly. A model had been developed based on the measurements to predict T stages. Then the tumor heterogeneity index (THI) was generated which could be used as a potential radiomic marker for MRI-guided diagnosis and radiotherapy on NSCLC.

Our results showed higher THI was associated with more advanced T stages. It was in accordance with the finding that the metabolic heterogeneity of primary tumor from ^{18}F-FDG PET showed a stepwise increase with the increase of T stages in NSCLC (8). As T stage is a well-established prognostic factor, imaging heterogeneity has been regarded as a potent biomarker for prognosis as well. The intratumor heterogeneity is a distinct predictive factor of response to radiation therapy, mainly due to the resistant subpopulations of cells (19). Non-enhancing tumor fraction assessed by DCE MRI subtraction was found to be a predictor of decrease in tumor volume in response to chemoradiotherapy (20). Texture features derived from MRI was reported to effectively predict tumor response after radiotherapy (18). Taken together, these findings implied that the heterogeneity assessed by MRI reflected the biological behaviors of tumors thus to provide important biological information for tumor diagnosis and treatment in NSCLC.

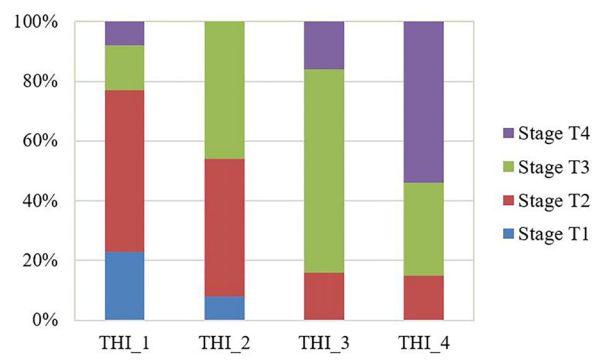


FIGURE 3 | The observed proportion of T stages in the testing group stratified by the tumor heterogeneity index (THI). THI was categorized by quantile to THI_1: 0.33~1.42; THI_2: 1.44~2.04; THI_3: 2.09~2.60; THI_4: 2.67~5.06. The proportion of more advanced stages grew gradually as the THI increased.

Although MRI emerges as an effective method for assessing tumor heterogeneity, the research of optimal sequences or method for quantifying heterogeneity are still processing. At present, studies on heterogeneity assessment by MRI mostly focus on DCE analysis in NSCLC. The semiquantitative perfusion, histogram, and texture parameters from DCE were shown to be prognostic of clinical outcomes (17, 21, 22). However, compared with routine sequences, DCE was not conveniently available in routine practice and relies on post-processing software for data analysis. Therefore, we explored the role of routine sequences in the assessment of intratumoral heterogeneity in NSCLC, which might provide practical information on MRI-guided diagnosis and radiation therapy (MRI simulation and MRI Linac). In our study, tumor on T1C showed the most remarkable signal heterogeneity among the four routine sequences. While unenhanced T1 sequence exhibited the most homogenous imaging. It may be expected that unenhanced images could not well differentiate regions of viable tumor, hypoxic tissues, necrosis or myxoid changes. Contrast-enhanced imaging allows for visualization of more diverse intratumoral components with heterogeneity in enhancement, which can be attributed to heterogeneity of intratumoral perfusion and permeability (23). Therefore, T1C might provide more spatially rich information, and be an important sequence for MRI-based radiomic analysis for intratumoral heterogeneity.

Other than T1C, the heterogeneity measurements from T2WI and T2fs also correlated with T stages. Combining analysis of multiple sequences might provide a more significant marker than single sequence alone. Therefore, we generated a model based on signal heterogeneity from T1C, T2WI, T2fs to predict T stages. By external validation, this model predicted T stages with an accuracy of nearly 62% in the testing group. It is noteworthy that more than 40% of patients were falsely predicted in the testing group. This

indicated that a model derived from merely MRI parameters could not yield a satisfactory performance in predicting T stages. Other factors, such as the location of primary tumor and the presence of atelectasis or obstructive pneumonitis, might also determine the T stages. For instance, a primary tumor of small size that invades mediastinum is staged as T4, while it might present with relatively homogenous appearance and therefore is predicted as T2 according to the model (**Figure 4**). Despite the suboptimal performance, it certainly added value in predictive models incorporating multiple clinical and imaging variables. From this point of view, we generated a novel heterogeneity score from the regression model, term as tumor heterogeneity index (THI), defined as the algebraic sum of the heterogeneity measurements from three sequences multiplied by their coefficients. THI represented the level of heterogeneity, which significantly correlated with T stages in the training and testing groups. Higher THI had relevance to more advanced T stage and possibly worse prognosis. THI could be easily accessed from routine MRI imaging, and therefore conveniently serves as an efficient MRI biomarker.

Besides quantifying the level of heterogeneity, another purpose of heterogeneity analysis is to identify the subregions with different biology that respond differentially to treatment. Kim *et al.* used clustering of PET and DWI to identify highly aggressive subregions in NSCLC. They found the volume of subregion with high aggressiveness was a negative prognostic factor of survival (16). Radiation therapy is the backbone of treatment for locally advanced NSCLC. High-dose radiation was associated with improved local-regional control. However, escalating the radiation dose to the whole tumor volume is limited by normal tissue toxicity. Nowadays, radiation planning using intensity modulated technique allows for different dose distributions inside a tumor volume. Therefore, improvement of local control could be achieved by taking into account intratumoral heterogeneity and delivering higher dose to resistant subregions (24). Kong *et al.* adapted the target volume based on mid-treatment PET and delivered higher-dose radiation to the FDG-avid areas of the tumor, which achieved favorable local-regional control. Compared with PET, MRI could be conveniently acquired before and during the course of radiation. MRI-guided radiotherapy with hybrid MR linear accelerator creates new perspectives towards an individualized planning and treatment approach (25). Therefore, fully depicting the intratumoral heterogeneity on MRI will help identify the resistant subregions and provide evidence for adaptation strategies.

This study has several limitations. Firstly, we quantified the parameters from the largest cross-sectional slice instead of the whole tumor volume. Although tumor heterogeneity measured by the two methods was similar (26), analysis of the whole tumor can theoretically capture more heterogeneous internal components. Secondly, in comparison with CT or PET, scanner and sequence acquisition parameters of MRI have great influence on signal intensity measurements and heterogeneity quantification. Therefore, in order to minimize the influence, we scanned all patients in the same scanner with the same parameters.

CONCLUSIONS

The aim of this study was to propose a method using basic texture measurements from T1- and T2-weighted MRI which can improve the assessment of primary pulmonary tumor heterogeneity and biological behaviors. We found the signal heterogeneity on T1C, T2WI, and T2fs sequences, in terms of SD/mean, correlated positively with T stages. The proposed ratio measurements and THI based on clinical routine MR images can serve as functional radiomic markers that correlated with T stages for evaluating heterogeneity of lung tumors, and provide more information on future MRI-guided radiation therapy. Further studies are warranted to validate the role of THI in response and survival prediction.

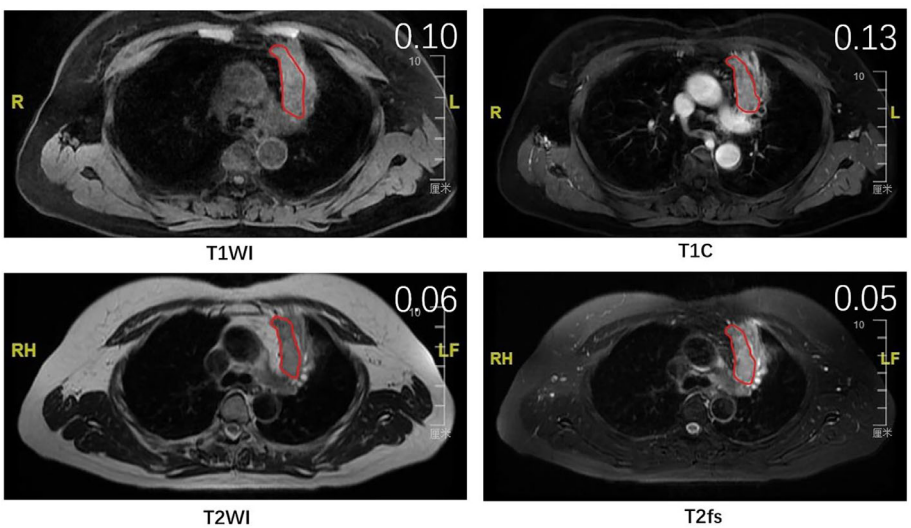


FIGURE 4 | The MR images of a patient with stage T4 that was falsely predicted as stage T2 by the model. Red lines delineate the ROI, which was delineated on the largest cross-sectional slice of primary lung tumor. The SD/mean value was shown on the right upper top of each image.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of cancer center of Sun Yat sen University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

BQ and HL are the main leaders of the project, making major contributions to the project design and article writing. NH is responsible for the main data statistics, image processing, and article writing. QL is responsible for part of the data statistics and image processing work. JG is responsible for data statistics. HH, SY, LZ, NG-J, PC, and CX helped with image processing and analysis. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Sci (New York NY)* (2014) 346:256–9. doi: 10.1126/science.1256930
2. Zito Marino F, Bianco R, Accardo M, Ronchi A, Cozzolino I, Morgillo F, et al. Molecular heterogeneity in lung cancer: From mechanisms of origin to clinical implications. *Int J Med Sci* (2019) 16:981–9. doi: 10.7150/ijms.34739
3. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the evolution of non-small-cell lung cancer. *N Engl J Med* (2017) 376:2109–21. doi: 10.1056/NEJMoa1616288
4. Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, et al. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat Med* (2016) 22:105–13. doi: 10.1038/nm.3984
5. Morris LG, Riaz N, Desrichard A, Senbabaoglu Y, Hakimi AA, Makarov V, et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget* (2016) 7:10051–63. doi: 10.18632/oncotarget.7067
6. O'Connor JP, Rose CJ, Waterton JC, Carano RA, Parker GJ, Jackson A. Imaging intratumor heterogeneity: Role in therapy response, resistance, and clinical outcome. *Clin Cancer Res* (2015) 21:249–57. doi: 10.1158/1078-0432.CCR-14-0990
7. Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA. Non-small cell lung cancer: Histopathologic correlates for texture parameters at ct. *Radiology* (2013) 266:326–36. doi: 10.1148/radiol.12112428
8. Li R, Lin J, Wang L, Zheng X, Tang K. The association between 18f-fluorodeoxyglucose pet intratumoral metabolic heterogeneity and pathological parameters in non-small cell lung cancer. *Nuclear Med Commun* (2019). doi: 10.1097/MNM.0000000000001058
9. Kim DH, Jung JH, Son SH, Kim CY, Hong CM, Oh JR, et al. Prognostic significance of intratumoral metabolic heterogeneity on 18f-fdg pet/ct in pathological n0 non-small cell lung cancer. *Clin nuclear Med* (2015) 40:708–14. doi: 10.1097/RLU.0000000000000867
10. Ciliberto M, Kishida Y, Seki S, Yoshikawa T, Ohno Y. Update of mr imaging for evaluation of lung cancer. *Radiol Clin North Am* (2018) 56:437–69. doi: 10.1016/j.rcl.2018.01.005
11. Koyama H, Ohno Y, Aoyama N, Onishi Y, Matsumoto K, Nogami M, et al. Comparison of stir turbo se imaging and diffusion-weighted imaging of the lung: Capability for detection and subtype classification of pulmonary adenocarcinomas. *Eur Radiol* (2010) 20:790–800. doi: 10.1007/s00330-009-1615-z
12. Matoba M, Tonami H, Kondou T, Yokota H, Higashi K, Toga H, et al. Lung carcinoma: Diffusion-weighted mr imaging–preliminary evaluation with apparent diffusion coefficient. *Radiology* (2007) 243:570–7. doi: 10.1148/radiol.2432060131
13. Ohno Y, Adachi S, Kono M, Kusumoto M, Motoyama A, Sugimura K. Predicting the prognosis of non-small cell lung cancer patient treated with conservative therapy using contrast-enhanced mr imaging. *Eur Radiol* (2000) 10:1770–81. doi: 10.1007/s003300000542
14. Huang YS, Chen JL, Hsu FM, Huang JY, Ko WC, Chen YC, et al. Response assessment of stereotactic body radiation therapy using dynamic contrast-enhanced integrated mr-pet in non-small cell lung cancer patients. *J Magnetic Resonance Imaging: JMRI* (2018) 47:191–9. doi: 10.1002/jmri.25758
15. Just N. Improving tumour heterogeneity mri assessment with histograms. *Br J Cancer* (2014) 111:2205–13. doi: 10.1038/bjc.2014.512
16. Kim J, Ryu SY, Lee SH, Lee HY, Park H. Clustering approach to identify intratumour heterogeneity combining fdg pet and diffusion-weighted mri in lung adenocarcinoma. *Eur Radiol* (2018). doi: 10.1007/s00330-018-5590-0
17. Yoon SH, Park CM, Park SJ, Yoon JH, Hahn S, Goo JM. Tumor heterogeneity in lung cancer: Assessment with dynamic contrast-enhanced mr imaging. *Radiology* (2016) 280:940–8. doi: 10.1148/radiol.2016151367
18. Mahon RN, Hugo GD, Weiss E. Repeatability of texture features derived from magnetic resonance and computed tomography imaging and use in predictive models for non-small cell lung cancer outcome. *Phys Med Biol* (2019). doi: 10.1088/1361-6560/ab18d3
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. Mannelli L, Patterson AJ, Zahra M, Priest AN, Graves MJ, Lomas DJ, et al. Evaluation of nonenhancing tumor fraction assessed by dynamic contrast-enhanced mri subtraction as a predictor of decrease in tumor volume in response to chemoradiotherapy in advanced cervical cancer. *AJR Am J roentgenology* (2010) 195:524–7. doi: 10.2214/AJR.09.3437
21. O'Connor JP, Rose CJ, Jackson A, Watson Y, Cheung S, Maders F, et al. Dce-mri biomarkers of tumour heterogeneity predict crc liver metastasis shrinkage following bevacizumab and folfox-6. *Br J Cancer* (2011) 105:139–45. doi: 10.1038/bjc.2011.191
22. Alic L, van Vliet M, van Dijke CF, Eggermont AM, Veenland JF, Niessen WJ. Heterogeneity in dce-mri parametric maps: A biomarker for treatment response? *Phys Med Biol* (2011) 56:1601–16. doi: 10.1088/0031-9155/56/6/006
23. Wang C, Subashi E, Yin FF, Chang Z. Dynamic fractal signature dissimilarity analysis for therapeutic response assessment using dynamic contrast-enhanced mri. *Med Phys* (2016) 43:1335–47. doi: 10.1118/1.4941739
24. Lambin P, Petit SF, Aerts HJ, van Elmpt WJ, Oberije CJ, Starmans MH, et al. The estro breur lecture 2009. From population to voxel-based radiotherapy: Exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. *Radiother Oncol* (2010) 96:145–52. doi: 10.1016/j.radonc.2010.07.001
25. Corradini S, Alongi F, Andratschke N, Belka C, Boldrini L, Cellini F, et al. Mr-guidance in clinical reality: Current treatment challenges and future perspectives. *Radiat Oncol* (2019) 14(1):92. doi: 10.1186/s13014-019-1308-y
26. Ng F, Kozarski R, Ganeshan B, Goh V. Assessment of tumor heterogeneity by ct texture analysis: Can the largest cross-sectional area be used as an alternative to whole tumor analysis? *Eur J Radiol* (2013) 82:342–8. doi: 10.1016/j.ejrad.2012.10.023

Conflict of Interest: Author NG-J was employed by the company Shanghai United Imaging Healthcare Co., Ltd.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 14 January 2021
doi: 10.3389/fonc.2020.608598

Machine Learning for Histologic Subtype Classification of Non-Small Cell Lung Cancer: A Retrospective Multicenter Radiomics Study

Fengchang Yang[1†], Wei Chen[2†], Haifeng Wei[3], Xianru Zhang[4], Shuanghu Yuan[5], Xu Qiao[4*] and Yen-Wei Chen[6,7]*

[1] Department of Radiology, Shandong Cancer Hospital and Institute, Cheeloo College of Medicine, Shandong University, Jinan, China, [2] Department of Implantology, School and Hospital of Stomatology, Cheeloo College of Medicine, Shandong University, Jinan, China, [3] First Clinical Medical College, Shandong University of Traditional Chinese Medicine, Jinan, China, [4] School of Control Science and Engineering, Shandong University, Jinan, China, [5] Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China, [6] Graduate School of Information Science and Engineering, Ritsumeikan University, Shiga, Japan, [7] Research Center for Healthcare Data Science, Zhejiang Lab, Hangzhou, China

OPEN ACCESS

Edited by:
Youyong Kong,
Southeast University, China

Reviewed by:
Jiandong Yin,
ShengJing Hospital of China Medical University, China
Yang Chen,
Southeast University, China

***Correspondence:**
Xu Qiao
qiaoxu@sdu.edu.cn
Shuanghu Yuan
yuanshuanghu@sina.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *21 September 2020*
Accepted: *30 November 2020*
Published: *14 January 2021*

Citation:
Yang F, Chen W, Wei H, Zhang X, Yuan S, Qiao X and Chen Y-W (2021) Machine Learning for Histologic Subtype Classification of Non-Small Cell Lung Cancer: A Retrospective Multicenter Radiomics Study. Front. Oncol. 10:608598. doi: 10.3389/fonc.2020.608598

Background: Histologic phenotype identification of Non-Small Cell Lung Cancer (NSCLC) is essential for treatment planning and prognostic prediction. The prediction model based on radiomics analysis has the potential to quantify tumor phenotypic characteristics non-invasively. However, most existing studies focus on relatively small datasets, which limits the performance and potential clinical applicability of their constructed models.

Methods: To fully explore the impact of different datasets on radiomics studies related to the classification of histological subtypes of NSCLC, we retrospectively collected three datasets from multi-centers and then performed extensive analysis. Each of the three datasets was used as the training dataset separately to build a model and was validated on the remaining two datasets. A model was then developed by merging all the datasets into a large dataset, which was randomly split into a training dataset and a testing dataset. For each model, a total of 788 radiomic features were extracted from the segmented tumor volumes. Then three widely used features selection methods, including minimum Redundancy Maximum Relevance Feature Selection (mRMR), Sequential Forward Selection (SFS), and Least Absolute Shrinkage and Selection Operator (LASSO) were used to select the most important features. Finally, three classification methods, including Logistics Regression (LR), Support Vector Machines (SVM), and Random Forest (RF) were independently evaluated on the selected features to investigate the prediction ability of the radiomics models.

Results: When using a single dataset for modeling, the results on the testing set were poor, with AUC values ranging from 0.54 to 0.64. When the merged dataset was used for modeling, the average AUC value in the testing set was 0.78, showing relatively good predictive performance.

Conclusions: Models based on radiomics analysis have the potential to classify NSCLC subtypes, but their generalization capabilities should be carefully considered.

Keywords: non-small cell lung cancer, radiomics, machine learning, feature selection, classification

INTRODUCTION

Lung cancer is the leading cause of cancer death in many countries (1, 2). Especially in China, lung cancer is the most common cancer with more than 430,000 deaths per year (3). According to the characteristics of cancer cells under the microscope, lung cancer is broadly classified into two types: small cell lung cancer (SCLC) and non-small cell lung cancers (NSCLC). NSCLC is the most common type of lung cancer, accounting for about 80% of all lung cancers. Squamous cell carcinoma (SCC) and adenocarcinoma (ADC) are the most common histological subtypes of NSCLC. The classification criteria is based on the histologic features, i.e., ADC appears as carcinoma of acinar/tubular structure or mucin production, while SCC appears as carcinoma with keratinization or intercellular bridges (4).

Since the treatment methods of SCC and ADC are quite different, it is of great significance to accurately distinguish SCC from ADC in patients with lung cancer (5, 6). For instance, pemetrexed (a multiple-enzyme inhibitor) has greater efficacy in ADC patients than in SCC patients (7). Pathological diagnosis is commonly regarded as the gold standard for distinguishing ADC from SCC. However, it is invasive and requires needle biopsy or surgery. The tumor may be heterogeneous, which may lead to sampling errors and affect biopsy results. In addition, the risk of complications is also an important factor that must be considered. These limitations of pathological diagnosis prompt us to develop non-invasive and accurate alternative methods.

Radiomics refers to extracting high-throughput features from medical images and mining the potential relationships between quantitative image features and pathophysiology characteristics. Radiomics analysis can be used for predicting various clinical outcomes, such as survival, distant metastases, and molecular characteristics classification (8–10). Several studies have focused on the identification of histologic subtype of NSCLC based on radiomics. Wu et al. (11) constructed two study cohorts with 350 patients and extracted 440 radiomic features for each sample. They applied 24 feature selection methods and 3 classification methods to identify SCC and ADC, of which the Naive Bayes method achieved the highest AUC of 0.72. Zhu et al. (12) retrospectively studied 129 patients with NSCLC and extracted 485 features from manually labeled tumor regions. Five features were selected to construct a radiomics signature by using a logistic regression method. This radiomic signature achieved an AUC of 0.893 in the test set. Chaunzwa et al. (13) retrospectively studied 157 patients with NSCLC to classify ADC or SCC. They used a VGG-16 neural network to extract deep features from CT images and classify them with fully connected layers. Besides, they also independently evaluated the extracted features using three machine learning classification models. The results showed that all models were able to classify tumor histology, of which the neural network achieved the highest performance with an AUC of 0.751.

Although these studies have achieved excellent results, there are still some critical problems that need to be solved: 1) Many radiomics studies generally have small size datasets, thus limiting the performance and the potential clinical applicability of these models. 2) The research methods are relatively simple, and there are few methods of feature selection and classifier for comparison. The differences between different research methods are not fully discussed, which reduces the credibility of the model and limits the application of the models in the clinic.

In order to solve the above problems, we collected three datasets from different centers. Each dataset was used as the training set to build a model and tested in the remaining two datasets. Then, we combined all the datasets into one large dataset to build a model; this dataset was randomly divided into a training set and a testing set. For each dataset, a total of 788 radiomic features were extracted from the segmented tumor volumes of corresponding CT images. Three widely used features selection methods, minimum Redundancy Maximum Relevance Feature Selection (mRMR), Sequential Forward Selection (SFS), Least Absolute Shrinkage and Selection Operator (Lasso) were used to select the most important features. Three widely used classification models were independently evaluated on the selected features: Logistics Regression (LR), Support Vector Machines (SVM), and Random Forest (RF). We aim to build models through multi-center datasets to thoroughly study the potential of radiomics in identifying SCC and ADC.

MATERIALS AND METHODS

Figure 1 presents the workflow of this study, including image acquisition and segmentation, feature extraction, feature selection, classifier construction and evaluation. In the following sections, we will describe these steps in detail.

Datasets

We utilized 3 independent datasets in this study that were collected from China institution and open-access online repositories.

Dataset 1

This retrospective study has been approved by our institutional review board and does not require patient informed consent. From June 2014 to June 2019, 324 patients with a diagnosis of lung cancer were retrospectively collected. The inclusion criteria are as follows: (1) pathologically confirmed lung cancer; (2) CT images can be obtained before treatment. Exclusion criteria were as follows: (1) small cell lung cancer (n = 3); (3) grade of

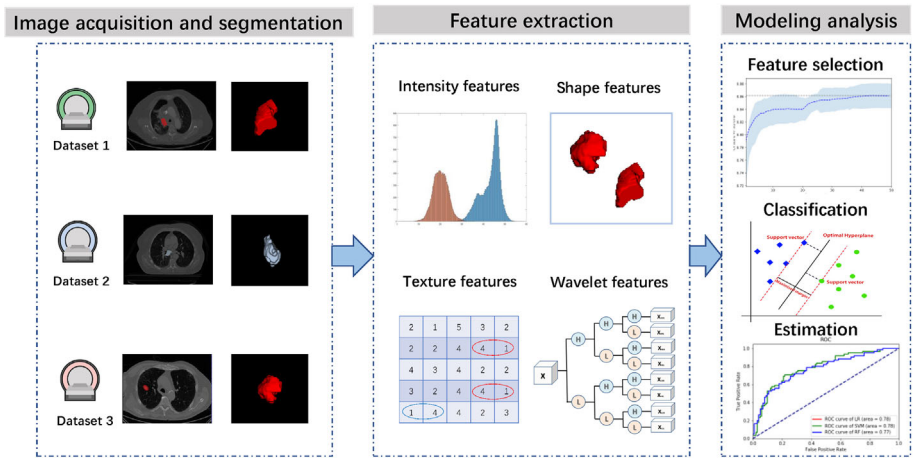


FIGURE 1 | Workflow of this study.

preoperative biopsy evidence was not available (n = 17). Finally, 302 patients were selected for this study. Tumors were classified into ADC or SCC based on pathological information. All pulmonary CT examinations were performed using four CT scanners, with tube voltage of 120 kVp, tube current of 220 mAs, and inter-layer slice thickness of 4–5 mm. For each patient, the tumor region was contoured in a slice-by-slice manner on CT images by an experienced radiologist (with eight years of experience) using Medical Imaging Interaction Toolkit (MITK) software (14) (version 2013.12.0; http://www.mitk.org/), and then confirmed by another experienced radiologist (with 15 years of experience). The final consensus was reached by group discussion if there were discrepant interpretations.

Dataset 2

This dataset was obtained from The Cancer Imaging Archive (TCIA) (15) and included 422 patients with NSCLC treated at Maastricht University Medical Center (16). All patients underwent a CT scan. Depending on the patient'sbody type, the scanning scheme was slightly different. The tube voltage was 120–140 kVp and the tube current was 40–80 mAs. The reconstructed pixel size was 0.977×977mm, the matrix size was 512×512, and the layer thickness between slices was 3 mm. For all CT images, the doctor performed manual tumor region segmentation. From all the samples, 203 samples that meet the requirements of this study were selected.

Dataset 3

This dataset was obtained from TCIA (15) and included 211 NSCLC patients (17). This is a retrospectively collected dataset through different CT equipments and different imaging parameters, with tube voltage of 80–140 kVp, tube current of 124–699 mAs, and inter-layer slice thickness of 0.625–3 mm. For all CT images, an undisclosed automatic segmentation algorithm was used for segmentation and then manually adjusted by the doctors. From all the samples, 140 samples that meet the requirements of this study were selected.

Some slices from the above datasets are displayed in **Figure 2** to show the variety in cancer locations, shapes, and appearances of the different datasets.

Feature Extraction and Selection

Before feature extraction, we resampled all the CT images to a $1\times1\times1\ mm^3$ voxel size. According to the radiomic features described by Imaging Biomarker Standardization Initiative (IBSI), a wide range of features including intensity features, shape features, texture features, and wavelet features were extracted from the segmented cancer regions (18).

Intensity features use first-order statistics (energy, entropy, standard deviation, skewness, kurtosis, etc.) to quantify the tumor intensity feature, which are calculated from the histogram of all tumor voxel intensity values. Shape features describe the shape of the tumor, such as sphericity or compactness of the tumor. Texture features can quantify intratumor heterogeneity differences in the texture that is observable within the tumor volume. These features are calculated in all three-dimensional directions within the tumor volume, thereby taking the spatial location of each voxel compared with the surrounding voxels into account. Texture features quantify the intratumor heterogeneity by using the Gray Level Cooccurrence (GLCM), Gray Level Run Length Matrices (GLRLM), Gray Level Size Zone Matrix (GLSZM), Neighbouring Gray Tone Difference Matrix (NGTDM) and Gray Level Dependence Matrix (GLDM). Wavelet features calculate the intensity and textural features from wavelet decompositions of the original image, thereby focusing the features on different frequency ranges within the tumor volume. All feature extraction algorithms were implemented in Pyradiomics toolkit (19).

To eliminate the differences in the value scales of the radiomic features, feature normalization was performed before feature selection. For features in the training group, each feature for a specific patient was subtracted by the mean value and divided by standard deviation value from this group. The same normalization method was applied to features in the validation group using the mean values and standard deviation values calculated based on the training group.

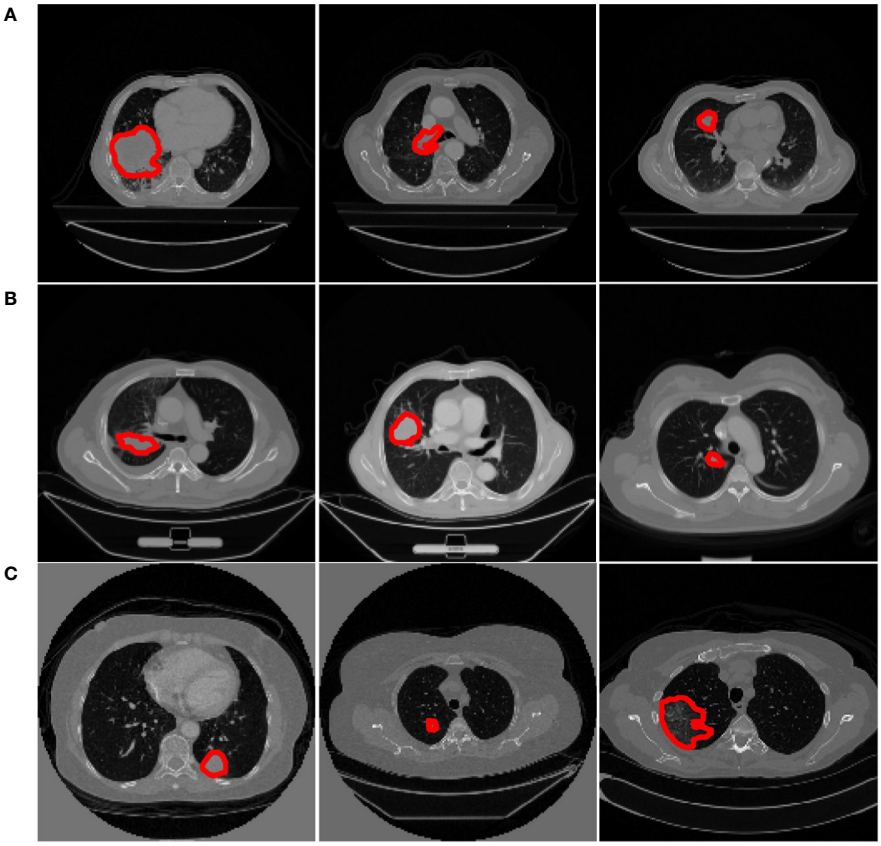


FIGURE 2 | Examples from different datasets. Each row represents three axial slices of different datasets; **(A)** Dataset 1. **(B)** Dataset 2. **(C)** Dataset 3. A red contour that outlined by the physician is displayed to identify the cancer area in each patient scan.

Too many features will increase the computational cost, and the redundancy between features will reduce the accuracy of the classification. Furthermore, the number of features is more than the number of samples in this work, which will increase the probability of overfitting. Therefore, feature selection is essential. There are three main types of feature selection algorithms: filter methods, wrapper methods, and embedded methods. Based on previous works, we selected three widely used feature selection methods, namely: minimum redundancy maximum correlation method (mRMR) (20), sequential forward selection method (SFS) (21), and least absolute shrinkage and selection operator (Lasso) (22).

mRMR is a multivariate filtering feature selection algorithm, which finds the optimal subset of features by considering both the importance of features and the correlation between them, that is, maximizing the correlation between features and categorical variables, while minimizing the redundancy between features. In the set S with N features, the correlation D of the features is calculated as follows:

$$D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c) \tag{1}$$

Where I represents mutual information and the redundancy between features is expressed as:

$$R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j) \tag{2}$$

The goal of mRMR is to find the feature subset where $D - R$ takes the maximum value. SFS is a kind of wrapper method that uses a bottom-up search strategy that starts from an empty feature set and gradually adds features selected by evaluation function. At each iteration, the feature to be added is selected from the remaining available features that have not been added to the feature set. Then, the final selected features should produce the best classification performance compared with the any other feature set (23). Lasso is a kind of embedded method that is widely used in radiomic feature selection of high dimensional data with relatively small sample size. It is based on ℓ_1-norm of the coefficient of a linear classifier. Some of the coefficients of the learned classifier may equal zero. Since each coefficient is associated with a feature, so feature selection is achieved by retaining features with non-zero coefficients.

Classifiers Construction

We evaluated three classification algorithms: logistic regression (LR), support vector machine (SVM), and random forest (RF). LR is a classical machine learning algorithm that was usually used for binary classification tasks. The model attempts to

estimate the probability $p(y = 1|x)$, that is, the probability of a positive result $y = 1$ under the given data x. The advantage of logistic regression is that the training speed is fast, and discrete and continuous variables can be used as inputs. The disadvantage is that it is a linear model, and the classification effect is not good enough in the face of complex data problems, but the logistic regression model can achieve good results on many datasets, and it is easy to implement and can be used as a basic modeling method (24). SVM is another widely used classification algorithm that attempts to separate the data by computing the decision boundary. This decision boundary, also called the hyperplane, is orientated in such a way that it is as far away as possible from the closest data points (support vectors) from each class (25). SVM is a powerful method for obtaining good classification results by using only a few data points (26). RF is an ensemble learning method, known for its high performance and generalizability. It uses bootstrap resampling to extract multiple samples from the original sample, and constructs a decision tree for each bootstrap sample, and then combines these decision trees together to obtain the final classification (27).

Statistical Analysis

The statistical analyses were performed with R 3.1.2 (http://www.R-project.org) and Python (version 3.6.4) machine learning library scikit-learn (version 0.19.1). Univariate analysis for clinical data was performed by using the Chi-square (χ^2) test or Mann-Whitney U test, as appropriate. The categorical variable (such as gender and category probability) was analyzed using the χ^2 test. The continuous variable (such as age) was analyzed using the Mann-Whitney U test.

RESULTS

Patients Statistics

Table 1 lists the clinical data of the patients in the three datasets. The percentages of SCC in dataset 1, 2, and 3 were 29, 75, and 20%, respectively. Among them, the category probability between dataset 1 and dataset 2 is statistically different ($P < 0.001$, χ^2 test); the category probability between dataset 1 and dataset 3 is statistically different ($P = 0.04959$, χ^2 test); the category probability between dataset 2 and dataset 3 is statistically different ($P < 0.001$, χ^2 test).

The percentages of males in dataset 1, 2, and 3 are 62, 71, and 74%, respectively. Among them, the probability of gender in dataset 1 is statistically significant difference ($P < 0.001$, χ^2 test); the probability of gender in dataset 2 is not statistically significant difference ($P = 0.01365$, χ^2 test); the probability of gender in dataset 3 is not statistically significant difference ($P = 0.01219$, χ^2 test).

There is a statistically significant difference between the age of SCC and ADC in dataset 1 ($P < 0.001$, Mann-Whitney U test); there is no statistical difference between the age of SCC and ADC in dataset 2 ($P = 0.06693$, Mann -Whitney U test); there is no statistical difference between the ages of SCC and ADC in dataset 3 ($P = 0.1501$, Mann-Whitney U test). As can be seen from the above results, there exist significant differences between different datasets.

Feature Extraction and Selection Results

A large number of features were extracted from the tumor volume, where each sample contains 788 features. To select the most distinguishing feature subset, we applied three widely used feature selection methods. For each method, we applied Grid search and 5-fold cross-validation to select the best hyper-parameters. The feature selection process is shown in **Figure 3**, where (I)–(IV) represent the feature selection process based on dataset 1, dataset 2, dataset 3, and merged dataset, respectively. The blue dashed line represents the average AUC value of the 5-fold cross-validation, and the shading represents the standard deviation. For mRMR and SFS, the hyperparameter is the number of features, and for Lasso, the hyperparameter is the regularization parameter α. Using the AUC value as the criterion for hyperparameter selection, it can be seen from the figure that for models 1–3, the AUC value is relatively low, and the standard deviation is larger, the Lasso feature selection method shows the best stability.

Table 2 lists the number of features selected by the three feature selection methods in the four training sets and the jointly selected features. For the model trained in dataset 1, mRMR, SFS, and Lasso selected 24, 28, and 28 features, respectively, and 1 feature was jointly selected. It is wavelet-LLH firstorder_Skewness, which measures the asymmetry of value distribution of the tumor area under wavelet transform. For the model trained in dataset 2, mRMR, SFS, and Lasso selected 17, 13, and 9 features, respectively, and there were no features that were jointly selected. For the model trained in dataset 3, mRMR, SFS, and Lasso selected 17, 24, and 7 features, respectively, and there were no features that were jointly selected. For the model trained in the merged dataset,

TABLE 1 | Patients statistics.

	Dataset 1		Dataset 2		Dataset 3	
	ADC (n = 215)	SCC (n = 87)	ADC (n = 51)	SCC (n = 152)	ADC (n = 112)	SCC (n = 28)
age						
range (median)	36–89 (59)	32–81 (65)	45–85 (68)	33–88 (70)	43–87 (68)	57–83 (71)
mead+std	59±10	64±9	67±9	70±10	68±9	71±6
sex						
male	108	79	32	112	80	24
female	107	8	19	40	32	4

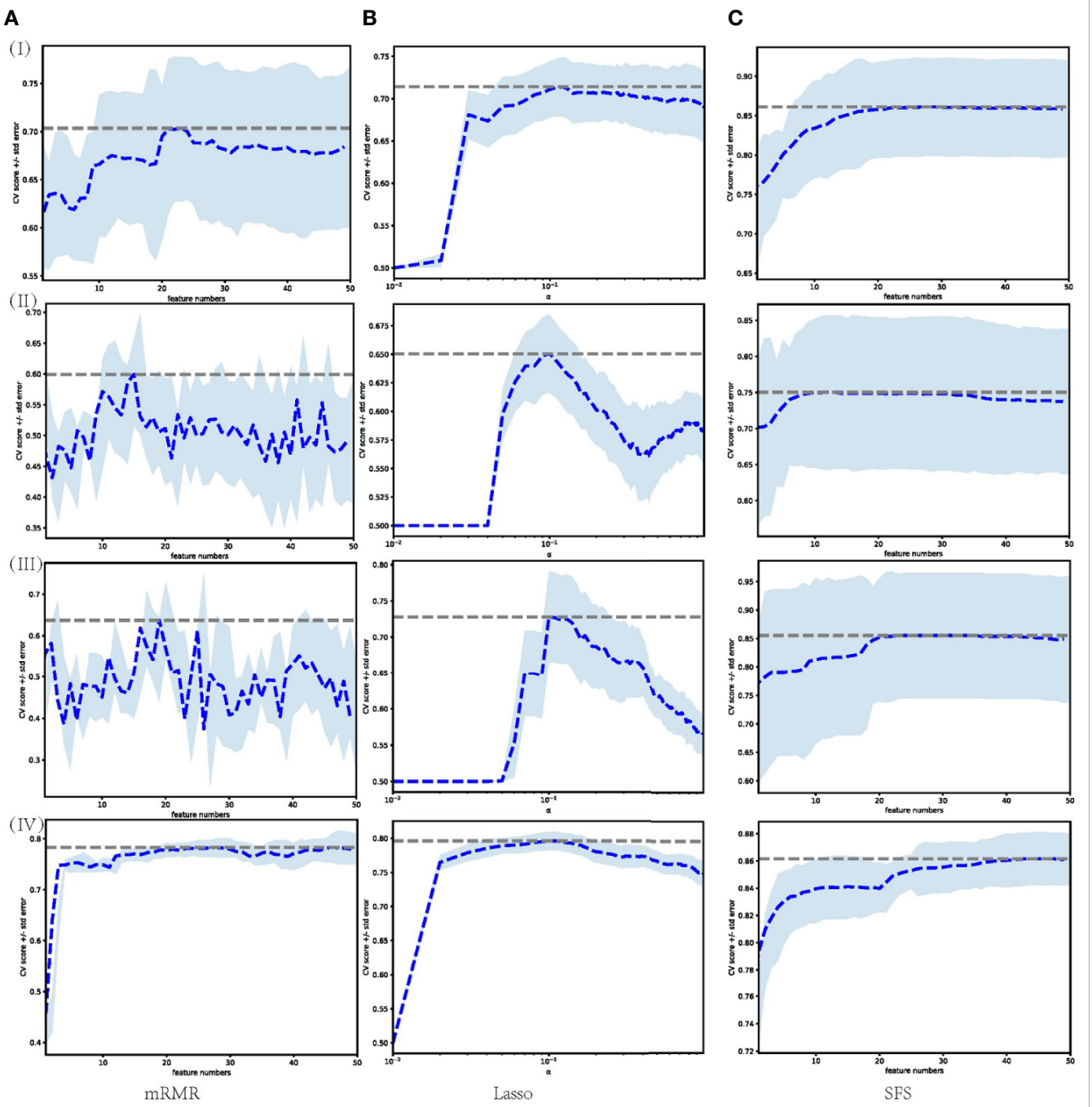


FIGURE 3 | Feature selection process performed on different datasets through different feature selection methods. Each row represents different datasets; (I) Dataset 1. (II) Dataset 2. (III) Dataset 3. (IV) Merged dataset. Each column represents feature selection methods; **(A)** mRMR. **(B)** Lasso. **(C)** SFS.

mRMR, SFS, and Lasso selected 29, 46, and 29 features, respectively, and 1 feature was jointly selected, which is wavelet-HHL_glcm_ClusterShade, that measures the skewness of GLCM features under wavelet transform. A higher ClusterShade indicates greater asymmetry. According to the above results, it can be seen that different feature selection methods select different features, and there is a great inconsistency between the selected features.

TABLE 2 | Commonly selected features.

	mRMR	SFS	Lasso	Common features
dataset 1	24	28	28	wavelet-LLH_firstorder_Skewness
dataset 2	17	13	9	None
dataset 3	17	24	7	None
merged dataset	29	46	29	wavelet-HHL_glcm_ClusterShade

Classification Result

It can be seen from the previous section that different feature selection methods selected distinct feature subsets. To evaluate these feature subsets, we used three classifiers for modeling and analysis. Taking the SVM method as an example, **Figures 4–6** show the ROC curves of different datasets that obtained by three feature selection methods. The left column of each figure shows the 5-fold cross-validated ROC curves of the training dataset, and the right column shows the ROC curves of the testing dataset.

When dataset 1 was used to build the model and mRMR method was used to select features, the average AUC value of 5-fold cross validation was 0.71, and the AUC value in dataset 2 and 3 were 0.56 and 0.59, respectively. When SFS feature selection method was used, the average AUC value of 5-fold cross validation was 0.87, and the AUC value in dataset 2 and dataset 3 were 0.56 and 0.56, respectively. When Lasso feature selection method was used, the average AUC value of 5-fold cross validation was 0.79, and the AUC values in dataset 2 and dataset 3 were 0.61 and 0.64, respectively.

When dataset 2 was used to build the model and mRMR method was used to select features, the average AUC value of 5-fold cross-validation was 0.54, and the AUC values in dataset 1 and dataset 3 were 0.54 and 0.55, respectively. When SFS feature

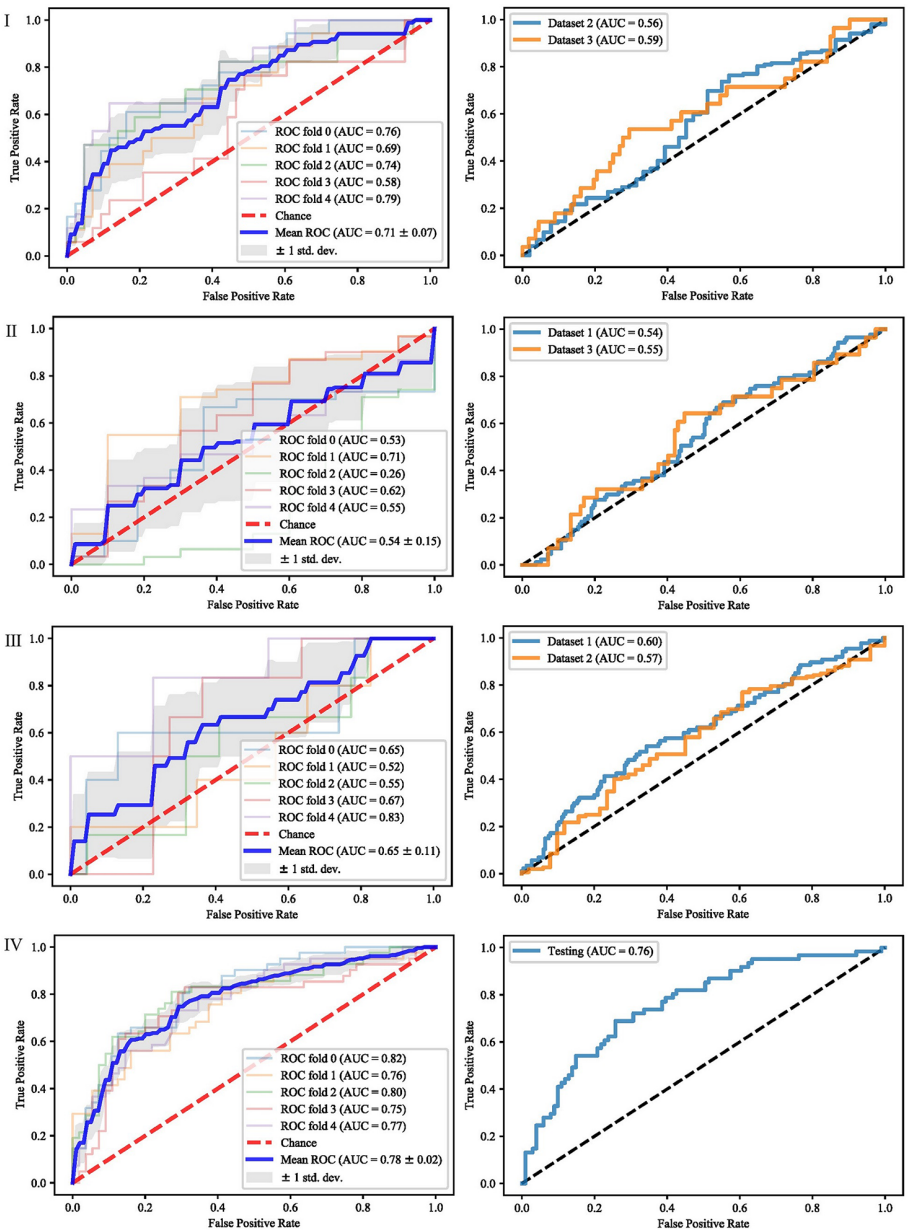


FIGURE 4 | ROCs of different datasets achieved through mRMR and SVM. Each row represents different datasets; (I) Dataset 1. (II) Dataset 2. (III) Dataset 3. (IV) Merged dataset.

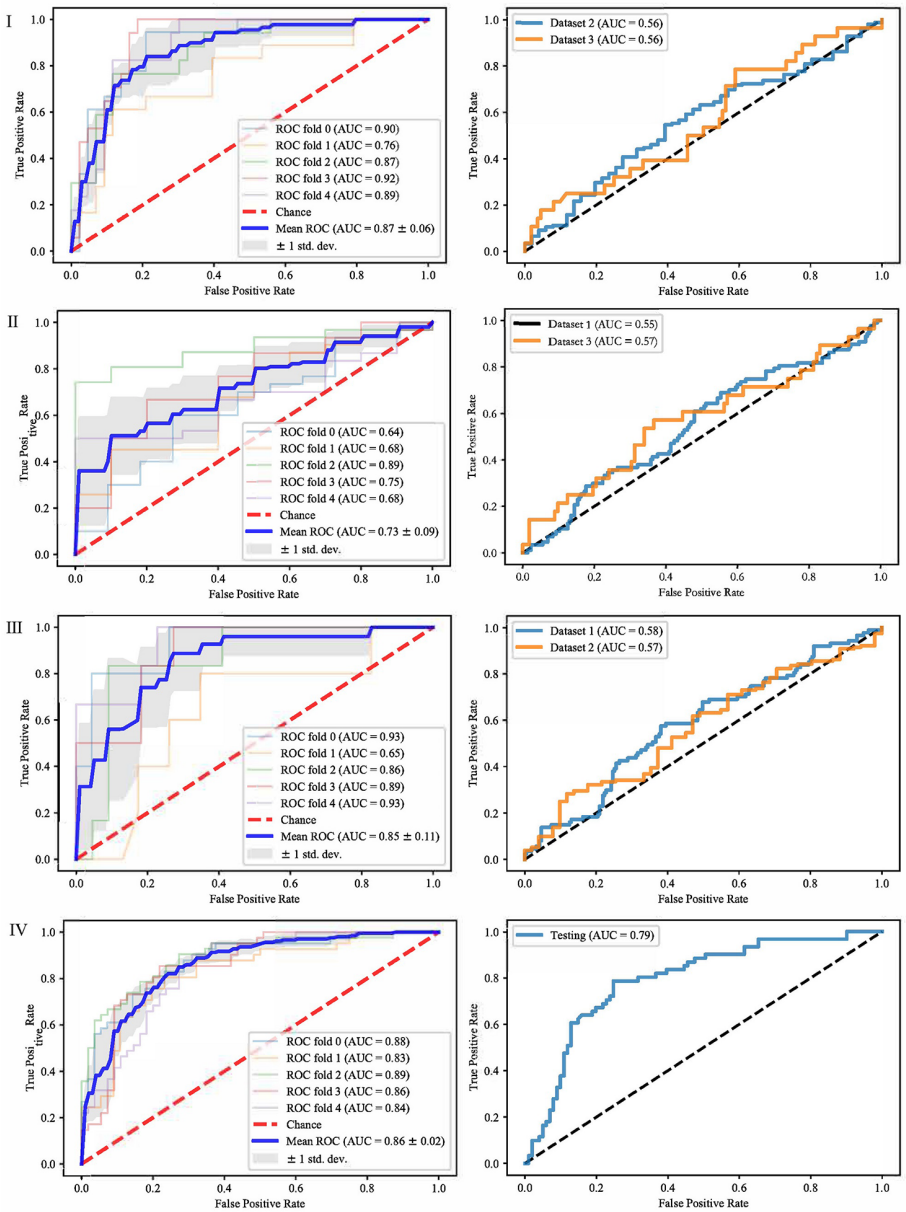


FIGURE 5 | ROCs of different datasets achieved through SFS and SVM. Each row represents different datasets; (I) Dataset 1. (II) Dataset 2. (III) Dataset 3. (IV) Merged dataset.

selection method was used, the average AUC value of 5-fold cross-validation was 0.73, and the AUC values in dataset 1 and dataset 3 were 0.55 and 0.57, respectively. When Lasso feature selection method was used, the average AUC value of 5-fold cross-validation was 0.70, and the AUC values in dataset 1 and dataset 3 were 0.64 and 0.57, respectively.

When dataset 3 was used to build the model and mRMR method was used to select features, the average AUC value of 5-fold cross-validation was 0.65, and the AUC values in dataset 1 and dataset 3 were 0.60 and 0.57, respectively. When SFS feature selection method was used, the average AUC value of 5-fold cross-validation was 0.85, and the AUC values in dataset 1 and dataset 3 were 0.58 and 0.57, respectively. When the Lasso feature selection method was used, the average AUC value of 5-fold cross-validation was 0.75, and the AUC values in dataset 1 and dataset 3 were 0.60 and 0.57, respectively.

When the merged dataset was used to build the model and mRMR method was used to select features, the average AUC value of 5-fold cross-validation was 0.78, and the AUC value in the testing set was 0.76. When the SFS feature selection method was used, the average AUC of 5-fold cross-validation was 0.86, and the AUC value in the test set was 0.79. When the Lasso feature

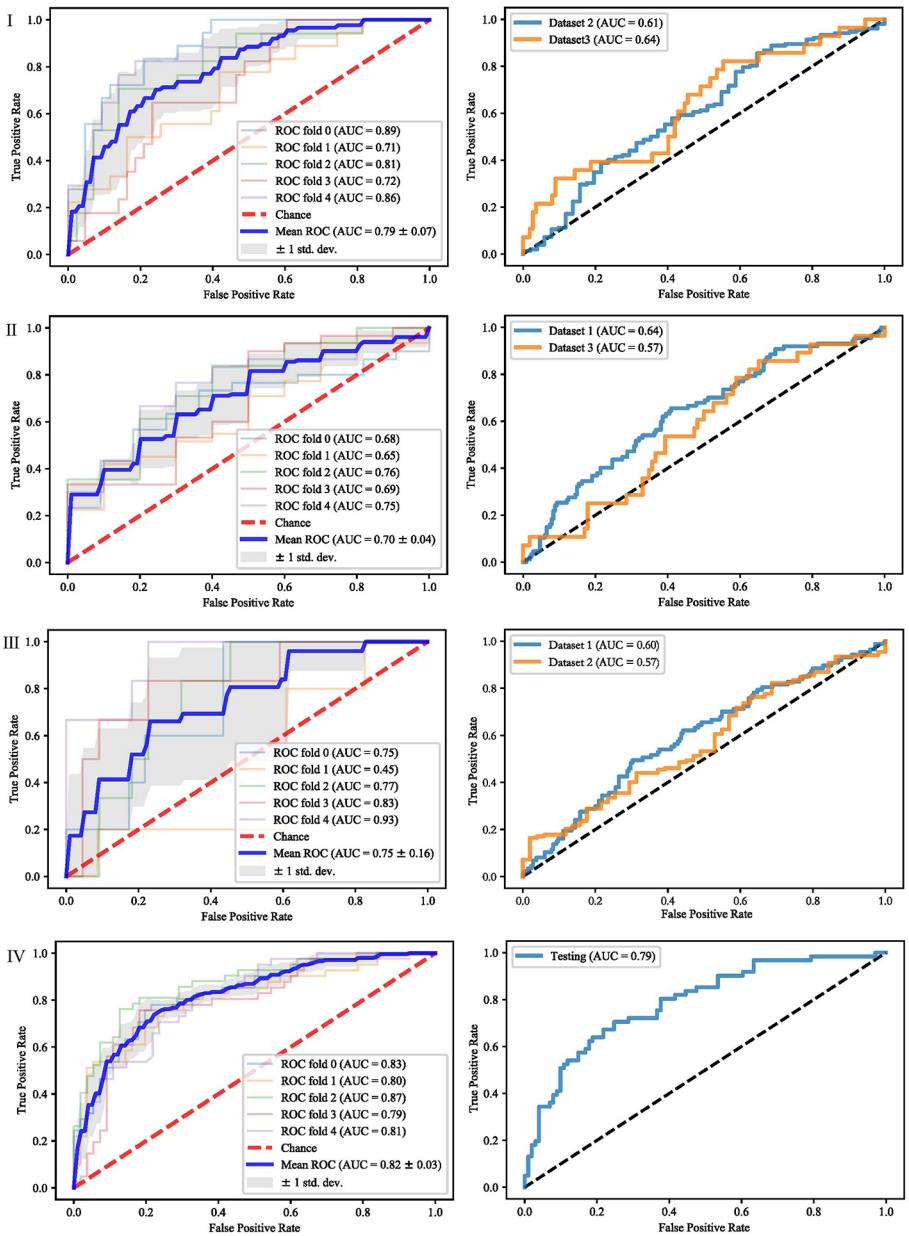


FIGURE 6 | ROCs of different datasets achieved through Lasso and SVM. Each row represents different datasets; (I) Dataset 1. (II) Dataset 2. (III) Dataset 3. (IV) Merged dataset.

selection method was used, the average AUC value of 5-fold cross-validation was 0.82, and the AUC value in the test set was 0.79.

It can be seen from the above results that although different feature selection methods selected different features, relatively consistent classification results were obtained under different classifiers, among which Lasso feature selection method achieved the best classification results. When using datasets 1–3 for modeling, the results on the testing set were poor, with AUC values ranging from 0.54 to 0.64. According to the definition of AUC value, in the range of [0.5–0.7], although the model has certain prediction ability, its prediction level is relatively poor. The AUC values of 0.76, 0.78, and 0.79 in the testing set of the merged dataset were obtained respectively, showing good predictive performance.

For further analysis, we evaluated the accuracy, sensitivity, specificity and AUC values of different models. The results are shown in **Table 3**. When modeling with dataset 1, the average accuracy, sensitivity, specificity, and AUC values on dataset 2 were 0.45, 0.73, 0.36, and 0.58, respectively; the average accuracy, sensitivity, specificity, and AUC values on dataset 3 were 0.64, 0.70, 0.43, and 0.59, respectively. When modeling with dataset 2, the average accuracy, sensitivity, specificity, and AUC values on

dataset 1 were 0.48, 0.40, 0.69, and 0.57, respectively; the average accuracy, sensitivity, specificity, and AUC values on dataset 3 were 0.46, 0.41, 0.67, and 0.56, respectively. When modeling with dataset 3, the average accuracy, sensitivity, specificity and AUC values on dataset 1 were 0.62, 0.72, 0.38, and 0.58, respectively; the average accuracy, sensitivity, specificity, and AUC values on dataset 2 were 0.47, 0.69, 0.40, and 0.58, respectively. When modeling with the merged dataset, the average accuracy, sensitivity, specificity and AUC values on the testing set were 0.74, 0.77, 0.68, and 0.78, respectively. Based on the above results, it can be seen that the classification results are the best when using the merged dataset for modeling.

DISCUSSION

This paper studied the subtype differentiation of NSCLC based on radiomics analysis. The identification of histological types of NSCLC is essential for treatment, and many studies have been conducted and demonstrated the potential ability of radiomics. However, existing studies usually focus on relatively small datasets and lack multi-center external validation datasets, resulting in a high risk of over-fitting, so the model's generalization performance cannot be adequately verified. Besides, the methods used in many studies are relatively simple, and few feature selection and classifier methods are compared. The differences between different methods are not fully discussed, which reduces the credibility of the model and limits its clinical application.

To solve the above problems, we retrospectively studied three independent datasets from different centers, where each dataset was used to train the model and tested in the remaining two datasets. Then all the datasets were combined into a large dataset and randomly divided into training and testing sets for modeling and analysis. For each form of dataset division, a corresponding radiomics model was constructed. The experimental results show that each model's performance is quite different, and the model based on the merged dataset obtains the best performance.

The feature subsets selected by different feature selection methods vary greatly, which is also the difficulty of radiology research. How to select the most effective feature subset is a complex feature engineering problem, especially in radiomics research. Besides, how to ensure the interpretability of features is another difficulty in applying radiomics models to the clinic. In future work, we will continue to conduct research in this field by combining doctors' qualitative semantic features and deep learning features, and using ensemble methods to select interpretable and distinguishable features.

Generalization ability is an important index in radiomics research. The samples we studied came from different centers, with different imaging methods and a wide range of demographic information. The experimental results show that the performance of the other centers' datasets was poor when only one dataset was used for modeling. The results put forward a requirement for future radiomics research; that is, to better apply it to clinical practice, it is necessary to collect as much multi-center datasets as possible in order to learn the common feature representation. When using multi-center datasets, the imaging quality is another issue that needs to pay attention. Some notable works have discussed imaging quality issues (28–30), which inspire us to carry out future work.

The curse of dimensionality is a huge challenge in the radiomics analysis, so feature selection is an essential step. Many studies have discussed the performance of different feature selection methods. Qian et al. (31) evaluated 12 feature selectionmethods and 7 classification methods to distinguish glioblastoma from solitary brain metastases and found that the Lasso feature selection and SVM classifier obtained the highest AUC. Wu et al. (11) investigated 24 filter-based feature selection methods and 3 classification methods for the classification of lung cancer histology, and found that the ReliefF feature selection method has higher prediction accuracy than other methods. In this study, we studied three widely used feature selection methods, namely mRMR, Lasso and SFS. The experimental results demonstrated that although different feature selection methods selected different features, relatively consistent classification results were obtained under different classifiers, among which Lasso feature selection method achieved the best classification results.

Since the samples we studied came from different institutions, the process of tumor segmentation by different radiologists and the repeatability of radiomic features may vary significantly. Subjectivity may occur in the determination of tumor volume and tumor boundaries, leading to uncertainties of tumor segmentation, which may adversely affect the repeatability of radiomic features. It is widely acknowledged that it is difficult to precisely delineate the tumor volume due to the similar characteristics between organs and tumors, as well as the differences in shape and size of the tumor. Moreover, medical images are far from perfect because they have limited resolution and may contain artifacts. Physicians often interpret tumors differently based on their skills and experiences. Since radiomic features are calculated based on the tumor masks, the uncertainties of the tumor segmentation significantly affect the features, resulting in poor generalization performance of the

TABLE 3 | Classification results of different testing sets.

Training set	Testing set	Accuracy	Sensitivity	Specificity	AUC
Dataset 1	Dataset 2	0.45	0.73	0.36	0.58
	Dataset 3	0.64	0.70	0.43	0.59
Dataset 2	Dataset 1	0.48	0.40	0.69	0.57
	Dataset 3	0.46	0.41	0.67	0.56
Dataset 3	Dataset 1	0.62	0.72	0.38	0.58
	Dataset 2	0.47	0.69	0.40	0.58
Merged training set	Merged testing set	0.74	0.77	0.68	0.78

prediction models (32). With the development of computer vision and deep learning, automatic tumor segmentation may help radiomics studies. One of our recent work shows that the classification results based on automated segmentation and ground truth segmentation have no significantdifferences in computer-aided glioma grading task (33). In the future work, we will combine the automatic segmentation method for radiomics research.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Ethics Committee of Shandong Cancer Hospital. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

FY and WC collected and analyzed the data, wrote the initial draft, and accomplished the final version. HW and XZ analyzed the data. SY, XQ, and YC conceptualized, designed, and supervised the study, and revised the article. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant nos. 81872475, 81372413, and U1806202), the Department of Science and Technology of Shandong Province (2017CXGC1502), and the Taishan Scholars Project in Shandong Province (ts201511106).

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* (2020) 70:7–30. doi: 10.3322/caac.21590
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in china, 2015. *CA Cancer J Clin* (2016) 66:115–32. doi: 10.3322/caac.21338
4. Inamura K. Lung cancer: Understanding its molecular pathology and the 2015 who classification. *Front Oncol* (2017) 7:193. doi: 10.3389/fonc.2017.00193
5. Zhan C, Yan L, Wang L, Sun Y, Wang X, Lin Z, et al. Identification of immunohistochemical markers for distinguishing lung adenocarcinoma from squamous cell carcinoma. *J Thorac Dis* (2015) 7:1398–405. doi: 10.3978/j.issn.2072-1439.2015.07.25
6. Kadota K, Nitadori JI, Rekhtman N, Jones DR, Adusumilli PS, Travis WD. Reevaluation and reclassification of resected lung carcinomas originally diagnosed as squamous cell carcinoma using immunohistochemical analysis. *Am J Surg Pathol* (2015) 39:1170–80. doi: 10.1097/pas.0000000000000439
7. Huang T, Li J, Zhang C, Hong Q, Jiang D, Ye M, et al. Distinguishing lung adenocarcinoma from lung squamous cell carcinoma by two hypomethylated and three hypermethylated genes: A meta-analysis. *PloS One* (2016) 11: e0149088. doi: 10.1371/journal.pone.0149088
8. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, 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
9. 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
10. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* (2017) 28:1191–206. doi: 10.1093/annonc/mdx034
11. Wu W, Parmar C, Grossmann P, Quackenbush J, Lambin P, Bussink J, et al. Exploratory study to identify radiomics classifiers for lung cancer histology. *Front Oncol* (2016) 6:71. doi: 10.3389/fonc.2016.00071
12. Zhu X, Dong D, Chen Z, Fang M, Zhang L, Song J, et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. *Eur Radiol* (2018) 28:2772–8. doi: 10.1007/s00330-017-5221-1
13. Chaunzwa TL, Christiani DC, Lanuti M, Shafer A, Diao N, Mak RH, et al. Using deep-learning radiomics to predict lung cancer histology. *J Clin Oncol* (2018) 36:8545–5. doi: 10.1200/JCO.2018.36.15_suppl.8545
14. Wolf I, Vetter M, Wegner I, Böttger T, Nolden M, Schöbinger M, et al. The Medical Imaging Interaction Toolkit. *Med Image Anal* (2005) 9:594–604. doi: 10.1016/j.media.2005.04.005
15. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The cancer imaging archive (tcia): Maintaining and operating a public information repository. *J Digit Imaging* (2013) 26:1045–57. doi: 10.1007/s10278-013-9622-7
16. Aerts H, Rios Velazquez E, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Data from nsclc-radiomics. *Cancer Imaging Arch* (2015). doi: 10.7937/K9/TCIA.2015.PF0M9REI
17. Bakr S, Gevaert O, Echegaray S, Ayers K, Zhou M, Shafiq M, et al. Data for nsclc radiogenomics collection. *Cancer Imaging Arch* (2017). doi: 10.7937/K9/TCIA.2017.7hs46erv
18. Zwanenburg A, Vallières M, Abdalah 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
19. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* (2017) 77:e104. doi: 10.1158/0008-5472.CAN-17-0339
20. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. In: . *Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003*. Stanford, CA, USA (2003). p. 523–8. doi: 10.1109/CSB.2003.1227396
21. Gheyas IA, Smith LS. Feature subset selection in large dimensionality domains. *Pattern Recognit* (2010) 43:5–13. doi: 10.1016/j.patcog.2009.06.009
22. Remeseiro B, Bolon-Canedo V. A review of feature selection methods in medical applications. *Comput Biol Med* (2019) 112:103375. doi: 10.1016/j.compbiomed.2019.103375
23. Marcano-Cedeño A, Quintanilla-Domínguez J, Cortina-Januchs MG, Andina D. Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network. In: . *IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society*, Glendale, AZ (2010). p. 2845–50. doi: 10.1109/IECON.2010.5675075

24. Forghani R, Savadjiev P, Chatterjee A, Muthukrishnan N, Reinhold C, Forghani B. Radiomics and artificial intelligence for biomarker and prediction model development in oncology. *Comput Struct Biotechnol J* (2019) 17:995–1008. doi: 10.1016/j.csbj.2019.07.001
25. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (svm) learning in cancer genomics. *Cancer Genomics Proteomics* (2018) 15:41–51. doi: 10.21873/cgp.20063
26. Zhou H, Dong D, Chen B, Fang M, Cheng Y, Gan Y, et al. Diagnosis of distant metastasis of lung cancer: Based on clinical and radiomic features. *Trans Oncol* (2018) 11:31–6. doi: 10.1016/j.tranon.2017.10.010
27. Abdoh SF, Rizka MA, Maghraby FA. Cervical cancer diagnosis using random forest classifier with smote and feature reduction techniques. *IEEE Access* (2018) 6:59475–85. doi: 10.1109/ACCESS.2018.2874063
28. Chen Y, Ma J, Feng Q, Luo L, Shi P, Chen W. Nonlocal Prior Bayesian Tomographic Reconstruction. *J Math Imaging Vision* (2008) 30:133–46. doi: 10.1007/s10851-007-0042-5
29. Yin X, Zhao Q, Liu J, Yang W, Yang J, Quan G, et al. Domain progressive 3d residual convolution network to improve low-dose ct imaging. *IEEE Trans Med Imaging* (2019) 38:2903–13. doi: 10.1109/TMI.2019.2917258
30. Hu D, Wu W, Xu M, Zhang Y, Liu J, Ge R, et al. Sister: Spectral-image similarity-based tensor with enhanced-sparsity reconstruction for sparse-view multi-energy ct. *IEEE Trans Comput Imaging* (2020) 6:477–90. doi: 10.1109/TCI.2019.2956886
31. Qian Z, Li Y, Wang Y, Li L, Li R, Wang K, et al. Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers. *Cancer Lett* (2019) 451:128–35. doi: 10.1016/j.canlet.2019.02.054
32. Owens CA, Peterson CB, Tang C, Koay EJ, Yu W, Mackin DS, et al. Lung tumor segmentation methods: Impact on the uncertainty of radiomics features for non-small cell lung cancer. *PloS One* (2018) 13:e0205003. doi: 10.1371/journal.pone.0205003
33. Chen W, Liu B, Peng S, Sun J, Qiao X. Computer-aided grading of gliomas combining automatic segmentation and radiomics. *Int J Biomed Imaging* (2018) 2018:2512037. doi: 10.1155/2018/2512037

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 25 February 2021
doi: 10.3389/fonc.2021.590836

Positron Emission Tomography-Based Short-Term Efficacy Evaluation and Prediction in Patients With Non-Small Cell Lung Cancer Treated With Hypo-Fractionated Radiotherapy

Yi-Qing Jiang[1†], Qin Gao[1†], Han Chen[1], Xiang-Xiang Shi[1], Jing-Bo Wu[1], Yue Chen[2], Yan Zhang[2], Hao-Wen Pang[1] and Sheng Lin[1,2*]*

[1] Department of Oncology, The Affiliated Hospital of Southwest Medical University, Luzhou, China, [2] Nuclear Medicine and Molecular Imaging Key Laboratory of Sichuan Province, The Affiliated Hospital of Southwest Medical University, Luzhou, China

OPEN ACCESS

Edited by:
Luigi Aloj, University of Cambridge, United Kingdom

Reviewed by:
Wei Huang, Shandong Academy of Medical Sciences (SDAMS), China
Leonardo Rundo, University of Cambridge, United Kingdom

***Correspondence:**
Hao-Wen Pang
haowenpang@foxmail.com
Sheng Lin
lslinsheng@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *03 August 2020*
Accepted: *18 January 2021*
Published: *25 February 2021*

Citation:
Jiang Y-Q, Gao Q, Chen H, Shi X-X, Wu J-B, Chen Y, Zhang Y, Pang H-W and Lin S (2021) Positron Emission Tomography-Based Short-Term Efficacy Evaluation and Prediction in Patients With Non-Small Cell Lung Cancer Treated With Hypo-Fractionated Radiotherapy. Front. Oncol. 11:590836. doi: 10.3389/fonc.2021.590836

Background: Positron emission tomography is known to provide more accurate estimates than computed tomography when staging non–small cell lung cancer. The aims of this prospective study were to contrast the short-term efficacy of the two imaging methods while evaluating the effects of hypo-fractionated radiotherapy in non-small cell lung cancer, and to establish a short-term efficacy prediction model based on the radiomics features of positron emission tomography.

Methods: This nonrandomized-controlled trial was conducted from March 2015 to June 2019. Thirty-one lesions of 30 patients underwent the delineation of the regions of interest on positron emission tomography and computed tomography 1 month before, and 3 months after hypo-fractionated radiotherapy. Each patient was evaluated for the differences in local objective response rate between the two images. The Kaplan Meier method was used to analyze the local objective response and subsequent survival duration of the two imaging methods. The 3D Slicer was used to extract the radiomics features based on positron emission tomography. Least absolute shrinkage and selection operator regression was used to eliminate redundant features, and logistic regression analysis was used to develop the curative-effect-predicting model, which was displayed through a radiomics nomogram. Receiver operating characteristic curve and decision curve were used to evaluate the accuracy and clinical usefulness of the prediction model.

Results: Positron emission tomography-based local objective response rate was significantly higher than that based on computed tomography [70.97% (22/31) and 12.90% (4/31), respectively ($p<0.001$)]. The mean survival time of responders and non-

responders assessed by positron emission tomography was 28.6 months vs. 11.4 months (p=0.29), whereas that assessed by computed tomography was 24.5 months vs. 26 months (p=0.66), respectively. Three radiomics features were screened to establish a personalized prediction nomogram with high area under curve (0.94, 95% CI 0.85–0.99, p<0.001). The decision curve showed a high clinical value of the radiomics nomogram.

Conclusions: We recommend positron emission tomography for evaluating the short-term efficacy of hypo-fractionated radiotherapy in non-small cell lung cancer, and that the radiomics nomogram could be an important technique for the prediction of short-term efficacy, which might enable an improved and precise treatment.

Registration number/URL: ChiCTR1900027768/http://www.chictr.org.cn/showprojen.aspx?proj=46057

Keywords: positron emission tomography, radiomics, non-small-cell lung cancer, computed tomography, hypo-fractionated radiotherapy

INTRODUCTION

Lung cancer is the most common cancer type and the leading cause of cancer-associated mortality worldwide (1). Hypo-fractionated radiotherapy (HFRT) includes stereotactic body radiation therapy (SBRT) and hypo-fractionated brachytherapy, both of which deliver a high biologically effective dose (BED) to the tumor while minimizing toxicity to the normal tissues. Therefore, HFRT can prompt superior local control and improved survival (2).

Currently, lung cancer treatment efficacy evaluation is mostly based on computed tomography (CT) (3). However, CT being a structural imaging, has limited value in the detection of an early response to therapy, and the tumors could be obscured by atelectasis and radiation pneumonitis (4). Instead, functional imaging with 18F-fludeoxyglucose positron emission tomography (^{18}F-FDG PET) may promote accurate and early assessment of therapy response (5–7). In addition, recent studies have demonstrated that radiomics has been successfully used to stage the tumor, assess the side effects, and predict the clinical endpoints in lung cancer (8–10). However, to our knowledge, few studies have focused on contrasting the efficacy of PET and CT in peripheral non-small cell lung cancer (NSCLC) after HFRT treatment as well as on the development of a prediction model of local short-term efficacy based on PET radiomics.

Therefore, the main objectives of this prospective trial were to investigate the differences between short-term efficacy of PET and CT while evaluating the effects of HFRT in peripheral NSCLC, and to screen the efficacy-related radiomics features of PET imaging and use those to establish a prediction model.

MATERIALS AND METHODS

Patients

PET imaging was performed as a part of the phase I/II clinical trial (Clinicaltrials.gov number: ChiCTR1900027768) that evaluated the efficacy of HFRT in patients with pathologically confirmed NSCLC. The patients who underwent HFRT for primary NSCLC (T2-4N0-3M0-1) at our hospital, between March 2015 and June 2019 were enrolled in this study. Eventually, 31 lesions of 30 patients underwent serial 18F-FDG PET/CT 1 month before HFRT and 3 months after HFRT on the same scanner. The TNM stage was designated according to the American Joint Committee on Cancer Staging 8th edition (AJCC) (11). The N and M staging was based on pre-treatment PET/CT and magnetic resonance imaging (MRI). Additionally, all patients underwent pathological diagnosis for the lung lesion before undergoing treatment. The first follow-up imaging examination was performed 4–12 weeks after radiotherapy. After that, patients were monitored every three months in the first year, every six months in the next two years, and once a year thereafter. Because the patients lived in remote places, the scan was occasionally performed outside these limits. The study was approved by the Research Ethics Committee of the Affiliated Hospital of Southwest Medical University (Date of approval by ethics committee and approval number: 2013-8-26, and KY2019276, respectively) and conducted in accordance with the Declaration of Helsinki (as revised in 2013) and its later amendments or comparable ethical standards. Written informed consent was obtained from all individual participants included in the study.

Treatment

All patients were treated with radiotherapy planning system (TPS) (Oncentra 4.3, Elekta, Sweden). Then, 19 out of 30 patients were treated with hypo-fractionated brachytherapy, delivered with an ^{192}Ir source from a MicroSelectron afterloader (Elekta Brachytherapy, Elekta AB, Stockholm, Sweden), and these patients were administered a single dose of 30Gy, as recommended by a previous clinical trial (12). The remaining 11 patients were treated with SBRT that was delivered in 3–5 fractions to a total of 23–50Gy. After radiotherapy, concurrent platinum-based doublet adjuvant chemotherapy was allowed. In addition to cisplatin (or carboplatin), a second concurrent nonplatinum agent was required (e.g., paclitaxel, etoposide).

PET and CT Scanning Acquisition and Processing

Patients fasted for at least 6 h before ^{18}F-FDG was administered. The patient's blood glucose level should have been ≤ 11 mmol/L. The PET-CT was performed according to the European Association of Nuclear Medicine (EANM) guidelines version 1.0 (13). A whole-body PET-CT (Philips Gemini TF/16; Philips, Cleveland, Ohio, USA) was performed after the intravenous administration of ^{18}F-FDG (5.55 MB q/kg). Then, low-dose helical CT transmission scanning (pitch, 0.813; current, 100 mAs; peak voltage, 120 kV; slice thickness, 5.0 mm) was performed with attenuation correction and lesion localization. PET was then performed at 1.5 min per bed position and used 19–21 bed positions. ^{18}F-PET/CT scanning was performed from the vertex of the head to the feet. In order to reduce the impact of respiratory motion on image acquisition and ensure the credibility of the research results, chest scans were conducted after having the patients hold their breath.

Assessment of Local Objective Response Rate

The PET-only and CT-only images were both sent to the three-dimensional (3D) radiotherapy planning system (TPS) (Oncentra 4.3, Elekta, Sweden) *via* the local area network. One physician with more than 15 years of experience with PET and CT in peripheral lung cancer and regions of interest (ROI) definition performed the analyses after being blinded to the patient outcome data. Nodal disease was not included in the analysis. PET-based ROI delineation was carried out before and after HFRT with standard uptake value (SUV) of 2.5 as the initial threshold (14, 15). Mean standard uptake value (SUVmean), maximum standard uptake value (SUVmax), metabolic tumor volume (MTV), and longest diameter (Dmax) were calculated for the ROI based on PET. CT-based ROI was manually drawn by the same physician with lung windows (window width, 1600 Hounsfield units [HU]; window level, 600 HU). Volume and longest diameter from the ROI based on CT were also calculated. CT-only scans were assessed for response using Response evaluation criteria in solid tumors (RECIST1.1) (16) after the treatment, and PET-based response criteria used in this study were according to the European Organization for Research and Treatment of Cancer (EORTC) (17) that were based on an assessment of the SUVmax measured through ROI analysis. The terms for the response categories were the same for both CT and PET: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD).

Radiomic Processing

The workflow for radiomic processing included the following four steps: image acquisition and reconstruction, image segmentation, feature extraction, and data analysis (18, 19). The first two steps involving collection of PET images and delineation of the ROI were described in the above section. The feature extraction and definition in this study were consistent with the Imaging Biomarker Standardization Initiative (IBSI) (20). The feature extraction process was divided into image processing and feature calculation. For each ROI based on PET image, a resampled 4×4×4 mm^3 voxel size and a bin width of 25 were applied (21). After image processing, Due to the characteristics of medical images, filter properties are important for image analysis methods. Filtering properties including geometric invariances for medical image analysis directional sensitivity, combining directional sensitivity and invariance to local rotations, spectral coverage (22). We used wavelet which is a filtering method based on a collection with eight combinations of applying either a high or a low pass filter and cover the entire image spectrum as study filter. Wavelets families contains the following wavelet groups: "haar," "db," "sym," "coif," "bior," "rbio," "dmey" (22). After those, all feature classes with the exception of shape can be calculated on the original image and a derived image which obtained by applying wavelets filter. Feature extraction was based on the 3Dslicer platform and used the pyradiomics package, which is available at: http://PyRadiomics.readthedocs.io/en/latest/, accessed June 30, 2019 (23).

Statistical Analysis

We assessed the differences between all parameters observed by CT and PET using the Wilcoxon signed-rank test or paired t test after ascertaining whether the parameters were normally distributed using the Shapiro-Wilks test. The delta-parameters of the CT and PET were identified by the prefix "Δ."

Tumor response to HFRT was analyzed using PET and CT independently. Patients were then grouped as responders (CR+PR) or non-responders (SD+PD). The difference of ORR between PET and CT was assessed using the Mann-Whitney U test. The Kaplan Meier method was used to analyze the local objective response and subsequent survival duration of the two imaging methods.

The least absolute shrinkage and selection operator (LASSO) regression analysis was used to screen out the radiomics features related to responders, and logistic regression analysis was used to develop the curative-effect-predicting model. To assess the probability of short-term efficacy in individuals, we built the radiomics nomogram based on multivariable logistic analysis. Receiver operating characteristic (ROC) curve and area under the curve (AUC) were used to assess the accuracy of the prediction model. To ensure radiomic robustness, we performed 1,000 bootstrap resamples to check the agreement between our prediction based on radiomics and actual observation. The net benefit was quantified by the decision curve analysis to determine the clinical applicability of the PET-based radiomics nomogram prediction model.

All statistical analyses were performed using SPSS, Version 17.0 software (SPSS, Inc., Chicago, IL) and R software, version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria), using the glmnet, rms, and pROC analysis packages for Windows. The level of statistical significance was defined as a p value less than 0.05 based on 2-sided tests.

RESULTS

Baseline Characteristics and Follow-Up

From March 2015 to June 2019, 30 patients with peripheral NSCLC were enrolled. The baseline clinical characteristics of these patients are listed in **Tables 1** and **2**.

Survival was measured from the date of completing radiotherapy to the date of death from any cause since April

TABLE 1 | Characteristics of the patients.

Characteristics	Median (Range) or Number (%)
Age (y)	55.5 (43–77)
Gender	
Male	24 (80.00)
Female	6 (20.00)
Stage	
IIIA~IIIC	8 (26.67)
IVA~IVB	22 (73.33)
Histology	
Adenocarcinoma	22 (73.33)
Squamous carcinoma	8 (26.67)
Smoking	
Yes	18 (60.00)
No	12 (40.00)
KPS	
70~80	15 (50.00)
90~100	15 (50.00)
ECOG	
0	15 (50.00)
1~2	15 (50.00)
Location	
Left	13 (41.94)
Right	18 (58.06)
Radiotherapy	
Stereotactic body radiation therapy	12 (38.71)
Hypo-fractionated brachytherapy	19 (61.29)

KPS, Karnofsky Performance Status; ECOG, Eastern Cooperative Oncology Group.

2020. Patients who were still alive at the date of last contact were censored at deadline. Only one patient was lost to follow-up before the deadline date; nine patients died and the rest were still alive. Estimated median follow-up duration was 16 months.

Comparison of PET- and CT-based Response Assessments

The PET based efficacy in all 31 lesions was assessed as CR (n=0), PR (n=22), SD (n=6), and PD (n=3), while the same according to CT was CR (n=0), PR (n=4), SD (n=23) and PD (n=4). Comparisons for all patients are shown in **Figure 1**. For these patients assessable by both PET and CT, there was a poor agreement between the two assessments, with a weighted kappa value of 0.114. A significantly higher number of patients were regarded as responders on PET than on CT in the patients assessable by both techniques (22 vs. four patients respectively; $p<0.001$). All differences between parameters observed by CT and PET are listed in **Table 3**. A representative case of an individual patient exemplifying the different CT and PET/CT response is shown in **Figure 2**.

Effect of Chemotherapy on Treatment Evaluation

In order to explore the effect of chemotherapy after radiotherapy on the evaluation of short-term efficacy, the patients were divided

TABLE 2 | Patient TNM staging, duration and type of chemotherapy, and duration of response.

Patient no.	Stage	Type of chemotherapy	Chemotherapy cycle	Duration of response (month)
1	T3N0M1	Pemetrexed+cisplatin	4	3
2	T4N2M0	Etoposide+ cisplatin	3	6
3	T3N2M0	None	\	3
4	T4N0M1	None	\	3
5	T4N2M1	None	\	3
6	T4N3M1	None	\	1
7	T4N2M1	None	\	3
8	T2N3M0	Paclitaxel+ carboplatin	2	3
9	T4N3M0	Etoposide+ cisplatin	4	6
10	T2N2M0	None	\	6
11	T4N2M1	None	\	6
12	T4N3M1	Etoposide+ cisplatin	3	6
13	T4N2M1	Paclitaxel+ carboplatin	1	1
14	T4N2M1	Paclitaxel+ carboplatin	1	1
15	T4N2M1	Paclitaxel+ carboplatin	1	3
16	T4N2M0	None	\	6
17	T2N3M0	None	\	1
18	T4N1M1	None	\	3
19	T4N3M1	None	\	1
20	T3N1M1	None	\	1
21	T3N1M1	None	\	3
22	T4N0M0	None	\	3
23	T3N3M1	None	\	6
24	T2N1M1	None	\	3
25	T4N3M1	None	\	3
26	T4N3M1	None	\	3
27	T4N3M1	None	\	6
28	T1N2M1	Etoposide+carboplatin	3	3
29	T4N2M1	None	\	1
30	T2N1M1	None	\	1

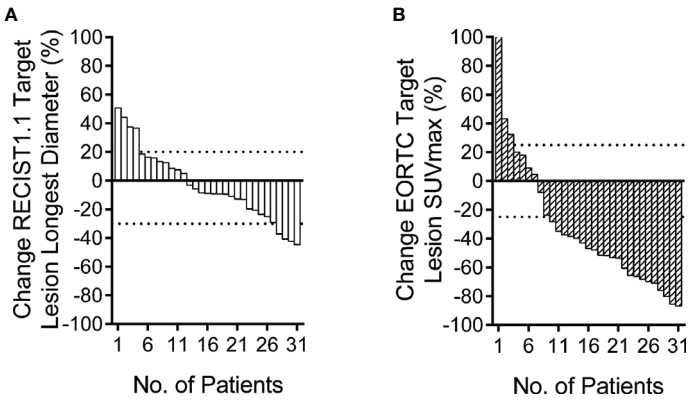


FIGURE 1 | Relative change rate of longest diameter and maximum standard uptake value (SUVmax) in each patient. Best overall response waterfall plots, in which computed tomography (CT) **(A)** is based on the rate of longest diameter changes, according to RECIST 1.1, and positron emission tomography (PET) **(B)** is based on the rate of SUVmax changes, according to European Organization for Research and Treatment of Cancer (EORTC). The top dotted line represents progressive disease, the bottom dotted line represents partial response, or complete response, while, stable disease is represented by the area between the two dotted lines.

TABLE 3 | Comparison of difference variables and short-term efficacy assessment based on PET and CT.

Variable	PET/CT	CT	p Value
Tumor volume (cm^3)			
V_B	25.31 (14.92,89.79)	28.54 (15.05,76.95)	0.337†
V_L	6.40 (2.14,14.46)	18.25 (8.59,67.02)	<.001†
ΔV	−18.07 (−63.04, −8.58)	−13.93 (−49.56, 5.03)	<.001†
D-max (cm)			
D_B	5.19 (3.53,6.80)	5.30 (3.99,7.38)	0.019†
D_L	3.19 (1.62,3.93)	4.59 (3.83,7.83)	<.001†
ΔD	−1.66 (−3.81, −0.73)	−0.35 ± 1.30	<.001†
$SUVmean_B$	4.78 (3.92,5.47)	/	/
$SUVmax_B$	9.92 (7.52,11.77)	/	/
$SUVmax_L$	4.72 (3.31,7.83)	/	/
ΔSUVmax	−4.17 ± 5.25	/	/
Efficacy§			<.001‡
Responders	22 (70.97%)	4 (12.90%)	
Non-responders	9 (29.03%)	27 (87.10%)	

PET, positron emission tomography/computed tomography; CT, computed tomography; the suffix B, value of volume/longest diameter before brachytherapy; the suffix L, value of volume/longest diameter after hypo-fractionated brachytherapy (HFBT). Use mean ± SD for normally distributed data and interquartile ranges (IQRs) for data that are not normally distributed.
$\Delta V=V_L-V_B$, $\Delta D=D_L-D_B$, $\Delta SUVmax=SUVmax_L-SUVmax_B$.
†p value of the Wilcoxon signed rank test.
‡p value of the McNemar test continuity correction.
§PET-based and CT-based evaluation of efficacy is according to the European Organization for Research and Treatment of Cancer (EORTC)and Response Evaluation Criteria in Solid Tumors (RECIST 1.1) respectively; complete response or partial response means responders, progressive disease or stable disease means non-responders.

into two groups: post radiotherapy chemotherapy group (Chemotherapy) and non-chemotherapy treatment group (None). The difference of curative effect between the two groups was compared. The results showed that there was no significant difference between the patients receiving chemotherapy and those who did not receive chemotherapy (**Figure 3**, p=0.374), that is, no matter whether chemotherapy was carried out after radiotherapy, the short-term efficacy evaluation of patients had no influence.

Prognostic Significance of Response Assessments

The mean survival time of responders and non-responders assessed by PET was 28.6 months vs. 11.4 months, whereas that assessed by CT was 24.5 months vs. 26 months, respectively. In PET assessment, the survival duration of responders was longer than that of non-responders, whereas CT assessment had the opposite result. Thus, the PET responses seem to be stronger prognostic indicators than the CT responses. However, neither the PET nor CT scan assessments of response had a significant difference in subsequent survival duration between the responders and non-responders (**Figures 4A, B**).

Radiomics Feature Selection and Response Prediction

All 851 radiomics features were extracted, including Shape features, First Order statistical features, Gray Level

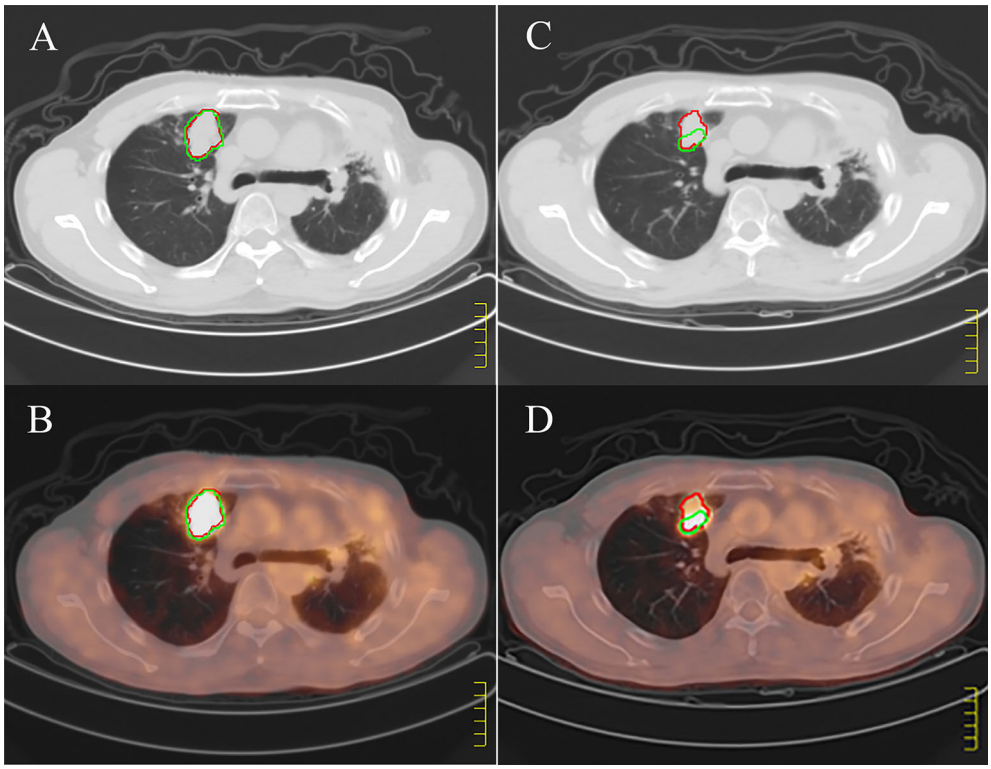


FIGURE 2 | Example of discordant positron emission tomography (PET/CT) and computed tomography (CT). **(A, B)** show pretreatment CT and PET/CT images respectively, and **(C, D)** show CT and PET/CT images of the large tumors 3 months after treatment. The red regions of interest (ROI) is manually segmented based on CT, and the green ROI is delineated according to SUV value higher than 2.5 based on PET. According to the images before and after treatment, CT shows stable disease; PET/CT shows partial response.

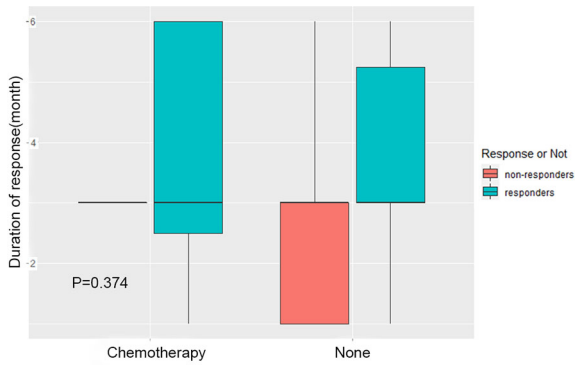


FIGURE 3 | Effect of chemotherapy on treatment evaluation. The patients were divided into two groups: chemotherapy group after radiotherapy (Chemotherapy) and group without chemotherapy after radiotherapy (None). The difference of short-term efficacy between the two groups was compared.

Co-occurrence Matrix (GLCM) features, Gray Level Dependence Matrix (GLDM) features, Gray Level Run Length Matrix (GLRLM) features, Gray Level Size Zone Matrix (GLSZM) features, and Neighboring Gray Tone Difference Matrix (NGTDM) features. Among the radiological features, 851 features were reduced to three potential predictors based on 31 lesions (**Figure 5**), which were Busyness of NGTDM of wavelet-LHL, Short Run High Gray Level Emphasis of GLRLM of wavelet-LHH, and Median of First order of wavelet-HHH. These three features were used to establish a model that is presented as a nomogram in **Figure 6**. Based on the ROC curve analysis, the model with high AUC (0.94, 95% CI 0.85–0.99, p<0.001) is presented in **Figure 7A**, and the decision curve analysis for the model is presented in **Figure 7B**. The decision curve indicates that if the threshold probability of a patient is 40%, then the use of radiographs from the current study to predict treatment outcomes would add more benefits than the "treat-all-patients" or "treat-none" options.

DISCUSSION

Currently, CT is the recommended method for response assessment in NSCLC as per the RECIST guidelines. There is insufficient standardization or evidence to abandon anatomical assessment of tumor burden, or to determine if it is appropriate to replace the unidimensional anatomic assessment with either volumetric anatomical assessment or functional assessment with

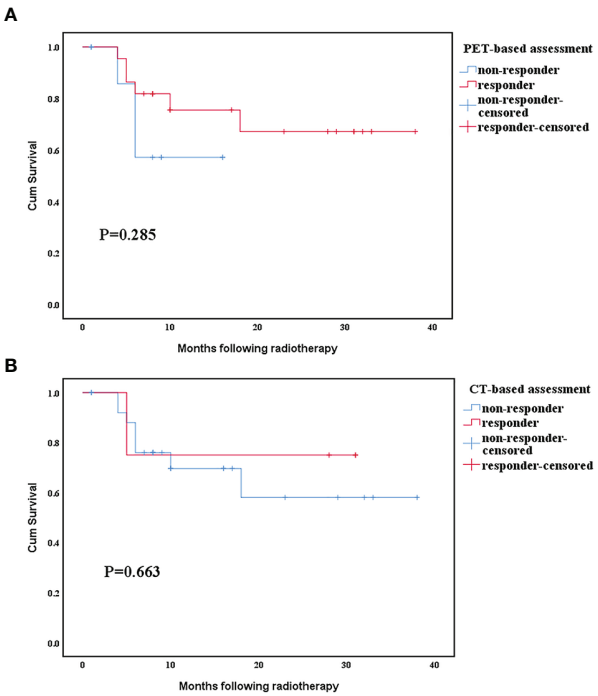


FIGURE 4 | Relationship between survival time and response evaluated by positron emission tomography (PET) and computed tomography (CT) scan. **(A)** Comparison of positron emission tomography scan response categories. **(B)** Comparison of computed tomography scan response categories.

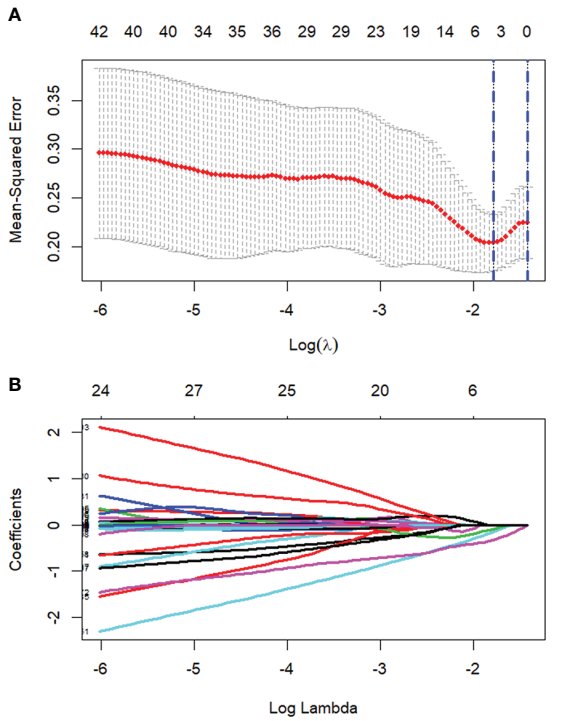


FIGURE 5 | Process of LASSO regression screening of radiomics features. Screening radiomics features using the least absolute shrinkage and selection operator (LASSO) binary logistic regression model. **(A)** Tuning parameter (λ) in the LASSO model used 5-fold cross-validation. Dotted vertical lines are drawn by using the minimum criteria and the 1 standard error of the minimum criteria. **(B)** 851 normalized lasso coefficient plots of radiomics features. When log (λ) takes the minimum criteria, three non-zero coefficients are selected.

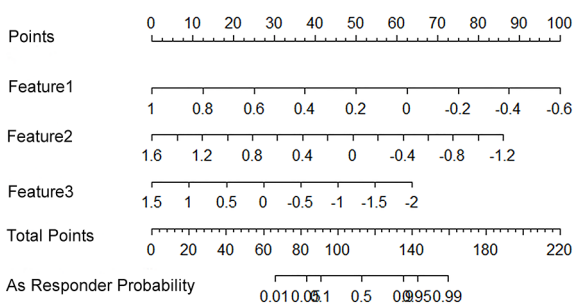


FIGURE 6 | A radiomics nomogram that predicts the probability of effective treatment. Feature1: Wavelet.LHL_NGTDM_Busyness; Feature2: Wavelet.LHH_GLRLM_ShortRunHighGrayLevelEmphasis; Feature3: Wavelet.HHH_Firstorder_Median.

PET (16). Some studies have revealed that PET has better evaluation validity than CT in patients with both NSCLC (7) and small cell lung cancer (24). Our data revealed that among such patients, who are assessable by both techniques, more patients were regarded as responders on PET than on CT. The local ORR of patients according to PET was 70.97%, whereas the corresponding CT evaluation found that to be 12.90%. Moreover, when the efficacy evaluation was compared between CT and PET, only eight lesions (25.81%) were evaluated uniformly, of which, four were evaluated as PR and four were evaluated as SD. Some researchers have previously reported that PET/CT-derived tumor volumes were smaller than those derived by CT alone in case of locally advanced-stage peripheral lung cancer before radiotherapy treatment (25). However, we found that the volumes of ROIs, delineated by PET and CT before HFRT, were similar. In contrast, Dmax of ROIs delineated using PET before and after HFRT, and the volumes of ROIs after HFRT delineated using PET were significantly smaller than that of ROIs outlined using CT, as shown in **Table 3**, mainly because CT was not very sensitive to distinguish between atelectasis and lung cancer. Moreover, radiation therapy may result in radiation-induced lung opacity (RILO) on CT (5), such as ground-glass opacity, scar or fibrotic changes, consolidation with air-bronchogram, consolidation alone, and nodule. These RILOs may result in a larger size and diameter of the tumor being observed on the CT, which may overestimate the progressive disease during exclusive assessment of efficacy *via* the CT. A previous study had also suggested that the combined PET/CT to evaluate post-treatment response would increase the correct identification of patients with progressive disease after lung SBRT (26). While investigating the differences in the PET and

CT based response assessment, we also report significant statistical differences between the two methods. As shown in **Figure 1**, only six patients were considered SD based on PET, whereas 23 patients were considered SD based on CT. According to our study, PET assessment is better than CT assessment in reflecting the prognosis of patients. Although there was no significant difference in the survival time between the two groups of patients, this may be due to the small sample size or short follow-up time. Based on the above data, we recommend that PET would be better than CT when evaluating the efficacy of HFRT in NSCLC, since the change in tumor volume may be slower than the metabolic change discernable by PET; additionally, CT scan may not accurately reflect the therapeutic effect on the tumor in time, which may lead to unnecessary overtreatment.

Since each patient's sensitivity to treatment is inconsistent, the prognosis may be completely different even in patients who are in the same cancer stage and receive the same treatment (27). Therefore, early prediction of treatment response is particularly important for identifying patients who may or may not benefit from treatment.

FDG uptake is not only related to increased metabolism, but also to other physiological parameters, such as cell proliferation (28), perfusion, invasiveness, and hypoxia (29). Therefore, radiomics can obtain several data contained in the PET image through non-invasive means. Many quantitative features of PET can be calculated during treatment of the patient. This principle of extracting image features is termed as "radiomics" that has been recently studied in esophageal cancer (30), NSCLC (31, 32), breast cancer (33), nasopharyngeal carcinoma (34), and rectal cancer (35), and demonstrated its potential in predicting treatment efficacy or patient prognosis.

PET-based radiomics had a high sensitivity in AUC for predicting the efficacy of radiochemotherapy in esophageal cancer (76%–92%) (30). In the prediction of the efficacy of adjuvant therapy for rectal cancer, the AUC of radiomics based on PET and MRI was up to 0.86 (35). Our model based on PET radiomics has an AUC of 0.94, indicating that PET radiomics plays a significant role in predicting the treatment efficacy in non-small cell lung cancer. The most important use of nomograms is based on explaining the individual's need for further treatment. Therefore, in order to prove its clinical value, we evaluated whether the radiomics nomogram assisted decision-making could add more benefits to the patients through the novel method of decision curve, which estimated the net benefit. (Net benefit was defined as the proportion of true positives minus the proportion of false positives, weighted by the relative harm of false-positive and false-negative results) (36).

The key to ensure the accuracy, generalization and repeatability of radiomics prediction is accurate and high repeatability ROI segmentation. In this study, all ROI are manually segmented. However, manual segmentation has the disadvantages of time-consuming and low repeatability of tumor volume description. Automatic or semi-automatic methods can make up for these defects. Recently, a lot of research on automatic segmentation is increasing rapidly. The fully automatic multi-mode PET/MRI

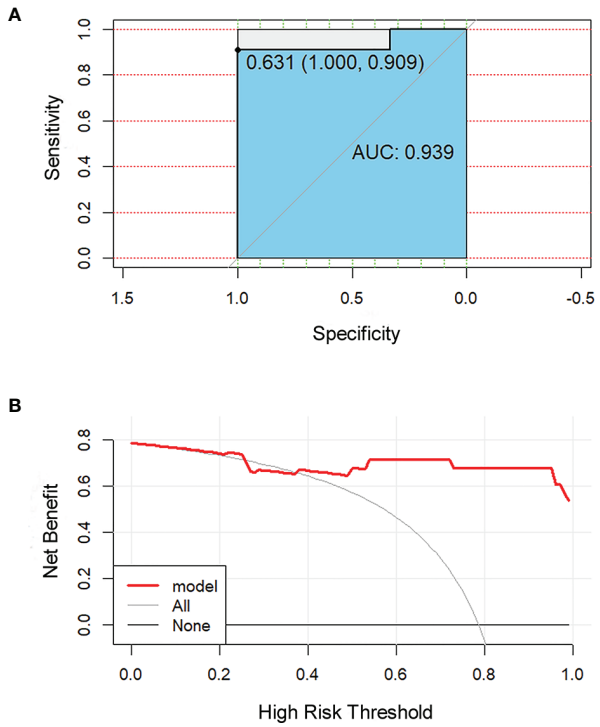


FIGURE 7 | Receiver operating characteristic (ROC) and decision curve for the model. The ROC curve **(A)** shows the prediction accuracy of the model (AUC 0.94, 95% CI 0.85–0.99, $p < 0.001$). In the decision curve **(B)**, the red line represents the radiomics nomogram. The gray line represents the assumption that all patients were responders. The thin black line represents the assumption that all patients were non-responders.

segmentation method proposed by some scholars is an operator independent method, which can help clinicians to outline containing both metabolic and morphological information (37). Some studies have shown that, compared with traditional manual segmentation radiomics approaches, the survival model of automatic tumor segmentation based on neural network segmentation shows significantly higher predictive power (38). However, there is not enough evidence to prove that automatic segmentation can replace manual segmentation. We have only preliminarily discussed the prediction performance of manual segmentation. We will test accurate automatic segmentation methods for reliable segmentation.

LIMITATIONS AND CONCLUSION

First, this study is based on a small-sample clinical trial to establish a training model. For the training sample size, some scholars suggest that for multiple regression, each prediction variable needs at least 10 observations to produce a reasonable and stable estimate (39). In our study, three features are selected as the final model, and the minimum data size is 30. Due to the limited sample size, in order to make the model more accurate, we use all the collected cases to

establish the training set. Like the previous small sample study (40), this study uses bootstrap resampling method to extract multiple samples from the original samples, generate simulation data, and compare these data results with the actual results to prove the robustness of the statistical data. In order to better understand whether the different short-term efficacy evaluations obtained by the two imaging techniques are related to the patient's long-term prognosis, we plan to increase the follow-up time to calculate the 3- or 5-year overall survival rates, and respective disease-free survival rates. Finally, our model needs to be validated through further prospective research, although it showed a high predictive power. Though we have accord with the requirement of the minimum sample size in training set and used LASSO regression to avoid overfitting (41), we still need to increase the sample size and set the training set to further avoid overfitting. We also plan to include more patients in further prospective studies, wherein some patients will continue to serve as the training set to increase the repeatability of the prediction performance of the model by expanding the sample size, while others will serve as the validation set to verify the accuracy of the verification model. The study provides further evidence to use PET to evaluate the efficacy in NSCLC. Our results show that early ^{18}F-FDG-PET could be particularly useful for identifying early responders, allowing clinicians to avoid overtreatment, and that the radiomics nomogram could be an important technique for the prediction of short-term efficacy in patients with NSCLC, which might enable an improved and precise treatment (41).

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Affiliated Hospital of Southwest Medical University Clinical trial ethics committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conception and design: SL, H-WP, and Y-QJ. Administrative support: SL and J-BW. Provision of study materials or patients: YC and YZ. Collection and assembly of data: HC, X-XS, and QG. Data analysis and interpretation: Y-QJ and HC. Manuscript writing: Y-QJ and H-WP. Final approval of manuscript: SL. All authors contributed to the article and approved the submitted version

FUNDING

This work was supported by the grants from the National Natural Science Foundation of China (no. 81201682), the Scientific Research Foundation of the Luzhou Science and Technology Bureau (no. 2016LZXNYD-J05), and the Southwest Medical University Foundation (no. 201617).

ACKNOWLEDGMENTS

We thank the patients who voluntarily participated in this study; Pei-Rong Ren for assessing the quality of the radiation therapy treatment plans; and Na Wang and Li-Shi Yang for providing language help and writing assistance. We would like to thank Editage (www.editage.cn) for the English language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.590836/full#supplementary-material

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
2. Ouyang W, Yu J, Nuerjiang S, Li Z, Wang D, Wang X, et al. Stereotactic body radiotherapy improves the survival of patients with oligometastatic non-small cell lung cancer. *Cancer Med* (2019) 8(10):4605–14. doi: 10.1002/cam4.2366
3. Ettinger DS, Wood DE, Akerley W, Bazhenova LA, Borghaei H, Camidge DR, et al. Non-small cell lung cancer, version 1.2015. *J Natl Compr Canc Netw* (2014) 12(12):1738–61. doi: 10.6004/jnccn.2014.0176
4. Linda A, Trovo M, Bradley JD. Radiation injury of the lung after stereotactic body radiation therapy (SBRT) for lung cancer: a timeline and pattern of CT changes. *Eur J Radiol* (2011) 79(1):147–54. doi: 10.1016/j.ejrad.2009.10.029
5. Takeda A, Kunieda E, Fujii H, Yokosuka N, Aoki Y, Oooka Y, et al. Evaluation for local failure by 18F-FDG PET/CT in comparison with CT findings after stereotactic body radiotherapy (SBRT) for localized non-small-cell lung cancer. *Lung Cancer* (2013) 79(3):248–53. doi: 10.1016/j.lungcan.2012.11.008
6. Pastis NJ Jr, Greer TJ, Tanner NT, Wahlquist AE, Gordon LL, Sharma AK, et al. Assessing the usefulness of 18F-fluorodeoxyglucose PET-CT scan after stereotactic body radiotherapy for early-stage non-small cell lung cancer. *Chest* (2014) 146(2):406–11. doi: 10.1378/chest.13-2281
7. Manus MPM, Hicks RJ, Matthews JP, McKenzie A, Rischin D, Salminen EK, et al. Positron Emission Tomography Is Superior to Computed Tomography Scanning for Response-Assessment After Radical Radiotherapy or Chemoradiotherapy in Patients With Non–Small-Cell Lung Cancer. *J Clin Oncol* (2003) 21(7):1285–92. doi: 10.1200/jco.2003.07.054
8. Yuan M, Zhang YD, Pu XH, Zhong Y, Li H, Wu JF, et al. Comparison of a radiomic biomarker with volumetric analysis for decoding tumour phenotypes of lung adenocarcinoma with different disease-specific survival. *Eur Radiol* (2017) 27(11):4857–65. doi: 10.1007/s00330-017-4855-3
9. Flechsig P, Frank P, Kratochwil C, Antoch G, Rath D, Moltz J, et al. Radiomic Analysis using Density Threshold for FDG-PET/CT-Based N-Staging in Lung Cancer Patients. *Mol Imag Biol* (2017) 19(2):315–22. doi: 10.1007/s11307-016-0996-z
10. Yu W, Tang C, Hobbs BP, Li X, Koay EJ, Wistuba II, 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* (2018) 102(4):1090–7. doi: 10.1016/j.ijrobp.2017.10.046
11. Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The Eighth Edition Lung Cancer Stage Classification. *Chest* (2017) 151(1):193–203. doi: 10.1016/j.chest.2016.10.010
12. Xiang L, Zhang JW, Lin S, Luo HQ, Wen QL, He LJ, et al. Computed Tomography-Guided Interstitial High-Dose-Rate Brachytherapy in Combination With Regional Positive Lymph Node Intensity-Modulated Radiation Therapy in Locally Advanced Peripheral Non-Small Cell Lung Cancer: A Phase 1 Clinical Trial. *Int J Radiat Oncol Biol Phys* (2015) 92(5):1027–34. doi: 10.1016/j.ijrobp.2015.04.019
13. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imag* (2010) 37(1):181–200. doi: 10.1007/s00259-009-1297-4
14. Hain SF, Curran KM, Beggs AD, Fogelman I, O'Doherty MJ, Maisey MN. FDG-PET as a "metabolic biopsy" tool in thoracic lesions with indeterminate biopsy. *Eur J Nucl Med* (2001) 28(9):1336–40. doi: 10.1007/s002590100563
15. Beggs AD, Hain SF, Curran KM, O'Doherty MJ. FDG-PET as a "metabolic biopsy" tool in non-lung lesions with indeterminate biopsy. *Eur J Nucl Med Mol Imag* (2002) 29(4):542–6. doi: 10.1007/s00259-001-0736-7
16. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026
17. Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lammertsma AA, et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography review and 1999 EORTC recommendations. *Eur J Cancer* (1999) 35(13):1773–82. doi: 10.1016/S0959-8049(99)00229-4
18. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* (2012) 30(9):1234–48. doi: 10.1016/j.mri.2012.06.010
19. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
20. Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145
21. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative-feature definitions. *arXiv preprint arXiv:1612.07003* (2016). doi: 10.17195/candat.2016.08.1
22. Depeursinge A, Andrearczyk V, Whybra P, van Griethuysen J, Müller H, Schaer R, et al. Standardised convolutional filtering for radiomics. *arXiv preprint arXiv:2006.05470* (2020). Available at: https://www.researchgate.net/publication/342093302_Standardised_convolutional_filtering_for_radiomics
23. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging* (2012) 30(9):1323–41. doi: 10.1016/j.mri.2012.05.001
24. Fischer BM, Mortensen J, Langer SW, Loft A, Berthelsen AK, Daugaard G, et al. PET/CT imaging in response evaluation of patients with small cell lung cancer. *Lung Cancer* (2006) 54(1):41–9. doi: 10.1016/j.lungcan.2006.06.012
25. Bradley J, Bae K, Choi N, Forster K, Siegel BA, Brunetti J, et al. A phase II comparative study of gross tumor volume definition with or without PET/CT fusion in dosimetric planning for non-small-cell lung cancer (NSCLC): primary analysis of Radiation Therapy Oncology Group (RTOG) 0515. *Int J Radiat Oncol Biol Phys* (2012) 82(1):435–41. doi: 10.1016/j.ijrobp.2010.09.033
26. Dunlap NE, Yang W, McIntosh A, Sheng K, Benedict SH, Read PW, et al. Computed Tomography-Based Anatomic Assessment Overestimates Local Tumor Recurrence in Patients With Mass-like Consolidation After Stereotactic Body Radiotherapy for Early-Stage Non-Small Cell Lung Cancer. *Int J Radiat OncologyBiologyPhysics* (2012) 84(5):1071–7. doi: 10.1016/j.ijrobp.2012.01.088
27. Birim O, Kappetein AP, van Klaveren RJ, Bogers AJ. Prognostic factors in non-small cell lung cancer surgery. *Eur J Surg Oncol* (2006) 32(1):12–23. doi: 10.1016/j.ejso.2005.10.001
28. Vesselle H, Schmidt RA, Pugsley JM, Li M, Kohlmyer SG, Vallires E, et al. Lung cancer proliferation correlates with [F-18]fluorodeoxyglucose uptake by positron emission tomography. *Clin Cancer Res* (2000) 6(10):3837–44.
29. Rajendran JG, Schwartz DL, O'Sullivan J, Peterson LM, Ng P, Scharnhorst J, et al. Tumor hypoxia imaging with [F-18] fluoromisonidazole positron emission tomography in head and neck cancer. *Clin Cancer Res* (2006) 12(18):5435–41. doi: 10.1158/1078-0432.CCR-05-1773
30. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med* (2011) 52(3):369–78. doi: 10.2967/jnumed.110.082404
31. Ahn HK, Lee H, Kim SG, Hyun SH. Pre-treatment (18)F-FDG PET-based radiomics predict survival in resected non-small cell lung cancer. *Clin Radiol* (2019) 74(6):467–73. doi: 10.1016/j.crad.2019.02.008
32. Ohri N, Duan F, Snyder BS, Wei B, Machtay M, Alavi A, et al. Pretreatment 18F-FDG PET Textural Features in Locally Advanced Non-Small Cell Lung Cancer: Secondary Analysis of ACRIN 6668/RTOG 0235. *J Nucl Med* (2016) 57(6):842–8. doi: 10.2967/jnumed.115.166934
33. Antunovic L, De Sanctis R, Cozzi L, Kirienko M, Sagona A, Torrisi R, et al. PET/CT radiomics in breast cancer: promising tool for prediction of pathological response to neoadjuvant chemotherapy. *Eur J Nucl Med Mol Imag* (2019) 46(7):1468–77. doi: 10.1007/s00259-019-04313-8
34. Peng H, Dong D, Fang MJ, Li L, Tang LL, Chen L, et al. Prognostic Value of Deep Learning PET/CT-Based Radiomics: Potential Role for Future Individual Induction Chemotherapy in Advanced Nasopharyngeal Carcinoma. *Clin Cancer Res* (2019) 25(14):4271–9. doi: 10.1158/1078-0432.Ccr-18-3065
35. Giannini V, Mazzetti S, Bertotto I, Chiarenza C, Cauda S, Delmastro E, et al. Predicting locally advanced rectal cancer response to neoadjuvant therapy with (18)F-FDG PET and MRI radiomics features. *Eur J Nucl Med Mol Imag* (2019) 46(4):878–88. doi: 10.1007/s00259-018-4250-6
36. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol* (2015) 16(4):e173–80. doi: 10.1016/S1470-2045(14)71116-7
37. Rundo L, Stefano A, Militello C, Russo G, Sabini MG, D'Arrigo C, et al. A fully automatic approach for multimodal PET and MR image segmentation in gamma knife treatment planning. *Comput Methods Programs BioMed* (2017) 144:77–96. doi: 10.1016/j.cmpb.2017.03.011
38. Baek S, He Y, Allen BG, Buatti JM, Smith BJ, Tong L, et al. Deep segmentation networks predict survival of non-small cell lung cancer. *Sci Rep* (2019) 9(1):17286. doi: 10.1038/s41598-019-53461-2
39. Chalkidou A, O'Doherty MJ, Marsden PK. False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review. *PloS One* (2015) 10(5):e0124165. doi: 10.1371/journal.pone.0124165
40. Gill AB, Rundo L, Wan JCM, Lau D, Zawaideh JP, Woitek R, et al. Correlating Radiomic Features of Heterogeneity on CT with Circulating Tumor DNA in Metastatic Melanoma. *Cancers (Basel)* (2020) 12(12):3493. doi: 10.3390/cancers12123493
41. Babyak M. What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosom Med* (2004) 66(3):411–21. doi: 10.1097/00006842-200405000-00021

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 02 March 2021
doi: 10.3389/fonc.2021.609054

Prognostic Value of Deep Learning-Mediated Treatment Monitoring in Lung Cancer Patients Receiving Immunotherapy

*Stefano Trebeschi[1,2,3], Zuhir Bodalal[1,2], Thierry N. Boellaard[1], Teresa M. Tareco Bucho[1,2], Silvia G. Drago[1], Ieva Kurilova[1,2], Adriana M. Calin-Vainak[1,4], Andrea Delli Pizzi[1,5], Mirte Muller[6], Karlijn Hummelink[7], Koen J. Hartemink[8], Thi Dan Linh Nguyen-Kim[1,2,9], Egbert F. Smit[4], Hugo J. W. L. Aerts[1,2,3,10,11] and Regina G. H. Beets-Tan[1,2,12]**

OPEN ACCESS

Edited by:
Ellen Ackerstaff,
Memorial Sloan Kettering Cancer Center, United States

Reviewed by:
Radka Stoyanova,
University of Miami, United States
Harini Veeraraghavan,
Memorial Sloan Kettering Cancer Center, United States

***Correspondence:**
Regina G. H. Beets-Tan
r.beetstan@nki.nl

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 22 September 2020
Accepted: 04 January 2021
Published: 02 March 2021

Citation:
Trebeschi S, Bodalal Z, Boellaard TN, Tareco Bucho TM, Drago SG, Kurilova I, Calin-Vainak AM, Delli Pizzi A, Muller M, Hummelink K, Hartemink KJ, Nguyen-Kim TDL, Smit EF, Aerts HJWL and Beets-Tan RGH (2021) Prognostic Value of Deep Learning-Mediated Treatment Monitoring in Lung Cancer Patients Receiving Immunotherapy. Front. Oncol. 11:609054. doi: 10.3389/fonc.2021.609054

[1] Department of Radiology, Netherlands Cancer Institute - Antoni vanLeeuwenhoek Hospital, Amsterdam, Netherlands, [2] GROW School for Oncology and Developmental Biology, Maastricht, Netherlands, [3] Artificial Intelligence in Medicine (AIM) Program, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States, [4] Affidea, Cluj-Napoca, Romania, [5] Department of Neuroscience, Imaging and Clinical Sciences, Gabriele D'Annunzio University of Chieti, Chieti, Italy, [6] Department of Thoracic Oncology, Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, Netherlands, [7] Department of Pathology, Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, Netherlands, [8] Department of Surgery, Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, Netherlands, [9] Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, Zürich, Switzerland, [10] Radiology and Nuclear Medicine, University of Maastricht, Maastricht, Netherlandsa, [11] CARIM School for Cardiovascular Diseases, University of Maastricht, Maastricht, Netherlands, [12] Department of Radiology, University of Southern Denmark, Odense, Denmark

Background: Checkpoint inhibitors provided sustained clinical benefit to metastatic lung cancer patients. Nonetheless, prognostic markers in metastatic settings are still under research. Imaging offers distinctive advantages, providing whole-body information non-invasively, while routinely available in most clinics. We hypothesized that more prognostic information can be extracted by employing artificial intelligence (AI) for treatment monitoring, superior to 2D tumor growth criteria.

Methods: A cohort of 152 stage-IV non-small-cell lung cancer patients (NSCLC) (73 discovery, 79 test, 903CTs), who received nivolumab were retrospectively collected. We trained a neural network to identify morphological changes on chest CT acquired during patients' follow-ups. A classifier was employed to link imaging features learned by the network with overall survival.

Results: Our results showed significant performance in the independent test set to predict 1-year overall survival from the date of image acquisition, with an average area under the curve (AUC) of 0.69 ($p < 0.01$), up to AUC 0.75 ($p < 0.01$) in the first 3 to 5 months of treatment, and 0.67 AUC ($p = 0.01$) for durable clinical benefit (6 months progression-free survival). We found the AI-derived survival score to be independent of clinical, radiological, PDL1, and histopathological factors. Visual analysis of AI-generated prognostic heatmaps revealed relative prognostic importance of morphological nodal changes in the mediastinum, supraclavicular, and hilar regions, lung and bone metastases, as well as pleural effusions, atelectasis, and consolidations.

Conclusions: Our results demonstrate that deep learning can quantify tumor- and non–tumor-related morphological changes important for prognostication on serial imaging. Further investigation should focus on the implementation of this technique beyond thoracic imaging.

Keywords: artificial intelligence, immunotherapy, checkpoint inhibitors, non small cell lung cancer, treatment monitoring

INTRODUCTION

Recent advancements in the understanding of the tumor-immune cell interactions (1, 2) have enabled the development of novel drugs for the treatment of advanced-stage lung cancer. Immune checkpoint inhibitors, in particular, have been shown to provide sustained clinical benefit to patients, especially in the metastatic setting (3–5).

Metastatic markers that can be used for patient selection (i.e., before the start of treatment), as well as for treatment monitoring (i.e., during treatment), are still under research (6–8). In the context of oncological research, most predictive/prognostic markers are derived from tissue samples, routinely-extracted blood (9), or non-invasive radiological imaging (surrogate imaging markers). Tissue samples derived from biopsies (usually taken from anatomically accessible locations) often fail to account for inter- and intra-lesion heterogeneity, and response assessed during evaluation of tissue samples of only a few lesions does not necessarily mean that all lesions have responded in the same way. Furthermore, serial biopsies during longitudinal follow-up are cumbersome for the patient but also impractical. Regardless of biomarker source, monitoring of response to therapy remains challenging. As such, they are not part of the routine clinical workflow of patients.

Standard clinical imaging provides a non-invasive overview of the entire tumor burden and has the potential to more accurately evaluate the overall response of the patient to the treatment. Yet, imaging evaluation is currently limited to 2-dimensional "subjective" measurements of tumor size changes (10), time-consuming ROI delineation (11, 12), and/or to values approximating metabolic activity (*i.e., SUV values in PET*) (6). By limiting the use of imaging for response evaluation to only these approaches, many (potentially prognostic) imaging characteristics are ignored. For example, as the disease evolves in multiple distal sites, traditional imaging assessment methods would not account for the microenvironment of each lesion, despite the fact that several potential prognostic factors (e.g., angiogenesis, inflammation, and lymphocytic infiltration) likely depend on that environment (13). Since immunotherapy is a systematic treatment modality, changes indicating response are not limited to one location but can occur all over the body. This is particularly relevant in patients treated with anti PD-1 blockade where lymphadenopathy (14, 15), parenchymal inflammations, edema (16, 17), and compression atelectasis (18), can be observed. Ideally, during image response evaluation these conditions, together with tumor growth, should be monitored and quantified as they might hold valuable prognostic information.

Using artificial intelligence (AI), treatment monitoring tools can be built, capable of rapidly assessing gross morphological changes between two (or more) follow-up images of the same patient (19), in a fully-automatic manner, completely independent of human input. In this context, image registration can be used as the basis for such a method. At its core, image-to-image registration is the process of establishing a voxel-wise match between two radiological images. By establishing a match, we can measure voxel-level differences between corresponding objects represented in the images quantitatively. While conventional registration techniques are very limited for this application, deep learning-based methods have shown promise in image-to-image registration (20). There are three main advantages to using deep learning-based image registration as the core technique. The first advantage is that registration networks are trained to match a pair of images, voxel-wise. This creates a network that is explicitly trained to quantify differences between two images. By leveraging its internal features, we can effectively obtain feature vectors that represent these voxel-wise changes. These vectors can be used for classification purposes. The second advantage of using image registration is that it can be trained on large unlabeled datasets (i.e., lacking any kind of manual annotation, such as segmentations or RECIST-like measurements), while not compromising its ability to model voxel-wise details, that are likely lost in a classical unsupervised autoencoder approach. The third advantage of using image-to-image registration is that, unlike standard RECIST, such a method could be fully automatic and not require any manual input (e.g., two-dimensional diameter measurements), and not be limited to changes in the tumor size, but it would also account for global morphological changes, whether tumor-associated or not, throughout the body. Applying an image-registration-based AI algorithm in oncological follow-up imaging enables us to develop a novel method that can accurately measure gross morphological changes during treatment. Quantitative measurements of these changes can then be used for prognostication.

This study aims to investigate the potential prognostic value of AI-mediated monitoring on CT scans in non-small cell lung cancer (NSCLC) patients receiving anti-PD-1 immune checkpoint blockade. Relying on existing technical research on image-to-image registration, we hypothesize the existence of quantitative imaging features describing a set of gross morphological changes during treatment that hold prognostic

value. To test this hypothesis, we developed a deep learning network for thoracic image-to-image registration and studied the prognostic value of features learned by the network in NSCLC patients being treated with PD-1 blockade.

MATERIALS AND METHODS

Study Cohort

For this study, we retrospectively included patients with stage IV NSCLC treated with anti-PD1 monotherapy within The Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital (NKI-AVL; Amsterdam, The Netherlands) between 2014 and 2016. All patients underwent standardized, imaging-based tumor response assessment with contrast-enhanced computed tomography (CT), with follow-up (FU) intervals of 8 to 12 weeks (**Supplement 1**). We retrieved all available FU scans within the first two years of treatment, together with a baseline scan (BL) performed 8 weeks before and up to 1 week after start of treatment. To encode pre-treatment tumor spread, a pre-baseline scan (PBL), defined as the first available scan before BL, was also retrieved when available. The exact dates of each scan were recorded with respect to the start of treatment (in days). Patients with only one scan available throughout the entire treatment regimen, or whose scan would not fully cover the thorax, were excluded from the analysis. The cohort was divided into a discovery and independent test set based on the patient identifier: patients with even ID numbers were assigned to the discovery set, patients with odd ID numbers were assigned to the independent test set. The study was carried out at the NKI-AVL with the approval of the local Institutional Review Board (IRBd19-083). This cohort is a longitudinal expansion of a previously described NSCLC cohort (11).

Image Acquisition

The CT scans were performed by either covering the chest or covering the chest and abdomen using multi-slice CT equipment (Toshiba Aquilion CX, Minato, Tokyo, Japan; Siemens Somatom Sensation Open, Erlangen, Germany) with a tube voltage of 120 kVp, slice thickness of 1 mm, and in-plane resolution of 0.75 x 0.75 mm. The bolus injection was performed at 3 ml/s (Omnipaque 300, GE Healthcare, Chicago, Illinois, US) not pre-warmed, with a total amount based on the patient weight + 40 cc (minimum of 90 cc and maximum of 130 cc) followed by a saline flush of 30 cc. The chest CT examinations were performed 40 s after contrast injection, whereas the chest and abdomen examinations were performed at 70 s.

Data Curation

Radiological datasets are often heterogeneous. To mitigate differences in radiological image acquisition, all CT scans were cropped between the liver and the lower neck region using the method proposed by Zhang et al. (21), and linearly resampled to 2 mm isotropic voxel size. Hounsfield units were clipped between −120 (fat) and 300 (cancellous bone) and rescaled between 0 and 1. CT scans were further cropped to 192 x 192 x 160 voxels from the center point in order to provide the network with regular image shapes during training.

AI-mediated Quantitative Treatment Monitoring

To harness AI for quantitative treatment monitoring, we developed a 3-dimensional convolutional neural network to perform image-to-image registration between subsequent follow-ups of the same patient (architecture shown in **Figure 1**), based on the research of Balakrishnan et al. (22) and Zaho et al. (23). The network comprised of two subsequent parts: the first performing affine registration aimed to provide alignment of the scans (i.e., to correct for different patient positions), the second section performing deformable registration and aimed to identify morphological changes during the course of the treatment (i.e., longitudinal tracking).

Architecture-wise, the first part of the network consisted of a VGG-like network comprised of a series of five convolutional blocks, and two fully-connected layers, regressing the 12 parameters of the affine transform. The output transform of the network was applied to the moving image, concatenated to the fixed image, and fed into the second part of the network. The second part of the network followed a U-Net architecture (24), and it aimed to quantify non-linear anatomical differences between the input scans. This consisted of an encoding section, comprising 4 convolutional blocks downsampling the images by half the size *via* striding, a convolutional latent space with stride of 1, and four deconvolutional blocks each upsampling the inputs by double the size *via* striding. Skip connections were implemented between encoding and decoding layers following the implementation in the original paper. The network was trained to minimize the correlation coefficient loss (23). Unlike standard measurements of classical registration procedures, this loss is easy to compute in the continuous case. Three penalties were also employed to mitigate for unlikely morphological deformations, each weighted 1/10 in the final loss. Adam optimizer was used during training, with an initial learning rate of 8×10^{-5}. A curriculum learning scheme was implemented during training, such that the loss would be computed on a smoothed version of the images. The smoothing was implemented *via* average pooling, starting with a kernel size of 9, and reduced by 3 at epochs 100, 150, and 175. Batch size was set to 2. To mitigate negative effects resulting from the small batch size, group normalization was employed instead of batch normalization. **Figure 1** shows a detailed overview of the model loss used. The network was trained on a publicly available dataset of 1010 patients of the lung image database consortium (25–27) with 10% hold out during training to control for overfitting (i.e., patients whose ID were multipliers of 10 were held out). Our code can be found online[1].

[1] Github: github.com/nki-radiology/PAM

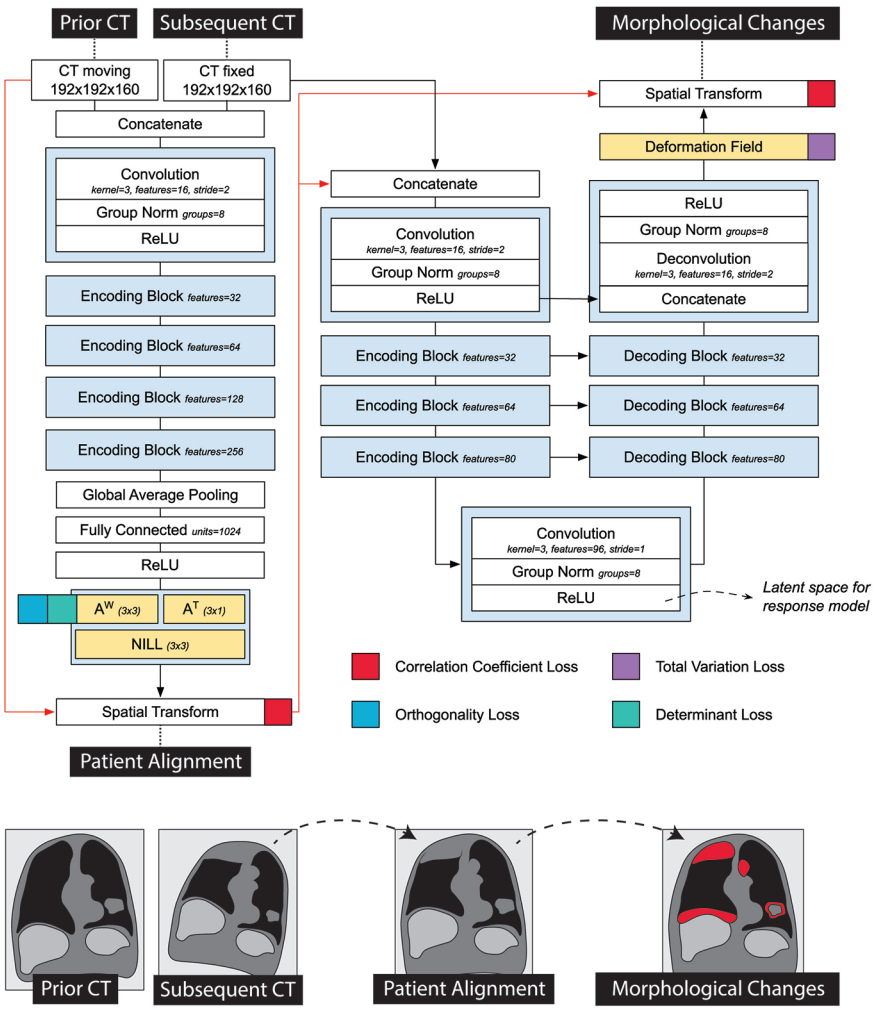


FIGURE 1 | Detailed representation of the registration network used in the prognostic AI-monitoring framework.

Prognostication Through Quantitative Monitoring

To explore the prognostic value of AI-mediated treatment monitoring, we trained a random forest classifier (28) (RFC), with wrapper feature selection, to predict survival based on network imaging features extracted from pairs of subsequent follow-up scans. More specifically, the RFC was trained longitudinally, on pairs of subsequent scans, to predict whether the patient would survive 1 year from the date of the latest of the two scans (see **Figure 2**). The input of the RFC consisted of 96 feature maps from the latent space of the decoder that represented the morphological changes between the prior and the subsequent scan. These are the deepest features found in the middle layer of the second section of the network—the one handling deformable registration. These features come in tensor shape, hence the name feature *maps*. For classification purposes, it is standard to transform the feature maps of the network to a feature vector, to be fed into a classifier. Global average pooling is the technique commonly used to create a feature vector out of a set of feature maps: each entry of the feature vector is the average value of the corresponding feature map. Alongside the global average pooling, we also included standard deviation, skewness and kurtosis, as we deemed the feature maps too large to be represented just by the mean activation—1,000 values per feature map, compared to 49 of a classical ResNet architecture.

To correct for temporal discrepancies (e.g., differences in time between follow-ups), the amount of days elapsed between the two scans, and the days elapsed since the start of treatment were also fed to the RFC. Furthermore, morphological changes should be order invariant: the differences estimated between image A and B should be the same as the differences between image B and A. To provide order invariance, we applied element-wise multiplication of the feature maps generated by swapping the input scans. More specifically, we computed the feature maps for the scan pair prior-to-subsequent, and the feature maps for the pair subsequent-to-prior. Then we multiplied them together, element-wise. The multiplication preserved only those changes that were detected in both directions, therefore providing order

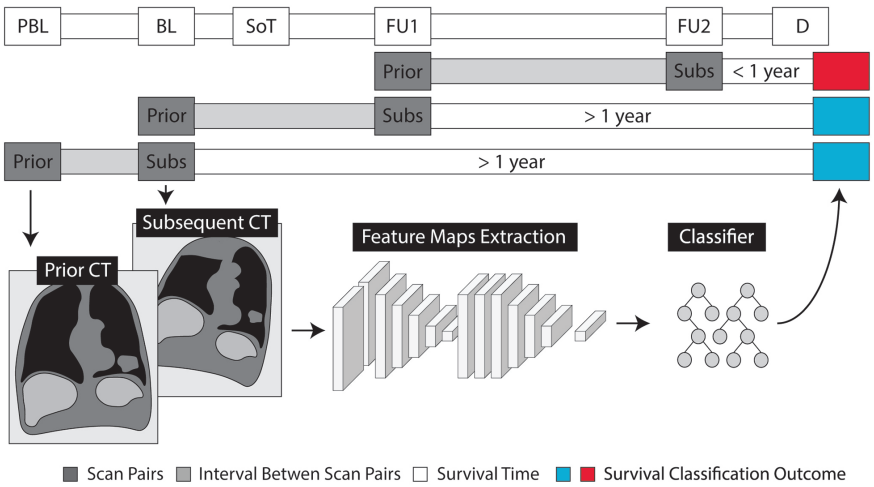


FIGURE 2 | Schematic representation of the evaluation of prognostic values through quantitative monitoring. Radiological examinations are shown as pre-baseline (PBL), baseline (BL) and follow-up (FU), with respect to the start of treatment (SoT). Prediction of survival is made based on the time of death (D). For each pair of subsequent scans, we label the earlier one as *prior* and the subsequent as *subsequent* (Subs).

invariance to our model. The discovery set was used for training, while testing was performed on the independent validation set. Both the registration network and the random forest classifier were trained on the partitioned data, at once, with their respective default parameters—no cross-validation or model selection was performed.

Prognostic Heatmaps

Occlusion sensitivity was employed to visualize the parts of the image that were deemed prognostic of the outcome (29). The main idea of the occlusion algorithm is based on the assumption that removing a *predictive* section/region from the original image will change the algorithm prediction substantially. In contrast, by removing a *non-predictive* section/region from the original image, the algorithm prediction will stay unchanged. We occluded a section (or patch) of the input image presented to the RFC. The prognostic value of that patch is then computed as the difference of the RFC survival score produced by the occluded image vs the original unoccluded one. The resulting prognostic map is the result of the algorithm scrolling the ROI through the image, and repeating the procedure. This was filtered with the gross morphological changes map to produce a prognostic map of the gross morphological changes used for visual interpretation. Details of the algorithm reported in **Supplement 1**. Visual assessment of the resulting prognostic maps was carried out by an expert reader (T.N.B., board-certified radiologist, 2 years experience in thoracic imaging at a tertiary oncologic center), blinded to all clinical parameters, including survival. All scan pairs were assessed with the prognostic maps overlaid on top. The reader was tasked to identify the areas of activation (i.e., hot spots) in the scan pair, and report them categorized as tumor-related areas, secondary comorbidities, and general anatomical areas. Tumor-related areas and secondary comorbidities, which were not highlighted in the prognostic map, were recorded separately.

Independence From Known Prognostic Factors

To test the independence of our AI model, we ran a multivariate analysis against known prognostic factors. Age and pathological cancer subtypes were extracted directly from the anonymized patient records. Changes in tumoral burden were computed based on the available manual segmentations of the total tumor—i.e., all visible and segmentable lesions in the body, except for bone and brain. To ensure comparability with 2D measurements from standard RECIST criteria, volumes were converted to pseudo-diameters *via* $d = \sqrt[3]{(6V/\pi)}$, where V is the total tumoral burden. This computes the diameter of the sphere equi-volumetric to the total tumor burden. Tumor PD-L1 expression scoring was performed according to the instruction manual of the qualitative immunohistochemical assay developed as a complementary diagnostic tool for nivolumab (PD-L1 IHC 22C3 pharmDx, Dako, Carpinteria, CA). PD-L1 expression levels were determined by observing complete circumferential or partial linear expression (at any intensity) of PD-L1 on the plasma cell membrane of viable tumor cells. In parallel, the pattern of staining in CD4 stained slides, which also stain $CD4^+$ lymphocytes and macrophages, was evaluated and compared to PD-L1 stained slides in order to avoid false positive assessment due to PD-L1 expressing macrophages in between tumor cells. Assessment of expression levels was performed in sections that included at least 100 tumor cells that could be evaluated.

Statistical Analysis

To assess prognostic performance, the area under the receiver operating curve (ROC-AUC) was used. Confidence intervals

were estimated *via* bootstrapping performed using repeated sampling with replacement (10000 times). Statistical significance was assessed *via* Mann-Whitney-U test. Kaplan Meier models were employed for survival analysis. Statistical significance of survival metrics was assessed *via* log-rank test. Prognostic (treatment monitoring) performance was quantified in terms of overall survival from the date of the scan. Biomarker performance was quantified in terms of overall survival and durable clinical benefit (complete or partial response, or stable disease, for at least 6 months) from the start of treatment. Cox-Hazards models were used for comparison of known prognostic factors.

RESULTS

Study Cohort

A total of 152 patients, 903 CT scans, and 611 scan matched pairs of subsequent CT scans were included in this study (see **Figure 3A**). The discovery set consisted of 73 patients (and 276 scan pairs), while the independent validation set had 79 patients (and 335 scan pairs). The median age of the entire cohort was 64.4 (IQR 57.8–68.9), with a higher prevalence of males (57.9%). Adenocarcinoma was the most common subtype, reported in 61% of the cohort. No differences in clinical characteristics were encountered between discovery and validation set, except for survival. In comparison to the discovery set, the independent validation set had 180 days longer overall survival, and 101 days longer progression-free survival. Imaging-wise, we collected 129 pre-baselines (PBL; 14.3%), 149 baselines (BL; 16.5%), 135 first follow-ups (FU1; 15.0%), and 103 second follow-ups (FU2; 11.4%). Subsequent follow-ups (FU3+) constituted the remaining 42.9% of the dataset (N = 387). Time-wise, BL scans were acquired on average 26 days before the start of treatment (IQR, 37–14), while the first FU scan, 68 days after (IQR, 46–77). Subsequent follow-ups were made on average every 77 days (IQR, 55–95). Acquisition of non-contrast-enhanced PET-CT instead of contrast-enhanced CT was the main reason for the lack of imaging during follow-up. Further patient characteristics in **Table 1**.

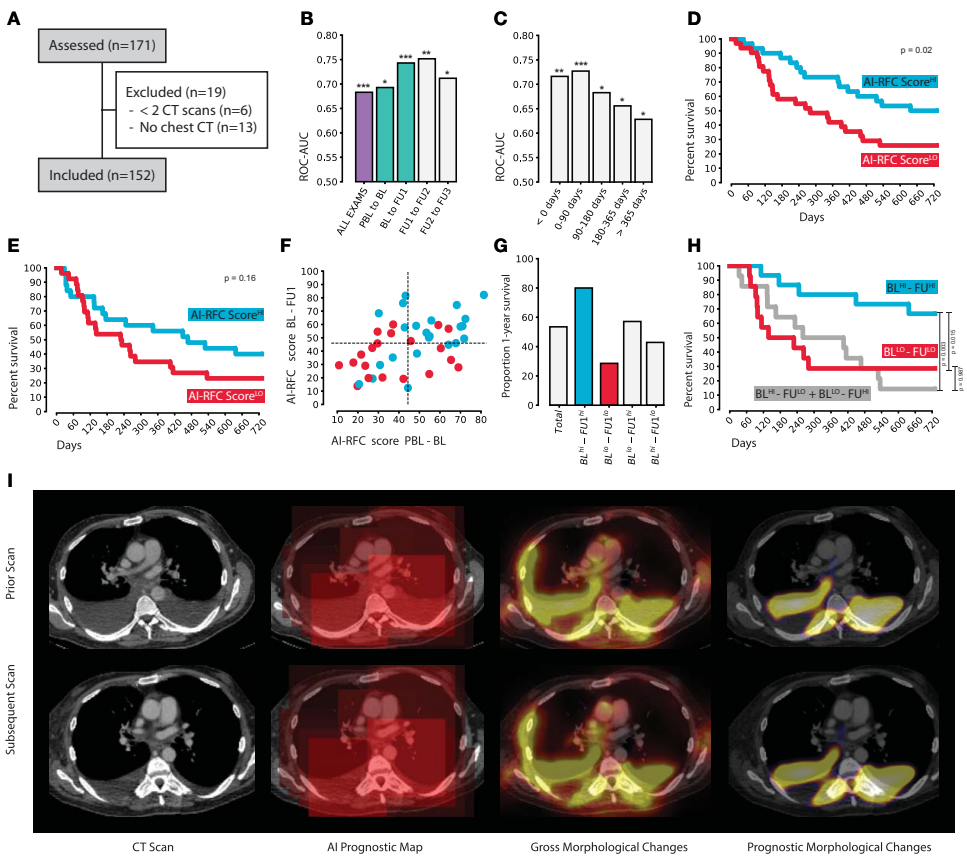


FIGURE 3 | **(A)** CONSORT diagram. **(B)** 1-year survival classification performance on the independent validation set, with respect to the clinical follow-up routine (highlighted in green the ROC-AUC of the scan pairs used for the 2-years survival analysis) and **(C)** corrected by time. **(D)** 2-years Kaplan-Meier curves of the RFC survival score of BL-FU1 and **(E)** PBL-BL. **(F)** Combination of the PBL-BL and BL-FU1 RFC survival scores with **(G)** enrichment of each of the four quadrants **(F, H)** survival of each of the four quadrants. **(I)** Example of the occlusion sensitivity method used for AI explainability and visualization. * indicates $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 1 | Patient and imaging data characteristics.

	Entire Dataset	Discovery Set	Validation Set
Patient Characteristics			
N	152	73	79
Age [median, IQR]	64.4 (57.8–68.9)	64.5 (58.3–69.2)	64.2 (56.2–68.2)
Gender [N, %]	88 Males (57.9%)	44 Males (60.3%)	44 Males (55.7%)
Survival [median days]	341	269	449
Adenocarcinoma [N, %]	92 (60.5%)	46 (63.0%)	19 (26.0%)
Squamous [N, %]	35 (23.0%)	46 (58.2%)	16 (20.3%)
Radiological Follow-up			
All scan pairs	611	276	335
— PB-BL to BL [N, %]	93 (15.2%)	42 (15.2%)	51 (15.2%)
— BL to FU1 [N, %]	116 (19.0%)	55 (19.9%)	61 (18.2%)
— FU1 to FU2 [N, %]	100 (16.4%)	50 (18.1%)	50 (14.9%)
Days b/w scans in any scan pairs [median, IQR]	77.0 (55.0–95.0)	77.0 (50.0–97.2)	77.0 (56.0–94.0)
— Pre-baseline to baseline [median, IQR]	76.0 (55.0–113.0)	75.0 (47.0–114.8)	76.0 (61.0–98.0)
— Baseline to follow-up 1 [median, IQR]	85.5 (69.0–105.0)	86.0 (68.5–107.0)	85.0 (70.0–104.0)
— Follow-up 1 to follow-up 2 [median, IQR]	57.0 (44.0–78.2)	53.5 (43.0–75.0)	72.0 (47.5–83.5)
Days b/w treatment start and BL [median, IQR]	−26.0 (−37.0 to −14.0)	−25.0 (−34.8 to −12.2)	−27.0 (−37.0 to −14.0)
Days b/w treatment start and FU1 [median, IQR]	68.0 (46.0–77.2)	67.0 (46.5–73.0)	68.0 (46.0–78.0)

Image Registration Performance

We evaluated the performance of the registration algorithm merely to identify the cases where the registration algorithm failed. The evaluation of a registration algorithm is usually performed by evaluating the distance between two known corresponding landmarks in the registered image. This can be done automatically, in a circular fashion. Namely, by selecting N random points in an image, we can transform them to their new coordinates in the target image, and back, using the registration functions T_{AB} to represent the transformation from source to target, and T_{BA} as the transformation from target to source. Ideally, these should be the inverse of one another. Practically however there is a registration error propagating from source to target and back. We estimate this error to be proportional to the euclidean distance between N and $T_{BA}(T_{AB}(N))$. It is not exactly the registration error, as this depends on two subsequent dependent registration steps. However, as registration is merely the auxiliary task in our model, a full evaluation of the registration procedure—also in terms of architecture and network components—is beyond the scope of this study. The purpose of this analysis is to analyze the worst cases, i.e., the failures of the algorithm.

We ran the evaluation for all scan pairs, with 100 randomly generated points that were transformed from prior to subsequent, and back to prior. The resulting error was 1.67 cm, on average (CI: 0.87–3.18). We selected for visual inspection the three worst cases, with error 4.54, 3.76 and 3.75 cm, respectively (see figure in **Supplement 2**). These can be considered the closest case of failure of the algorithm. In each of these cases, we can notice the presence of unlikely deformation, like in the heart or the thoracic wall. Although a penalty was set to deter this behavior, we would refrain from increasing it, as it might limit the ability of the network to model other deformations. The strength of the algorithm is represented by the classifier able to distinguish informative deformations from non-informative ones. Overall, in other locations of the image, the registration was still successful in matching anatomical structures properly.

Prognostic Performance

We fed pairs of subsequent follow-up scans to our network trained for (CT chest) image-to-image registration, and trained a random forest classifier (RFC) on its feature maps to investigate the prognostic value of the imaging features learned by the network. Overall results of the RFC survival score on the independent validation set show an AUC of 0.68 (N = 335, CI: 0.62–0.74, $p < 0.001$) to predict 1-year overall survival from the date of the later scan of the scan pair (see **Figure 3B**). The highest prognostic value can be found for the scan pair BL-FU1, reaching an AUC of 0.74 (N = 61, CI: 0.61–0.86, $p < 0.001$), and for the scan pair FU1-FU2, reaching an AUC of 0.75 (N = 42, CI: 0.58–0.89, $p = 0.002$). A decrease in performance is observed during follow-ups, with a 0.71 AUC (N = 42, CI: 0.50–0.89, $p = 0.02$) for the pair FU2-FU3. None of these differences however reached statistical significance. Interestingly, RFC survival scores on the pair PBL-BL also showed prognostic value (0.69 AUC, N = 51, CI: 0.54–0.83, $p = 0.01$). After the fourth follow-up image, the prognostic performance of the model dropped (0.57 AUC, N = 131, CI: 0.47–0.67, $p = 0.11$). This trend becomes evident when looking at the performance with respect to the days between the later scan in the scan pair, and the start of the treatment (see **Figure 3C**). In this respect, we divided the exam pairs in five groups, based on the time between the day of the later scan, and the day of start of treatment (i.e., before start of treatment, 0–90 days from start, 90–180 days and >365 days), and tested the performance in each group individually. Exam pairs performed before start of treatment showed an AUC of 0.72 (N = 48, CI: 0.57–0.86, $p = 0.006$), between start and 90 days after start of treatment showed an AUC of 0.73 (N = 64, CI: 0.59–0.84, $p < 0.001$), between 90 and 180 days showed an AUC of 0.68 (N = 59, CI: 0.51–0.83, $p = 0.01$), between 180 and 365 days an AUC of 0.66 (N = 89, CI: 0.51–0.79, $p = 0.01$). Exam pairs performed in

the second year of treatment showed an AUC of 0.63 (N = 75, CI: 0.50–0.75, $p = 0.04$). Results summary in **Table 2**.

Biomarker Performance

To investigate the prognostic value of AI-monitoring as a biomarker we ran a survival analysis on the scan pairs closest to the date of treatment start, i.e., PBL-BL and BL-FU1. High and low-risk groups were defined for each scan pair by splitting the RFC survival scores on the median value. The scan pair BL-FU1 offered the highest prognostic performance ($p = 0.02$), with a median survival difference of 357 days (637 vs 280 days median survival respectively, $p = 0.02$, see **Figure 3D**). A similar trend was observed for the PBL-BL pair, with a median survival difference of 239 days (467 vs 228 days median survival, respectively, see **Figure 3E**). This, however, did not reach statistical significance ($p = 0.16$). For durable clinical benefit (6 months progression-free survival from start of treatment), we ran a classification analysis on the same scan pairs. This yielded a significant performance of 0.67 AUC (CI: 0.52–0.80, $p = 0.01$) for the BL-FU1 pair, and a similar trend for the PBL-BL pair (0.61 AUC, CI: 0.44–0.77, $p = 0.10$).

Combination of Multiple Time-points

To investigate the prognostic value of AI-monitoring across multiple time points, we combined the prognostic scores of PBL-BL monitoring, and BL-FU1 monitoring (see **Figures 3F, G**). For this particular analysis, we chose the start of treatment as reference, as differences in follow-up schemas might magnify when combining multiple time-points. Across the subset of patients analyzed (with PBL, BL and FU1 scans available, N = 43), 53% survived 1 year after start of treatment (N = 23). Patients with high expression of prognostic features during the monitoring of both PBL-BL and BL-FU1 (N = 15) showed the highest increase in survival, with enrichment from the baseline of 27% (80% survived 1 year after start of treatment). On the contrary, patients with low prognostic features on both PBL-BL and BL-FU1 (N = 14) showed a diminution from baseline of 24% (29% survived 1-year after start of treatment). A point of interest is to be made for patients showing conflicting prognostic scores between PBL-BL and BL-FU1 (positive-negative and negative-positive, N = 7, respectively). While these groups do not seem to show any deviation from the baseline (50% survived 1-year after start of treatment), further analysis on OS showed comparable results to the negative-negative group ($p = 0.99$) over a longer time span (2 year, see **Figure 3H**). The positive-positive group, on the other hand, kept showing significantly higher OS compared to both negative-negative ($p = 0.01$) and negative-positive ($p = 0.003$) groups.

TABLE 2 | Performance of the AI model in predicting 1 year survival after the date of the CT scan.

	N -	N +	*p*-value	Area under the ROC curve
All	128	207	<0.001	0.68 (CI: 0.62–0.74)
PBL-BL	27	24	0.010	0.69 (CI: 0.54–0.83)
BL-FU1	30	31	<0.001	0.74 (CI: 0.61–0.86)
FU1-FU2	18	32	0.002	0.75 (CI: 0.58–0.89)
FU2-FU3	14	28	0.015	0.71 (CI: 0.50–0.89)
FU3 +	39	92	0.112	0.57 (CI: 0.47–0.67)
With respect to days from start of treatment				
< 0	25	23	0.0057	0.72 (CI: 0.56–0.86)
0–90	33	31	<0.001	0.73 (CI: 0.60–0.84)
90–180	19	40	0.013	0.68 (CI: 0.51–0.83)
180–365	26	63	0.011	0.66 (CI: 0.51–0.79)
365 +	25	50	0.037	0.63 (CI: 0.50–0.75)

Comparison With Known Prognostic Factors

To compare the prognostic value of AI-monitoring against other known clinical prognostic factors, we ran a multivariate cox-hazards survival analysis. Specifically, we compared the RFC prognostic scores to age, cancer subtype, volumetric changes in total tumor burden between BL and FU1, and PDL1 expression at baseline. To mitigate collinearity, we reduced PBL-BL/BL-FU1 scores to a single score by principal component analysis. Complete data was available for 22 patients in the independent validation set. Results showed our RFC survival score preserved statistical significance (0.35 HR, CI: 0.12–0.97, $p = 0.04$) against age (2.69 HR, CI: 1.20–6.05, $p = 0.02$), volumetric change of total tumor burden (2.36 HR, CI: 0.67–8.22, $p = 0.18$), >1% PDL-1 expression (0.26 HR, CI: 0.03–2.22, $p = 0.22$), adenocarcinoma (0.34 HR, CI: 0.03–4.43, $p = 0.41$) and squamous subtype (0.14 HR, CI: 0.01–3.01, $p = 0.21$).

Visual Inspection of Prognostic Maps

The main idea behind predictive maps was to evaluate the predictive value of different regions of the image by removing those regions, one at a time, and estimating the difference in predicted survival. **Figure 3I** shows an example. The input scans are displayed in the first column. The second column shows the prognostic map generated by the occlusion algorithm (**Supplement 1**). The patchy look of the overlay is the result of the cubic ROI, being scrolled around the image. Its intensity values are proportional to the change in predicted survival resulting from occluding that region. The third column is the deformation map, where hotspots correspond to regions of gross morphological changes (i.e., pleural effusion). The fourth column was the visualization presented to the reader. It is the result of the fusion between the prognostic map and the deformation map, and highlights the prognostic changes identified by the algorithm.

At visual inspection, lymph node metastases and lung lesions were common hotspots in the prognostic maps. Nodal metastases were present in 58% of scan pairs (N = 57), and highlighted as prognostic in 81% of the cases (N = 46). The mediastinum contained the most nodal hotspots, being highlighted in 80% of cases, followed by supraclavicular and hilar nodal metastases, highlighted in 67% and 57% of cases respectively. Axillary and pericardial nodal metastases were hotspots in 75% and 50% of cases, but found only in 4 and 2 scan pairs respectively. Large lung masses were found in 45% of scan pairs (N = 39), and highlighted as prognostic in 85% of cases. The same rate was observed for small lung nodules, while being less frequent, found in 30% of the scan pairs (N = 26). Bone metastases were found in 20% of scan pairs (N = 17). Nonetheless, they were deemed prognostic by the algorithm in

82% of cases. Pleural masses, liver metastases and subcutaneous lesions, while being almost exclusively hotspots in the prognostic maps, accounted together for only 13 scan pairs. Among secondary comorbidities, pleural effusion, consolidations and atelectasis were the most common, accounting for 31%, 28% and 20% of scans pairs (N = 27, 24, and 17, respectively). Hotspots were found in 94% cases of atelectasis (N = 16), 93% cases of pleural effusions (N = 25), and 83% cases of non-specific consolidation (N = 20). Pericardial effusions were hotspots in 75% of the times, but found only in 8 cases. Only one case of ascites was reported, which the algorithm also highlighted as prognostic. Hotspots in anatomical regions included the spine in 56% of cases, the thoracic wall in 55% of cases, and various regions in the upper thorax, including periscapular (51%), shoulders (49%), neck (48%), and supraclavicular (45%), with the exception of the axilla, highlighted only in 13% of scan pairs. Normal lung parenchyma was highlighted in 28% of cases. The remaining hotspots include the great vessels (9%) and the breast (4%). Detailed summary reported in **Table 3**.

DISCUSSION

Advanced treatment monitoring through more detailed quantitative descriptors of the overall status of the patient, as visualized on routine imaging scans, could provide valuable prognostic information. Our aim was to investigate the potential prognostic value gained by AI-based treatment monitoring on imaging in NSCLC patients treated by PD-1 checkpoint inhibitors. To test this, we implemented a convolutional neural network for image-to-image registration, and trained it on a large public dataset of chest CT scans. The trained network was then used to longitudinally model gross morphological changes between subsequent scans of NSCLC patients receiving PD1 checkpoint inhibitors. Morphological changes identified by the network were then used to train a classifier to predict 1-year OS from the date of the latest scan.

TABLE 3 | Highlighted areas in the AI-generated prognostic maps.

	ALL	PBL-BL	BL-FU1
Tumor Related			
Lymph Nodes	46/57 (80.70%)	21/27 (77.78%)	25/30 (83.33%)
— Pericardial	1/2 (50.00%)	1/1 (100.00%)	0/1 (0.00%)
— Mediastinal	42/53 (79.25%)	18/25 (72.00%)	24/28 (85.71%)
— Hilar	16/28 (57.14%)	7/12 (58.33%)	9/16 (56.25%)
— Supraclavicular	16/24 (66.67%)	5/10 (50.00%)	11/14 (78.57%)
— Axillary	3/4 (75.00%)	1/2 (50.00%)	2/2 (100.00%)
Large Lung Masses	33/39 (84.62%)	16/20 (80.00%)	17/19 (89.47%)
Small Lung Nodules	22/26 (84.62%)	8/11 (72.73%)	14/15 (93.33%)
Bone Metastases	14/17 (82.35%)	7/7 (100.00%)	7/10 (70.00%)
Pleural Masses	6/6 (100.00%)	3/3 (100.00%)	3/3 (100.00%)
Liver Metastases	5/6 (83.33%)	2/3 (66.67%)	3/3 (100.00%)
Subcutaneous Lesions	1/1 (100.00%)	—	1/1 (100.00%)
Secondary Comorbidities			
Pleural Effusion	25/27 (92.59%)	12/12 (100.00%)	13/15 (86.67%)
Consolidation	20/24 (83.33%)	10/12 (83.33%)	10/12 (83.33%)
— Post-radiation	3/3 (100.00%)	2/2 (100.00%)	1/1 (100.00%)
Atelectasis	16/17 (94.12%)	9/9 (100.00%)	7/8 (87.50%)
— Post-obstructive	7/8 (87.50%)	4/4 (100.00%)	3/4 (75.00%)
Pericardial Effusion	6/8 (75.00%)	2/3 (66.67%)	4/5 (80.00%)
Ascites	1/1 (100.00%)	—	1/1 (100.00%)
General Anatomical Areas			
Spine	48/86 (55.81%)	26/43 (60.47%)	22/43 (51.16%)
Thoracic Wall	47/86 (54.65%)	25/43 (58.14%)	22/43 (51.16%)
Periscapular	44/86 (51.16%)	20/43 (46.51%)	24/43 (55.81%)
Shoulder	42/86 (48.84%)	23/43 (53.49%)	19/43 (44.19%)
Neck	41/86 (47.67%)	20/43 (46.51%)	21/43 (48.84%)
Periclavicular	39/86 (45.35%)	19/43 (44.19%)	20/43 (46.51%)
Lung Parenchyma	24/86 (27.91%)	13/43 (30.23%)	11/43 (25.58%)
Axilla	11/86 (12.79%)	6/43 (13.95%)	5/43 (11.63%)
Great Vessels	8/86 (9.30%)	5/43 (11.63%)	3/43 (6.98%)
Breast	3/86 (3.49%)	1/43 (2.33%)	2/43 (4.65%)

Our results showed significant performance in the independent test set for the prediction of 1-year OS from the date of image acquisition, with an average AUC of 0.69, and up to 0.75 AUC for the first 3 to 5 months after start of treatment, and 0.67 AUC for durable clinical benefits, suggesting the presence of (AI-quantified) gross morphological changes encoding prognostic value. These results are comparable to state-of-the-art methods, which currently employs laborious and time-consuming segmentation procedures (11, 12). While the field of research has been focusing on single-lesion analysis—leveraging different known factors in cancer growth, including vascularity (30), oxygenation (31), and metabolic activity (32)—our approach offers a novel fully automatic procedure which completely eradicates the need of time-consuming segmentations, and simultaneously offers a way to provide a full picture of the patient status as seen on chest imaging. While this does not preclude the usefulness of the single-lesion approach, it proposes a way for future multi-scale solutions that leverage both single lesion imaging biomarkers as well as whole image approaches that provide general quantitative information about the status of the patient receiving treatment. Research efforts, however, have to be made in order to overcome the bottleneck of manual ROI delineation procedures, either in the form of automatic segmenters (33), or with implicit AI representations of the total tumor burden.

In addition to the statistical analysis of the performance, we investigated the choices the AI made by means of sensitivity occlusion (29). This resulted in a set of prognostic heatmaps, highlighting regions of morphological changes that the AI deemed prognostic relevance. Gross morphological changes in nodal and lung lesions held the highest prognostic value, especially nodal lesions in the mediastinum, hilum, and supraclavicular region. Further results suggested additional prognostic value for morphological changes affecting the lungs, either in the form of compression from the thoracic wall (due to pleural effusion or pleural masses), non-specific consolidations, or atelectasis. These results also seemed to extend to other regions, with ascites and pericardial effusions also being highlighted as prognostic, despite their rare occurrence. The AI seemed to pay particular attention to the skeleton, with the spine being the anatomical region most commonly highlighted by the AI in the prognostic maps, and bone metastases deemed prognostic in most cases where those were present. As common imaging follow-up schemas, such as RECIST (29, 34), do not account for tumor

burden in the bones, our findings suggest that, on the contrary, such phenomena should not be ignored. Further investigations should lead to novel guidelines, which can provide valuable contribution from the imaging beyond diametrical measurements.

Particular attention should also be paid to nodal metastases and nodal growths during treatment. Imaging features of nodal metastases were found already to be correlated with disease progression for NSCLC, melanoma, and head and neck cancer (11, 35), though no distinction was made between the location of the lymph nodes. However, both our findings and the current literature suggest that this information may be of value. This would be especially interesting in the light of regional (tumor-draining) lymph nodes which play a critical role in terms of anti-tumor immunity and priming (36), increased expression of cytokines and checkpoint markers (37), and changes in the immune compartments resulting in a tumor favorable microenvironment (38). A major hurdle that remains in the analysis of lymph nodes is represented by the radiological assessment, often in contrast with the pathological one. Most radiomics studies so far focused on the detection of positive nodal metastases rather than their prognostic values (39–44).

The analysis of lung lesions is far more common. Imaging features from lung lesions have been reported to hold prognostic value for patients receiving immunotherapy in several studies (11, 32, 45–48). Indeed our findings confirm the association between lung lesions and treatment outcome, with about 85% percent of them being hotspots in the AI-generated prognostic maps, independent of size. Most of the studies published so far focus on the analysis of the tumor region and/or the peritumoral boundary, which may hold valuable information regarding tumor vascularization and inflammatory environment. In this study, the proposed AI model monitors the whole image including both the healthy tissue as well as the tumor(s). As the growth of a cancer lesion does not uniquely depend on the genetic makeup, but rather a complex interaction of microenvironmental features and favorable location for seeding, it would not be surprising to establish a link between a comprehensive modeling tool of cancer growth and its biological features. Even in this case however, further research is needed to establish any link between imaging features and tumor biology.

Following our results, we observed an increase in the prognostic performance of the AI resulting from the combination of multiple time points, namely pre-baseline, baseline and first follow-up. This analysis showed good OS for patients with higher AI-survival scores (AI-RFC^{hi}) in both pre-baseline to baseline scan pair, and baseline to first follow-up—and worse OS for the opposite case (AI-RFC^{lo}). Interestingly, patients with contradicting scores (AI-RFC^{hi} for pre-baseline to baseline scan pair, and AI-RFC^{lo} for the baseline to follow-up, and vice versa) showed worse survival, similar to the double negative group. These results suggest the existence of a prognostic combination of pretreatment and early-treatment characteristics, both of which should be accounted for during patient stratification. Further insights could be achieved by more advanced AI methods that would account for larger time spans, or even the entirety of patients' treatment history.

The combined score was demonstrated to be an independent prognostic parameter even when corrected for other known prognostic parameters. This is of particular interest when we consider the possible role of such a tool, for example as an additional input to the tumor board during treatment decision making. Further research is required to study its implementation in the clinical settings.

Limitations and Future Outlook

Our study aimed to monitor AI-measured gross morphological changes between imaging follow-up for survival prediction in NSCLC patients receiving PD1 checkpoint antibodies. In this study, we pre-trained a neural network on a large dataset of chest CT scans, and fine-tuned it for survival on our smaller local immunotherapy data set. Under the current settings, we limited the analysis to chest imaging which, in addition to the chest, frequently included the lower neck and the upper abdomen. While this limitation could hold for lung malignancy, an extension to other cancer types would require this technique to be extended to include the whole body—i.e., the abdomen and, when available, the brain. Moreover, due to the limited amount of data, it was not possible to explore more complex machine learning algorithms for the prediction of survival, nor for more precise visualization of the prognostic maps. Expansion of the dataset, both in terms of patients and in terms of time points, would certainly allow for an increase in performance and better explainability of the AI algorithm. Specifically, an extension of the field of view of the algorithm to the whole body, as well as the usage of parameters other than imaging, could potentially improve the performance of the algorithm to be usable in the clinics. Further clinical validation of the method is also needed. While this study presented a comparison of this method with response evaluation criteria (e.g., changes in total tumor burden) and biomarkers (e.g., PD-L1 expression), the primary objective for future studies should be a comparison with the clinical standard, namely the RECIST criteria. It remains to be investigated whether this method would be complementary to the current radiological response evaluation (i.e., RECIST). Furthermore, additional investigations are required to link biological features to tumor growth and gross morphological changes. Further analysis should also study the effects of different machine acquisition parameters, and the sensitivity of the method to imaging acquisition parameter variability. Looking into the future, we envision that an AI solution could be set up as a clinical decision support system capable of providing information to the treating physician complementary to traditional clinical and pathological input data.

CONCLUSIONS

In this study, we aimed to investigate the potential prognostic value of AI-mediated monitoring in NSCLC patients receiving PD-1 blockade. We hypothesized the existence of quantitative imaging features describing a set of gross morphological changes happening during treatment that hold prognostic information. Our results demonstrate the existence of such factors (as

described by the AI on imaging), that are tumor-related, such as nodal, lung and bone lesions, as well as non-tumor related, such as pleural effusions, atelectasis and non-specific consolidations. Further investigation should focus on the development of more flexible models that can extend beyond thoracic imaging, as well as on external validations.

DATA AVAILABILITY STATEMENT

The data sets presented in this article are not readily available because the restrictions of agreement with the local ethical committee. We can, however, share the feature space on request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

ST: software development. ST, ZB, TT, TDLN-K: conceptualization, experimental design. ZB, TB, TDLN-K: clinical results validation and interpretation. ST, ZB, SD, IK, AC-V, AD, TDLN-K: radiological and clinical data collection and curation. MM, KH: pathological and clinical data collection and curation. KJH, TDLN-K, ES, HA, RB-T: project supervision, resource acquisition. All authors: results and manuscript editing and validation. All authors contributed to the article and approved the submitted version.

FUNDING

This work was also carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. The authors would also like to thank NVIDIA for their kind donation *via* the Academic GPU Grant Program as well as the Maurits en Anna de Kock Stichting for its financial support. TN-K was funded by the Oncologic Imaging Fellowship Grant from the European Society of Radiology.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.609054/full#supplementary-material

REFERENCES

1. Leach DR, Krummel MF, Allison JP. Enhancement of antitumor immunity by CTLA-4 blockade. *Science* (1996) 271:1734–6. doi: 10.1126/science.271.5256.1734
2. Ishida Y, Agata Y, Shibahara K, Honjo T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J* (1992) 11:3887–95. doi: 10.1002/j.1460-2075.1992.tb05481.x
3. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *N Engl J Med* (2015) 373:1627–39. doi: 10.1056/NEJMoa1507643
4. Brahmer J, Reckamp KL, Baas P, Crinò L, Eberhardt WEE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *N Engl J Med* (2015) 373:123–35. doi: 10.1056/NEJMoa1504627
5. Herbst RS, Baas P, Kim D-W, Felip E, Pérez-Gracia JL, Han J-Y, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet* (2016) 387:1540–50. doi: 10.1016/S0140-6736(15)01281-7
6. Teng F, Meng X, Kong L, Yu J. Progress and challenges of predictive biomarkers of anti PD-1/PD-L1 immunotherapy: A systematic review. *Cancer Lett* (2018) 414:166–73. doi: 10.1016/j.canlet.2017.11.014
7. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat Rev Cancer* (2019) 19:133–50. doi: 10.1038/s41568-019-0116-x
8. Ros J, Baraibar I, Vivancos A, Rodon J. Review of immunogenomics and the role of tumor mutational burden as a biomarker for immunotherapy response. *J Immunother Precis Oncol* (2019) 0. doi: 10.4103/jipo.jipo_19_19
9. Wang Z, Duan J, Cai S, Han M, Dong H, Zhao J, et al. Assessment of Blood Tumor Mutational Burden as a Potential Biomarker for Immunotherapy in Patients With Non-Small Cell Lung Cancer With Use of a Next-Generation Sequencing Cancer Gene Panel. *JAMA Oncol* (2019) 5:696–702. doi: 10.1001/jamaoncol.2018.7098
10. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* (2009) 45:228–47. doi: 10.1016/j.ejca.2008.10.026
11. Trebeschi S, Drago SG, Birkbak NJ, Kurilova I, Călin AM, Delli Pizzi A, et al. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann Oncol* (2019) 30:998–1004. doi: 10.1093/annonc/mdz108
12. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* (2018) 19:1180–91. doi: 10.1016/S1470-2045(18)30413-3
13. García-Figueiras R, Baleato-González S, Padhani AR, Luna-Alcalá A, Vallejo-Casas JA, Sala E, et al. How clinical imaging can assess cancer biology. *Insights Imaging* (2019) 10:1–35. doi: 10.1186/s13244-019-0703-0
14. Nishino M, Hatabu H, Hodi FS. Imaging of Cancer Immunotherapy: Current Approaches and Future Directions. *Radiology* (2019) 290:9–22. doi: 10.1148/radiol.2018181349
15. Tirumani SH, Ramaiya NH, Keraliya A, Bailey ND, Ott PA, Hodi FS, et al. Radiographic Profiling of Immune-Related Adverse Events in Advanced Melanoma Patients Treated with Ipilimumab. *Cancer Immunol Res* (2015) 3:1185–92. doi: 10.1158/2326-6066.CIR-15-0102
16. Alessandrino F, Sahu S, Nishino M, Adeni AE, Tirumani SH, Shinagare AB, et al. Frequency and imaging features of abdominal immune-related adverse events in metastatic lung cancer patients treated with PD-1 inhibitor. *Abdom Radiol (NY)* (2019) 44:1917–27. doi: 10.1007/s00261-019-01935-2
17. Johnson DB, Balko JM, Compton ML, Chalkias S, Gorham J, Xu Y, et al. Fulminant Myocarditis with Combination Immune Checkpoint Blockade. *N Engl J Med* (2016) 375:1749–55. doi: 10.1056/NEJMoa1609214
18. Pang Z, Ding N, Dong W, Ni Y, Zhang T, Qu X, et al. Prognostic effects of preoperative obstructive pneumonitis or atelectasis and comparison with tumor size in non-small cell lung cancer. *J Thorac Dis* (2017) 9:768–78. doi: 10.21037/jtd.2017.02.88

19. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin* (2019) 69:127–57. doi: 10.3322/caac.21552
20. Haskins G, Kruger U, Yan P. Deep learning in medical image registration: a survey. *Mach Vision Appl* (2020) 31:1–18. doi: 10.1007/s00138-020-01060-x
21. Zhang P, Wang F, Zheng Y. Self supervised deep representation learning for fine-grained body part recognition. In: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE (2017). doi: 10.1109/isbi.2017.7950587
22. Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. VoxelMorph: A Learning Framework for Deformable Medical Image Registration. *IEEE Trans Med Imaging* (2019) 8:1788–1800. doi: 10.1109/TMI.2019.2897538
23. Zhao S, Lau T, Luo J, Chang EI-C, Xu Y. Unsupervised 3D End-to-End Medical Image Registration with Volume Tweening Network. *IEEE J BioMed Health Inform* (2019) 24.5:1394–1404. doi: 10.1109/JBHI.2019.2951024
24. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer, Cham. 234–41.
25. McNitt-Gray MF, Armato SG III, Meyer CR, Reeves AP, McLennan G, Pais R, et al. The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation. *Med Imaging 2007: Computer-Aided Diagnosis* (2007) 14.12:1464–74. doi: 10.1117/12.713754
26. Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. *Med Phys* (2011) 38:915–31. doi: 10.1118/1.3528204
27. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* (2013) 26:1045–57. doi: 10.1007/s10278-013-9622-7
28. Breiman - Machine learning L. Random forests. In: *Machine learning*. Springer (2001). Available at: http://www.springerlink.com/index/U0P06167N6173512.pdf.
29. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. *Comput Vision – ECCV 2014* (2014) 8689:818–33. doi: 10.1007/978-3-319-10590-1_53
30. Alilou M, Bera K, Vaidya P, Zagouras A, Patil P, Khorrami M, et al. Quantitative vessel tortuosity radiomics on baseline non-contrast lung CT predict response to immunotherapy and are prognostic of overall survival. *Med Imaging 2019: Computer-Aided Diagnosis* (2019) 10950:109501F. doi: 10.1117/12.2513648
31. Tunali I, Tan Y, Gray JE, Katsoulakis E, Eschrich SA, Saller J, et al. Hypoxia-related radiomics predict immunotherapy response of non-small cell lung cancer patients. *Proceedings: AACR Annual Meeting* (2020). doi: 10.1101/2020.04.02.020859
32. Mu W, Qi J, Lu H, Balagurunathan Y, Gillies RJ, Tunali I, et al. Radiomic biomarkers from PET/CT multi-modality fusion images for the prediction of immunotherapy response in advanced non-small cell lung cancer patients. *Med Imaging 2018: Computer-Aided Diagnosis* (2018) 10575:105753S. doi: 10.1117/12.2293376
33. Yan K, Wang X, Lu L, Summers RM. DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. *J Med Imaging (Bellingham)* (2018) 5:036501. doi: 10.1117/1.JMI.5.3.036501
34. Schwartz LH, Litière S, de Vries E, Ford R, Gwyther S, Mandrekar S, et al. RECIST 1.1—Update and clarification: From the RECIST committee. *Eur J Cancer* (2016) 62:132–7. doi: 10.1016/j.ejca.2016.03.081
35. Yu Y, Chen L, Riaz N, Kang JJ, Thor M, Veeraraghavan H, et al. Predictors of Early Response to Immunotherapy in Head and Neck Cancer: A Secondary Clinical and Radiomic Analysis of a Prospective Randomized Trial with Nivolumab. *Int J Radiat Oncol Biol Phys* (2019) 105:S211. doi: 10.1016/j.ijrobp.2019.06.286
36. Zhao X, Kassaye B, Wangmo D, Lou E, Subramanian S. Chemotherapy but not the tumor-draining lymph nodes determine the immunotherapy response in secondary tumors. *iScience* (2020) 23.5:101056. doi: 10.1101/664912
37. Ho WJ, Yarchoan M, Charmsaz S, Munday RM, Danilova L, Sztein MB, et al. Multipanel mass cytometry reveals anti–PD-1 therapy–mediated B and T cell compartment remodeling in tumor-draining lymph nodes. *JCI Insight* (2020) 5:e132286. doi: 10.1172/jci.insight.132286
38. Jones D, Pereira ER, Padera TP. Growth and Immune Evasion of Lymph Node Metastasis. *Front Oncol* (2018) 8:36. doi: 10.3389/fonc.2018.00036
39. Ho T-Y, Chao C-H, Chin S-C, Ng S-H, Kang C-J, Tsang N-M. Classifying Neck Lymph Nodes of Head and Neck Squamous Cell Carcinoma in MRI Images with Radiomic Features. *J Digit Imaging* (2020) 16:1–6. doi: 10.1007/s10278-019-00309-w
40. Sha X, Gong G, Qiu Q, Duan J, Li D, Yin Y. Discrimination of mediastinal metastatic lymph nodes in NSCLC based on radiomic features in different phases of CT imaging. *BMC Med Imaging* (2020) 20:12. doi: 10.1186/s12880-020-0416-3
41. Tan X, Ma Z, Yan L, Ye W, Liu Z, Liang C. Radiomics nomogram outperforms size criteria in discriminating lymph node metastasis in resectable esophageal squamous cell carcinoma. *Eur Radiol* (2019) 29:392–400. doi: 10.1007/s00330-018-5581-1
42. Li M, Zhang J, Dan Y, Yao Y, Dai W, Cai G, et al. A clinical-radiomics nomogram for the preoperative prediction of lymph node metastasis in colorectal cancer. *J Transl Med* (2020) 18:46. doi: 10.1186/s12967-020-02215-0
43. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun* (2020) 11:1–9. doi: 10.1038/s41467-020-15027-z
44. Chen J, He B, Dong D, Liu P, Duan H, Li W, et al. Noninvasive CT radiomic model for preoperative prediction of lymph node metastasis in early cervical carcinoma. *Br J Radiol* (2020) 93:20190558. doi: 10.1259/bjr.20190558
45. Tang C, Hobbs B, Amer A, Li X, Behrens C, Canales JR, et al. Development of an Immune-Pathology Informed Radiomics Model for Non-Small Cell Lung Cancer. *Sci Rep* (2018) 8:1922. doi: 10.1038/s41598-018-20471-5
46. Tunali I, Gray JE, Qi J, Abdalah M, Jeong DK, Guvenis A, et al. Novel clinical and radiomic predictors of rapid disease progression phenotypes among lung cancer patients treated with immunotherapy: An early report. *Lung Cancer* (2019) 129:75–9. doi: 10.1016/j.lungcan.2019.01.010
47. Sun Z, Hu S, Ge Y, Wang J, Duan S, Hu C, et al. Radiomics study for predicting the expression of PD-L1 in non-small cell lung cancer based on CT images and clinicopathologic features. *J Xray Sci Technol* (2020) 1–11. doi: 10.3233/xst-200642
48. Patil P, Bera K, Velu P, Khorrami M, Prasanna P, Velcheti V, et al. A combination of intra- and peritumoral features on baseline CT scans is associated with overall survival in non-small cell lung cancer patients treated with immune checkpoint inhibitors: a multi-agent multi-site study. *Med Imaging 2019: Computer-Aided Diagnosis* (2019) 10950:109500R. doi: 10.1117/12.2513001

Conflict of Interest: Author HJWLA is an advisor for Onc.Ai and Bristol Myers Squibb (BMS), outside of submitted work. Author AD was employed by company Affidea.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 18 March 2021
doi: 10.3389/fonc.2021.559623

Correlation Between Dual-Time-Point FDG PET and Tumor Microenvironment Immune Types in Non-Small Cell Lung Cancer

Jianyuan Zhou[1†], Sijuan Zou[1†], Siyuan Cheng[1], Dong Kuang[2], Dan Li[1], Lixing Chen[1], Cong Liu[2], Jianhua Yan[3] and Xiaohua Zhu[1*]*

[1] Department of Nuclear Medicine and PET, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2] Department of Pathology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3] Shanghai Key Laboratory of Molecular Imaging, Shanghai University of Medicine and Health Sciences, Shanghai, China

OPEN ACCESS

Edited by:
Georgios S. Limouris, National and Kapodistrian University of Athens, Greece

Reviewed by:
Kuangyu Shi, University of Bern, Switzerland
Katsuhiko Shimizu, Kawasaki Medical School Hospital, Japan
Lucia Nogova, University of Cologne, Germany

***Correspondence:**
Xiaohua Zhu
evazhu@vip.sina.com
Jianhua Yan
jianhua.yan@gmail.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *06 May 2020*
Accepted: *01 March 2021*
Published: *18 March 2021*

Citation:
Zhou J, Zou S, Cheng S, Kuang D, Li D, Chen L, Liu C, Yan J and Zhu X (2021) Correlation Between Dual-Time-Point FDG PET and Tumor Microenvironment Immune Types in Non-Small Cell Lung Cancer. Front. Oncol. 11:559623. doi: 10.3389/fonc.2021.559623

Purpose: Dual-time-point ^{18}F-fluorodeoxyglucose positron emission tomography (DTP ^{18}F-FDG PET), which reflects the dynamics of tumor glucose metabolism, may also provide a novel approach to the characterization of both cancer cells and immune cells within the tumor immune microenvironment (TIME). We investigated the correlations between the metabolic parameters (MPs) of DTP ^{18}F-FDG PET images and the tumor microenvironment immune types (TMITs) in patients with non-small cell lung cancer (NSCLC).

Methods: A retrospective analysis was performed in 91 patients with NSCLC who underwent preoperative DTP ^{18}F-FDG PET/CT scans. MPs in the early scan (eSUVmax, eSUVmean, eMTV, eTLG) and delayed scan (dSUVmax, dSUVmean, dMTV, dTLG) were calculated, respectively. The change in MPs (ΔSUVmax, ΔSUVmean, ΔMTV, ΔTLG) between the two time points were calculated. Tumor specimens were analyzed by immunohistochemistry for PD-1/PD-L1 expression and $CD8^+$ tumor-infiltrating lymphocytes (TILs). TIME was classified into four immune types (TMIT I ~ IV) according to the expression of PD-L1 and $CD8^+$ TILs. Correlations between MPs with TMITs and the immune-related biomarkers were analyzed. A composite metabolic signature (Meta-Sig) and a combined model of Meta-Sig and clinical factors were constructed to predict patients with TMIT I tumors.

Results: eSUVmax, eSUVmean, dSUVmax, dSUVmean, ΔSUVmax, ΔSUVmean, and ΔTLG were significantly higher in PD-L1 positive patients (p = 0.0007, 0.0006, < 0.0001, < 0.0001, 0.0002, 0.0002, 0.0247, respectively), and in TMIT-I tumors (p = 0.0001, < 0.0001, < 0.0001, < 0.0001, 0.0009, 0.0009, 0.0144, respectively). Compared to stand-alone MP, the Meta-Sig and combined model displayed better performance for assessing TMIT-I tumors (Meta-sig: AUC = 0.818, sensitivity = 86.36%, specificity = 73.91%; Model: AUC = 0.869, sensitivity = 77.27%, specificity = 82.61%).

Conclusion: High glucose metabolism on DTP ^{18}F-FDG PET correlated with the TMIT-I tumors, and the Meta-Sig and combined model based on clinical and metabolic information could improve the performance of identifying the patients who may respond to immunotherapy.

Keywords: DTP ^{18}F-FDG PET, PD-L1, tumor microenvironment immune types, NSCLC, metabolic parameters

INTRODUCTION

Lung cancer is the leading cause of cancer death in China (1). Non-small cell lung cancer (NSCLC) accounts for more than 80% of all lung cancer cases. Over the last decade, the immune checkpoint inhibitors (ICIs) targeting the programmed death protein 1 (PD-1)/programmed death ligand 1 (PD-L1) axis have shown significant clinical benefit for advanced NSCLC patients. The expression of PD-L1 on tumor cells is considered as a predictive biomarker for the response to anti-PD-1/PD-L1 ICIs (2). However, not all patients with positive PD-L1 expression respond well to immunotherapy. It suggests that other tumor immune microenvironment (TIME) factors may also affect the response to the ICIs (3). In addition to PD-L1 expression, $CD8^+$ tumor-infiltrating lymphocytes (TILs) might play an important role in anti–PD-1/PD-L1 therapies (2). Without $CD8^+$ TILs, it's unlikely that blocking PD-1 or PD-L1 causes any tumor inhibition (4). Characterized by high infiltration of $CD8^+$ cytotoxic lymphocytes, the infiltrated–inflamed TIME has significantly better responses to ICIs (5). Therefore, it was proposed that TIME could be classified into four subtypes based on PD-L1 and $CD8^+$ TILs status (4). Tumors with high PD-L1 expression and the presence of $CD8^+$ TILs are classified as tumor microenvironment immune type I (TMIT-I), a immunologically 'hot' subtype that would likely benefit from anti-PD-1/PD-L1 therapies (6). However, there is no noninvasive method to identify TMIT I tumors, and up to now the overall and dynamic detection of TIME biomarkers is still challenging.

Among the image-based modalities for non-invasive tumor assessment, ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) is the most common one of patients with NSCLC (7). Glucose metabolism is closely related to the characteristics of TIME. As a nutrient, glucose is actively entrapped in neoplastic tissue and tumor-related activated immune cells (8). Previous studies have established the correlation between the MPs [maximum and mean standard uptake value (SUVmax and SUVmean)] of ^{18}F-FDG PET and the expression of immune markers (PD-1, PD-L1 and CD8) in patients with NSCLC (9–16). However, little attention was paid to $CD8^+$ TILs and tumor immune types. The predictive value of tumor metabolism solely based on SUVmax remains weak in patients undergoing ICI immunotherapy. Dual-time-point (DTP) ^{18}F-FDG PET, which reflects the dynamics of glucose metabolism, is expected to be a potential imaging method to reveal the TIME information. Up to date, the correlation between metabolic parameters (MPs) on DTP FDG PET and TMITs in pretreated NSCLC remains unclear.

This retrospective study was conducted to correlate a number of MPs of DTP ^{18}F-FDG PET with immune markers and TMITs in a cohort of pretreated NSCLC patients. We hypothesize that the abundant metabolic information on DTP FDG PET imaging defines the TMITs of NSCLC and helps optimize patient selection for ICIs treatment.

MATERIALS AND METHODS

Patient Population

Patients who underwent pretreatment DTP ^{18}F-FDG PET/CT scans in Tongji hospital for NSCLC diagnosis and staging from December 2014 to December 2017, were retrospectively reviewed. Eligible patients were histologically confirmed with NSCLC, underwent initial PET scan less than 30 days from surgery (or biopsy), with tumor size ≥ 1 cm in diameter. Key exclusion criteria were: patients that received anti-tumor therapy before surgery (or biopsy), and patients whose tumor specimens were not available for immunohistochemistry. This retrospective study was approved by the institutional review board.

DTP ^{18}F-FDG PET/CT Acquisition Protocol and Image Analysis

In each patient, 3.7 MBq/kg FDG was intravenously administered after fasting for at least 6 h. The blood glucose concentration was lower than 200 mg/dL before injecting. PET/CT images were obtained by a PET/CT scanner (Discovery 690 PET/CT, GE) at approximately 60 ± 5 min (early) and 120 ± 5 min (delayed) after FDG administration. Whole body images were obtained from the base of the skull to mid-thigh by means of an integrated PET/CT tomography (5 to 7 bed positions with 3 min per bed position). PET images were attenuation-corrected and anatomically fused with low-dose CT images, and reconstructed onto a 128 × 128 matrix. A low-dose helical CT scan (120 kV, 120 mA, slice thickness, 3.75mm) was performed for anatomical correlation and attenuation correction.

Images were analyzed by two board-certified nuclear medicine physicians. Tumor mass area of increased radiotracer uptake was first identified, then a semi-automated, ellipsoid 3D-isocontour volume of interest (VOI) with threshold of 40%SUVmax was marked around the tumor for the measurement of SUVmax, SUVmean and metabolic tumor volume (MTV). For tumors with low uptake, VOI was obtained by manually delineating the boundary layer by layer along the tumors, then SUVs and MTV (with threshold of 40%SUVmax) were automatically calculated within each VOI. Total lesion glycolysis (TLG) and the change of SUVmax, SUVmean, MTV and TLG were calculated according to the following formula: TLG = SUVmean × MTV, ΔMP = dMP – eMP.

Immunohistochemistry Analysis

Immunohistochemistry was performed using 4 μm sections from a paraffin-embedded tissue block as previously described (17). Briefly, the sections were deparaffinized in xylene and rehydrated in graded ethanol and distilled water. Slides were auto-stained with primary antibodies raised against $CD8^+$ (ZA-0508, ZSGB-BIO, China), PD-1 (Abcam, EPR4877(2), ab137132), PD-L1 (ZA-0629, ZSGB-BIO, China). The PD-L1 immunostaining results were classified into two groups based on staining intensity and proportion of tumor cell positivity (18). Staining intensity was scored as follows: 0, negative staining; 1, weak staining; 2, moderate staining; and 3, strong staining (more intense than alveolar macrophages). Case in which more than 5% of tumor cells displayed a staining intensity ≥2 was considered positive. Case with staining intensity <2 or less than 5% of tumor cells was defined negative. The expressions of PD-1 and $CD8^+$ TIL were evaluated according to the average number of positively stained cells in 3 randomly selected high-power fields in each case. The numbers of $CD8^+$ TILs were classified into two groups based on the median value: $CD8^+$ $TILs^+$ (n ≤100), $CD8^+$ $TILs^-$ (n> 100).

Four tumor microenvironment immune types was classified as reported (3, 6, 19): TMIT-I ($PD\text{-}L1^+$, $CD8^+$ $TILs^+$); TMIT-II ($PD\text{-}L1^-$, $CD8^+$ $TILs^-$); TMIT-III ($PD\text{-}L1^+$, $CD8^+$ $TILs^-$) and TMIT-IV ($PD\text{-}L1^-$, $CD8^+$ $TILs^+$).

Statistical Analysis

Data was analyzed with the SPSS statistical package, MedCalc and R software. The distribution of variables was checked using Shapiro-Wilk test. For continuous data, the differences between two groups were assessed using Mann-Whitney *U* test or Student's t-test. Differences among multi-group were compared using one-way analysis of variance (ANOVA) (with least significant difference method) or Kruskal-Wallis H test, when appropriate. Spearman's correlation coefficients were calculated. The least absolute shrinkage and selection operator (LASSO) algorithm method using 10-fold cross-validation was employed to select the optimal features. Features with non-zero coefficients at the minimum of lambda were selected from the candidate MPs to construct a metabolic signature (Meta-Sig). Multivariate logistic regression analysis with backward stepwise elimination method was performed to construct a combined model, based on clinical factors and the Meta-Sig. The receiver operating characteristic (ROC) curves and DeLong test were used to compare the area under the curves (AUCs) for predicting TMIT I tumors. Furthermore, decision curve analysis (DCA) was used to evaluate the clinical usefulness of the combined model by quantifying the net benefits at different threshold probabilities. $P<0.05$ was considered to be statistically significant.

RESULTS

Patient Characteristics

A total of 91 patients (68 with adenocarcinoma, 22 with squamous cell carcinoma and 1 with adenosquamous carcinoma; 47 male, 44 female) were included. The median age of these patients was 59 years (range 36~78). Patients' demographics and the median value of twelve MPs of DTP ^{18}F-FDG PET images were shown in **Table 1**.

Positive PD-L1 immunostaining was observed on the membranes and/or in the cytoplasm of tumor cells. Positive PD-L1 expressions were noted in 33 of the 91 (36.26%) patients. The median value of $CD8^+$ TILs and PD-1 TILs was 100 (range 0~300) and 74 (range 0~282) respectively. The percentage of TMIT-I tumors was 24.18% (22/91) in this study.

Characterization of TMITs

The percentages of four TMITs were as follows: 22 (24.18%) TMIT I, 36 (39.56%) TMIT II, 11 (12.09%) TMIT III and 22 (24.18%) TMIT IV. PD-1 expression was significantly higher in TMIT I tumors than TMIT II and III tumors ($p < 0.001$, 0.015), but no statistically significant difference was found between TMIT I and IV groups ($p = 0.584$) (**Figure 1**). The spearman's analysis showed a statistically significant correlation between the PD-1 and $CD8^+$ TILs ($rho = 0.543$, $p < 0.001$). PD-1 and $CD8^+$ TILs were significantly higher in PD-L1 positive patients ($p = 0.015$ and 0.004, respectively).

Correlations Between PD-L1 Expression and MPs on DTP FDG PET

By Mann-Whitney *U* test, PD-L1 positive patients showed higher early SUVmax (eSUVmax) ($p = 0.0007$), early SUVmean

TABLE 1 | Demographic and clinical data of all patients.

Characteristics		No. (%)
Age	Median (range)	59 (36~78) years
Gender	Male	47 (51.65%)
	Female	44 (48.35%)
Smoking status	Smoker	37 (40.70%)
	Non-smoker	54 (59.30%)
CEA	Median (range)	2.52 (0.50~397.51) ug/L*
CYFRA21	Median (range)	2.49 (0.94~28.07) ng/ml*
SCC Ag	Median (range)	0.8 (0.30~47.99) μg/L*
Histology	SCC	22 (24.18%)
	ADC	68 (74.73%)
	ASC	1 (1.1%)
Pathologic stage	I	37 (40.70%)
	II	22 (24.2%)
	III	23 (25.30%)
	IV	9 (9.90%)
eSUVmax	Median (range)	8.40 (0.9~20.2)
eSUVmean	Median (range)	5.00 (0.6~12.00)
eMTV	Median (range)	5.78 (0.83~96.14) cm^3
eTLG	Median (range)	24.68 (0.85~663.37) g
dSUVmax	Median (range)	13.10 (0.80~27.90)
dSUVmean	Median (range)	7.60 (0.60~16.00)
dMTV	Median (range)	5.43 (0.61~89.45) cm^3
dTLG	Median (range)	30.69 (0.79~796.11) g
ΔSUVmax	Median (range)	2.80 (-0.2~8.500)
ΔSUVmean	Median (range)	1.90 (-0.1~4.30)
ΔMTV	Median (range)	-0.65 (-13.34~4.30) cm^3
ΔTLG	Median (range)	5.75 (-13.88~151.76) g

**Ten patients' data on tumor markers were absence.*

ADC, adenocarcinoma; SCC, squamous cell carcinoma; ASC, adenosquamous carcinoma; CEA, carcinoembryonic antigen; CYFRA21, cytokeratin 19-fragments; SCC Ag, squamous cell carcinoma antigen; SUVmax, maximum standard uptake value; SUVmean, mean standard uptake value; MTV, metabolic tumor volume; TLG, total lesion glycolysis.

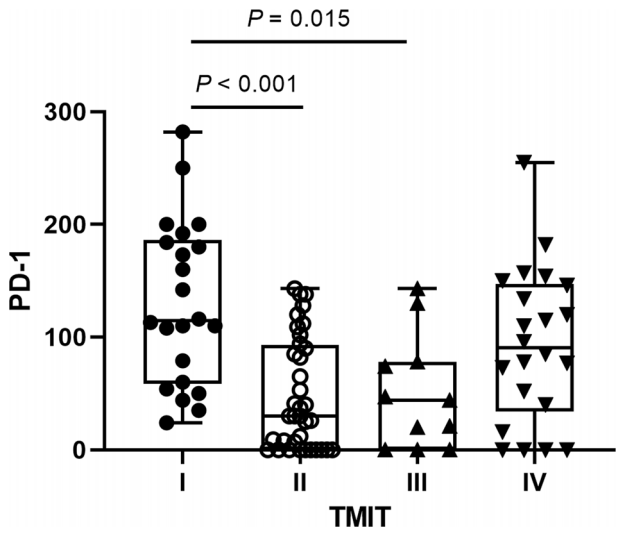


FIGURE 1 | The distributions of programmed cell death 1 (PD-1) expression according to different tumor microenvironment immune types (TMITs). The PD-1 expression was significantly higher in TMIT I tumors than TMIT II and III tumors.

(eSUVmean) (p = 0.0006), delayed SUVmax (dSUVmax) (p < 0.0001), delayed SUVmean (dSUVmean) (p<0.0001), ΔSUVmax (p = 0.0002), ΔSUVmean (p = 0.0002) and ΔTLG (p = 0.0247) in **Figure 2**. ROC curve for MPs showed moderate ability for predicting PD-L1 expression in **Table 2**.

Correlations Between TILs and MPs on DTP FDG PET

Spearman's correlation coefficients revealed poor correlations between TILs and MPs. Specifically, CD8+ TILs in NSCLC were weakly correlated with dSUVmean (rho = 0.212, p = 0.044) and ΔSUVmean (rho = 0.209, p = 0.047). Similarly, weak correlations were found between PD-1 TILs and eSUVmax (rho = 0.234, p = 0.026), eSUVmean (rho = 0.242, p = 0.021), dSUVmax (rho = 0.225, p = 0.032) and dSUVmean (rho = 0.235, p = 0.025), respectively (**Figure 3**).

Correlations Between Different TMITs and MPs on DTP FDG PET

Figure 4 showed that most MPs were significantly different between TMIT I and other immune types (TMIT II, III, IV) respectively. Meanwhile, the least p value was shown in TMIT I vs II, increasing gradually in TIMT I vs IV and TIMT I vs III. Moreover, TMIT I tumors exhibited higher eSUVmax, eSUVmean, eTLG, dSUVmax, dSUVmean, dTLG, ΔSUVmax, ΔSUVmean and ΔTLG (p = 0.0001, < 0.0001, 0.0453, < 0.0001, <0.0001, 0.0231, 0.0009, 0.0009, 0.0144) than other types together (TMIT II+III+IV). **Figure 5** showed a representative patient with a TMIT-I tumor exhibited hypermetabolic tumors on DTP PET, characterized by high expression of PD-L1 and high density of PD-1, $CD8^+$ TILs. Of the nine metabolic features above, eSUVmax, eSUVmean, eTLG, dSUVmax and ΔTLG were selected in the LASSO model with 10-fold cross-validation (**Figure 6**). The Meta-Sig (metabolic signature) was constructed as follows:

$$Meta-Sig = 0.125626063 \times eSUVmax + 0.066810617 \times eSUV—mean - 0.006982037 \times eTLG + 0.066701772 \times dSUVmax + 0.024456035 \times \Delta TLG - 3.584450396.$$

Meta-Sig had a higher AUC value than the stand-alone MP, although the differences were not statistically significant according to the Delong test. When compared to eSUVmax (AUC: 0.775, sensitivity: 77.27%, specificity: 73.91%) and dSUVmax (AUC: 0.788, sensitivity: 86.36%, specificity: 65.22%), Meta-Sig enhanced the performance to predict TMIT I tumors with higher sensitivity or specificity (AUC: 0.818, sensitivity: 86.36%, specificity: 73.91%) (**Table 3**).

With multivariate logistic regression analysis (using backward stepwise elimination method), the combined model was constructed based on the clinical and metabolic information. The formula was as follows:

$$Model = 1.490 \times Meta-Sig + 2.435 \times smoking - 0.840 \times stage - 1.689 \times gender + 1.883$$

The preferable model was assessed using ROC analyses and decision curve analysis (DCA). Compared to the combined model (AUC: 0.869, sensitivity: 77.27%, specificity: 82.61%), AUCs for eSUVmax, dSUVmax, gender, and smoking were relatively lower according to the Delong test (p = 0.0187, 0.0182, <0.0001, 0.0004) (**Table 3**; **Figure 7**). DCA for the combined model was shown in **Figure 8**. Using the combined model to predict TMIT I tumors added more benefit than eSUVmax or dSUVmax.

DISCUSSION

The past decade was marked by a revolution in the field of cancer treatment. Recently, antibody-based immunotherapy that modulates immune responses against tumors has been approved as first-line treatment option for selected advanced or metastatic lung cancer (20). However, the response of NSCLC patients to immunotherapy is affected by the TIME. Notably, patients with TMIT I tumors, regarded as immunologically 'hot', are most likely to benefit from ICI therapy. Our study provided a new insight into the underlying correlation between TIME types and DTP FDG PET imaging. To our knowledge, this is the first study to identify TMIT I type using DTP FDG PET scan in pretreated NSCLC patients.

In general, TMIT I tumors are characterized by somatic tumor mutations, PD-L1/PD-1 expression, and $CD8^+$ TILs. In a previous report, TMIT I tumors were found to harbor significantly more somatic tumor mutations (6), therefore presenting more neoantigens (4). Endogenous $CD8^+$ T cells can recognize these neoantigens, increase the TILs density, and trigger an immune response by the host (21). Meanwhile, glucose transporter 1 is reported to upregulate in these activated $CD8^+$ T

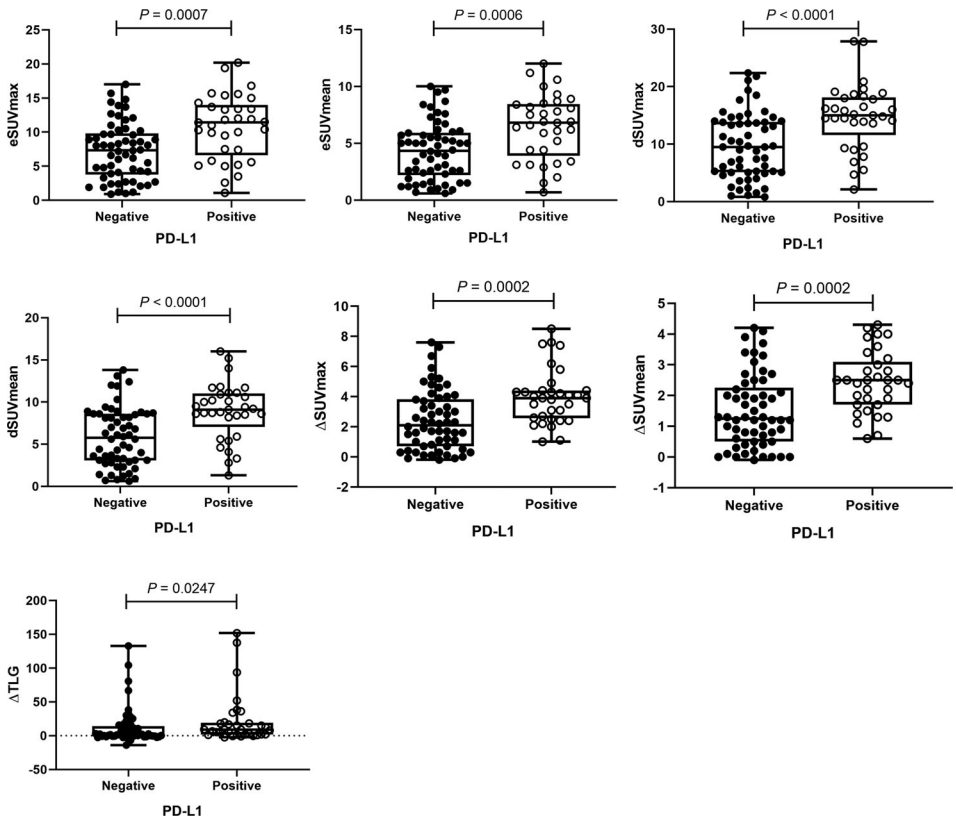


FIGURE 2 | The distributions of metabolic parameters (MPs) according to programmed cell death-ligand 1 (PD-L1) protein expression. The MPs were significantly higher in patients with PD-L1 positivity than those with PD-L1 negativity.

TABLE 2 | ROC analysis for the metabolic parameters according to PD-L1 expression.

Parameters	sensitivity	specificity	cutoff	AUC	95%CI
eSUVmax	66.67%	75.86%	9.7	0.714	0.610 - 0.804
eSUVmean	63.64%	77.59%	6.0	0.716	0.612 - 0.806
dSUVmax	66.67%	79.31%	14.0	0.750	0.648 - 0.835
dSUVmean	72.73%	72.41%	8.2	0.747	0.645 - 0.832
ΔSUVmax	93.94%	48.28%	1.9	0.732	0.629 - 0.820
ΔSUVmean	84.85%	56.90%	1.3	0.738	0.636 - 0.825

CI, confidence interval; SUVmax, maximum standard uptake value; SUVmean, mean standard uptake value; eSUVmax, early SUVmax; eSUVmean, early SUVmean; dSUVmax, delayed SUVmax; dSUVmean, delayed SUVmean; PD-L1, programmed death ligand 1.

cells, leading to increased glucose uptake (8). Furthermore, during therapeutic PD-1 blockade, pre-treatment samples obtained from responders exhibited higher $CD8^+$ cell densities compared to those from non-responders (22). According to results in our study, tumors with high infiltration of $CD8^+$ T cells had a tendency for high FDG uptake in comparison with the $TILs^-$ tumors. It suggested that FDG PET could serve as a non-invasive tool to assess the tumor microenvironment, and might help to identify responders ahead of treatment. However, CD8+ TILs alone had low correlations (p = 0.044, 0.047) with MPs in our work, proving the necessity of simultaneously interpreting multiple immune biomarkers within the complex system of tumor immune microenvironment. Indeed, there was a metabolic competition between tumors and immune cells. The high levels of glycolysis within tumor cells consumed extracellular glucose, which in turn impaired the glycolysis in T cells (23). We hypothesized that the FDG uptake of TMIT I tumors depends primarily on PD-L1+ tumor cells rather than $CD8^+$ T cells. Chang et al. (24) reported that PD-L1 expression maintained Akt/mTOR signaling, which in turn promoted metabolic pathway through the translation of glycolysis enzymes. Lopci et al. (9) first investigated the correlation between PD-L1 expression and FDG uptake in NSCLC. Although no correlation was found in this particular study, the negative results may arise from the small size of studied patients. Later, Takada et al. (10) found a positive correlation between FDG uptake with PD-L1 expression in a larger group of patients. In addition to PD-L1 expression and $CD8^+$ TILs, PD-1

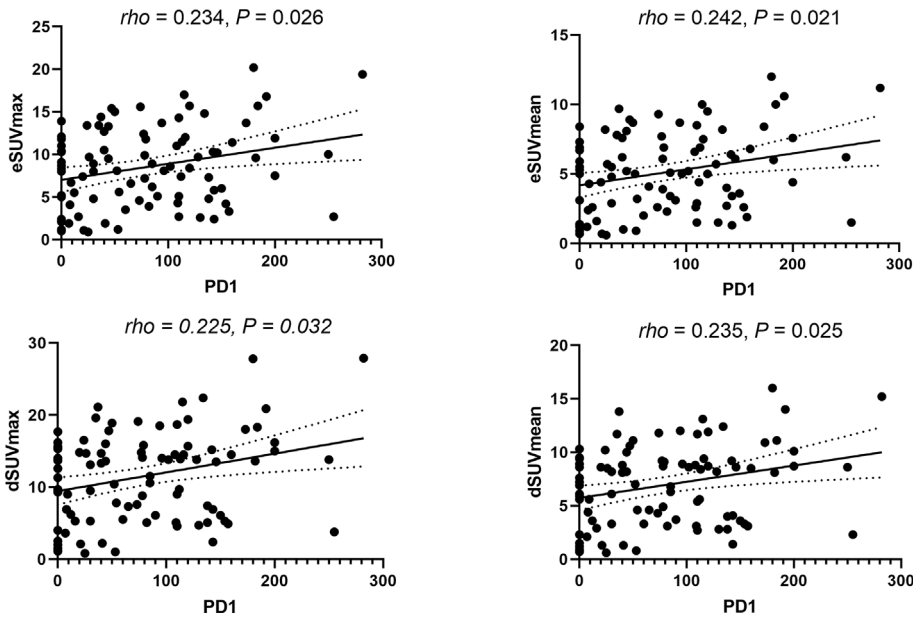


FIGURE 3 | Correlations between metabolic parameters and PD-1 expression.

expression is also a biomarker of ICIs treatment. PD-1 expression was significantly higher in TMIT I tumors than type II and III, probably because PD-1 is of upregulated in CD8+ TILs, where the binding of PD-1 and its ligand PD-L1 can inhibit a cytotoxic T-cell response.

CD8+ TILs and the tumor immune microenvironment may also affect patients' responses to immunotherapy, since not all patients with positive PD-L1 expression respond well. We hereby categorized the tumor immune microenvironment into four subgroups to assist the stratification of NSCLC patients. We strived to identify responders from the perspective of different tumor immune types, and for the first time established the correlation between dynamic glucose metabolism and TMIT-I tumors. One remarkable finding in the present study was that TMIT I tumors present higher MP values than other types (TMIT II, III, IV), and the smallest *p* value was shown between TMIT I vs II. Contrary to TMIT I tumors, TMIT II tumors, regarded as "cold" tumors, lack tumors cells expressing PD-L1 and $CD8^+$ T cells. While TMIT I tumors are most likely to benefit from single-agent anti-PD-1/L1 blockade, TMIT II tumors had significantly worse prognosis due to the lack of detectable immune responses (4). The semi-quantification of MPs on DTP FDG PET was a decision support methodology for the complex clinical decisions to differentiate TMIT I and II tumors. Compared to TMIT I vs II, the *p* values increased gradually in TMIT I vs IV. TMIT IV tumors are enriched with $CD8^+$ TILs. However, the lower uptake FDG may be attributed to PD-L1 negativity. Interestingly, TMIT III tumors also showed relatively lower FDG uptake than TMIT I, and the *p* value increased even further in TMIT I vs III. Since TMIT III tumors were also characterized by PD-L1 positivity like TMIT I tumors, $CD8^+$ T cells have less effect on the FDG uptake of TMIT III tumors. It should be noted that only a low proportion of tumors in this study belonged to TMIT III (12.1%, 11/91). A similarly low occurrence of TMIT III was also observed in melanoma. (2%) (25). Nevertheless, the presence of TMIT III tumors indicates that not all patients with PD-L1 positive expression respond well to the ICIs. Our work demonstrated the importance of highlighting this subset of patients as probable non-responders to immunotherapy. The identification of different tumor immune types is meaningful from the perspective of clinical practice, since using the same strategy for all patients will be inefficient, costly, and unreasonable (26).

For the identification of TMIT I tumors, DTP FDG PET not only provided a non-invasive work-up, but also improved the capability for pre-selecting patients that are likely responsive to ICIs. Cancer cells continuously uptake ^{18}F-FDG and trap them intracellularly in the form of ^{18}F-FDG-6-phosphate (27). In contrast, in benign tissue, the uptake of ^{18}F-FDG decreases or plateaued after reaching a maximum within 30 min of FDG administration. Although it is generally accepted that FDG-PET/CT images obtained during the delayed phase reflect the dynamics of tumor glucose metabolism, our study revealed that eSUVmax and dSUVmax had similarly unsatisfactory performances for the assessment of TMIT I tumors since the sensitivity of eSUVmax (77.27%) and specificity of dSUVmax (65.22%) were relatively low when compared with each other. The application of DTP FDG PET provided a new perspective for the assessment of tumor immune microenvironment, and can make up the deficiencies of a single MP. Compared with the MP at a single time point, DTP PET exhibited improved sensitivity (Meta-Sig vs eSUVmax: 86.36% vs 77.27%) and specificity (Meta-Sig vs dSUVmax: 73.91% vs 65.22%). Therefore, DTP FDG PET might reduce the false positive rates on early scan and false negative rates on delayed scan, which would help to facilitate accurate treatment strategies and reduce unnecessary medical cost. Furthermore, the combined model

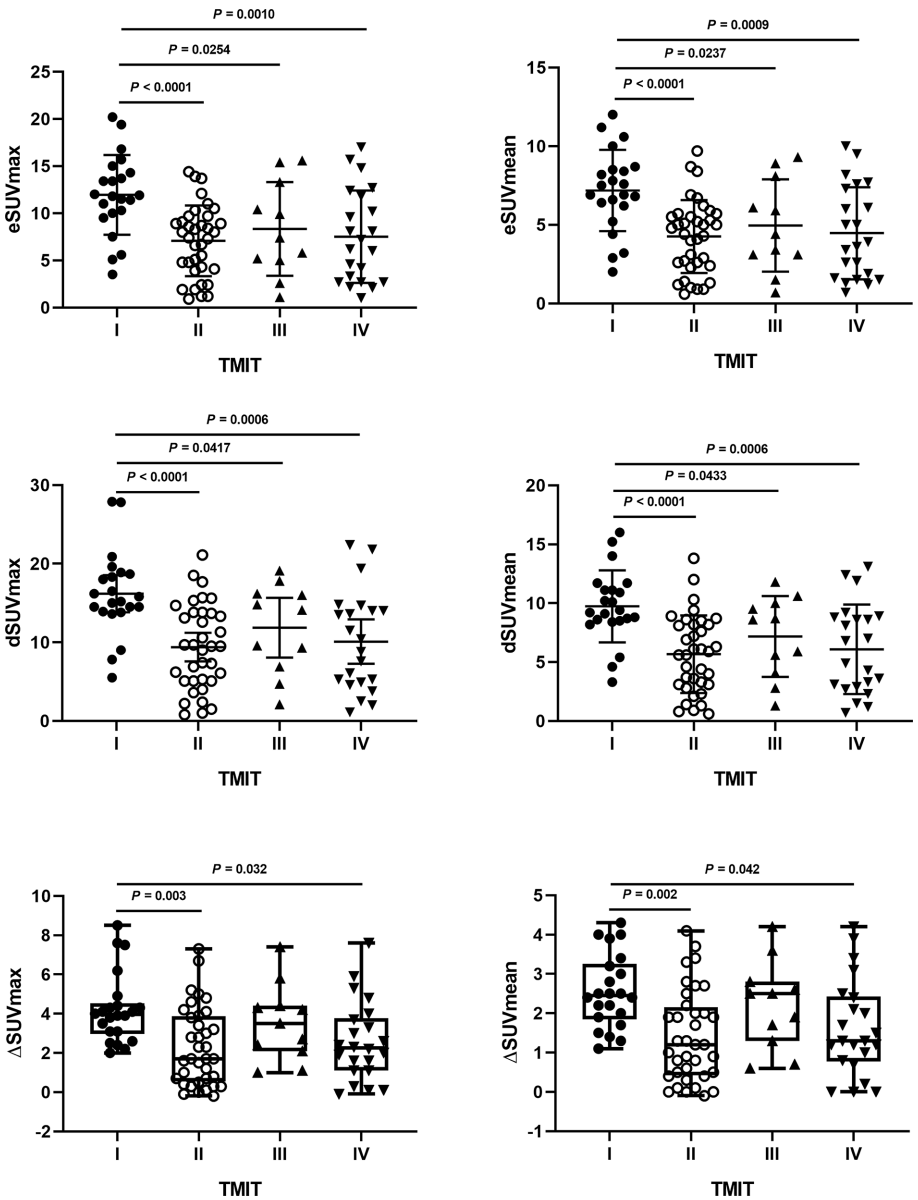


FIGURE 4 | The differences of metabolic parameters (MPs) according to different tumor microenvironment immune types (TMITs). The MPs were significantly higher in TMIT I tumors than other immune types.

achieved a performance for the identification of TMIT I tumors with AUC of 0.869, better than that of Meta-Sig (0.818). As confirmed by DCA, the combined model had a better clinical usefulness than Meta-Sig at a wide range of threshold probability (except the range of 20%~25%).

This present study also has a couple of limitations. First, this study was a single institutional retrospective study with a limit number of subjects. Second, further understanding of the complex and volatile TIME is still needed because TIME consists of various cells types, which display metabolic interactions in the tumor microenvironment. Third, as the analysis of histology slices was performed with a semi-quantitative method in the present study, some clustering analyses are needed to make it more quantitative. The clustering analysis may allocate the cohort into distinct and discrete subgroups, and clearly delineate patients who are suitable for immunotherapy. Fourth, none of the patients in the retrospective cohort received immunotherapy within one month of the 18F-FDG PET/CT scan. Therefore, it is unknown whether the patients with significantly changed MPs could benefit from immunotherapy. A recent study on 89 patients with advanced or recurrent NSCLC showed that patients with higher baseline SUVmax values had higher response rate to immunotherapy than those with lower baseline SUVmax values (28). Future studies are warranted to enroll patients that receive subsequent 18F-FDG PET/

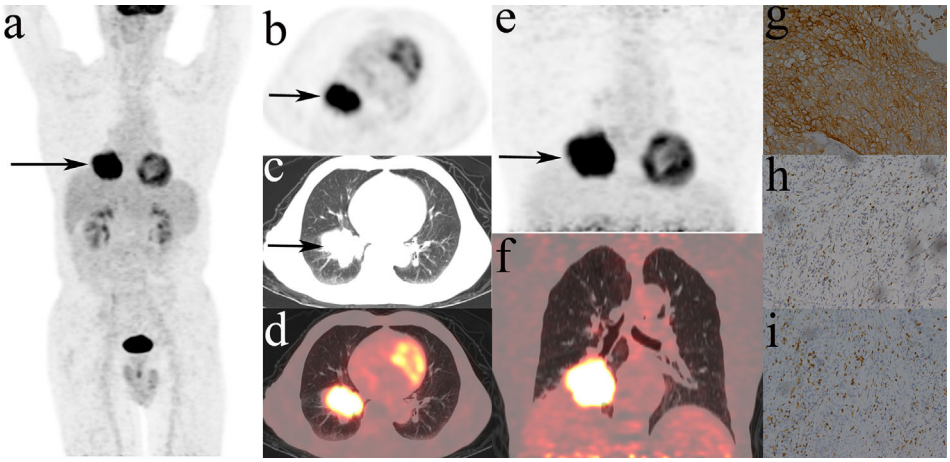


FIGURE 5 | Representative DTP ^{18}F-FDG PET/CT imagings of a 71y male patients, defined as TMIT I tumor. **(A–D)**: early images, **(E, F)**: delayed images, **(A)**: MIP figure, **(B–D)**: PET, lung window, PET/CT fusion image. A mass was in the lower lobe of right lung (arrow) with markedly increased radioactivity, eSUVmax: 20.3, eSUVmean: 12.0, eMTV: 52.9 cm^3, eTLG: 634.8 g, dSUVmax: 27.7, dSUVmean: 16.0, dMTV: 49.16 cm^3, dTLG: 786.56 g. The surgical pathology: moderately differentiated squamous cell carcinoma. **(G)** high PD-L1 expression. **(H)** PD-1 TIL high density. **(I)** CD8$^+$ TIL high density.

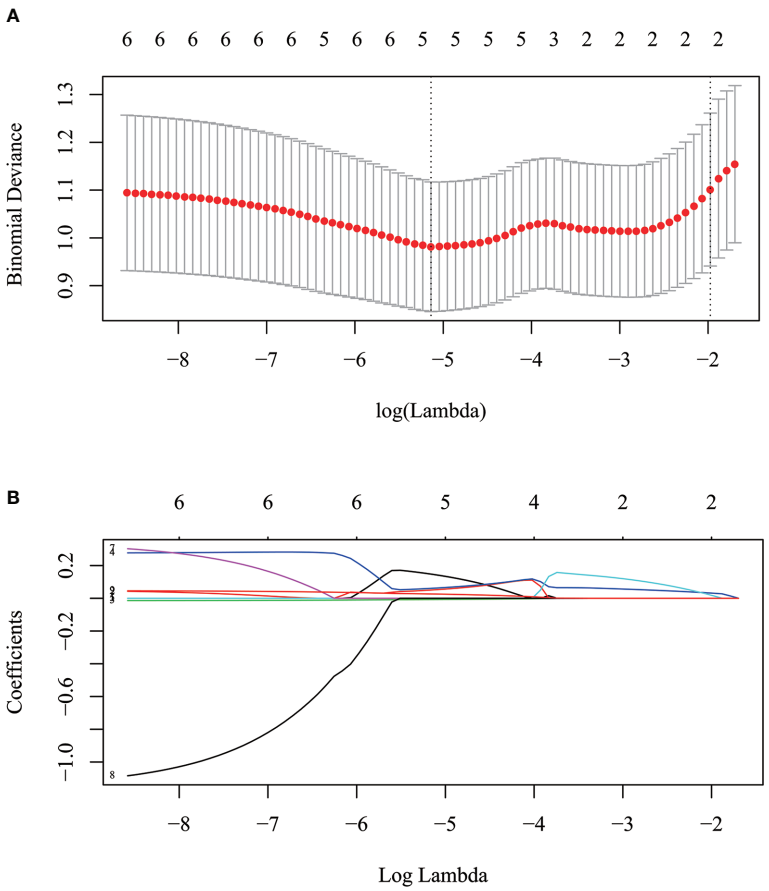


FIGURE 6 | The selection of optimal MPs using the LASSO algorithm. **(A)** The optimal tuning parameter (Lambda) in the LASSO model was selected using 10-fold cross-validation at the minimum of lambda. **(B)** LASSO coefficient profiles of the 9 parameters. According to the 10-fold cross-validation in **(A)**, Five parameters with non-zero coefficients were included for metabolic signature construction.

TABLE 3 | Performance of the combined model and other factors according to TMIT I tumors.

Factors	Sensitivity	Specificity	AUC	ΔAUC	95% CI	*Z*	*P*
Model	77.27%	82.61%	0.869				
dSUVmax	86.36%	65.22%	0.788	0.0814	0.0135 - 0.149	2.351	0.0187*
eSUVmax	77.27%	73.91%	0.775	0.0935	0.0159 - 0.171	2.361	0.0182*
Gender	68.18%	53.62%	0.609	0.260	0.140 - 0.380	4.250	< 0.0001*
Smoking	68.18%	68.12%	0.681	0.188	0.0840 - 0.291	3.551	0.0004*
Meta-Sig	86.36%	73.91%	0.818	0.0514	-0.0134 - 0.116	1.553	0.1203

DeLong method were used to compared the AUC of combined model with other factors.
**Statistically significant.*
SUVmax, maximum standard uptake value; dSUVmax, delayed SUVmax; eSUVmax, early SUVmax; Meta-Sig, metabolic signature; CI, confidence interval; TMIT I: 3tumor microenvironment immune type I.

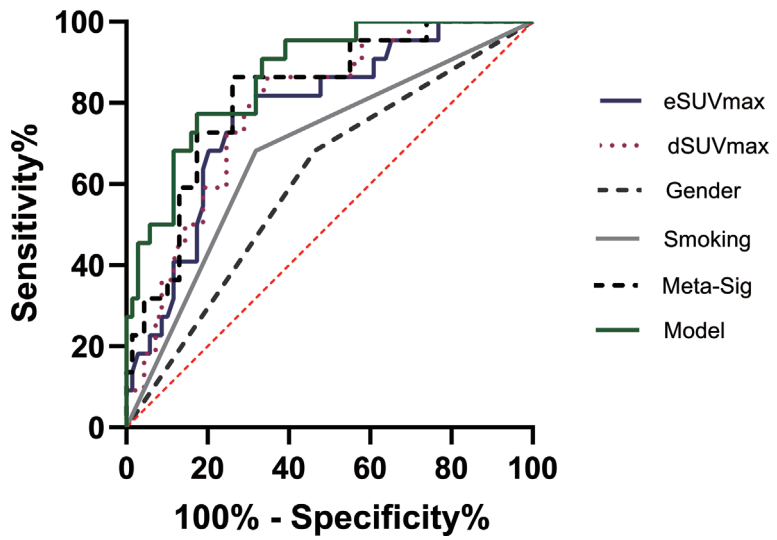


FIGURE 7 | Representative image of receiver operating characteristic (ROC) curves for various factors in the analyses of TMIT I tumors. The combined model had the highest AUC than other factors.

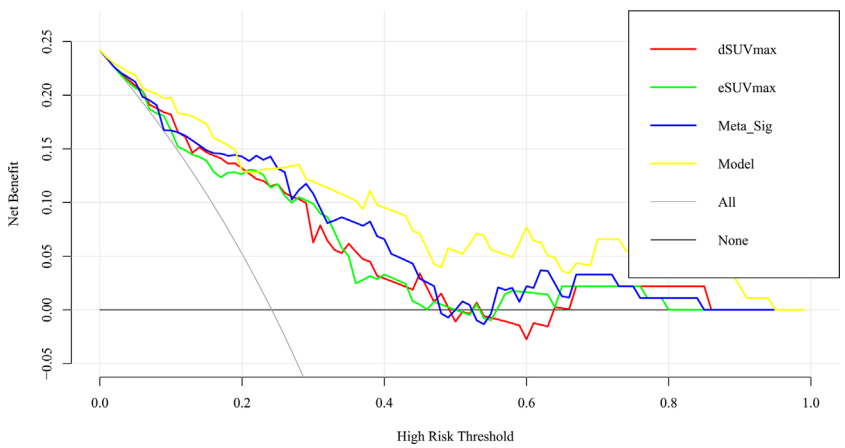


FIGURE 8 | Decision curve analysis for the model and other factors. The y axis measures the net benefit. The x axis shows the threshold probability. The yellow line represents the combined model. The blue line represents the Meta-Sig only. The thin gray line represents the assumption that all patients were with TMIT I tumors. The black line represents the assumption that no patients have a TMIT I tumor.

CT and immunotherapy within one month of each other, and to evaluate whether the patients with significantly changed MPs could benefit from immunotherapy.

CONCLUSION

High glucose metabolism on DTP ^{18}F-FDG PET is relevant to TMIT-I tumors, and the Meta-Sig and combined model based on clinical and metabolic information could improve the performance of identifying patients who may respond to ICIs treatment.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

All procedures performed in studies involving human participants were approved by the Institutional Review Board of Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology (TJ-IRB 20181202) and in accordance with the principles of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. This article does not describe any studies with animals performed by any of the authors.

AUTHOR CONTRIBUTIONS

XZ, JZ and SZ contributed to the conception and design of the study. JZ, SC, DK, DL, LC and CL carried out the research. JZ, SZ, JY and XZ performed the data analysis. JZ wrote the first draft of the manuscript. XZ, SZ, DK and JY made the comments. XZ and SZ critically reviewed and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (81671718, 81873903, 91959119), and Natural Science Foundation of Hubei Province of China (2016CFB687).

REFERENCES

1. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin* (2016) 66(2):115–32. doi: 10.3322/caac.21338
2. Zhang M, Yang J, Hua W, Li Z, Xu Z, Qian Q. Monitoring checkpoint inhibitors: predictive biomarkers in immunotherapy. *Front Med* (2019) 13 (1):32–44. doi: 10.1007/s11684-018-0678-0
3. Lin Z, Gu J, Cui X, Huang L, Li S, Feng J, et al. Deciphering Microenvironment of NSCLC based on CD8+ TIL Density and PD-1/PD-L1 Expression. *J Cancer* (2019) 10(1):211–22. doi: 10.7150/jca.26444
4. Teng MW, Ngiow SF, Ribas A, Smyth MJ. Classifying Cancers Based on T-cell Infiltration and PD-L1. *Cancer Res* (2015) 75(11):2139–45. doi: 10.1158/0008-5472.CAN-15-0255
5. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med* (2018) 24(5):541–50. doi: 10.1038/s41591-018-0014-x
6. Ock C-Y, Keam B, Kim S, Lee J-S, Kim M, Kim TM, et al. Pan-Cancer Immunogenomic Perspective on the Tumor Microenvironment Based on PD-L1 and CD8 T-Cell Infiltration. *Clin Cancer Res* (2016) 22(9):2261–70. doi: 10.1158/1078-0432.ccr-15-2834
7. Takeuchi S, Khiewvan B, Fox PS, Swisher SG, Rohren EM, Bassett RL Jr., et al. Impact of initial PET/CT staging in terms of clinical stage, management plan, and prognosis in 592 patients with non-small-cell lung cancer. *Eur J Nucl Med Mol Imaging* (2014) 41(5):906–14. doi: 10.1007/s00259-013-2672-8
8. Scharping NE, Delgoffe GM. Tumor Microenvironment Metabolism: A New Checkpoint for Anti-Tumor Immunity. *Vaccines (Basel)* (2016) 4(4):46. doi: 10.3390/vaccines4040046
9. Lopci E, Toschi L, Grizzi F, Rahal D, Olivari L, Castino GF, et al. Correlation of metabolic information on FDG-PET with tissue expression of immune markers in patients with non-small cell lung cancer (NSCLC) who are candidates for upfront surgery. *Eur J Nucl Med Mol Imaging* (2016) 43 (11):1954–61. doi: 10.1007/s00259-016-3425-2
10. Takada K, Toyokawa G, Okamoto T, Baba S, Kozuma Y, Matsubara T, et al. Metabolic characteristics of programmed cell death-ligand 1-expressing lung cancer on (18) F-fluorodeoxyglucose positron emission tomography/computed tomography. *Cancer Med* (2017) 6(11):2552–61. doi: 10.1002/cam4.1215
11. Takada K, Toyokawa G, Tagawa T, Kohashi K, Akamine T, Takamori S, et al. Association Between PD-L1 Expression and Metabolic Activity on (18)F-FDG PET/CT in Patients with Small-sized Lung Cancer. *Anticancer Res* (2017) 37 (12):7073–82. doi: 10.21873/anticanres.12180
12. Zhang M, Wang D, Sun Q, Pu H, Wang Y, Zhao S, et al. Prognostic significance of PD-L1 expression and (18)F-FDG PET/CT in surgical pulmonary squamous cell carcinoma. *Oncotarget* (2017) 8(31):51630–40. doi: 10.18632/oncotarget.18257
13. Kasahara N, Kaira K, Bao P, Higuchi T, Arisaka Y, Erkhem-Ochir B, et al. Correlation of tumor-related immunity with 18F-FDG-PET in pulmonary squamous-cell carcinoma. *Lung Cancer* (2018) 119:71–7. doi: 10.1016/j.lungcan.2018.03.001
14. Kaira K, Shimizu K, Kitahara S, Yajima T, Atsumi J, Kosaka T, et al. 2-Deoxy-2-[fluorine-18] fluoro-d-glucose uptake on positron emission tomography is associated with programmed death ligand-1 expression in patients with pulmonary adenocarcinoma. *Eur J Cancer* (2018) 101:181–90. doi: 10.1016/j.ejca.2018.06.022
15. Chen R, Zhou X, Liu J, Huang G. Relationship between the expression of PD-1/PD-L1 and (18)F-FDG uptake in bladder cancer. *Eur J Nucl Med Mol Imaging* (2019) 46(4):848–54. doi: 10.1007/s00259-018-4208-8
16. Choi SH, Chang JS, Koo JS, Park JW, Sohn JH, Keum KC, et al. Differential Prognostic Impact of Strong PD-L1 Expression and 18F-FDG Uptake in Triple-negative Breast Cancer. *Am J Clin Oncol* (2018) 4(11):1049–57. doi: 10.1097/coc.0000000000000426
17. Castino GF, Cortese N, Capretti G, Serio S, Di Caro G, Mineri R, et al. Spatial distribution of B cells predicts prognosis in human pancreatic adenocarcinoma. *Oncoimmunology* (2016) 5(4):e1085147. doi: 10.1080/2162402X.2015.1085147

18. Koh J, Go H, Keam B, Kim MY, Nam SJ, Kim TM, et al. Clinicopathologic analysis of programmed cell death-1 and programmed cell death-ligand 1 and 2 expressions in pulmonary adenocarcinoma: comparison with histology and driver oncogenic alteration status. *Mod Pathol* (2015) 28(9):1154–66. doi: 10.1038/modpathol.2015.63
19. Koh J, Nam SK, Roh H, Kim J, Lee BC, Kim JW, et al. Somatic mutational profiles of stage II and III gastric cancer according to tumor microenvironment immune type. *Genes Chromosomes Cancer* (2019) 58 (1):12–22. doi: 10.1002/gcc.22683
20. Ettinger DS, Aisner DL, Wood DE, Akerley W, Bauman J, Chang JY, et al. NCCN Guidelines Insights: Non-Small Cell Lung Cancer, Version 5.2018. *J Natl Compr Canc Netw* (2018) 16(7):807–21. doi: 10.6004/jnccn.2018.0062
21. Baretti M, Le DT. DNA mismatch repair in cancer. *Pharmacol Ther* (2018) 189:45–62. doi: 10.1016/j.pharmthera.2018.04.004
22. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* (2014) 515(7528):568–71. doi: 10.1038/nature13954
23. Sukumar M, Roychoudhuri R, Restifo NP. Nutrient Competition: A New Axis of Tumor Immunosuppression. *Cell* (2015) 162(6):1206–8. doi: 10.1016/j.cell.2015.08.064
24. Chang CH, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, et al. Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. *Cell* (2015) 162(6):1229–41. doi: 10.1016/j.cell.2015.08.016
25. Taube JM, Anders RA, Young GD, Xu H, Sharma R, McMiller TL, et al. Colocalization of inflammatory response with B7-h1 expression in human melanocytic lesions supports an adaptive resistance mechanism of immune escape. *Sci Transl Med* (2012) 4(127):127ra37. doi: 10.1126/scitranslmed.3003689
26. Sanmamed M F, Chen L. A Paradigm Shift in Cancer Immunotherapy: From Enhancement to Normalization. *Cell* (2018) 175(2):313–26. doi: 10.1016/j.cell.2018.09.035
27. Parghane R V, Basu S. Dual-time point (18)F-FDG-PET and PET/CT for Differentiating Benign From Malignant Musculoskeletal Lesions: Opportunities and Limitations. *Semin Nucl Med* (2017) 47(4):373–91. doi: 10.1053/j.semnuclmed.2017.02.009
28. Takada K, Toyokawa G, Yoneshima Y, Tanaka K, Okamoto I, Shimokawa M, et al. ^{18}F-FDG uptake in PET/CT is a potential predictive biomarker of response to anti-PD-1 antibody therapy in non-small cell lung cancer. *Sci Rep* (2019) 9(1):13362. doi: 10.1038/s41598-019-50079-2

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 05 July 2021
doi: 10.3389/fonc.2021.661244

DeepCUBIT: Predicting Lymphovascular Invasion or Pathological Lymph Node Involvement of Clinical T1 Stage Non-Small Cell Lung Cancer on Chest CT Scan Using Deep Cubical Nodule Transfer Learning Algorithm

Kyongmin Sarah Beck[1], Bomi Gil[1], Sae Jung Na[1], Ji Hyung Hong[2], Sang Hoon Chun[2], Ho Jung An[2], Jae Jun Kim[3], Soon Auck Hong[4], Bora Lee[5], Won Sang Shim[5], Sungsoo Park[5] and Yoon Ho Ko[2,6]**

[1] Department of Radiology, College of Medicine, The Catholic University of Korea, Seoul, South Korea, [2] Division of Oncology, Department of Internal Medicine, College of Medicine, The Catholic University of Korea, Seoul, South Korea, [3] Department of Thoracic and Cardiovascular Surgery, Uijeongbu St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, South Korea, [4] Department of Pathology, College of Medicine, Chung-Ang University, Seoul, South Korea, [5] Deargen Inc., Daejeon, South Korea, [6] Cancer Research Institute, College of Medicine, The Catholic University of Korea, Seoul, South Korea

The prediction of lymphovascular invasion (LVI) or pathological nodal involvement of tumor cells is critical for successful treatment in early stage non-small cell lung cancer (NSCLC). We developed and validated a Deep Cubical Nodule Transfer Learning Algorithm (DeepCUBIT) using transfer learning and 3D Convolutional Neural Network (CNN) to predict LVI or pathological nodal involvement on chest CT images. A total of 695 preoperative CT images of resected NSCLC with tumor size of less than or equal to 3 cm from 2008 to 2015 were used to train and validate the DeepCUBIT model using five-fold cross-validation method. We also used tumor size and consolidation to tumor ratio (C/T ratio) to build a support vector machine (SVM) classifier. Two-hundred and fifty-four out of 695 samples (36.5%) had LVI or nodal involvement. An integrated model (3D CNN + Tumor size + C/T ratio) showed sensitivity of 31.8%, specificity of 89.8%, accuracy of 76.4%, and AUC of 0.759 on external validation cohort. Three single SVM models, using 3D CNN (DeepCUBIT), tumor size or C/T ratio, showed AUCs of 0.717, 0.630 and 0.683, respectively on external validation cohort. DeepCUBIT showed the best single model compared to the models using only C/T ratio or tumor size. In addition, the DeepCUBIT model could significantly identify the prognosis of resected NSCLC patients even in stage I. DeepCUBIT using transfer learning and 3D CNN can accurately predict LVI or nodal involvement in cT1 size NSCLC on CT images. Thus, it can provide a more accurate selection of candidates who will benefit from limited surgery without increasing the risk of recurrence.

Keywords: deep learning, non-small cell lung cancer, prognosis, computed tomography, lobectomy

OPEN ACCESS

Edited by:
Ulas Bagci,
Northwestern University,
United States

Reviewed by:
Huaiqiang Sun,
Sichuan University, China
Suyash P. Awate,
Indian Institute of Technology Bombay, India

***Correspondence:**
Sungsoo Park
sspark@deargen.me
Yoon Ho Ko
koyoonho@catholic.ac.kr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *30 January 2021*
Accepted: *09 June 2021*
Published: *05 July 2021*

Citation:
Beck KS, Gil B, Na SJ, Hong JH, Chun SH, An HJ, Kim JJ, Hong SA, Lee B, Shim WS, Park S and Ko YH (2021) DeepCUBIT: Predicting Lymphovascular Invasion or Pathological Lymph Node Involvement of Clinical T1 Stage Non-Small Cell Lung Cancer on Chest CT Scan Using Deep Cubical Nodule Transfer Learning Algorithm. Front. Oncol. 11:661244. doi: 10.3389/fonc.2021.661244

INTRODUCTION

Lung cancer is one of the most prevalent lethal diseases in the world. Over the past decade, the percentage of early-stage lung cancer has also increased; clinical stage IA disease has increased to account for 15% of non-small cell lung cancer (NSCLC) patients in developed countries (1). The standard treatment for stage I NSCLC is lobectomy with mediastinal lymph node dissection for the best chance of cure (2). However, due to an increase of screen-detected, indolent cancers appearing as subsolid nodules, there has been a shift of surgical treatment modality toward limited resection, such as sublobar resection. Yet, randomized trials are ongoing and the results for limited resection are pending (3, 4). For limited surgery to be successful, a careful selection of patients who would most benefit from limited surgery is one of the most important steps. However, there are no definite selection criteria for limited resection as of yet; a few studies have suggested possible candidates for limited resection with their study results.

Predominance of ground-glass opacity (GGO) in a lung nodule on computed tomography (CT) has been widely recognized to correlate with less invasive pathological findings of cancer cells replacing the alveolar epithelial cells (5). NSCLC patients with predominantly GGO appearance showed extremely good prognosis following surgical resection (6), suggesting that they are good candidates for limited surgery. In addition, a previous study has also suggested that NSCLC without lymphovascular invasion (LVI) or nodal involvement to be suitable candidates for limited surgical resection. They concluded that consolidation to tumor ratio (C/T ratio) less than 0.25 or 0.5 on a CT scan could accurately predict the absence of LVI or nodal involvement with a very high specificity (96.4%) (7). However, C/T ratio requires a few extra steps of manually measuring the size of the tumor and its solid component and calculating the ratio, which could be time-consuming and increase work burden for the radiologist. If LVI or nodal involvement can be accurately identified with a simpler method, it would prove useful in the selection of candidates for limited resection.

Deep learning has emerged as a powerful tool of representation learning, drastically changing the landscape of feature engineering from hand-crafted manner to a self-taught, machine-driven manner (8). This has been proven to be useful in the field of medical image analysis. Moreover, many studies have successfully demonstrated various applications of deep learning, including nodule detection on chest radiographs (9) and prediction of malignancy in lung nodules (10). A deep learning system can identify features that cannot be assessed by the human eyes. Therefore, we hypothesized that it would be possible to develop a system which would classify and predict pathological features of a nodule on chest CT images to make an accurate selection of possible candidates for limited surgery with a simpler method. Thus, we planned to develop, train, and validate a deep learning algorithm in predicting LVI or pathological nodal involvement using chest CT images without manual measurements.

MATERIALS AND METHODS

Study Population

The clinical and pathological data of NSCLC patients who had undergone curative resection between 2011 and 2015 at two different hospitals (Seoul St. Mary's Hospital and Incheon St. Mary's Hospital) of the Catholic University of Korea were reviewed. The inclusion criteria were as follows: (i) pathologically confirmed stage I - III NSCLC; (ii) tumor size of ≤ 3 cm on pathology report; (iii) availability of pathology report; (iv) no preoperative radiation or chemotherapy; (v) availability of chest CT scan (axial images) prior to surgical treatment; and (vi) measurable cancer lesion on preoperative CT images. This study was approved by the institutional review board of Catholic Medical Center (No. UC17SESI0073), and was performed in accordance with the guideline of human research. The requirement for written informed consent was waived by the institutional review board (Catholic Medical Center) because of this study's character of retrospective analysis.

One thousand seventy-six patients underwent lung cancer surgery between 2011 and 2015, and the final 600 patients who met the inclusion criteria were identified at Seoul St. Mary's Hospital (cohort I). Patients in cohort I were used for training, validation and test, and 95 patients from Incheon St. Mary's Hospital (cohort II) were used for external validation (**Figure 1**). Six hundred ninety-five patients (mean age, 63.0 years ± 9.7) were enrolled in this study; the patient cohort consisted of 361 males and 334 females. Two hundred fifty-four (36.5%) had LVI or nodal involvement and 441 (63.5%) no LVI or nodal involvement (**Table 1**).

Data Preparation and Lesion Labeling

The preoperative CT images at Seoul St. Mary's Hospital were acquired from Siemens (Somtatom; Erlangen, Germany), with a tube voltage of 120 kVp and tube-current time product of 35-290 mAs, and the images were reconstructed with a slice thickness of 3-5 mm and increment of 3-5 mm. The preoperative CT images at Incheon St. Mary's Hospital were acquired from Toshiba (Aquilion; Tochigi-ken, Japan), with a tube voltage of 120 kVp and tube-current time product of 30-108 mAs, and the images were reconstructed with a slice thickness of 3-5 mm and increment of 3-5 mm.

Two board-certified radiologists (K.S.B. and B.M.K.), who were blinded to clinical data of all patients, manually drew a rectangular region of interest (ROI) (smallest possible rectangle that could encompass the entire tumor) around the cancer lesion on axial CT images on the PACS workstation (Maroview 5.4; Infinitt, Seoul, Korea) independently. ROI was drawn on the contrast-enhanced images, if available. The chest CT images were extracted in DICOM (Digital Imaging and Communications in Medicine) format with ROI information to develop prediction models.

Data Splitting and Pretreatment

The cohort I was randomized to maintain the ratio of training (64%), validation (16%), and test (20%). The training and

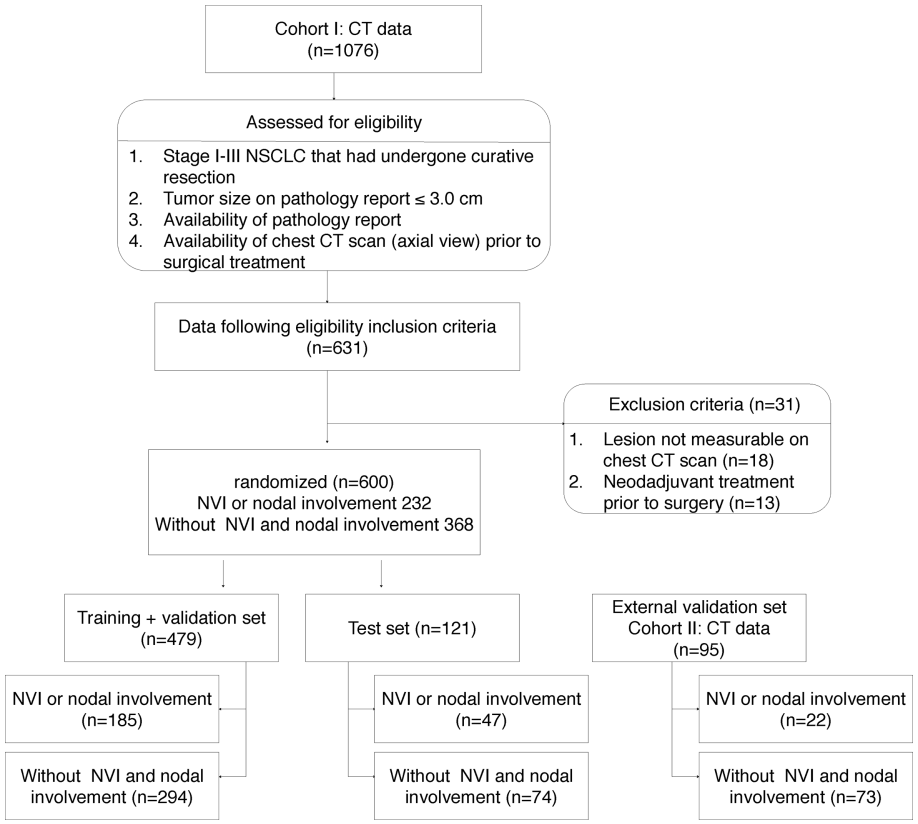


FIGURE 1 | Data criteria and specification.

TABLE 1 | Baseline characteristics of the training, validation, and external validation cohorts.

Characteristic		Total (n = 695)	Training and validation cohort (n = 600)	External validation cohort (n = 95)	*p-value*
Age (years)*	Mean ± SD	63.0 ± 9.7	63.0 ± 9.8	63.6 ± 9.5	0.551
Sex	Male	361	308 (51.3%)	53 (55.8%)	0.419
	Female	334	292 (48.7%)	42 (44.2%)	
Smoking history	Never	412	349 (58.2%)	63 (66.3%)	0.281
	Current	128	112 (18.7%)	16 (16.8%)	
	Former	155	139 (23.1%)	16 (16.8%)	
Histology	AC	471	395 (65.8%)	76 (80.0%)	0.005
	SqCC	123	108 (18.0%)	15 (15.8%)	
	Others	101	97 (16.2%)	4 (4.2%)	
Tumor size (cm)†		2.0 (1.6-2.6)	2.0 (1.5-2.6)	2.1 (1.7-2.6)	0.184
C/T ratio†		1.0 (0.5-1.0)	1.0 (0.5-1.0)	1.0 (0.7-1.0)	0.008
LVI or nodal involvement	Yes	254	232 (38.7%)	22 (23.2%)	0.004
	No	441	368 (61.3%)	73 (76.8%)	

**Data are mean ± SD.*

†Measured on CT image and data are median (with interquartile range in parentheses).

C/T ratio, consolidation to tumor ratio; LVI, lymphovascular invasion; AC, adenocarcinoma; SqCC, squamous cell carcinoma.

validation sets were used for model learning and optimal model selection, and the test set was used to evaluate the performance of the model. Five-fold stratified cross validation was adopted for training and validation. **Figure 1** indicates a detailed number of the lesions for training, validation, and testing, whereas **Figure 2** shows the overall evaluation pipeline. In the development of the deep learning system, each data sample was defined as (1): A 3D patch of 32mm×32mm×32mm, cropped from the CT scan at the center of a nodule; (2) The pathologically identified label of LVI or nodal involvement; and (3) Manually labeled voxel-wise nodule mask. Online data augmentation (randomly flipping the images on x, y, z axes) was performed for efficient training of the networks.

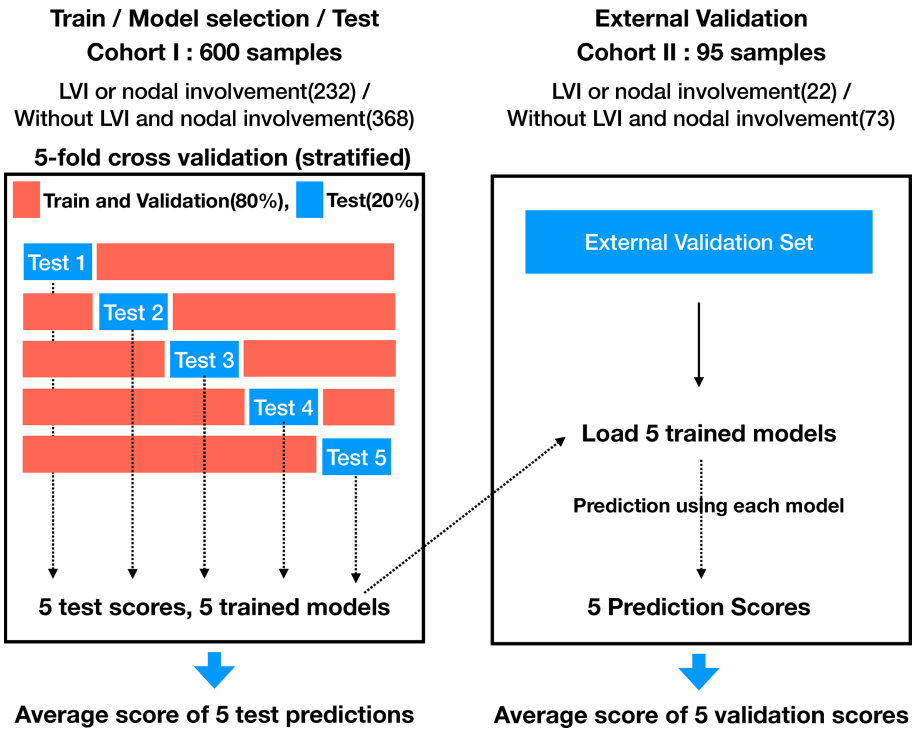


FIGURE 2 | Evaluation pipeline for proposed model.

Deep Cubical Nodule Transfer Learning (DeepCUBIT) Model

In order to overcome the limited number of samples of medical data in this study, which we considered insufficient to learn from scratch, we used a transfer learning method. Transfer learning is a machine learning technique for predictive modeling on related tasks that can be reused to accelerate the training and improve the performance of a model. This is done by fine-tuning the weights from a pre-trained network (11). We named the overall process consisting of pre-training, deep transfer learning, and actual prediction of LVI or nodal involvement as DeepCUBIT (Deep CUBical nodule Transfer learning) algorithm (**Figure 3**).

Preprocessing

Images were obtained as three-dimensional CT image data by layering two-dimensional slice images. Preprocessing was required to apply the data to the DeepCUBIT model, because the relative size of one voxel is different between samples and the entire CT image cannot be used as an input for the DeepCUBIT model. The CT images were preprocessed in the following steps. In CT scans, each 3D voxel intensity is expressed as Hounsfield Units (HU), which represent a measure of radiodensity. For example, HU value of -1000 represents the air and HU value from -500 to -600 represents the lungs. Inconsistency of cylindrical scanning boundary and image boundary results in an abnormal HU value of - 2000 HU. However, since these are noises and because we thought there is no need to differentiate air and noise, we changed the value of noise to be -1000HU to represent air, instead of original -2000HU. In order to have comparability among samples, we have rescaled the CT images so that one voxel represents size of 1mm×1mm×1mm by linear interpolation, and the voxel values were normalized using min-max normalization for each sample. We extracted 3D nodule cubes in the ROI according to manually annotated center of nodules for the test set, as in the training set. Whole raw image data were used instead of segmenting the lung region because there were nodules on the boundary between lung and mass.

Model Architecture

The structure of the DeepCUBIT consists of four 3D CNN units and one classifier unit. **Figure 4** shows the entire architecture of the DeepCUBIT. In the 3D CNN units, the channels are increased through the convolutional operation of CNN layers. The max-pooling layer is used to reduce resolution (12), and the batch normalization layer (13) is used to speed up the learning time and facilitate the learning process. The kernel size of the CNN layer is all (3, 3, 3), which means 3 pixels depth, 3 pixels height and 3 pixels width. In a 3D CNN layer, adding one kernel increases one channel, and the value computed through convolutional operation using the kernel forms one 3D cube channel. 3D CNN units are stacked to increase the channels and reduce the resolution. Thus, the DeepCUBIT model learns from the detailed features to the high-level abstract features. The classifier unit was made by adding a 64 nodes dense layer and sigmoid layer on the CNN layer. In the transfer learning process, the architecture of all units was unchanged except for the

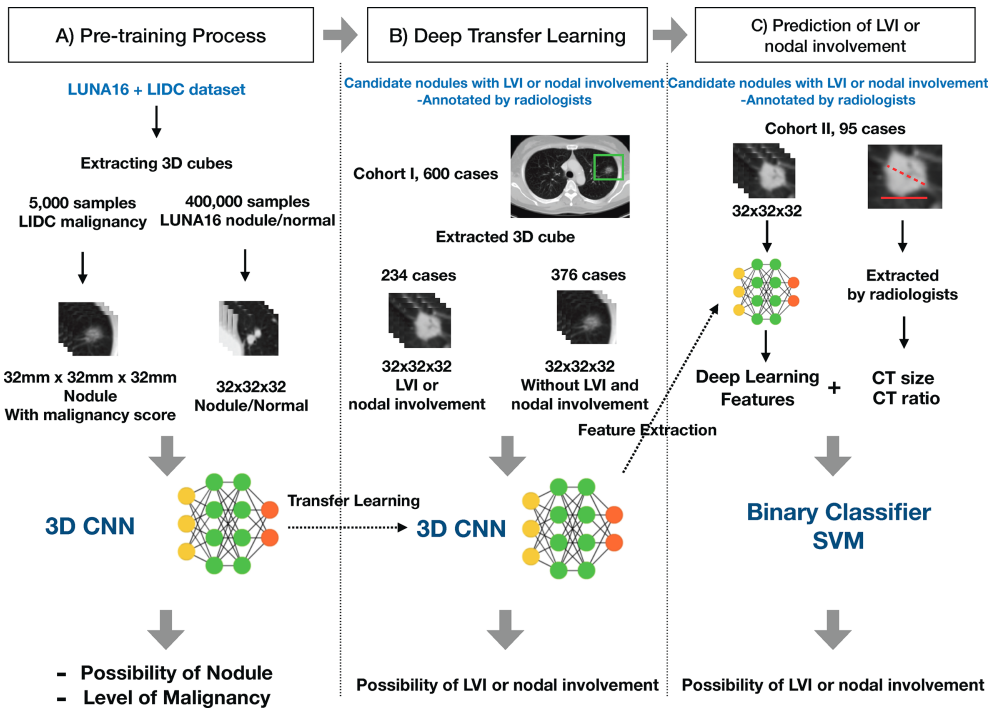


FIGURE 3 | Overall process of the DeepCUBIT algorithm. **(A)** Pre-training Process: Nodule samples and malignancy samples are presented to pre-train. **(B)** Deep Transfer Learning: Predicting the LVI or nodal involvement by fine-tuning the model with weights of pre-trained weights. **(C)** Prediction of LVI or nodal involvement: Feature integration and prediction of LVI or nodal involvement for extra validation cohort.

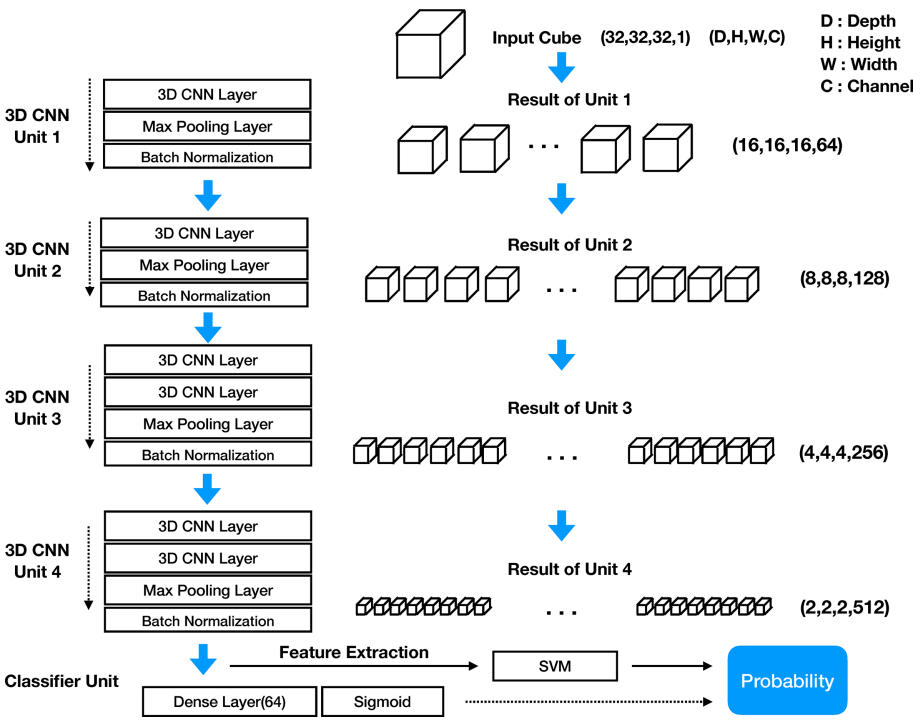


FIGURE 4 | Architecture of DeepCUBIT model.

classifier unit. After the training of the model, we changed the classifier unit to support vector machine (SVM) algorithm (14).

3D Convolutional Neural Network

CNN is a type of deep learning model, and some CNN models have the capability of end to end learning (15). In two-dimensional still image data, CNN image filter becomes a 2D kernel acting as a spatial feature extractor. However, in the CT data, we have to reflect the volume information as well as 2D spatial information; therefore, it is necessary to extract the features through the 3D kernels and 3D CNN. The 3D CNN obtains a feature map through a Width × Height × Depth filter called a kernel. The CNN layer creates a new feature map from neighboring pixels centered on the location where the kernel mapping to the previous CNN layer. The size of the feature map itself is reduced by attaching the max pooling layer to the end of the CNN layer. Moreover, stacking more than 2 layers of CNN layer will increase the feature map channel numbers (16). We can get the n_{th} channel feature maps value of (x, y, z) position of L_{th} CNN layer from the $L{-}1_{th}$ CNN layer using the following formula (12).

$$X_{nL}^{xyz} = f\left(b_{nL} + \sum_{m} \sum_{p=0}^{Pn-1} \sum_{q=0}^{Qn-1} \sum_{r=0}^{Rn-1} W_{nLm}^{pqr} X_{(L-1)m}^{(x+p)(y+q)(z+r)}\right)$$

In this formula, f is the activation function of the each node, b_{nL} is the bias values mapping kernel weights, P_n, Q_n, R_n are the size of the 3D kernel consist of width, height, and volume dimension, respectively, and w_{nLm}^{pqr} is the $(p, q, r)_{th}$ value of the kernel connected to the m_{th} feature map in the $L{-}1_{th}$ layer (previous layer). One 3D kernel has a feature map that extracts one feature because one 3D kernel with $P \times Q \times R$ weights will apply the same weight to all input CT data sliding with fixed stride hops. Therefore, we have to create a large number of kernels to extract various types of features.

Deep Transfer Learning

Pre-training domains related to the fine-tuning domain, which predicts LVI or nodal involvement of a nodule in this case, need to be selected for transfer learning to work. Pre-training was on the 3D ROIs. We selected two types of domain for pre-training: a nodule detection domain and a nodule malignancy prediction domain. LUNA16 (LUng Nodule Analysis 2016 challenge, https://luna16.grand-challenge.org/) data set of about 400,000 samples was trained for the first pre-training step to predict the presence of a nodule for the presented 3D cube. LIDC (The Lung Image Database Consortium image collection, https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI) data set of about 5,000 samples was trained for the second pre-training step to predict malignancy scores for the presented nodule cubes. The number of samples represents the number of nodule cubes. DeepCUBIT was then fine-tuned using nodule samples of cohort I based on the weights of pre-trained results.

Predicting Lymphovascular Invasion or Nodal Involvement and Addition of Clinical Data

To assess how well the features are extracted through the deep learning model and how well LVI or nodal involvement of nodule is predicted, we compared the prediction using DeepCUBIT model with prediction using tumor size and C/T ratio. The tumor size and C/T ratio are known indicators for determining the invasiveness of a nodule, both measured on CT and then calculated by the radiologists. We measured the performance of the learned model by using each feature set independently. We also measured the performance of the learned model by combining all the feature sets. We trained the SVM classifier to predict LVI or nodal involvement of a nodule. Radial Basis Function (RBF) kernel was used to train the SVM. The integrated features were used in the training of the SVM classifier after concatenating 64 size vector of DeepCUBIT features, tumor size, and C/T ratio. The process of the final LVI or nodal involvement prediction is presented in a part C) on the **Figure 3**.

Model performance was evaluated by averaging the scores of five stratified hold-out test sets on three different classifiers (SVM, xgboost, random forest). All classifiers were trained using default parameters, and software package versions were as follows: "scikit-learn" python package 0.21.3 for SVM and random forest and "xgboost" python package version 0.90 for xgboost. SVM classifier was finally selected, because it showed the best performance score among the three classifiers.

Statistical Analysis

Clinicopathologic characteristics are presented as median (range) for continuous variables or numbers (percentage) for categorical variables. Comparisons between the two groups were performed using the Students unpaired t-test or chi-square test. Recurrence-free survival (RFS) was defined as the duration between the date of diagnosis and the date of recurrence date or death from any causes. The performance of the algorithms was evaluated using Harrells concordance-index (C-index), which is a non-parametric statistic that measures concordance between predicted risk and actual survival (17). The predictive performance of all models was compared based on the mean AUC. The evaluation matrix includes Accuracy, Sensitivity, Specificity, PPV (Positive Predictive Value) and NPV (Negative Predictive Value). Kaplan-Meier method and log-rank test were used to determine the differences of estimated survival curves according to the classifier.

RESULTS

Performance of Transfer Learning

DeepCUBIT model was pre-trained using a total of 405,000 samples from LUNA16 and LIDC models. As depicted in **Table 2**, DeepCUBIT showed a much better performance than Deep 3D CNN without transfer learning. This result demonstrates that transfer learning is a critical step in training the domains that predict LVI or nodal involvement of nodules.

Model Performance

After transfer learning, nodule image features were extracted by the proposed deep network (DeepCUBIT), and clinical features were integrated into the model. The integrated model was performed to predict the performance of the best result in the test sets, and the

TABLE 2 | Performance comparison for transfer learning in Cohort 1 and 2.

Cohort	P-value	Model	CLF	AUC	Cis (95%)
1	**2.447e-09**	**Deep 3D CNN with TL**	NN	**0.682**	**0.587 - 0.772**
		Deep 3D CNN without TL	NN	0.606	0.503 - 0.707
2	**7.485e-11**	**Deep 3D CNN with TL**	NN	**0.669**	**0.553 - 0.78**
		Deep 3D CNN without TL	NN	0490	0.364 - 0.625

CNN, Convolutional Neural Network; CLF, Classifier; NN, Neural Network; TL, Transfer Learning; AUC, area under the curve; CIs, Confidence Intervals for AUC score.
Variables with DeepCUBIT model are shown in bold type.

results of the comparison are reported in **Table 3**, **Supplementary Table 1**, and **Supplementary Figure 1**. Three single SVM classifier models, using 3D CNN, tumor size, and C/T ratio showed AUCs of 0.723, 0.657 and 0.742, respectively. Evaluation scores improved (sensitivity of 75.8%, and specificity of 67.6%, accuracy of 70.8% and AUC of 0.77) after applying SVM to the merged features of DeepCUBIT model, C/T ratio, and tumor size.

We also did a subgroup analysis in patients with C/T ratio < 1.0 to compare the probability scores of DeepCUBIT and SVM classifier models using C/T ratio in predicting LVI or nodal involvement. The performance of DeepCUBIT model was superior to the C/T ratio model in this subgroup, and the detailed results can be seen in the **Supplementary Table 2**. These findings indicate that deep learning features and clinical features are complementary to each other.

External Validation

To further analyze the robustness, reproducibility, and reliability of the model, we performed an additional validation using data from cohort II. Similar to the results of cohort I, DeepCUBIT showed the best single model performance (AUC of 0.717) compared to the models using only C/T ratio or tumor size on SVM classifier (**Table 4**, **Supplementary Table 3** and **Supplementary Figure 2**). Applying SVM using all the features, including DeepCUBIT features, tumor size, and C/T ratio, showed the best predictive performance (sensitivity of 31.8%, and specificity of 89.8%, accuracy of 76.4% and AUC of 0.759).

Clinical Significance of DeepCUBIT Model

To identify the area in the CT image most responsible for the DeepCUBIT model in predicting LVI or nodal involvement, GradCAM (Gradient-weighted Class Activation Mapping) (18) was adapted to visualize the 3D CNN result, creating heatmaps (**Figure 5**). Grad-cam was applied after the DeepCUBIT network was trained. DeepCUBIT model was made of four 3D CNN blocks. Last output activations and gradients of third 3D CNN layers block was used in the analysis. Heatmaps showed different results for tumors with C/T ratio 1.0 and those with C/T ratio lower than 1.0. In solid tumors with C/T ratio 1.0, the area most responsible for the prediction of LVI or nodal involvement was the tumor itself.

TABLE 3 | Performance evaluation for test data (Cohort I, average of 5 fold hold-out test set).

Classifier	SVM		Xgboost		Random Forest	
Feature Type	**AUC**	**CIs (95%)**	**AUC**	**CIs (95%)**	**AUC**	**CIs (95%)**
3D CNN (DeepCUBIT)	**0.723**	**0.633 - 0.814**	**0.730**	**0.642 - 0.816**	**0.715**	**0.622 - 0.802**
Tumor size	0.657	0.558 - 0.751	0.621	0.522 - 0.720	0.577	0.473 - 0.684
C/T Ratio	0.742	0.663 - 0.817	0.726	0.644 - 0.803	0.631	0.538 - 0.721
Tumor Size + C/T Ratio	0.754	0.669 - 0.834	0.735	0.658 - 0.817	0.686	0.591 - 0.777
3D CNN + Tumor size + C/T Ratio	0.770	0.681 - 0.852	0.752	0.663 - 0.833	0.725	0.635 - 0.813

CNN, Convolutional Neural Network; DeepCUBIT, Deep Cubical Nodule Transfer Learning Algorithm; C/T Ratio, consolidation to tumor ratio; SVM, Support Vector Machine; AUC, area under the curve; CIs, Confidence Intervals for AUC score.
Variable with DeepCUBIT model is shown in bold type.

TABLE 4 | Performance evaluation for external validation data (Cohort II, external hold-out set).

Classifier	SVM		Xgboost		Random Forest	
Feature Type	**AUC**	**CIs (95%)**	**AUC**	**CIs (95%)**	**AUC**	**CIs (95%)**
3D CNN (DeepCUBIT)	**0.717**	**0.601 - 0.819**	**0.685**	**0.566 - 0.797**	**0.660**	**0.528 - 0.779**
Tumor size	0.630	0.502 - 0.749	0.634	0.510 - 0.752	0.606	0.476 - 0.729
C/T Ratio	0.683	0.614 - 0.743	0.682	0.612 - 0.743	0.658	0.591 - 0.733
Tumor size + C/T Ratio	0.716	0.606 - 0.813	0.715	0.613 - 0.812	0.663	0.544 - 0.776
3D CNN + Tumor size + C/T Ratio	0.759	0.646 - 0.855	0.757	0.654 - 0.843	0.716	0.607 - 0.820

CNN, Convolutional Neural Network; DeepCUBIT, Deep Cubical Nodule Transfer Learning Algorithm; C/T Ratio, consolidation to tumor ratio; SVM, Support Vector Machine; AUC, area under the curve; CIs, Confidence Intervals for AUC score.
Variable with DeepCUBIT model is shown in bold type.

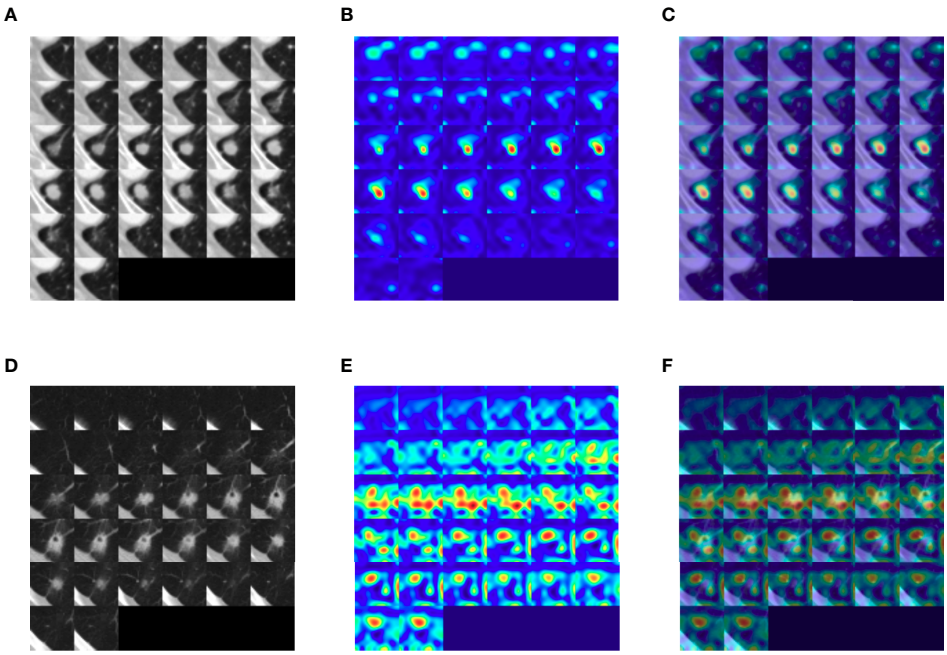


FIGURE 5 | Gradient-weighted class activation heatmaps of nodule cubes. **(A)** Raw intensity, **(B)** gradient heatmap, and **(C)** overaly heatmap of a solid tumor with C/T ratio 1.0 show the area most responsible for the prediction of LVI or nodal involvement to be the solid tumor itself, rather than pleural tag. **(D)** Raw intensity, **(E)** gradient heatmap and **(F)** overlay heatmap of a part-solid tumor with C/T ratio 0.75 show that the area most responsible for the prediction of LVI or nodal involvement to be the interface of the tumor with the adjacent lung parenchyma.

However, in part-solid tumors with C/T ratio lower than 1.0, the area most responsible for the prediction of LVI or nodal involvement was the periphery of the tumor, which is the interface between the tumor and the adjacent lung parenchyma.

To ascertain the clinical significance of DeepCUBIT model, a novel 3D CNN using the deep cubical transfer learning algorithm, the survival analysis for relapse free survival (RFS) was performed on patients with stage I. These patients were of special interest since postoperative treatment for stage I disease is controversial in external cohort. We assumed that the samples with high invasion probability score will have high risk probability, so we sorted samples according to the invasion probability scores based on the median probability. Despite the small number of patients with stage I disease (105 patients; cohort I, those not used in the training, n = 62; cohort II, n = 43), the RFS of patients with high and low risk scores using DeepCUBIT alone (P = 0.019) and SVM model using DeepCUBIT features with tumor size and C/T ratio (P = 0.223) was significantly different. However, SVM model using tumor size or C/T ratio alone did not demonstrate any significant difference between high and low risk score for the 3-year RFS (**Figure 6**).

DISCUSSION

Lobectomy is the standard surgical care for patients with resectable NSCLC; recently, pulmonary function-preserving limited surgery has become more prevalent due to the increased number of early and small sized lung cancer, owing to advances in CT technology, more widespread use of CT, and implementation of low-dose CT screening programs worldwide (4). However, previous randomized controlled study had failed to demonstrate the efficacy and validity of limited surgery for clinical early-stage NSCLC (19). Therefore, a careful selection of early stage NSCLC patients for limited resection is of paramount importance to achieve favorable clinical outcomes. One of the factors for favorable clinical outcomes is the absence of LVI or nodal involvement (20). If early-stage NSCLC population without LVI or nodal invasion could be accurately identified before surgery, they could undergo limited resection expecting favorable outcomes. To our knowledge, this the first study incorporating deep learning with preoperative CT images of primary tumor to identify LVI or nodal involvement. We developed a Deep 3D CNN with transfer learning algorithm, the DeepCUBIT, that showed similar performance to the C/T ratio, which is a strong indicator for LVI or nodal involvement of early lung cancer in previous studies (7, 21). Adding C/T ratio and tumor size to the deep learning algorithm further improved the predicting capability of deep learning algorithm. However, even with DeepCUBIT alone, the prediction of LVI or nodal involvement in cT1 stage NSCLC using CT images has become much simpler yet accurate.

The performance of our model, as represented by specificity of 92.6% and sensitivity of 27.6%, is similar to the results of a previous study using C/T ratio to predict LVI or pathological nodal involvement, which resulted in specificity of 96.4% and sensitivity of 30.4% (7). The results cannot be directly compared,

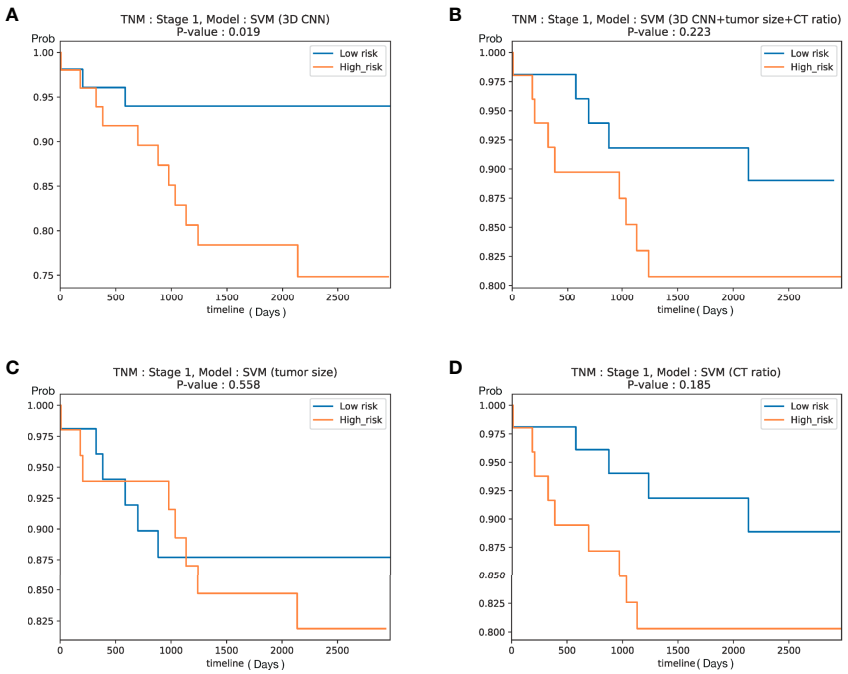


FIGURE 6 | Kaplan–Meier curves according to predicted risk of recurrence for NSCLC patients with stage I in Cohort I (test set only) and Cohort II (105 patients). Curves obtained using **(A)** DeepCUBIT model, **(B)** SVM classifier using DeepCUBIT features with tumor size and C/T ratio, **(C)** SVM classifier using tumor size alone, and **(D)** SVM classifier using C/T ratio alone.

because our study population and theirs are slightly different: that study consisted almost entirely of adenocarcinomas (97.1%, to be exact), but our study consisted of 65.8% of adenocarcinomas, and 34.2% were NSCLCs other than adenocarcinomas, and this could have influenced the accuracy.

Lymph node involvement or LVI are known to have higher recurrence rate and mortality risk (20, 22). In previous studies, tumor size and C/T ratio have been identified as well-known risk factors for mediastinal nodal involvement (23). Another study has also shown that the size of the consolidation or solid portion of the primary tumor measured on CT images is one of the independent predictors of lymph node metastasis (24). However, conventional CT images, which rely on lymph node size alone, have a low accuracy in predicting nodal involvement (25). Although the specificity of 18F-FDG PET/CT for detecting lymph node metastasis is high, accuracy of PET-CT is insufficient because of its low sensitivity, especially in the tuberculosis endemic countries (26). Unlike lymph node metastasis, LVI cannot be clinically and preoperatively determined based on CT imaging features. Thus, the fact that LVI or nodal metastasis has been incorporated in the prediction of invasiveness using CT images seems to be an encouraging step. Therefore, our success in using the CT images of the primary tumor to predict LVI or nodal metastasis is in line with the results of previous studies (20, 22).

There have been many conflicting reports dealing with proper indication and efficacy of limited surgery for early stage NSCLC. However, based on the long-term results of the JCOG 0201 trial, limited surgery would lead to satisfactory prognoses in patients with predominantly GGO lung cancers with C/T ratio of 0.5 or less and tumor sizes exceeding 2 cm but 3 cm or less (21, 27). In our study, the performance of the DeepCUBIT model predicting LVI or nodal involvement was similar to that of using C/T ratio, in tumors less than 3cm in size. Of note, this deep learning model can potentially identify patients at high recurrence risk even in stage I patients in a simpler way, which may reflect the biology of primary tumor and provide additional beneficial prognostic information. The benefit from adjuvant chemotherapy is now widely accepted stage in II or III NSCLC (28), but there is no agreement on the use of adjuvant chemotherapy in stage I NSCLC. Previous retrospective study indicated that adjuvant chemotherapy might be beneficial to stage I NSCLC patients with high risk features, such as LVI (29). However, choosing optimal candidates for adjuvant treatment according to conventional single risk factor might be insufficient because it does not consider all clinical or biologic factors and the varying weight of each factor. Thus, we could identify the subgroup harboring a high risk of recurrence that might benefit from adjuvant therapy by applying the novel deep algorithm developed in this study.

In general, to apply the machine learning to CT images, a known feature is extracted from radiologists using the domain knowledge in the CT images or using feature extraction software tools. Then, the machine running is applied to that extracted features. However, in this way, the performance is limited, because new hidden features are difficult to find and learning occurs only within the known existing feature set. To overcome these drawbacks, we used an end-to-end learning method by

extracting a feature set from CT images directly using 3D CNN. Moreover, we used transfer learning to improve generality, robustness, and performance. We observed that the addition of transfer learning to 3D CNN improved the performance of 3D CNN, which means that transfer learning process, consisting of pre-training and fine-tuning, is a necessary step in optimization of learning features. Heatmaps generated in order to identify the areas in the CT image most responsible for the DeepCUBIT model in predicting LVI or nodal involvement showed that the DeepCUBIT model seemed to work differently in tumors with C/T ratio 1.0 and those with C/T ratio lower than 1.0. In solid tumors with C/T ratio 1.0, the area most responsible for the prediction of LVI or nodal involvement was the tumor itself. However, in part-solid tumors with C/T ratio lower than 1.0, the area most responsible for the prediction of LVI or nodal involvement was the periphery of the tumor, which is the interface between the tumor and the adjacent lung parenchyma. This suggests that there may be factors not identifiable to the human eye influencing the invasiveness of a part-solid tumor at the interface between the tumor and the adjacent lung parenchyma. Further studies may be needed to expand on this idea. We believe such approach integrating the deep learning models and readily available clinical or radiological data can be used to develop other models in the medical field.

The clinical relevance of the findings in this study has several limitations. First, the number of patients in an independent external validation dataset is relatively small, but both training and validation cohort data, such as radiological findings, standardized surgical treatment, and the detailed records of clinical parameters, were well obtained and of good quality. Second, this is a retrospective study, but the inclusion and exclusion criteria were strictly applied to ensure the inclusion of definite LVI or nodal involvement in the study. Third, there were 18 (2.9%) out of 631 cases with primary cancer lesion "not measurable" on chest CT scans, which were excluded from the study population. All of these lesions were endobronchial lesions in the lobar or intermediate bronchi, and 15 of these tumors could be seen on CT but the exact extent of tumor was not entirely clear because they were blended with distal atelectasis. Three other lesions were diagnosed through transbronchoscopic biopsy but could not be located on CT. However, they only represent a relatively small population (2.9%). Moreover, endobronchial tumors are considered as relative contraindications for performing sublobar resection in NSCLC (30, 31), so we don't believe excluding these patients would create major limitation for using our approach in the clinical practice for selecting candidates for limited resection. Fourth, the single model of DeepCUBIT did not outperform the boosting model using C/T ratio. This result might be due to the limited number of samples. However, the deep learning model has the advantage that CT images can be directly used to predict LVI or nodal involvement of the tumor without increasing the workload of a radiologist. Fifth, the CT images used in this study are comprised of a heterogeneous mixture of CTs from different vendors, machines, and protocols. Contrast-enhancement images were used if available because we hypothesized enhancement pattern may be useful in determining LVI or nodal involvement. However, because contrast-enhancement images were not available in all patients, the data is heterogeneous, and this could have affected the ability of the deep learning algorithm in predicting LVI or nodal involvement. Nonetheless, we believe such heterogeneity in CT images accurately reflects the real world, and a trained deep learning system with such data may be more fit for the real-world clinical practice.

CONCLUSIONS

The authors have shown that the DeepCUBIT algorithm using transfer learning and 3D CNN based on CT scan images can accurately predict LVI or nodal involvement of primary NSCLC. This deep learning algorithm prediction may be convenient and useful for individualizing treatment modality. In order to predict LVI or nodal involvement of the tumor before surgery, an integrated deep learning approach that combines the multimodal imaging data with clinical data may be more useful.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethical committee at Seoul St. Mary's Hospital and Incheon St. Mary's Hospital of the Catholic University of Korea. Written informed consent was waived by the institutional review board because of this study's character of retrospective analysis.

AUTHOR CONTRIBUTIONS

YK and SP generated the concept of the work. BG, SN, SC, JH, JK, SH, WS, HA, and KB performed the data acquisition, analysis, and interpretation. WS, BL, and SP created the new algorithm used in this work. KB drafted the manuscript. HA and YK substantively revised it. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by a grant from the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (1720100).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.661244/full#supplementary-material

Supplementary Figure 1 | Comparison of performance of three models (DeepCUBIT, tumor size, and C/T ratio) *via* ROC AUC. AUC scores were calculated using the hold-out test data in cohort I (n = 121 patients, **(A)** and cohort II (n = 95 patients, **(B)**.

REFERENCES

1. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol* (2016) 11(1):39–51. doi: 10.1016/j.jtho.2015.09.009
2. Lackey A, Donington JS. Surgical Management of Lung Cancer. *Semin Intervent Radiol* (2013) 30(2):133–40. doi: 10.1055/s-0033-1342954
3. Donington JS, Kim YT, Tong B, Moreira AL, Bessich J, Weiss KD, et al. Progress in the Management of Early-Stage Non-Small Cell Lung Cancer in 2017. *J Thorac Oncol* (2018) 13(6):767–78. doi: 10.1016/j.jtho.2018.04.002
4. Khullar OV, Liu Y, Gillespie T, Higgins KA, Ramalingam S, Lipscomb J, et al. Survival After Sublobar Resection Versus Lobectomy for Clinical Stage IA Lung Cancer: An Analysis From the National Cancer Data Base. *J Thorac Oncol* (2015) 10(11):1625–33. doi: 10.1097/JTO.0000000000000664
5. Taylor ML, Carmona F, Thiagarajan RR, Westgate L, Ferguson MA, del Nido PJ, et al. Mild Postoperative Acute Kidney Injury and Outcomes After Surgery for Congenital Heart Disease. *J Thorac Cardiovasc Surg* (2013) 146(1):146–52. doi: 10.1016/j.jtcvs.2012.09.008
6. Ohde Y, Nagai K, Yoshida J, Nishimura M, Takahashi K, Suzuki K, et al. The Proportion of Consolidation to Ground-Glass Opacity on High Resolution CT Is a Good Predictor for Distinguishing the Population of Non-Invasive Peripheral Adenocarcinoma. *Lung Cancer* (2003) 42(3):303–10. doi: 10.1016/j.lungcan.2003.07.001
7. 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(4):751–6. doi: 10.1097/JTO.0b013e31821038ab
8. Shen D, Wu G, Suk HI. Deep Learning in Medical Image Analysis. *Annu Rev BioMed Eng* (2017) 19:221–48. doi: 10.1146/annurev-bioeng-071516-044442
9. Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, et al. Development and Validation of Deep Learning-Based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* (2019) 290(1):218–28. doi: 10.1148/radiol.2018180237
10. Causey JL, Zhang J, Ma S, Jiang B, Qualls JA, Politte DG, et al. Highly Accurate Model for Prediction of Lung Nodule Malignancy With CT Scans. *Sci Rep* (2018) 8(1):9286. doi: 10.1038/s41598-018-27569-w
11. Shao L, Zhu F, Li X. Transfer Learning for Visual Categorization: A Survey. *IEEE Trans Neural Netw Learn Syst* (2015) 26(5):1019–34. doi: 10.1109/TNNLS.2014.2330900
12. Ji S, Yang M, Yu K. 3D Convolutional Neural Networks for Human Action Recognition. *IEEE Trans Pattern Anal Mach Intell* (2013) 35(1):221–31. doi: 10.1109/TPAMI.2012.59
13. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. (2015).
14. Cortes C, Vapnik V. Support-Vector Networks. *Mach Learn* (1995) 20 (3):273–97. doi: 10.1007/BF00994018
15. LeCun Y, Bottou L, Yoshua B, Haffner P. Gradient-Based Learning Applied to Document Recognition. *Proc IEEE* (1998) 86:2278–324. doi: 10.1109/5.726791
16. Yosinski J, Clune J, Nguyen AM, Fuchs TJ, Lipson H. Understanding Neural Networks Through Deep Visualization. *CoRR* (2015). arXiv preprint arXiv:1506.06579.
17. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the Yield of Medical Tests. *JAMA* (1982) 247(18):2543–6. doi: 10.1001/jama.247.18.2543
18. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations From Deep Networks *Via* Gradient-Based Localization. *Proc IEEE Int Conf Comput Vision (ICCV)* (2017), 618–26. doi: 10.1109/ICCV.2017.74
19. Ginsberg RJ, Rubinstein LV. Randomized Trial of Lobectomy Versus Limited Resection for T1 N0 Non-Small Cell Lung Cancer. Lung Cancer Study Group. *Ann Thorac Surg* (1995) 60(3):615–22; discussion 22-3. doi: 10.1016/0003-4975(95)00537-U
20. Higgins KA, Chino JP, Ready N, D'Amico TA, Berry MF, Sporn T, et al. Lymphovascular Invasion in Non-Small-Cell Lung Cancer: Implications for Staging and Adjuvant Therapy. *J Thorac Oncol* (2012) 7(7):1141–7. doi: 10.1097/JTO.0b013e3182519a42
21. Asamura H, Hishida T, Suzuki K, Koike T, Nakamura K, Kusumoto M, et al. Radiographically Determined Noninvasive Adenocarcinoma of the Lung: Survival Outcomes of Japan Clinical Oncology Group 0201. *J Thorac Cardiovasc Surg* (2013) 146(1):24–30. doi: 10.1016/j.jtcvs.2012.12.047
22. Wang J, Wang B, Zhao W, Guo Y, Chen H, Chu H, et al. Clinical Significance and Role of Lymphatic Vessel Invasion as a Major Prognostic Implication in Non-Small Cell Lung Cancer: A Meta-Analysis. *PloS One* (2012) 7(12):e52704. doi: 10.1371/journal.pone.0052704
23. Koike T, Koike T, Yamato Y, Yoshiya K, Toyabe S. Predictive Risk Factors for Mediastinal Lymph Node Metastasis in Clinical Stage IA Non-Small-Cell Lung Cancer Patients. *J Thorac Oncol* (2012) 7(8):1246–51. doi: 10.1097/JTO.0b013e31825871de
24. Zang RC, Qiu B, Gao SG, He J. A Model Predicting Lymph Node Status for Patients With Clinical Stage T1aN0-2M0 Nonsmall Cell Lung Cancer. *Chin Med J (Engl)* (2017) 130(4):398–403. doi: 10.4103/0366-6999.199838
25. Toloza EM, Harpole L, McCrory DC. Noninvasive Staging of Non-Small Cell Lung Cancer: A Review of the Current Evidence. *Chest* (2003) 123(1 Suppl): 137S–46S. doi: 10.1378/chest.123.1_suppl.137S
26. Pak K, Park S, Cheon GJ, Kang KW, Kim IJ, Lee DS, et al. Update on Nodal Staging in Non-Small Cell Lung Cancer With Integrated Positron Emission Tomography/Computed Tomography: A Meta-Analysis. *Ann Nucl Med* (2015) 29(5):409–19. doi: 10.1007/s12149-015-0958-6
27. Aokage K, Yoshida J, Ishii G, Matsumura Y, Haruki T, Hishida T, et al. Identification of Early T1b Lung Adenocarcinoma Based on Thin-Section Computed Tomography Findings. *J Thorac Oncol* (2013) 8(10):1289–94. doi: 10.1097/JTO.0b013e31829f6d3b
28. Pignon JP, Tribodet H, Scagliotti GV, Douillard JY, Shepherd FA, Stephens RJ, et al. Lung Adjuvant Cisplatin Evaluation: A Pooled Analysis by the LACE Collaborative Group. *J Clin Oncol* (2008) 26(21):3552–9. doi: 10.1200/JCO.2007.13.9030
29. Tsuchiya T, Akamine S, Muraoka M, Kamohara R, Tsuji K, Urabe S, et al. Stage IA Non-Small Cell Lung Cancer: Vessel Invasion Is a Poor Prognostic Factor and a New Target of Adjuvant Chemotherapy. *Lung Cancer* (2007) 56 (3):341–8. doi: 10.1016/j.lungcan.2007.01.019
30. Ketchedjian A, Daly B, Landreneau R, Fernando H. Sublobar Resection for the Subcentimeter Pulmonary Nodule. *Semin Thorac Cardiovasc Surg* (2005) 17 (2):128–33. doi: 10.1053/j.semtcvs.2005.04.003
31. Sihoe AD, Van Schil P. Non-Small Cell Lung Cancer: When to Offer Sublobar Resection. *Lung Cancer* (2014) 86(2):115–20. doi: 10.1016/j.lungcan.2014.09.004

Conflict of Interest: Authors BL, WSS, and SP were employed by company Deargen Inc.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 17 November 2021
doi: 10.3389/fonc.2021.769272

A Novel Approach Using FDG-PET/CT-Based Radiomics to Assess Tumor Immune Phenotypes in Patients With Non-Small Cell Lung Cancer

Jianyuan Zhou[1], *Sijuan Zou*[1], *Dong Kuang*[2], *Jianhua Yan*[3], *Jun Zhao*[1,4†] *and Xiaohua Zhu*[1*†]

[1] Department of Nuclear Medicine and PET, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2] Department of Pathology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3] Shanghai Key Laboratory of Molecular Imaging, Shanghai University of Medicine and Health Sciences, Shanghai, China, [4] School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

OPEN ACCESS

Edited by:
Yiyan Liu,
University of Louisville, United States

Reviewed by:
Zehua Zhu,
Xiangya Hospital, China
Hongzan Sun,
Shengjing Hospital of China Medical University, China

***Correspondence:**
Xiaohua Zhu
evazhu@vip.sina.com
orcid.org/0000-0003-0495-9510

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 01 September 2021
Accepted: 25 October 2021
Published: 17 November 2021

Citation:
Zhou J, Zou S, Kuang D, Yan J, Zhao J and Zhu X (2021) A Novel Approach Using FDG-PET/CT-Based Radiomics to Assess Tumor Immune Phenotypes in Patients With Non-Small Cell Lung Cancer. Front. Oncol. 11:769272. doi: 10.3389/fonc.2021.769272

Purpose: Tumor microenvironment immune types (TMITs) are closely related to the efficacy of immunotherapy. We aimed to assess the predictive ability of ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT)-based radiomics of TMITs in treatment-naive patients with non-small cell lung cancer (NSCLC).

Methods: A retrospective analysis was performed in 103 patients with NSCLC who underwent ^{18}F-FDG PET/CT scans. The patients were randomly assigned into a training set (n = 71) and a validation set (n = 32). Tumor specimens were analyzed by immunohistochemistry for the expression of programmed death-ligand 1 (PD-L1), programmed death-1 (PD-1), and CD8+ tumor-infiltrating lymphocytes (TILs) and categorized into four TMITs according to their expression of PD-L1 and CD8+ TILs. LIFEx package was used to extract radiomic features. The optimal features were selected using the least absolute shrinkage and selection operator (LASSO) algorithm, and a radiomics signature score (rad-score) was developed. We constructed a combined model based on the clinical variables and radiomics signature and compared the predictive performance of models using receiver operating characteristic (ROC) curves.

Results: Four radiomic features (GLRLM_LRHGE, GLZLM_SZE, SUVmax, NGLDM_Contrast) were selected to build the rad-score. The rad-score showed a significant ability to discriminate between TMITs in both sets ($p < 0.001$, $p < 0.019$), with an area under the ROC curve (AUC) of 0.800 [95% CI (0.688–0.885)] in the training set and that of 0.794 [95% CI (0.615–0.916)] in the validation set, while the AUC values of clinical variables were 0.738 and 0.699, respectively. When clinical variables and

radiomics signature were combined, the complex model showed better performance in predicting TMIT-I tumors, with the AUC values increased to 0.838 [95% CI (0.731–0.914)] in the training set and 0.811 [95% CI (0.634–0.927)] in the validation set.

Conclusion: The FDG-PET/CT-based radiomic features showed good performance in predicting TMIT-I tumors in NSCLC, providing a promising approach for the choice of immunotherapy in a clinical setting.

Keywords: radiomics, tumor microenvironment immune types, non-small cell lung cancer, ^{18}F-FDG PET/CT, PD-L1

INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths in the United States (1). Among the common subtypes of lung cancer, non-small cell lung cancer (NSCLC) represents approximately 85% of lung cancer cases. Most of the patients with NSCLC are already at an advanced stage upon diagnosis, whose 5-year survival rate is below 5% (2). Recently, immune checkpoint inhibitors targeting the programmed death-1 (PD-1)/programmed death-ligand 1 (PD-L1) axis have become standard treatments for patients with advanced NSCLC. Biomarkers, including the tumor proportion score of PD-L1, are being tested in clinical trials for its ability to identify patients who are most likely to benefit from immunotherapy (3). However, the predictive ability of PD-L1 expression is still under debate (4), since the majority of patients with PD-L1-positive tumors did not respond to PD-1/PD-L1 blockade. In addition to cancer cells, tumor immune microenvironment also plays a critical role in immunotherapy. Recent studies have demonstrated that tumor tissue dampened the host immune response by upregulation of PD-L1, which subsequently ligated to PD-1 on the antigen-specific CD8+ T cells (5). Therefore, without the preexistence of CD8+ tumor-infiltrating lymphocytes (TILs), blockade of PD-L1 or PD-1 is unlikely to achieve any antitumor efficacy. Tumor immune microenvironment could be classified into four types according to the status of PD-L1 expression and CD8+ TIL abundance (6, 7), while the tumors with tumor microenvironment immune type I (TMIT-I), i.e., with high PD-L1 expression and presence of CD8+ TILs, are more likely to benefit from anti-PD-L1/PD-1 therapies (6). An accurate identification of the TMIT-I subset not only can maximize the therapeutic efficacy of anti-PD-1/PD-L1 therapy but also can minimize the adverse effects of treatments. However, to date, there are no noninvasive methods to specifically identify the TMITs of NSCLC tumors.

Medical imaging allows a noninvasive evaluation of tumor and its microenvironment, as well as a longitudinal assessment of tumor progression. ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG-PET/CT) is one of the most commonly used diagnostic imaging modalities in oncology (8). ^{18}F-FDG PET monitors the metabolism of glucose that is actively entrapped as nutrients in neoplastic tissues and tumor-associated activated immune cells (9). Therefore, ^{18}F-FDG PET signals depicting the glucose metabolism are closely related to the characteristics of tumor immune microenvironment. Previous studies have shown a direct association between the maximum standardized uptake value (SUVmax) of ^{18}F-FDG-PET and the expression of tumor-related immunity markers within the tumor immune microenvironment (10, 11). However, the SUVmax does not account for the spatial heterogeneity in the metabolism and biological features of tumor. Its predictive value on patients treated with immune checkpoint inhibitors remains weak. Tumor heterogeneity poses a significant challenge to personalized cancer medicine. The heterogeneity in the tumor uptake of FDG is of clinical importance as evidenced by a number of clinical trials (12). However, little attention is paid to the association between tumor immune microenvironment and the intratumoral heterogeneity of ^{18}F-FDG uptake. Radiomics is a rapidly evolving field of research that is focused on the extraction and quantification of patterns within medical images (13). Unlike biopsies that only take a snapshot within a small tumor portion, radiomics captures heterogeneity across the entire tumor volume.

This retrospective study was conducted to establish a correlation between the intratumoral heterogeneity of ^{18}F-FDG PET signals and tumor immune phenotype in a cohort of treatment-naive NSCLC patients. We hypothesized that radiomic features would provide insights into TMIT categorization and help optimize patient selection for immunotherapy.

MATERIALS AND METHODS

Patients

With approval from the institutional review board, we retrospectively analyzed consecutive patients who had been diagnosed pathologically with NSCLC between December 2014 and December 2017 at our institution. Enrollment eligibility: patients histologically confirmed to present NSCLC and underwent initial ^{18}F-FDG PET/CT scan within 30 days of surgery or biopsy; tumor size ≥1 cm in diameter. Exclusion criteria: patients who received antitumor therapy before surgery or biopsy due to the concern of therapy-induced alteration in PD-L1 expression. Patients without available tumor specimens for immunohistochemistry were also excluded.

Immunohistochemistry Analysis

Immunohistochemistry was performed using protocols described in a previous study (14). In brief, 4-μm continuous sections were prepared from formalin-fixed, paraffin-embedded (FFPE) tissue blocks. Slides were autostained by the Leica Bond-

Max automation (www.leica-microsystems.com) with primary antibodies against CD8 (ZA-0508, ZSGB-BIO), PD-1 [Abcam, EPR4877(2), ab137132], and PD-L1 (ZA-0629, ZSGB-BIO). The analysis of Immunohistochemistry results was performed as our previous study (15). The PD-L1 immunostaining results were classified into two groups based on staining intensity and proportion of tumor cell positivity. Staining intensity was scored as 0-3: 0, negative staining; 1, weak staining; 2, moderate staining; and 3, strong staining (more intense than alveolar macrophages). Cases with more than 5% of tumor cells and staining intensity ≥2 were defined as positive. Cases with staining intensity <2 or with positive staining in less than 5% of tumor cells were defined as negative. The expressions of PD-1 and CD8+ TILs were evaluated according to the average number of positively stained cells in three randomly selected high-power fields (HPFs) in each case. The numbers of CD8+ TILs were classified into two groups based on the median value (n=99): CD8+ TILs- (n ≤ 99) and CD8+ TILs+ (n > 99).

Four TMITs were classified as reported (6): TMIT-I (PD-L1+, CD8+ TILs+), TMIT-II (PD-L1-, CD8+ TILs-), TMIT-III (PD-L1+, CD8+ TILs-), and TMIT-IV (PD-L1-, CD8+ TILs+).

^{18}F-FDG PET/CT Acquisition Protocol and Image Analysis

^{18}F-FDG was intravenously administered at a dose of 3.7 MBq/kg after fasting for at least 6 h. The blood glucose concentration was lower than 11 mmol/L before injecting ^{18}F-FDG. PET/CT imaging was performed on a PET/CT scanner (Discovery 690 PET/CT, GE) at 60 ± 5 min after FDG administration. Whole-body images were obtained from the base of the skull to mid-thigh by means of an integrated PET/CT tomography (5–7 bed positions with 2 min per bed position). A low-dose helical CT scan (120 kV; 120 mA; slice thickness, 3.75 mm) was performed for anatomical correlation and attenuation correction. Reconstructed images were then displayed on a GE ADW4.5 workstation. Tumor mass was identified as the volume with elevated ^{18}F-FDG uptake compared to normal lung parenchyma or other mediastinal structures. SUVmax was defined as the highest pixel value of PET imaging. Tumor burden was calculated by drawing a three-dimensional volume of interest (VOI) on the volume of tumor-related metabolic activity and applying a percentage threshold at 30% of SUVmax.

Radiomic Feature Extraction

The feature extraction was performed as previously described (16). Briefly, LIFEx package (version 5.10, http://www.lifexsoft.org) was used to extract the texture features of ^{18}F-FDG PET/CT images of lesions in the same VOI. The ^{18}F-FDG PET/CT images of the patient in the DICOM format were imported into the software. Two experienced PET/CT diagnostic physicians semiautomatically delineated the VOI of the target lesion using a threshold at 30% of SUVmax. The interobserver reliability between the two physicians was analyzed. Then, the software program automatically calculates and extracts 52 PET radiomic features and 51 CT radiomic features, which are provided in **Supplemental Table 1**.

Radiomic Feature Selection and Model Establishment

Radiomic features with significant differences among different TMITs were selected in the training set using the Mann–Whitney U test with a p-value <0.05. The least absolute shrinkage and selection operator (LASSO) algorithm with 10-fold cross-validation was then used to select the optimal predictive features in the training set. The selected features with non-zero coefficients at the minimum of lambda were selected to construct a radiomics signature score (rad-score). Finally, rad-score and clinical variables were combined to establish a complex model using multivariate logistic regression analysis.

Model performance was tested in the validation set. Briefly, receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) were used to evaluate the model performance in the training and validation sets. A nomogram was developed to display the prediction results for each patient using the rad-score and clinical variables, and calibration curves were plotted to improve the nomogram's prediction accuracy. Furthermore, decision curve analysis (DCA) was performed to evaluate the clinical usefulness of the combined model by quantifying the net benefits at different threshold probabilities.

Statistical Analysis

All statistical tests were performed using SPSS statistical package (version 22.0, IBM, Armonk, NY, USA), MedCalc (MedCalc Software bvba, Ostend, West Flanders, Belgium), and R version 3.6.2 (http://www.r-project.org).

Feature reliability was analyzed using an intraclass correlation coefficient (ICC), where ICC ≥0.75 is generally considered to indicate good repeatability of the measured results. Mann–Whitney U test and Fisher's exact test were used to test the differences between continuous variables or categorical variables, respectively. Relations between two variable distributions were analyzed with the Spearman rank correlation coefficient (rho).

R package "glmnet" was used to perform LASSO binary logistic regression analysis, "rms" package to create the nomogram and calibration curve, "rmda" package to plot the DCA, "ggplot" package to plot the bar graph, and the "pROC" package to analyze ROC curves. A p-value <0.05 was considered statistically significant.

RESULTS

Patient Characteristics

In total, 103 patients were eligible for the retrospective analysis. The median age of the patients was 59 years old (range: 33–78 years old). The patients were randomly assigned to training or validation set at a ratio of 7 to 3, with 71 cases in the training set and 32 in the validation set. The baseline characteristics of the patients are summarized in **Table 1**.

Feature Reliability

Feature extraction was performed by two physicians to ensure the validity and reproducibility of the procedure. After

TABLE 1 | Demographic and clinical data of all patients.

Variables		All patients (n = 103)	Training set (n = 71)	Validation set (n = 32)	*p*
Age (years)	Range	59 (33–78)	56 (33–78)	63 (49–76)	0.03
Gender	Male	57	35	22	0.09
	Female	46	36	10	
Smoking	Smoker	45	26	19	0.03
	Non-smoker	58	45	13	
Histology	SCC	28	17	11	0.208
	No-SCC	75	54	21	
Stage	I	37	23	14	0.20
	II	24	16	8	
	III	30	23	7	
	IV	12	9	3	
SUVmax	Range	9.49 (0.88–23.5)	9.61 (0.88–23.5)	9.00 (1.19–21.0)	0.932

SCC, squamous cell carcinoma.

examining the inter-set differences with Manny–Whitney U-test, as well as the interobserver reliability with ICC, it was concluded that none of the features was significantly different from each other ($p > 0.05$), suggesting that all the features were reliable and reproducible (ICC > 0.75).

Correlations Between Radiomic Features and Immune Variables

By Mann–Whitney U test, 51 radiomic features were significantly different between PD-L1+ and PD-L1- patients ($p < 0.05$) (**Supplemental Table 2**). ROC for these indices showed moderate ability for predicting PD-L1 expression (AUC < 0.710), and the preferable features in differentiating PD-L1 status include SUVmax (AUC = 0.704) among the basic features and GLRLM_LRHGE (AUC = 0.702) and GLRLM_HGRE (AUC = 0.700) among the texture features.

Thirty-seven radiomic features correlated with CD8+ TILs in NSCLC (rho = -0.289 to 0.310, $p < 0.05$), among which NGLDM_Contrast has a strong correlation with CD8+ TILs with the largest linear correlation coefficient (rho = 0.310, $p = 0.001$; **Supplemental Table 3**).

In addition, PD-1+ TILs correlated with abundant radiomics indices, including 40 PET features and 28 CT features (rho = -0.317 to 0.356, $p < 0.05$; **Supplemental Table 3**). The strongest correlation was between SUVpeak (1 ml) and PD-1 expression (rho = 0.356, $p < 0.001$; **Supplemental Table 4**).

Feature Extraction

To avoid model overfitting, radiomic features with *p*-values <0.05 were first selected by the Mann–Whitney U test. Seventy-three features (42 PET features, 31 CT features) were found significantly different among TMIT groups in the training set. All of these features showed moderate power for predicting the TMIT-I tumors (**Figure 1**).

Construction of the Radiomics Signature and Complex Model

LASSO algorithm and 10-fold cross-validation were used to extract the optimal subset of radiomic features from the 73 features above. Four radiomic features were then selected to build the radiomics signature score based on the 71 patients in the training set (**Figure 2**) as follows: GLRLM_LRHGE,

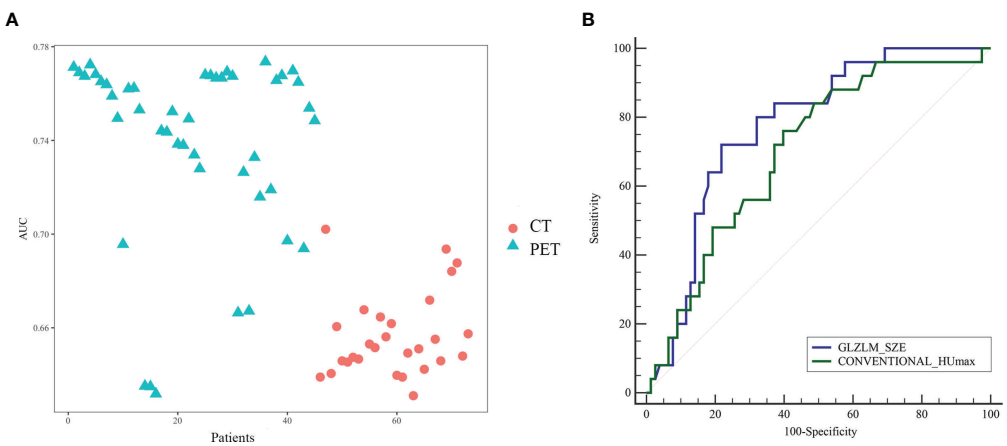


FIGURE 1 | **(A)** The performance of radiomic features for the evaluation of tumor microenvironment immune type (TMIT)-I tumors. All the features showed moderate power for predicting the TMIT-I tumors, and PET features have better ability than CT features with higher area under the receiver operating characteristic (ROC) curve (AUC). Panel **(B)** was the ROC curve of the optimal PET (GLZLM_SZE) and CT (Conventional_HUmax) features to distinguish TMIT-I from other groups.

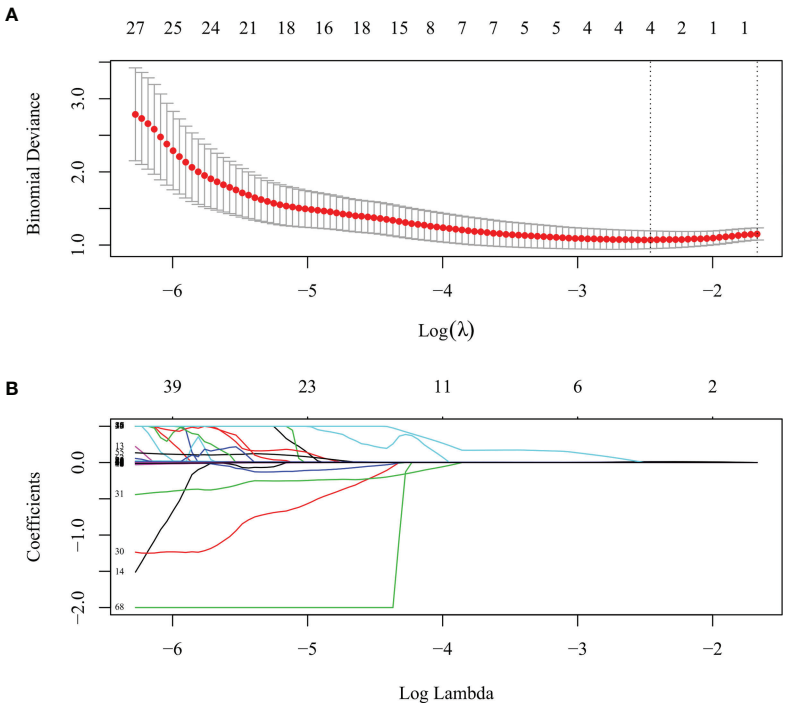


FIGURE 2 | The least absolute shrinkage and selection operator (LASSO) algorithm and 10-fold cross-validation were used to extract the optimal subset of radiomic features. **(A)** Tuning parameter (lambda, λ) selection in the LASSO model used 10-fold cross validation for the training set. The mean deviance (goodness-of-fit statistics, red dots) was plotted vs. log (λ), error bars displaying the range of standard error. Dotted vertical lines were drawn at the point of minimum deviance and at the point where maximum λ was obtained among errors smaller than the standard error of minimum deviance. **(B)** LASSO coefficient profiles of the 73 texture features.

GLZLM_SZE, SUVmax, and NGLDM_Contrast. The first three were PET features, and the last one was a CT feature. A rad-score for each patient was calculated using the following formula:

$$Rad-score = GLZLM\_SZE \times 0.6929504962 + GLRLM\_LRHGE \times 0.0001966283 + SUVmax \times 0.0707030000 + NGLDM\_Contrast \times 0.0086261012 - 2.5406198360 \quad (1)$$

The median and the range for the four selected radiomic features and the calculated rad-score are shown in **Table 2**. The rad-score and the four selected features were significantly different among the TMITs in both the training and the validation sets ($p < 0.05$). The rad-score for each patient in the two sets was displayed as a bar graph in **Figure 3**.

With multivariate logistic regression analysis (using backward stepwise elimination method), the combined model was constructed based on the clinical variables (age, gender, smoking history, stage) and radiomics signature. The formula was as follows:

$$Model-score = 1.668 \times rad-score + 1.481 \times smoking - 0.121 \quad (2)$$

TABLE 2 | The differences of four selected radiomic features and the calculated rad-score between TMITs.

Variables	Training set (n = 71)		*p*	Validation set (n = 32)		*p*
	TMIT-I (n = 18)	TMIT-II~IV (n = 53)		TMIT-I (n = 7)	TMIT-II~IV (n = 25)	
Rad-Score	-0.645 (-1.67 to 0.64)	-1.328 (-2.47 to -0.08)	0.000157	-0.734 (-1.223 to 0.123)	-1.286 (-2.39 to 0.289)	0.018895
PET features						
SUVmax	12.65 (5.10–23.50)	8.03 (0.875–19.50)	0.000634	13.70 (8.01–19.40)	8.13 (1.19–21.00)	0.024031
GLRLM_LRHGE	617 (85.90–2,680)	242 (18.50–1,440)	0.000750	664.0 (254–1,180)	276 (25.8–1770)	0.030368
GLZLM_SZE	0.697 (0.514–0.797)	0.548 (0.001–0.872)	0.000438	0.707 (0.421–0.763)	0.535 (0.017–0.791)	0.040220
CT feature						
NGLDM_Contrast	37.40 (0–72.80)	15.10 (0–62.60)	0.009743	30.7 (14.6–61.6)	21.30 (0–78.4)	0.171421

TMIT, tumor microenvironment immune type.

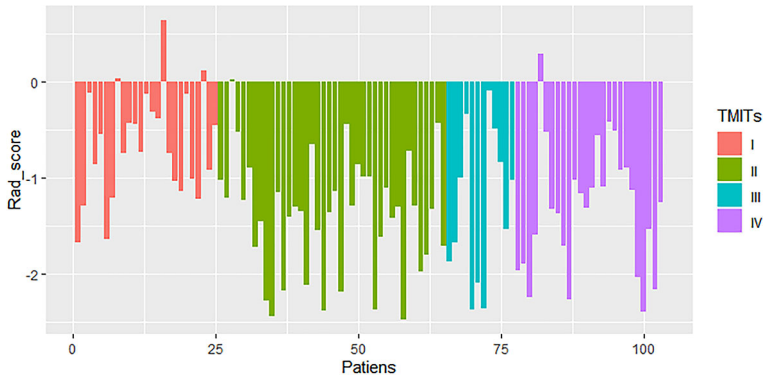


FIGURE 3 | Rad-score of patients in the cohort of patients with non-small cell lung cancer (NSCLC). Generally, rad-scores in the tumor microenvironment immune type (TMIT)-I tumors were higher than other TMIT-II~IV.

Performance of the Radiomics Signature and Clinical Features

We evaluated the models based on radiomics signature, clinical variables (smoking history), and the complex model in terms of their ability to predict TMIT-I tumors. The complex model had good predictive ability, and its AUCs in differentiating TMIT groups were 0.838 [95% CI (0.731–0.914)] in the training set and 0.811 [95% CI (0.634–0.927)] in the validation set. The predictive abilities of the four models, including sensitivity and specificity, were shown in **Table 3**. The differences of AUC values in different variables were shown in **Figure 4**. Notably, the AUCs of the complex model and smoking history were significantly different in the training set and validation set ($p = 0.0156$, $p = 0.0250$). The AUC values of the complex model and radiomics signature were not significantly different in either the training set or the validation set ($p > 0.05$).

Individualized Nomogram Construction and Validation

Given that the complex model based on both rad-score and clinical variables had better ability to predict TMIT-I tumors, we created a nomogram representing the individualized predictions based on the training set, which visualized the prediction result and the proportion of each factor (**Figure 5A**) . The calibration curves of the nomogram in the training and validation sets were presented in **Figures 5B, C** and showed good agreement between the predicted and observed values in the training set. DCA for the combined model (**Figure 6**) showed that prediction of TMIT-I tumors with the complex model added more benefit to SUVmax and the clinical variable (smoking history) in the training set. **Figure 7** showed that a representative patient with a TMIT-I type exhibited a hypermetabolic and heterogeneous tumor on ^{18}F-FDG PET, characterized by high expression of PD-L1 and high density of PD-1, CD8+ TILs.

DISCUSSION

The past decade was marked by a revolution in the treatment of NSCLC, including the variety of immunotherapy strategies targeting the tumor immune microenvironment (17–19). Biomarkers, such as TMIT-I, can identify the patient population that are more likely to respond to the immunotherapy (6). Consequently, novel approaches to assess the tumor immune microenvironment are of particular interest in clinical practice. We strived to address this need by proposing an ^{18}F-FDG-PET/CT-based radiomics to assess TMITs, especially TMIT-I tumors in pretreatment NSCLC patients. To the best of our knowledge, this is the first attempt to identify this type of immune "hot" tumors using PET/CT-based radiomics in pretreatment NSCLC patients.

Among the selected features in our work, numerous indices, including basic and texture features, were associated with PD-L1/PD-1 expression and CD8+ TILs. The metabolic characteristics of PD-L1/PD-1 expression in lung cancer were revealed in the previous study (10, 11). It seems that PD-L1-positive cells take up more glucose. Tumor microenvironment with high PD-L1

TABLE 3 | Predictive performance of variables in the training and validation sets.

Variables	Training set			Validation set		
	AUC (95% CI)	Sensitivity (%)	Specificity (%)	AUC (95% CI)	Sensitivity (%)	Specificity (%)
Model-score	0.838 (0.731–0.914)	72.22%	88.68%	0.811 (0.634–0.927)	85.71%	76.00%
Rad-Score	0.800 (0.688–0.885)	66.67%	81.13%	0.794 (0.615–0.916)	100%	56.00%
Smoking	0.738 (0.621–0.836)	72.22%	75.47%	0.699 (0.481–0.824)	85.71%	48.00%
SUVmax	0.771 (0.656–0.862)	72.22%	73.58%	0.783 (0.602–0.908)	85.71%	72.00%

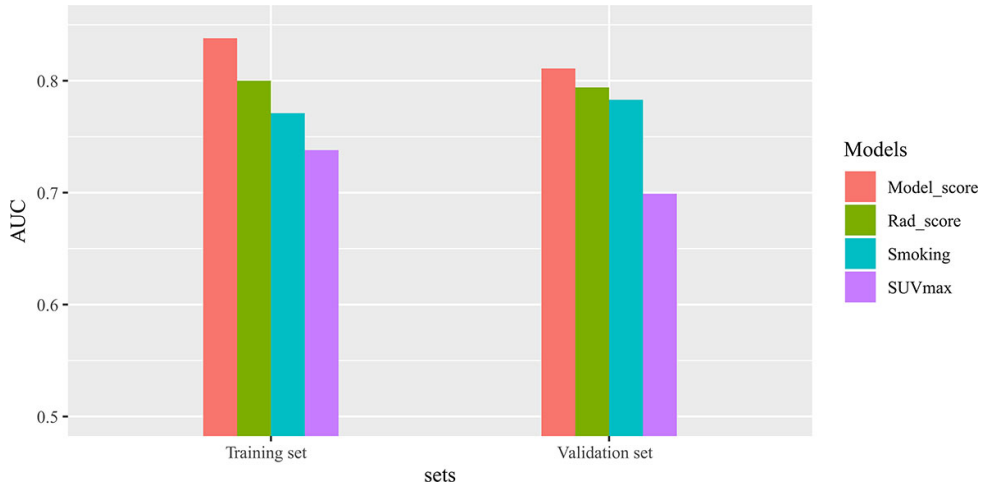


FIGURE 4 | The differences of area under the receiver operating characteristic (ROC) curve (AUC) values in different variables. There are significant differences in AUCs between Model-score and smoking in both the training and validation sets.

expression is often accompanied with dysfunctional anti-tumor immune responses, and therefore can foster immune tolerance that is favorable for tumor progression (20). However, the molecular mechanism between glucose metabolism and PD-L1 expression has not been fully revealed. Chang et al. (21) reported that PD-L1 expression maintained Akt/mammalian target of rapamycin (mTOR) signaling, which in turn promoted metabolic pathway through the translation of glycolysis enzymes. It might partly explain that the PD-L1/PD-1-positive tumors were more heterogeneous with more ^{18}F-FDG involvement in tumor cells. Moreover, the PD-L1 protein expression has been noted to be heterogeneous within different intertumoral regions, and the distribution of expression was frequently present near stromal tumor interfaces (22). Some tumors may display heterogeneous PD-L1 expression at different biopsied sites, which may partly explain the reason of mixed response to anti-PD-L1/PD-1 therapy (23, 24). On the other hand, the distribution of PD-L1 expression may cause different metabolic distributions of tumor cells. As is well known, the heterogeneity of image voxel intensities can be quantified by different image processing and analysis methods, including texture analysis, thus texture features describe the uniformity and heterogeneity of the PET images. These metabolic patterns could be representative of the intratumoral heterogeneous expression of PD-L1/PD-1.

Previously, quantitative CT radiomic features were extracted to predict PD-L1 expression in advanced-stage lung adenocarcinoma, yet their ability to predict PD-L1 positivity was weak (AUC = 0.661) (25). Recently, radiomics models of PET/CT demonstrated good performance in classifying a group of patients with PD-L1 expression, either ≥1% or ≥50%; however, TILs were not included in their research (26). Meanwhile, Jiang et al. (26) found that the performance of PET features was still unsatisfying, although the radiomics-based signatures from CT data achieved significant and robust individualized estimation of specific PD-L1 status. In this study, we used ^{18}F-FDG-PET/CT-based radiomics to analyze the correlation between radiomic features and PD-L1 expression. Among the numerous parameters, GLRLM_LRHGE derived from the PET images is the preferable feature to discriminate the PD-L1 status and achieved a moderate performance of predicting PD-L1. GLRLM reflects the comprehensive information of the image gray scale with respect to direction, adjacent interval, and variation amplitude (27), which is a set of statistical features extracted from medical images and frequently applied in radiomics (28, 29). Long-Run High Gray-level Emphasis (LRHGE) is the distribution of the long homogeneous runs with high gray levels. This may reflect that intertumoral regions with high PD-L1 expression are associated with high gray levels (or high metabolic distribution) on PET images.

Interestingly, NGLDM_Contrast had a strong correlation with CD8+TILs. NGLDM_Contrast measures the intensity difference between neighboring regions. Several studies have shown that preexisting tumoral and peritumoral immune infiltration correlates with patient response to anti-PD-1 and anti-PD-L1 therapy (30). CD8+ TILs are not evenly distributed within the tumor, where both T cell-infiltrated and T cell-excluded regions are present (31). It is likely that the heterogeneous distribution of CD8+ TILs contributed to the heterogeneity pattern of tumor metabolism, which was depicted by NGLDM_Contrast.

Rad-signature and complex model showed better predictive performances for TMIT-I tumors compared to the conventional features (SUVmax) and clinical variables probably because SUVmax alone does not accurately recapitulate the spatial heterogeneity of tumor metabolism (32). Radiomics aims to extract quantitative information from medical images that are difficult to be recognized or quantified by human eyes (33). Until recently, Sun et al. (34) developed a radiomics signature predictive of immunotherapy response by combining contrast-enhanced CT images and RNA-seq genomic data. The signature was able to discriminate inflamed tumors from immune-desert

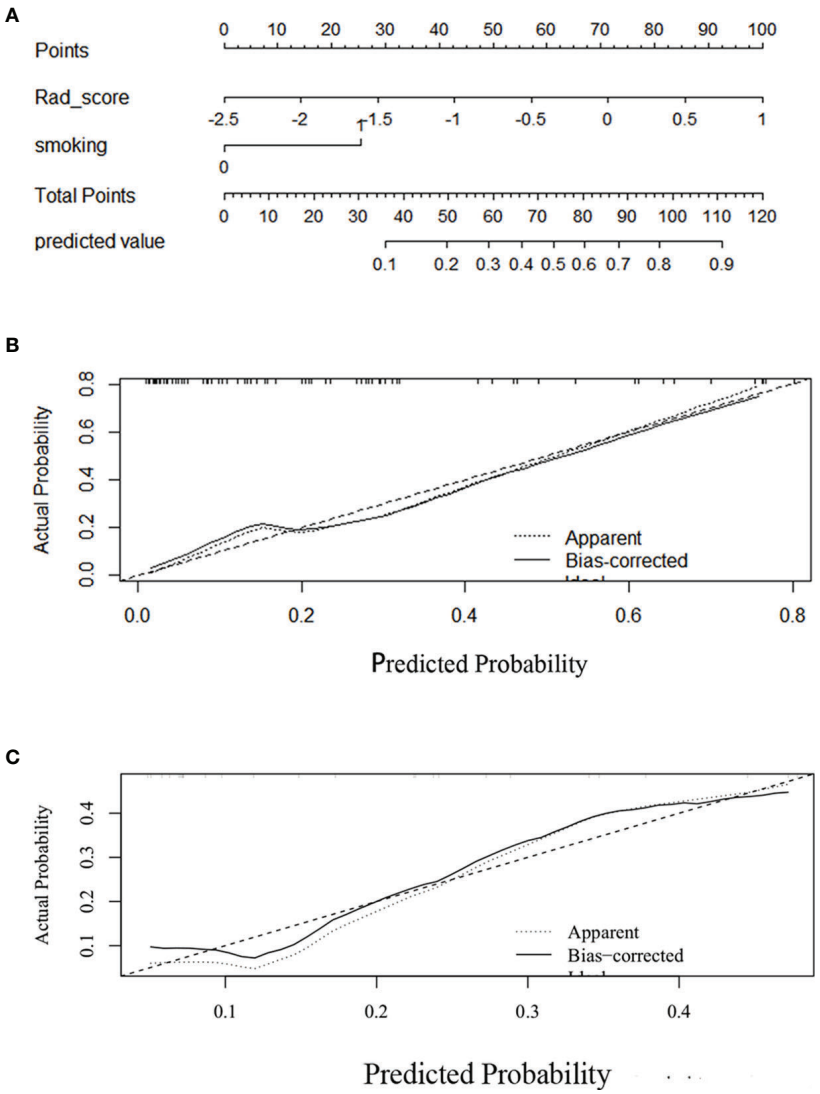


FIGURE 5 | Development and performance of a nomogram. **(A)** Nomogram based on rad-score and clinical factors (smoking history). Calibration curves **(B, C)** of the nomogram in the training set. The horizontal axis is the predicted incidence of the tumor microenvironment immune type (TMIT)-I tumors. The vertical axis is the observed incidence of the TMIT-I tumors. The diagonal line is the reference line, indicating that the predicted value is equal to the actual value. The prediction results and diagonals were basically coincident, indicating that the prediction results were accurate.

tumors, although with a modest AUC value of 0.76. Still, the ability of the radiomics signature to predict the gene expression signature of CD8 cells is unsatisfactory in the validation set (AUC = 0.67), underlying the importance of developing more and better imaging modality-based radiomics. We assessed the tumor immune microenvironment with ^{18}F-FDG-PET/CT radiomics and provided a promising way to predict the tumor immune phenotype. The nomogram included the radiomics signature score and clinical variables, which visualized the prediction results and provided an easy-to-use method for individualized prediction of TMIT-I tumors. In addition, radiomics-based signature could provide predicting outcomes at the time of image acquisition, providing a real-time guidance for patient stratification and therapeutic efficacy prediction.

DCA was used to facilitate the comparison between different prediction models. The utility of risk models may be evaluated with cost-effectiveness studies in clinical practice (35). DCA focuses on net benefit, which combines the number of true positives and false positives into a single "net" number (36, 37). In the TMIT example, the "net" values were calculated by subtracting the false positives (inconsistent biopsies showing other types of TMITs from the true positives TMIT I tumors confirmed by biopsies).

As seen in **Figure 6**, the clinical usefulness of each model was evaluated using DCA method by plotting the "net" benefit of using the model to stratify patients (y axis) against the continuum of potential thresholds for the probability of TMIT-1 tumors (x axis) (38). This study developed and validated a

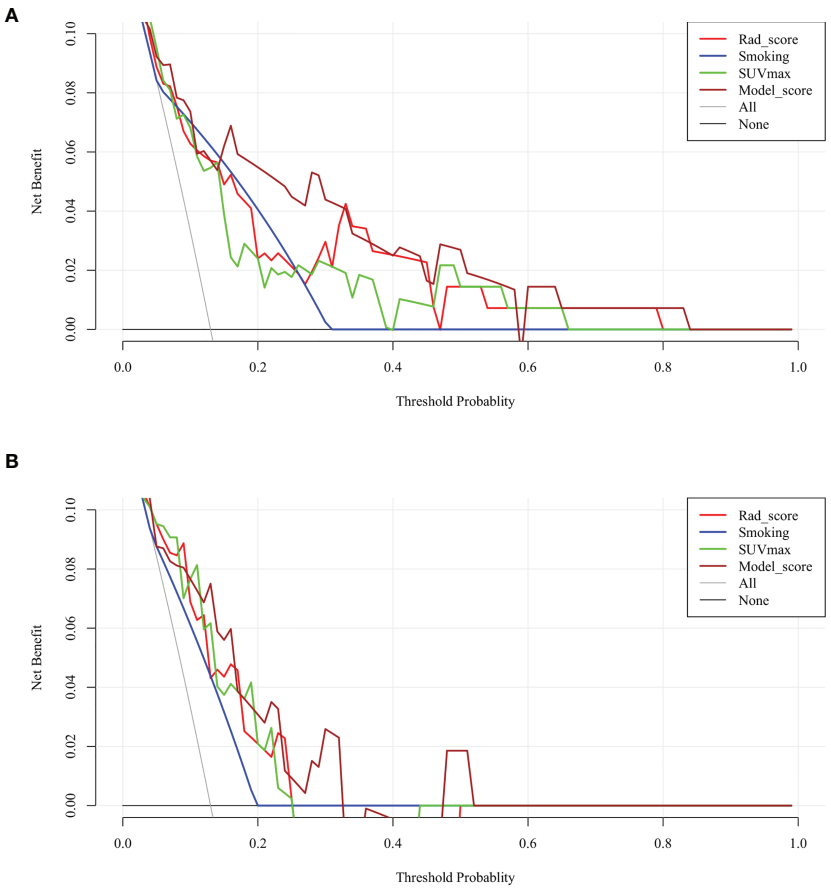


FIGURE 6 | Decision curve analysis (DCA) of each model in predicting tumor microenvironment immune type (TMIT)-I for non-small cell lung cancer (NSCLC). The vertical axis measures standardized net benefit. The horizontal axis shows the corresponding risk threshold. In the training set **(A)**, the DCA showed that if the threshold probability is between 0.1 and 0.8, using the complex model (brown line) provided a greater benefit than the clinical model (blue curve) and basic PET parameter (green curve). In the validation set **(B)**, the DCA showed that if the threshold probability is between 0.1 and 0.5, using the complex model provided a greater benefit than the clinical model.

complex model to identify NSCLC patients with Type-I TMIT. The novel approach was based on radiomic features, clinical variable, and ^{18}F-FDG uptake. The ^{18}F-FDG uptake accounted for intratumoral heterogeneity that correlated with underlying biological processes. The model described in our study showed good discriminative ability in both training and validation sets and exhibited higher predictive accuracy than conventional PET parameters (e.g., SUVmax). Within the range from 0.15 to 0.4 of the threshold probabilities, the model obviously showed a higher curve than that of SUVmax in **Figure 6**, indicating a much higher net benefit of our complex model than that of SUVmax. Therefore, this complex model obtained more true-positive cases of TMIT-I tumors and avoid more false-negative cases of other immune types. Considering the low probability of TMIT-I in clinical practice, it indicates that our DCA curve has a promising potential for clinical application. We agree that the DCA curve of the verification set is less optimal than that of the training set, which may require expansion of sample size and further optimization of the training model. Nevertheless, the DCA curves demonstrated advantages of complex model over radiomics, indicating clinical variable is also important.

To the best of our knowledge, there are no consensus cutoff values of PD-L1 and CD8+ TILs, even though the Food and Drug Administration (FDA) approved the cutoff of 50% tumor proportion score for first-line therapy with pembrolizumab and 1% tumor proportion score for second-line therapy with pembrolizumab/atezolizumab/bevacizumab (39). For the expression of PD-L1, we referred to a previous literature with a relatively large cohort and thus more reliable results (40). Koh et al. (39) evaluated PD-L1 immunohistochemistry based on the intensity and proportion of membranous and/or cytoplasmic staining in tumor cells. For CD8+ TILs, median or mean values were often used for classification of high or low infiltration (7, 41). Lin et al. (42) transferred continual variables like CD8+ T-cell infiltrating density and PD-1/PD-L1 mRNA expression level into categorical variables (high vs. low) with median value as cutoff point. Similarly, a recent assessment for PD-L1 was performed by Noh et al. (43), where PD-L1 expression was

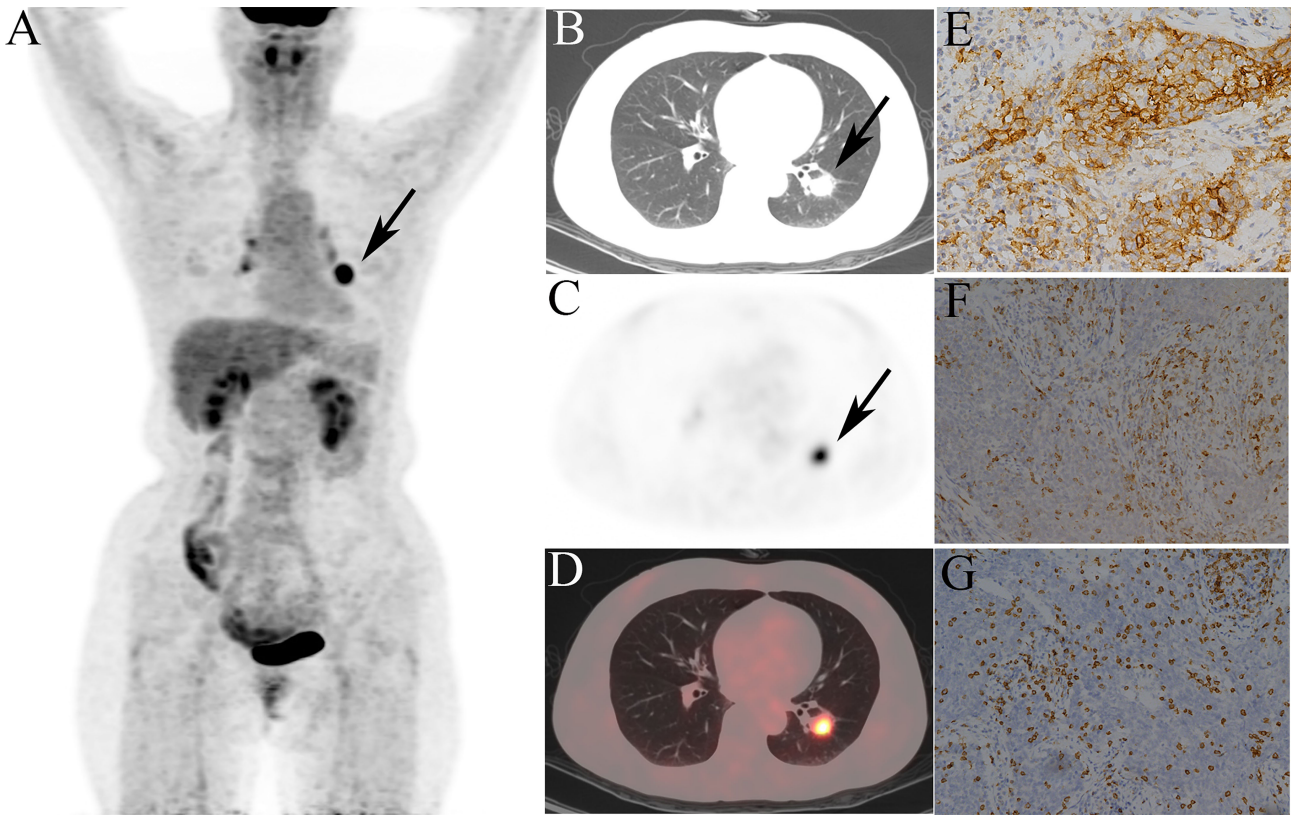


FIGURE 7 | Representative ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) images of a 52-year-old female patient, defined as tumor microenvironment immune type (TMIT)-I tumor. **(A)** MIP figure, **(B–D)** CT, PET, PET/CT fusion images. A nodule was in the lower lobe of the left lung (arrow) with markedly increased radioactivity, SUVmax: 10.3, GLRLM_LRHGE: 382, GLZLM_SZE: 0.682, and GLCM_Contrast: 14.6. The surgical pathology: poorly moderately differentiated squamous cell carcinoma. **(E)** High programmed death-ligand 1 (PD-L1) expression. **(F)** Programmed death-1 (PD-1) tumor-infiltrating lymphocytes (TILs) high density. **(G)** CD8+ TILs high density.

interpreted based on the proportion and intensity (graded as 0–3) of positive tumor cells. Besides, they utilized mean values as the cutoff threshold to categorize the CD8 TILs as "high" or "low." Based on the above, PD-L1+ was defined as more than 5% of tumor cells with staining intensity ≥2, and median value >99 was for CD8+ TILs in our study.

Our study has some limitations. First, it was of a single-center design and the relatively small sample size may influence the predictive ability of radiomics signature. Therefore, it is necessary to carry out multicenter studies and test multicenter data to ensure better robustness of the model. Second, patients with both lung squamous cell carcinoma and adenocarcinoma were enrolled and investigated, and the predictive performance of each tumor subtype should be further validated separately in a larger cohort. Third, with the development of quantitative imaging methods along with machine learning, it provides powerful modeling tools to mine the huge amount of image data available and reveal underlying complex biological mechanisms (44). Therefore, more advanced radiomics approaches, such as machine learning and deep learning, should be established to develop a model with optimal prediction performance.

CONCLUSION

In conclusion, a radiomics signature and complex model were developed and validated in patients with NSCLC. ^{18}F-FDG-PET/CT radiomics may provide a noninvasive method for predicting tumor immune phenotypes, which can assist in clinical practice to identify candidates for immunotherapy.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Huazhong

University of Science and Technology, Tongji Medical College affiliated Tongji Hospital (TJ-IRB 20181202).

AUTHOR CONTRIBUTIONS

XZ, JYZ, and SZ contributed to the conception and design of the study. JYZ and DK carried out the acquisition of data. JYZ, SZ, JZ, JY, and XZ performed the data analysis. JYZ wrote the first draft of the article. XZ, SZ, DK, JZ, and JY made the comments. XZ and JZ critically reviewed and revised the article. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the National Natural Science Foundation of China (91959119, 81873903, 81671718).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.769272/full#supplementary-material

REFERENCES

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin* (2021) 71(1):7–33. doi: 10.3322/caac.21654
2. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. *Adv Exp Med Biol* (2016) 893:1–19. doi: 10.1007/978-3-319-24223-1_1
3. Zhang M, Yang J, Hua W, Li Z, Xu Z, Qian Q, et al. Monitoring Checkpoint Inhibitors: Predictive Biomarkers in Immunotherapy. *Front Med* (2019) 13 (1):32–44. doi: 10.1007/s11684-018-0678-0
4. Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol Cancer Ther* (2015) 14(4):847–56. doi: 10.1158/1535-7163.MCT-14-0983
5. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, et al. PD-1 Blockade Induces Responses by Inhibiting Adaptive Immune Resistance. *Nature* (2014) 515(7528):568–71. doi: 10.1038/nature13954
6. Teng MW, Ngiow SF, Ribas A, Smyth MJ. Classifying Cancers Based on T-Cell Infiltration and PD-L1. *Cancer Res* (2015) 75(11):2139–45. doi: 10.1158/0008-5472.CAN-15-0255
7. Ock CY, Keam B, Kim S, Lee JS, Kim M, Kim TM, et al. Pan-Cancer Immunogenomic Perspective on the Tumor Microenvironment Based on PD-L1 and CD8 T-Cell Infiltration. *Clin Cancer Res* (2016) 22(9):2261–70. doi: 10.1158/1078-0432.CCR-15-2834
8. Takeuchi S, Khiewvan B, Fox PS, Swisher SG, Rohren EM, Bassett RL Jr, et al. Impact of Initial PET/CT Staging in Terms of Clinical Stage, Management Plan, and Prognosis in 592 Patients With Non-Small-Cell Lung Cancer. *Eur J Nucl Med Mol Imaging* (2014) 41(5):906–14. doi: 10.1007/s00259-013-2672-8
9. Scharping NE, Delgoffe GM. Tumor Microenvironment Metabolism: A New Checkpoint for Anti-Tumor Immunity. *Vaccines (Basel)* (2016) 4(4):1–15. doi: 10.3390/vaccines4040046
10. Lopci E, Toschi L, Grizzi F, Rahal D, Olivari L, Castino GF, et al. Correlation of Metabolic Information on FDG-PET With Tissue Expression of Immune Markers in Patients With Non-Small Cell Lung Cancer (NSCLC) Who Are Candidates for Upfront Surgery. *Eur J Nucl Med Mol Imaging* (2016) 43 (11):1954–61. doi: 10.1007/s00259-016-3425-2
11. Kaira K, Shimizu K, Kitahara S, Yajima T, Atsumi J, Kosaka T, et al. 2-Deoxy-2-[Fluorine-18] Fluoro-D-Glucose Uptake on Positron Emission Tomography Is Associated With Programmed Death Ligand-1 Expression in Patients With Pulmonary Adenocarcinoma. *Eur J Cancer* (2018) 101:181–90. doi: 10.1016/j.ejca.2018.06.022
12. Lee JW, Lee SM. Radiomics in Oncological PET/CT: Clinical Applications. *Nucl Med Mol Imaging* (2018) 52(3):170–89. doi: 10.1007/s13139-017-0500-y
13. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to Radiomics. *J Nucl Med* (2020) 61(4):488–95. doi: 10.2967/jnumed.118.222893
14. Inaguma S WZ, Lasota J, Sarlomo-Rikala M, McCue PA, Ikeda H, Miettinen M. Comprehensive Immunohistochemical Study of Programmed Cell Death Ligand 1 (PD-L1): Analysis in 5536 Cases Revealed Consistent Expression in Trophoblastic Tumors. *Am J Surg Pathol* (2016) 40(8):1133–42. doi: 10.1097/PAS.0000000000000653
15. Zhou J, Zou S, Cheng S, Kuang D, Li D, Chen L, et al. Correlation Between Dual-Time-Point FDG PET and Tumor Microenvironment Immune Types in Non-Small Cell Lung Cancer. *Front Oncol* (2021) 11(807):559623. doi: 10.3389/fonc.2021.559623
16. Kirienko M, Cozzi L, Rossi A, Voulaz E, Antunovic L, Fogliata A, et al. Ability of FDG PET and CT Radiomics Features to Differentiate Between Primary and Metastatic Lung Lesions. *Eur J Nucl Med Mol Imaging* (2018) 45 (10):1649–60. doi: 10.1007/s00259-018-3987-2
17. Mok TSK, Wu YL, Kudaba I, Kowalski DM, Cho BC, Turna HZ, et al. Pembrolizumab Versus Chemotherapy for Previously Untreated, PD-L1-Expressing, Locally Advanced or Metastatic Non-Small-Cell Lung Cancer (KEYNOTE-042): A Randomised, Open-Label, Controlled, Phase 3 Trial. *Lancet* (2019) 393(10183):1819–30. doi: 10.1016/S0140-6736(18)32409-7
18. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Pembrolizumab Versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *N Engl J Med* (2016) 375(19):1823–33. doi: 10.1056/NEJMoa1606774
19. Langer CJ, Gadgeel SM, Borghaei H, Papadimitrakopoulou VA, Patnaik A, Powell SF, et al. Carboplatin and Pemetrexed With or Without Pembrolizumab for Advanced, Non-Squamous Non-Small-Cell Lung Cancer: A Randomised, Phase 2 Cohort of the Open-Label KEYNOTE-021 Study. *Lancet Oncol* (2016) 17(11):1497–508. doi: 10.1016/S1470-2045(16)30498-3
20. Grizzi F, Castello A, Lopci E. Is It Time to Change Our Vision of Tumor Metabolism Prior to Immunotherapy? *Eur J Nucl Med Mol Imaging* (2018) 45 (6):1072–5. doi: 10.1007/s00259-018-3988-1
21. Chang CH, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, et al. Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. *Cell* (2015) 162(6):1229–41. doi: 10.1016/j.cell.2015.08.016
22. McLaughlin J, Han G, Schalper KA, Carvajal-Hausdorf D, Pelekanou V, Rehman J, et al. Quantitative Assessment of the Heterogeneity of PD-L1 Expression in Non-Small-Cell Lung Cancer. *JAMA Oncol* (2016) 2(1):46–54. doi: 10.1001/jamaoncol.2015.3638
23. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, Activity, and Immune Correlates of Anti-PD-1 Antibody in Cancer. *N Engl J Med* (2012) 366(26):2443–54. doi: 10.1056/NEJMoa1200690
24. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, et al. Safety and Activity of Anti-PD-L1 Antibody in Patients With Advanced Cancer. *N Engl J Med* (2012) 366(26):2455–65. doi: 10.1056/NEJMoa1200694
25. Yoon J, Suh YJ, Han K, Cho H, Lee HJ, Hur J, et al. Utility of CT Radiomics for Prediction of PD-L1 Expression in Advanced Lung Adenocarcinomas. *Thorac Cancer* (2020) 11(4):993–1004. doi: 10.1111/1759-7714.13352
26. Jiang M, Sun D, Guo Y, Guo Y, Xiao J, Wang L, et al. Assessing PD-L1 Expression Level by Radiomic Features From PET/CT in Nonsmall Cell Lung Cancer Patients: An Initial Result. *Acad Radiol* (2020) 27(2):171–9. doi: 10.1016/j.acra.2019.04.016
27. Gao J, Huang X, Meng H, Zhang M, Zhang X, Lin X, et al. Performance of Multiparametric Functional Imaging and Texture Analysis in Predicting Synchronous Metastatic Disease in Pancreatic Ductal Adenocarcinoma Patients by Hybrid PET/MR: Initial Experience. *Front Oncol* (2020) 10:198. doi: 10.3389/fonc.2020.00198
28. Aide N, Talbot M, Fruchart C, Damaj G, Lasnon C, et al. Diagnostic and Prognostic Value of Baseline FDG PET/CT Skeletal Textural Features in Diffuse Large B Cell Lymphoma. *Eur J Nucl Med Mol Imaging* (2018) 45 (5):699–711. doi: 10.1007/s00259-017-3899-6

29. Liang B, Yan H, Tian Y, Chen X, Yan L, Zhang T, et al. Dosiomics: Extracting 3D Spatial Features From Dose Distribution to Predict Incidence of Radiation Pneumonitis. *Front Oncol* (2019) 9:269. doi: 10.3389/fonc.2019.00269
30. Chen DS, Mellman I. Elements of Cancer Immunity and the Cancer-Immune Set Point. *Nature* (2017) 541(7637):321–30. doi: 10.1038/nature21349
31. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the Tumor Immune Microenvironment (TIME) for Effective Therapy. *Nat Med* (2018) 24(5):541–50. doi: 10.1038/s41591-018-0014-x
32. Shen L-F, Zhou S-H, Yu Q. Predicting Response to Radiotherapy in Tumors With PET/CT: When and How? *Trans Cancer Res* (2020) 9(4):2972–81. doi: 10.21037/tcr.2020.03.16
33. 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(4):441–6. doi: 10.1016/j.ejca.2011.11.036
34. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A Radiomics Approach to Assess Tumour-Infiltrating CD8 Cells and Response to Anti-PD-1 or Anti-PD-L1 Immunotherapy: An Imaging Biomarker, Retrospective Multicohort Study. *Lancet Oncol* (2018) 19(9):1180–91. doi: 10.1016/S1470-2045(18)30413-3
35. Hlatky MA, Greenland P, Arnett DK, Ballantyne CM, Criqui MH, Elkind MS, et al. Criteria for Evaluation of Novel Markers of Cardiovascular Risk: A Scientific Statement From the American Heart Association. *Circulation* (2009) 119(17):2408–16. doi: 10.1161/CIRCULATIONAHA.109.192278
36. Vickers AJ, Van Calster B, Steyerberg EW. Net Benefit Approaches to the Evaluation of Prediction Models, Molecular Markers, and Diagnostic Tests. *BMJ* (2016) 352:i6. doi: 10.1136/bmj.i6
37. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Making* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361
38. Hijazi Z, Oldgren J, Lindbäck J, Alexander JH, Connolly SJ, Eikelboom JW, et al. The Novel Biomarker-Based ABC (Age, Biomarkers, Clinical History)-Bleeding Risk Score for Patients With Atrial Fibrillation: A Derivation and Validation Study. *Lancet* (2016) 387(10035):2302–11. doi: 10.1016/S0140-6736(16)00741-8
39. Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman JR, Bharat A, et al. Nccn Guidelines Insights: Non-Small Cell Lung Cancer, Version 2.2021. *J Natl Compr Canc Netw* (2021) 19(3):254–66. doi: 10.6004/jnccn.2021.0013
40. Koh J, Go H, Keam B, Kim MY, Nam SJ, Kim TM, et al. Clinicopathologic Analysis of Programmed Cell Death-1 and Programmed Cell Death-Ligand 1 and 2 Expressions in Pulmonary Adenocarcinoma: Comparison With Histology and Driver Oncogenic Alteration Status. *Mod Pathol* (2015) 28 (9):1154–66. doi: 10.1038/modpathol.2015.63
41. Noh BJ KJ, Eom DW. Prognostic Significance of Categorizing Gastric Carcinoma by PD-L1 Expression and Tumor Infiltrating Lymphocytes. *Ann Clin Lab Sci* (2018) 48(6):695–706 http://www.annclinlabsci.org/content/48/6/695.long.
42. Lin Z, Gu J, Cui X, Huang L, Li S, Feng J, et al. Deciphering Microenvironment of NSCLC Based on CD8+ TIL Density and PD-1/PD-L1 Expression. *J Cancer* (2019) 10(1):211–22. doi: 10.7150/jca.26444
43. Noh BJ, Kwak JY, Eom DW. Immune Classification for the PD-L1 Expression and Tumour-Infiltrating Lymphocytes in Colorectal Adenocarcinoma. *BMC Cancer* (2020) 20(1):58. doi: 10.1186/s12885-020-6553-9
44. Avanzo M, Wei L, Stancanello J, Vallières M, Rao A, Morin O, et al. Machine and Deep Learning Methods for Radiomics. *Med Phys* (2020) 47(5):e185–202. doi: 10.1002/mp.13678

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

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 02 December 2021
doi: 10.3389/fonc.2021.788424

Preoperative CT-Based Radiomics Combined With Nodule Type to Predict the Micropapillary Pattern in Lung Adenocarcinoma of Size 2 cm or Less: A Multicenter Study

Meirong Li[1], *Yachao Ruan*[1], *Zhan Feng*[1]*, *Fangyu Sun*[2], *Minhong Wang*[3] *and Liang Zhang*[4]

[1] Department of Radiology, The First Affiliated Hospital, Zhejiang University, School of Medicine, Hangzhou, China, [2] Department of Radiology, Xiaoshan Hospital of Traditional Chinese Medicine, Hangzhou, China, [3] Department of Radiology, First Affiliated Hospital of Wannan Medical College, Wuhu, China, [4] Department of Radiology, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, China

OPEN ACCESS

Edited by:
Yiyan Liu,
University of Louisville, United States

Reviewed by:
Beatrice Aramini,
University Hospital of Modena, Italy
Jing Shen,
Affiliated Zhongshan Hospital of Dalian University, China

***Correspondence:**
Zhan Feng
gerxyuan@zju.edu.cn

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 02 October 2021
Accepted: 15 November 2021
Published: 02 December 2021

Citation:
Li M, Ruan Y, Feng Z, Sun F, Wang M and Zhang L (2021) Preoperative CT-Based Radiomics Combined With Nodule Type to Predict the Micropapillary Pattern in Lung Adenocarcinoma of Size 2 cm or Less: A Multicenter Study. Front. Oncol. 11:788424. doi: 10.3389/fonc.2021.788424

Purpose: To construct an optimal radiomics model for preoperative prediction micropapillary pattern (MPP) in adenocarcinoma (ADC) of size ≤ 2 cm, nodule type was used for stratification to construct two radiomics models based on high-resolution computed tomography (HRCT) images.

Materials and Methods: We retrospectively analyzed patients with pathologically confirmed ADC of size ≤ 2 cm who presented to three hospitals. Patients presenting to the hospital with the greater number of patients were included in the training set (n = 2386) and those presenting to the other two hospitals were included in the external validation set (n = 119). HRCT images were used for delineation of region of interest of tumor and extraction of radiomics features; dimensionality reduction was performed for the features. Nodule type was used to stratify the data and the random forest method was used to construct two models for preoperative prediction MPP in ADC of size ≤ 2 cm. Model 1 included all nodule types and model 2 included only solid nodules. The receiver operating characteristic curve was used to assess the prediction performance of the two models and independent validation was used to assess its generalizability.

Results: Both models predicted ADC with MPP preoperatively. The area under the curve (AUC) of prediction performance of models 1 and 2 were 0.91 and 0.78, respectively. The prediction performance of model 2 was lower than that of model 1. The AUCs in the external validation set were 0.81 and 0.72, respectively. The DeLong test showed statistically significant differences between the training and validation sets in model 1 (p = 0.0296) with weak generalizability. There was no statistically significant difference between the training and validation sets in model 2 (p = 0.2865) with some generalizability.

Conclusion: Nodule type is an important factor that affects the performance of radiomics predictor model for MPP with ADC of size ≤ 2 cm. The radiomics prediction model constructed based on solid nodules alone, can be used to evaluate MPP and may contribute to proper surgical planning in patients with ADC of size ≤ 2 cm.

Keywords: lung adenocarcinoma, radiomics model, micropapillary pattern (MPP), multicenter, computed tomography

INTRODUCTION

With widespread use of high-resolution CT (HRCT), lung cancer is increasingly being detected at an early stage and small peripheral lung cancers are increasingly treated with surgical resection. Adenocarcinoma (ADC) is the most common lung cancer type. According to the World Health Organization classification of lung ADC (1), invasive adenocarcinoma (IAC) is divided into five pathological types, namely lepidic, acinar, papillary, micropapillary, and solid subtypes. The micropapillary pattern (MPP) is a marker for poor outcome (2–6). Lee (7) found that the overall survival and disease-free survival were worse in cases with minimal areas of MPP (accounting for 1-5% of the entire tumor). Since this adverse prognostic effect, identification of MPP may potentially be instructive for surgical plan and further aggressive adjuvant treatment.

Tumor size and MPP are important prognostic factors for surgical outcome in patients with early stage lung ADC. The use of limited resection has gradually increased for patients with non-small cell lung cancer with size ≤ 2 cm. Notably, MPP accounting for more than 5% of the entire tumor is an independent risk factor for recurrence and poor outcome of lung ADC with size ≤ 2 cm (3, 8, 9), suggesting that limited resection may not be the optimal surgical approach for such patients. Thus, preoperative confirmation of MPP (constituting> 5%) in ADC with size ≤ 2 cm has importance for selection of surgical procedure. Due to technical reasons, preoperative histological examination cannot be performed for many peripheral small tumors.

Radiomics can objectively and quantitatively analyze imaging features that reflect tumor heterogeneity. Previous studies showed that radiomics can be used to preoperatively detect ADC with MPP or MPP/solid pattern (10–18). However, previous reports of predicting MPP subtypes using radiomic analyses are with some limitations. First, previous studies have been conducted on the predictor MPP for T1-stage ADC (17, 18), including patients of tumors with size ≥ 2 cm. Because the preferred surgical procedures for ADC with sizes ≤ 2 cm or > 2 cm are different, prediction of MPP in ADC with size ≤ 2 cm can aid in determining the optimal resection method. In addition, previous studies considered only tumor size and overlooked different imaging characteristics between ground glass opacity (GGO) and solid nodule. These radiomics studies combined solid and GGO nodules but the unequal number of these two nodule types may have introduced bias in these study results. Furthermore, these studies were performed on small populations and lacked external validation.

Therefore, the purpose of our retrospective study was to develop an optimal radiomics model for preoperative prediction of MPP with ADC of size ≤ 2 cm. First, we included a large number of ADC patients with peripheral tumor size ≤ 2 cm from three institutions. Second, nodule type was used for stratification and it was combined with HRCT radiomics characteristics to construct two models. Model 1 included all nodule types (solid and GGO). Because most ADC with MPP of size ≤ 2 cm were solid (19), the GGO type was excluded for model 2; only solid nodules were used to construct model. Third, independent external validation was used to validate the two model's generalization ability.

MATERIALS AND METHODS

Patients

This was a retrospective study, which was approved by the institutional review board of the First Affiliated Hospital of Zhejiang University School of Medicine, the First Affiliated Hospital of Wannan Medical College and Zhejiang Cancer Hospital. Informed consent by the patients was waived by the hospital ethics committee. Patients diagnosed with ADC between September 2019 and January 2021 were selected from three study centers.

Patients were included in the study if they had a tumor with size on CT images < 2 cm, peripheral nodules on chest CT, no marked cavitation of lesions, histologic subtype without solid component, and surgery within 1 week of CT examination.

We excluded patients with history of neoadjuvant chemotherapy or radiotherapy, lung cancer surgery in the past 2 years, simultaneous multiple cancers, and patients with multifocal lesions.

We included 2,386 patients (mean age: 51.62 ± 13.13 years; range: 27–76 years) from Hospital 1 in this study. The proportion of female patients was 76.3%, 194 patients had MPP and 2,192 did not have MPP. Of the 2386 patients, 400 had solid nodule and 1,792 had GGO.

The independent external validation database consisted of 119 patients from the other two hospitals. The proportion of female patients was 66.4% and the average age was 60.56 ± 9.51 years (range: 29–80 years). There were 65 patients with MPP and 54 without MPP. Of the 119 patients, 80 had solid nodule and 39 had GGO.

To imbalance data numbers between the two groups, when model 1 (including GGO and solid nodules) was constructed, 2,192 ADC patients without MPP from hospital 1 were matched

to 195 controls. Finally, Model 1 included 389 patients from hospital 1 and 119 patients from the other two hospitals in the training and validation sets, respectively. When constructing model 2 (only including solid nodules), 400 and 108 patients from hospitals 1 and the other two hospitals respectively, were included in the training and validation sets. **Figure 1** shows the model construction flow chart. The clinical variables were retrospectively reviewed from the electronic medical records.

CT Imaging Acquisition

Pulmonary CT imaging was performed for all patients, using six CT scanners from Philips Healthcare (iCT 256 and 16-slice CT), Siemens Healthineers (64-slice CT), and GE (64-slice CT). All scans were acquired with a deep inhaled breath held in the supine position, without contrast. The scan covered from the thoracic inlet to bilateral adrenal glands. A sharp reconstruction algorithm was used. CT imaging parameters used in the three institutions were as follows: tube voltage of 100 or 120 kev, automatic tube current modulation; reconstructed slice thickness 1-3 mm. and reconstruction interval 0.625-1 mm. The images were displayed in two gray scales for interpretation of lung (width 1500 HU, level −430 HU) and mediastinal (width 200 HU, level 40 HU) windows.

Histologic Evaluation and CT Features

Histological subtype was independently evaluated by two experienced radiologists and discrepancies were resolved through consensus. According to the 2015 World Health Organization classification of lung tumors (1), ADC histologic subtypes were recorded using a semiquantitative assessment of each subtype in 5% increments. In our study, according to the amount of micropapillary component, patients were divided into group 1 (ADC with MPP constituting > 5% of the entire tumor) and group 2 (ADC without MPP or MPP < 5% of the entire tumor). Based on the guidelines from the Fleischer society (20) nodule type (pure ground glass opacity, mixture ground glass opacity or solid) was determined by two radiologists on the lung window setting. In our study, the pure ground glass opacity (pGGO) and mixture ground glass opacity(mGGO) nodules were classified as GGO.

Tumor Segmentation and Radiomics Feature Extraction

In this study, the Dr. Wise® research platform was used for radiomics analysis. All patient images were downloaded and processed in the raw DICOM format and images were transferred to the post-processing workstation. One radiologist manually labeled the lesion region in thin-layer HRCT using the raw dataset (**Figure 2**) to avoid bronchovascular bundles and normal lung parenchyma. Then, the second radiologist confirmed the final regions of interest with consensus.

Image Preprocessing

The radiomics features included first-order features, tumor morphologic features, gray-level co-occurrence matrix describing the tumor internal and surface textures, gray-level run-length matrix, gray-level size zone matrix, and gray level dependence matrix texture features. A total of 105 radiomics features were extracted from every ROI and Z-score standardization was performed.

Feature Dimensionality Reduction and Model Construction

Spearman's correlation analysis was performed for radiomics features using a correlation coefficient threshold of 0.8. Then, a tree-based algorithm was used for feature dimensionality reduction. The random forest method was used to construct two prediction models, based on the nodule types combined with radiomics features extracted from plain HRCT images. Model 1 contained solid nodules and GGO, while model 2 contained only solid nodules. Ten cross-validations were performed on the training set to analyze model stability. The ROC curve was used to evaluate the diagnostic performance of two models in predicting ADC with MPP in the training and validation sets.

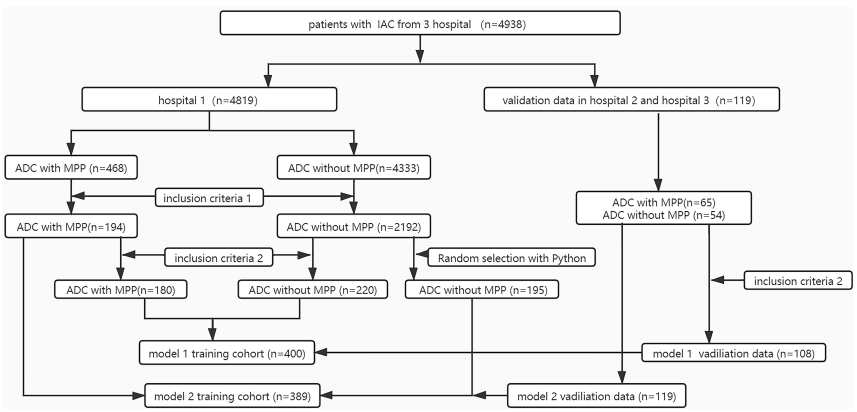


FIGURE 1 | Flowchart selection patients of the two-model dataset. Inclusion criteria 1: tumor size measured in CT images <2 cm;2) the presence of peripheral nodules on chest CT; 3) the lesions without marked cavity;4) histologic subtype without solid pattern; Inclusion criteria 2: the nodule manifest as solid on chest CT.

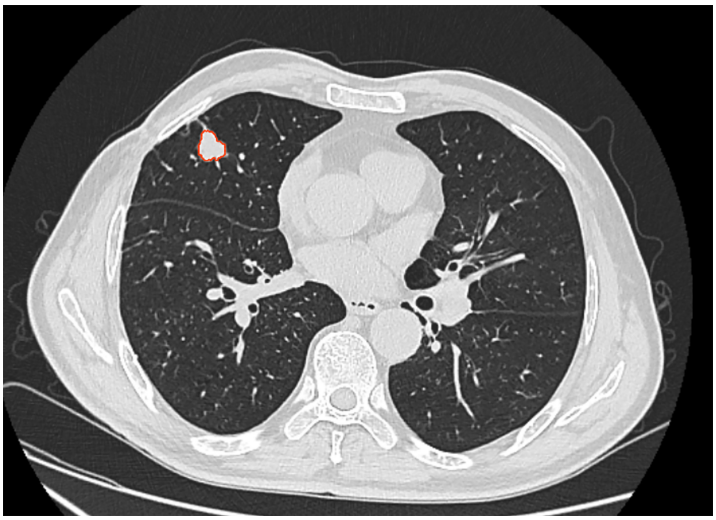


FIGURE 2 | 52 years-old patient, male, the lesion located in right middle lung. The area inside the red line represents the ROI for the tumor.

Finally, the generalizability of the two models was evaluated using the external validation set.

Statistical Analysis

R software (version: 3.4.1; R Foundation for Statistical Computing, Vienna, Austria; http://www.Rproject.org) was used for the statistical analyses. Quantitative data were expressed as mean ± standard deviation (SD), while qualitative data were expressed as frequency (%). Qualitative variables (gender, nodule type, and ADC with/without MPP) were compared using the Chi-square test. Continuous variables (age) were evaluated using a two-sample t-test or Mann-Whitney U test.

The area under the ROC curve (AUC), 95% confidence interval (CI), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the two models were calculated for the training and validation sets. DeLong test was used to compare AUC differences between the training and validation sets for the two models. A two-tailed difference of $p < 0.05$ was deemed to be statistically significant.

RESULT

Dataset Characteristics

The clinical characteristics of the two model databases are summarized in **Tables 1**, **2**.

Model 1 included 389 patients from the training set. ADC patients with MPP were aged 31–87 years (61.8 ± 10.9), 46% were females, 93% had solid nodules, and 7% had GGOs. ADC patients without MPP were aged 20–81 years (60.8 ± 11.1), 71% were females, 31% had solid nodules, and 69% had GGOs. There were 119 patients in the validation set. In the validation set, the age of ADC patients with MPP was 35–81 years (60.5 ± 9.3), 51% were females, 92% had solid nodules, and 8% had GGOs. ADC patients without MPP were aged 29–84 years (60.56 ± 9.91), 72% were females, 37% had solid nodules, and 63% had GGOs.

There were 400 patients in Model 2 from the training set. In the training set, ADC patients with MPP were aged 31–83 years (62 ± 10.8) and 47% were females. ADC patients without MPP were aged 26–83 years (61.4 ± 10.9) and 68% were females. There were 108 patients in the validation set. In the validation set,

TABLE 1 | Demographic and clinical characteristics of patients on different datasets of model 1.

Model 1	Training		P	validation		P
	MPP (n = 194)	Without MIP (n = 195)		MPP (n = 65)	Without MPP (n = 54)	
Age	61.8 ± 10.9	60.8 ± 11.1	0.25	60.5 ± 9.3	60.6 ± 9.9	0.29
Gender						0.045
Man	104 (54%)	56 (29%)	<0.001	31 (48%)	15 (32%)	
Woman	90 (46%)	139 (71%)		34 (51%)	39 (72%)	
Nodule type			<0.001			<0.001
Solid	180 (93%)	61 (31%)		60 (92%)	20 (37%)	
GGO	14 (7%)	134 (69%)		5 (8%)	34 (63%)	

TABLE 2 | Demographic and clinical characteristics of patients on different datasets of model 2.

Model 2	Training		P	validation		P
	MPP (n=180)	Without MPP (n = 220)		MPP (n = 60)	Without MPP (n = 48)	
Age	62 ± 10.8	61.4 ± 10.9	0.28	61.1 ± 9.2	60.8 ± 9.2	0.19
Gender			<0.001			0.012
Man	96 (53%)	70 (32%)		31 (52%)	11 (29%)	
Woman	84 (47%)	150 (68%)		29 (48%)	37 (71%)	

ADC patients with MPP were aged 35–81 years (61 ± 9.2) and 48% were females. ADC patients without MPP were aged 29–84 years (60.8 ± 9.2) and 68% were females.

No statistically significant difference was observed in the age between the training and validation sets in the two models, but statistically significant differences in gender were observed. The difference in solid nodule ratio was statistically significant between the two groups.

Feature Selection

Tables 3, **4** shows the radiomics characteristics used in the two models. Eight radiomics characteristics were included in model 1: gray level matrix (GLDM; n = 4) and first order (n = 4). Eighteen optimal radiomics characteristics were included in model 2: gray level co-occurrence matrix (GLCM; n = 6); first order (n = 4), shape feature (n = 4), neighborhood gray tone difference matrix (NGTDM; n = 2) and gray level size zone matrix (GLSZM; n = 2).

Evaluation of Model Prediction Performance

The AUC values for the two radiomics models in training and validation cohort were shown in **Table 5**. The ROC curve results showed that model 1 had excellent preoperative prediction for ADC with MPP. In the training set, AUC was 0.91 (95% CI 0.88–0.94), accuracy was 0.79, sensitivity was 0.73, specificity was 0.86, PPV was 0.79, and NPV was 0.82. In the validation set, AUC was 0.82 (95% CI 0.74–0.89), accuracy was 0.83, sensitivity was 0.76, specificity was 0.88, PPV was 0.83, and NPV was 0.83.

Model 2 showed good preoperative prediction performance for ADC with MPP. In the training set, AUC was 0.78 (95%CI 0.74–0.82), accuracy was 0.73, sensitivity was 0.77, specificity was 0.6, PPV was 0.73, and NPV was 0.65. In the training set, AUC was 0.76 (95% CI 0.63–0.82), accuracy was 0.76, sensitivity was 0.49, specificity was 0.85, PPV was 0.75, and NPV was 0.65.

The DeLong test showed that there were statistically significant differences between the training and validation sets in model 1 (p = 0.0296) (**Figure 3A**). There was no statistically significant difference between the training and validation sets in model 2 (p = 0.2865) (**Figure 3B**).

DISCUSSION

In this multicenter study, stratification was based on nodule types to construct two radiomics models for preoperative prediction the MPP in peripheral lung ADC with size ≤ 2 cm. The AUC of model 1 (which included solid nodules and GGO) and model 2 (solid nodules only) were 0.91 and 0.78, respectively. Both models had good prediction ability but the prediction performance of model 2 was lower compared to that of model 1, suggesting that GGO affects prediction model performance. In the external validation set, the AUCs for the two models were 0.82 and 0.72, respectively. The DeLong test suggested a difference in AUC between the training and validation sets for model 1, showing poor generalizability. However, no significant difference was noted in AUC between the training and validation sets in model 2, showing good generalizability. Therefore, radiomics can be used as a convenient and non-invasive biomarker for preoperative prediction MPP in peripheral lung ADC with size ≤ 2 cm and to guide the diagnosis and treatment.

Previous studies proved that tumor size is an independent factor for postoperative prognosis in non-small cell lung cancer. Additionally, lung ADC with size 2.1–3 cm has a significantly higher local recurrence rate than ADC < 2 cm (21). Tumor size

TABLE 3 | Selected radiomic features for the prediction model 1.

Feature class	Feature name	Feature coefficient	Weight
First order	Mean	0.207	1
GLDM	LDHGLE	0.1884	0.9104
First order	Energy	0.1824	0.8815
First order	10 Percentile	0.1016	0.491
First order	Minimum	0.0833	0.4027
GLDM	SDLGLE	0.0809	0.391
GLDM	LDLGLE	0.0784	0.3786
GLDM	Contrast	0.0779	0.3765

TABLE 4 | Selected radiomic features for the prediction model 2.

Feature class	Feature name	Feature coefficient	Weight
First order	Energy	0.0911	1
GLCM	Imc2	0.0844	0.9271
GLCM	Imc1	0.0756	0.8296
GLCM	Dependence Non-Uniformity Normalized	0.0718	0.7886
shape	Sphericity	0.0657	0.7211
First order	Kurtosis	0.0583	0.6406
shape	Least Axis Length	0.0578	0.6347
GLCM	Correlation	0.0577	0.6335
GLCM	Joint Entropy	0.0519	0.5704
GLSZM	Large Area High Gray Level Emphasis	0.0507	0.5571
GLCM	Maximum Probability	0.0495	0.5433
First order	10 Percentile	0.0472	0.5184
GLSZM	SZN	0.047	0.5161
NGTDM	Busyness	0.0442	0.4849
First order	Minimum	0.0429	0.471
NGTDM	Coarseness	0.0406	0.4462
shape	Major Axis Length	0.032	0.3517
shape	Maximum 3D Diameter	0.0315	0.3463

GLCM, Gray Level Co-occurrence Matrix; GLDM, L Gray Level Dependence Matrix; GLSZM, Gray Level Size Zone Matrix; NGTDM, Neighborhood Gray Tone Difference Matrix; DHGLE, Large Dependence High Gray Level Emphasis; SDLGLE, Small Dependence Low Gray Level Emphasis; LDLGLE, Large Dependence Low Gray Level Emphasis; Dependence Non-Uniformity Normalized; SZN, Size Zone Non-Uniformity Normalized.

also affects the choice of surgical procedure for T1-stage peripheral lung ADC. Su (22) found that in ADC patients with tumor size ≤ 2 cm and MPP < 0–5%, limited resection is suitable (3, 8), whereas lobectomy may be required for patients with MPP > 5% (9). Therefore, prediction MPP in lung ADC with size ≤ 2 cm can guide surgical resection strategy.

Many previous studies have showed that CT image-based radiomics analysis can be used to predict ADC with MPP with a prediction performance of 0.7–0.98 (11–18). However, there are inconsistencies and significant differences among the results of previous studies. MPP and solid pattern appear mostly solid on CT, while solid, mGGO, and pGGO types are present in other ADC subtypes. Radiomics identification of GGO and solid lesions is simple. We speculate that the ratio of GGO and solid nodules in ADC without MPP dataset may affected the performance of the prediction model. Many previous studies using radiomics analysis combined GGO and solid nodules in the non-MPP dataset, and the proportion of GGOs was unknown. Chen (11) achieved a prediction performance of 0.86, but GGO and solid nodules accounted for 78% and 22%, respectively, in the ADC without MPP group. Park (12) obtained a prediction performance of 0.98 but the subtypes in non-MPP dataset only consisted of lepidic subtype. The prediction performance decreased to 0.84 when the dataset consisted of papillary and acinar subtype. The dataset by Wang (13) consisted of entire pGGO type for predictor MPP in ADC and the performance was below 0.80. In the current study, the AUC for the prediction performance of model 1 (solid nodules and GGO) was 0.91, but AUC decreased to 0.78 in model 2 after excluding GGO type (69%). In the final external validation, the AUC of model 1 was significantly decreased, may be due to the different ratio of GGO in the training and validation sets. Importantly, the proportion of GGO was lower in the validation set than the training set. These results suggest that the GGO type in the non-MPP dataset contributed to bias in the study results and may explain the differences in results from previous studies.

Notably, there are significant differences in the characteristics of the two prediction models. First, in model 1, the most important characteristic was mean, which was not included in model 2. Mean is a first-order characteristic related to the CT density of the lesion. The CT value of solid nodules is significantly higher than of GGO. Second, first-order characteristic parameters accounted for around half of the parameters in model 1. This may be because GGO is rare in ADC with MPP, resulting in higher contribution to first-order parameters in the histogram. In contrast, there were fewer first-order characteristics in model 2, which means that the role of density characteristics in non-solid nodule prediction is not significant. The prediction of model 1 relied more on texture features and other high-order parameters; therefore, the prediction performance of model 2 was lower than that of model 1. This emphasizes that nodule type is an important factor affecting the prediction MPP in early stage ADC, as well as the necessity of stratifying nodule type when constructing models to predict ADC with MPP. However, the proportion of solid nodules was significantly higher than GGOs (pGGO and mGGO) in ADC with MPP of size ≤ 2 cm (19). Therefore, preoperative prediction of ADC with MPP for solid nodules with 2 cm or less has great clinical value for guiding surgical treatment.

This was a multicenter study and external validation was performed for both models. The results were significantly different between the models, and the generalizability of model 2 was validated in external validation. Although the AUC of model 1 was 0.91 in the training set, which was significantly lower in external validation. There were statistically significant differences in AUC between the training set and external

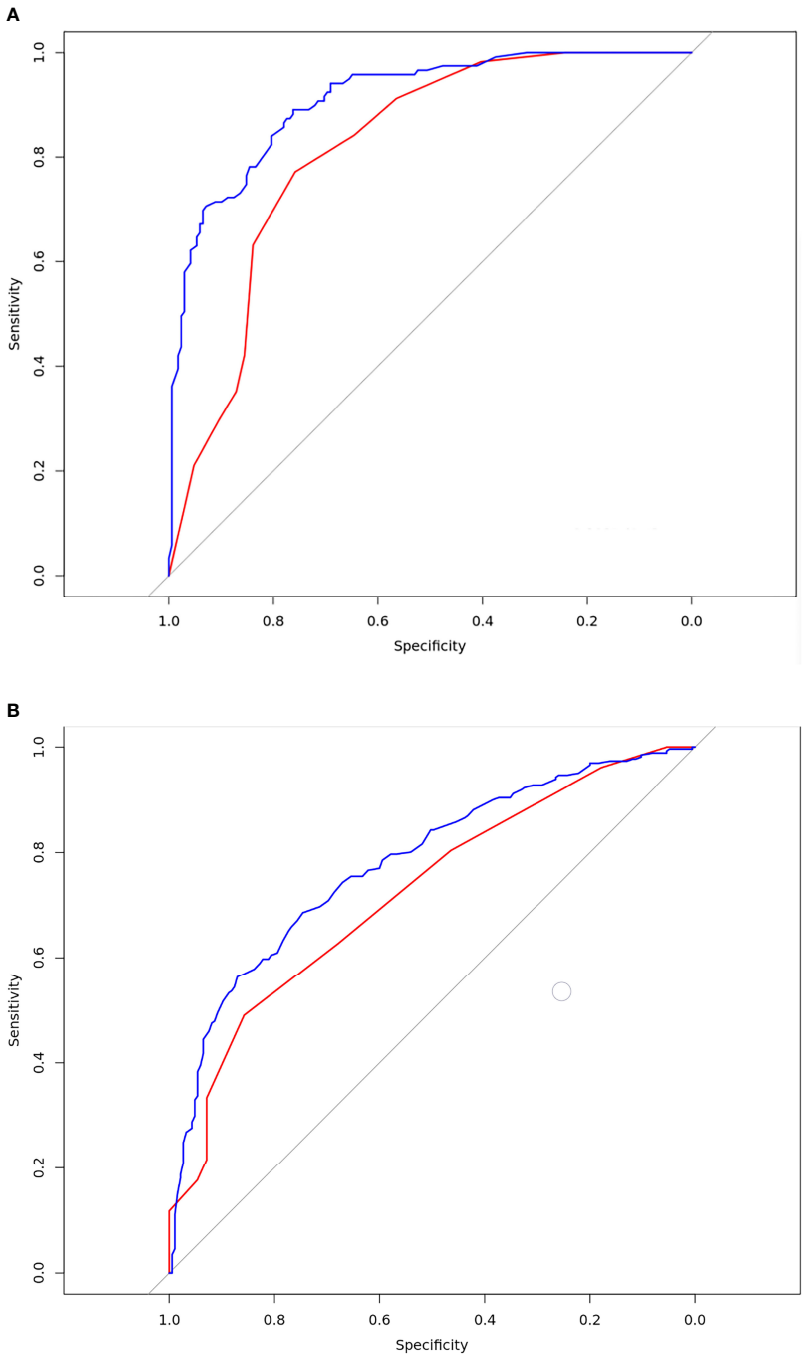


FIGURE 3 | Results of the receiver-operating characteristic curve analysis for the two models. **(A)** The ROC curves for the model 1 in the training and validation database. The blue line was training set. the AUC value was 0.91[95% confidence intervals (CI):0.88-0.94]; the red line was validation set. The AUC value was 0.82 [CI:0.74-0.89];Delong test p=0.0296; **(B)** The ROC curves for the model 2 in the training and validation database; The blue line was training set. The AUC value was 0.78[CI:0.74-0.82]; The red line was validation set. The AUC value was 0.72[CI:0.63-0.82];Delong test p=0.2865.

validation set, thereby raising concern over its generalizability for model 1. These results highlight the importance of external validation. In previous prediction studies of T1 stage ADC with MPP, only He et al. (18) performed external validation, while most others were single-center studies that lacked external validation. There are limitations to single-center studies because of over-fitting in the prediction model. Over-fitting was observed in model 1 after external validation, resulting in poor reliability

TABLE 5 | Predictive probabilities for the two radiomic model on the training and validation cohort.

	AUC_CI	Accuracy	Sensitivity	Specificity	PPV	NPV
Model 1: the lesion including solid, part-solid and GGO nodule						
Training	0.91[0.88-0.94]	0.79	0.73	0.86	0.79	0.82
Validation	0.82[0.74-0.89]	0.83	0.76	0.88	0.83	0.83
Model 2: the lesion including pure solid nodule						
Training	0.78[0.74-0.82]	0.73	0.77	0.6	0.73	0.65
Validation	0.72[0.63-0.82]	0.76	0.49	0.85	0.75	0.65

of study results and limited clinical applicability. Multicenter data must be used for constructing and testing radiomics models for better clinical application.

There were several limitations to this study. First, retrospective studies have inherent weaknesses and potential bias. In future prospective works, we will strengthen the research of ratio of GGO and solid type. Second, the study included a small number of patients with MPP in ADC. Although this study included three hospitals, our sample size was small. This may be because we included patients with peripheral lung ADC with size ≤ 2 cm, and the incidence of MPP in ADC increases with tumor size. A larger sample size is required in future studies. Third, the incidence of MPP in ADC in our study is different from previous studies. As pathologists have different expertise, potential subjective deviation may be present when semi-quantitative analysis is used to record the ratio of each pathological subtype, which may affect the study results.

CONCLUSION

In summary, the constructed two models based on nodule type stratification has potential to predict MPP in lung ADC of size ≤ 2 cm. We found that GGO nodule type in the without MPP dataset will affect the performance of the prediction model. Thus, the pure solid nodules (model 2) prediction had moderate stable generalizability. This model may contribute to an auxiliary method for preoperative prediction of MPP peripheral lung ADC of size ≤ 2 cm with proper treatment planning.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the First Affiliated Hospital of Zhejiang University School of Medicine; the First Affiliated Hospital of Wannan Medical College; Zhejiang Cancer Hospital. The ethics committee waived the requirement of written informed consent for participation.

AUTHOR CONTRIBUTIONS

ML and ZF: Study design, manuscript writing and data analysis. ML, YR, and FS: labeled the image. LZ and MW: CT raw data collection. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol* (2015) 10:1243–60. doi: 10.1097/jto.0000000000000630
2. Zhao Y, Wang R, Shen X, Pan Y, Cheng C, Li Y, et al. Minor Components of Micropapillary and Solid Subtypes in Lung Adenocarcinoma Are Predictors of Lymph Node Metastasis and Poor Prognosis. *Ann Surg Oncol* (2016) 23:2099–105. doi: 10.1245/s10434-015-5043-9
3. Hung JJ, Yeh YC, Wu YC, Chou TY, Hsu WH. Prognostic Factors in Completely Resected Node-Negative Lung Adenocarcinoma of 3 Cm or Smaller. *J Thorac Oncol* (2017) 12:1824–33. doi: 10.1016/j.jtho.2017.07.009
4. Qian F, Yang W, Wang R, Xu J, Wang S, Zhang Y, et al. Prognostic Significance and Adjuvant Chemotherapy Survival Benefits of a Solid or Micropapillary Pattern in Patients With Resected Stage IB Lung Adenocarcinoma. *J Thorac Cardiovasc Surg* (2018) 155:1227–35.e2. doi: 10.1016/j.jtcvs.2017.09.143
5. Pani E, Kennedy G, Zheng X, Ukert B, Jarrar D, Gaughan C, et al. Factors Associated With Nodal Metastasis in 2-Centimeter or Less Non-Small Cell Lung Cancer. *J Thorac Cardiovasc Surg* (2020) 159:1088–96.e1. doi: 10.1016/j.jtcvs.2019.07.089
6. Xu F, Zhu W, Shen Y, Wang J, Xu R, Qutesh C, et al. Radiomic-Based Quantitative CT Analysis of Pure Ground-Glass Nodules to Predict the Invasiveness of Lung Adenocarcinoma. *Front Oncol* (2020) 10:872. doi: 10.3389/fonc.2020.00872
7. Lee G, Lee HY, Jeong JY, Han J, Cha MJ, Lee KS, et al. Clinical Impact of Minimal Micropapillary Pattern in Invasive Lung Adenocarcinoma: Prognostic Significance and Survival Outcomes. *Am J Surg Pathol* (2015) 39:660–6. doi: 10.1097/pas.0000000000000399
8. Qiu C, Wang G, Xu J, Cui L, Dong W, Ni Y, et al. Sublobectomy Versus Lobectomy for Stage I Non-Small Cell Lung Cancer in the Elderly. *Int J Surg* (2017) 37:1–7. doi: 10.1016/j.ijsu.2016.11.090
9. Nitadori J, Bograd AJ, Kadota K, Sima CS, Rizk NP, Morales EA, et al. Impact of Micropapillary Histologic Subtype in Selecting Limited Resection vs Lobectomy for Lung Adenocarcinoma of 2cm or Smaller. *J Natl Cancer Inst* (2013) 105:1212–20. doi: 10.1093/jnci/djt166

10. Song SH, Park H, Lee G, Lee HY, Sohn I, Kim HS, et al. Imaging Phenotyping Using Radiomics to Predict Micropapillary Pattern Within Lung Adenocarcinoma. *J Thorac Oncol* (2017) 12:624–32. doi: 10.1016/j.jtho.2016.11.2230
11. Chen LW, Yang SM, Wang HJ, Chen YC, Lin MW, Hsieh MS, et al. Prediction of Micropapillary and Solid Pattern in Lung Adenocarcinoma Using Radiomic Values Extracted From Near-Pure Histopathological Subtypes. *Eur Radiol* (2021) 31:5127–38. doi: 10.1007/s00330-020-07570-6
12. Park S, Lee SM, Noh HN, Hwang HJ, Kim S, Do KH, et al. Differentiation of Predominant Subtypes of Lung Adenocarcinoma Using a Quantitative Radiomics Approach on CT. *Eur Radiol* (2020) 30:4883–92. doi: 10.1007/s00330-020-06805-w
13. Wang X, Zhang L, Yang X, Tang L, Zhao J, Chen G, et al. Deep Learning Combined With Radiomics may Optimize the Prediction in Differentiating High-Grade Lung Adenocarcinomas in Ground Glass Opacity Lesions on CT Scans. *Eur J Radiol* (2020) 129:109150. doi: 10.1016/j.ejrad.2020.109150
14. Yang SM, Chen LW, Wang HJ, Chen LR, Lor KL, Chen YC, et al. Extraction of Radiomic Values From Lung Adenocarcinoma With Near-Pure Subtypes in the International Association for the Study of Lung Cancer/the American Thoracic Society/the European Respiratory Society (IASLC/ATS/ERS) Classification. *Lung Cancer* (2018) 119:56–63. doi: 10.1016/j.lungcan.2018.03.004
15. Kim H, Lee D, Cho WS, Lee JC, Goo JM, Kim HC, et al. CT-Based Deep Learning Model to Differentiate Invasive Pulmonary Adenocarcinomas Appearing as Subsolid Nodules Among Surgical Candidates: Comparison of the Diagnostic Performance With a Size-Based Logistic Model and Radiologists. *Eur Radiol* (2020) 30:3295–305. doi: 10.1007/s00330-019-06628-4
16. Wu G, Woodruff HC, Sanduleanu S, Refaee T, Jochems A, Leijenaar R, et al. Preoperative CT-Based Radiomics Combined With Intraoperative Frozen Section Is Predictive of Invasive Adenocarcinoma in Pulmonary Nodules: A Multicenter Study. *Eur Radiol* (2020) 30:2680–91. doi: 10.1007/s00330-019-06597-8
17. Ding H, Xia W, Zhang L, Mao Q, Cao B, Zhao Y, et al. CT-Based Deep Learning Model for Invasiveness Classification and Micropapillary Pattern Prediction Within Lung Adenocarcinoma. *Front Oncol* (2020) 10:1186. doi: 10.3389/fonc.2020.01186
18. He B, Song Y, Wang L, Wang T, She Y, Hou L, et al. A Machine Learning-Based Prediction of the Micropapillary/Solid Growth Pattern in Invasive Lung Adenocarcinoma With Radiomics. *Transl Lung Cancer Res* (2021) 10:955–64. doi: 10.21037/tlcr-21-44
19. Yoshida Y, Nitadori JI, Shinozaki-Ushiku A, Sato J, Miyaji T, Yamaguchi T, et al. Micropapillary Histological Subtype in Lung Adenocarcinoma of 2 Cm or Less: Impact on Recurrence and Clinical Predictors. *Gen Thorac Cardiovasc Surg* (2017) 65:273–9. doi: 10.1007/s11748-017-0747-3
20. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner Society: Glossary of Terms for Thoracic Imaging. *Radiology* (2008) 246:697–722. doi: 10.1148/radiol.2462070712
21. Katsumata S, Aokage K, Nakasone S, Sakai T, Okada S, Miyoshi T, et al. Radiologic Criteria in Predicting Pathologic Less Invasive Lung Cancer According to TNM 8th Edition. *Clin Lung Cancer* (2019) 20:e163–70. doi: 10.1016/j.cllc.2018.11.001
22. Su H, Xie H, Dai C, Zhao S, Xie D, She Y, et al. Procedure-Specific Prognostic Impact of Micropapillary Subtype may Guide Resection Strategy in Small-Sized Lung Adenocarcinomas: A Multicenter Study. *Ther Adv Med Oncol* (2020) 12:1758835920937893. doi: 10.1177/1758835920937893

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

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 03 January 2022
doi: 10.3389/fonc.2021.801213

Radiomics Study for Discriminating Second Primary Lung Cancers From Pulmonary Metastases in Pulmonary Solid Lesions

Feiyang Zhong[1†], *Zhenxing Liu*[2†], *Wenting An*[1], *Binchen Wang*[1], *Hanfei Zhang*[1], *Yumin Liu*[2] *and Meiyan Liao*[1*]

[1] *Department of Radiology, Zhongnan Hospital of Wuhan University, Wuhan, China,* [2] *Department of Neurology, Zhongnan Hospital of Wuhan University, Wuhan, China*

Background: The objective of this study was to assess the value of quantitative radiomics features in discriminating second primary lung cancers (SPLCs) from pulmonary metastases (PMs).

Methods: This retrospective study enrolled 252 malignant pulmonary nodules with histopathologically confirmed SPLCs or PMs and randomly assigned them to a training or validation cohort. Clinical data were collected from the electronic medical records system. The imaging and radiomics features of each nodule were extracted from CT images.

Results: A rad-score was generated from the training cohort using the least absolute shrinkage and selection operator regression. A clinical and radiographic model was constructed using the clinical and imaging features selected by univariate and multivariate regression. A nomogram composed of clinical-radiographic factors and a rad-score were developed to validate the discriminative ability. The rad-scores differed significantly between the SPLC and PM groups. Sixteen radiomics features and four clinical-radiographic features were selected to build the final model to differentiate between SPLCs and PMs. The comprehensive clinical radiographic–radiomics model demonstrated good discriminative capacity with an area under the curve of the receiver operating characteristic curve of 0.9421 and 0.9041 in the respective training and validation cohorts. The decision curve analysis demonstrated that the comprehensive model showed a higher clinical value than the model without the rad-score.

Conclusion: The proposed model based on clinical data, imaging features, and radiomics features could accurately discriminate SPLCs from PMs. The model thus has the potential to support clinicians in improving decision-making in a noninvasive manner.

Keywords: second primary lung cancers, pulmonary metastases, clinical-radiographic factor, radiomics, lung cancer

OPEN ACCESS

Edited by:
Yiyan Liu,
University of Louisville, United States

Reviewed by:
Chen Gao,
Zhejiang Provincial Hospital of Traditional Chinese Medicine, China
Weiling Ma,
Shanghai Jiaotong University, China
Guojin Zhang,
Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, China

***Correspondence:**
Meiyan Liao
liaomy@whu.edu.cn

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *25 October 2021*
Accepted: *06 December 2021*
Published: *03 January 2022*

Citation:
Zhong F, Liu Z, An W, Wang B, Zhang H, Liu Y and Liao M (2022) Radiomics Study for Discriminating Second Primary Lung Cancers From Pulmonary Metastases in Pulmonary Solid Lesions. Front. Oncol. 11:801213. doi: 10.3389/fonc.2021.801213

INTRODUCTION

Over the last few decades, owing to advancements in cancer screening and treatment, the life expectancy of cancer survivors continues to improve. It was estimated that approximately 16.9 million Americans were living with cancer as of January 1, 2019, and this number is expected to increase to 20 million by January 1, 2030 (1). Cancer survivors have a higher risk of developing new primary malignant tumors than the general population. The most common newly developed primary malignant tumor is lung cancer (2). Lung cancer remains the leading cause of cancer-related death worldwide (3). Meanwhile, the lungs are the sites most frequently affected by metastasis. Approximately 30% of cancer survivors develop lung metastases (4).

The distinction of second primary lung cancers (SPLCs) from pulmonary metastases (PMs) is of great clinical interest because of the vastly different survival outcomes between them. On account of the close clinical monitoring and regular follow-up of cancer survivors, SPLCs are often diagnosed at an early stage. Compared with primary lung cancer, SPLCs have a fair prognosis after surgical resection of the lesion (5). Both radiotherapy and chemotherapy are regarded as effective methods of treatment (6, 7). Metastasis is the leading cause of mortality among tumor patients (8). The occurrence of metastasis is considered to represent the terminal, incurable stage of a tumor. Early differential diagnosis between these two disorders may help clinicians decide whether aggressive treatment or palliative care is appropriate.

Pathologic assessment remains the gold standard for distinguishing between SPLCs and PMs. Histologically distinct primary tumors are presumed to have diverse origins in a single patient. When tumors are categorized as the same histologic type, immunohistochemistry and genetic testing can assist in confirming the diagnosis (9). However, pathological results cannot be obtained preoperatively. Histopathology specimen acquisition relies on invasive lung biopsy, which may cause several complications, such as pneumothorax, pneumorrhachis, or air embolism. Not all patients are suitable for a needle biopsy or surgical resection. In addition, pathological specimens are typically acquired from one or more separate focal areas and cannot completely characterize the whole tumor.

The value of clinical and imaging characteristics in differential diagnosis between SPLCs and PMs has been reported in our previous study (10) and other articles (11–14). However, there is a lack of radiomics studies concerning the distinctions. Radiomics is an emerging science that extracts a large number of imaging features from radiographic images. It converts images into quantitative parameters and subsequently performs statistical analysis to support decision-making. Previous studies have shown that radiomics can play an important role in diagnosing malignancy, assessing treatment efficacy, and predicting clinical outcomes (15–17). In particular, radiomics has been used to discriminate different pathological types of lung cancer (18, 19). The present study thus aimed to assess whether radiomics features can discriminate SPLCs from PMs and to develop a comprehensive model based on clinical imaging and radiomics to guide clinical decisions.

MATERIALS AND METHODS

Patients

This retrospective, single-center study was approved by the Institutional Review Board of Zhongnan Hospital of Wuhan University and was conducted in accordance with the Declaration of Helsinki. The requirement for informed consent was waived owing to the retrospective nature of this study. The inclusion criteria were as follows: (1) pathological confirmation of malignant pulmonary lesions based on the histopathological evaluation of surgical resection and percutaneous biopsy; (2) thin-section chest CT (section thickness ≤1.5 mm) examination performed within one week before needle biopsy or surgery, and (3) history of malignant tumors. The exclusion criteria were as follows: (1) insufficient image quality for analysis (20), (2) any anti-tumor treatment received before the CT scan, (3) ground-glass opacity (GGO) lesions, (4) uncertainty of whether the lesion was primary or metastatic, and (5) a previous history of multiple primary tumors in separate organs. All pathologically confirmed lesions for included patients were examined unless they had three or more lesions. In this case, the two largest focuses of tumor were selected. Based on the above criteria, 252 lesions (97 SPLCs and 155 PMs) from 245 patients of the given institution from January 2017 to June 2020 were included. Patients included in this study partially were described previously (10). The lesions were randomly assigned to a training cohort (n = 137) or validation cohort (n = 115).

The clinicopathological data, including pathologic assessment, sex, age, history of smoking, family history of malignancy, the recurrence status of the initial tumor, and serum tumor markers [neuron-specific enolase (NSE), carcinoembryonic antigen (CEA), and carbohydrate antigen 125 (CA125)] were obtained by reviewing the electronic medical record system. The upper limit of each tumor marker was the following: NSE, 15.2 ng/mL; CEA, 5 ng/mL; and CA125, 35 U/mL. The above tumor markers were considered positive if their values were higher than the upper limit. Two authors (ZFY and LZX) independently extracted the data.

CT Scanning

The chest CT images were obtained from the following CT systems: SOMATOM definition scanner (Siemens Healthineers, Forchheim, Germany), and GE discovery 750HD scanner (GE Medical Systems, Milwaukee, WI, USA). The scanning parameters of the above devices were as follows: tube voltage, 120 kV; automatic tube current adjustment technology, 100–350 mAs; matrix size, 512×512; slice interval, 0 mm; standard soft-tissue algorithm reconstruction; reconstructed section thickness, 1 mm.

Evaluation of Subjective Radiographic Characteristics

The subjective radiographic characteristics were independently analyzed by two thoracic radiologists (FZ, with three years of chest radiological experience, and HZ, with seven years of chest radiological experience), who were blinded to the final pathological results. The CT images were reviewed in the lung window setting (width, 1500 HU; level, −700 HU) and mediastinal window setting

(width, 300 HU; level, 40 HU) windows. Discrepancies in the evaluations were resolved through consultation. The imaging characteristics of each pulmonary malignant lesion included (1) lesion size (maximum diameter), (2) distribution of the lesions (single or multiple), (3) central or peripheral type, (4) density (homogeneous or heterogeneous), (5) air bronchogram (absent, present), (6) bubble lucency (absent, present), (7) calcification (absent, present), (8) vessel convergence sign (absent, present), (9) margin (clear, unclear), (10) contour (round, irregular), (11) lobulation (absent, present), (12) spiculation (absent, present), (13) pleural effusion (absent, present), and (14) enlarged mediastinal lymph node (absent, present).

Region of Interest (ROI) Segmentation and Radiomics Feature Extraction

The pulmonary lesions were semi-automatically segmented using ITK-SNAP (version 3.8.0, http://www.itk-snap.org). The original Digital Imaging and Communications in Medicine files were imported into the in-house software (Analysis Kit, version 3.1.5.R, GE Healthcare) for pre-processing, and the lesions were segmented in standard images, slice by slice, under the lung window setting (width, 1500 HU; level, −700 HU). The lesions were delineated to avoid large vessels, bronchi, and chest walls, if possible.

Radiomics feature extraction was applied to the chest CT images using AK software. Finally, from one segmented ROI, a total of 402 imaging texture features were extracted: 42 histogram features, 144 Gy-level co-occurrence matrix features, 11 Gy-level size zone matrix features, 180 Gy-level run-length matrix features, and 25 shape- and size-based features. Details of the extraction features are provided in the **Supplementary Material**. Each image was normalized to eliminate the impact of different quantization levels on the texture features.

Feature Selection and Model Building

Dice Similarity Coefficient was used to describe inter reader segmentation variability, which ranged from 0.74 to 0.98 (**Table S2**) (21). The intraclass correlation coefficient (ICC) was used to evaluate inter-observer agreement of quantitative radiomics parameters. In 20 randomly sampled cases, two chest radiologists (FZ and HZ) independently drew the ROI and extracted the radiomics features. Radiomics features with an ICC higher than 0.75 were regarded as consistent (21) and were included for further analysis. A radiologist (FZ) sketched ROIs in the remaining cases.

In the training set, a minimum redundancy–maximum relevance (mRMR) algorithm was employed to rank the importance of the selected features. Finally, the 100 highest mRMR-ranked features were input to the least absolute shrinkage and selection operator (LASSO) classifier to select the most predictive features. The rad-score was calculated for each lesion based on the final selected features (22), and a receiver operating characteristic (ROC) curve was constructed to evaluate the discriminatory ability of the rad-scores *via* the area under the curve (AUC) in the training and validation cohorts. The clinical data and subjective radiographic characteristics were evaluated using univariate analysis. Significant factors were included in the multivariate analysis to build a clinical–radiographic model.

The clinical–radiographic and rad-scores were combined to construct an individualized discriminatory nomogram based on a multivariate logistic regression algorithm. Internal validation was performed using a calibration curve, which was verified by the Hosmer–Lemeshow test. The AUCs of the ROC were calculated to evaluate the above models in the respective training and validation sets. Decision curve analysis (DCA) was performed to compare the clinical value of the models (23).

Statistical Analysis

Categorical variables were expressed as frequency rates and compared using the χ^2 test or Fisher's exact test. Continuous variables were described as the median (interquartile range [IQR]) and compared using the t-test or Wilcoxon rank-sum test. A two-sided α value of less than 0.05 was considered statistically significant. Statistical analysis was performed using R software (version 4.1.2, http://www.R-project.org). The "mRMRe" package was used to conduct the mRMR algorithm, "irr" package for Intra-class correlation coefficient (ICC) algorithm, "tableone" for comparison of clinical baseline data between groups, "tidyverse" for data collation and exploration, "glmnet" for LASSO regression, "rms" for Nomogram, "rmda" for decision curve analysis(DCA), "pROC" for receiver operating characteristic curve analysis(ROC), "ResourceSelection" for goodness of fit test, and "ggpubr" for data result visualization, respectively.

RESULTS

Patient Information

A total of 245 patients (134 men [54.7%] and 111 women [45.3%]) who met the inclusion criteria were included in this retrospective study. The median patient age was 62 years (IQR, 55–67 years; range, 31–93 years). The baseline characteristics of the patients are presented in **Table 1**. Of the 245 patients, 252 solid pulmonary lesions were pathologically diagnosed as malignant foci, including 97 primary lesions and 155 metastatic lesions. There are 21 synchronous SPLCs and 76 metachronous ones. 55SPLCs and 82 PMs were included in the training set, while 42 SPLCs and 73 PMs were included in the validation set. Univariate analysis revealed no difference between the clinical data and subjective radiographic characteristics in the training and validation sets.

Comparison of Clinical and Radiographic Features Between SPLC and PM Groups

The univariate logistic regression analysis of the clinical data showed a statistically significant difference in sex, history of smoking, and CEA level between the SPLC and PM groups in the training set ($p < 0.05$). Other clinical data, such as age, recurrence status of the initial tumor, family history of malignancy, NSE level, and CA125 level, were not statistically significant.

Among the visual imaging findings, the maximum diameter of lesions, the distribution of lesions, central or peripheral type of lesions, air bronchogram, calcification, vessel convergence sign,

TABLE 1 | The clinical and radiographic factors of patients in SPLC[†] and PM[§] Groups.

Variables	SPLC Group (N=97) No. of patient (%)	PM Group (N=155) No. of patient (%)	P value
Sex			0.029
Male	62 (63.9)	76 (49.0)	
Female	35 (36.1)	79 (51.0)	
Age (years)			<0.001
(Median [IQR])	64.00 [59.00, 68.00]	59.00 [52.00, 65.00]	
History of smoking			0.003
Yes	54 (55.7)	116 (74.8)	
No	43 (44.3)	39 (25.2)	
Family history of malignancy (%)			0.43
Yes	88 (90.7)	146 (94.2)	
No	9 (9.3)	9 (5.8)	
Recurrence status of the initial tumor			0.085
Yes	97 (100.0)	149 (96.1)	
No	0 (0.0)	6 (3.9)	
Maximal lesion size (mm)			<0.001
(Median [IQR])	30.00 [20.00, 49.00]	19.00 [13.00, 28.50]	
NSE			0.003
Normal	74 (76.3)	141 (91.0)	
Abnormal	23 (23.7)	14 (9.0)	
CEA			0.002
Normal	59 (60.8)	123 (79.4)	
Abnormal	38 (39.2)	32 (20.6)	
CA125			0.034
Normal	69 (71.1)	129 (83.2)	
Abnormal	28 (28.9)	26 (16.8)	
The distribution of lesions			<0.001
Single	80 (82.5)	72 (46.5)	
Multiple	17 (17.5)	83 (53.5)	
Central or peripheral type			<0.001
Peripheral	75 (77.3)	149 (96.1)	
Central	22 (22.7)	6 (3.9)	
Density			0.546
Homogeneous	79 (81.4)	120 (77.4)	
Heterogeneous	18 (18.6)	35 (22.6)	
Air bronchogram			<0.001
Absent	47 (48.5)	134 (86.5)	
Present	50 (51.5)	21 (13.5)	
Bubble lucency			0.283
Absent	81 (83.5)	138 (89.0)	
Present	16 (16.5)	17 (11.0)	
Calcification			0.001
Absent	79 (81.4)	147 (94.8)	
Present	18 (18.6)	8 (5.2)	
Vessel convergence sign			<0.001
Absent	62 (63.9)	137 (88.4)	
Present	35 (36.1)	18 (11.6)	
Margin			0.003
Clear	79 (81.4)	146 (94.2)	
Unclear	18 (18.6)	9 (5.8)	
Contour			<0.001
Round	22 (22.7)	118 (76.1)	
Irregular	75 (77.3)	37 (23.9)	
Lobulation			<0.001
Absent	6 (6.2)	40 (25.8)	
Present	91 (93.8)	115 (74.2)	
Spiculation			<0.001
Absent	30 (30.9)	130 (83.9)	
Present	67 (69.1)	25 (16.1)	
Pleural effusion			<0.001
Absent	25 (25.8)	83 (53.5)	
Present	72 (74.2)	72 (46.5)	
Enlarged mediastinal lymph node			0.158
Absent	70 (72.2)	125 (80.6)	
Present	27 (27.8)	30 (19.4)	

[†]SPLC indicates second primary lung cancer; [§]PM, pulmonary metastasis.

contour, lobulation, spiculation, and pleural effusion presented significant differences in the univariate analysis ($p < 0.05$). Other radiographic features, such as density, bubble lucency, margin, and enlarged mediastinal lymph nodes, were not statistically significant in the two groups. All clinical–radiographic variables that achieved statistical significance were enrolled in the multivariate logistic regression analysis.

In the multivariate analysis, the distribution of lesions (odds ratio [OR], 6.52; 95% confidence interval [CI], 1.92–26.84; p-value = 0.005), central or peripheral type (OR, 0.05; 95% CI, 0–0.66; p-value = 0.031), contour (OR, 0.23; 95% CI, 0.06–0.76; p-value = 0.018), and spiculation (OR, 0.12; 95% CI, 0.03–0.42; p-value <0.001) were identified as independent variables and included in the clinical–radiographic model.

Feature Selection and Radiomics Model Building

A total of 233 radiomics features with an ICC higher than 0.75 were enrolled for the next feature extraction step. Then, the mRMR program was used to select the 100 highest-ranked features in the training set. Finally, LASSO logistic regression was used to reduce the 100 features to 16 features with nonzero coefficients, as shown in **Figure 1**. The rad-score of each lesion was calculated using the following formula:

$$\begin{aligned}\text{Rad}-\text{Score} &= 0.585917936 + 0.300121336 \times \text{InverseDifferenceMoment\_AllDirection\_offset1\_SD} \\ &+ 0.139511141 \times \text{HaralickCorrelation\_angle90\_offset7} + 0.117363006 \\ &\times \text{LongRunEmphasis\_angle0\_offset7} - 0.286189566 \times \text{LongRunEmphasis\_angle0\_offset4} \\ &+ 0.807308895 \times \text{Compactness2} - 0.095952964 \times \text{HaralickCorrelation\_angle0\_offset7} - 0.09620897 \\ &\times \text{SphericalDisproportion} + 0.410874754 \\ &\times \text{ShortRunHighGreyLevelEmphasis\_AllDirection\_offset7\_SD} - 0.109558249 \times \text{SmallAreaEmphasis} \\ &- 0.05680965 \times \text{ZonePercentage} + 0.102430665 \times \text{GLCMEntropy\_angle135\_offset4} + 0.148896463 \\ &\times \text{LongRunHighGreyLevelEmphasis\_AllDirection\_offset1\_SD} + 0.057282871 \times \text{Elongation} \\ &+ 0.202108196 \times \text{LongRunLowGreyLevelEmphasis\_angle0\_offset7} - 0.244924038 \\ &\times \text{GLCMEnergy\_angle90\_offset7} + 0.310948511 \times \text{LongRunEmphasis\_AllDirection\_offset7\_SD}\end{aligned}$$

The rad-scores were significantly different between the SPLC and the PM groups in both the training and validation sets ($p < 0.05$); PMs had higher rad-scores than SPLCs. The rad-scores for both the training and validation sets are shown in **Figure 2**.

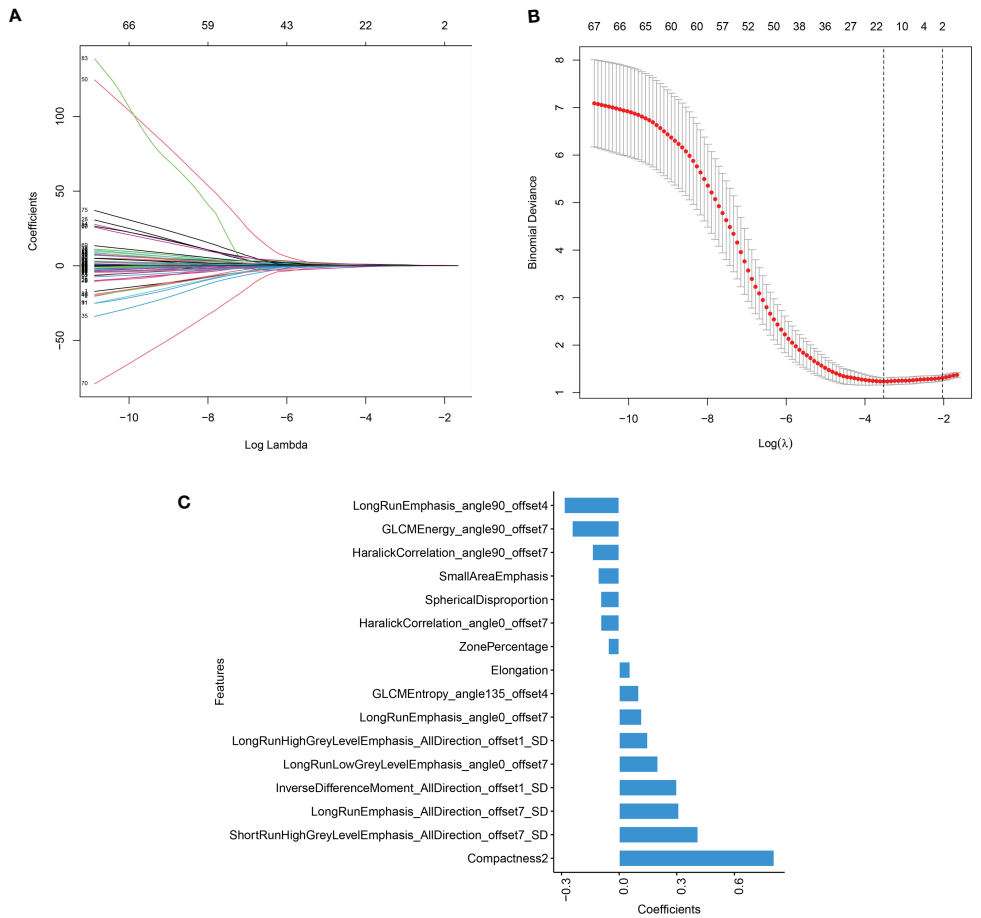


FIGURE 1 | The result of LASSO model **(A)** LASSO coefficient profiles of the candidate predictors. **(B)** The features with nonzero coefficients are shown in the model. **(C)** The y-axis indicates the selected radiomics features, and the x-axis represents the coefficient of radiomics.

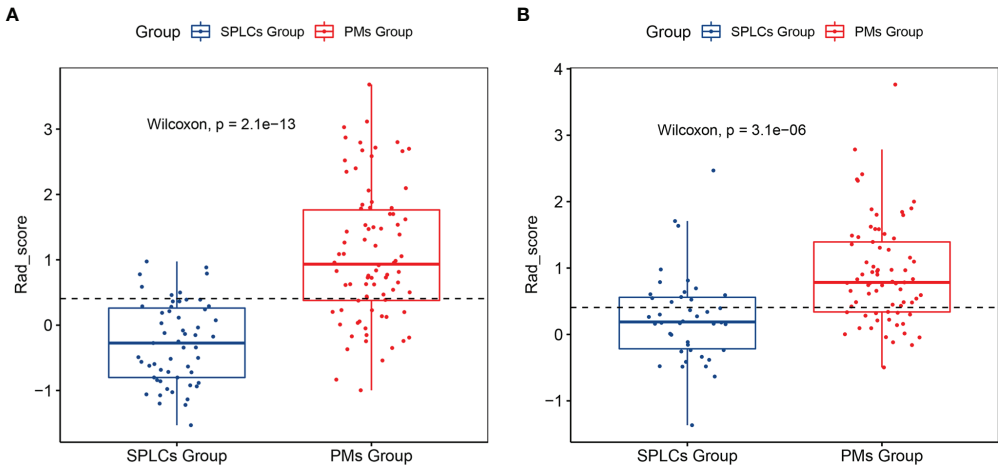


FIGURE 2 | The Rad-score of each lesion in the training set **(A)** and validation set **(B)**.

Nomogram Model Construction and Validation

Based on the training set, the radiomics scores, distribution of lesions, central or peripheral type, contour, and spiculation were incorporated into the comprehensive nomogram model construction (**Figure 3A**). **Figure 3** shows the calibration curves of the nomogram model in training cohort (B) and validation cohort (C). The Hosmer–Lemeshow test finding was not significant (P = 0.4612); it showed good calibration in the training set. **Figure 4** shows the discriminative abilities of each model. The radiomics model had good discriminative performance, with AUCs of 0.8707 (95% CI, 0.8138–0.9277) in the training set and 0.7622 (95% CI, 0.6702–0.8543) in the validation set. The clinical–radiographic model had AUCs of 0.8989 (95% CI, 0.8475–0.9503) and 0.9035 (95% CI, 0.8489–0.9581) in the training and validation cohorts, respectively. The comprehensive model achieved a slightly higher AUC in the training (0.9421; 95% CI, 0.9056–0.9786) and validation sets (0.9041; 95% CI, 0.8417–0.9665). **Figure 5** presents the DCA of the nomogram. The DCA showed that in most circumstances, using the comprehensive model to distinguish between SPLCs and PMs would be more clinically beneficial than using other models in training cohort (A) and validation cohort (B).

DISCUSSION

In this study, the ability of CT-based radiomics to discriminate SPLCs from PMs was investigated. An individual nomogram model integrated with clinical data, radiographic characteristics, and radiomics features was constructed. It achieved an excellent discriminative capability and has the potential to support clinicians in improving decision-making in a noninvasive manner.

In recent years, radiomics studies have attracted increasing attention because they can reflect quantitative intratumoral pathophysiological information in a noninvasive manner (24). Radiomics features represented tumor heterogeneity and were extracted from the entire ROI; they were not just limited to the biopsy site (17). Previous studies demonstrated that radiomics plays a role in differentiating between primary and metastatic tumors (25–29). In particular, CT radiomics features combined with positron emission tomography (PET) features can accurately distinguish between primary and metastatic lung cancers (26, 27). However, neither of these previous studies focused on the patient's history of neoplastic disease, leading to insufficient clinical value. To the best of the present authors' knowledge, this is the first study to apply radiomics studies to SPLCs and to build a radiomics model for distinguishing SPLCs from PMs. In this study, the rad-scores in the SPLC group were significantly lower than those in the PM group (−0.005 [IQR −0.516–0.365] *vs*. 0.830 [IQR 0.357–1.53]), thereby showing strong diagnostic efficacy.

Several studies have demonstrated that combining radiographic and radiomics provides a higher prognostic performance than radiomics alone in lung lesions (20, 30, 31). Despite some observer bias, the subjective radiographic characteristic is the most frequently used method for describing pulmonary lesions. Pulmonary GGO lesions have shown that tumor cells grow along the alveolar wall and are known to be a key sign of primary lung adenocarcinoma (11, 14, 32). This was also observed at the present research institution (10); therefore, only solid lesions were employed in this study. A prior study argue that central-type pulmonary lesion strongly prompts to be the SPLC because endobronchial metastasis is a rare event (13), and the same result was obtained in this study. In the final model, four independent imaging characteristics, including the distribution of lesions, central or peripheral type, contour, and spiculation, were in accordance with the authors' clinical experience. The final model did not include clinical variables because they were not statistically significant in the multiple logistic regression analysis. In the past, the event-

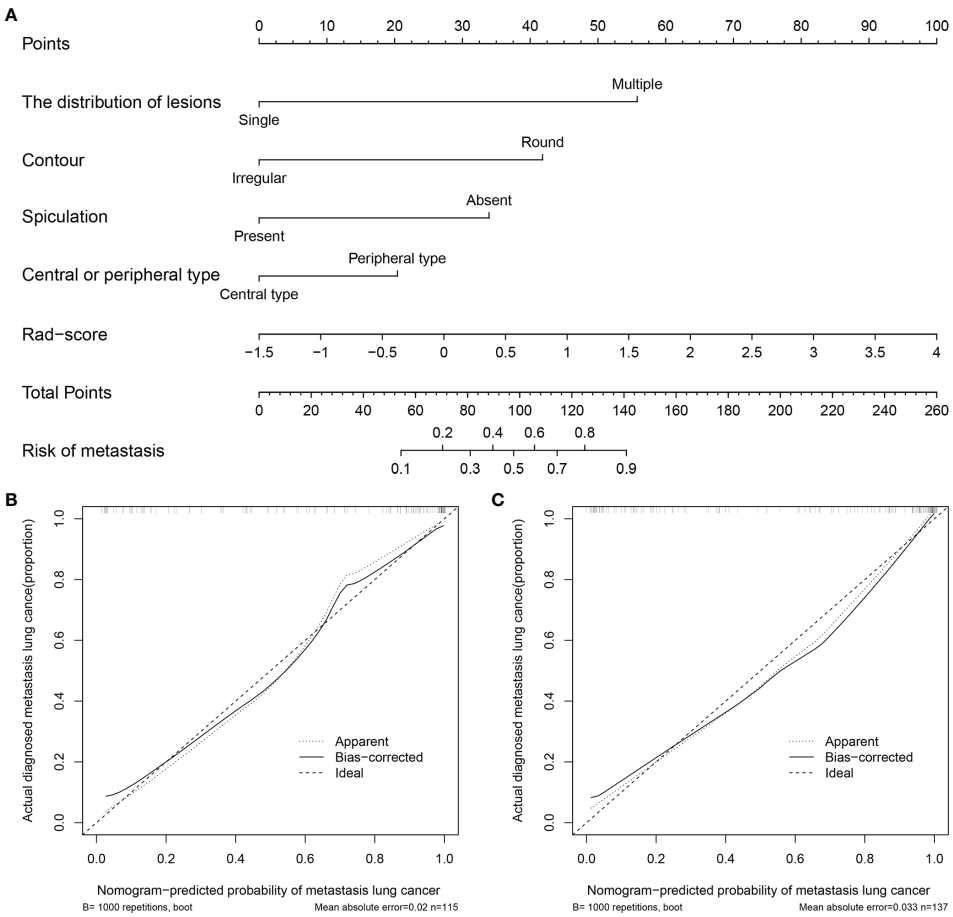


FIGURE 3 | **(A)** Nomogram for predicting SPLCs and PMs. For each patient, draw a vertical line between the variable value and the corresponding point line, and then assign a score for each variable based on the clinical and imaging characteristics to obtain a total score. The risk of metastasis can be predicted according to the total score. **(B)** Calibration curve for the nomogram in training cohort. **(C)** Calibration curve for the nomogram in validation cohort.

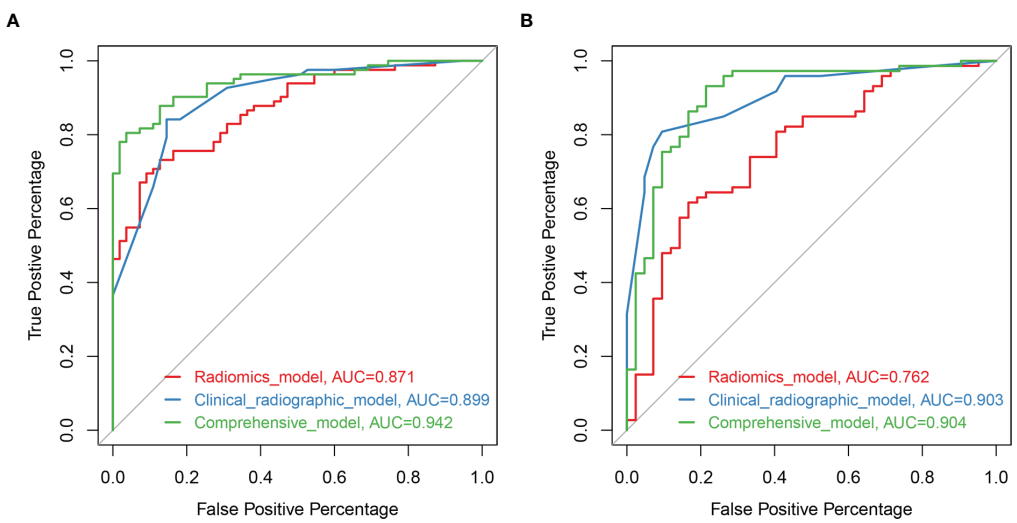


FIGURE 4 | Receiver operating characteristic (ROC) curves of the models based on clinical-radiographic factors (blue), radiomics features alone (red), and comprehensive clinical-radiography-radiomics features (green) in the training set **(A)** and validation set **(B)**.

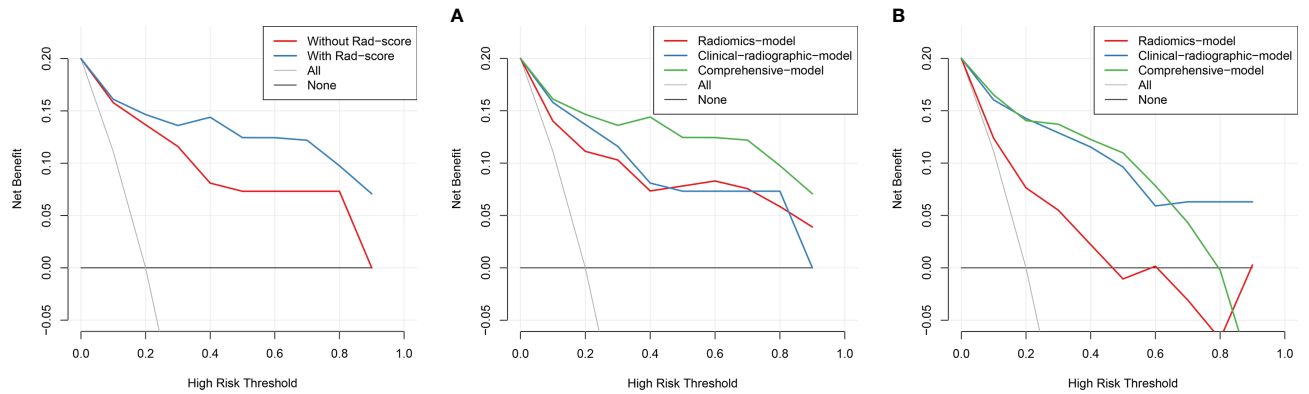


FIGURE 5 | The DCA showed that in most circumstances, using the comprehensive model to distinguish between SPLCs and PMs would be more clinically beneficial than other models in training cohort **(A)** and validation cohort **(B)**.

free duration is considered to be an important differentiating factor (13). However, it is difficult to accurately measure the time between primary and secondary tumor. So, this characteristic was not included in this study. Smoking history used to be a risk factor for SPLC (14, 33) but it has not been seen in this research. This may be because traditional Chinese women smoke less. Thus, combining the radiomics model with the radiographic features improved the prognostic performance.

However, there were some limitations to this study. First, it was a retrospective, single-center study with a small sample size. As the number of cases was small, different organs of initial primary cancers were not distinguished, which may have led to bias. Therefore, larger sample sizes from multiple centers are required for further studies. Second, radiomics feature extraction was performed only on plain CT scanning images. Enhanced CT or PET images may contain added valuable information. Third, in this study, a semi-automatic method was adopted to segment ROIs, which could lead to artificial differences. An accurate automatic segmentation method should be considered in future studies (34). Last, the relationship between radiomics signatures and subjective radiographic characteristics was not assessed. This aspect will be explored in future work.

In conclusion, the model developed using clinical–radiographic factors and CT-based radiomics features shows good performance discriminating between SPLCs and PMs. Therefore, for pulmonary malignancy patients with a history of other malignant tumors, the individual nomogram model may guide therapeutic decisions. With the development of artificial intelligence and machine learning, quantitative radiomics may have promising clinical applications.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

FZ and ZL designed the study, analyzed the data, and wrote the manuscript. WA, BW, and HZ participated in data collection and processing. ML and YL reviewed the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.cn) for English language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.801213/full#supplementary-material

REFERENCES

1. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, et al. Cancer Treatment and Survivorship Statistics, 2019. *CA: Cancer J Clin* (2019) 69(5):363–85. doi: 10.3322/caac.21565
2. Araujo-Filho JB, Halpenny D, Mcquade C, Puthoff G, Chiles C, Nishino M, et al. Management of Pulmonary Nodules in Oncologic Patients: Expert Panel Narrative Review. *AJR Am J Roentgenol* (2021) 216(6):1423–31. doi: 10.2214/AJR.20.24907
3. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2020. *CA: Cancer J Clin* (2020) 70(1):7–30. doi: 10.3322/caac.21590

4. Hong Y, Li Z, Zhang Q. A Circulating Tumor Cell Cluster-Based Model for Tumor Metastasis (Hypothesis). *Oncol Lett* (2016) 12(6):4891–5. doi: 10.3892/ol.2016.5358
5. Ko K-H, Huang H-K, Chen Y-I, Chang H, Tsai W-C, Huang T-W. Surgical Outcomes of Second Primary Lung Cancer After the Extrapulmonary Malignancy. *J Cancer Res Clin Oncol* (2020) 146(12):3323–32. doi: 10.1007/s00432-020-03310-x
6. Wu Y, Han C, Chong Y, Liu J, Gong L, Wang Z, et al. Prognostic Study for Survival Outcome Following the Treatment of Second Primary Lung Cancer in Patients With Previously Resected Non-Small Cell Lung Cancer. *Thorac Cancer* (2020) 11(10):2840–51. doi: 10.1111/1759-7714.13610
7. Wang Z, Wu Y, Wang L, Gong L, Han C, Xie F, et al. Role of Chemotherapy for Survival in Patients With Second Primary Non-Small Cell Lung Cancer. *Thorac Cancer* (2021) 12(4):426–43. doi: 10.1111/1759-7714.13762
8. Li Y-L, Chen C-H, Chen J-Y, Lai YS, Wang SC, Jiang SS, et al. Single-Cell Analysis Reveals Immune Modulation and Metabolic Switch in Tumor-Draining Lymph Nodes. *Oncoimmunology* (2020) 9(1):1830513. doi: 10.1080/2162402X.2020.1830513
9. Detterbeck FC, Franklin WA, Nicholson AG, Girard N, Arenberg DA, Travis WD, et al. The IASLC Lung Cancer Staging Project: Background Data and Proposed Criteria to Distinguish Separate Primary Lung Cancers From Metastatic Foci in Patients With Two Lung Tumors in the Forthcoming Eighth Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol Off Publ Int Assoc Study Lung Cancer* (2016) 11(5):651–65. doi: 10.1016/j.jtho.2016.01.025
10. Zhong F, Liu Z, Wang B, An W, Zhang H, Liao M, et al. A Predictive Model to Differentiate Between Second Primary Lung Cancers and Pulmonary Metastasis. *Acad Radiol* (2021) S1076-6332(21)00254–3. doi: 10.1016/j.acra.2021.05.015
11. Ohtaki Y, Shimizu K, Nagashima T, Nakazawa S, Obayashi K, Azuma Y, et al. Clinical and Radiological Discrimination of Solitary Pulmonary Lesions in Colorectal Cancer Patients. *World J Surg* (2018) 42(4):1161–70. doi: 10.1007/s00268-017-4243-9
12. Dijkman BG, Schuurbiers OCJ, Vriens D, Looijen-Salamon M, Bussink J, Timmer-Bonte JN, et al. The Role of (18)F-FDG PET in the Differentiation Between Lung Metastases and Synchronous Second Primary Lung Tumours. *Eur J Nucl Med Mol Imaging* (2010) 37(11):2037–47. doi: 10.1007/s00259-010-1505-2
13. Ge J, Gou H-F, Chen Y, Cheng K, Li LH, Dong H, et al. Clinical Characteristics of Patients With Solitary Pulmonary Mass After Radical Treatment for Primary Cancers: Pulmonary Metastasis or Second Primary Lung Cancer? *Cancer Invest* (2013) 31(6):397–403. doi: 10.3109/07357907.2013.800092
14. Nakadate A, Nakadate M, Sato Y, Nakagawa T, Yoshida K, Suzuki Y, et al. Predictors of Primary Lung Cancer in a Solitary Pulmonary Lesion After a Previous Malignancy. *Gen Thorac Cardiovasc Surg* (2017) 65(12):698–704. doi: 10.1007/s11748-017-0825-6
15. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
16. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial Intelligence in Cancer Imaging: Clinical Challenges and Applications. *CA: Cancer J Clin* (2019) 69(2):127–57. doi: 10.3322/caac.21552
17. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypinski P, Gibbs P, et al. Introduction to Radiomics. *J Nucl Med* (2020) 61(4):488–95. doi: 10.2967/jnumed.118.222893
18. Chen BT, Chen Z, Ye N, Mambetsariev I, Fricke J, Daniel E, et al. Differentiating Peripherally-Located Small Cell Lung Cancer From Non-Small Cell Lung Cancer Using a CT Radiomic Approach. *Front Oncol* (2020) 10:593–3. doi: 10.3389/fonc.2020.00593
19. Liu S, Liu S, Zhang C, Yu H, Liu X, Hu Y, et al. Exploratory Study of a CT Radiomics Model for the Classification of Small Cell Lung Cancer and Non-Small-Cell Lung Cancer. *Front Oncol* (2020) 10:1268–8. doi: 10.3389/fonc.2020.01268
20. Kim HY, Shim YM, Lee KS, Han J, Yi CA, Kim YK. Persistent Pulmonary Nodular Ground-Glass Opacity at Thin-Section CT: Histopathologic Comparisons. *Radiology* (2007) 245(1):267–75. doi: 10.1148/radiol.2451061682
21. Wu S, Zheng J, Li Y, Yu H, Shi S, Xie W, et al. A Radiomics Nomogram for the Preoperative Prediction of Lymph Node Metastasis in Bladder Cancer. *Clin Cancer Res* (2017) 23(22):6904–11. doi: 10.1158/1078-0432.CCR-17-1510
22. 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(3):947–57. doi: 10.1148/radiol.2016152234
23. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Making* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361
24. Zhu X, Dong D, Chen Z, Fang M, Zhang L, Song J, et al. Radiomic Signature as a Diagnostic Factor for Histologic Subtype Classification of Non-Small Cell Lung Cancer. *Eur Radiol* (2018) 28(7):2772–8. doi: 10.1007/s00330-017-5221-1
25. Hu Y, Weng Q, Xia H, Chen T, Kong C, Chen W, et al. A Radiomic Nomogram Based on Arterial Phase of CT for Differential Diagnosis of Ovarian Cancer. *Abdom Radiol (NY)* (2021) 46(6):2384–92. doi: 10.1007/s00261-021-03120-w
26. Kirienko M, Cozzi L, Rossi A, Voulaz E, Antunovic L, Fogliata A, et al. Ability of FDG PET and CT Radiomics Features to Differentiate Between Primary and Metastatic Lung Lesions. *Eur J Nucl Med Mol Imaging* (2018) 45 (10):1649–60. doi: 10.1007/s00259-018-3987-2
27. Zhou Y, Ma X-L, Zhang T, Wang J, Zhang T, Tian R. Use of Radiomics Based on F-FDG PET/CT and Machine Learning Methods to Aid Clinical Decision-Making in the Classification of Solitary Pulmonary Lesions: An Innovative Approach. *Eur J Nucl Med Mol Imaging* (2021) 48(9):2904–13. doi: 10.1007/s00259-021-05220-7
28. Lenga L, Bernatz S, Martin SS, Booz C, Solbach C, Mulert-Ernst R, et al. Iodine Map Radiomics in Breast Cancer: Prediction of Metastatic Status. *Cancers (Basel)* (2021) 13(10):2431. doi: 10.3390/cancers13102431
29. Mao B, Ma J, Duan S, Xia Y, Tao Y, Zhang L. Preoperative Classification of Primary and Metastatic Liver Cancer *via* Machine Learning-Based Ultrasound Radiomics. *Eur Radiol* (2021) 31(7):4576–86. doi: 10.1007/s00330-020-07562-6
30. Wen Q, Yang Z, Dai H, Feng A, Li Q. Radiomics Study for Predicting the Expression of PD-L1 and Tumor Mutation Burden in Non-Small Cell Lung Cancer Based on CT Images and Clinicopathological Features. *Front Oncol* (2021) 11:620246–6. doi: 10.3389/fonc.2021.620246
31. Shao X, Niu R, Shao X, Jiang Z, Wang Y. Value of (18)F-FDG PET/CT-Based Radiomics Model to Distinguish the Growth Patterns of Early Invasive Lung Adenocarcinoma Manifesting as Ground-Glass Opacity Nodules. *EJNMMI Res* (2020) 10(1):80–0. doi: 10.1186/s13550-020-00668-4
32. Suh YJ, Lee H-J, Sung P, Yoen H, Kim S, Han S, et al. A Novel Algorithm to Differentiate Between Multiple Primary Lung Cancers and Intrapulmonary Metastasis in Multiple Lung Cancers With Multiple Pulmonary Sites of Involvement. *J Thorac Oncol: Off Publ Int Assoc Study Lung Cancer* (2020) 15(2):203–15. doi: 10.1016/j.jtho.2019.09.221
33. Boyle JM, Tandberg DJ, Chino JP, D'amico TA, Ready NE, Kelsey CR. Smoking History Predicts for Increased Risk of Second Primary Lung Cancer: A Comprehensive Analysis. *Cancer* (2015) 121(4):598–604. doi: 10.1002/cncr.29095
34. Huang L, Xia W, Zhang B, Qiu B, Gao X. MSFCN-Multiple Supervised Fully Convolutional Networks for the Osteosarcoma Segmentation of CT Images. *Comput Methods Programs BioMed* (2017) 143:67–74. doi: 10.1016/j.cmpb.2017.02.013

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 03 February 2022
doi: 10.3389/fonc.2022.803824

Positron Emission Tomography/ Magnetic Resonance Imaging Radiomics in Predicting Lung Adenocarcinoma and Squamous Cell Carcinoma

Xin Tang[1,2†], Jiangtao Liang[3†], Bolin Xiang[4], Changfeng Yuan[2], Luoyu Wang[5], Bin Zhu[4], Xiuhong Ge[6], Min Fang[4] and Zhongxiang Ding[6]*

[1] The Fourth Clinical College, Zhejiang Chinese Medical University, Hangzhou, China, [2] Department of Radiology, Hangzhou Wuyunshan Hospital (Hangzhou Health Promotion Research Institute), Hangzhou, China, [3] Department of Radiology, Hangzhou Universal Medical Imaging Diagnostic Center, Hangzhou, China, [4] Department of Radiology, Zhejiang Quhua Hospital, Quzhou, China, [5] Centre for Cognition and Brain Disorders, Hangzhou Normal University, Hangzhou, China, [6] Department of Radiology, Key Laboratory of Clinical Cancer Pharmacology and Toxicology Research of Zhejiang Province, Affiliated Hangzhou First People's Hospital, Cancer Center, Zhejiang University School of Medicine, Hangzhou, China

Objective: To investigate the diagnostic value of positron emission tomography (PET)/ magnetic resonance imaging (MRI) radiomics in predicting the histological classification of lung adenocarcinoma and lung squamous cell carcinoma.

Methods: PET/MRI radiomics and clinical data were retrospectively collected from 61 patients with lung cancer. According to the pathological results of surgery or fiberscope, patients were divided into two groups, lung adenocarcinoma and squamous cell carcinoma group, which were set as positive for adenocarcinoma (40 cases) and negative for squamous cell carcinoma (21 cases). The radiomics characteristics most related to lung cancer classification were calculated and selected using radiomics software, and the two lung cancer groups were randomly assigned into a training set (70%) and a test set (30%). Maximum relevance and minimum redundancy (mRMR) and least absolute shrinkage and selection operator (LASSO) methods in the uAI Research Portal software (United Imaging Intelligence, China) were used to select the desired characteristics from 2600 features extracted from MRI and PET. Eight optimal features were finally retained through 5-fold cross-validation, and a PET/MRI fusion model was constructed. The predictive ability of this model was evaluated by the difference in area under the curve (AUC) obtained from the receiver operating characteristic (ROC) curve.

Results: AUC of PET/MRI model for the training group and test group were 0.886 (0.787-0.985) and 0.847 (0.648-1.000), respectively. PET/MRI radiomics features revealed different degrees of correlation with the classification of lung adenocarcinoma and squamous cell carcinoma, with significant differences.

OPEN ACCESS

Edited by:
Yiyan Liu,
University of Louisville, United States

Reviewed by:
Jun Shen,
Sun Yat-sen University, China
Wenli Cai,
Massachusetts General Hospital and Harvard Medical School, United States

***Correspondence:**
Zhongxiang Ding
hangzhoudzx73@126.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *28 October 2021*
Accepted: *03 January 2022*
Published: *03 February 2022*

Citation:
Tang X, Liang J, Xiang B, Yuan C, Wang L, Zhu B, Ge X, Fang M and Ding Z (2022) Positron Emission Tomography/Magnetic Resonance Imaging Radiomics in Predicting Lung Adenocarcinoma and Squamous Cell Carcinoma. Front. Oncol. 12:803824. doi: 10.3389/fonc.2022.803824

Conclusion: The prediction model constructed based on PET/MRI radiomics features can predict the preoperative histological classification of lung adenocarcinoma and squamous cell carcinoma without seminality and repeatability. It can also provide an objective basis for accurate clinical diagnosis and individualized treatment, thus having important guiding significance for clinical treatment.

Keywords: lung, cancer, PET/MRI, radiomic, adenocarcinoma, squamous

INTRODUCTION

In 2020, there were 19.29 million new cancer cases and 9.96 million cancer-related deaths worldwide, among which lung cancer accounted for 2.2 million (11.4%) and 1.8 million (18%), respectively. Thus, lung cancer has become the second most common cancer and the leading cause of death globally. In addition, lung cancer occurs more frequently in men with the highest incidence and mortality in males compared to other tumors (1). Lung cancer is divided into adenocarcinoma, squamous cell carcinoma, small cell carcinoma, etc., of which lung adenocarcinoma and squamous cell carcinoma are the main types, accounting for about 75% (2, 3). The treatment methods vary for different pathological types of lung cancer, and early diagnosis is of great significance for the diagnosis, overall treatment, and personalized treatment of patients with lung cancer.

Over recent years, the diagnosis and treatment of lung cancer have been further improved by the integration of radiomics, molecular biology, clinical and other disciplines. With the progress of imaging technology and the continuous development of drugs, especially the popularization and application of PET/CT and PET/MRI molecular radiomics technology, the performance level of clinical diagnosis and treatment efficacy evaluation of lung cancer have been greatly advanced. Sepehri et al. (4) found that the PET/CT radiomics-based model outperformed the standard clinical staging by retrospectively analyzing 138 patients with stage II-III non-small cell lung cancer. Ehman et al. (5) found that PET/CT had superiority in terms of use opportunity, application cost, examination speed and clinical awareness, but PET/MRI produced less radiation and was more advantageous in the detection of soft tissue tumors. In addition, for the staging of breast cancer, compared to PET/CT, PET/MRI can better distinguish the invasion of chest wall, diaphragm and mediastinum/distant soft tissues, which affected the TMN staging. As a result of the fusion of the metabolic information by PET with the high soft-tissue resolution and functional information by MRI, PET/MRI has gained more advantages in detecting primary soft tissue lesions, histopathological classification, TMN staging, prognosis prediction, efficacy evaluation, and recurrence detection. At the same time, as the fusion radiomics can determine the accurate location of the lesion and the anatomical relationship with the surrounding tissues, it has obvious advantages in determining the biological target area for lung cancer radiotherapy and formulating the extent of surgical resection. Thus, in the treatment of lung cancer, PET/MRI can be used for early observation of the tumor's response to treatment, timely adjustment and optimization of the treatment plan, avoidance of ineffective treatment or toxic side effects, gaining treatment time for patients, improving the therapeutic effect, prolonging the survival time of patients and improving the quality of life.

The current diagnosis and treatment of lung cancer still mainly rely on the subjective experience of physicians and clinicians, and there is a lack of systematic analysis of the data information generated by radiomics examinations. Needle biopsy is the gold standard for pathological diagnosis of patients; still, it is invasive, reproducible, has potential complications, and is difficult to perform when the lesion is deep or adjacent to blood vessels. Therefore, this method has certain limitations and may even lead to fatal outcomes (3). However, radiomics methods use automated data characterization algorithms to transform medical images into high-resolution graphics, excavate feature spatial data, and quantify lesion morphological characteristics and internal heterogeneity (6–9). Deep mining of radiomics data can obtain many quantitative radiomics characteristics that the human eye cannot perceive.

This study aimed to find new radiomics quantitative parameters for histological classification of lung adenocarcinoma and squamous cell carcinoma based on PET/MRI radiomics method, construct a prediction model, and explore the diagnostic value of this technique in predicting the classification of lung adenocarcinoma and squamous cell carcinoma without seminality.

MATERIALS AND METHODS

Subjects

A total of 61 patients with lung adenocarcinoma or squamous cell carcinoma confirmed by surgery or puncture, including 40 with lung adenocarcinoma and 21 with squamous cell carcinoma, who were initially diagnosed by PET/MRI examination in Hangzhou Universal Medical Imaging Diagnostic Center between October 2018 and August 2021 were retrospectively included in the study. The research protocol met the requirements of medical ethics (Scientific Research Medical Ethics, No. 2021-008), and all methods were implemented in accordance with the Declaration of Helsinki.

Inclusion criteria were the following: all patients underwent PET/MR examination before treatment and were pathologically confirmed to have adenocarcinoma or squamous cell carcinoma; no chemotherapy or radiotherapy and surgical anti-tumor

therapy were performed; clear whole-body and chest PET/MR could be obtained before treatment; PET/MR examination was performed 40 – 60 min after injection of 18F-fluorodeoxyglucose (18F-FDG).

Exclusion criteria were: patients whose PET/MRI image failed to meet the diagnostic criteria (such as obvious metal or motion artifact); patients with contraindications to MRI examination or inability to tolerate the examination; patients with a history of other thoracic malignant tumors or other systemic malignancies; patients who had received any form of treatment before PET/MR examination (such as radiotherapy, chemotherapy, etc.); pathologically confirmed adenocarcinoma and squamous cell carcinoma of other histopathological types.

Instruments and Equipments

Imaging data were acquired using integrated time-of-flight (TOF) PET/MR from GE (GE SIGNA, WI, USA). The system consisted of a PET detector with TOF technology (TOF-PET) and the latest generation of 750W 3.0T magnetic resonance. The TOF-PET detector is constructed with a state-of-the-art solid-phase array photoelectric converter (SiPM) and a new generation of LBS crystal. Simultaneous PET and MRI scanning were performed with the thinnest acquisition slice thickness of 2.8 mm (Transverse FOV: 60 cm, axial FOV: 25 cm, transverse resolution (1 cm from the center): 4.2 mm, axial resolution (1 cm from the center): 5.8 mm, temporal resolution: 385 ps, energy resolution: 11%, and sensitivity: 21 cps/kBq).

Patients Preparation

Patients were required to fast for more than 6 hours, and the blood glucose concentration was controlled to be < 7.8 mmol/L before injection of ^{18}F-FDG. On the examination day, patients wore clothes that did not have any accessories or were easy to take off. They were injected with ^{18}F-FDG at a dose of 3.7 Mbq/kg and underwent whole-body PET/MRI 40 min later. Written informed consent was obtained from all patients or legal guardians before the examination.

PET/MRI Scan

The patient was placed in the supine position. After performing attenuation correction, whole-body PET/MR scans were performed from the top of the head to the middle of the femur, and if necessary, sweeping to the sole of the foot. A total of 5 – 6 beds were collected, with an acquisition time of 6 minutes per bed. PET images was acquired and reconstructed using 3D mode, TOF technique, and point spread function (PSF) with ordered subset expectation maxima (OSEM) algorithm, which used two iterations, 28 subsets, and a 5 mm Gaussian post-processing filter with a 192 × 192 matrix. PET data acquisition was performed during a whole-body MRI examination. A regional PET/MR scan of the chest was then carried out, ranging from the apex to the base of the lung, and radiomics were obtained using dedicated MRI coils for the chest region, resulting in whole-body and regional PET, MRI, and PET/MR fusion radiomics. All data were acquired from the same PET/MR instrument. MRI sequences included LAVA-Flex T1WI, fs-PROPELLER T2WI, DWI (b = 800 mm^2/s), and coronal fs-PROPELLER T2WI. In this study, chest local Axial T2WI radiomics and PET radiomics were selected as radiomics feature extraction sequences (10, 11).

Radiomics Data Processing

Conjoined uAI Research Portal software (United Imaging Intelligence, China) that was embedded into the widely used package-PyRadiomics (https://pyradiomics.readthedocs.io/en/latest/index.html) was used for radiomics analysis on the region of interest (ROI) of the subject's primary tumor. The workflow of radiomics mainly included the following steps: lesion segmentation, feature extraction, feature selection, and machine learning modeling (12–15).

Lesion Segmentation

Chest PET and MRI data in DICOM format were imported into ITK-SNAP software (http://www.itksnap.org), which was used to delineate the region of interest (ROI) of the patient on PET and MRI axial data, manually delineate along the edge of the primary tumor of lung cancer, exclude adjacent normal tissues and lymph nodes, overlap the segmentation boundaries of PET and MRI data, and finally export the three-dimensional segmentation results obtained by PET and MRI radiomics sequentially into the original map and the corresponding ROI map (**Figures 1–4**). The segmentation result was saved as nii file. Two radiologists with 15 to 20 years of experience in thoracic PET/MR diagnosis simultaneously segmented ROIs on MRI and PET images of the primary lesion to obtain the corresponding ROI segmented graphics, respectively. When the results were inconsistent, the third radiologist with twenty years of experience performed ROI delineation again and checked until the results were unified.

Feature Extraction

All ROI data and the original images of PET and MRI were imported into the uAI Research Portal software in batch.

Feature Selection

Data Import

The radiomics of PET and MRI were imported into the R software (version 4.0.5, http://www.Rproject.org) for feature selection.

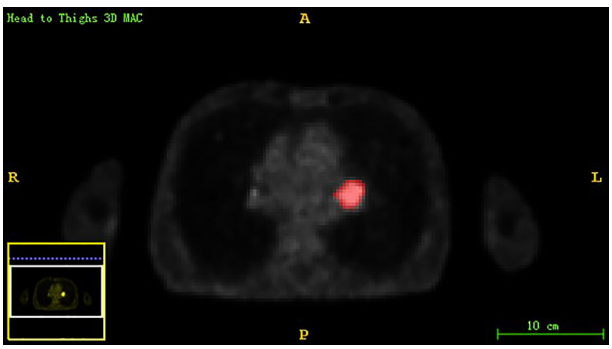


FIGURE 1 | The lesion ROIs of axial PET sequence.

FIGURE 2 | The lesion ROIs of coronal PET sequence.

FIGURE 4 | The lesion ROIs of axial MRI sequence.

Feature Selection

The patients were randomly assigned into a training set (70%) and a test set (30%) (12, 16, 17). We used two feature selection methods, mRMR and LASSO, to select the features. Firstly, mRMR was performed to eliminate the redundant and irrelevant features; then LASSO was conducted to choose the optimized subset of features to construct the final model.

1. LASSO analysis included choosing the regular parameter λ, determining the number of the feature (**Figure 5**). After the number of feature was determined, the most predictive feature subset was chosen and the corresponding coefficients were evaluated (**Figure 6**).

2. After the number of feature was determined, the most predictive feature subset was chosen and the corresponding coefficients were evaluated (**Figure 7**).

The radiomic signature (Rad score) was calculated by summing the selected texture features, which were weighted by their respective coefficients. All rad scores between lung adenocarcinoma and squamous cell carcinoma group were compared in the training set and test set respectively.

The final formula for the PET/MRI rad score was: "Radscore=-1.161*PET_wavelet_glszm_wavelet-HHL-GrayLevelNonUniformity+-0.147*PET_boxsigmaimage_glcm_ClusterShade+-0.516*PET_normalize_glrlm_ShortRunLowGrayLevelEmphasis+

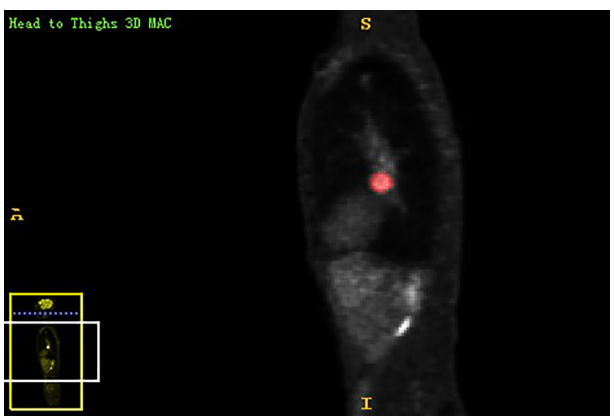


FIGURE 3 | The lesion ROIs of sagittal PET sequence.

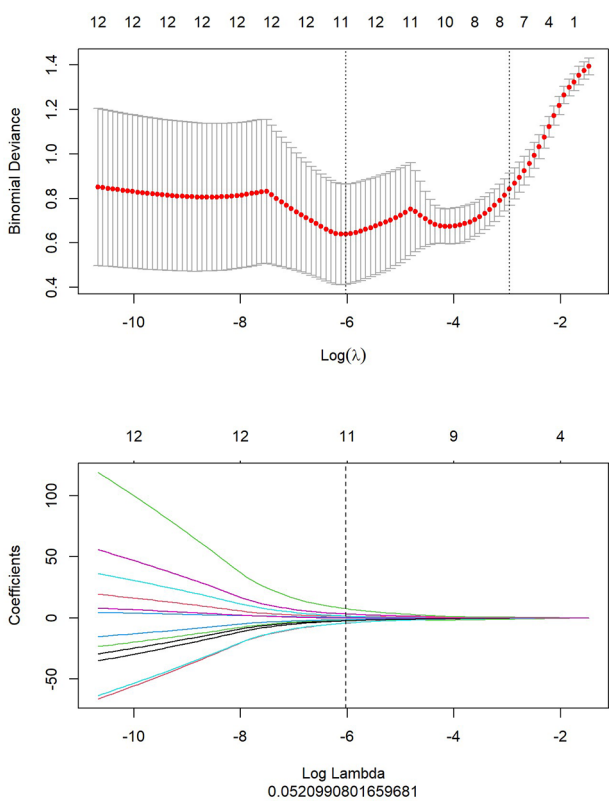


FIGURE 5 | LASSO analysis of PET/MRI.

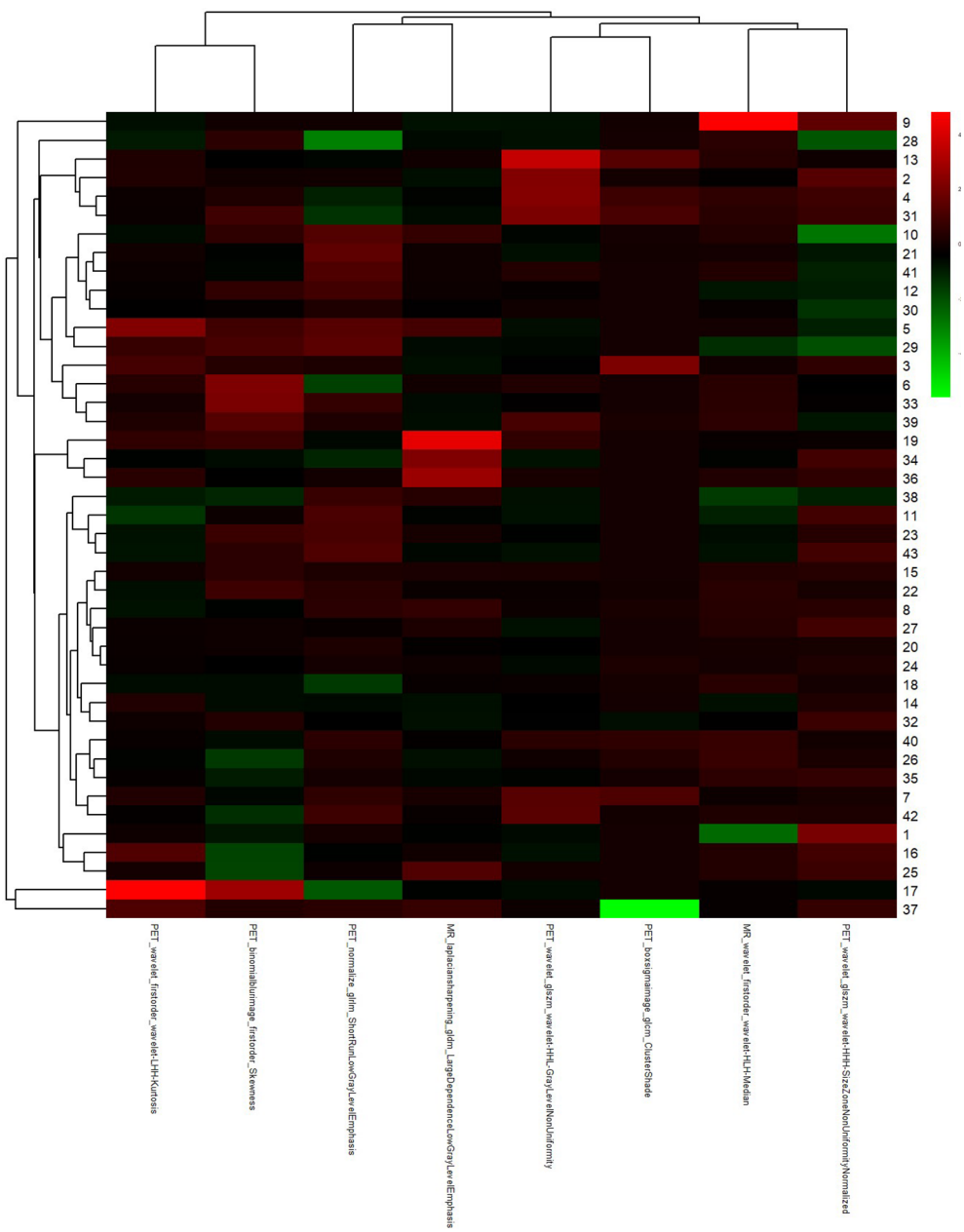


FIGURE 6 | The most predictive feature subset of PET/MRI.

0.311*MR_wavelet_firstorder_wavelet-HLH-Median+-0.332*PET_wavelet_firstorder_wavelet-LHH-Kurtosis+0.336*PET_wavelet_glszm_wavelet-HHH-SizeZoneNonUniformityNormalized+0.491*MR_laplaciansharpening_gldm_LargeDependenceLowGrayLevelEmphasis+-0.021*PET_binomialblurimage_firstorder_Skewness + 0.449"

Radiomics Validation

We used ROC analysis to evaluate the performance of the model (**Figure 10**).

Nomogram Building

Statistical Analysis

Statistical comparisons of gender and age were performed using SPSS software (version 26). In addition, feature selection and radiomics signature construction and validation were conducted with R software. The statistical significance was set at a *P*-value of 0.05 with two-tailed analyses (3, 18). Feature extraction ROC measured the evaluation consistency between radiologists using the inter-correlation coefficients (ICC). All statistical methods of

FIGURE 7 | The radiomic coefficients of each feature in the most predictive feature subset in PET/MRI radiomics signature construction.

the radiomics analysis were conducted with uAI Research Portal software and R software.

RESULTS

In this study, the ICC value of > 0.86, which was considered to be in good agreement of the ROC. We selected ROC results from the senior radiologist to extract features.

Comparison of Clinical Data

Table 1 showed the results of statistical analysis of the demographic and clinical data. There were no statistically significant differences in age and gender between the lung adenocarcinoma and lung squamous cell carcinoma groups (**Table 1**).

Radiomics Analysis Results

Figure 5 showed that top 20 imaging features were ranked and used as candidate features for LASSO regression analysis, according to the results with mRMR algorithm. In LASSO regression analysis, when Log (In) was -6 and Log Lambda was 0.521, the PET/MRI prediction model showed the best diagnostic performance, and the eight optimal imaging features were determined at this point.

Figures 6, **7** showed that eight characteristic parameters were obtained from the PET/MRI prediction model, which were PET_wavelet_glszm_wavelet-HHL-GrayLevelNonUniformity; PET_boxsigmaimage_glcm_ClusterShade; PET_normalize_glrlm_ShortRunLowGrayLevelEmphasis; MR_wavelet_firstorder_wavelet-HLH-Median; PET_wavelet_firstorder_wavelet-LHH-Kurtosis; PET_wavelet_glszm_wavelet-HHH-SizeZoneNonUniformityNormalized; MR_laplaciansharpening_gldm_LargeDependenceLowGrayLevelEmphasis; and PET_binomialblurimage_firstorder_Skewness.

Figures 8A, B showed that the predicted values of the training and test group were very close to the actual values, and they showed that the prediction ability of the nomogram was good.

Figure 9 showed that labels "0" and "1" were added to the rad scores of the training and test groups, respectively, where adenocarcinoma was labeled "1" and scale-cell carcinoma was labeled "0".

Figure 10 showed that AUC of PET/MRI model in the training and test group was 0.886 (0.787-0.985) and 0.847 (0.648-1.000), respectively. Based on Youden Index, other parameters were calculated as follows (**Table 2**):

Figures 11A, B showed the Hosmer-Lemeshow test result.

Finally, we used decision curve to evaluate the clinical usefulness of the model (**Figure 12**).

TABLE 1 | Summary of original data of cases.

Characteristic	Adenocarcinoma	squamous cell carcinoma	Statistical analysis	*P*-value
Gender			chi-square	$p = 0.918$
Male (case)	10	16		
Female (case)	30	5		
Age (years)	23-90	51-80	A two-sample T-test	$P = 0.834$

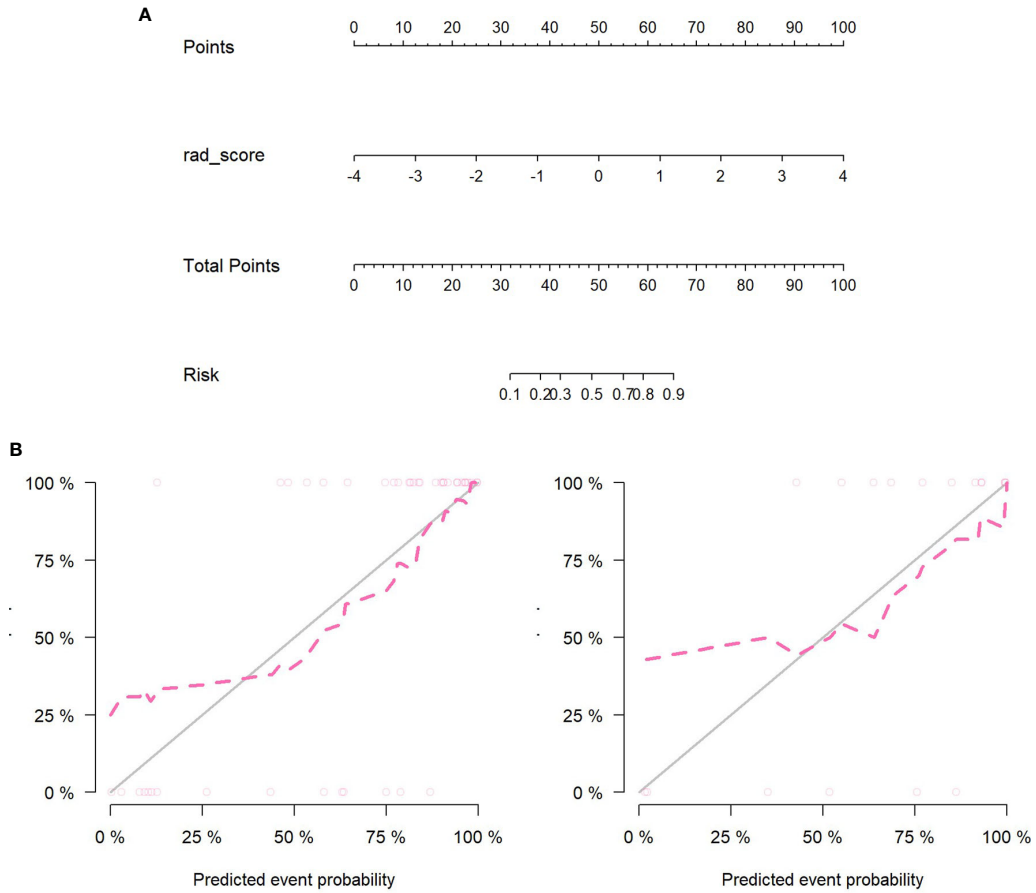


FIGURE 8 | **(A)** Nomogram construction in the PET/MRI training set. **(B)** The left figure is the training set calibration curve; the right figure is the test set calibration curve; the solid black line represents the theoretical curve, and the red dashed line represents the deviation correction curve. The formula of PET/MRI nomoscore is defined as follows: "Nomoscore = (Intercept)*0.188+rad_score*1.648".

DISCUSSION

In this study, PET/MRI prediction models established based on chest MRI and PET radiology were used to analyze 40 patients with lung adenocarcinoma and 21 patients with squamous cell carcinoma, and finally 8 optimal characteristic parameters were obtained, of which 3 belonged to intensity characteristics and 5 belonged to texture characteristics, 6 characteristics were from PET radiology and 2 characteristics were from MRI radiology. These results indicated that texture characteristics were more related to lung adenocarcinoma and lung squamous cell carcinoma classification, of which the characteristics with the highest feature weight ratio was wavelet GLSZM-HHL-GLNU, indicating that it had the most significant predictive effect on NSCLC patient subtypes in PET/MRI models, and GLSZM was a standardized distribution of regional counts relative to gray values. The lower the value was, the more uniform the intensity value became. Moreover, GLSZM is negatively correlated with survival and helps identify hypoxic or necrotic areas with poor prognosis. Yang et al. (19) showed that gray level run length matrix (GLRLM) and wavelet characteristics were related to the survival time of lung cancer. Additionally, the AUC of the PET/MRI model in the training and test set were 0.886 (0.787 – 0.985) and 0.847 (0.648 – 1.000), respectively, which were very close. The results showed that the PET/MRI prediction model had good fit, good consistency and stability, and also showed that the PET/MRI prediction model could effectively and non-invasively classify the pathological types of lung adenocarcinoma and lung squamous cell carcinoma. The diagnostic value of PET/MR prediction model in the training and test set suggested that the model constructed based on PET/MRI radiomics characteristics had a high predictive value for preoperative pathological classification of lung adenocarcinoma and squamous cell carcinoma, providing an objective basis for accurate clinical diagnosis and individualized treatment, and having an important guiding significance for clinical treatment. The Hosmer-Lemeshow test indicated that the PET/MRI prediction model could effectively differentiate the pathological subtypes of lung adenocarcinoma from lung squamous cell carcinoma. Decision curve analysis of different variables for clinical application of PET/MRI model showed that the net benefit of the PET/MRI radiomics model under different

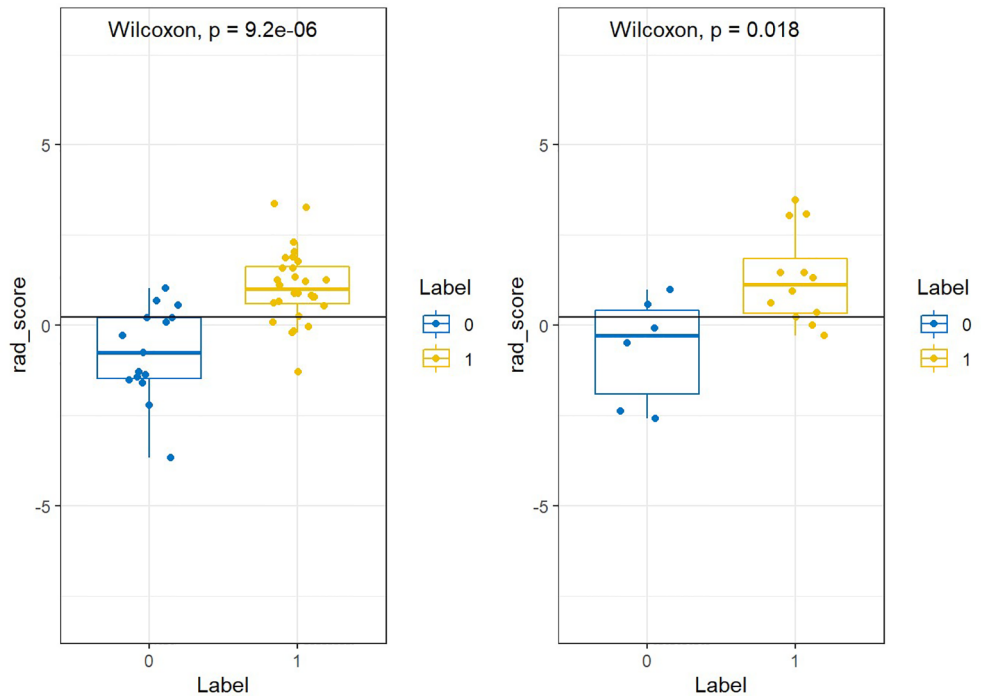


FIGURE 9 | The label of PET/MRI.

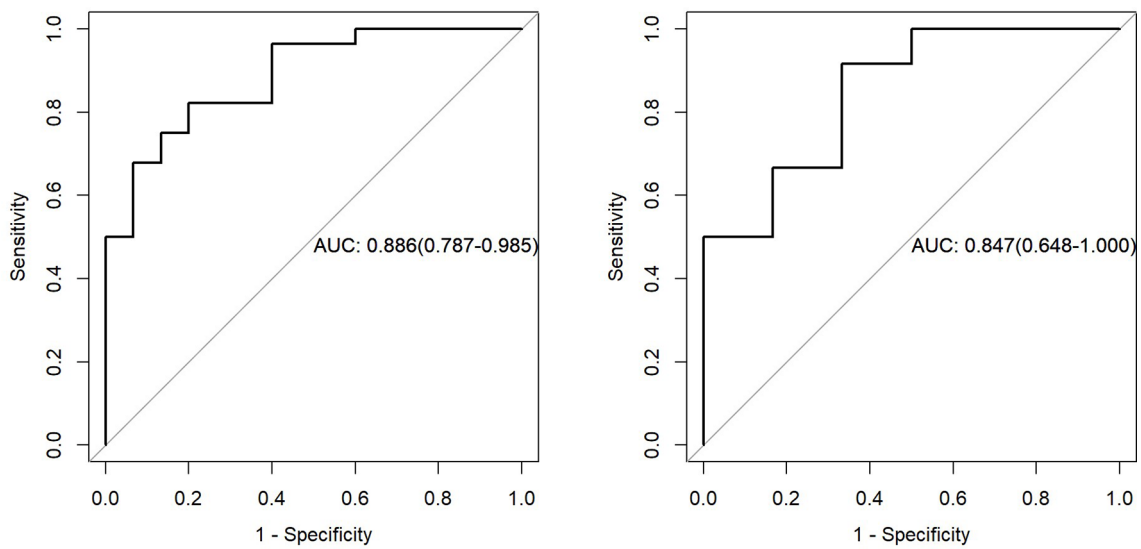


FIGURE 10 | The left figure represents the AUC of PET/MRI model in the training set, and the right figure represents the AUC of PET/MRI model in the test set.

TABLE 2 | Results of PET/MRI radiomics.

Group	Accuracy	Accuracy Lower	Accuracy Upper	Sensitivity	Specificity	Pos.Pred.Value	Neg.Pred.Value
Training	0.814	0.666	0.916	0.800	0.821	0.706	0.885
Test	0.833	0.586	0.964	0.667	0.917	0.800	0.846

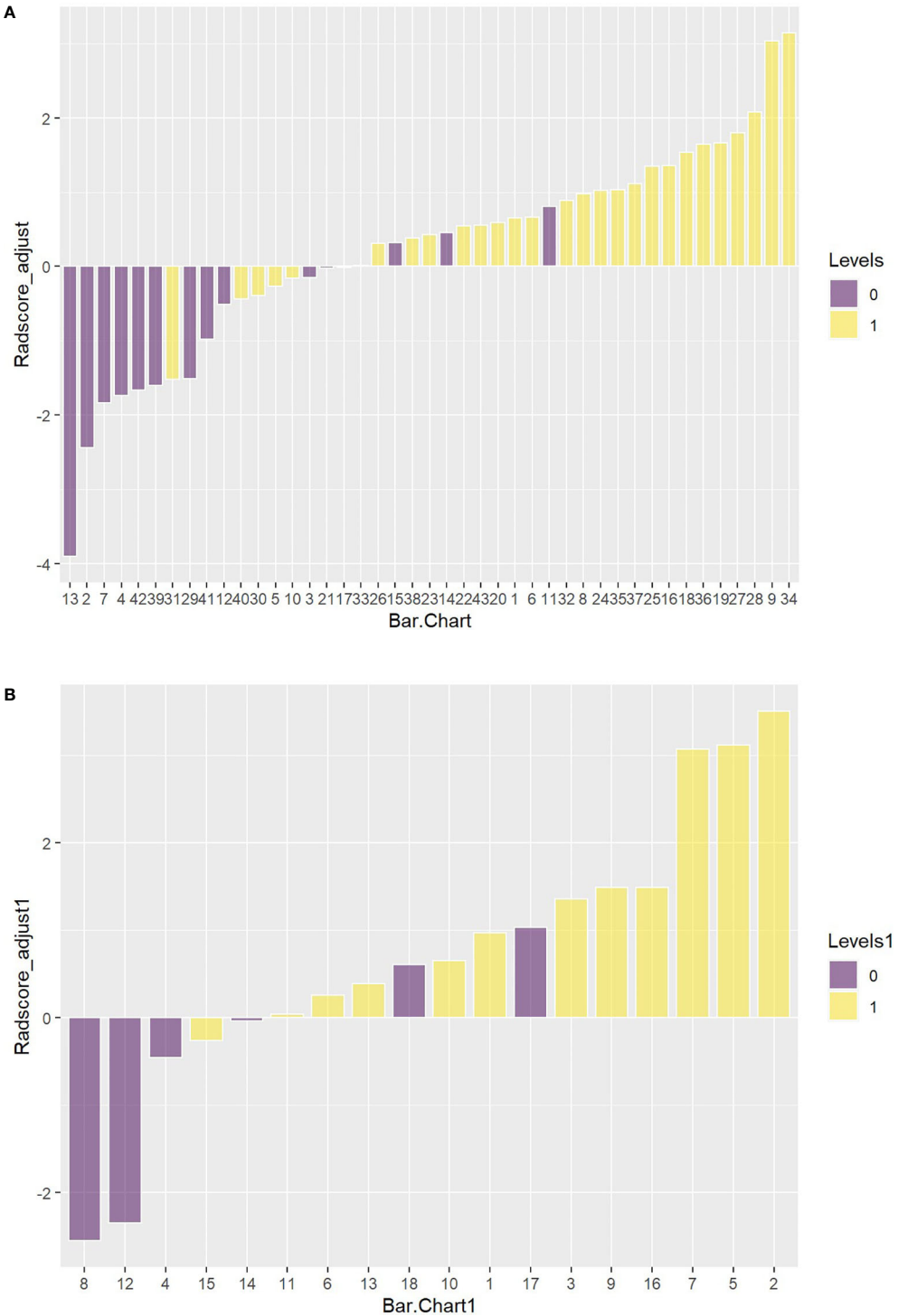


FIGURE 11 | **(A)** The Hosmer-Lemeshow result for the PET/MRI model in the training set. **(B)** The Hosmer-Lemeshow results for PET/MRI model in the test set.

threshold probabilities outweighed the clinical variables. We found that these curves could further reflect the clinical utility and the higher predictive efficiency of PET/MRI radiomics models.

Coroller et al. (20) analyzed the CT radiomics of 85 patients with locally advanced NSCLC and extracted the radiomics parameters of lymph nodes and primary tumors using radiomics methods. They reported that the phenotypic

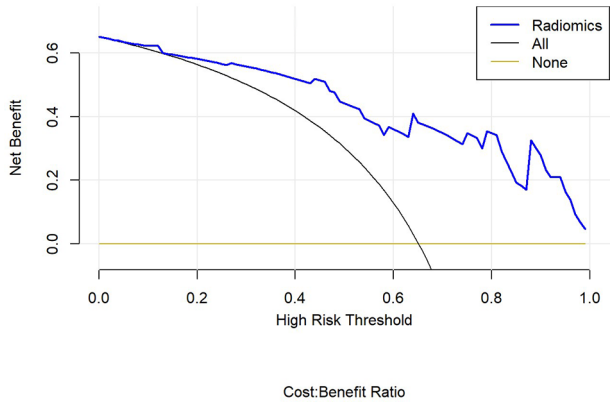


FIGURE 12 | Decision curve analysis for different variables in PET/MRI models.

information of lymph nodes was more effective than that of primary tumors in predicting the pathological responses. Kirienko et al. (21) retrospectively analyzed the radiomics data of PET and CT of 534 lung lesions cases and found that the texture characteristics of PET radiomics using the method of linear discriminant analysis could distinguish primary lung cancer from metastatic tumors (AUC > 0.90), and could classify the histological subtypes of primary lung cancer (AUC 0.61, 0.97). Other studies have also evaluated the relationship between tumor subtypes, histopathological grades, diagnosis, treatment, and prognosis of lung cancer and radiomics characteristics. Orlhac and colleagues (22) compared the texture characteristics of adenocarcinoma and squamous cell carcinoma and showed differences in most texture characteristics, while squamous cell carcinoma had lower homogeneity and higher entropy, reflecting its higher heterogeneity than adenocarcinoma. Yang et al. (23) retrospectively analyzed PET/CT radiology data of 315 NSCLC patients and found that the radiological nomograms based on 18F-FDG PET/CT rad score and clinicopathological factors had a good predictive performance for survival outcomes, providing feasible and practical guidance for individualized management of NSCLC patients. Szyszko et al. (24) and Grosse et al. (25) reported that the main advantages of PET/MRI were the reduction of radiation dose and the improvement of anatomical details in soft tissue areas, making it suitable for pediatric patients and patients requiring repeated radiomics. Hyun et al. (26) established a PET prediction model that successfully predicted histological subtypes of lung cancer and found that the logistic regression model outperformed all other classifiers (AUC = 0.859, accuracy = 0.769). In the present study, we used mRMR and LASSO methods to construct a combined PET/MRI prediction model based on chest MRI and PET, and showed that the results of the training group, with an AUC of 0.886 and an accuracy of 0.814. We concluded that the PET/MRI prediction model was superior to the PET prediction model in terms of predicting the pathological types of lung adenocarcinoma and squamous cell carcinoma, and the PET/MRI radiomics might be more helpful for clinicians to improve the histopathological diagnosis. As a non-invasive radiomics method, PET/MRI can significantly reduce the radiation dose compared with PET/CT and provide morphological, functional, and molecular radiomics information of the tumor in one examination. Compared with histopathological and genetic testing methods, PET/MR examination can overcome sampling deviations and complications caused by biopsy. It can also provide more comprehensive and accurate information in predicting biomarkers. This study showed that PET/MRI could be used for non-seminal evaluation and prediction of lung adenocarcinoma and squamous cell carcinoma, which was conducive to develop a specific and individualized treatment plan for lung cancer patients in clinical practice.

The present study had some limitations. Firstly, this was a single-center retrospective study with a limited sample size. Furthermore, the cost of PET/MRI was too expensive and PET/MRI examination required the use of multi-sequence MR radiomics. The examination time was too long, the patient could not fully cooperate statically, and the rapidly changing gradient of MRI produced great noise, resulting in the inability to directly assess tissue density, especially for the lung and bone. Therefore, an accurate attenuation correction map was obtained, and the interference caused by different radiomics acquisition parameters and respiratory motion displacement might reduce the diagnostic accuracy of the model. In the future, more data on lung cancer patients' samples can be obtained, new MRI examination sequences can be developed, respiratory and motion artifacts can be improved, more comfortable bed surfaces and audio and video equipments for PET/MRI machines can be improved, and examination time can be shortened without affecting radiomics quality. It is necessary to select appropriate machine learning algorithms, build multi-modal and multi-center cooperation, improve the prediction efficiency, minimize the risk of overfitting, perform more refined sample data processing, and construct artificial intelligence classification models with higher complexity that may have an important role in the accurate classification and prediction of lung adenocarcinoma and lung squamous cell carcinoma. At present, the advantages of PET/MRI do not exceed their disadvantages. Nevertheless, PET/MRI is equivalent or complementary to PET/CT for lung tumor detection. At the same time, the radiomics research is still in its infancy, and there is a lot of possibility for future development.

With the development of artificial intelligence and radiomics, a multi-modal combination of clinical, radiomics, and pathological data will be adopted. Therefore, it is believed that imaging-based PET/MRI prediction for lung cancer classification will have a promising future in the clinical auxiliary diagnosis of lung cancer.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Scientific Research Medical Ethics, No. 2021-008. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Guarantor of integrity of the entire study: ZD. Study concepts: XT, JL, BX. Study design: XT, JL, CY. Literature research: XT, BX. Data acquisition: XT, JL, BZ, MF, XG. Statistical analysis: LW. Manuscript preparation: XT, JL. Manuscript editing: XT. Manuscript review: ZD. All authors contributed to the article and approved the submitted version.

FUNDING

National Natural Science Foundation of China (81871337), Natural Science Foundation of Zhejiang Province (Y22H185692, LY16H180007), and Science Foundation from the Health Commission of Zhejiang Province (A20200507).

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
2. Travis WD. Pathology of Lung Cancer. *Clin Chest Med* (2011) 32(4):669–92. doi: 10.1016/j.ccm.2011.08.005
3. Li H, Gao L, Ma H, Arefan D, He J, Wang J, et al. Radiomics-Based Features for Prediction of Histological Subtypes in Central Lung Cancer. *Front Oncol* (2021) 11:658887. doi: 10.3389/fonc.2021.658887
4. Sepehri S, Tankyevych O, Upadhaya T, Visvikis D, Hatt M, Cheze Le Rest C. Comparison and Fusion of Machine Learning Algorithms for Prospective Validation of PET/CT Radiomic Features Prognostic Value in Stage II-III Non-Small Cell Lung Cancer. *Diagn (Basel)* (2021) 11(4):675. doi: 10.3390/diagnostics11040675
5. Ehman EC, Johnson GB, Villanueva-Meyer JE, Cha S, Leynes AP, Larson PEZ, et al. PET/MRI: Where Might it Replace PET/Ct? *J Magn Reson Imaging* (2017) 46(5):1247–62. doi: 10.1002/jmri.25711
6. Schacht DV, Drukker K, Pak I, Abe H, Giger ML. Using Quantitative Image Analysis to Classify Axillary Lymph Nodes on Breast MRI: A New Application for the Z 0011 Era. *Eur J Radiol* (2015) 84(3):392–7. doi: 10.1016/j.ejrad.2014.12.003
7. Andersen MB, Harders SW, Ganeshan B, Thygesen J, Torp Madsen HH, Rasmussen F. CT Texture Analysis can Help Differentiate Between Malignant and Benign Lymph Nodes in the Mediastinum in Patients Suspected for Lung Cancer. *Acta Radiol* (2016) 57(6):669–76. doi: 10.1177/0284185115598808
8. Kunimatsu A, Kunimatsu N, Kamiya K, Watadani T, Mori H, Abe O. Comparison Between Glioblastoma and Primary Central Nervous System Lymphoma Using MR Image-Based Texture Analysis. *Magn Reson Med Sci* (2018) 17(1):50–7. doi: 10.2463/mrms.mp.2017-0044
9. Zhou W, Zhang L, Wang K, Chen S, Wang G, Liu Z, et al. Malignancy Characterization of Hepatocellular Carcinomas Based on Texture Analysis of Contrast-Enhanced MR Images. *J Magn Reson Imaging* (2017) 45(5):1476–84. doi: 10.1002/jmri.25454
10. Fan L, Sher A, Kohan A, Vercher-Conejero J, Rajiah P. PET/MRI in Lung Cancer. *Semin Roentgenol* (2014) 49(4):291–303. doi: 10.1053/j.ro.2014.07.002
11. Feng Q, Liang J, Wang L, Niu J, Ge X, Pang P, et al. Radiomics Analysis and Correlation With Metabolic Parameters in Nasopharyngeal Carcinoma Based on PET/MR Imaging. *Front Oncol* (2020) 10:1619. doi: 10.3389/fonc.2020.01619
12. Chong H, Gong Y, Pan X, Liu A, Chen L, Yang C, et al. Peritumoral Dilation Radiomics of Gadoxetate Disodium-Enhanced MRI Excellently Predicts Early Recurrence of Hepatocellular Carcinoma Without Macrovascular Invasion After Hepatectomy. *J Hepatocell Carcinoma* (2021) 8:545–63. doi: 10.2147/JHC.S309570
13. Wang W, Gu D, Wei J, Ding Y, Yang L, Zhu K, et al. A Radiomics-Based Biomarker for Cytokeratin 19 Status of Hepatocellular Carcinoma With Gadoxetic Acid-Enhanced MRI. *Eur Radiol* (2020) 30(5):3004–14. doi: 10.1007/s00330-019-06585-y
14. Ji GW, Zhu FP, Xu Q, Wang K, Wu MY, Tang WW, et al. Machine-Learning Analysis of Contrast-Enhanced CT Radiomics Predicts Recurrence of Hepatocellular Carcinoma After Resection: A Multi-Institutional Study. *EBioMedicine* (2019) 50:156–65. doi: 10.1016/j.ebiom.2019.10.057
15. Kim J, Choi SJ, Lee SH, Lee HY, Park H. Predicting Survival Using Pretreatment CT for Patients With Hepatocellular Carcinoma Treated With Transarterial Chemoembolization: Comparison of Models Using Radiomics. *AJR Am J Roentgenol* (2018) 211(5):1026–34. doi: 10.2214/AJR.18.19507
16. Wang XH, Long LH, Cui Y, Jia AY, Zhu XG, Wang HZ, et al. MRI-Based Radiomics Model for Preoperative Prediction of 5-Year Survival in Patients With Hepatocellular Carcinoma. *Br J Cancer* (2020) 122(7):978–85. doi: 10.1038/s41416-019-0706-0
17. Gu D, Xie Y, Wei J, Li W, Ye Z, Zhu Z, et al. MRI-Based Radiomics Signature: A Potential Biomarker for Identifying Glypican 3-Positive Hepatocellular Carcinoma. *J Magn Reson Imaging* (2020) 52(6):1679–87. doi: 10.1002/jmri.27199
18. Feng Q, Niu J, Wang L, Pang P, Wang M, Liao Z, et al. Comprehensive Classification Models Based on Amygdala Radiomic Features for Alzheimer's Disease and Mild Cognitive Impairment. *Brain Imaging Behav* (2021) 15(5):2377–86. doi: 10.1007/s11682-020-00434-z
19. Yang L, Yang J, Zhou X, Huang L, Zhao W, Wang T, et al. Development of a Radiomics Nomogram Based on the 2D and 3D CT Features to Predict the Survival of non-Small Cell Lung Cancer Patients. *Eur Radiol* (2019) 29(5):2196–206. doi: 10.1007/s00330-018-5770-y
20. Coroller TP, Agrawal V, Huynh E, Narayan V, Lee SW, Mak RH, et al. Radiomic-Based Pathological Response Prediction From Primary Tumors and Lymph Nodes in NSCLC. *J Thorac Oncol* (2017) 12(3):467–76. doi: 10.1016/j.jtho.2016.11.2226
21. Kirienko M, Cozzi L, Rossi A, Voulaz E, Antunovic L, Fogliata A, et al. Ability of FDG PET and CT Radiomics Features to Differentiate Between Primary and Metastatic Lung Lesions. *Eur J Nucl Med Mol Imaging* (2018) 45(10):1649–60. doi: 10.1007/s00259-018-3987-2
22. Orlhac F, Soussan M, Chouahnia K, Martinod E, Buvat I. 18f-FDG PET-Derived Textural Indices Reflect Tissue-Specific Uptake Pattern in Non-Small Cell Lung Cancer. *PloS One* (2015) 10(12):e0145063. doi: 10.1371/journal.pone.0145063
23. Yang B, Zhong J, Zhong J, Ma L, Li A, Ji H, et al. Development and Validation of a Radiomics Nomogram Based on 18F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography and Clinicopathological Factors to Predict the Survival Outcomes of Patients With Non-Small Cell Lung Cancer. *Front Oncol* (2020) 10:1042. doi: 10.3389/fonc.2020.01042
24. Szyszko TA, Cook GJR. PET/CT and PET/MRI in Head and Neck Malignancy. *Clin Radiol* (2018) 73(1):60–9. doi: 10.1016/j.crad.2017.09.001
25. Grosse J, Dirk H. PET/CT Und PET/MRT Bei Tumoren Des Kopf-Hals-Bereichs" [PET/CT and PET/MRI in Head and Neck Cancer]. *Laryngo-rhino-otologie* (2020) 99(1):12–21. doi: 10.1055/a-1057-1244

26. Hyun SH, Ahn MS, Koh YW, Lee SJ. A Machine-Learning Approach Using PET-Based Radiomics to Predict the Histological Subtypes of Lung Cancer. *Clin Nucl Med* (2019) 44(12):956–60. doi: 10.1097/RLU.0000000000002810

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

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 04 March 2022
doi: 10.3389/fonc.2022.828505

Investigating the Clinico-Molecular and Immunological Evolution of Lung Adenocarcinoma Using Pseudotime Analysis

Hyunjong Lee[1,2] *and Hongyoon Choi*[3]*

[1] Department of Nuclear Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea, [2] Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, South Korea, [3] Department of Nuclear Medicine, Seoul National University College of Medicine, Seoul, South Korea

OPEN ACCESS

Edited by:
Yiyan Liu,
University of Louisville, United States

Reviewed by:
Pankaj Kumar Ahluwalia,
Augusta University, United States
Yanye Wang,
Tianjin Medical University General Hospital, China
Qiang Xu,
GenomiCare Biotechnology Shanghai Co., Ltd, China
Ahreum Kim,
CHA University, South Korea

***Correspondence:**
Hongyoon Choi
chy1000@snu.ac.kr

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 03 December 2021
Accepted: 07 February 2022
Published: 04 March 2022

Citation:
Lee H and Choi H (2022) Investigating the Clinico-Molecular and Immunological Evolution of Lung Adenocarcinoma Using Pseudotime Analysis. Front. Oncol. 12:828505. doi: 10.3389/fonc.2022.828505

Introduction: As the molecular features of lung adenocarcinoma (LUAD) have been evaluated as a cross-sectional study, the course of tumor characteristics has not been modeled. The temporal evolution of the tumor immune microenvironment (TIME), as well as the clinico-molecular features of LUAD, could provide a precise strategy for immunotherapy and surrogate biomarkers for the course of LUAD.

Methods: A pseudotime trajectory was constructed in patients with LUAD from the Cancer Genome Atlas and non-small cell lung cancer radiogenomics datasets. Correlation analyses were performed between clinical features and pseudotime. Genes associated with pseudotime were selected, and gene ontology analysis was performed. F-18 fluorodeoxyglucose positron emission tomography images of subjects were collected, and imaging parameters, including standardized uptake value (SUV), were obtained. Correlation analyses were performed between imaging parameters and pseudotime. Correlation analyses were performed between the enrichment scores of various immune cell types and pseudotime. In addition, correlation analyses were performed between the expression of PD-L1, tumor mutation burden, and pseudotime.

Results: Pseudotime trajectories of LUAD corresponded to clinical stages. Molecular profiles related to cell division and natural killer cell activity were changed along the pseudotime. The maximal SUV of LUAD tumors showed a positive correlation with pseudotime. Type 1 helper T (Th1) cells showed a positive correlation, whereas M2 macrophages showed a negative correlation with pseudotime. PD-L1 expression showed a negative correlation, whereas tumor mutation burden showed a positive correlation with pseudotime.

Conclusion: The estimated pseudotime associated with the stage suggested that it could reflect the clinico-molecular evolution of LUAD. Specific immune cell types in the TIME as well as cell division and glucose metabolism were dynamically changed according to the progression of the pseudotime. As a molecular progression of LUAD, different cellular targets should be considered for immunotherapy.

Keywords: lung adenocarcinoma, stage, glucose metabolism, tumor immune microenvironment, pseudotime analysis

INTRODUCTION

Lung adenocarcinoma (LUAD) is the most frequent histological type among lung malignancies (1). The diagnosis and treatment of LUAD are based on initial evaluation of disease progression. TNM stage has been recognized as the most basic and critical factor to evaluate the status of disease (2). In addition, the progression of lung cancer has been assessed by histological features and imaging findings as well as staging. Among imaging modalities, ^{18}F-fluorodeoxyglucose positron emission tomography (FDG PET) is widely used to evaluate the extent of metastasis and aggressiveness of tumors (3, 4). However, there is a limitation to investigating the natural progression of tumors based on conventional diagnostic information, as pathological staging and imaging examinations are performed at the timing of initial diagnosis. Therefore, the biological and molecular progression of LUAD has hardly been modeled on a continuous scale.

The tumor immune microenvironment (TIME) plays a crucial role in tumor progression and metastasis. Because of recent broad-range indications for immune checkpoint inhibitors, tumor immune profiles in addition to staging have been suggested for predicting prognosis and considering appropriate treatment plans (5). Among immune cells, natural killer (NK) cells and T cells are known to have a role in antitumor immunity (6, 7). In addition, tumor-associated macrophages (TAMs) exert various functions in lung cancer by differentiating into different subtypes: M1 and M2 macrophages, with M1 macrophages mainly contributing to antitumor activity, and M2 macrophages mainly contributing to protumor activity (8). In the clinical aspect, characterization of the TIME is important to explore therapeutic targets and predict the response to immunotherapy (9, 10). Therefore, it is important to investigate the evolution of the TIME during disease progression. More specifically, recent trends in immunotherapy suggest a strategy according to the characteristics and the progression pattern of the TIME (11).

Pseudotime analysis, also called trajectory inference analysis, is a spotlighted method to explore changes in cell or tissue characteristics based on transcriptomic expression (12). It provides a numerical scale to reflect where a cell or tissue is in the course of disease, other than the TNM staging system. Although there have been several studies to apply pseudotime analysis in lung cancer samples, the scopes of those studies were limited to only single-cell RNA-sequencing (RNA-seq) data from small numbers of patients (13, 14). Pseudotime analysis in large numbers of subjects may provide a model to explore the course of biological progression of lung cancer.

In this study, we aimed to reveal the evolution of the TIME along with the molecular progression of LUAD. Pseudotime trajectories were estimated in the LUAD cohorts from The Cancer Genome Atlas (TCGA) and a non-small cell lung cancer (NSCLC) radiogenomics dataset. Associations between TIME cell types as well as clinico-molecular features and pseudotime were analyzed. We expected to find appropriate targets of the TIME according to the evolution of the TIME along the pseudotime.

METHODS

Pseudotime Estimation

A pseudotime trajectory was constructed based on the sum of two publicly available datasets: TCGA-LUAD and TCGA-lung squamous cell carcinoma (LUSC). The datasets were obtained using the "TCGAbiolinks" package in R (15). Legacy data of gene expression quantification were downloaded using the "GDCdownload" function. There were 600 LUAD samples and 553 LUSC samples. RNA-seq data were prepared as a matrix format and normalized by log2 transformation. Highly variable genes were selected using the "DESeq2" package in R (16, 17). First, variance and coefficients of variation for each gene expression were calculated in 1153 total samples. Subsequently, a generalized linear model was fitted to set a reference for variability of each gene expression using "glmgam.fit" function in R. The fitted curve was hypothesized as an expected distribution of estimates of variance and coefficients of variation. Chi-squared tests were performed to evaluate deviation from the fitted curve. Finally, genes showing lower p-value than 0.001 were selected as highly variable genes (HVGs). A total of 8589 genes were selected among a total of 21,022 genes. A pseudotime trajectory was generated using the "Phenopath" package in R (18). PhenoPath, an analytic tool for pseudotime, provides an ordering of gene expression measurements across individual objects. It employs a Bayesian statistics and models latent progression of gene expression (18). Among various pseudotime analysis tools, PhenoPath was the only method to be utilized in bulk tissue RNA-seq dataset. Therefore, it was selected in the present study. The input data were a gene expression matrix of HVGs from TCGA-LUAD and TCGA-LUSC datasets. We chose the evidence lower bound as 10^{-6} and computed it thinned by 2 iterations. Ultimately, pseudotime as a reference value for latent progression of disease was estimated.

Applying Pseudotime Into a New Dataset

LUAD samples of the NSCLC radiogenomics dataset were employed to perform additional correlative analysis with imaging-derived variables. An RNA-seq dataset (GSE103584) was downloaded from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) (19). There were 96 LUAD samples in the NSCLC radiogenomics dataset. To translate pseudotime estimated by the TCGA dataset to the new dataset, a lasso regression model was used to estimate pseudotime in LUAD samples of the NSCLC radiogenomics dataset. There were three reasons to apply lasso regression model. First, there was a technical limitation that 'Phenopath' tool could not generate a proper model due to small sample size of the NSCLC radiogenomics dataset. Second, it was necessary to predict pseudotime based on genes which were revealed to have significant association with pseudotime in TCGA dataset. Third, lasso regression provides better interpretability and prevents overfitting as it deals some of the estimated coefficients as zero (20). To develop a model to predict pseudotime using a lasso regression, two hundred genes were selected from genes that showed a significant association between pseudotime: the top 100 genes with highest correlation coefficients in the positive correlation group and the top 100 genes with highest correlation coefficients in the negative correlation group. An expression matrix of those 200 genes was constructed from the TCGA dataset. It was divided into two groups at a 2:1 ratio: training and internal validation data. The lambda with the least error was selected *via* cross-validation. A lasso regression model was obtained. The alpha was 1. The model was applied to LUAD samples of the NSCLC radiogenomics dataset to predict a pseudotime trajectory.

Clinical Feature Analysis

Clinical data of the TCGA dataset were downloaded from cBioPortal (http://www.cbioportal.org/) using the "cgdsr" package in R. TNM stage, disease-free survival (DFS), overall survival (OS), and duration of DFS/OS were selected as representative clinical factors. A heatmap was plotted to visualize the associations between genes, clinical factors, and pseudotime using the "Complexheatmap" package in R. The pseudotimes of each TNM stage group were compared using t-tests or ANOVA. Survival analyses for DFS and OS were performed using the Kaplan–Meier method between subjects with early and late pseudotime. Cutoff values of pseudotime were explored using the "cutoff" package in R. Clinical data of the NSCLC radiogenomics dataset were downloaded from The Cancer Imaging Archive (TCIA, https://www.cancerimagingarchive.net/). The pseudotime of each TNM stage group was compared using the Wilcoxon rank-sum test or the Kruskal–Wallis test.

Genetic Feature Analysis

Principal component analysis (PCA) was performed to visualize the temporal evolution of the genetic characteristics of LUAD and LUSC using the "PCA" function included in the "factoextra" package in R. Phenopath analysis provided four output values: *alpha*: degree of differential expression, *beta*: degree of covariate-pseudotime interaction, *lambda*: degree of pseudotime dependency, *z*: estimates of pseudotime. A Bayesian significance test was applied to select genes showing significant pseudotime dependency (nonzero lambda) and significant covariate pseudotime dependency (nonzero beta). Gene ontology (GO) analysis was conducted on genes showing significant pseudotime dependency to investigate which functions were upregulated or downregulated along the pseudotime trajectory using the "enrichGO" function included in the "clusterProfiler" package in R. A cut-off of p-value was 0.05 and that of q-value was 0.1 to select significant GO terms.

Glucose Metabolism Analysis

FDG PET images of LUAD subjects from both the TCGA dataset and the NSCLC radiogenomics dataset were downloaded from The Cancer Imaging Archive (TCIA, https://www.cancerimagingarchive.net/). There were 16 and 93 samples with both RNA-seq data and FDG PET images in LUAD samples of the TCGA and NSCLC radiogenomics datasets, respectively. Tumor margins were delineated using the Nestle adaptive threshold method provided by "LifeX" software (21, 22). An adaptive threshold to define tumor margins was applied, and the deterministic parameter *beta* was set to 0.3. The maximal standardized uptake value (SUV), mean SUV, and metabolic tumor volume (MTV) were obtained from the region of interest. Total lesion glycolysis (TLG) was calculated from the mean SUV and MTV. Correlation coefficients of expression of FDG PET parameters, solute carrier family 2 member 1 (SLC2A1) expression, and pseudotime were calculated by Spearman and Pearson correlation tests in the TCGA and NSCLC radiogenomics datasets, respectively.

Immune Profile Analysis

In both the TCGA and NSCLC radiogenomics datasets, the enrichment scores of 64 immune and stromal cell types were estimated using the "xCellAnalysis" function in the "xCell" package in R (23). Correlation coefficients between enrichment scores and pseudotime were calculated by the Pearson correlation test. The false discovery rate was calculated from p values with the Bonferroni method. Volcano plots, heatmaps, and scatter plots were generated to describe the association between the enrichment scores of immune cells and pseudotime. The expression of PD-L1 and the tumor mutation burden (TMB) are well-known indicators of the immune profiles of tumors (24, 25). In the TCGA dataset, gene mutation data were downloaded from genomic data commons (https://gdc.cancer.gov/), and a mutation annotation format file was then constructed using the "read.maf" function included in the "maftools" package (26). TMB was calculated by the number of non-synonymous somatic mutations using the "tmb" function included in the "maftools" package. Correlation coefficients of expression of PD-L1 and TMB with pseudotime were calculated by the Pearson correlation test. All statistical analyses were performed using R software (v4.0.4, Vienna, Austria). A p value of < 0.05 was considered statistically significant.

RESULTS

Clinical Features Related to Pseudotime

Demographic and clinical characteristics of the patients were described in **Table 1**. A heatmap was constructed to visualize the clinical factors of each sample with the top 10 genes upregulated and those downregulated over pseudotime in TCGA-LUAD samples (**Figure 1**). Notably, histone coding genes showed upregulation along pseudotime. Boxplots represent the association between TNM stage and pseudotime in LUAD samples (**Figures 2A–D**). There was a significant difference in pseudotime in each T stage ($p < 0.001$), especially in T1-T2 (mean: -0.08013 vs. 0.03092, $p < 0.001$) and T1-T3 (mean: -0.08013 vs. 0.08310, $p < 0.001$). Pseudotime in different N stages and M stages showed no difference. There was a significant difference in pseudotime in each overall TNM stage ($p < 0.001$), especially in I-II (mean: -0.02835 vs. 0.03565, $p = 0.021$) and I-III (mean: -0.02835 vs. 0.04780, $p = 0.019$). Disease-free survival and overall survival were well discriminated according to pseudotime (**Figure 3**, $p = 0.002$ and $p < 0.001$, respectively).

Pseudotime was estimated in the NSCLC radiogenomics dataset based on a lasso regression model from the TCGA dataset. The association between clinical factors and pseudotime was evaluated in LUAD samples of the NSCLC radiogenomics dataset. There was a tendency of increasing T stage along pseudotime, especially in early T stages (**Supplementary Figure 2A**, $p = 0.097$). There was no association between N/M stage and pseudotime (**Supplementary Figures 2B, C**). Histological grade showed an association with pseudotime (**Supplementary Figure 2D**, $p = 0.017$). There was no significant association between overall TNM stage and pseudotime (**Supplementary Figure 2E**).

TABLE 1 | Demographic and clinical characteristics of the patients.

Characteristics	TCGA-LUAD Patients, n (%)	LUAD in the NSCLC radiogenomics dataset Patients, n (%)
Total	600	96
Sex		
Female	325 (54.2)	29 (30.2)
Male	275 (45.8)	67 (69.8)
Age, median (range), years	66 (33-88)	68 (43-85)
Smoking history		
Current	131 (21.8)	19 (19.8)
Former	364 (60.7)	57 (59.4)
Never	85 (14.2)	20 (20.8)
Unknown	20 (3.3)	0
Location		
Right upper lobe	220 (36.7)	31 (32.3)
Right middle lobe	24 (4)	8 (8.3)
Right lower lobe	106 (17.7)	12 (12.5)
Left upper lobe	144 (24)	30 (31.3)
Left lower lobe	88 (14.7)	15 (15.6)
Unknown	18 (3)	0
Pathological T stage		
Tis	0	5 (5.2)
T1	195 (32.5)	38 (39.6)
T2	331 (55.2)	38 (39.6)
T3	51 (8.5)	11 (11.5)
T4	20 (3.3)	4 (4.2)
Unknown	3 (0.5)	0
Pathological N stage		
N0	381 (63.5)	78 (81.3)
N1	110 (18.3)	7 (7.3)
N2	87 (14.5)	11 (11.5)
N3	2 (0.3)	0
Unknown	20 (3.3)	0
Pathological M stage		
M0	407 (67.8)	91 (94.8)
M1	27 (4.5)	5 (5.2)
Unknown	166 (27.7)	0
Pathological stage		
0	0	5 (5.2)
I	322 (53.7)	57 (59.4)
II	138 (23)	16 (16.7)
III	97 (16.2)	13 (13.5)
IV	28 (4.7)	5 (5.2)
Unknown	15 (2.5)	0

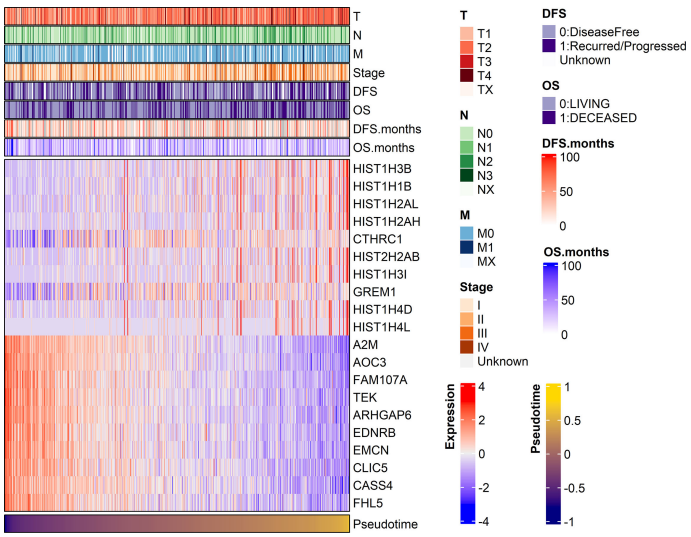


FIGURE 1 | A heatmap visualizing clinical features and gene expression along pseudotime in the TCGA-LUAD dataset. Clinical features, including TNM stage and the expression of the top 10 genes showing significant positive or negative associations with pseudotime, are displayed. In particular, histone-coding genes showed significant upregulation over pseudotime.

Genetic and Functional Features Related to Pseudotime

We investigated gene expression features and their functional relevance according to pseudotime. On the dimension reduction plot using PCA, the LUAD and LUSC samples seemed to be in the same position at the beginning of pseudotime (**Supplementary Figure 3**). As pseudotime passed, the LUAD and LUSC samples were clearly discriminated in the PCA plot. We investigated genes regulated over pseudotime in total lung cancer, LUAD, and LUSC samples (**Figure 4A**). A total of 603 genes showed significantly positive correlations with pseudotime in total lung cancer samples, 2594 genes showed negative correlations in total cancer samples, 178 genes showed positive correlations in LUAD samples, 853 genes showed negative correlations in LUAD samples, 479 genes showed positive correlations in LUSC samples, and 647 genes showed negative correlations in LUSC samples. GO analysis was performed to determine which biological pathways were related (**Figures 4B, C**). In total lung cancer samples, molecular functions related to cell division were upregulated over pseudotime. In LUAD samples, molecular functions related to cell division, such as nucleosome assembly and DNA packaging, were upregulated over pseudotime, as in total lung cancer samples. Those related to NK cell function are downregulated over pseudotime.

Evolution of Glucose Metabolism Along Pseudotime

As increased glucose metabolism measured by FDG PET is associated with poor prognosis by reflecting biological aggressiveness, we tested the association between FDG uptake and pseudotime (27). In the TCGA-LUAD dataset, there was a significant positive correlation between maximal SUV and pseudotime (**Supplementary Figure 4**, rho = 0.518, p = 0.042). There was also a significant positive correlation between mean SUV and pseudotime (**Supplementary Figure 4**, rho = 0.517, p = 0.049). However, MTV and TLG showed no association with pseudotime. In the NSCLC radiogenomics dataset, there was a significant positive correlation between maximal SUV and pseudotime (**Figure 5A**, r = 0.259, p = 0.005). There was also a significant positive correlation between mean SUV and pseudotime (**Figure 5B**, r = 0.227, p = 0.029). However, MTV and TLG showed no association with pseudotime.

Evolution of Immune Profiles Along Pseudotime

As GO terms related to pseudotime included downregulation of immune-related functions according to pseudotime, we further analyzed tumor immune microenvironment profiles related to pseudotime. A volcano plot and a heatmap were constructed to illustrate the immune and stromal cells associated with pseudotime in the TCGA dataset. (**Figures 6A, B**). Among cell types, the enrichment score of type 1 helper T (Th1) cells showed a positive correlation (**Figure 7A**, r = 0.524, p < 0.001), and that of M2 macrophages showed a negative correlation (**Figure 7B**, r = -0.545, p < 0.001). PD-L1, the most representative immunotherapy target in lung cancer, showed a weakly negative correlation with pseudotime (**Figure 7C**, r = -0.289, p < 0.001). TMB showed a weakly positive correlation with pseudotime (**Figure 7D**, r = 0.243, p < 0.001). In the NSCLC radiogenomics dataset, Th1 cells showed a positive correlation (**Supplementary Figure 5A**, r = 0.444, p < 0.001), and M2 macrophages showed a negative correlation (**Supplementary Figure 5B**, r = -0.367, p = 0.020). PD-L1 showed no significant correlation with pseudotime (**Supplementary Figure 5C**, r = 0.041, p = 0.698).

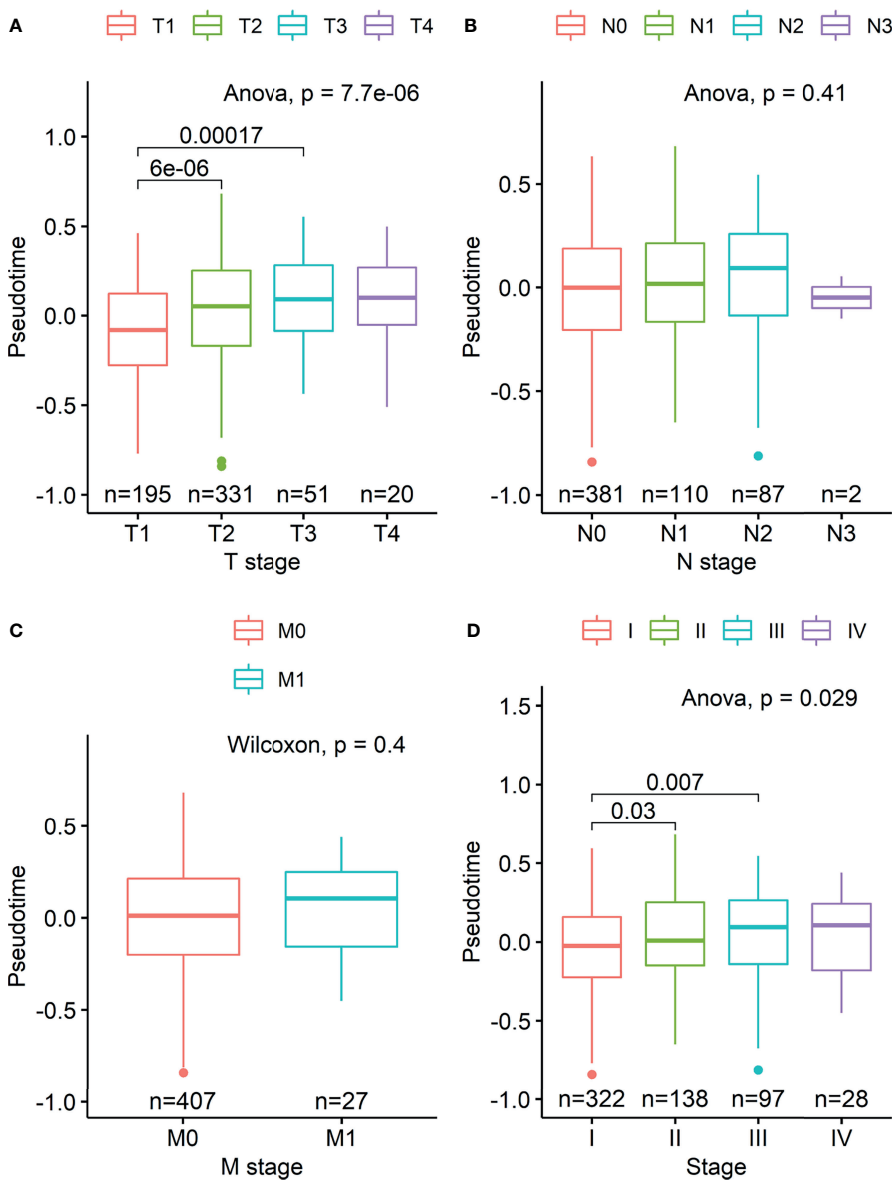


FIGURE 2 | Boxplots visualizing differences in pseudotime according to TNM stage in the TCGA-LUAD dataset. **(A)** There was a significant difference in pseudotime in each T stage, especially in T1-T2 (mean: -0.08013 vs. 0.03092, $p < 0.001$) and T1-T3 (mean: -0.08013 vs. 0.08310, $p < 0.001$). **(B, C)** Pseudotime in each N and M stage showed no difference. **(D)** There was a significant difference in pseudotime in each overall TNM stage, especially in IA-IB (mean: -0.09326 vs. 0.02969, $p < 0.001$), IA-IIB (mean: -0.09326 vs. 0.05374, $p < 0.001$), and IA-IIIA (mean: -0.09326 vs. 0.03749, $p < 0.001$).

DISCUSSION

TNM staging in lung cancer is a well-established system to evaluate disease progression status, predict prognosis, and select appropriate treatment options (28, 29). However, it is the result of cross-sectional observation *via* clinical/pathologic/radiologic findings at the timing of initial diagnosis. Therefore, there is a limitation to investigating the temporal evolution of tumor biology longitudinally based on TNM staging as a reference scale. We attempted to construct a temporal model for biological progression and change of TIME from genetic profiles of a bulk RNA-seq dataset using pseudotime analysis. For scRNA-seq dataset, pseudotime analysis orders cells along a hypothetic trajectory based on patterns of gene expression (30). It is based on the hypothesis that multiple cross-sectional data is integrated into sequential data demonstrating temporal evolution (31). As bulk RNA-seq can be used for the trajectory analysis with a same manner, Campbell and Yau uncovered temporal evolution of tumor tissues of colorectal cancer and breast cancer (18). In this study, continuous scale for temporal evolution of tumor tissue was generated using previously known analytic tool, PhenoPath (18). The validity of the generated

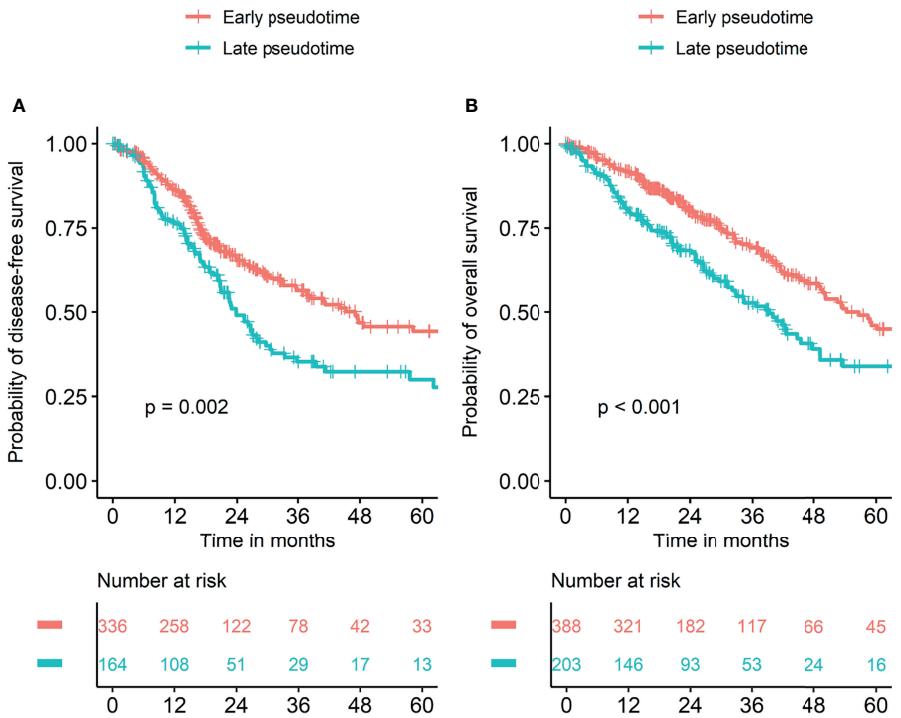


FIGURE 3 | Survival curves according to pseudotime. Survival analyses were performed in two groups divided by cutoff values of pseudotime which were explored using the "cutoff" package in R. **(A)** Disease-free survival was significantly different between the two groups. **(B)** Overall survival was significantly different between the two groups.

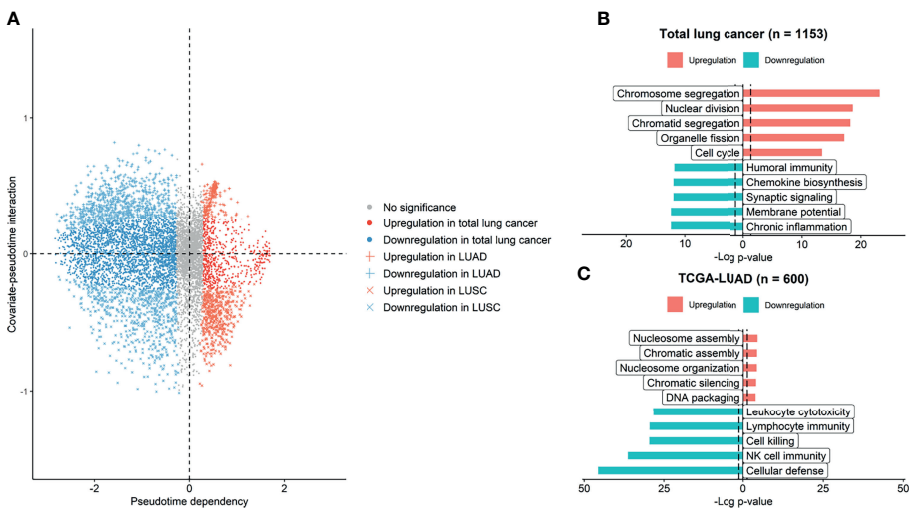


FIGURE 4 | Genes and molecular functions upregulated and downregulated over pseudotime. **(A)** Genes upregulated and downregulated over pseudotime were selected and plotted. A total of 3197 genes showed significant correlations with pseudotime in total lung cancer. A total of 1031 genes showed significant correlations in LUAD samples, and 1126 genes showed significant correlations in LUSC samples. **(B)** In total lung cancer samples, molecular functions related to cell division were upregulated over pseudotime. **(C)** In LUAD samples, those related to natural killer cell activity were downregulated over pseudotime, and those related to cell division were upregulated.

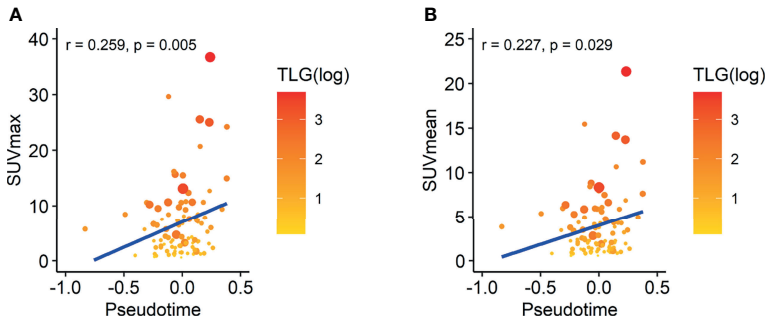


FIGURE 5 | Correlation between SUV and pseudotime in LUAD samples of the NSCLC radiogenomics dataset. The size of the dot represents the metabolic tumor volume. The color of the dot represents total lesion glycolysis as a log scale. **(A)** The maximal SUV showed a weakly positive correlation with pseudotime (r = 0.259, p = 0.005). **(B)** The mean SUV showed a weakly positive correlation with pseudotime (r = 0.227, p = 0.029). However, MTV and TLG showed no association with pseudotime.

model as a reference of disease progression was examined by correlation analyses between clinico-molecular features and pseudotime. Based on the generated model, we investigated the temporal evolution of the tumor immune microenvironment in lung adenocarcinoma.

In this study, we successfully estimated a pseudotime trajectory in TCGA-LUAD and TCGA-LUSC datasets. In PCA, LUAD and LUSC samples in the early phase are revealed to have similar genetic characteristics and to differentiate into LUAD and LUSC along pseudotime order. In the tumorigenesis

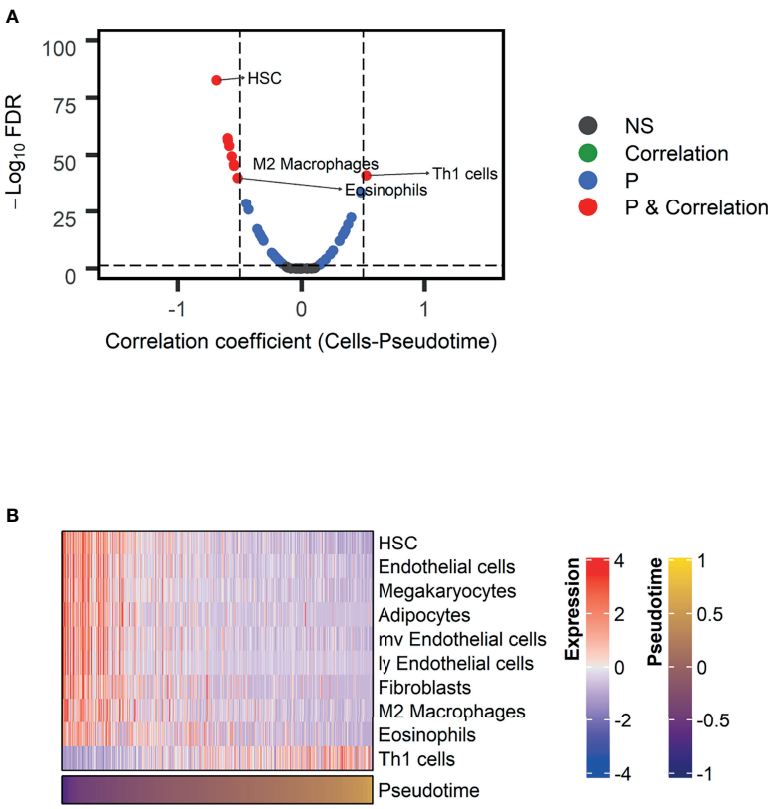


FIGURE 6 | Immune and stromal cells associated with pseudotime in the TCGA-LUAD dataset. **(A)** A volcano plot representing immune cells associated with pseudotime is shown. Cell types with correlation coefficients above 0.5 are plotted as red dots. Cell types with FDR below 0.05 and correlation coefficients below 0.5 are plotted as blue dots. Among them, immune cells were annotated. **(B)** A heatmap representing immune and stromal cells associated with pseudotime is shown.

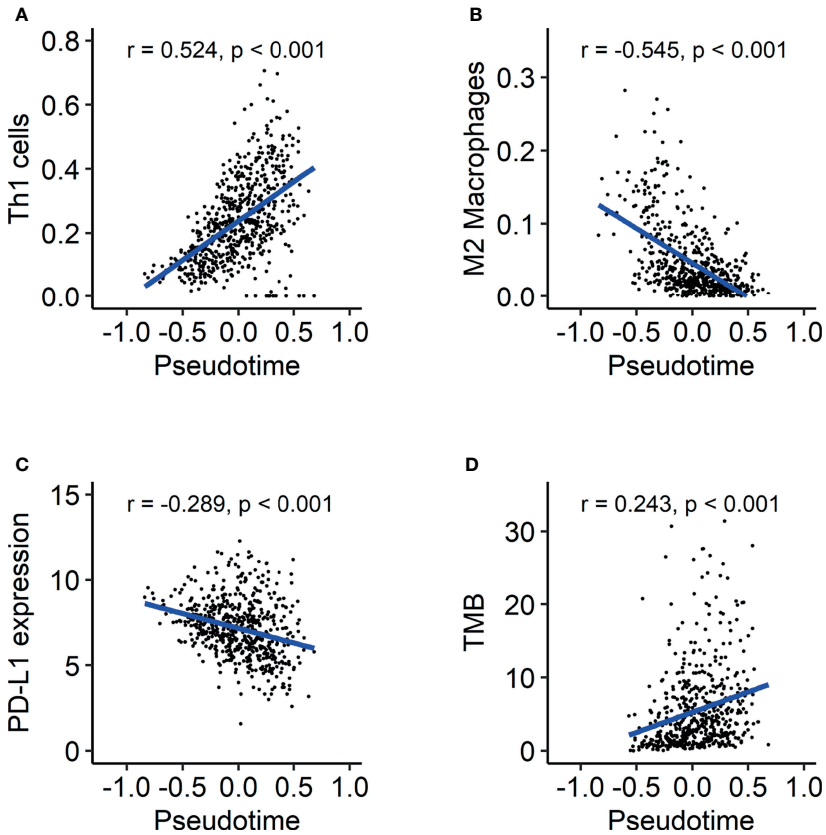


FIGURE 7 | Evolution of the tumor immune microenvironment over pseudotime in the TCGA-LUAD dataset. **(A)** Th1 cells showed a positive correlation with pseudotime ($r = 0.524$, $p < 0.001$). **(B)** M2 macrophages showed a negative correlation with pseudotime ($r = -0.545$, $p < 0.001$). **(C)** PD-L1 expression showed a weakly negative correlation with pseudotime ($r = -0.289$, $p < 0.001$). **(D)** TMB showed a weakly positive correlation with pseudotime ($r = 0.243$, $p < 0.001$).

of NSCLC, molecular events such as 3p allele loss and telomerase activation are observed in most NSCLCs (32–34). The similarity of genetic characteristics in early LUAD and LUSC may be caused by common pathogenesis mechanisms. This result implies that tumors showing specific characteristics of LUAD or LUSC have a high possibility of progressed disease.

In both total lung cancer and LUAD samples, GO terms related to cell division were selected as significantly upregulated terms over pseudotime. In particular, histone-related genes showed a high correlation with pseudotime in LUAD. These results can be interpreted as the result of either the presence of a large number of tumor cells or enhanced mitotic activity of tumor cells in the late phase. Similarly, T stage and overall TNM stage demonstrated good association with pseudotime. This is consistent with the current TNM staging system. Notably, there were significant associations in the early T stages (T1-T2, T2-T3) and early overall stages (IA-IB, IA-IIB, IA-IIIA). In the current TNM staging system, T2-T4 stages include not only size criteria but also criteria involving other structures, such as bronchi or chest walls (35). Thus, a small tumor with involvement of other structures can be diagnosed as a high T stage. If there is lymph node metastasis, it is highly likely to be classified as above stage IIIA. These characteristics of the current TNM staging system cause associations between early T stage/overall stage and pseudotime. The probability of disease-free survival and overall survival showed significant differences according to pseudotime. This indicates that pseudotime may have clinical usability to classify patients based on prognosis as TNM staging. Of course, further study is warranted to explore the clinical significance of pseudotime trajectories.

In both datasets, the maximal SUV and mean SUV demonstrated significantly positive correlations with pseudotime. These findings are consistent with a previously revealed relationship between FDG uptake and tumor stage (3, 4, 36). Furthermore, the tendency of increasing FDG uptake along pseudotime is closely related to the molecular function of cell division, showing the same tendency along pseudotime. Proliferative activity is known to be a significant factor affecting FDG uptake in tumors in lung cancer (37, 38). In brief, changes in genetic features, clinical features, and glucose metabolism over pseudotime were revealed to be consistent with previous knowledge about tumor progression. Therefore, the estimated pseudotime was hypothesized to be an appropriate temporal reference of disease progression.

The present study demonstrated the temporal evolution of immune profiles in LUAD. It is noteworthy that the enrichment

score of Th1 cells represented a significantly positive correlation with pseudotime. It is generally believed that Th1 cells contribute to the antitumor response, inducing cytotoxicity (7, 39). It is also remarkable that M2 macrophages showed a significantly negative correlation with pseudotime. M2 macrophages exert protumor activity *via* tissue remodeling and angiogenesis (40, 41). Briefly, antitumor immunity seems to strengthen along pseudotime, whereas protumor immunity seems to weaken along pseudotime. These results indicate two possible mechanisms. First, protumor immunity *via* M2 macrophages prepares an appropriate environment for the survival and proliferation of tumor cells in the early phase of lung cancer. Second, antitumor immunity *via* Th1 cells is induced by a high tumor burden to attempt to control and suppress disease in the late phase of lung cancer. These results are also consistent with a previous report documenting that the proportion of high stage was larger in samples with high immune scores and cytolytic scores (42). The present study showed heterogeneous evolution of specific immune cells over pseudotime. This implies that selecting immunotherapy options for appropriate targets may be considered during the disease progression of LUAD. Pharmaceuticals such as resveratrol and imatinib were found to suppress cancer progression *via* inhibition of M2 macrophage activation (43, 44). The application of those drugs may be utilized in patients with early pseudotime.

Interestingly, PD-L1 expression showed a negative correlation with pseudotime, whereas TMB showed a positive correlation in the TCGA dataset. These are well-known biomarkers predicting the response to cancer immunotherapy (25, 45). This implies that the response to immunotherapy, such as pembrolizumab, may not represent any tendency according to the molecular progression of lung adenocarcinoma. Referring to this heterogeneous finding of opposite tendencies of PD-L1 and TMB, it is necessary to establish a more precise immunotherapy strategy. Considering that both the enrichment of Th1 cells and TMB showed positive correlations with pseudotime, it is supposed that immunogenic antigens are enriched in progressed LUAD. LUAD with a relatively early pseudotime associated with high PD-L1 suggested early anti-PD-1/PD-L1 treatment before tumor evolution (46). As LUAD with late pseudotime showed high TMB and low PD-L1 expression, immune checkpoint inhibitors targeting molecules other than PD-L1 can be proposed as an appropriate immunotherapy option for tumors in progressed LUAD. Our model of the temporal evolution of the TIME and biomarkers related to immunotherapy suggested that a more precise strategy of immunotherapy could be needed according to the biological progression of lung cancer. In this respect, further study can be planned for selecting appropriate immunotherapy regimens and evaluating treatment response using pseudotime concept. Based on estimated pseudotime from transcriptomics, patients can be classified into those with early and late phase of disease. Application of different immunotherapy regimen and evaluation prognosis can be performed in each group. The further study is expected to expand usefulness and clinical significance of pseudotime in LUAD patients.

This study has clinical implication and benefits as followings. We attempted to generate simple and continuous scale of disease progression in LUAD. The genomic landscape of LUAD has been investigated in many previous studies and varying genetic characteristics were known to associate with prognosis of LUAD. It is needed to suggest more simplified and available value integrating diverse genetic information. In recent clinical filed, there are a few approaches to provide genetic profile information to LUAD patients with microarray or RNA sequencing. However, it is too complicated and difficult to deliver its clinical implication to patients. It is expected that simplified scale of disease progression is helpful to communicate with patients for discussing disease progression status, further treatment plan, and prognosis. Of course, further study is needed to construct a reference model from a larger cohort to validate and utilize pseudotime.

There are some limitations in this study. First, FDG PET examination of subjects was performed in different institutes so that there were differences in image acquisition and reconstruction methods. However, the purpose of analyzing the association between SUV and pseudotime was not to predict accurate SUV or pseudotime but to assess the overall tendency of SUV along pseudotime. Furthermore, the image acquisition protocol of each sample was not identified in the obtained clinical data. Therefore, all the data were included in a single correlation study. Further study is warranted to analyze the evolution of glucose metabolism over pseudotime more accurately using FDG PET image data from the same institute. Second, pseudotime trajectory from RNA-seq has a limitation to apply to the clinical field due to the complexity of obtaining tumor tissue and analyzing transcriptomic data from each patient. To facilitate the application of pseudotime in clinical situations, further study is underway to construct pseudotime trajectories from FDG PET images. Third, estimation of immune cell infiltration using deconvolutional method for transcriptomic data has a limitation to exactly estimate immune cell fraction of tumor tissue. In contrast, it has an advantage that quantification for large-scale dataset is available. Further study is warranted to validate actual TIME of tumor tissue with experimental methods such as immunofluorescence.

Taken together, pseudotime trajectories were successfully estimated in lung adenocarcinoma subjects from the TCGA dataset and the NSCLC radiogenomics dataset. These results show fair correlations with TNM stage, clinical outcome, and glucose metabolism, suggesting the feasibility of a new scale evaluating disease progression status. There were heterogeneous findings in the evolution of tumor immune microenvironment components over pseudotime. The present study suggested that individualized immunotherapy strategies should be selected according to different molecular characteristics evolving during disease progression.

DATA AVAILABILITY STATEMENT

RNA-sequencing data from The Cancer Genome Atlas datasets are available in National Cancer Insititute Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov/). Clinical data of The Cancer Genome Atlas datasets are available in cBioPortal (https://www.cbioportal.org/). RNA-sequencing data from non-small cell lung carcinoma radiogenomics dataset

(GSE103584) are available in Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/).

AUTHOR CONTRIBUTIONS

HL and HC designed the study, performed analysis and interpretation, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Research Foundation of Korea (NRF) and funded by the Korean government (MSIT) (No.2020M3A9B6038086, NRF-2019R1F1A1061412). This research was also funded by the Korea Medical Device Development Fund grant funded by the Korean government (Ministry of Science and ICT, Ministry of Trade, Industry and Energy, Ministry of Health & Welfare, Ministry of Food and Drug Safety) (Project Number: 202011A06).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.828505/full#supplementary-material

REFERENCES

1. Barta JA, Powell CA, Wisnivesky JP. Global Epidemiology of Lung Cancer. *Ann Glob Health* (2019) 85(1):8. doi: 10.5334/aogh.2419
2. Mountain CF, Carr DT, Anderson W. A System for the Clinical Staging of Lung Cancer. *Am J Roentgenol Radium Ther Nucl Med* (1974) 120(1):130–8. doi: 10.2214/ajr.120.1.130
3. Sunnetcioglu A, Arısoy A, Demir Y, Ekin S, Dogan E. Associations Between the Standardized Uptake Value of 18F-FDG PET/CT and Demographic, Clinical, Pathological, Radiological Factors in Lung Cancer. *Int J Clin Exp Med* (2015) 8(9):15794–800.
4. Li M, Sun Y, Liu Y, Han A, Zhao S, Ma L, et al. Relationship Between Primary Lesion FDG Uptake and Clinical Stage at PET–CT for Non-Small Cell Lung Cancer Patients: An Observation. *Lung Cancer* (2010) 68(3):394–7. doi: 10.1016/j.lungcan.2009.07.009
5. Bruni D, Angell HK, Galon J. The Immune Contexture and Immunoscore in Cancer Prognosis and Therapeutic Efficacy. *Nat Rev Cancer* (2020) 20 (11):662–80. doi: 10.1038/s41568-020-0285-7
6. Takanami I, Takeuchi K, Giga M. The Prognostic Value of Natural Killer Cell Infiltration in Resected Pulmonary Adenocarcinoma. *J Thorac Cardiovasc Surg* (2001) 121(6):1058–63. doi: 10.1067/mtc.2001.113026
7. Nishimura T, Nakui M, Sato M, Iwakabe K, Kitamura H, Sekimoto M, et al. The Critical Role of Th1-Dominant Immunity in Tumor Immunology. *Cancer Chemother Pharmacol* (2000) 46(Suppl):S52–61. doi: 10.1007/pl00014051
8. Quatromoni JG, Eruslanov E. Tumor-Associated Macrophages: Function, Phenotype, and Link to Prognosis in Human Lung Cancer. *Am J Transl Res* (2012) 4(4):376–89.
9. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the Tumor Immune Microenvironment (TIME) for Effective Therapy. *Nat Med* (2018) 24(5):541–50. doi: 10.1038/s41591-018-0014-x
10. Taube JM, Galon J, Sholl LM, Rodig SJ, Cottrell TR, Giraldo NA, et al. Implications of the Tumor Immune Microenvironment for Staging and Therapeutics. *Mod Pathol* (2018) 31(2):214–34. doi: 10.1038/modpathol.2017.156
11. Vitale I, Sistigu A, Manic G, Rudqvist N-P, Trajanoski Z, Galluzzi L. Mutational and Antigenic Landscape in Tumor Progression and Cancer Immunotherapy. *Trends Cell Biol* (2019) 29(5):396–416. doi: 10.1016/j.tcb.2019.01.003
12. Trapnell C. Defining Cell Types and States With Single-Cell Genomics. *Genome Res* (2015) 25(10):1491–8. doi: 10.1101/gr.190595.115
13. Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, et al. Single-Cell RNA Sequencing Demonstrates the Molecular and Cellular Reprogramming of Metastatic Lung Adenocarcinoma. *Nat Commun* (2020) 11(1):2285. doi: 10.1038/s41467-020-16164-1
14. Song Q, Hawkins GA, Wudel L, Chou PC, Forbes E, Pullikuth AK, et al. Dissecting Intratumoral Myeloid Cell Plasticity by Single Cell RNA-Seq. *Cancer Med* (2019) 8(6):3072–85. doi: 10.1002/cam4.2113
15. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. Tcgabiolinks: An R/Bioconductor Package for Integrative Analysis of TCGA Data. *Nucleic Acids Res* (2016) 44(8):e71. doi: 10.1093/nar/gkv1507
16. Love MI, Huber W, Anders S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data With Deseq2. *Genome Biol* (2014) 15(12):1–21. doi: 10.1186/s13059-014-0550-8
17. Department of Biomedical Informatics, Harvard Medical School. *Identifying Highly Variable Genes"* (2014). Available at: http://pklab.med.harvard.edu/scw2014/subpop_tutorial.html (Accessed Dec 20, 2020).
18. Campbell KR, Yau C. Uncovering Pseudotemporal Trajectories With Covariates From Single Cell and Bulk Expression Data. *Nat Commun* (2018) 9(1):2442. doi: 10.1038/s41467-018-04696-6
19. Bakr S, Gevaert O, Echegaray S, Ayers K, Zhou M, Shafiq M, et al. A Radiogenomic Dataset of non-Small Cell Lung Cancer. *Sci Data* (2018) 5:180202. doi: 10.1038/sdata.2018.202
20. Bellazzi R, Ferrazzi F, Sacchi L. Predictive Data Mining in Clinical Medicine: A Focus on Selected Methods and Applications. *Wiley Interdiscip Rev Data Min Knowl Discov* (2011) 1(5):416–30. doi: 10.1002/widm.23
21. Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rübe C, et al. Comparison of Different Methods for Delineation of 18F-FDG PET–Positive Tissue for Target Volume Definition in Radiotherapy of Patients With non–Small Cell Lung Cancer. *J Nucl Med* (2005) 46(8):1342–8.
22. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. Lifex: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res* (2018) 78(16):4786–9. doi: 10.1158/0008-5472.CAN-18-0125
23. Aran D, Hu Z, Butte AJ. Xcell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape. *Genome Biol* (2017) 18(1):220. doi: 10.1186/s13059-017-1349-1
24. Auslander N, Zhang G, Lee JS, Frederick DT, Miao B, Moll T, et al. Robust Prediction of Response to Immune Checkpoint Blockade Therapy in Metastatic Melanoma. *Nat Med* (2018) 24(10):1545–9. doi: 10.1038/s41591-018-0157-9
25. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of Tumor Mutation Burden as an Immunotherapy Biomarker: Utility for the Oncology Clinic. *Ann Oncol* (2019) 30(1):44–56. doi: 10.1093/annonc/mdy495
26. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118
27. Higashi K, Ueda Y, Ayabe K, Sakurai A, Seki H, Nambu Y, et al. FDG PET in the Evaluation of the Aggressiveness of Pulmonary Adenocarcinoma: Correlation With Histopathological Features. *Nucl Med Commun* (2000) 21 (8):707–14. doi: 10.1097/00006231-200008000-00002
28. Chansky K, Sculier J-P, Crowley JJ, Giroux D, Van Meerbeeck J, Goldstraw P. The International Association for the Study of Lung Cancer Staging Project: Prognostic Factors and Pathologic TNM Stage in Surgically Managed Non-Small Cell Lung Cancer. *J Thorac Oncol* (2009) 4(7):792–801. doi: 10.1097/JTO.0b013e3181a7716e

29. Collins LG, Haines C, Perkel R, Enck RE. Lung Cancer: Diagnosis and Management. *Am Fam Physician* (2007) 75(1):56–63.
30. Cannoodt R, Saelens W, Saeys Y. Computational Methods for Trajectory Inference From Single-Cell Transcriptomics. *Eur J Immunol* (2016) 46 (11):2496–506. doi: 10.1002/eji.201646347
31. Saelens W, Cannoodt R, Todorov H, Saeys Y. A Comparison of Single-Cell Trajectory Inference Methods. *Nat Biotechnol* (2019) 37(5):547–54. doi: 10.1038/s41587-019-0071-9
32. Fong KM, Sekido Y, Minna JD. Molecular Pathogenesis of Lung Cancer. *J Thorac Cardiovasc Surg* (1999) 118(6):1136–52. doi: 10.1016/S0022-5223 (99)70121-2
33. Lantuéjoul S, Salameire D, Salon C, Brambilla E. Pulmonary Preneoplasia–Sequential Molecular Carcinogenetic Events. *Histopathology* (2009) 54(1):43–54. doi: 10.1111/j.1365-2559.2008.03182.x
34. Travis WD. Pathology of Lung Cancer. *Clin Chest Med* (2011) 32(4):669–92. doi: 10.1016/j.ccm.2011.08.005
35. Brierley JD, Gospodarowicz MK, Wittekind C. *TNM Classification of Malignant Tumours*. Hoboken, NJ, USA: John Wiley & Sons (2017).
36. de Geus-Oei L-F, van Krieken JHJ, Aliredjo RP, Krabbe PF, Frielink C, Verhagen AF, et al. Biological Correlates of FDG Uptake in non-Small Cell Lung Cancer. *Lung Cancer* (2007) 55(1):79–87. doi: 10.1016/j.lungcan.2006.08.018
37. Higashi K, Ueda Y, Yagishita M, Arisaka Y, Sakurai A, Oguchi M, et al. FDG PET Measurement of the Proliferative Potential of non-Small Cell Lung Cancer. *J Nucl Med* (2000) 41(1):85–92.
38. Vesselle H, Schmidt RA, Pugsley JM, Li M, Kohlmyer SG, Vallières E, et al. Lung Cancer Proliferation Correlates With [F-18] Fluorodeoxyglucose Uptake by Positron Emission Tomography. *Clin Cancer Res* (2000) 6(10):3837–44.
39. Becker JC, Andersen MH, Schrama D, thor Straten P. Immune-Suppressive Properties of the Tumor Microenvironment. *Cancer Immunol Immunother* (2013) 62(7):1137–48. doi: 10.1007/s00262-013-1434-6
40. Ruffell B, Affara NI, Coussens LM. Differential Macrophage Programming in the Tumor Microenvironment. *Trends Immunol* (2012) 33(3):119–26. doi: 10.1016/j.it.2011.12.001
41. Chanmee T, Ontong P, Konno K, Itano N. Tumor-Associated Macrophages as Major Players in the Tumor Microenvironment. *Cancers (Basel)* (2014) 6 (3):1670–90. doi: 10.3390/cancers6031670
42. Seo J-S, Kim A, Shin J-Y, Kim YT. Comprehensive Analysis of the Tumor Immune Micro-Environment in non-Small Cell Lung Cancer for Efficacy of Checkpoint Inhibitor. *Sci Rep* (2018) 8(1):14576. doi: 10.1038/s41598-018-32855-8
43. Sun L, Chen B, Jiang R, Li J, Wang B. Resveratrol Inhibits Lung Cancer Growth by Suppressing M2-Like Polarization of Tumor Associated Macrophages. *Cell Immunol* (2017) 311:86–93. doi: 10.1016/j.cellimm.2016.11.002
44. Yao Z, Zhang J, Zhang B, Liang G, Chen X, Yao F, et al. Imatinib Prevents Lung Cancer Metastasis by Inhibiting M2-Like Polarization of Macrophages. *Pharmacol Res* (2018) 133:121–31. doi: 10.1016/j.phrs.2018.05.002
45. Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol Cancer Ther* (2015) 14(4):847–56. doi: 10.1158/1535-7163.MCT-14-0983
46. Vansteenkiste J, Wauters E, Reymen B, Ackermann CJ, Peters S, De Ruysscher D. Current Status of Immune Checkpoint Inhibition in Early-Stage NSCLC. *Ann Oncol* (2019) 30(8):1244–53. doi: 10.1093/annonc/mdz175

Conflict of Interest: HC is a co-founder and CTO of Portrai, Inc. and scientific advisor of AitheNutrigene Inc.

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

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 29 April 2022
doi: 10.3389/fonc.2022.889293

Development of a Nomogram Based on 3D CT Radiomics Signature to Predict the Mutation Status of EGFR Molecular Subtypes in Lung Adenocarcinoma: A Multicenter Study

Guojin Zhang[1,2†], Liangna Deng[3†], Jing Zhang[4†], Yuntai Cao[5†], Shenglin Li[3], Jialiang Ren[6], Rong Qian[1,2], Shengkun Peng[1,2], Xiaodi Zhang[7], Junlin Zhou[3], Zhuoli Zhang[8], Weifang Kong[1,2*] and Hong Pu[1,2*]*

[1] Department of Radiology, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China, [2] Department of Radiology, Chinese Academy of Sciences Sichuan Translational Medicine Research Hospital, Chengdu, China, [3] Department of Radiology, Lanzhou University Second Hospital, Lanzhou, China, [4] Department of Radiology, Fifth Affiliated Hospital of Zunyi Medical University, Zhuhai, China, [5] Department of Radiology, Affiliated Hospital of Qinghai University, Xining, China, [6] Department of Pharmaceuticals Diagnosis, GE Healthcare, Beijing, China, [7] Clinical Science Department, Philips (China) Investment Co., Ltd., Chengdu, China, [8] Department of Radiology and BME, University of California Irvine, Irvine, CA, United States

Background: This study aimed to noninvasively predict the mutation status of epidermal growth factor receptor (EGFR) molecular subtype in lung adenocarcinoma based on CT radiomics features.

Methods: In total, 728 patients with lung adenocarcinoma were included, and divided into three groups according to EGFR mutation subtypes. 1727 radiomics features were extracted from the three-dimensional images of each patient. Wilcoxon test, least absolute shrinkage and selection operator regression, and multiple logistic regression were used for feature selection. ROC curve was used to evaluate the predictive performance of the model. Nomogram was constructed by combining radiomics features and clinical risk factors. Calibration curve was used to evaluate the goodness of fit of the model. Decision curve analysis was used to evaluate the clinical applicability of the model.

Results: There were three, two, and one clinical factor and fourteen, thirteen, and four radiomics features, respectively, which were significantly related to each EGFR molecular subtype. Compared with the clinical and radiomics models, the combined model had the highest predictive performance in predicting EGFR molecular subtypes [Del-19

Edited by:
Yiyan Liu,
University of Louisville, United States

Reviewed by:
Wenjuan Ma,
Tianjin Medical University Cancer Institute and Hospital, China
Cheng Dong,
The Affiliated Hospital of Qingdao University, China

***Correspondence:**
Guojin Zhang
zhanggj15@lzu.edu.cn
Hong Pu
pqs19951211@126.com
Weifang Kong
13936034@qq.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *04 March 2022*
Accepted: *04 April 2022*
Published: *29 April 2022*

Citation:
Zhang G, Deng L, Zhang J, Cao Y, Li S, Ren J, Qian R, Peng S, Zhang X, Zhou J, Zhang Z, Kong W and Pu H (2022) Development of a Nomogram Based on 3D CT Radiomics Signature to Predict the Mutation Status of EGFR Molecular Subtypes in Lung Adenocarcinoma: A Multicenter Study. Front. Oncol. 12:889293. doi: 10.3389/fonc.2022.889293

mutation *vs.* wild-type, AUC=0.838 (95% CI, 0.799-0.877); L858R mutation *vs.* wild-type, AUC=0.855 (95% CI, 0.817-0.894); and Del-19 mutation *vs.* L858R mutation, AUC=0.906 (95% CI, 0.869-0.943), respectively], and it has a stable performance in the validation set [AUC was 0.813 (95% CI, 0.740-0.886), 0.852 (95% CI, 0.790-0.913), and 0.875 (95% CI, 0.781-0.929), respectively].

Conclusion: Our combined model showed good performance in predicting EGFR molecular subtypes in patients with lung adenocarcinoma. This model can be applied to patients with lung adenocarcinoma.

Keywords: NSCLC, lung adenocarcinoma, EGFR, computed tomography, radiomics

INTRODUCTION

Targeted therapy has brought based on recognizing the importance of acquired gene driver mutations, such as epidermal growth factor receptor (EGFR) mutations, kristen rat sarcoma (KRAS) mutations and anaplastic lymphoma kinase (ALK) rearrangements, in non-small cell lung cancer (NSCLC) new hope to patients with these gene mutations. In the Asian population, about 50% of lung adenocarcinoma patients have known carcinogenic driver genes (1, 2). There are currently targeted drugs used in clinical practice for these mutations, such as gefitinib and osimertinib for EGFR mutations. In contrast, patients without these mutations are not candidates for targeted therapy (3). Furthermore, there are molecular differences between each molecular mutation and molecular subtype, and these differences lead to different therapeutic effects after using other targeted drugs (4). EGFR mutations mainly include exon 18-21 mutations. Among them, exon 19 deletion (Del-19) mutation and 21 L858R point (L858R) mutation are the two most common activating mutations, and they are also the two most sensitive mutation sites for tyrosine kinase inhibitors (TKI) treatment (5). In a single targeted therapy, patients with Del-19 mutation benefited more from osimertinib (6), while patients with L858R mutation benefited significantly from dacomitinib (7); in addition, combination therapy and immunotherapy brought patients with L858R for more potential benefits (8, 9). Therefore, the detection of specific EGFR mutation subtypes can make targeted therapies more precise and allow patients receiving these treatments to benefit the most.

Currently, the detection of EGFR mutation status from histological specimens is the most common detection method. However, in clinical practice, these detection techniques also have some limitations. For example, tissue samples are obtained through invasive methods such as biopsy or surgery; sometimes the amount of tissue samples obtained due to operational errors is insufficient; biopsy can increase the risk of tumor metastasis; in addition, a small part of the tissue obtained does not represent the heterogeneity of the entire tumor, etc. (10–12). In addition, another noninvasive detection strategy for EGFR mutations is 'liquid biopsy', which is a biological detection method on the blood. For patients with advanced NSCLC, 'liquid biopsy' is a promising method to isolate circulating tumor DNA from blood samples (13). However, 'liquid biopsy' has a high risk of false-negative results (30%) (14). Therefore, until this defect is effectively resolved, 'liquid biopsy' is far from substitute for histological testing. Because of this, there is an urgent need for a simple and noninvasive method to detect EGFR mutation subtypes before targeted drug therapy.

The radiological features have been shown to reflect EGFR mutation status in lung adenocarcinoma (12, 15–17). However, the clinical applicability of these studies needs to be confirmed by further research. Compared with traditional CT, radiomics converts medical images into mineable data and extracts a large number of features that cannot be observed by the human naked eye system, thereby reflecting more characteristics of tumors (18). To our knowledge, some studies have used radiomics to predict EGFR mutation status (19–22). Although the prediction performance of these studies is different, this shows that it is feasible to predict EGFR mutations noninvasively through radiomics. However, only a few studies have used radiomics methods to predict the mutation status of EGFR molecular subtypes (23–26). Unfortunately, the sample size included in these studies is limited, and the accuracy of the obtained prediction model was only 65.5-79.0%. In this study, we retrospectively collected a relatively large data set and constructed a model based on CT radiomics signature to noninvasively predict the mutation status of EGFR molecular subtype in lung adenocarcinoma.

MATERIALS AND METHODS

Patient Population

This retrospective study was ethically approved by the Institutional Review Board of the Sichuan Provincial People's Hospital and Lanzhou University Second Hospital, and the need for patient informed consent was waived. Clinical data and chest CT images of these patients were obtained from the picture archiving and communication system (PACS). The inclusion criteria were as follows: (1) patients with the histologic type of lung adenocarcinoma; (2) patients with complete CT thin-slice images (1.25 mm) and clinical data; (3) patients who did not receive lung cancer-related treatment before CT scan; (4) patients who underwent biopsy or surgery within one month

after CT scan; (5) patients with EGFR exon Del-19 mutation, exon L858R mutation, and wild-type. The exclusion criteria were as follows: (1) patients whose tumor boundary is difficult to be recognized by the naked eye on CT images; (2) patients younger than 18 years old.

According to the above inclusion and exclusion criteria, 728 patients (median age, 57.0 years, age range, 21-82 years, 370 males and 358 females) were finally selected from 2,557 patients in the two hospitals. Among them, a total of 540 patients from Sichuan Provincial People's Hospital were used as the training set from January 2018 to March 2021, and 188 patients from Lanzhou University Second Hospital were identified as the external validation set from January 2019 to September 2020. The patient recruitment flowchart is shown in **Figure 1**.

Demographic and clinical data include the patient's sex, age, smoking history [including non-smoking (never smoked) and smoking (former or current smoking)], carcinoembryonic antigen (CEA) level, and tumor lobe location of the tumor (including right upper, right middle, right lower, left upper and left lower lobes). If the tumor crosses the fissure, the lobe location is defined as the lobe in which the tumor predominates.

EGFR Mutation Status Detection

The polymerase chain reaction-amplified refractory mutation system (PCR-ARMS) detected EGFR mutation status. The human EGFR gene detection kit (Beijing SinoMD Gene Detection Technology Co., Ltd., China; Amoy Diagnostics, Xiamen, China) detected EGFR exon 18 to 21 mutation status.

CT Image Acquisition

CT scans ranged from the thoracic inlet to the level of the lower edge of the 12th rib were completed by three spiral CT scanners (Discovery CT750 HD, GE Healthcare; Philips iCT 256, Koninklijke Philips N.V.; Somatom Sensation 64, Siemens Healthineers). Scanning parameters were as follows: (1) tube voltage 120 kVp, tube current adjusted automatically for the Sensation 64 scanner, and (2) tube voltage 120 kVp, tube current 150 to 200 mA for the other two scanners. For all scanners, 0.5-1.0 second tube rotation time, and field of view (FOV): 350 mm; matrix, 512 × 512; the layer thickness and spacing were both 5 mm; the reconstruction layer thickness and spacing were both 1.25 mm. All images were exported in DICOM format to facilitate feature extraction.

Tumor Segmentation and Radiomics Feature Extraction

Radiomics feature extraction and analysis workflow are shown in **Figure 2**. To ensure the accuracy and consistency of the data, two readers (radiologists with 6 and 4 years of experience in chest CT diagnosis, respectively) independently used the open-source software ITK-SNAP 3.8.0 (http://www.itksnap.org) to segment the tumor on the thin-slice CT lung window (window width: 1500HU; window level: -500HU). Training cases were segmented by reader 1 (G.J.Z), and validation cases were segmented by reader 2 (L.N.D). Both readers were blinded to all patients' clinical data, pathological records, and EGFR status results. When the two readers were unsure, a consultant radiologist (J.Z) confirmed the segmentation with 17 years of experience. The region of interest (ROI) was manually segmented on CT axial images with tumor tissue and confirmed on sagittal and coronal images.

In order to evaluate the robustness and repeatability of the radiomics feature extraction process, one month later, 40 patients were randomly selected from the training set and segmented again by readers 1 and 2 to construct a re-segmentation set, and 40 patients were randomly selected from each CT scanner to construct different CT scanner sets for calculating the intra-/interclass correlation coefficients (ICC), respectively. ICC values > 0.8 reflected good consistency (26).

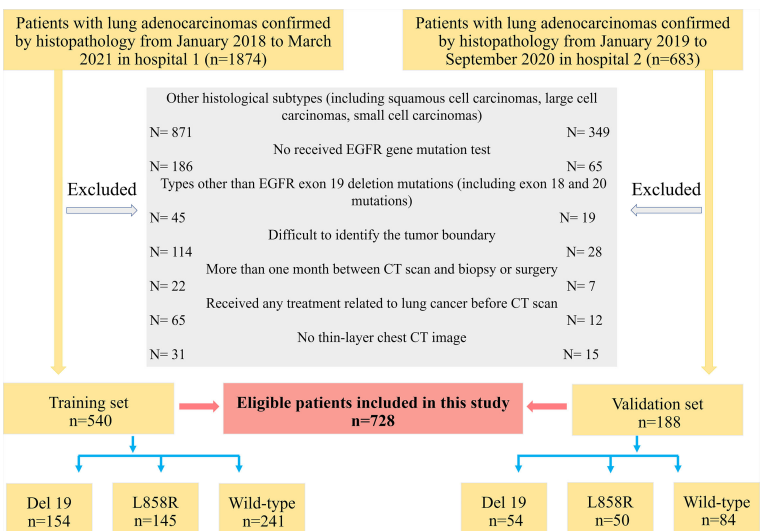


FIGURE 1 | The flowchart of the inclusion and exclusion criteria.

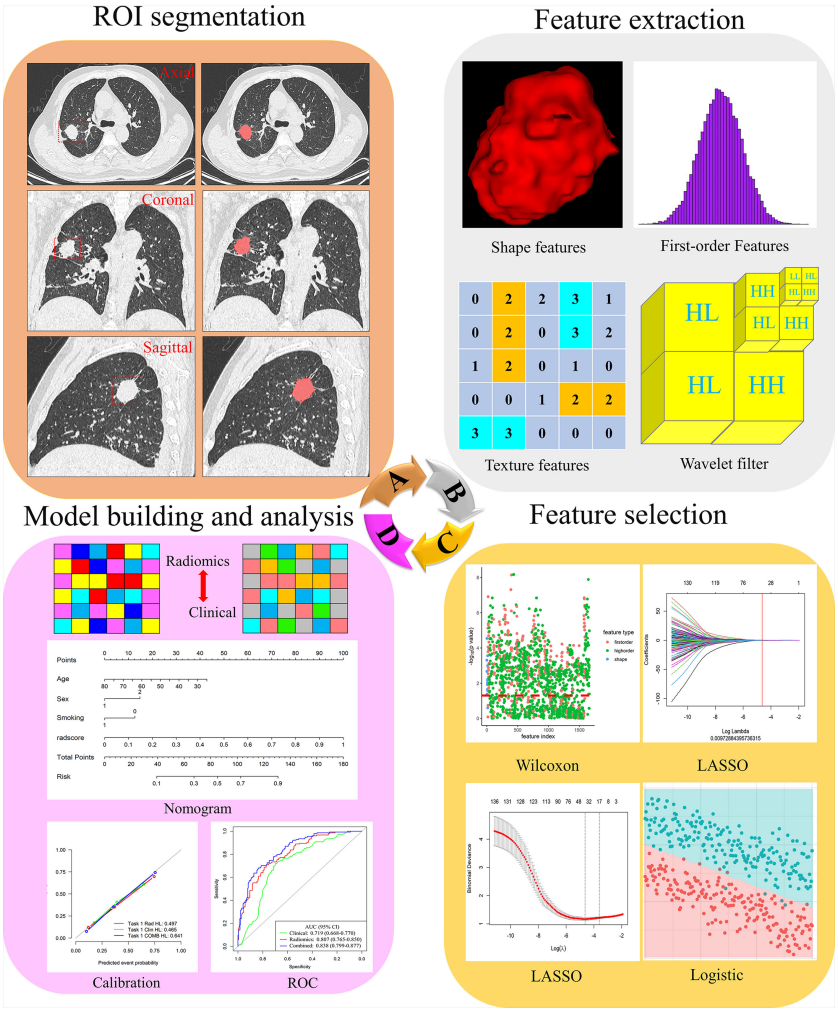


FIGURE 2 | Flowchart of the process of radiomics. **(A)** The tumours were segmented on CT images to form the region of interest (ROI). **(B)** Radiomics feature extraction from the ROI. **(C)** Radiomics feature dimensionality reduction process. **(D)** Construct a radiomics model.

The open-source Python software package PyRadiomics 3.0.1 automatically extracted radiomics features from the three-dimensional (3D) tumor volume after segmentation. Radiomics features were divided into three main categories: 16 shape features, 324 first-order features and 1387 texture features. Details of radiomics features were included in the **Supplementary Material** (Methods). According to the recommendations of International Symposium on Biomedical Imaging (ISBI), we only resampled the image and set the bin width of gray discretization to 25. We performed z-score preprocessing on the extracted radiomics features.

Radiomics Feature Selection

To avoid model overfitting and improve accuracy, we used three steps for feature selection to obtain the optimal feature subset. First, Wilcoxon test was used to retain the features with *P*-value less than 0.05. Secondly, the least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation was used to eliminate the collinearity features. LASSO is a recognized algorithm that has been used for feature selection of higher-dimensional variables (27). Finally, multiple logistic regression was used to select the features, and the minimum features of akaike information criterion (AIC) were retained.

For clinical factors, the Chi-square and Student's *t*-tests were first used in the training set to screen for clinical characteristics that were correlated between each group. *P*-value was set to 0.05. Clinical factors with a *P*-value greater than 0.05 were excluded. Next, use logistic regression to further analyze the most relevant variables.

Radiomics Model Establishment

Logistic regression was used in the training set to build a model for predicting Del-19 or L858R mutations, and its performance

was evaluated in the external validation set. Logistic regression is a classic method in radiomics research. It is easy to understand, explain, and combine discrete and continuous variables (28, 29). To identify Del 19 and L858R mutations, we used logistic regression, support vector machine (SVM), and random forest (RF) to construct prediction models. The clinical and radiomics models were constructed based on clinical factors and radiomics features, respectively, while the combined model was constructed based on clinical and radiomics models. Additionally, clinical models were constructed using logistic regression.

Statistical analysis

All statistical analyses were performed using R 3.6.0 (http://www.R-project.org). Two-sided *P*-values less than 0.05 were considered to be statistically different. Kolmogorov-Smirnov test was used to evaluate the normal distribution of the data. Categorical variables were expressed as percentiles, and the Chi-square test or Fisher's exact test was used to analyze groups. Continuous variables were expressed as mean ± standard deviation (SD), and Student's *t*-test or Mann-Whitney U test was used for analysis between groups. Receiver operating characteristic (ROC) curve was used to evaluate the performance of the model, and the area under the curve (AUC), sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) were calculated. Delong test was used to compare the performance differences of the prediction models.

Based on the above-screened clinical factors and radiomics features, we constructed a personalized nomogram to predict the mutation status of the EGFR molecular subtype. Calibration curve and Hosmer-Lemeshow (H-L) test were used to evaluate the model's goodness of fit. Decision curve analysis (DCA) was used to assess the clinical applicability of the model.

RESULTS

Clinical Characteristics of Patients

There were no significant differences in clinical factors (including sex, smoking history, and CEA), the mutation rate of Del-19 or L858R, and tumor location in each EGFR mutant subtype group (all $P > 0.05$), while there were significant differences in age between the two EGFR mutant subtype groups (Del-19 *vs.* wild type, Del-19 *vs.* L858R) (**Supplementary Tables S1–S3**).

Univariate analysis revealed that age, sex and smoking history were significantly different between the Del-19 mutation and wild-type groups ($P < 0.05$), sex and smoking history were significantly different between the L858R mutation and wild-type groups ($P < 0.05$), and age was significantly different between the Del-19 mutation and L858R mutation groups ($P < 0.05$). Multivariate analysis revealed that age (OR, 0.972; 95% CI, 0.948-0.996; $P = 0.021$) and sex (OR, 3.193; 95% CI, 1.836-5.565; $P < 0.001$) were correlated independently with the task of Del-19 *vs.* wild-type (**Table 1**), sex (OR, 2.612; 95% CI, 1.548-4.457; $P < 0.001$) and smoking history (OR, 0.427; 95% CI, 0.238-0.761; $P = 0.001$) were correlated independently with the task of L858R *vs.* wild-type (**Table 2**), and age (OR, 1.050; 95% CI, 1.022-1.081; $P < 0.001$) was correlated independently with the task of Del-19 *vs.* L858R (**Table 3**). Based on multivariate analysis results, clinical factors with $P < 0.05$ in each task were incorporated in the clinical model.

Radiomics Feature Selection and Model Establishment

In total, 1727 radiomics features were extracted from the 3D images of each ROI. The ICC values of the radiomics features extracted from two readers and different CT scanners were all greater than 0.80, reflecting good consistency. Fourteen radiomics features were highly correlated with Del-19 mutation (**Table S4; Figure S1**), thirteen radiomics features were highly correlated with L858R mutation (**Table S4; Figure S2**). For Del-19 mutation *vs.* L858R mutation, only four radiomics features were screened after using the Wilcoxon test (**Table S4; Figure S3**). Therefore, we retained these four features to construct the prediction model.

Correlation analysis showed that the correlation between each feature is weak and independent in the training and validation sets (**Figures S4–S6**).

Based on the above-screened radiomics features and clinical factors, the clinical, radiomics, and combined models were established in the training set, respectively, to predict the EGFR molecular subtype mutation status.

Predictive performance and Validation Based on Clinical, Radiomics, and Combined Models

The predictive performance of different models in the training and validation sets is shown in **Figure 3** and **Table 4**. The predictive performance of the combined model was higher than that of other single models. In the training set, the AUC of the combined model was 0.838 (95% CI, 0.799-0.877), 0.855 (95% CI, 0.817-0.894), and 0.906 (95% CI, 0.869-0.943), respectively. In addition, we used an external validation set to verify the accuracy of the combined model, and the AUC was 0.813 (95% CI, 0.740-0.886), 0.852 (95% CI, 0.790-0.913), and 0.875 (95% CI, 0.781-0.929), respectively. In addition, when distinguishing between Del-19 and L858R mutations, the prediction model's performance constructed using random forest was higher than that of other single models. The AUC of the training and validation sets were 0.881 (95% CI, 0.840-0.921) and 0.871 (95% CI, 0.802-0.941), respectively.

Delong test showed that there were significant differences in AUC values of the three models in the training set between EGFR Del-19 mutation or L858R mutation and wild-type groups (all $P < 0.05$); However, only the AUC value of combined model and clinical model was significantly different in the validation set ($P < 0.05$), and the AUC values between other models were not statistically significant ($P > 0.05$) (**Figures S7A–D**). There were significant differences in AUC values of the combined model and clinical, SVM or Logistic models in the training and validation sets between EGFR Del-19 mutation and L858R mutation groups (all $P < 0.05$). However, the AUC value between combined model and RF model was not statistically significant in the both sets ($P > 0.05$) (**Figures S7E, F**).

TABLE 1 | The relationship between clinical variables of patients and EGFR molecular subtypes (Del-19 mutation *vs.* Wild-type) in the training set.

Variable	Total (n = 395)	Del-19 mutation (n =154)	Wild-type (n = 241)	`Univariate analysis	Multivariate analysis	
				P value	OR (95%CI)	*P* value
Age (years)				<0.001	0.972 (0.948-0.996)	0.021
- Mean ± SD	56.70 ± 9.19	54.87 ± 8.13	58.87 ± 9.65			
- Median (Q_1, Q_3)	56.0 (50.0,63.0)	55.0(49.0,61.0)	58.0(51.0,65.0)			
- Range	26-79	32-78	26-79			
Sex (%)				<0.001		<0.001
- Male	221 (55.9%)	51 (33.1%)	170 (70.5%)		Reference	
- Female	174 (44.1%)	103 (66.9%)	71 (29.5%)		3.193 (1.836-5.656)	
Smoking history (%)				<0.001	NA	
- No	240 (60.8%)	122 (79.2%)	118 (49.0%)			
- Yes	155 (39.2%)	32 (20.8%)	123 (51.0%)			
CEA (%)				0.391	NA	
- Normal	167 (42.3%)	61 (39.6%)	106 (44.0%)			
- High	228 (57.7%)	93 (60.4%)	135 (56.0%)			
Lobe location (%)				0.959	NA	
- Right upper lobe	135 (34.2%)	54 (35.1%)	81 (33.6%)			
- Right middle lobe	17 (4.3%)	6 (3.9%)	11 (4.6%)			
- Right lower lobe	97 (24.6%)	35 (22.7%)	62 (25.7%)			
- Left upper lobe	81 (20.5%)	33 (21.4%)	48 (19.9%)			
- Left lower lobe	65 (16.5%)	26 (16.9%)	39 (16.2%)			

CEA, Carcinoembryonic antigen; CI, Confidence interval; Del 19, Exon-19 deletion mutation; EGFR, Epidermal growth factor receptor; NA, not applicable; OR, Odds ratio; SD, Standard deviation. vs., versus.

Clinical Application of the Combined Model

Based on radiomics score and clinical risk factors, we constructed two user-friendly nomograms to predict the mutation status of EGFR molecular subtypes (**Figures 4A** and **5A**). The detailed formula for calculating the radiomics score is shown in the **Supplementary Material** (Result). The calibration curve analysis showed that the probability of Del-19 mutation or L858R mutation predicted by the combined model was highly consistent with the actual possibility, indicating that the model had the best discriminant ability (**Figures 4B, C**, and **5B, C**).

TABLE 2 | The relationship between clinical variables of patients and EGFR molecular subtypes (L858R mutation *vs.* Wild-type) in the training set.

Variable	Total (n = 386)	L858R mutation (n = 145)	Wild-type (n = 241)	Univariate analysis	Multivariate analysis	
				P value	OR (95%CI)	*P* value
Age (years)				0.686	NA	
- Mean ± SD	58.12 ± 9.46	58.50 ± 9.15	58.87 ± 9.65			
- Median (Q_1, Q_3)	58.0 (52.0,65.0)	58.0(53.0,64.0)	58.0(51.0,65.0)			
- Range	21-82	21-82	26-79			
Sex (%)				<0.001		<0.001
- Male	223 (57.8%)	53 (36.6%)	170 (70.5%)		Reference	
- Female	174 (44.2%)	92 (63.4%)	71 (29.5%)		2.612 (1.548-4.457)	
Smoking history (%)				<0.001		0.004
- No	234 (60.6%)	116 (80.0%)	118 (49.0%)		Reference	
- Yes	152 (39.4%)	29 (20.0%)	123 (51.0%)		0.427 (0.238-0.761)	
CEA (%)				0.301	NA	
- Normal	162 (42.0%)	56 (38.6%)	106 (44.0%)			
- High	224 (58.0%)	89 (61.4%)	135 (56.0%)			
Lobe location (%)				0.262	NA	
- Right upper lobe	124 (32.1%)	43 (29.7%)	81 (33.6%)			
- Right middle lobe	24 (6.2%)	13 (9.0%)	11 (4.6%)			
- Right lower lobe	99 (25.6%)	37 (25.5%)	62 (25.7%)			
- Left upper lobe	83 (21.5%)	35 (24.1%)	48 (19.9%)			
- Left lower lobe	56 (14.5%)	17 (11.7%)	39 (16.2%)			

CEA, Carcinoembryonic antigen; CI, Confidence interval; EGFR, Epidermal growth factor receptor; L858R, Exon-21 L858R point mutation; NA, not applicable; OR, Odds ratio; SD, Standard deviation. vs., versus.

TABLE 3 | The relationship between clinical variables of patients and EGFR molecular subtypes (Del-19 mutation *vs.* L858R mutation) in the training set.

Variable	Total (n = 299)	Del-19 mutation (n = 154)	L858R mutation (n = 145)	Univariate analysis	Multivariate analysis	
				P value	OR (95%CI)	*P* value
Age (years)				<0.001	1.050 (1.022-1.081)	<0.001
- Mean ± SD	56.63 ± 8.81	54.87 ± 8.13	58.50 ± 9.15			
- Median (Q_1, Q_3)	56.0 (50.5,62.5)	55.0(49.0,61.0)	58.0(53.0,64.0)			
- Range	21-82	32-78	21-82			
Sex (%)				0.533	NA	
- Male	104 (34.8%)	51 (33.1%)	53 (36.6%)			
- Female	195 (65.2%)	103 (66.9%)	92 (63.4%)			
Smoking history (%)				0.867	NA	
- No	238 (79.6%)	122 (79.2%)	116 (80.0%)			
- Yes	61 (20.4%)	32 (20.8%)	29 (20.0%)			
CEA (%)				0.861	NA	
- Normal	117 (39.1%)	61 (39.6%)	56 (38.6%)			
- High	182 (60.9%)	93 (60.4%)	89 (61.4%)			
Lobe location (%)				0.235	NA	
- Right upper lobe	97 (32.4%)	54 (35.1%)	43 (29.7%)			
- Right middle lobe	19 (6.4%)	6 (3.9%)	13 (9.0%)			
- Right lower lobe	72 (24.1%)	35 (22.7%)	37 (25.5%)			
- Left upper lobe	68 (22.7%)	33 (21.4%)	35 (24.1%)			
- Left lower lobe	43 (14.4%)	26 (16.9%)	17 (11.7%)			

CEA, Carcinoembryonic antigen; CI, Confidence interval; Del 19, Exon-19 deletion; EGFR, Epidermal growth factor receptor; L858R, Exon-21 L858R point mutation; NA, not applicable; OR, Odds ratio; SD, Standard deviation. vs., versus.

Decision curve analysis showed that the combined model threshold in range of 0.18-0.77 have higher net benefit for Del-19 *vs.* wild type and the cutoff value was 0.440 fall in this rang; the combined model threshold in range of 0.16-0.715 have higher net benefit for L858R *vs.* wild type and the cutoff value was 0.389 fall in this range (**Figures 4D, E** and **5D, E**).

The precision-recall curves showed that the combined model constructed by the RF model combined with clinical factors had better performance than other single models in predicting Del-19 and L858R mutations (**Figures 6A, B**).

DISCUSSION

Preoperative noninvasive prediction of EGFR mutant subtypes is a new field that attracts researchers' attention. It can well overcome some shortcomings of molecular mutation detection based on histology and provide critical information for the rational formulation of targeted therapy in clinical practice. This study established different models based on clinical factors and radiomics features to predict EGFR mutation subtypes. Among them, the combined model showed good predictive performance in the training set. It also had good stability when evaluating the model's performance in the external validation set, which reveals the feasibility of predicting EGFR molecular subtypes through radiomics features.

Previous studies have found that some clinical variables such as female, non-smokers, patients with histological type of lung adenocarcinoma, and East Asian populations are significantly associated with EGFR mutations (16, 17, 20, 30, 31). Our previous research has also confirmed this (12, 15). However, these studies did not further analyze the correlation between EGFR mutation subtypes and clinical variables. In our study, sex, smoking history, and age were significantly different in the EGFR mutation subtypes group. Compared with EGFR wild-type patients, Del-19 mutation patients were more common in females and non-smokers, and L858R mutation patients were more common in females and non-smokers. Compared with patients with Del-19 mutation, patients with L858R mutation were relatively older. Only patients with Del-19 mutation and L858R mutation were selected because they are the most common mutations in EGFR mutation subtypes. The two mutation sites most related to the sensitivity of EGFR TKI treatment.

Some previous studies have predicted the mutation status of EGFR molecular subtypes based on CT radiomics features and achieved promising results. For example, Li and colleagues (26) retrospectively collected 312 patients with NSCLC, and 580 radiomics features were extracted from each patient's CT images to construct a model to predict EGFR mutation subtypes (Del 19 and L858R). The test set's AUC for predicting Del 19 and L858R mutations were 79.3% and 77.5%, respectively. Similarly, Zhao et al. (25) included 637 patients with lung adenocarcinoma in their study to predict EGFR mutation subtypes, and extracted 475 radiomics features to construct a model. The results showed that the AUC in the training and validation datasets were 68.9% and 75.7%, respectively. However, these studies did not distinguish between the Del-19 mutation and the L858R mutation, limiting the clinical applicability of these studies. In this study, we not only distinguished between EGFR Del-19 mutation or L858R mutation and EGFR wild-type. More importantly, we further distinguished the Del-19 mutation and the L858R mutation, and achieved good prediction performance. The training and validation sets' AUC was 90.6% and 87.5%, respectively. Therefore, our research may be more in line with actual clinical needs.

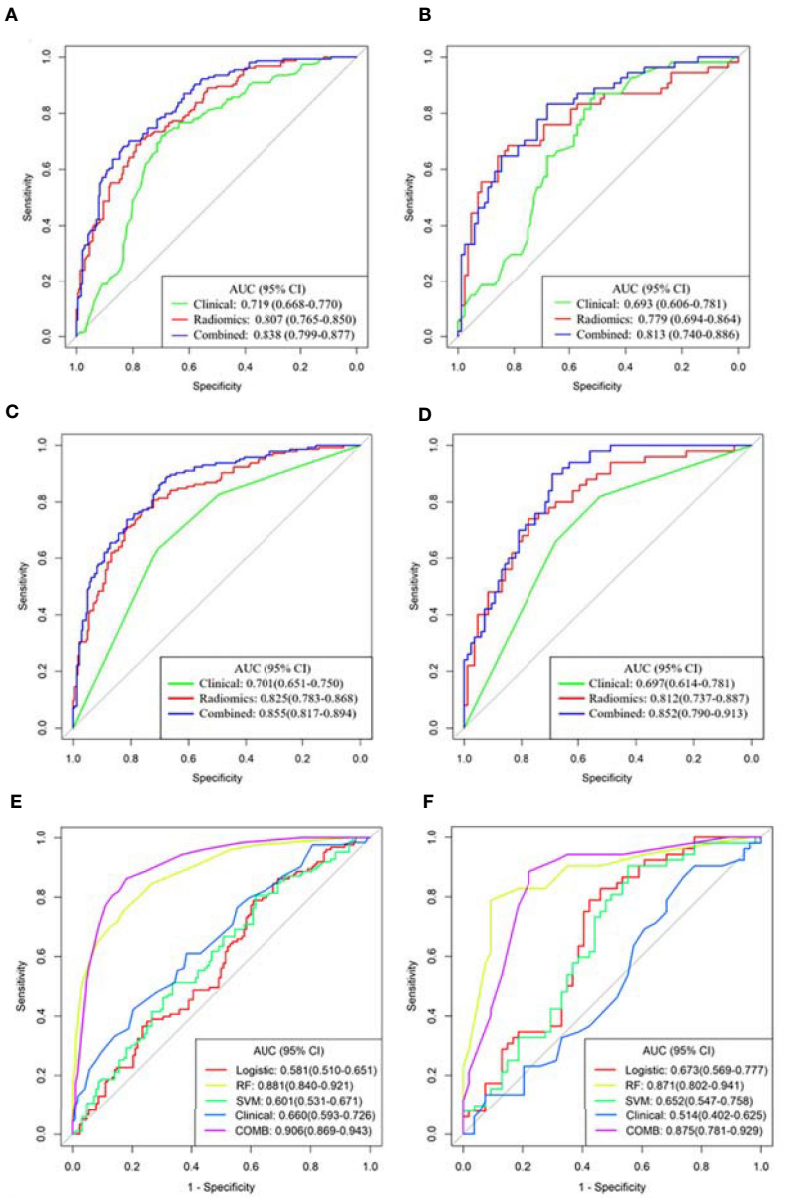


FIGURE 3 | Receiver operating characteristic (ROC) curves of the three models were used to predict the mutant status of EGFR molecular subtypes. **(A, B)** Del-19 mutation *vs.* wild-type. **(C, D)** L858R mutation *vs.* wild-type. **(E, F)** Del-19 mutation *vs.* L858R mutation. **(A, C, E)** Training set. **(B, D, F)** Validation set.

In this study, whether in the training or validation sets, the combined model established by clinical factors combined with radiomics features can improve the diagnostic performance of identifying EGFR molecular subtypes. Liu et al. (24) included 263 patients with lung adenocarcinoma in their study to detect EGFR mutation status and its molecular subtypes. Among the 6 models established, the combined model had better distinguishing ability than the model that only uses radiomics features or clinical factors. Tu et al. (20) included 404 NSCLC patients in their study to predict EGFR mutation status, and the comprehensive model showed higher predictive performance than any other single model. Jia et al. (32) also showed that a comprehensive model with radiomics features combined with clinical factors had better diagnostic performance than a single model. It shows that adding clinical factors to the radiomics model can improve the diagnostic performance of the model.

In recent years, the study of radiomics in predicting tumor gene mutations has attracted extensive attention from researchers (20, 22, 32, 33). The intrinsic relationship between the radiomics features and EGFR mutation status in patients with lung adenocarcinoma can be further explored through data mining to guide clinical decision-making, predict prognosis and evaluate

TABLE 4 | The prediction performance of different models in the training and validation sets.

Models	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV
Del-19 mutation *vs.* wild-type						
Training set						
Clinical model	0.719 (0.668-0.770)	0.706 (0.659-0.751)	0.721 (0.589-0.792)	0.697 (0.574-0.743)	0.603 (0.554-0.626)	0.796 (0.763-0.806)
Radiomics model	0.807 (0.765-0.850)	0.747 (0.701-0.789)	0.708 (0.597-0.779)	0.772 (0.651-0.834)	0.665 (0.626-0.686)	0.805 (0.777-0.817)
Combined model	0.838 (0.799-0.877)	0.775 (0.30-0.815)	0.682 (0.565-0.760)	0.834 (0.705-0.896)	0.724 (0.685-0.745)	0.804 (0.776-0.815)
Validation set						
Clinical model	0.693 (0.606-0.781)	0.667 (0.581-0.745)	0.648 (0.268-0.787)	0.679 (0.512-0.762)	0.565 (0.349-0.612)	0.750 (0.694-0.771)
Radiomics model	0.779 (0.694-0.864)	0.732 (0.650-0.804)	0.685 (0.592-0.815)	0.762 (0.571-0.952)	0.649 (0.615-0.687)	0.790 (0.738-0.825)
Combined model	0.813 (0.740-0.886)	0.768 (0.689-0.836)	0.648 (0.407-0.759)	0.845 (0.678-0.929)	0.729 (0.629-0.759)	0.789 (0.750-0.804)
L858R mutation *vs.* wild-type						
Training set						
Clinical model	0.701 (0.651-0.750)	0.679 (0.630-0.725)	0.634 (0.507-0.722)	0.705 (0.600-0.769)	0.564 (0.509-0.596)	0.762 (0.732-0.778)
Radiomics model	0.825 (0.783-0.868)	0.764 (0.719-0.806)	0.772 (0.662-0.842)	0.759 (0.647-0.826)	0.659 (0.623-0.678)	0.847 (0.825-0.858)
Combined model	0.855 (0.817-0.894)	0.756 (0.710-0.798)	0.890 (0.793-0.945)	0.676 (0.531-0.747)	0.623 (0.596-0.637)	0.911 (0.889-0.918)
Validation set						
Clinical model	0.697 (0.614-0.781)	0.672 (0.585-0.750)	0.660 (0.458-0.820)	0.679 (0.552-0.798)	0.550 (0.459-0.603)	0.770 (0.732-0.798)
Radiomics model	0.812 (0.737-0.887)	0.746 (0.664-0.817)	0.760 (0.580-0.880)	0.738 (0.547-0.845)	0.633 (0.568-0.667)	0.838 (0.793-0.855)
Combined model	0.852 (0.790-0.913)	0.739 (0.656-0.811)	0.920 (0.740-1.000)	0.631 (0.500-0.774)	0.597 (0.544-0.617)	0.930 (0.913-0.942)
Del-19 mutation *vs.* L858R mutation						
Training set						
Logistic model	0.581 (0.510-0.651)	0.587 (0.524-0.649)	0.789 (0.593-0.854)	0.395 (0.256-0.473)	0.554 (0.483-0.574)	0.662 (0.559-0.701)
RF model	0.881 (0.840-0.921)	0.786 (0.730-0.835)	0.715 (0.621-0.825)	0.853 (0.769-0.948)	0.822 (0.801-0.842)	0.759 (0.739-0.778)
SVM model	0.601 (0.531-0.671)	0.591 (0.528-0.653)	0.805 (0.634-0.870)	0.388 (0.240-0.465)	0.556 (0.497-0.575)	0.676 (0.564-0.714)
Clinical model	0.660 (0.593-0.726)	0.599 (0.536-0.660)	0.797 (0.645-0.872)	0.411 (0.278-0.506)	0.563 (0.511-0.585)	0.679 (0.589-0.723)
Combined model†	0.906 (0.869-0.943)	0.833 (0.781-0.877)	0.821 (0.738-0.916)	0.845 (0.758-0.923)	0.835 (0.819-0.849)	0.832 (0.816-0.844)
Validation set						
Logistic model	0.673 (0.569-0.777)	0.679 (0.582-0.767)	0.827 (0.481-0.924)	0.537 (0.315-0.686)	0.632 (0.500-0.658)	0.763 (0.654-0.804)
RF model	0.871 (0.802-0.941)	0.849 (0.766-0.911)	0.788 (0.394-0.904)	0.907 (0.654-0.981)	0.891 (0.804-0.904)	0.817 (0.763-0.828)
SVM model	0.652 (0.547-0.758)	0.651 (0.552-0.741)	0.808 (0.538-0.962)	0.500 (0.370-0.648)	0.609 (0.509-0.649)	0.730 (0.667-0.778)
Clinical model	0.514 (0.402-0.625)	0.538 (0.438-0.635)	0.692 (0.378-0.885)	0.389 (0.248-0.546)	0.522 (0.373-0.582)	0.568 (0.455-0.648)
Combined model†	0.875 (0.781-0.929)	0.830 (0.745-0.896)	0.885 (0.490-0.962)	0.778 (0.536-0.889)	0.793 (0.680-0.806)	0.875 (0.828-0.889)

AUC, Area under the curve; Del 19, Exon-19 deletion; L858R, Exon-21 L858R point mutation; NPV, Negative predictive value; PPV, Positive predictive value; RF, Random forest; SVM, Support vector machine; vs., versus.

†Combined model: RF model combined Clinical model.

efficacy (19, 22, 32). This study investigated the relationship between radiomics features and EGFR molecular subtypes. Among these features, most of them were texture features, indicating that texture features were more closely related to EGFR molecular subtypes. The human visual system cannot recognize these features, nor can they be interpreted as specific meanings (34, 35). We observed that radiomics features, including logarithm_glcm_Correlation, wavelet.LLL_glszm_Zone Entropy (ZE), and gradient_glszm_Gray Level Non-Uniformity Normalized (GLNN), etc. were associated significantly with Del-19 mutation. Among them, logarithm_glcm_Correlation and wavelet.LLL_glszm_ZE reflected the image texture of the tumor

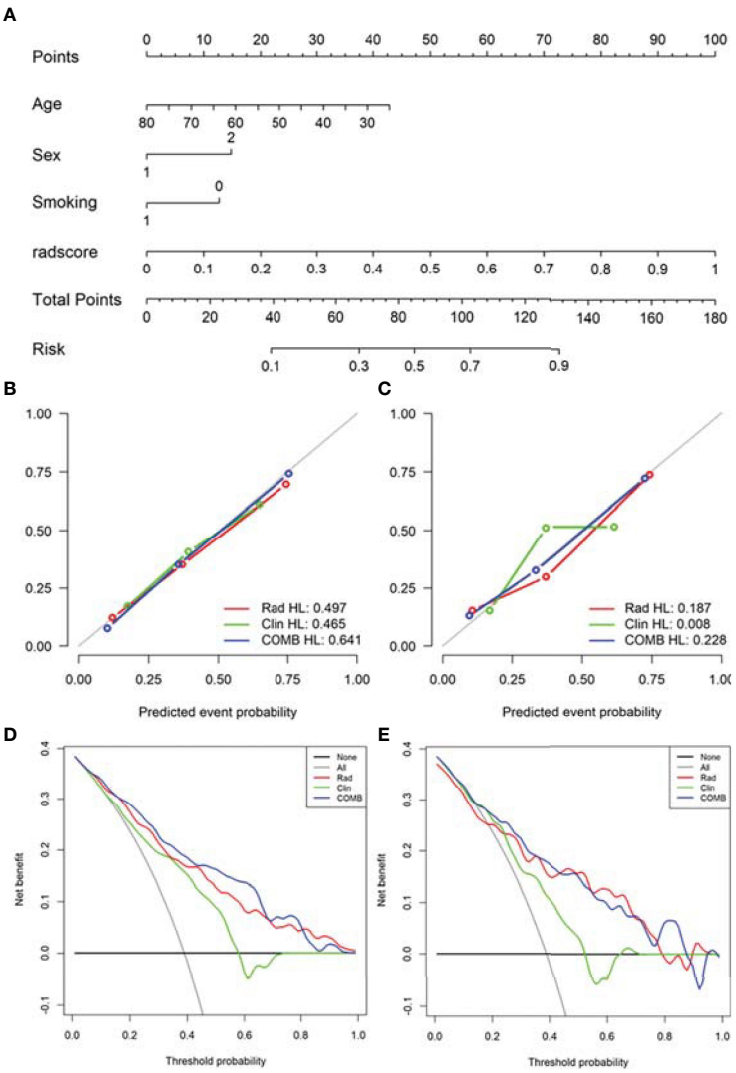


FIGURE 4 | Nomogram was used to identify Del-19 mutation and wild-type. **(A)** Construct a nomogram in the training set based on the combined model. **(B, C)** Calibration curve of the combined model in the training **(B)** and validation **(C)** sets. The x-axis represents the use of the combined model to predict the risk of Del-19 mutation. The y-axis represents the actual Del-19 mutation rate. The green, red, and blue lines represent the distinguishing ability of the clinical, radiomics, and combined models, respectively, while the gray diagonal line represents the ideal evaluation of the ideal model. The closer the fit to the diagonal line indicates the better discrimination ability. **(D, E)** Decision curve analysis for the combined model in the training **(D)** and validation **(E)** sets. The x-axis shows the threshold probability, and the y-axis measures the net benefit. The gray line represents all patients with Del-19 mutation, and the black line represents all patients without Del-19 mutation. The green, red, and blue lines represent the clinical, radiomics, and combined models, respectively.

area, and gradient_glszm_GLNN reflected the gray image value of the tumor area (36). Compared with the EGFR wild-type group, the values of these features were higher in Del-19 mutation, indicating that the image texture and gray image values were related to Del-19 mutation. Lbp.3D.m1_firstorder_10Percentile and Lbp.3D.m1_firstorder_Skewness, etc. were associated significantly with L858R mutation. They reflected the voxel intensity of the image (36). Compared with the EGFR wild-type group, the values of these features were higher in L858R mutation, indicating that the image voxel intensity was related to L858R mutation. Therefore, radiomics features as a new auxiliary tool can predict EGFR molecular subtypes.

Compared with the radiomics model based on only containing radiomics features, incorporating preoperative clinical factors of the nomogram showed the best predictive performance. This user-friendly nomogram will help clinicians easily predict EGFR molecular subtypes in clinical practice. The results were more practical than a single model and can be used for clinical applications in patients with lung adenocarcinoma undergoing CT scans. The task of Del-19 vs. wild-type and L858R vs. wild-type build with linear model (logistic regression) could obtain a satisfactory result, and the linear model is easy for application. Such we didn't applied nonlinear model. The task of Del-19 vs. L858R was hard, the performance of linear model was not satisfactory, so we

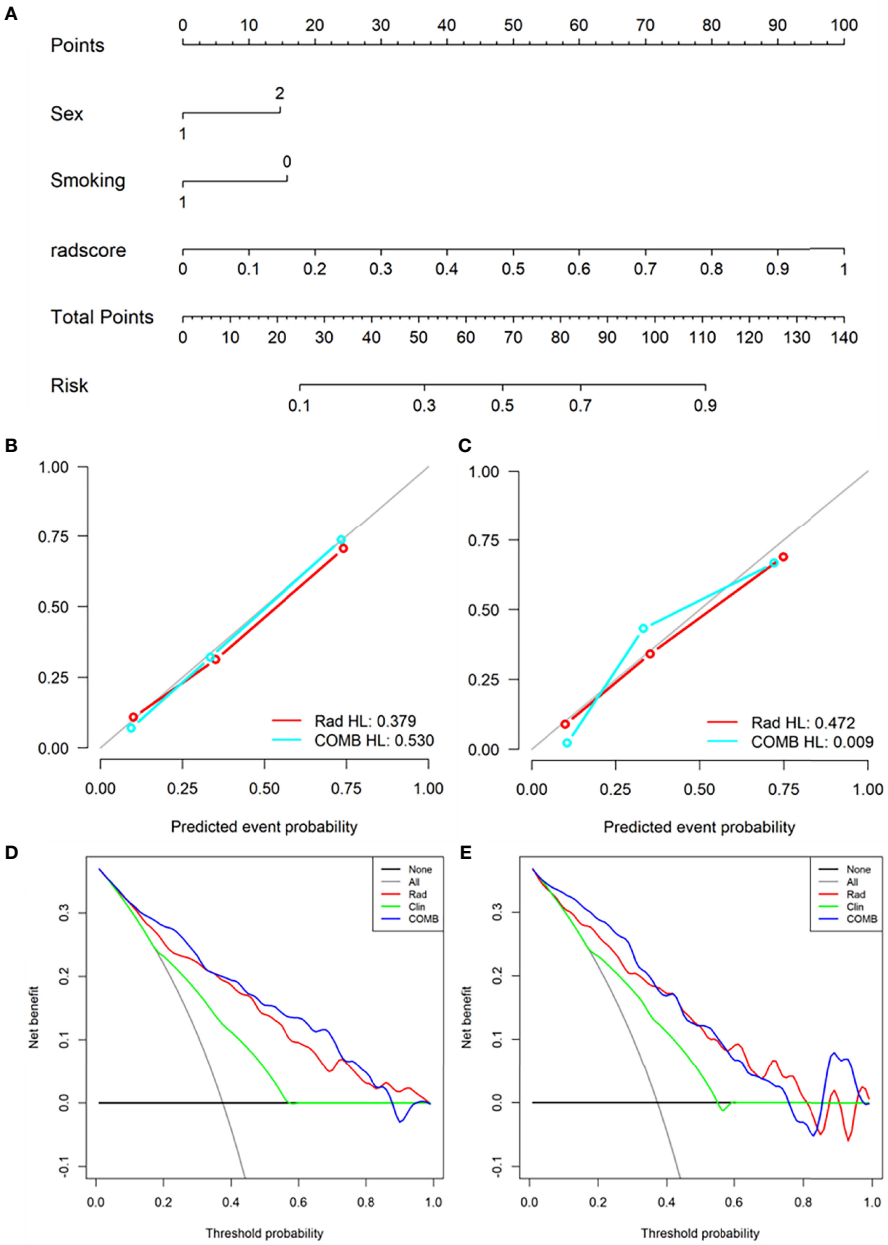


FIGURE 5 | Nomogram was used to identify L858R mutation and wild-type. **(A)** Construct a nomogram in the training set based on the combined model. **(B, C)** Calibration curve of the combined model in the training **(B)** and validation **(C)** sets. **(D, E)** Decision curve analysis for the combined model in the training **(D)** and validation **(E)** sets.

add nonlinear model for comparison and selected best model for radiomics score construction.

Our study had several limitations. First, although we collected data from two large medical centers, this was a retrospective study and there may be inevitable selection bias. The conclusions of this study need to be prospectively verified in more centers to improve the clinical applicability of our model. Second, although our study included 728 patients, increasing the sample size will further improve the accuracy of the results of this study. Finally, two radiologists spent a lot of time manually segmenting ROI. Therefore, ROI can be segmented automatically and effectively in future research.

CONCLUSION

In conclusion, demonstrated the feasibility of identifying EGFR molecular subtypes through the radiomics features of patients

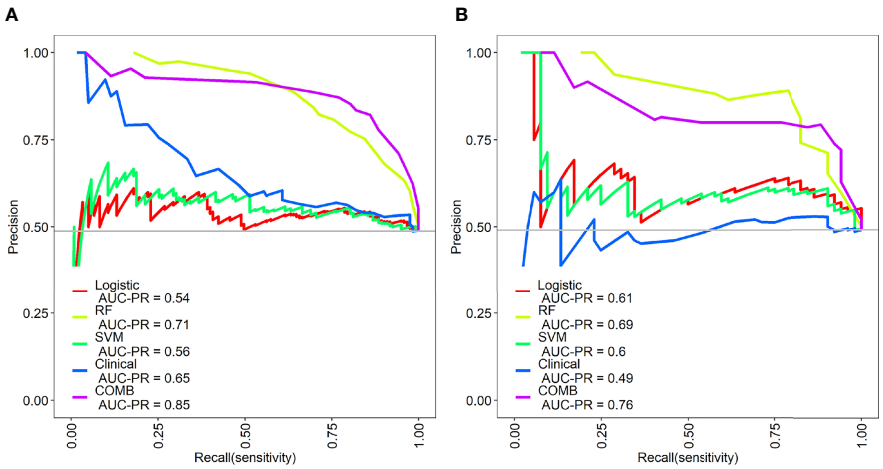


FIGURE 6 | Precision-recall (PR) curves of the different models in the training **(A)** and validation sets **(B)**. PR represents the relationship between precision and recall.The larger the area under the PR curve, the better the model performance.

with lung adenocarcinoma, making the formulation of clinically individualized targeted therapy programs more precise and more in line with actual clinical needs, so as to benefit the patients with candidate targeted therapy the most.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the medical ethics committees of Sichuan Provincial People's Hospital and Lanzhou University Second Hospital. The ethics committee waived the requirement to participate in written informed consent.

AUTHOR CONTRIBUTIONS

GZ, LD, WK, JZ and HP contributed to conception and design of the study. GZ, LD, SL, RQ, SP and XZ organized the database. GZ, YC, JZ and JR performed the statistical analysis. GZ, LD, JZ and YC wrote the first draft of the manuscript. YC, JZ and ZZ wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study received funding from the Sichuan Provincial Cadre Health Research Project (No. Chuan Gan Yan 2022-208), Qinghai Province "Kunlun Talents High-end Innovation and Entrepreneurial Talents" Top Talent Cultivation Project, Medical Science and Technology Research Fund Project of Guangdong Province (B2022144), Science and Technology Plan Fund of Guizhou Provincial (Qiankehe Foundation-ZK [2022] General 634), and Doctoral research start-up fund project of Zunyi Medical University (BS2021-03).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.889293/full#supplementary-material

REFERENCES

1. Harrison PT, Vyse S, Huang PH. Rare Epidermal Growth Factor Receptor (EGFR) Mutations in non-Small Cell Lung Cancer. *Semin Cancer Biol* (2020) 61:167–79. doi: 10.1016/j.semcancer.2019.09.015
2. D'Angelo SP, Pietanza MC, Johnson ML, Riely GJ, Miller VA, Sima CS, et al. Incidence of EGFR Exon 19 Deletions and L858R in Tumor Specimens From Men and Cigarette Smokers With Lung Adenocarcinomas. *J Clin Oncol* (2011) 29(15):2066–70. doi: 10.1200/jco.2010.32.6181
3. Li WQ, Cui JW. Non-Small Cell Lung Cancer Patients With Ex19del or Exon 21 L858R Mutation: Distinct Mechanisms, Different Efficacies to Treatments. *J Cancer Res Clin Oncol* (2020) 146(9):2329–38. doi: 10.1007/s00432-020-03296-6
4. Jänne PA, Yang JC, Kim DW, Planchard D, Ohe Y, Ramalingam SS, et al. AZD9291 in EGFR Inhibitor-Resistant non-Small-Cell Lung Cancer. *N Engl J Med* (2015) 372(18):1689–99. doi: 10.1056/NEJMoa1411817
5. Lee DH. Treatments for EGFR-Mutant non-Small Cell Lung Cancer (NSCLC): The Road to a Success, Paved With Failures. *Pharmacol Ther* (2017) 174:1–21. doi: 10.1016/j.pharmthera.2017.02.001
6. Ramalingam SS, Vansteenkiste J, Planchard D, Cho BC, Gray JE, Ohe Y, et al. Overall Survival With Osimertinib in Untreated, EGFR-Mutated Advanced NSCLC. *N Engl J Med* (2020) 382(1):41–50. doi: 10.1056/NEJMoa1913662

7. Mok TS, Cheng Y, Zhou X, Lee KH, Nakagawa K, Niho S, et al. Improvement in Overall Survival in a Randomized Study That Compared Dacomitinib With Gefitinib in Patients With Advanced Non-Small-Cell Lung Cancer and EGFR-Activating Mutations. *J Clin Oncol* (2018) 36(22):2244–50. doi: 10.1200/jco.2018.78.7994
8. Mok TS, Wu YL, Ahn MJ, Garassino MC, Kim HR, Ramalingam SS, et al. Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer. *N Engl J Med* (2017) 376(7):629–40. doi: 10.1056/NEJMoa1612674
9. Soria JC, Ohe Y, Vansteenkiste J, Reungwetwattana T, Chewaskulyong B, Lee KH, et al. Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. *N Engl J Med* (2018) 378(2):113–25. doi: 10.1056/NEJMoa1713137
10. Yang X, Dong X, Wang J, Li W, Gu Z, Gao D, et al. Computed Tomography-Based Radiomics Signature: A Potential Indicator of Epidermal Growth Factor Receptor Mutation in Pulmonary Adenocarcinoma Appearing as a Subsolid Nodule. *Oncologist* (2019) 24(11):e1156–e64. doi: 10.1634/theoncologist.2018-0706
11. Zhang G, Cao Y, Zhang J, Zhao Z, Zhang W, Zhou J. Epidermal Growth Factor Receptor Mutations in Lung Adenocarcinoma: Associations Between Dual-Energy Spectral CT Measurements and Histologic Results. *J Cancer Res Clin Oncol* (2021) 147(4):1169–78. doi: 10.1007/s00432-020-03402-8
12. Zhang G, Zhang J, Cao Y, Zhao Z, Li S, Deng L, et al. Nomogram Based on Preoperative CT Imaging Predicts the EGFR Mutation Status in Lung Adenocarcinoma. *Transl Oncol* (2021) 14(1):100954. doi: 10.1016/j.tranon.2020.100954
13. Rolfo C, Mack PC, Scagliotti GV, Baas P, Barlesi F, Bivona TG, et al. Liquid Biopsy for Advanced Non-Small Cell Lung Cancer (NSCLC): A Statement Paper From the IASLC. *J Thorac Oncol* (2018) 13(9):1248–68. doi: 10.1016/j.jtho.2018.05.030
14. Goldman JW, Noor ZS, Remon J, Besse B, Rosenfeld N. Are Liquid Biopsies a Surrogate for Tissue EGFR Testing? *Ann Oncol* (2018) 29(suppl_1):i38–46. doi: 10.1093/annonc/mdx706
15. Zhang G, Zhao Z, Cao Y, Zhang J, Li S, Deng L, et al. Relationship Between Epidermal Growth Factor Receptor Mutations and CT Features in Patients With Lung Adenocarcinoma. *Clin Radiol* (2021) 76(6):473.e17–24. doi: 10.1016/j.crad.2021.02.012
16. Liu Y, Kim J, Qu F, Liu S, Wang H, Balagurunathan Y, et al. CT Features Associated With Epidermal Growth Factor Receptor Mutation Status in Patients With Lung Adenocarcinoma. *Radiology* (2016) 280(1):271–80. doi: 10.1148/radiol.2016151455
17. Lee HJ, Kim YT, Kang CH, Zhao B, Tan Y, Schwartz LH, et al. Epidermal Growth Factor Receptor Mutation in Lung Adenocarcinomas: Relationship With CT Characteristics and Histologic Subtypes. *Radiology* (2013) 268 (1):254–64. doi: 10.1148/radiol.13112553
18. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
19. Zhang G, Cao Y, Zhang J, Ren J, Zhao Z, Zhang X, et al. Predicting EGFR Mutation Status in Lung Adenocarcinoma: Development and Validation of a Computed Tomography-Based Radiomics Signature. *Am J Cancer Res* (2021) 11(2):546–60.
20. Tu W, Sun G, Fan L, Wang Y, Xia Y, Guan Y, et al. Radiomics Signature: A Potential and Incremental Predictor for EGFR Mutation Status in NSCLC Patients, Comparison With CT Morphology. *Lung Cancer* (2019) 132:28–35. doi: 10.1016/j.lungcan.2019.03.025
21. Hong D, Xu K, Zhang L, Wan X, Guo Y. Radiomics Signature as a Predictive Factor for EGFR Mutations in Advanced Lung Adenocarcinoma. *Front Oncol* (2020) 10:28. doi: 10.3389/fonc.2020.00028
22. Liu Y, Kim J, Balagurunathan Y, Li Q, Garcia AL, Stringfield O, et al. Radiomic Features Are Associated With EGFR Mutation Status in Lung Adenocarcinomas. *Clin Lung Cancer* (2016) 17(5):441–48.e6. doi: 10.1016/j.cllc.2016.02.001
23. Mei D, Luo Y, Wang Y, Gong J. CT Texture Analysis of Lung Adenocarcinoma: Can Radiomic Features be Surrogate Biomarkers for EGFR Mutation Statuses. *Cancer Imaging* (2018) 18(1):52. doi: 10.1186/s40644-018-0184-2
24. Liu G, Xu Z, Ge Y, Jiang B, Groen H, Vliegenthart R, et al. 3D Radiomics Predicts EGFR Mutation, Exon-19 Deletion and Exon-21 L858R Mutation in Lung Adenocarcinoma. *Transl Lung Cancer Res* (2020) 9(4):1212–24. doi: 10.21037/tlcr-20-122
25. Zhao W, Wu Y, Yn X, Sun Y, Gao P, Tan M, et al. The Potential of Radiomics Nomogram in Non-Invasively Prediction of Epidermal Growth Factor Receptor Mutation Status and Subtypes in Lung Adenocarcinoma. *Front Oncol* (2020) 9:1485(1485). doi: 10.3389/fonc.2019.01485
26. Li S, Ding C, Zhang H, Song J, Wu L. Radiomics for the Prediction of EGFR Mutation Subtypes in Non-Small Cell Lung Cancer. *Med Phys* (2019) 46 (10):4545–52. doi: 10.1002/mp.13747
27. Zhang L, Chen B, Liu X, Song J, Fang M, Hu C, et al. Quantitative Biomarkers for Prediction of Epidermal Growth Factor Receptor Mutation in Non-Small Cell Lung Cancer. *Transl Oncol* (2018) 11(1):94–101. doi: 10.1016/j.tranon.2017.10.012
28. Ji GW, Zhang YD, Zhang H, Zhu FP, Wang K, Xia YX, et al. Biliary Tract Cancer at CT: A Radiomics-Based Model to Predict Lymph Node Metastasis and Survival Outcomes. *Radiology* (2019) 290(1):90–8. doi: 10.1148/radiol.2018181408
29. 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(3):947–57. doi: 10.1148/radiol.2016152234
30. Zhou JY, Zheng J, Yu ZF, Xiao WB, Zhao J, Sun K, et al. Comparative Analysis of Clinicoradiologic Characteristics of Lung Adenocarcinomas With ALK Rearrangements or EGFR Mutations. *Eur Radiol* (2015) 25(5):1257–66. doi: 10.1007/s00330-014-3516-z
31. Dearden S, Stevens J, Wu YL, Blowers D. Mutation Incidence and Coincidence in Non Small-Cell Lung Cancer: Meta-Analyses by Ethnicity and Histology (Mutmap). *Ann Oncol* (2013) 24(9):2371–6. doi: 10.1093/annonc/mdt205
32. Jia TY, Xiong JF, Li XY, Yu W, Xu ZY, Cai XW, et al. Identifying EGFR Mutations in Lung Adenocarcinoma by Noninvasive Imaging Using Radiomics Features and Random Forest Modeling. *Eur Radiol* (2019) 29 (9):4742–50. doi: 10.1007/s00330-019-06024-y
33. Park JE, Park SY, Kim HJ, Kim HS. Reproducibility and Generalizability in Radiomics Modeling: Possible Strategies in Radiologic and Statistical Perspectives. *Korean J Radiol* (2019) 20(7):1124–37. doi: 10.3348/kjr.2018.0070
34. Wu Q, Yao K, Liu Z, Li L, Zhao X, Wang S, et al. Radiomics Analysis of Placenta on T2WI Facilitates Prediction of Postpartum Haemorrhage: A Multicentre Study. *EBioMedicine* (2019) 50:355–65. doi: 10.1016/j.ebiom.2019.11.010
35. Zhang J, Yao K, Liu P, Liu Z, Han T, Zhao Z, et al. A Radiomics Model for Preoperative Prediction of Brain Invasion in Meningioma non-Invasively Based on MRI: A Multicentre Study. *EBioMedicine* (2020) 58:102933. doi: 10.1016/j.ebiom.2020.102933
36. Zwanenburg A, Vallières M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145

Conflict of Interest: Author JR was employed by GE Healthcare. Author XZ was employed by Philips (China) Investment Co., Ltd.

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

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 25 May 2022
doi: 10.3389/fonc.2022.876264

A Nomogram Combined Radiomics and Clinical Features as Imaging Biomarkers for Prediction of Visceral Pleural Invasion in Lung Adenocarcinoma

Xinyi Zha[1†], *Yuanqing Liu*[1†], *Xiaoxia Ping*[1,2], *Jiayi Bao*[1], *Qian Wu*[1], *Su Hu*[1,2*] *and Chunhong Hu*[1,2*]

[1] Department of Radiology, The First Affiliated Hospital of Soochow University, Suzhou, China, [2] Institute of Medical Imaging, Soochow University, Suzhou, China

OPEN ACCESS

Edited by:
Yiyan Liu,
University of Louisville, United States

Reviewed by:
Xiaohu Li,
First Affiliated Hospital of Anhui Medical University, China
Guojin Zhang,
Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, China

***Correspondence:**
Su Hu
husu@suda.edu.com
Chunhong Hu
sdhuchunhong@sina.com

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

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: 15 February 2022
Accepted: 21 April 2022
Published: 25 May 2022

Citation:
Zha X, Liu Y, Ping X, Bao J, Wu Q, Hu S and Hu C (2022) A Nomogram Combined Radiomics and Clinical Features as Imaging Biomarkers for Prediction of Visceral Pleural Invasion in Lung Adenocarcinoma. Front. Oncol. 12:876264. doi: 10.3389/fonc.2022.876264

Objectives: To develop and validate a nomogram model based on radiomics features for preoperative prediction of visceral pleural invasion (VPI) in patients with lung adenocarcinoma.

Methods: A total of 659 patients with surgically pathologically confirmed lung adenocarcinoma underwent CT examination. All cases were divided into a training cohort (n = 466) and a validation cohort (n = 193). CT features were analyzed by two chest radiologists. CT radiomics features were extracted from CT images. LASSO regression analysis was applied to determine the most useful radiomics features and construct radiomics score (radscore). A nomogram model was developed by combining the optimal clinical and CT features and the radscore. The model performance was evaluated using ROC analysis, calibration curve and decision curve analysis (DCA).

Results: A total of 1316 radiomics features were extracted. A radiomics signature model with a selection of the six optimal features was developed to identify patients with or without VPI. There was a significant difference in the radscore between the two groups of patients. Five clinical features were retained and contributed as clinical feature models. The nomogram combining clinical features and radiomics features showed improved accuracy, specificity, positive predictive value, and AUC for predicting VPI, compared to the radiomics model alone (specificity: training cohort: 0.89, validation cohort: 0.88, accuracy: training cohort: 0.84, validation cohort: 0.83, AUC: training cohort: 0.89, validation cohort: 0.89). The calibration curve and decision curve analyses suggested

that the nomogram with clinical features is beyond the traditional clinical and radiomics features.

Conclusion: A nomogram model combining radiomics and clinical features is effective in non-invasively prediction of VPI in patients with lung adenocarcinoma.

Keywords: CT, lung adenocarcinomas, radiomics, Nomogram, prediction, visceral pleural invasion

INTRODUCTION

Lung cancer is currently the second most common cancer in the world and remains the leading cause of death among malignant tumors (1). Over 83% of lung cancers are non-small cell lung cancer (NSCLC) (2). Visceral pleural invasion (VPI), defined as tumor extension beyond the elastic layer of viscera pleura, is one of the most important adverse prognostic factors in non-small cell lung cancers with tumor sizes ≤ 3 cm (3, 4). In the eighth edition of TNM classification for NSCLC, VPI increases the T staging of lung cancer with diameters ≤ 3 cm: the presence of VPI leads to upstaging T1 tumor to T2 and stage IA tumor to IB (5, 6).

Several studies have evaluated the morphological characteristics of VPI in NSCLC based on CT images (7–9). However, there is no definite morphological feature that can reliably predict VPI, especially when the tumor is far from the pleura without pleura indentation or pleural attachment (7–9). Radiomics extracts a large amount of quantitative information from medical images (10, 11). Radiomics have been utilized for clinical-decision support systems in lung cancer, including diagnose and prognostic prediction (12–14). However, few studies have been reported to assess for the presence of VPI in patients with NSCLC using radiomics methods (15).

Therefore, the purpose of this study was to construct a nomogram model based on radiomics features, and determine whether VPI of lung adenocarcinoma can be predicted using the model.

MATERIALS AND METHODS

Patients

This retrospective study was approved by the institutional review board of the First Affiliated Hospital of Soochow University (Suzhou, China), and the requirement for patient informed consent was waived. Patients with peripheral lung adenocarcinoma who underwent chest CT scans with thin-section (1–1.25 mm) images from January 2016 to December 2020 were reviewed. Inclusion criteria were as follows: (a) all the patients were confirmed as lung adenocarcinoma by pathological examination, and whether pleural invasion or not was evaluated pathologically; (b) the peripheral lesion was determined as N0M0 stage with the largest diameter smaller than 3.0 cm; (c) thin-section CT scan was performed within 30 days before surgery; (d) available results for clinical data, including age, sex, smoking history; (e) at least one of the following features were presented on CT images: pleural depression, pleural attachment or pleural closeness. 750 patients were excluded because of the following reasons: (a) histological diagnosis of SCLC (n=89); (b) tumor size > 3 cm (n=185); (c) whether pleural invasion or not cannot be assessed pathologically (n=136); (d) the lesion is far from the pleura without any of the above three features (n =301). (e) poor imaging quality due to respiratory artifact during examination (n=39). Finally, A total of 659 patients met all the inclusion criteria and included in this study.

CT Scans

Patients underwent preoperative unenhanced CT scanning using various multidetector row scanners: Brilliance 16 or Brilliance iCT (Philips Healthcare, Best, the Netherlands), Somatom Sensation 64 or Somatom Definition (Siemens Healthineers, Erlangen, Germany), GE revolution or Discovery CT 750 HD (GE Healthcare, Chicago, USA), Aquilion One (Toshiba Medical Systems, Tokyo, Japan). The imaging parameters for thin-section CT were as follows: tube voltage 100-120 kV, automatic tube current modulation, matrix 512 × 512, field of view (FOV) of 400 mm (Brilliance 16 scanner) and 500 mm (other machines), slice thickness of 1-2 mm, the iterative reconstruction algorithm. All CT images were obtained in the supine position during inspiratory breath-hold.

Imaging Analysis

Two experienced radiologists analyzed the CT images independently with a lung window (window width, 1500 HU; window level, −500 HU) and mediastinum window (window width, 400 HU; window level, 60 HU). Consensus was reached by discussion in case of disagreement. Image features included the following (1): tumor density (solid/part-solid) (2); maximum diameter (3); margin (lobulated, spiculated) (4); air bronchogram (5); pleura indentation, pleural attachment, or pleural closeness (6); distance from the pleura. A part-solid nodule was defined as a tumor that included both GGO and solid components (0<CTR<1.0). A pure-solid nodule was defined as a tumor that included only consolidation without GGO (CTR=1.0) (16). In the current study, pure GGO was excluded since VPI was never observed in these lesions due to its minimally invasive nature and inability to penetrate the thick elastic layer (17). Pleural indentation was defined as tumor indentation of the visceral pleura on CT images at the lung window. Pleural attachment was defined as no visible space between the nodule and the visceral pleura on CT images at the lung window or tumor attachment to the interlobar pleura at the lung window. Pleural closeness was defined as tumor located within 1.0 cm of the pleura (16).

Histologic Evaluation

Surgically resected specimens were stained with hematoxylin and eosin, and examined to determine the presence or absence of VPI. VPI was defined as invasion beyond the elastic layer of the visceral pleura according to the 8th edition of the TNM classification criteria (5). Histologic evaluation was performed by one experienced pathologist.

Tumor Segmentation and Radiomics Feature Extraction

CT images of enrolled patients were exported from the picture archiving and communication system (PACS), and segmented semi-automatically using ITK-SNAP software (version 3.6.0, www.itk-snap.org) (18). The workflow of the analysis is summarized in **Figure 1**. All images were automatically segmented and adjusted by a radiologist with 8 years of experience. After 4 weeks this radiologist segmented the images of 30 randomly selected patients for intra-observer reproducibility. In addition, another radiologist with 20 years of experience segmented 30 randomly selected patient images for inter-observer reproducibility. The inter- and intra-observer reproducibility of feature extraction was evaluated by intraclass correlation coefficients (ICCs). ICCs greater than 0.75 were considered as good consistency.

All images were performed image normalization before feature extraction (19). Radiomics features were extracted from the ROI by the pyradiomic package Python software (version 3.7.12, www.python.org). A total of 1316 high-dimensional features were extracted from each sample and these were classified into seven categories: first order statistics (n = 252), shape (n = 14), neighborhood gray-tone difference matrix (n = 70), grey level dependence matrix (GLDM) (n = 196), grey level co-occurrence matrix (GLCM) (n = 336), run-length matrix (RLM) (n = 224), and grey level zone size matrix (GLZSM) (n = 224).

Radiomics Feature-Based Prediction Model Construction

Radiomics signature model based on selected features from the training cohort was constructed. Two feature selection methods were used to select the features. First, maximum relevance minimum redundancy (mRMR) was performed to eliminate redundant and irrelevant features. Then, least absolute shrinkage and selection operator (LASSO) was used to select the most useful features. A radiomics score (Radscore) was computed for each patient through a linear combination of selected features weighted by their respective coefficients. The final formula for the Radscore was as follows: "Radscore=0.085*image_wavelet-LLL_glszm_LargeAreaHighGrayLevelEmphasis+-0.034*image_exponential_firstorder_TotalEnergy+-1.071*image_wavelet-LLL_ngtdm_Coarseness+0.21*image_exponential_glszm_LargeAreaLowGrayLevelEmphasis+0.083*image_square_glszm_ZoneVariance+-0.771*image_squareroot_firstorder_Skewness + -1.266". Furthermore, the Radscore was compared between lung adenocarcinoma with VPI and those without VPI in both the training and validation cohorts.

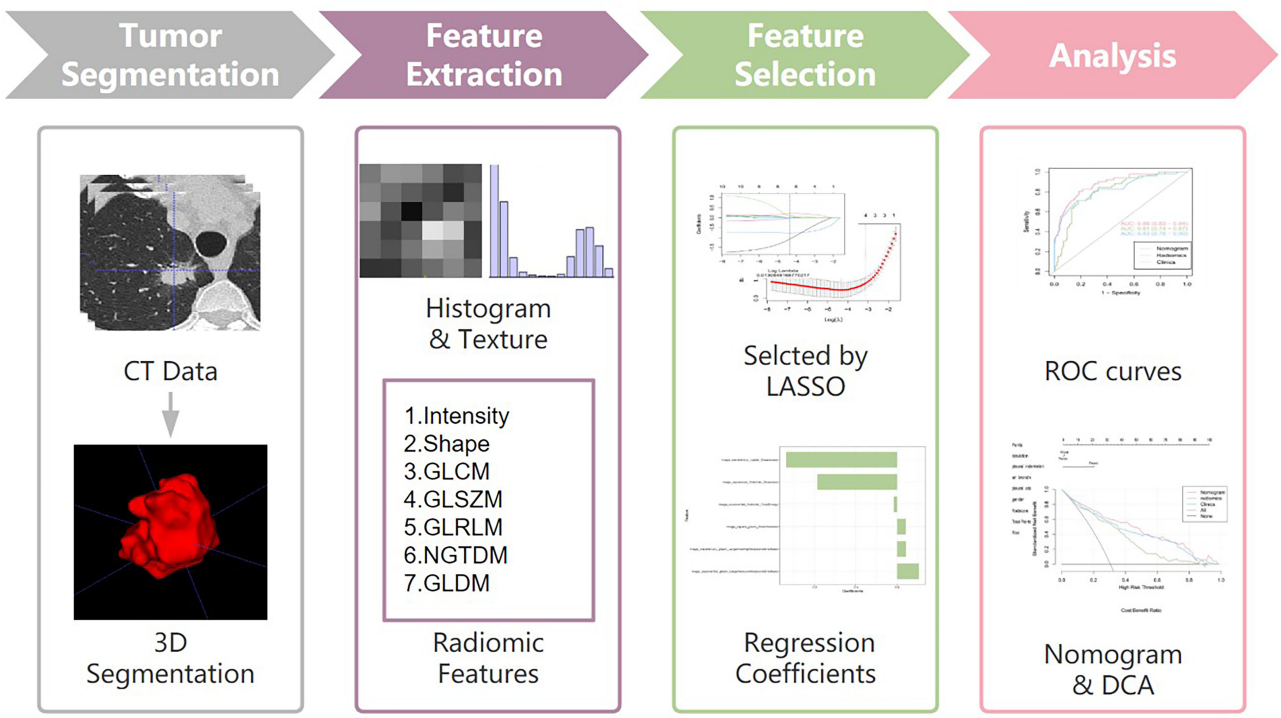


FIGURE 1 | Workflow of the study. Workflow can be divided into four parts: tumor segmentation, feature extraction, feature selection and analysis.

Logistic regression was performed to select the independent clinical predictors in the training cohort. Prediction models combining radiomics features and clinical variables were established. Finally, a radiomics nomogram based on the multivariate logistic regression model in the training cohort was constructed, and receiver operating characteristic (ROC) curves were developed to evaluate the discriminatory ability of the nomogram. The calibration curve and Hosmer-Lemeshow test was used to assess the goodness-of-fit of nomogram (20, 21). Decision curve analysis was performed to assessed the clinical value of nomogram. The net benefit is calculated within a threshold probability, defined as the minimum probability of a disease requiring further intervention (22).

Statistical Analysis

Statistical analyses were performed using R software (version 4.1.0) for quantitative characterisation. The characteristics of patients with VPI and without VPI were compared by Student's t-test for normally distributed data, otherwise the Mann-Whitney u-test was used. The intra-observer reproducibility of tumor segmentation and feature extraction were evaluated by intraclass correlation coefficients (ICCs). ICCs greater than 0.75 were considered to have good consistency. A multivariate binary logistic regression was implemented using the "rms" package. The nomogram was created and the calibration plots were created using the "rms" package. ROC curves were plotted to evaluate the diagnostic efficiency of the nomogram model. The area under the ROC curve (AUC) was calculated. P-values < 0.05 were considered to be significant.

RESULTS

Clinical Characteristics

A total of 659 patients were included in this study, of whom 193 (29.3%) were diagnosed with VPI and 466 (70.7%) were diagnosed without VPI (**Table 1**).

There were significantly differences between VPI-presence and VPI-absence group in gender, pleural indentation, pleural attachment, air bronchogram, and lobulation (all P < 0.05). Gender, pleural indentation, pleural attachment, air bronchogram, and lobulation were independent risk factors for predicting VPI after logistic regression analysis (**Table 2**). A clinical model was developed based on these characteristics.

Reproducibility Analysis

The average ICCs of intra-observer was 0.96, indicating satisfactory agreement. The number of features with fair consistency (0.75 > ICC ≥ 0.4) and poor consistency (ICC <0.4) were 4 (0.3%) and 26 (2.0%), respectively. The average ICCs of inter-observer was 0.95, indicating satisfactory agreement. The number of features with fair consistency (0.75 > ICC ≥ 0.4) and poor consistency (ICC <0.4) were 11 (0.8%) and 32 (2.4%), respectively.

Radiomics Feature Selection Signature Construction, Validation, and Evaluation

30 features were retained after eliminating the redundant and irrelevant features with mRMR. Then, 6 features were selected as the most predictive subset after LASSO (**Figures 2A, B**). The corresponding coefficients were evaluated (**Figure 2C**) and a predictive model was constructed. Radscore was calculated by summing the selected features weighted by their coefficients. There was a significant difference in radscore between lung adenocarcinoma with VPI and without VPI in the training and validation groups (**Figure 3**).

As shown in **Figure 4**, the radiomics feature model had an AUC of 0.83 in the training cohort and 0.81 in the validation cohort. Then, clinical indicators (**Table 1**) with p values less than 0.01 in the logistic regression analysis with radscore were used to constructed a combined model, which showed an AUC of 0.89 (95% CI, 0.86-0.92) in the training cohort (**Figure 4A**) and an AUC of 0.88 (95% CI, 0.83-0.94) in the validation cohort (**Figure 4B**). The predictive performance of the combined model was shown in **Table 3**. In both the training and validation cohorts, the accuracy, specificity, positive predictive value, and AUC of the combined model outperformed both the radiomics feature model and the clinical feature-based model.

Subsequently, a nomogram model was created (**Figure 5A**). The calibration curve of the nomogram for predicting VPI matched well with the estimated and actual observed values of the radiomics nomogram. The p-value for the predictive power of the nomogram obtained by the Hosmer-Lemeshow test was 0.94 in the training cohort (**Figure 5B**) and 0.86 in the validation cohort (**Figure 5C**).

The DCA showed that the net benefit of the combined nomogram outperformed the clinical and radiomics feature models (**Figure 5D**). The decision curve showed that the combined nomogram established in this study has more benefit for predicting VPI if the threshold probability of a patient is between 0 to 55%, 60% to 80% and 90% to 100%.

DISCUSSION

This study constructed and validated a nomogram model based on clinical and radiomics features extracted from CT imaging for identifying VPI in lung adenocarcinoma less than or equal to 3 cm. The nomogram model was able to classify stage I lung adenocarcinoma into those with VPI and without VPI, with AUC values greater than those of the radiomics model and the clinical model. The results demonstrated that the combined model can reliability predict VPI.

In this study, we included patients with lung adenocarcinoma to construct a nomogram to predict with or without VPI. This is because research showed that no significant difference in survival rates associated with VPI in NSCLC (23). As for NSCLC, squamous cell carcinoma and adenocarcinoma showed significantly different biological behaviors (24). Whereas the heterogeneity in biological behavior between lung adenocarcinoma and squamous cell carcinoma can be reflected by radiomics, radiomics can predict their histological subtypes (25, 26). In addition, lung adenocarcinoma is the most common subtype of lung cancer, so in this study we only discussed lung adenocarcinoma.VPI is a poor prognostic factor for lung

TABLE 1 | Characteristics of 659 lung adenocarcinoma patients, according to the presence of the visceral pleural invasion.

Characteristics	Total (n=659)	Univariate logistic regression			Multivariate logistic regression
		VPI (–) (n=466)	VPI (+) (n=193)	*P* value	*P* value
Gender				**<0.001**	**0.01**
Male	244	150	94		
Famale	415	316	99		
Age(years)	61 (53-67)	60 (52-66)	63 (57-69)	**<0.001**	NA
Smoking status				**<0.001**	NA
Active	553	413	140		
Inactive	106	53	53		
lobulation				**<0.001**	**0.04**
Present	481	315	166		
Absent	178	151	27		
spiculation				0.528	
Present	479	342	137		
Absent	180	124	56		
air bronchogram				**<0.001**	**<0.001**
Present	352	298	54		
Absent	307	168	139		
Radiological tumor type				**<0.001**	NA
Pure-solid	381	217	164		
Part-solid	278	249	29		
Pleura indentation				**<0.001**	**<0.001**
Present	422	266	156		
Absent	237	200	37		
Pleural attachment				**<0.001**	**<0.001**
Present	327	193	134		
Absent	332	273	59		

Age is expressed as Median (interquartile range). Otherwise, data are number of patients. The P value marked bold indicated statistical significance.
NA means that the characteristic is not included in the logistic regression.

adenocarcinoma (27–29), since VPI has been associated with increased overall mortality and decreased disease-free survival (30). The visceral pleura is rich in lymphatic vessels and forms an intercommunicating network on the lung surface. This network penetrates the lung parenchyma, connects to the bronchial lymphatics and flows into the hilar lymph nodes (31), which may progressively develop into metastatic disease (lymph node metastasis or distant metastasis). According to the 8th edition of the AJCC staging manual, a tumour size of 0-3 cm with VPI (including PL1 and PL2) is considered IB stage (31). Some previous studies have shown that patients with stage IB NSCLC can benefit from adjuvant chemotherapy treatment (32–34).

The correlation between CT morphological features and VPI has been reported previously (30, 35, 36). The present study concluded that lobulation and air bronchogram were not significant indicators of VPI in lung adenocarcinoma, which was consistent with previous study (30, 35). A lobulated contour implies uneven growth, which is associated with malignancy. However, lobulation also occurs in up to 25% of benign nodules (37). In our study, the lobulation sign was not an independent risk factor for predicting VPI, which may be due to a selection bias resulting from the small size of our enrolled tumors and the fact that the number of patients in this study was not large enough. The air bronchogram sign is the result of tumor cells spreading along the wall of fine bronchus and alveolar wall in a volvulus-like growth pattern without destroying the lung scaffold structure, and the residual gas in the bronchus and alveoli is visualized (38). In previous studies, air bronchogram signs were associated with low invasiveness and helped to distinguish with or without VPI of lung adenocarcinoma (39). Many studies have suggested that the node-pleura relationship is an important predictor of positive VPI. In lung adenocarcinoma, pleural indentation is generally considered to be a positive predictor of VPI (39). Indentation increases the risk of tumor invasion of the visceral pleura (40). Pleural attachment is another known factor for local recurrence and poor survival of lung adenocarcinoma after radiotherapy for non-small cell lung cancer (41, 42). Although most studies have evaluated the morphologic features of VPI on CT images, the accuracy of studies based on morphologic features of CT images remained low, and the morphological features identified are dependent on the experience of the radiologists (7–10).

Currently, radiomics allows for the non-invasive evaluation of internal tumor heterogeneity by extracting and analyzing a

TABLE 2 | Variables and coefficients of clinical model.

Variable	Adjusted OR	95%CI	*P* value
Gender	1.95	1.17-3.25	0.011
Lobulation	0.32	0.19-0.53	0.040
Air bronchogram	2.11	1.03-4.31	<0.0001
Pleura indentation	19.07	9.38-38.76	<0.0001
Pleural attachment	10.10	5.33-19.14	<0.0001

OR, odds ratio; CI, confidence interval.

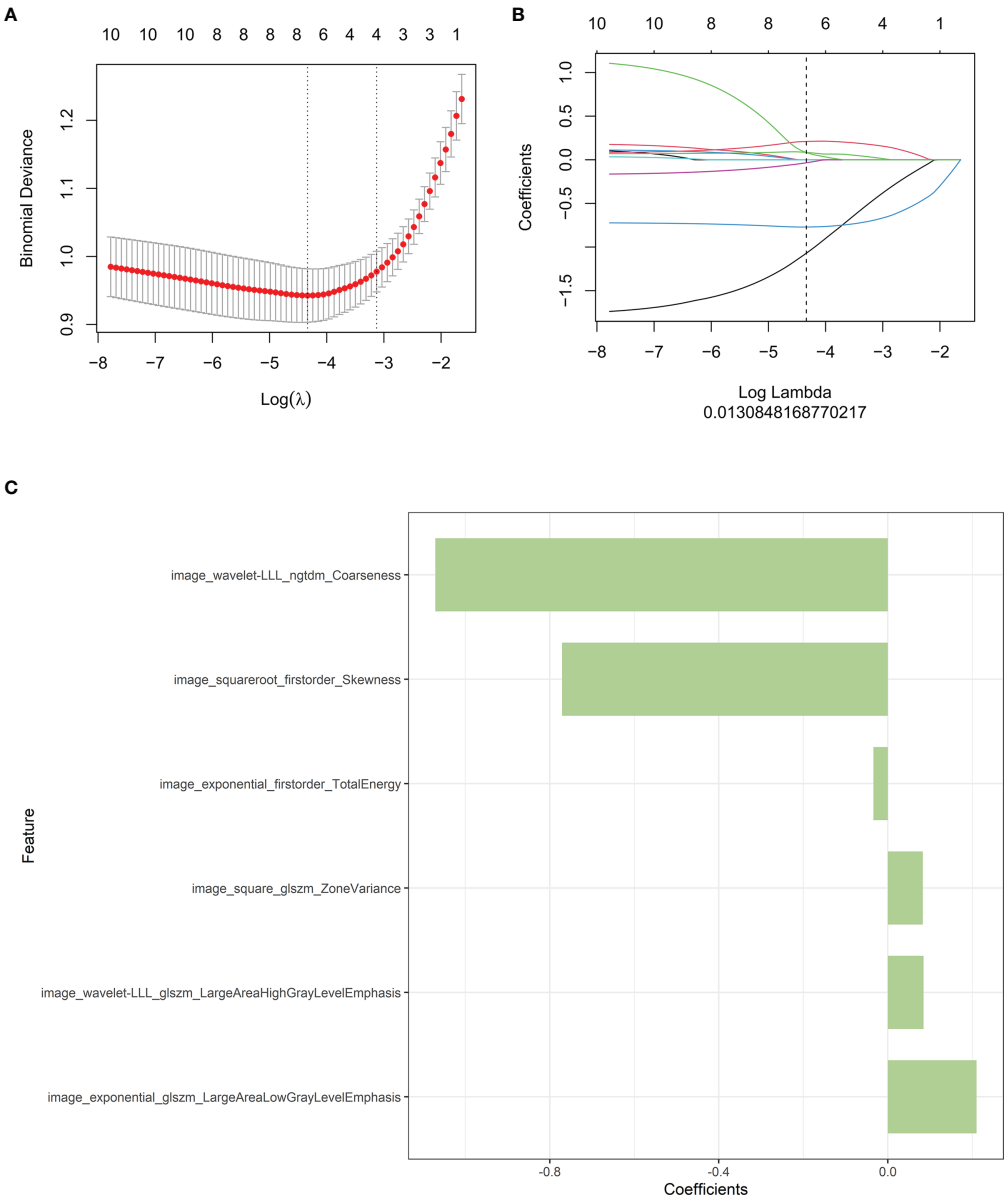


FIGURE 2 | Radiomics features associated with VPI were selected using LASSO regression models. **(A)** Cross-validation curve. An optimal log lambda (0.013) was selected, and 6 non-zero coefficients were chosen. **(B)** LASSO coefficient profiles of the 1316 radiomics features against the deviance explained. **(C)** Histogram shows the contribution of the selected parameters with their regression coefficients in the signature construction.

large number of advanced quantitative imaging features from CT images (12, 43). Yuan et al. proposed a support vector machine (SVM) based deep learning model to predict the status of VPI from preoperative CT scans, with a high AUC in the validation cohort (10). However, the model could only distinguish patients with or without VPI based on radiomics models, without incorporating relevant clinical characteristic parameters, and did not take the relations of tumor to adjacent pleura into account.

In this study, six optimal quantitative radiomics features (including Coarseness, Skewness, TotalEnergy, ZoneVariance, LAHGLE and LALGLE) were extracted Coarseness is a parameter for the neighbouring gray tone difference matrix (NGTDM). The lower coarseness values in the present study indicated more heterogeneous textures of the lesion. Skewness and total energy are both the first order parameters. Lower skewness and total energy values indicated higher heterogeneity of the lesion. LAHGLE and LALGLE are parameters for the gray level size zone matrix (GLSZM). Zone variance is also a parameter for the GLSZM. In this study, the high LAHGLE, LALGLE and zone variance values indicated high heterogeneity of the lesion.

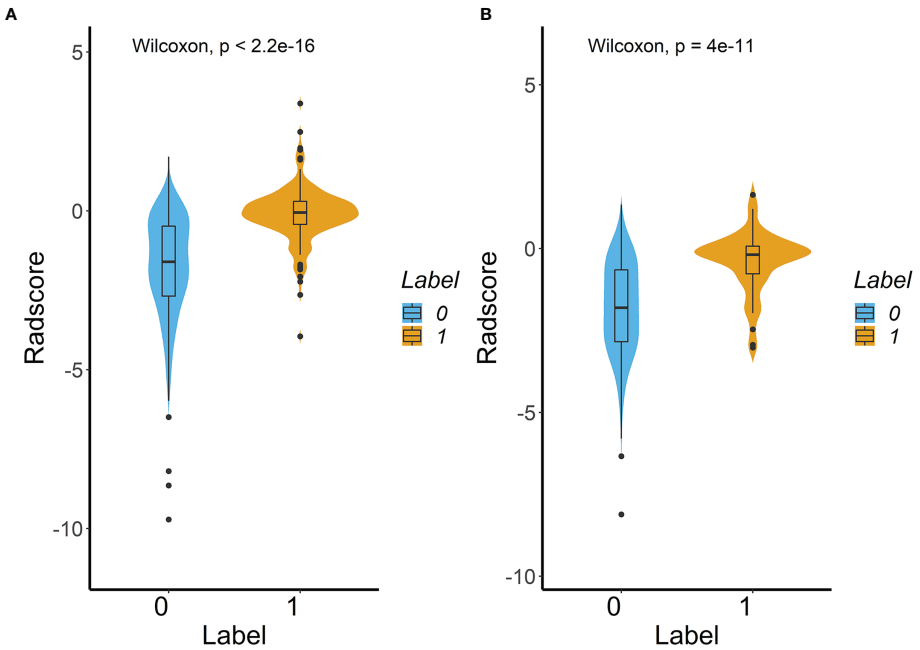


FIGURE 3 | Difference in the Radscore between lung adenocarcinoma with VPI and without VPI in training cohort **(A)** and validation cohort **(B)**. (Label 0: No VPI; label 1: VPI).

According to our findings, the model that combines radiomics features and clinical features is more effective.The clinical features model included four semantic features (the signs of lobulation, air bronchogram, pleural attachment and pleural indentation) among the five features, all describing the perimeter and morphology of tumor. Because of the difficulty in outlining peritumoral ROI due to the proximity of the tumor to the pleura, this imaging histology study focused only on the interior of the tumor. However, we included some of the imaging signs to reflect the peritumoral situation as described previously. Previous studies have demonstrated that the above features were the risk factors of VPI. Furthermore, the radiomics

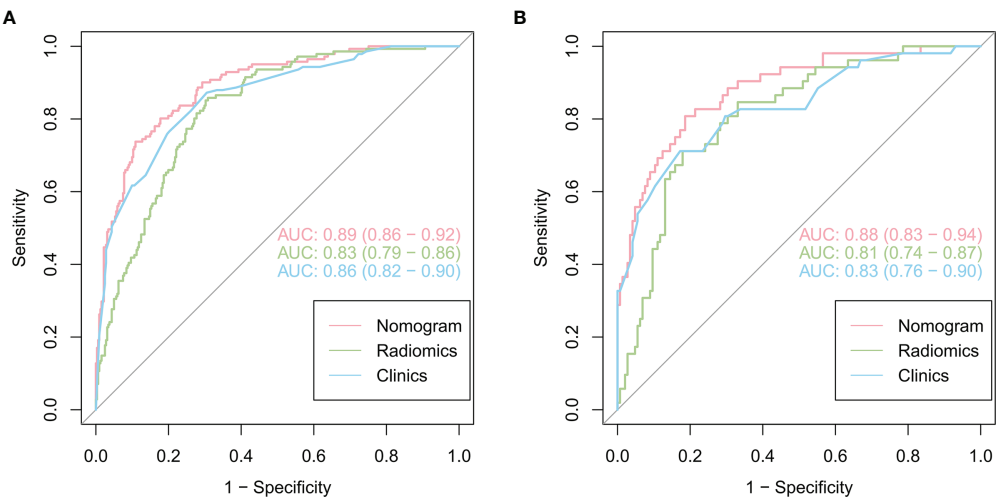


FIGURE 4 | Comparison of the performance of three models for predicting VPI in lung adenocarcinoma. ROC curves for clinical features alone, radiomics features alone and combined features for the training **(A)** and validation **(B)** cohorts.

TABLE 3 | Predictive performance of the three models in the training and validation cohorts.

Model	Accuracy [95%CI]	AUC [95%CI]	Sensitivity	Specificity	PPV	NPV
Training cohort						
Radiomics features	0.74 [0.70-0.78]	0.83 [0.79-0.86]	0.86	0.69	0.55	0.91
Clinics features	0.75 [0.71-0.79]	0.86 [0.82-0.90]	0.87	0.69	0.56	0.93
Joint features	0.84 [0.81-0.88]	0.89 [0.86-0.92]	0.74	0.89	0.75	0.89
Validation cohort						
Radiomics features	0.73 [0.66-0.79]	0.81 [0.74-0.87]	0.75	0.72	0.49	0.88
Clinics features	0.73 [0.66-0.79]	0.83 [0.76-0.90]	0.79	0.71	0.49	0.90
Joint features	0.83 [0.78-0.89]	0.88 [0.83-0.94]	0.71	0.88	0.65	0.90

AUC, area under the curve; 95%CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

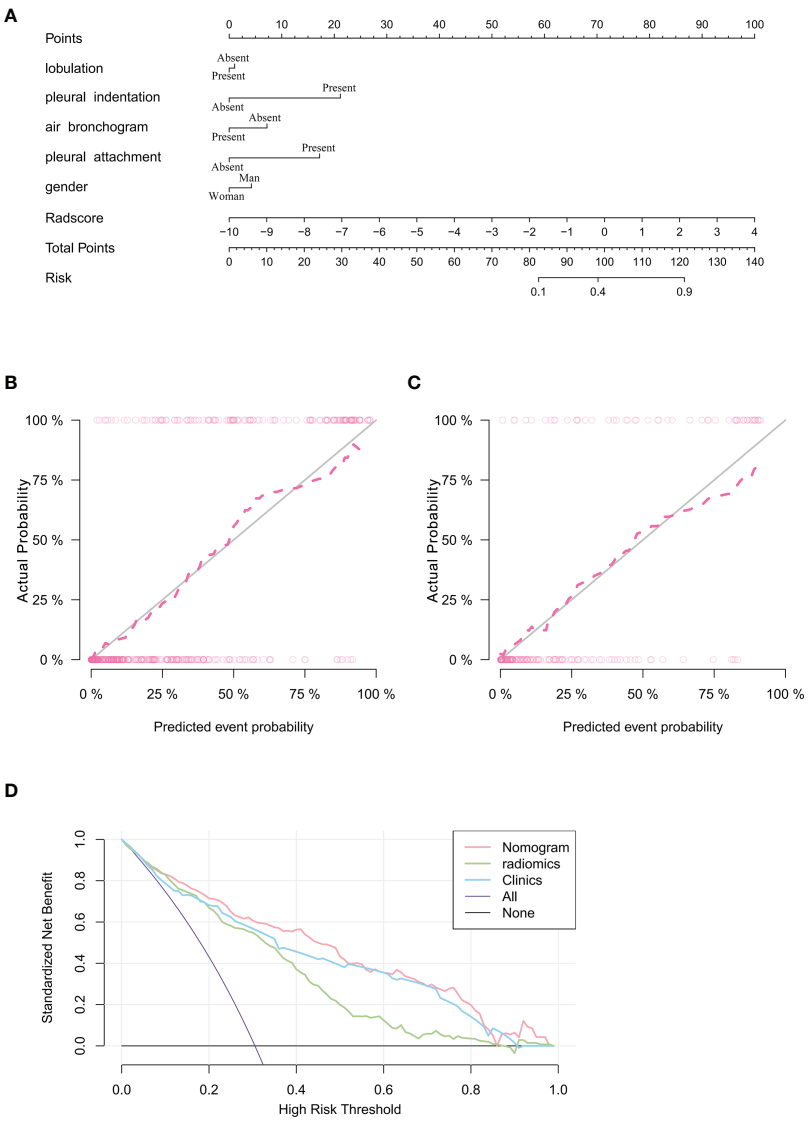


FIGURE 5 | Nomogram for prediction of VPI based on training cohort and the model evaluation of calibration curve. **(A)** Radiomics nomogram based on clinical characteristics and Radscore. The calibration curves were used to evaluate the consistency of the probability of VPI predicted by the nomogram with the actual fraction of visceral pleural invasion in the training **(B)** and validation **(C)** cohorts. **(D)** DCA for the prediction of VPI in lung adenocarcinoma for each model. X-axis represents the threshold probability and Y-axis represents the net benefit. The red curve represents the nomogram. The blue curve represents the clinical features model. The green curve represents the radiomics features model.

focused on the heterogeneity within the tumor, the two models were complementary to each other. Although the diagnostic performances of the radiomic model and the clinical model were similar, the combination of the two models can obtain a higher diagnostic efficiency. The AUC values of the combined model were higher than those of the radiomics and clinical models ($p < 0.05$). Moreover, the DCA results showed that the nomogram was superior to both the clinical features model and the radiomics model for most ranges of reasonable threshold probabilities.

There are several limitations in this study. Firstly, it was a retrospective study and there may have been selection bias. Secondly, tumour serum indicators may be missing due to the small size of the tumour. Thirdly, multiple different CT scanning devices were used, using different acquisition protocols. Multicentre studies should be conduted to validate the reliability of Nomogram.

In summary, a CT image-based nomogram model combining radiomics features and clinical features was developed for predicting VPI in lung adenocarcinoma. A nomogram based on radiomics features may provide a non-invasive method to evaluate the prognosis of early lung adenocarcinoma.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

XZ and CH designed the research. XZ and JB helped collect patient information. YL and XP analyzed data. YL and QW prepared figures and tables. XZ and YL wrote the paper. CH and SH conceived the project and supervised and coordinated all aspects of the work. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by Gusu health talent project of Suzhou (GSWS2020003).

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer Statistics in China, 2015. *CA: Cancer J Clin* (2016) 66:115–32. doi: 10.3322/caac.21338
3. Shimizu K, Yoshida J, Nagai K, Nishimura M, Ishii G, Morishita Y, et al. Visceral Pleural Invasion Is an Invasive and Aggressive Indicator of Non-Small Cell Lung Cancer. *J Thorac Cardiovasc Surg* (2005) 130:160–5. doi: 10.1016/j.jtcvs.2004.11.021
4. Liu Q, Deng X, Zhou D, Li J, Min J, Dai J. Visceral Pleural Invasion Impacts the Prognosis of Non-Small Cell Lung Cancer: A Meta-Analysis. *Eur J Surg Oncol* (2016) 42:1707–13. doi: 10.1016/j.ejso.2016.03.012
5. Rami-Porta R, Bolejack V, Crowley J, Ball D, Kim J, Lyons G, 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* (2015) 10:990–1003. doi: 10.1097/JTO.0000000000000559
6. Rami-Porta R, Bolejack V, Giroux DJ, Chansky K, Crowley J, Asamura H, et al. The IASLC Lung Cancer Staging Project: The New Database to Inform the Eighth Edition of the TNM Classification of Lung Cancer. *J Thorac Oncol* (2014) 9:1618–24. doi: 10.1097/JTO.0000000000000334
7. Hsu J, Han I, Tsai T, Lin S, Jaw T, Liu G, et al. Pleural Tags on CT Scans to Predict Visceral Pleural Invasion of Non-Small Cell Lung Cancer That Does Not Abut the Pleura. *Radiol* (2016) 279:590–6. doi: 10.1148/radiol.2015151120
8. Qi LP, Li XT, Yang Y, Chen JF, Wang J, Chen ML, et al. Multivariate Analysis of Pleural Invasion of Peripheral Non-Small Cell Lung Cancer-Based Computed Tomography Features. *J Comput Assist Tomogr* (2016) 40:757–62. doi: 10.1097/RCT.0000000000000439
9. Ebara K, Takashima S, Jiang B, Numasaki H, Fujino M, Tomita Y, et al. Pleural Invasion by Peripheral Lung Cancer: Prediction With Three-Dimensional CT. *Acad Radiol* (2015) 22:310–9. doi: 10.1016/j.acra.2014.10.002
10. 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
11. 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
12. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep Learning for Lung Cancer Prognostication: A Retrospective Multi-Cohort Radiomics Study. *PLos Med* (2018) 15:e1002711. doi: 10.1371/journal.pmed.1002711
13. Shi L, He Y, Yuan Z, Benedict S, Valicenti R, Qiu J, et al. Radiomics for Response and Outcome Assessment for Non-Small Cell Lung Cancer. *Technol Cancer Res T* (2018) 17:1077050436. doi: 10.1177/1533033818782788
14. Collins LG, Haines C, Perkel R, Enck RE. Lung Cancer: Diagnosis and Management. *Am Fam Phys* (2007) 75:56–63. doi: 10.1186/1471-2296-8-1
15. Yuan M, Liu J, Zhang T, Zhang Y, Li H, Yu T. Prognostic Impact of the Findings on Thin-Section Computed Tomography in Stage I Lung Adenocarcinoma With Visceral Pleural Invasion. *Sci Rep-Uk* (2018) 8:4743. doi: 10.1038/s41598-018-22853-1
16. Zhao L, Xie H, Zhang L, Zha J, Zhou F, Jiang G, et al. Visceral Pleural Invasion in Lung Adenocarcinoma ≤3 Cm With Ground-Glass Opacity: A Clinical, Pathological and Radiological Study. *J Thorac Dis* (2016) 8:1788–97. doi: 10.21037/jtd.2016.05.90
17. Hattori A, Suzuki K, Matsunaga T, Takamochi K, Oh S. Visceral Pleural Invasion Is Not a Significant Prognostic Factor in Patients With a Part-Solid Lung Cancer. *Ann Thorac Surg* (2014) 98:433–8. doi: 10.1016/j.athoracsur.2014.04.084
18. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability. *Neuroimage* (2006) 31:1116–28. doi: 10.1016/j.neuroimage.2006.01.015
19. Collewet G, Strzelecki M, Mariette F. Influence of MRI Acquisition Protocols and Image Intensity Normalization Methods on Texture Classification. *Magn Reson Imaging* (2004) 22:81–91. doi: 10.1016/j.mri.2003.09.001
20. Jin C, Cao J, Cai Y, Wang L, Liu K, Shen W, et al. A Nomogram for Predicting the Risk of Invasive Pulmonary Adenocarcinoma for Patients With Solitary Peripheral Subsolid Nodules. *J Thorac Cardiovasc Surg* (2017) 153:462–9. doi: 10.1016/j.jtcvs.2016.10.019
21. Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. *Theranostics* (2019) 9:1303–22. doi: 10.7150/thno.30309
22. Vickers AJ, van Calster B, Steyerberg EW. A Simple, Step-by-Step Guide to Interpreting Decision Curve Analysis. *Diagn Prognostic Res* (2019) 3:18. doi: 10.1186/s41512-019-0064-7

23. Nicholson AG, Chansky K, Crowley J, Beyruti R, Kubota K, Turrisi A, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals for the Revision of the Clinical and Pathologic Staging of Small Cell Lung Cancer in the Forthcoming Eighth Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol* (2016) 11:300–11. doi: 10.1016/j.jtho.2015.10.008
24. Schuurbiers OC, Meijer TW, Kaanders JH, Looijen-Salamon MG, de Geus-Oei LF, van der Drift MA, et al. Glucose Metabolism in NSCLC Is Histology-Specific and Diverges the Prognostic Potential of 18FDG-PET for Adenocarcinoma and Squamous Cell Carcinoma. *J Thorac Oncol* (2014) 9:1485–93. doi: 10.1097/JTO.0000000000000286
25. Zhu X, Dong D, Chen Z, Fang M, Zhang L, Song J, et al. Radiomic Signature as a Diagnostic Factor for Histologic Subtype Classification of Non-Small Cell Lung Cancer. *Eur Radiol* (2018) 28:2772–8. doi: 10.1007/s00330-017-5221-1
26. Wu W, Parmar C, Grossmann P, Quackenbush J, Lambin P, Bussink J, et al. Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology. *Front Oncol* (2016) 6:71. doi: 10.3389/fonc.2016.00071
27. Yoshida J, Nagai K, Asamura H, Goya T, Koshiishi Y, Sohara Y, et al. Visceral Pleura Invasion Impact on Non-Small Cell Lung Cancer Patient Survival: Its Implications for the Forthcoming TNM Staging Based on a Large-Scale Nation-Wide Database. *J Thorac Oncol* (2009) 4:959–63. doi: 10.1097/JTO.0b013e3181a85d5e
28. Kawase A, Yoshida J, Ishii G, Hishida T, Nishimura M, Nagai K. Visceral Pleural Invasion Classification in Non-Small Cell Lung Cancer. *J Thorac Oncol* (2010) 5:1784–8. doi: 10.1097/JTO.0b013e3181eedd9c
29. Wu LL, Liu X, Jiang WM, Huang W, Lin P, Long H, et al. Stratification of Patients With Stage IB NSCLC Based on the 8th Edition of the American Joint Committee on Cancer (AJCC) Staging Manual. *Front Oncol* (2020) 10:571. doi: 10.3389/fonc.2020.00571
30. Heidinger BH, Schwarz-Nemec U, Anderson KR, de Margerie-Mellon C, Monteiro Filho AC, Chen Y, et al. Visceral Pleural Invasion in Pulmonary Adenocarcinoma: Differences in CT Patterns Between Solid and Subsolid Cancers. *Radiol Cardiothorac Imaging* (2019) 1:e190071. doi: 10.1148/ryct.2019190071
31. Kang J, Kim KD, Chung KY. Prognostic Value of Visceral Pleura Invasion in Non-Small Cell Lung Cancer. *Eur J Cardio-Thorac Surg* (2003) 23:865–9. doi: 10.1016/S1010-7940(03)00119-2
32. Woodard GA, Wang SX, Kratz JR, Zoon-Besselink CT, Chiang C, Gubens MA, et al. Adjuvant Chemotherapy Guided by Molecular Profiling and Improved Outcomes in Early Stage, Non-Small-Cell Lung Cancer. *Clin Lung Cancer* (2018) 19:58–64. doi: 10.1016/j.cllc.2017.05.015
33. Moon H, Zhao Y, Pluta D, Ahn H. Subgroup Analysis Based on Prognostic and Predictive Gene Signatures for Adjuvant Chemotherapy in Early-Stage Non-Small-Cell Lung Cancer Patients. *J Biopharm Stat* (2018) 28:750–62. doi: 10.1080/10543406.2017.1397006
34. Williams CD, Gajra A, Ganti AK, Kelley MJ. Use and Impact of Adjuvant Chemotherapy in Patients With Resected Non-Small Cell Lung Cancer. *Cancer-Am Cancer Soc* (2014) 120:1939–47. doi: 10.1002/cncr.28679
35. Ahn SY, Park CM, Jeon YK, Kim H, Lee JH, Hwang EJ, et al. Predictive CT Features of Visceral Pleural Invasion by T1-Sized Peripheral Pulmonary Adenocarcinomas Manifesting as Subsolid Nodules. *AJR Am J Roentgenol* (2017) 209:561–6. doi: 10.2214/AJR.16.17280
36. Kim HJ, Cho JY, Lee YJ, Park JS, Cho YJ, Yoon HI, et al. Clinical Significance of Pleural Attachment and Indentation of Subsolid Nodule Lung Cancer. *Cancer Res Treat* (2019) 51:1540–8. doi: 10.4143/crt.2019.057
37. Erasmus JJ, Connolly JE, McAdams HP, Roggli VL. Solitary Pulmonary Nodules: Part I. Morphologic Evaluation for Differentiation of Benign and Malignant Lesions. *Radiographics* (2000) 20:43–58. doi: 10.1148/radiographics.20.1
38. Lederlin M, Puderbach M, Muley T, Schnabel PA, Stenzinger A, Kauczor HU, et al. Correlation of Radio- and Histomorphological Pattern of Pulmonary Adenocarcinoma. *Eur Respir J* (2013) 41:943–51. doi: 10.1183/09031936.00056612
39. Wang JC, Sone S, Feng L, Yang ZG, Takashima S, Maruyama Y, et al. Rapidly Growing Small Peripheral Lung Cancers Detected by Screening CT: Correlation Between Radiological Appearance and Pathological Features. *Br J Radiol* (2000) 73:930–7. doi: 10.1259/bjr.73.873.11064644
40. Antony VB, Sahn SA, Mossman B, Gail DB, Kalica A. NHLBI Workshop Summaries. Pleural Cell Biology in Health and Disease. *Am Rev Respir Disease* (1992) 145:1236–9. doi: 10.1164/ajrccm/145.5.1236
41. Yamamoto T, Kadoya N, Shirata Y, Koto M, Sato K, Matsushita H, et al. Impact of Tumor Attachment to the Pleura Measured by a Pretreatment CT Image on Outcome of Stage I NSCLC Treated With Stereotactic Body Radiotherapy. *Radiat Soncol (Lond Engl)* (2015) 10:35. doi: 10.1186/s13014-015-0343-6
42. Wang H, Schabath MB, Liu Y, Berglund AE, Bloom GC, Kim J, et al. Semiquantitative Computed Tomography Characteristics for Lung Adenocarcinoma and Their Association With Lung Cancer Survival. *Clin Lung Cancer* (2015) 16:e141-63. doi: 10.1016/j.cllc.2015.05.007
43. Weng Q, Hui J, Wang H, Lan C, Huang J, Zhao C, et al. Radiomic Feature-Based Nomogram: A Novel Technique to Predict EGFR-Activating Mutations for EGFR Tyrosin Kinase Inhibitor Therapy. *Front Oncol* (2021) 11:590937. doi: 10.3389/fonc.2021.590937

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 04 July 2022
doi: 10.3389/fonc.2022.924223

Comparison of ^{68}Ga-FAPI and ^{18}F-FDG PET/CT in the Evaluation of Patients With Newly Diagnosed Non-Small Cell Lung Cancer

Junhao Wu[1,2,3†], *Hao Deng*[1,2,3†], *Haoshu Zhong*[4,5†], *Tao Wang*[6], *Zijuan Rao*[1,2,3], *Yingwei Wang*[1,2,3], *Yue Chen*[1,2,3] *and Chunyin Zhang*[1,2,3*]

[1] Department of Nuclear Medicine, The Affiliated Hospital of Southwest Medical University, Luzhou, China, [2] Nuclear Medicine and Molecular Imaging Key Laboratory of Sichuan Province, Luzhou, China, [3] Academician (Expert) Workstation of Sichuan Province, Luzhou, China, [4] Department of Hematology, Clinical Medicine, Affiliated Hospital of Southwest Medical University, Luzhou, China, [5] Stem Cell Laboratory, The Clinical Research Institute, Affiliated Hospital of Southwest Medical University, Luzhou, China, [6] Department of the General Surgery, The Affiliated Hospital of Southwest Medical University, Luzhou, China

Purpose: Several studies have demonstrated that ^{68}Ga-FAPI PET/CT shows high intratumoral tracer uptake and low normal tissue uptake, allowing for excellent visualization of cancer. The purpose of this study was to compare the ability of ^{68}Ga-FAPI and ^{18}F-FDG PET/CT for the evaluation of newly diagnosed NSCLC.

Materials and Methods: A prospective analysis of 28 individuals with histopathologically newly confirmed NSCLC that underwent ^{68}Ga-FAPI and ^{18}F-FDG PET/CT was conducted. The performance of two imaging modalities was compared based upon visual assessment, rates of cancer detection, and semi-quantitative parameters (target-to-background ratio [TBR], maximum standard uptake value [SUVmax]) for both primary tumors and metastases.

Results: In total, this study enrolled 28 participants (13 male, 15 female; median age: 60.5 years, range: 34 – 78 years. <u>For primary tumors, ^{68}Ga-FAPI and ^{18}F-FDG PET/CT have similar detection performance (28 vs. 27). However, ^{68}Ga-FAPI PET/CT was found to more effectively evaluate most metastases as compared to ^{18}F-FDG PET/CT. ^{68}Ga-FAPI PET/CT detecting more metastases present within the lymph nodes (53 vs. 49), pleura (8 vs. 7), liver (4 vs. 1), and bone (41 vs. 35).</u> The SUVmax and TBR values for ^{68}Ga-FAPI were substantially superior to those for ^{18}F-FDG in lymph node, pleural, and bone metastases. While the SUVmax for these two imaging approaches was comparable for hepatic metastases, ^{68}Ga-FAPI exhibited a significantly higher TBR in relation to that of ^{18}F-FDG. In addition, ^{68}Ga-FAPI PET/CT demonstrates excellent N (80% [8/10]) and M (92.9% [26/28]) staging accuracy in NSCLC patients.

OPEN ACCESS

Edited by:
Ryogo Minamimoto, National Center For Global Health and Medicine, Japan

Reviewed by:
Hubing WU, Southern Medical University, China
Filippo Lococo, Agostino Gemelli University Polyclinic (IRCCS), Italy

***Correspondence:**
Chunyin Zhang
zhangchunyin345@sina.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *20 April 2022*
Accepted: *03 June 2022*
Published: *04 July 2022*

Citation:
Wu J, Deng H, Zhong H, Wang T, Rao Z, Wang Y, Chen Y and Zhang C (2022) Comparison of ^{68}Ga-FAPI and ^{18}F-FDG PET/CT in the Evaluation of Patients With Newly Diagnosed Non-Small Cell Lung Cancer. Front. Oncol. 12:924223. doi: 10.3389/fonc.2022.924223

Conclusions: ^{68}Ga-FAPI PET/CT as an examination modality is excellent for evaluation of newly diagnosed NSCLC. ^{68}Ga-FAPI PET/CT improves the detection rates of most metastases and facilitating the superior staging of patients with newly diagnosed NSCLC, relative to that achieved by ^{18}F-FDG PET/CT.

Keywords: 68 Ga-FAPI, 18 F-FDG, non-small-cell lung cancer (NSCLC), fibroblast activation protein (FAP), metastases

INTRODUCTION

Cancer is one of the fundamental threats to human health and well-being, with lung cancer in particular remaining among the most common and deadliest tumors (1). Lung cancer is a heterogeneous classification of epithelial malignancy with a range of pathological and clinical manifestations. Broadly speaking, lung cancer cases are subdivided into non-small-cell lung cancer (NSCLC) and small cell lung cancer (SCLC) (2–4). For individuals with stage I – IIIA NSCLC, surgical resection is the optimal therapeutic intervention, but just 20-25% of patients are suited to undergo curative surgical resection (2, 3, 5, 6). The eligibility of newly diagnosed patients for such treatment is generally dependent on the degree of tumor involvement such that accurate tumor staging is essential and can affect both the prognostic evaluation and treatment of patients (5, 7). 18F-fluorodeoxyglucose (18F-FDG) positron emission tomography/computed tomography (PET/CT) imaging has emerged as the most widely used modality for diagnosing and systemically staging NSCLC. However, the utility of this approach can be limited by insufficient soft-tissue contrast and by elevated levels of physiological background activity in specific organs (1, 8, 9). Cancer-associated fibroblasts (CAF) are commonly linked to a poor cancer patient prognosis (10–16). CAFs frequently express elevated levels of the type II transmembrane serine protease fibroblast-activated protein (FAP) (17–19), which plays key roles in migratory, invasive, and angiogenic activity in oncogenic contexts (20–24). Recently, novel quinoline FAP-specific inhibitor-based PET tracers have been developed that can be used to precisely target fibrotic and tumor-associated stromal tissue (19, 25, 26). 68Ga-FAPI PET/CT exhibits a high degree of intratumoral tracer uptake, low normal tissue uptake, and rapid clearance, thus resulting in excellent tumor visibility and a great target to background ratio (26–29). In multiple recent research (18, 19, 25, 27, 28, 30–34), 68Ga-FAPI PET/CT was demonstrated to aid in the visualization of a diversity of tumors in addition to offering clear advantages as compared to 18F-FDG PET/CT when discerning lymph node, pleural, brain, and bone metastases.

Current research advances suggest that 68Ga-FAPI may be a more accurate and convenient alternative to 18F-FDG PET/CT for the diagnosis and staging of lung cancer. Therefore, this study was conducted to examine the performance of 68Ga-FAPI and 18F-FDG PET/CT for the evaluation of newly diagnosed NSCLC.

MATERIALS AND METHODS

Patients

The Ethics Committee of Southwest Medical University Hospital approved the present study, which was conducted from July 2020 - October 2021 (Ethics committee approval No.: 2020035), and all patients signed a written informed consent form. The inclusion criteria for this study were as follows: (1) individuals ≥ 18 years of age; (2) individuals newly diagnosed with NSCLC that had not undergone any previous antitumor treatment; (3) individuals who underwent both ^{68}Ga-FAPI and ^{18}F-FDG PET/CT at a 1-week interval. Contributors were excluded if they: (1) underwent < 3 months of follow-up; (2) had undergone antitumor treatment prior to PET/CT imaging; or (3) harbored any other non-NSCLC primary tumors.

PET/CT Imaging

Contributors were asked to fast, not received intravenous glucose, and avoid strenuous activity or prolonged exercise for a minimum of 6 h before intravenous ^{18}F-FDG (3.7 MBq/kg) infusion, and patients also needed to have normal blood glucose levels. ^{68}Ga-FAPI injection (1.85–2.59 MBq/kg) did not necessitate any specific fasting or glycemic preparation. A hybrid PET/CT scanner (uMI780, United Imaging Healthcare, Shanghai, China) was used to conduct all PET/CT imaging ~1 h following radiotracer administration. With the contributor's arms raised above their head, an initial spiral CT scan was conducted from the top of the skull to the upper portion of the mid-thigh (current 120 mA; tube voltage 120 kV; matrix 512 × 512 pixels; slice thickness 3.00 mm; window width 300–500 HU; window level 40–60 HU). PET scanning was subsequently conducted using the same bed position utilized for CT scanning, with 1.5 min/position in 3D acquisition mode and 5–6 bed positions. The resultant outcomes were transferred to a post-processing workstation (v R002, uWS-MI, United Imaging Healthcare, Shanghai, China). PET attenuation correction was performed using CT data, with PET data reconstruction being conducted based upon an ordered subset estimation maximization algorithm (20 subsets, 2 iterations). The overall condition of each case, such as their body temperature, heart rate, blood pressure, and mental status, was assessed by a nuclear medicine physician within 2 h following injection.

Image Review

Two experienced nuclear medicine physicians independently conducted visual, qualitative, and semi-quantitative

interpretation of all ^{18}F-FDG and ^{68}Ga-FAPI PET/CT. Discrepancies were resolved through discussion and consensus. Patient PET/CT images were assessed in the coronal, axial, and sagittal planes. Positive lesions were identified by areas of non-physiological uptake above background in ^{68}Ga-FAPI or ^{18}F-FDG PET images. Positive lesions were combined with data from the corresponding CT scan images for further diagnosis, and their length were measured and recorded. Positive PET/CT lesions were further categorized as non-malignant lesions, primary tumors, distant metastases, or lymph node metastases. ^{18}F-FDG PET/CT and ^{68}Ga-FAPI imaging results were initially compared *via* a visual assessment in which the two images for each patient were assessed to establish their relative inferiority or superiority when detecting primary tumors (based upon tumor size and conspicuousness) and metastatic lesions (based upon numbers, involvement, and conspicuousness). Semi-quantitative analyses were then conducted by comparing ^{18}F-FDG and ^{68}Ga-FAPI radiotracer uptake within the same lesions. SUVmax was measured using the analytical workstation after the region of interest (ROI) surrounding the lesion had been defined by a physician. The TBR was defined as the difference in radiotracer uptake between the lesion and background, and was measured *via* dividing the SUVmax for a given lesion by the mean normalized uptake (SUVmean) for normal background tissue.

Diagnostic Criteria

Histopathological findings were used for final diagnostic determinations for all primary tumors. When histopathological results were not available for metastases, final diagnosis was made based upon the results derived from multiple imaging modalities (MRI, enhanced CT, ultrasound, bone scan, PET/CT) and corresponding follow-up imaging. During follow-up, a suspicious lesion was considered to be malignant if it exhibited progressive growth or the number and/or size of suspect lesions declined following antitumor treatment.

Statistical Analysis

Statistical evaluations were executed using SPSS (v 26.0; IBM, NY, USA). General data were compared through descriptive analyses, with categorical variables being listed as numbers with percentages, while continuous variables were listed as the mean ± SD. Chi-squared tests were used to compare numbers of positive lesions. Student's t-tests were employed for comparing SUVmax and TBR values for specific lesions associated with ^{68}Ga-FAPI and ^{18}F-FDG PET/CT. Correlations between lesion length and metabolic parameters (TBR and SUVmax) were assessed through Spearman's rank correlation analyses. A two-tailed $P < 0.05$ was the threshold of significance.

Results

Generally, this study enrolled 28 cases (13 male, 15 female; median age: 60.5 years, range: 34-78 years). The basic features of these cases are detailed in **Table 1**.

The patients had been newly diagnosed with NSCLC, including 24 patients diagnosed with adenocarcinomas and 5 diagnosed with squamous cell carcinomas, with one patient (patient 17) having been simultaneously diagnosed with two primary tumors. In total, 16 patients underwent surgical resection, with 10 having simultaneously undergone mediastinal lymph node dissection. The remaining 12 patients underwent non-surgical antitumor treatment.

Adverse Event

No patients developed any adverse events, discomfort, or abnormalities with respect to heart rate, body temperature, blood pressure, or mental status within 2 h following imaging agent injection.

Comparison of Visual Assessment Outcomes

Upon visual assessment, ^{68}Ga-FAPI PET/CT enabled clearer metastatic and primary tumor visualization as compared to ^{18}F-FDG PET/CT in a majority of patients. Specifically, ^{68}Ga-FAPI outperformed ^{18}F-FDG PET/CT for the visual evaluation of primary tumors (14/28 [50.0%] vs. 9/28 [32.1%]) (**Figure 1**), lymph node metastases (9/15 [60.0%] vs. 5/15 [33.3%]), pleural metastases (3/3 [100.0%] vs. 0/3 [0%]) (**Figure 2**), hepatic metastases (2/2 [100.0%] vs. 0/2 [0%]), and bone metastases (6/7 [85.7%] vs. 0/7 [0%]), but it performed less effectively for pulmonary (0/2 [0%] vs. 2/2 [100.0%]) and adrenal metastases (0/2 [0%] vs. 2/2 [100.0%]) (**Figure 3**).

Lesion Detection Analysis

^{68}Ga-FAPI PET/CT outperformed ^{18}F-FDG PET/CT in a lesion-based analysis when detecting hepatic (100% [4/4] vs. 25% [1/4]) and bone metastases (97.6% [41/42] vs. 83.3% [35/42]) (**Figure 4**), whereas ^{68}Ga-FAPI was inferior to ^{18}F-FDG PET/CT when utilized to detect adrenal metastases (0% [0/2] vs. 100% [2/2]). ^{68}Ga-FAPI and ^{18}F-FDG PET/CT performed similarly when used to detect primary tumors (96.6% [28/29] vs. 93.1% [27/29]), as well as lymph node (93.0% [53/57] vs. 86.0% [49/57]), pulmonary (100% [3/3] vs. 100% [3/3]), and pleural metastases (100% [8/8] vs. 87.5% [7/8]) (**Table 2**).

Comparison of Different Pathological Types

Evaluation of metabolism of primary tumors and lymph node metastases based on pathological type. For primary lung adenocarcinoma, there was no statistically significant difference in SUVmax (9.4 ± 4.8 vs. 8.7 ± 6.2, P = 0.572) and TBR (26.5 ± 19.9 vs. 19.7 ± 18.6, P = 0.131) between ^{68}Ga-FAPI and ^{18}F-FDG PET/CT. For primary lung squamous cell carcinoma, there was also no statistically significant difference in SUVmax (9.0 ± 4.1 vs. 15.8 ± 8.0, P = 0.156) and TRB (25.5 ± 14.1 vs. 44.2 ± 25.6, P = 0.16) between the two examination. For lymph node metastasis, The SUVmax (8.4 ± 4.3 vs. 5.9 ± 4.2, P =0.001) and TBR (10.8 ± 6.4 vs. 5.7 ± 4.6, P = 0.001) of lymph node metastases from adenocarcinoma were significantly higher in ^{68}Ga-FAPI than in ^{18}F-FDG PET/CT. In contrast, SUVmax (9.0 ± 5.3 vs. 11.2 ± 6.9, P = 0.077) and TBR (9.1 ± 5.2 vs. 10.1 ± 6.3, P = 0.227) for lymph node metastases from squamous cell carcinoma were not statistically significantly different between the two examination modalities.

TABLE 1 | Basic patient characteristics.

NO.	Sex	Age	Pathology	Primary tumor site	Length (cm)	Metastases site	Staging
1	F	44	ADC	right upper lobe	5.9	LNM; LM; PM	IVA
2	M	61	ADC	left upper lobe	2.7	LNM; AM	IVA
3	M	66	ADC	right upper lobe	1.2	None	IA
4	F	46	ADC	left upper lobe	1.8	LNM; BM	IVB
5	F	48	SCC	left lower lobe	2.2	None	IA
6	F	57	ADC	left lower lobe	3.2	LNM; HM; BM	IVB
7	F	53	ADC	left upper lobe	1.1	None	IA
8	F	72	ADC	right upper lobe	1.2	None	IA
9	M	70	ADC	right upper lobe	2.3	None	IA
10	F	78	ADC	left lower lobe	8.1	LM; PM; BM	IVA
11	M	68	ADC	left upper lobe	2.1	LNM	IIIA
12	F	57	ADC	right middle lobe	2.8	LNM	IIIB
13	M	69	SCC	right lower lobe	3	None	IA
14	M	49	ADC	right upper/lower lobe	8.9	LNM, Pancreas, Kidney	IVB
15	F	46	ADC	right middle lobe	3.2	LNM	IIIB
16	M	63	ADC	left lower lobe	3.1	BM	IVB
17	F	68	ADC	right lower lobe	1.9	LNM	IIIA
			ADC	right middle lobe	1.8		
18	M	63	ADC	left upper lobe	1.2	LNM;HM; BM	IVB
19	M	71	SCC	right upper lobe	3.3	LNM; BM	IVA
20	M	67	SCC	right upper lobe	2.7	AM	IVA
21	M	34	ADC	right lower lobe	3.3	LNM; BM	IVB
22	F	58	ADC	right lower lobe	3.1	LNM; PM	IVA
23	F	61	ADC	left upper lobe	2.2	LNM	IIB
24	F	60	ADC	right upper lobe	3.4	None	IB
25	M	56	SCC	left lower lobe	4.3	LNM	IIIA
26	F	45	ADC	right middle lobe	1.9	None	IA
27	F	53	ADC	right upper lobe	2.3	None	IA
28	M	68	ADC	right upper lobe	2.5	None	IA

SCC, squamous cell carcinoma; ADC, adenocarcinoma; LNM, lymph node metastasis; LM, lung metastasis; PM, Pleural metastasis; AM, adrenal metastasis; BM, bone metastasis; HM, hepatic metastases.

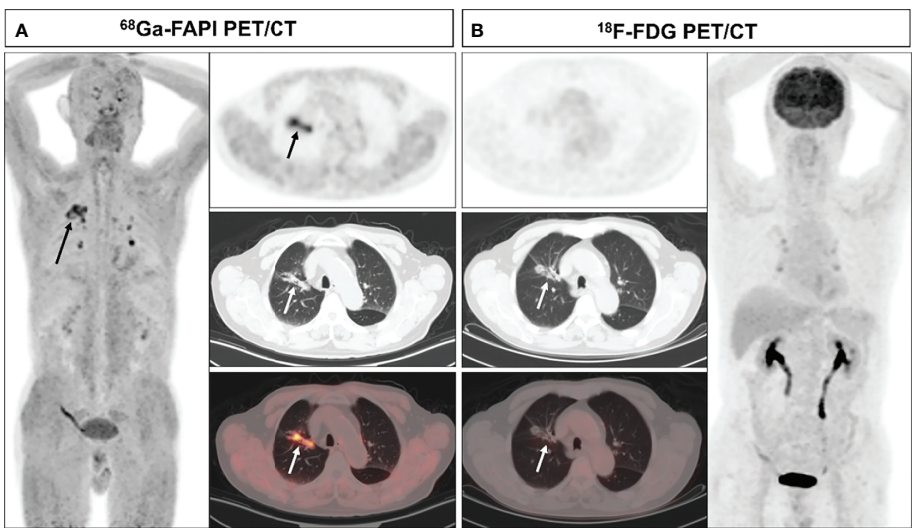


FIGURE 1 | A 70-year-old male (patient 9) diagnosed with adenocarcinoma. ^{68}Ga-FAPI PET/CT **(A)** revealed an adenocarcinoma lesion with increased FAPI uptake (solid arrows, SUVmax=6.3), while ^{18}F-FDG PET/CT did not reveal any significant uptake in the primary lesion **(B)** solid arrows).

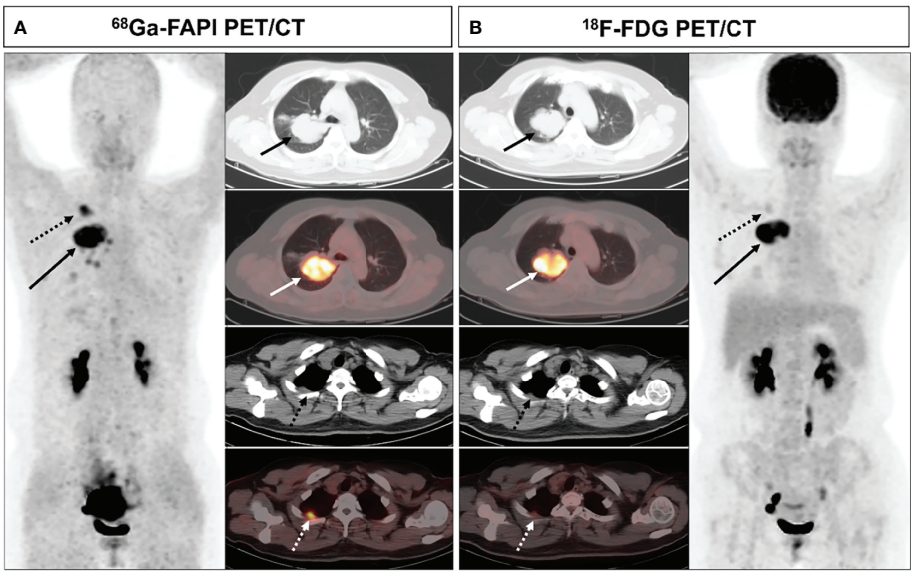


FIGURE 2 | A 44-year-old female (patient 1) diagnosed with adenocarcinoma. ^{68}Ga-FAPI PET/CT **(A)** revealed increased FAPI uptake in the primary lesion (solid arrows, SUVmax = 11.7) and pleural lesion (dashed arrows, SUVmax = 7.0). ^{18}F-FDG PET/CT **(B)** also showed high FDG uptake in the primary lesion (solid arrows, SUVmax = 12.4), while the pleural lesion with only mild FDG uptake (dashed arrow, SUVmax = 2.2). The pleural lesion was deemed likely to be metastatic, as confirmed upon subsequent follow-up.

Comparison of Semi-Quantitative Parameters

The SUVmax and TBR values for ^{68}Ga-FAPI and ^{18}F-FDG PET/CT did not differ significantly when used for detecting primary tumors, pulmonary metastases, and adrenal metastases, while the SUVmax and TBR of ^{68}Ga-FAPI were substantially superior to those for ^{18}F-FDG PET/CT when used to detect lymph node, pleural, and bone metastases. Although there was no significant difference in SUVmax between these two imaging modalities in detecting liver metastases (P = 0.062), ^{68}Ga-FAPI had significantly greater TBR values relative to ^{18}F-FDG (P = 0.027) (**Table 2**).

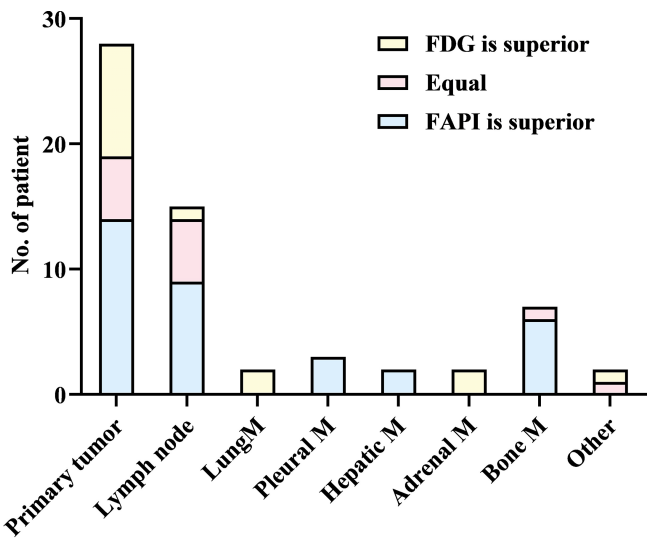


FIGURE 3 | Visual assessment comparison for ^{68}Ga-FAPI and ^{18}F-FDG PET/CT. M = metastases.

TABLE 2 | Comparison of ^{68}Ga-FAPI and ^{18}F-FDG PET/CT semi-quantitative imaging parameters.

Parameter	Imaging method	Primary tumor	Lymph node metastasis	Lung metastasis	Pleural metastasis	Hepatic metastasis	Adrenal metastasis	Bone metastasis
Number of lesions		29	57	3	8	4	2	42
Positive detection	^{68}Ga-FAPI	28	53	3	8	4	0	41
	^{18}F-FDG	27	49	3	7	1	2	35
	P	0.554	0.222	1.000	0.302	0.028	0.046	0.026
SUVmax	^{68}Ga-FAPI	9.3 ± 4.6	8.4 ± 4.3	2.4 ± 1.6	10.8 ± 3.6	6.2 ± 2.1	1.2 ± 0.4	11.2 ± 5.5
	^{18}F-FDG	9.9 ± 6.9	6.4 ± 4.7	2.9 ± 1.9	5.5 ± 3.0	3.4 ± 0.27	6.4 ± 3.3	6.5 ± 3.9
	P	0.631	0.003	0.192	<0.001	0.062	0.237	<0.001
TBR	^{68}Ga-FAPI	26.3 ± 18.8	10.6 ± 6.3	3.4 ± 1.8	9.1 ± 2.8	11.4 ± 5.3	1.6 ± 0.1	16.2 ± 11.2
	^{18}F-FDG	24.0 ± 21.6	6.1 ± 4.9	4.8 ± 3.2	6.2 ± 3.3	1.3 ± 0.3	3.3 ± 2.5	5.9 ± 5.8
	P	0.589	<0.001	0.215	0.001	0.027	0.500	<0.001

The Relationship Between Lesion Length and Suvmax

Significant correlations between lesion length and FAPI-SUVmax were noted for primary tumors, lymph node metastases, and bone metastases in Spearman's correlation analyses, while FDG-SUVmax values were only correlated with lesion length for primary tumors and lymph node metastases but not for bone metastases (**Figure 5**).

N and M Staging

Ten of all patients underwent mediastinal lymph node dissection. A total of 180 lymph nodes underwent pathological biopsy, of which 11 were malignant and 169 were benign. The sensitivity and specificity of ^{68}Ga-FAPI and ^{18}F-FDG PET/CT for detecting lymph node metastasis were 81.8% (9/11), 97.6% (165/169) and 72.7% (8/11), 88.8% (150/169), respectively (**Figure 6**). ^{68}Ga-FAPI PET/CT led to a lower N-stage in 1 patient owing to overlooked lymph node metastases and a higher N-stage in 1 patient owing to the detection of additional false-positive lymph nodes. In contrast, ^{18}F-FDG PET/CT detected additional false-positive lymph nodes in 5 patients resulting in higher N staging for these individuals. Overall, N-staging of NSCLC patients based on ^{68}Ga-FAPI-FAPI results was more accurate than ^{18}F-FDG PET/CT results in these same patients (80% [8/10] vs. 50% [5/10]), but there was no significant difference between the two values (p=0.16).

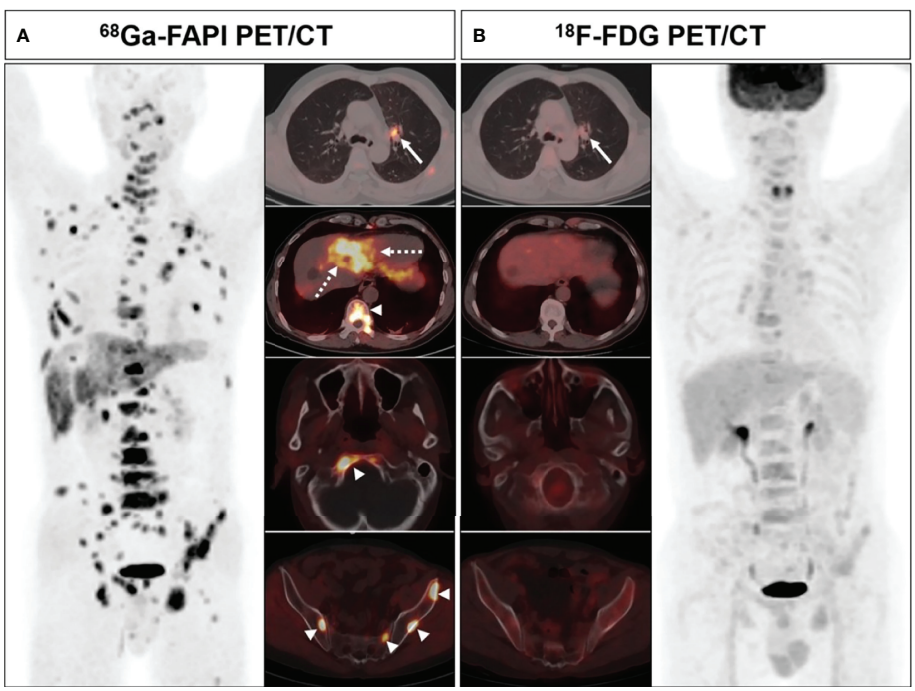


FIGURE 4 | A 63-year-old male (patient 18) diagnosed with adenocarcinoma. ^{68}Ga-FAPI PET/CT **(A)** showed intense tracer uptake in the primary tumor (solid arrows, SUVmax=10.0), liver metastasis (dashed arrows, SUVmax=7.6) and bone metastases (arrows, SUVmax=8.3-8.5). ^{18}F-FDG PET/CT **(B)** showed primary lesion with mild FDG uptake (solid arrows, SUVmax=3.6), while no significant FDG uptake was showed in liver metastasis and multiple bone metastases.

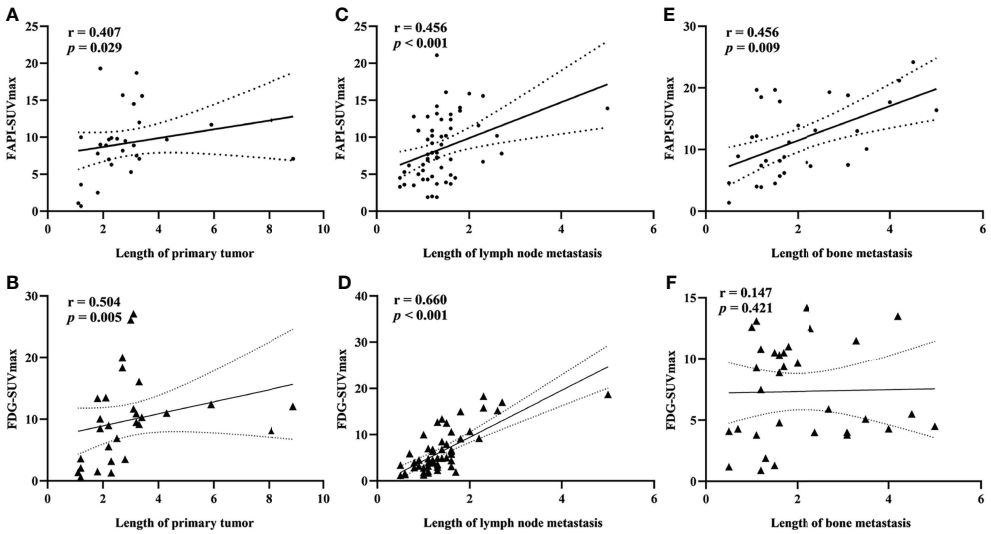


FIGURE 5 | Spearman rank correlation analysis of the relationship between SUVmax value and lesion length for primary tumors **(A)**. FAPI-SUVmax; **(B)** FDG-SUVmax), lymph node metastases **(C)**. FAPI-SUVmax; **(D)** FDG-SUVmax), and bone metastases **(E)**. FAPI-SUVmax; **(F)** FDG-SUVmax).

Distant metastatic lesions were found in 12 of all patients (42.9%). ^{68}Ga-FAPI PET/CT failed to detect adrenal metastases in two patients, resulting in decreased M stage. ^{18}F-FDG PET/CT resulted in incorrect M staging due to 1 false positive adrenal lesion and 2 false negatives for bone metastases. The overall M-staging accuracy of ^{68}Ga-FAPI PET/CT was slightly higher than that of ^{18}F-FDG PET/CT (92.9% [26/28] vs. 89.3% [25/28]), but the values were not statistically different between them (P=0.64).

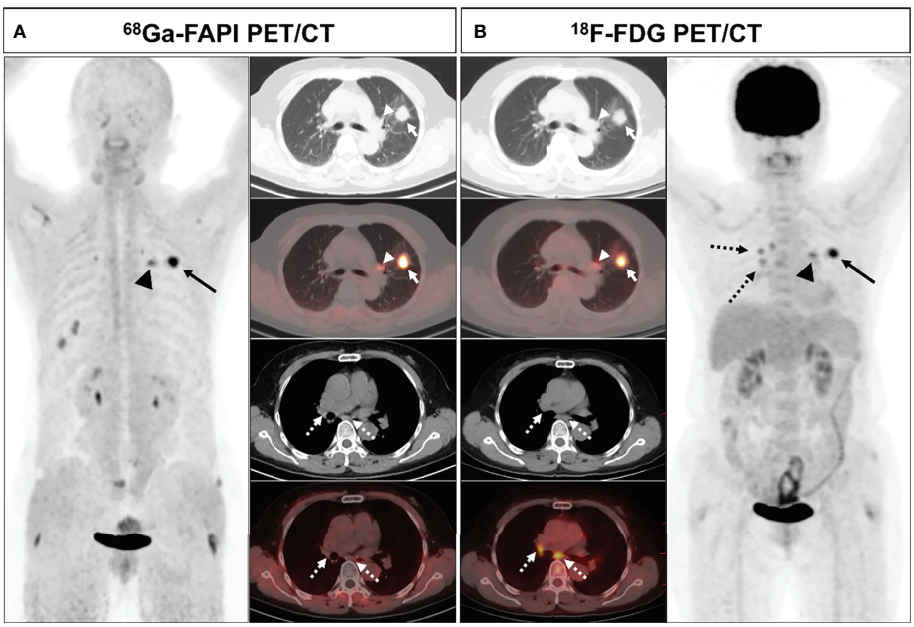


FIGURE 6 | A 61-year-old female (patient 23) diagnosed with adenocarcinoma. ^{68}Ga-FAPI PET/CT **(A)** revealed intense FAPI uptake in the primary tumor (solid arrows, SUVmax = 9.7) along with moderately increased uptake in the left pulmonary hilar lymph node (arrows, SUVmax = 5.0), whereas there was no FAPI uptake in the subcarinal and right pulmonary hilar lymph nodes (dashed arrows). ^{18}F-FDG PET/CT **(B)** demonstrated intense FDG uptake in the primary tumor (solid arrows, SUVmax = 9.0) with moderate uptake in the left pulmonary hilar (arrows, SUVmax = 4.5), subcarinal, and right pulmonary hilar (dashed arrows, SUVmax = 3.5) lymph nodes. Pathological biopsy confirmed metastasis in the left hilar lymph node, while no metastasis was found in the subcardiac or right hilar lymph nodes.

DISCUSSION

The current exploration is to explore the comparative performance of ^{68}Ga-FAPI and ^{18}F-FDG PET/CT in the evaluation of patients with newly diagnosed NSCLC. Overall, these results show that ^{68}Ga-FAPI PET/CT provides better lesion visualization and staging accuracy than ^{18}F-FDG PET/CT in NSCLC.

One recent analysis (31) comparing ^{68}Ga-FAPI and ^{18}F-FDG reported no significant differences between these two technologies with respect to primary lung cancer detection rates or associated SUVmax or TBR values, in line with our findings. In contrast, Wang et al. (35) reported that ^{68}Ga-FAPI yielded significantly higher SUVmax and TBR values as compared to ^{18}F-FDG PET/CT, leading them to conclude that this former technology is better suited to the detection of early-stage lung cancer. AS their analysis specifically included individuals with large tumors (Mean size: 3.3 cm) and advanced disease, this may account for their inconsistent results. In addition, no significant differences in SUVmax and TBR were found in ^{68}Ga-FAPI and ^{18}F-FDG PET/CT for different pathological subtypes of primary tumors.

At present, surgical tumor resection is the benchmark of care for early-stage NSCLC patients. The capability of predicting and detecting regional lymph node metastases in these patients performs a central task in treatment planning and associated management efforts (36, 37). While ^{18}F-FDG PET/CT imaging is frequently employed as a screening tool to stage lung cancer patients, it exhibits relatively low sensitivity for small metastatic lesions located within lymph nodes (8, 9). In contrast, we discovered that ^{68}Ga-FAPI PET/CT was capable of detecting lymph node metastases more reliably than was ^{18}F-FDG PET/CT, yielding higher SUVmax and TBR values for these metastases relative to the latter imaging modality. As ^{68}Ga-FAPI PET/CT imaging can detect lymph node metastases at an earlier stage, it has the potential to increase occult lymph node metastasis detection, guiding the more accurate staging of NSCLC patients. However, for lymph node metastases from squamous cell carcinoma, SUVmax and TBR of ^{68}Ga-FAPI were not significantly different compared to ^{18}F-FDG. The ability of ^{68}Ga-FAPI PET/CT to detect lymph node metastasis in squamous cell carcinoma still requires further and larger data studies. In patients undergoing mediastinal lymph node dissection, ^{68}Ga-FAPI detected fewer mediastinal false-positive lymph nodes relative to ^{18}F-FDG PET/CT, indicating that ^{68}Ga-FAPI is more specific and has the potential to reduce the rate of unnecessary treatment in patients with NSCLC.

Our analyses additionally revealed ^{68}Ga-FAPI to be superior to ^{18}F-FDG PET/CT when used for the detection of hepatic, pleural, and bone metastases, in line with prior evidence (28, 35, 38, 39). This is ascribable to the reduced physiological uptake of the ^{68}Ga-FAPI radiotracer and associated sensitivity gains. High levels of hepatic glucose metabolism have the potential to mask FDG uptake by metastatic lesions within this organ, while the use of ^{68}Ga-FAPI PET/CT may enable the more reliable detection of these lesions. ^{68}Ga-FAPI PET/CT is also capable of facilitating the early detection of occult bone and pleural metastases to guide more appropriate patient staging and treatment efforts. Unfortunately, we found discovered that ^{68}Ga-FAPI PET/CT exhibited low sensitivity when used to detect adrenal metastases, suggesting that such lesions may be not associated with substantial fibrotic activity. However, ^{18}F-FDG PET/CT is also not effective in diagnosing adrenal metastases due to its low specificity, suggesting that a combination of CT imaging and other modalities is necessary to ensure an accurate diagnosis. High physiological uptake in normal organs masks lesions, or metastases with low FDG uptake or small size are difficult to detect on ^{18}FDG-PET/CT, which may lead to low detection rates on ^{18}F-FDG PET/CT. The superiority of ^{68}Ga-FAPI over ^{18}F-FDG PET/CT for visual assessment and detection of most metastases may be attributed to the higher FAPI accumulation in the lesion and lower FAPI accumulation in normal organs.

There are multiple limitations to the present analysis. For one, the number of included contributors was relatively small, and the variety of NSCLC pathological types was limited, thus potentially contributing to some degree of bias in the overall study results. Second, accurate pathological results were not available for many suspicious metastatic lesions in individuals with advanced NSCLC as it is generally impractical and unethical to conduct biopsies of these samples. Third, the minimum follow-up duration for patients in this study was just 3 months, and future studies should thus utilize an extended follow-up interval.

CONCLUSION

In summary, these results indicate that ^{68}Ga-FAPI PET/CT imaging demonstrates desirable performances when used for the initial staging of newly diagnosed NSCLC. Moreover, ^{68}Ga-FAPI exhibits significantly better diagnostic efficacy relative to that of ^{18}F-FDG PET/CT imaging when used to detect metastatic lesions in the lymph nodes, pleura, liver, and bone. Therefore, ^{68}Ga-FAPI PET/CT is expected to be a viable imaging modality for staging and management of patients with NSCLC, and may be an ideal alternative to ^{18}F-FDG PET/CT.

DATA AVAILABILITY STATEMENT

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study. Requests to access these datasets should be directed to the corresponding author.

ETHICS STATEMENT

The Ethics Committee of Southwest Medical University Hospital approved the present study, which was conducted from July 2020 - October 2021 (Ethics committee approval No.: 2020035), and all patients signed a written informed consent form.

AUTHOR CONTRIBUTIONS

JW and CZ conceived and designed the study, as well as desiged the figures and tables. HD and HZ contributed to the statistical analysis, TW, ZR, YW, YC drafted and correcting the manuscript. All authors read and approved the final manuscript.

REFERENCES

1. Thai AA, Solomon BJ, Sequist LV, Gainor JF, Heist RS. Lung Cancer. *Lancet* (2021) 398(10299):535–54. doi: 10.1016/S0140-6736(21)00312-3
2. Group NM-aC, Arriagada R, Auperin A, Burdett S, Higgins JP, Johnson DH, et al. Adjuvant Chemotherapy, With or Without Postoperative Radiotherapy, in Operable Non-Small-Cell Lung Cancer: Two Meta-Analyses of Individual Patient Data. *Lancet* (2010) 375(9722):1267–77. doi: 10.1016/S0140-6736(10)60059-1
3. Zhong WZ, Wang Q, Mao WM, Xu ST, Wu L, Shen Y, et al. Gefitinib Versus Vinorelbine Plus Cisplatin as Adjuvant Treatment for Stage Ii-Iiia (N1-N2) Egfr-Mutant Nsclc (Adjuvant/Ctong1104): A Randomised, Open-Label, Phase 3 Study. *Lancet Oncol* (2018) 19(1):139–48. doi: 10.1016/S1470-2045(17)30729-5
4. Han Y, Ma Y, Wu Z, Zhang F, Zheng D, Liu X, et al. Histologic Subtype Classification of Non-Small Cell Lung Cancer Using Pet/Ct Images. *Eur J Nucl Med Mol Imaging* (2021) 48(2):350–60. doi: 10.1007/s00259-020-04771-5
5. Dwamena BA, Sonnad SS, Angobaldo JO, Wahl RL. Metastases From Non-Small Cell Lung Cancer: Mediastinal Staging in the 1990s–Meta-Analytic Comparison of Pet and Ct. *Radiology* (1999) 213(2):530–6. doi: 10.1148/radiology.213.2.r99nv46530
6. Le Pechoux C, Pourel N, Barlesi F, Lerouge D, Antoni D, Lamezec B, et al. Postoperative Radiotherapy Versus No Postoperative Radiotherapy in Patients With Completely Resected Non-Small-Cell Lung Cancer and Proven Mediastinal N2 Involvement (Lung Art): An Open-Label, Randomised, Phase 3 Trial. *Lancet Oncol* (2022) 23(1):104–14. doi: 10.1016/S1470-2045(21)00606-9
7. Machado Medeiros T, Altmayer S, Watte G, Zanon M, Basso Dias A, Henz Concatto N, et al. 18f-Fdg Pet/Ct and Whole-Body Mri Diagnostic Performance in M Staging for Non-Small Cell Lung Cancer: A Systematic Review and Meta-Analysis. *Eur Radiol* (2020) 30(7):3641–9. doi: 10.1007/s00330-020-06703-1
8. Reck M, Heigener DF, Mok T, Soria JC, Rabe KF. Management of Non-Small-Cell Lung Cancer: Recent Developments. *Lancet* (2013) 382(9893):709–19. doi: 10.1016/S0140-6736(13)61502-0
9. Huellner MW, de Galiza Barbosa F, Husmann L, Pietsch CM, Mader CE, Burger IA, et al. Tnm Staging of Non-Small Cell Lung Cancer: Comparison of Pet/Mr and Pet/Ct. *J Nucl Med* (2016) 57(1):21–6. doi: 10.2967/jnumed.115.162040
10. Chen L, Chen M, Han Z, Jiang F, Xu C, Qin Y, et al. Clinical Significance of Fap-Alpha on Microvessel and Lymphatic Vessel Density in Lung Squamous Cell Carcinoma. *J Clin Pathol* (2018) 71(8):721–8. doi: 10.1136/jclinpath-2017-204872
11. Sandberg TP, Stuart M, Oosting J, Tollenaar R, Sier CFM, Mesker WE. Increased Expression of Cancer-Associated Fibroblast Markers at the Invasive Front and Its Association With Tumor-Stroma Ratio in Colorectal Cancer. *BMC Cancer* (2019) 19(1):284. doi: 10.1186/s12885-019-5462-2
12. Wong PF, Wei W, Gupta S, Smithy JW, Zelterman D, Kluger HM, et al. Multiplex Quantitative Analysis of Cancer-Associated Fibroblasts and Immunotherapy Outcome in Metastatic Melanoma. *J Immunother Cancer* (2019) 7(1):194. doi: 10.1186/s40425-019-0675-0
13. Kim MG, Shon Y, Kim J, Oh YK. Selective Activation of Anticancer Chemotherapy by Cancer-Associated Fibroblasts in the Tumor Microenvironment. *J Natl Cancer Inst* (2016) 109(1):djw186. doi: 10.1093/jnci/djw186
14. Sukowati CH, Anfuso B, Croce LS, Tiribelli C. The Role of Multipotent Cancer Associated Fibroblasts in Hepatocarcinogenesis. *BMC Cancer* (2015) 15:188. doi: 10.1186/s12885-015-1196-y
15. Kieffer Y, Hocine HR, Gentric G, Pelon F, Bernard C, Bourachot B, et al. Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discovery* (2020) 10(9):1330–51. doi: 10.1158/2159-8290.CD-19-1384
16. Paulsson J, Micke P. Prognostic Relevance of Cancer-Associated Fibroblasts in Human Cancer. *Semin Cancer Biol* (2014) 25:61–8. doi: 10.1016/j.semcancer.2014.02.006
17. Shangguan C, Gan G, Zhang J, Wu J, Miao Y, Zhang M, et al. Cancer-Associated Fibroblasts Enhance Tumor (18)F-Fdg Uptake and Contribute to the Intratumor Heterogeneity of Pet-Ct. *Theranostics* (2018) 8(5):1376–88. doi: 10.7150/thno.22717
18. Chen H, Zhao L, Ruan D, Pang Y, Hao B, Dai Y, et al. Usefulness of [(68)Ga]Ga-Dota-Fapi-04 Pet/Ct in Patients Presenting With Inconclusive [(18)F]Fdg Pet/Ct Findings. *Eur J Nucl Med Mol Imaging* (2021) 48(1):73–86. doi: 10.1007/s00259-020-04940-6
19. Loktev A, Lindner T, Mier W, Debus J, Altmann A, Jager D, et al. A Tumor-Imaging Method Targeting Cancer-Associated Fibroblasts. *J Nucl Med* (2018) 59(9):1423–9. doi: 10.2967/jnumed.118.210435
20. Chen WT, Kelly T. Seprase Complexes in Cellular Invasiveness. *Cancer Metastasis Rev* (2003) 22(2-3):259–69. doi: 10.1023/a:1023055600919
21. Huang Y, Wang S, Kelly T. Seprase Promotes Rapid Tumor Growth and Increased Microvessel Density in a Mouse Model of Human Breast Cancer. *Cancer Res* (2004) 64(8):2712–6. doi: 10.1158/0008-5472.can-03-3184
22. Yang X, Lin Y, Shi Y, Li B, Liu W, Yin W, et al. Fap Promotes Immunosuppression by Cancer-Associated Fibroblasts in the Tumor Microenvironment *Via* Stat3-Ccl2 Signaling. *Cancer Res* (2016) 76(14):4124–35. doi: 10.1158/0008-5472.CAN-15-2973
23. Nawrocka D, Krzyscik MA, Opalinski L, Zakrzewska M, Otlewski J. Stable Fibroblast Growth Factor 2 Dimers With High Pro-Survival and Mitogenic Potential. *Int J Mol Sci* (2020) 21(11):4108. doi: 10.3390/ijms21114108
24. Bainbridge TW, Dunshee DR, Kljavin NM, Skelton NJ, Sonoda J, Ernst JA. Selective Homogeneous Assay for Circulating Endopeptidase Fibroblast Activation Protein (Fap). *Sci Rep* (2017) 7(1):12524. doi: 10.1038/s41598-017-12900-8
25. Kratochwil C, Flechsig P, Lindner T, Abderrahim L, Altmann A, Mier W, et al. (68)Ga-Fapi Pet/Ct: Tracer Uptake in 28 Different Kinds of Cancer. *J Nucl Med* (2019) 60(6):801–5. doi: 10.2967/jnumed.119.227967
26. Giesel FL, Kratochwil C, Lindner T, Marschalek MM, Loktev A, Lehnert W, et al. (68)Ga-Fapi Pet/Ct: Biodistribution and Preliminary Dosimetry Estimate of 2 Dota-Containing Fap-Targeting Agents in Patients With Various Cancers. *J Nucl Med* (2019) 60(3):386–92. doi: 10.2967/jnumed.118.215913
27. Pang Y, Zhao L, Luo Z, Hao B, Wu H, Lin Q, et al. Comparison of (68)Ga-Fapi and (18)F-Fdg Uptake in Gastric, Duodenal, and Colorectal Cancers. *Radiology* (2021) 298(2):393–402. doi: 10.1148/radiol.2020203275
28. Chen H, Pang Y, Wu J, Zhao L, Hao B, Wu J, et al. Comparison of [(68)Ga]Ga-Dota-Fapi-04 and [(18)F] Fdg Pet/Ct for the Diagnosis of Primary and Metastatic Lesions in Patients With Various Types of Cancer. *Eur J Nucl Med Mol Imaging* (2020) 47(8):1820–32. doi: 10.1007/s00259-020-04769-z
29. Koerber SA, Staudinger F, Kratochwil C, Adeberg S, Haefner MF, Ungerechts G, et al. The Role of (68)Ga-Fapi Pet/Ct for Patients With Malignancies of the Lower Gastrointestinal Tract: First Clinical Experience. *J Nucl Med* (2020) 61(9):1331–6. doi: 10.2967/jnumed.119.237016
30. Wei Y, Cheng K, Fu Z, Zheng J, Mu Z, Zhao C, et al. [(18)F]Alf-Nota-Fapi-04 Pet/Ct Uptake in Metastatic Lesions on Pet/Ct Imaging Might Distinguish Different Pathological Types of Lung Cancer. *Eur J Nucl Med Mol Imaging* (2022) 49(5):1671–81. doi: 10.1007/s00259-021-05638-z
31. Lan L, Liu H, Wang Y, Deng J, Peng D, Feng Y, et al. The Potential Utility of [(68) Ga]Ga-Dota-Fapi-04 as a Novel Broad-Spectrum Oncological and Non-Oncological Imaging Agent-Comparison With [(18)F]Fdg. *Eur J Nucl Med Mol Imaging* (2022) 49(3):963–79. doi: 10.1007/s00259-021-05522-w
32. Gu B, Xu X, Zhang J, Ou X, Xia Z, Guan Q, et al. The Added Value of (68)Ga-Fapi-04 Pet/Ct in Patients With Head and Neck Cancer of Unknown Primary With (18)F-Fdg Negative Findings. *J Nucl Med* (2022) 63(6):875–81. doi: 10.2967/jnumed.121.262790
33. Syed M, Flechsig P, Liermann J, Windisch P, Staudinger F, Akbaba S, et al. Fibroblast Activation Protein Inhibitor (Fapi) Pet for Diagnostics and Advanced Targeted Radiotherapy in Head and Neck Cancers. *Eur J Nucl Med Mol Imaging* (2020) 47(12):2836–45. doi: 10.1007/s00259-020-04859-y
34. Zhao L, Chen S, Lin L, Sun L, Wu H, Lin Q, et al. [(68)Ga]Ga-Dota-Fapi-04 Improves Tumor Staging and Monitors Early Response to Chemoradiotherapy in a Patient With Esophageal Cancer. *Eur J Nucl Med Mol Imaging* (2020) 47(13):3188–9. doi: 10.1007/s00259-020-04818-7
35. Wang L, Tang G, Hu K, Liu X, Zhou W, Li H, et al. Comparison of (68)Ga-Fapi and (18)F-Fdg Pet/Ct in the Evaluation of Advanced Lung Cancer. *Radiology* (2022) 303(1):191–9. doi: 10.1148/radiol.211424

36. Zhu L, Yin G, Chen W, Li X, Yu X, Zhu X, et al. Correlation Between Egfr Mutation Status and F(18) -Fluorodeoxyglucose Positron Emission Tomography-Computed Tomography Image Features in Lung Adenocarcinoma. *Thorac Cancer* (2019) 10(4):659–64. doi: 10.1111/1759-7714.12981
37. Li ZM, Ding ZP, Luo QQ, Wu CX, Liao ML, Zhen Y, et al. Prognostic Significance of the Extent of Lymph Node Involvement in Stage Ii-N1 Non-Small Cell Lung Cancer. *Chest* (2013) 144(4):1253–60. doi: 10.1378/chest.13-0073
38. Zhao L, Pang Y, Zheng H, Han C, Gu J, Sun L, et al. Clinical Utility of [(68)Ga] Ga-Labeled Fibroblast Activation Protein Inhibitor (Fapi) Positron Emission Tomography/Computed Tomography for Primary Staging and Recurrence Detection in Nasopharyngeal Carcinoma. *Eur J Nucl Med Mol Imaging* (2021) 48(11):3606–17. doi: 10.1007/s00259-021-05336-w
39. Wu J, Wang Y, Liao T, Rao Z, Gong W, Ou L, et al. Comparison of the Relative Diagnostic Performance of [(68)Ga]Ga-Dota-Fapi-04 and [(18)F]Fdg Pet/Ct for the Detection of Bone Metastasis in Patients With Different Cancers. *Front Oncol* (2021) 11:737827. doi: 10.3389/fonc.2021.737827

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

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 21 October 2022
DOI 10.3389/fonc.2022.870556

OPEN ACCESS

EDITED BY
Yiyan Liu,
University of Louisville, United States

REVIEWED BY
Fan Yingwei,
Beijing Institute of Technology, China
Yuping Li,
First Affiliated Hospital of Wenzhou Medical University, China

*CORRESPONDENCE
Zhen Yang
yztoogetyou@163.com
Liangan Chen
chenliangan301@163.com

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 07 February 2022
ACCEPTED 27 September 2022
PUBLISHED 21 October 2022

CITATION
Zhu Q, Yu H, Liang Z, Zhao W, Zhu M, Xu Y, Guo M, Jia Y, Zou C, Yang Z and Chen L (2022) Novel image features of optical coherence tomography for pathological classification of lung cancer: Results from a prospective clinical trial.
Front. Oncol. 12:870556.
doi: 10.3389/fonc.2022.870556

Novel image features of optical coherence tomography for pathological classification of lung cancer: Results from a prospective clinical trial

Qiang Zhu[1], Hang Yu[1], Zhixin Liang[1], Wei Zhao[1], Minghui Zhu[2], Yi Xu[1], Mingxue Guo[1], Yanhong Jia[1], Chenxi Zou[1], Zhen Yang[1*] and Liangan Chen[1*]

[1]Department of Respiratory Medicine, The First Medical Center of Chinese People Liberation Army (PLA) General Hospital, Beijing, China, [2]Department of Pulmonary and Critical Care Medicine, Zhongnan Hospital of Wuhan University, Wuhan, China

Background: This study aimed to explore the characteristics of optical coherence tomography (OCT) imaging for differentiating between benign and malignant lesions and different pathological types of lung cancer in bronchial lesions and to preliminarily evaluate the clinical value of OCT.

Methods: Patients who underwent bronchoscopy biopsy and OCT between February 2019 and December 2019 at the Chinese PLA General Hospital were enrolled in this study. White-light bronchoscopy (WLB), auto-fluorescence bronchoscopy (AFB), and OCT were performed at the lesion location. The main characteristics of OCT imaging for the differentiation between benign and malignant lesions and the prediction of the pathological classification of lung cancer in bronchial lesions were identified, and their clinical value was evaluated.

Results: A total of 135 patients were included in this study. The accuracy of OCT imaging for differentiating between benign and malignant bronchial lesions was 94.1%, which was significantly higher than that of AFB (67.4%). For the OCT imaging of SCC, adenocarcinoma, and small-cell lung cancer, the accuracies were 95.6, 94.3, and 92%, respectively. The accuracy, sensitivity, and specificity of OCT were higher than those of WLB. In addition, these main OCT image characteristics are independent influencing factors for predicting the corresponding diseases through logistic regression analysis between the main

OCT image characteristics in the study and the general clinical features of patients (p<0.05).

Conclusion: As a non-biopsy technique, OCT can be used to improve the diagnosis rate of lung cancer and promote the development of non-invasive histological biopsy.

KEYWORDS

bronchoscopy, lung cancer, OCT, AFB, WLB

Background

Lung cancer is the leading cause of cancer deaths in the world (1–3), which is a serious threat to human health. At present, the "gold standard" for the diagnosis of lung cancer is histopathological results, which could be obtained by bronchoscopic biopsy, CT-guided lung biopsy, and surgical operation. These biopsy methods have high diagnostic accuracy; however, patients might suffer from complications such as bleeding, pneumothorax, and infection. Moreover, some of the biopsy methods are high risk, traumatic, and expensive (4, 5). In recent years, the "non-invasive histology" biopsy technology, which can effectively avoid the possible complications of tissue biopsy, has been developing by leaps and bounds. Recent studies (6–10) have also reported that non-invasive histological biopsy has high accuracy, sensitivity, and specificity for the diagnosis of lung cancer, in which optical coherence tomography (OCT) examination can achieve similar histopathological diagnosis without biopsy.

OCT is a high-resolution optical imaging technology, which has the characteristics of non-invasiveness, non-radiation, simple operation, and high repeatability. OCT integrates new technologies such as optics, supersensitive detection, and computer image processing. It utilizes low-energy near-infrared harmless light as the light source, and it detects the microstructure of biological tissue using optical interference principles. The resolution of the OCT image is 30 μm, and the depth of tissue that it detects is 3 mm. Mucous layer, submucosa, alveoli, glands, cartilage, and other structures of the bronchial wall are clearly presented in OC images, which are highly matched with histopathological images. OCT imaging could be applied to distinguish between benign and malignant bronchial lesions, predict the histological classification of lung cancer, and precisely detect small precancerous lesions that are unable to be seen with naked eyes.

In 1998, Pitris (11, 12) has confirmed the feasibility of OCT in human airway *in vitro*. Since then, studies (13–15) have proposed the OCT imaging features for different types of lung cancer with an accuracy of more than 80%. However, there are limitations with the present studies. First of all, the OCT images of adenocarcinoma, squamous cell carcinoma, and poorly differentiated cancer were only completed *in vitro* in these studies. It is known that the tissue degeneration and change of blood flow could largely affect the OCT images, leading to the inaccuracy of characteristics that were summarized. Second, the sample size of these studies was also small, so that the results need to be further verified. Third, the main OCT image features of small-cell lung cancer were only reported in case reports, which still need further verification (13). In addition, the OCT image features summarized in these studies were few and lacked unified image feature evaluation standards. Therefore, we aimed to explore the main OCT image features for the differentiation between benign and malignant lesions and for the prediction of pathological classification of lung cancer including central lung cancer and peripheral lung cancer and to evaluate their clinical value *in vivo*. The results of this study might be useful for the future application of OCT in the diagnosis, evaluation, and prognosis of lung cancer.

Abbreviations: SCC, Squamous cell carcinoma; SCLC, Small-cell lung cancer; OCT, Optical coherence tomography; CT, Computed tomography; WLB, White-light bronchoscopy; AFB, Autofluorescence bronchoscopy; EBUS, Endobronchial ultrasound; BAL, Broncho alveolar lavage; HE, Hematoxylin–eosin; WHO, World Health Organization; IASLC, International Association for the Study of Lung Cancer; ATS, American Thoracic Society; ERS, European Respiratory Society; COPD, Chronic obstructive pulmonary disease; APC, Argon plasma coagulation; FNA, Fine-needle aspiration.

Materials and methods

Research population

This prospective study collected and analyzed the data of patients who underwent bronchoscopy biopsy and OCT examination at the Interventional Diagnosis and Treatment Center for Lung Cancer and Respiratory Diseases of the

Chinese PLA General Hospital from February 2019 to December 2019. Prior to the study, all subjects had met the examination indications in the guidelines for the application of diagnostic flexible bronchoscopy in adults. This study was reviewed and approved by the Ethics Committee of CPLAGH (Ethics No. 2018-232-01). Patients were fully informed of the possible risks of the study and signed the informed consent form before the study began.

The admission criteria of this study were as followed: (1) voluntary participation and written informed consent signed by patients, (2) age ≥ 18 years old, (3) patients with routine diagnostic bronchoscopy through clinical evaluation, (4) normal ECG, and (5) adequate hematopoietic function of bone marrow and organs confirmed by blood examination. The exclusion criteria of this study were as follows: (1) patients with contraindication of bronchoscopy (such as respiratory failure and acute cardio-cerebrovascular events), (2) patients who refuse bronchoscopy because of physical reasons or personal wishes, (3) patients who are not suitable for bronchoscopy or OCT examination by researchers, (4) patients who cannot tolerate operation during bronchoscopy resulting in the uncompleted examination, (5) patients with unfound abnormal lesion during routine bronchoscopy or those who could not complete the biopsy, (6) patients who are participating in other clinical studies, (7) patients with poor compliance who are believed by the researchers to be unable to cooperate for the completion of OCT examination and follow-up, and (8) women who were pregnant.

OCT

OCT system: The OCT system (Yongshida Medical Technology, Guangdong, China) has been approved by the FDA (K102599) for medical research and application. OCT system consists of image analysis system mainframe and aseptic removable probe. The probe of OCT is a cylindrical catheter with 1.7 mm in diameter and 150 cm in length, which is sealed with transparent sheath of 1 mm in length near its head and used for image scanning and acquisition. There is a flexible optical fiber axis in the sheath rotating at the speed of 600–1200 rpm; the working wavelength of the optical fiber passing the light source is 1300 nm and the frequency is 50kHz, the image acquisition speed is 10 frames per second, the axial and longitudinal resolution is 15 and 25 μm, respectively, and the detection depth of gray scale and color mode is 3 mm.

Procedure

In this study, the relevant bronchoscopic procedures were performed by a respiratory physician with 7 years of experience in respiratory endoscopic diagnosis and treatment, including preoperative evaluation, preparation, and anesthesia. The respiratory endoscopic doctors completed the capability training on OCT examination following the learning curve prior to the formal inception of this study.

The respiratory endoscopic doctors had examined bronchial lesions with white-light bronchoscopy (WLB), auto-fluorescence bronchoscopy (AFB), and OCT successively after the completion of preoperative anesthesia. Routine tissue biopsy was performed on the same location after OCT examination, and the biopsy samples were fixed in 10% formalin for subsequent histopathological examination. If the lesion failed to be found or biopsied during the bronchoscopy, the OCT examination will not be performed, and the subsequent operation such as EBUS-TBNA and BAL will be carried out routinely by the doctor. The patients were followed up by telephone or outpatient service on the 3rd and 7th day after the examination, and adverse events and auxiliary examination results were recorded. The data of all patients were recorded in detail as shown in the table of case report.

Histopathology

After the OCT examination, the tissue specimens were immediately fixed in 10% formalin, processed and sliced according to the standard histological procedure, and stained with HE. Two independent pathologists analyzed histological sections of each subject according to the lung tumor classification strategy proposed by WHO, IASLC, ATS, and ERS (16, 17). Adenocarcinoma, SCC and small-cell lung cancer were diagnosed by micro-endoscopy (magnification: 10–20).

Analysis on the images of WLB, AFB, and OCT

The images of WLB and AFB were analyzed by respiratory endoscope operators without knowing the histopathological results, whereas OCT images were analyzed by the professional analysts and three clinicians with professional training experiences. The clinical values of WLB, AFB, and OCT in differentiating benign bronchial lesions from malignant ones and predicting different types of lung cancer were further evaluated based on the histopathological results.

Statistics

SPSS23.0 statistical software was used for statistical analysis. The measurement data that accord with the normal distribution was expressed by $X \pm SD$ and the abnormal distribution by the median value. The counting data were expressed by the

percentage. Paired *t*-test was used to compare the intra-group data in accordance with normal distribution. Independent sample *t*-test was used to compare inter-group data. Paired rank sum test was used to compare intra-group data in accordance with abnormal distribution, and rank-sum test for independent sample was used to compare inter-group data.

Results

Clinical information

A total of 135 patients were enrolled in this study. One hundred five (77.78%) were men and 30 (22.22%) were women; the mean age of the patients was 59.25 ± 10.35 (27–78) years. Ninety-eight (72.59%) patients had a history of smoking. All patients were examined by bronchoscopy biopsy and OCT successfully under local or basic anesthesia, including 126 (93.33%) cases with local anesthesia and 9 (6.67%) cases with basic anesthesia. There were 66 cases with left bronchial lesions and 69 cases with right bronchial lesions. Histopathological results of bronchial lesions included 30 (22.22%) cases with benign lesions and 105 (77.78%) cases with malignant lesions. There were 52 (49.52%) cases of lung squamous cell carcinoma including 19 (18.10%) cases with lung adenocarcinoma and 34 (32.38%) cases with small-cell lung cancer. The average time of OCT examination was 2.8 ± 1.6 min (Table 1). There were no adverse events directly related to OCT examination in this study.

TABLE 1 The clinical data of patients.

Characteristic	Values
Gender——n (%)	
Male	105 (77.78)
Female	30 (22.22)
Age——Years	59.25±10.35
Smoking——n (%)	
Yes	98 (72.59)
No	37 (27.41)
Anesthetic method——n (%)	
Local anesthetic	126 (93.33)
Basic anesthetic	9 (6.67)
Location of lesion——n (%)	
RMB	4 (2.96)
RUB	20 (14.81)
RISB	11 (8.15)
RMIB	12 (8.89)
RLB	19 (14.07)
LMB	10 (7.41)
LUB	10 (7.41)
LSLB	9 (6.67)
LLB	5 (3.70)
LLLB	35 (25.93)
Histopathology——n (%)	
Benign	30 (22.22)
Inflammation of mucosa	23 (76.67)
Granulomatous inflammation	3 (10)
Necrotizing inflammation	4 (13.33)
Malignant	105 (77.78)
SCC	52 (49.52)
adenocarcinoma	19 (18.10)
SCLC	34 (32.38)
Time (min)	
Total	18.6±6.5
WLB	8.7±1.2
AFB	7.1±1.4
OCT	2.8±1.6

The image features and clinical values of OCT, WLB, and AFB

The main OCT image features of benign bronchial lesions was the integrity of normal structure with or without mucosal edema, whereas the image of malignant bronchial lesions was the destruction of normal structural layers including mucous layer, submucosa, and adventitia (Figure 1). The accuracy, sensitivity, and specificity of distinguishing benign from malignant bronchial lesions based on the main OCT image features were 94.1%, 97.1% and 83.3%, respectively. Receiver operator characteristic curve (AUC) = 0.902 ± 0.041 (95% Cl = [0.821, 0.983], $P < 0.001$). The AFB images of benign bronchial lesions are green, whereas those of malignant bronchial lesions are pink (Figure 2). The accuracy, sensitivity, and specificity of distinguishing benign from malignant bronchial lesions through AFB images were 67.4, 78.1, and 30%, respectively. AUC = 0.540 ± 0.061 (95% Cl = [0.421, 0.660], $P < 0.001$). It can be seen that the accuracy, sensitivity, and specificity of OCT in differential diagnosis of benign and malignant bronchial lesions are higher than those of AFB.

The main OCT image features of squamous cell carcinoma include (1) "cyst-like" structure; (2) "round" or "irregularly shaped" high-signal nests, with or without low-density shadow; and (3) scaly protuberance of the epithelial layer and a darkened gray color of the submucosa layer. The accuracy of image features (2) for predicting squamous cell carcinoma was 95.6%. The sensitivity and specificity were 96.2 and 93.8%, respectively. AUC = 0.964 ± 0.025 (95% Cl = [0.916, 1.000], $P < 0.001$). The specificity was 100% by combining the image features (1) and (3), but the sensitivity was reduced to 75% (Figure 3A).

The main OCT image features of adenocarcinoma include (1) "round" or "angulated" poor signal structures in mucous layer and (2) continuous low-density shadow in submucous layer. The accuracy was 94.3% by combining image features (1)

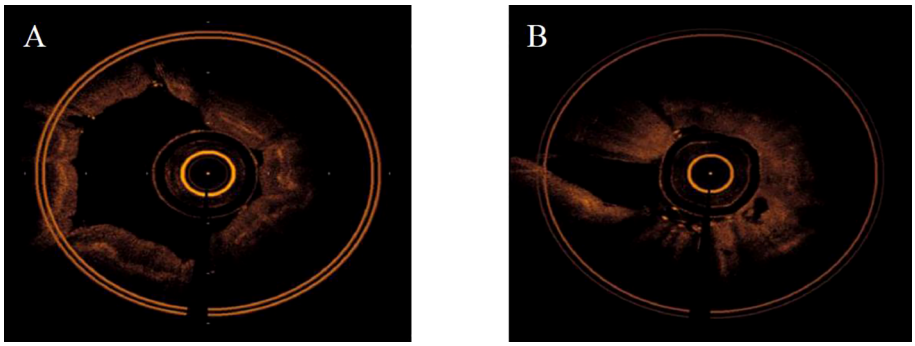


FIGURE 1
The main OCT feature of benign and malignant bronchial lesions. **(A)**: benign bronchial lesions. **(B)**: malignant bronchial lesions.

and (2) for predicting adenocarcinoma, the sensitivity and specificity was 89.5 and 100%, respectively, and AUC = 0.947 ± 0.042 (95% Cl = [0.865, 1.000], $P < 0.001$) (Figure 3A).

The main OCT image features of small-cell lung cancer include (1) low signal lines, which resembled "fracture line," (2) a significantly thickened lamina propria with a darkened gray color of the submucosa layer. The accuracy for predicting small-cell lung cancer was 92% by combining OCT image features (1) and (2); the sensitivity and specificity was 91.2 and 93.8% respectively; AUC = 0.939 ± 0.034 (95% Cl = [0.872, 1.000], $P < 0.001$) (Figure 3C).

The accuracy was verified by incorporating the main OCT image features for the three types of lung cancer into the total sample size and drawing ROC curves, respectively. The AUC was 0.945 ± 0.023 for squamous cell carcinoma (95% Cl= [0.900, 0.989], $P < 0.001$), 0.947 ± 0.041 for adenocarcinoma (95% Cl = [0.866, 0.989], $P < 0.001$), and 0.941 ± 0.030 for small-cell lung cancer (95% Cl = [0.882, 1.000], $P < 0.001$). Therefore, the main OCT image features for the three types of lung cancer in this study have been verified with high accuracy (Figure 4). In addition, the study showed that the accuracy of WLB in predicting squamous cell carcinoma, adenocarcinoma, and small-cell lung cancer was about 70%, which was significantly lower than that of OCT in predicting the pathological classification of lung cancer.

We have conducted a further statistical analysis on the main OCT image features for benign and malignant lesions and for the three types of lung cancer by combining with the total sample size and drawing ROC curves to evaluate its accuracy for the diagnosis of bronchial lesions. The AUC was 0.902 ± 0.023 for benign bronchial lesions (95% Cl= [0.821, 0.983], $P < 0.001$), 0.945 ± 0.041 for squamous cell carcinoma (95% Cl= [0.900, 0.989], $P < 0.001$), 0.947 ± 0.041 for adenocarcinoma (95% Cl= [0.866, 1.000], $P < 0.001$), and 0.941 ± 0.030 for small-cell lung cancer (95% Cl= [0.882, 0.983], $P < 0.001$) (Figure 5). In addition, these main OCT image features are independent

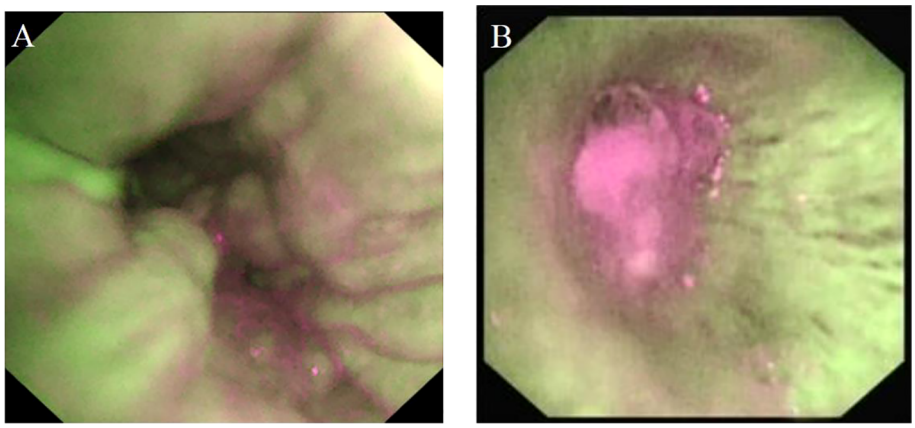


FIGURE 2
The AFB image feature of benign and malignant bronchial lesions. **(A)**: benign bronchial lesions. **(B)**: malignant bronchial lesions.

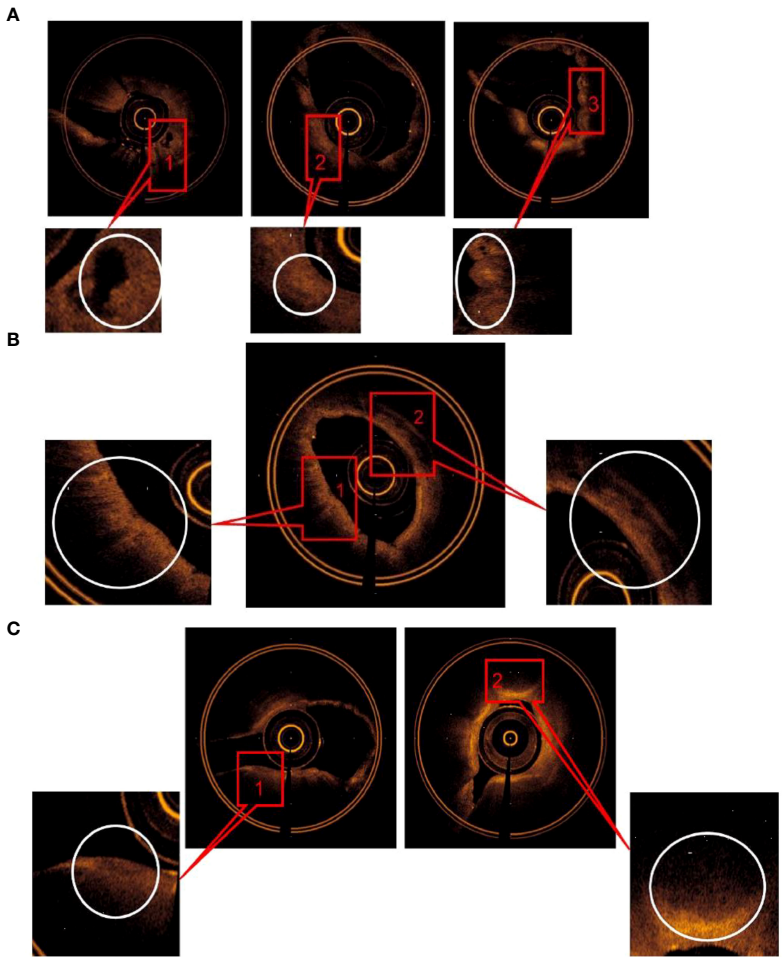


FIGURE 3
(A) The main OCT image features for predicting SCC: (1) "Cyst-like" structure, (2) "round" or irregularly shaped" high signal nests, with or without low-density shadow, (3) Scaly protuberance of the epithelial layer, and a darkened gray level of the submucosa layer was found. **(B)** The main OCT image features for predicting adenocarcinoma: (1) "Round" or "angulated" poor signal structures was found in mucous layer, (2) continuous low density shadow was found in submucous layer. **(C)** The main OCT image features for predicting SCLC: (1)Low signal lines, which resembled "fracture line", (2) a significantly thickened lamina propria with a darkened gray level of the submucosa layer w2as foundcontinuous low density shadow was found.

influencing factors for predicting the corresponding diseases through constructing Logistic regression analysis between the main OCT image characteristics in the study and the general clinical features of patients (Table 2). Therefore, the main OCT image that features each has shown a good differentiation from the other three features in predicting the corresponding diseases, that is, different OCT image features have specific corresponding diagnostic significance.

Discussion

OCT is a newly developed technology with real-time imaging *in vivo*. Researches about OCT have been conducted in various fields of medicine (18), especially in ophthalmology, cardiovascular, and digestive tract diseases. OCT outperformed CT and ultrasound with its high resolution and non-radiation process, indicating its potential clinical application.

At present, OCT has been applied in respiratory diseases such as COPD, bronchial asthma, and lung cancer based on the principle that tissue images can be constructed by computer processing after infrared scattering, and the microstructure of bronchial wall can be recognized in real time. Unlike ultrasound, light waves of OCT do not need liquid medium for propagation, indicating that OCT is more suitable to be used in airway imaging. Moreover, OCT does not require a catheter to contact the tissue, which in turn minimizes adverse reactions that are commonly seen in invasive examination procedures. There is also no risk of exposing to infrared light in a short period of time. In addition, the OCT catheter probe reaches the end of the bronchoscope through the working channel of the bronchoscope and then is further sent to the lesion. During the

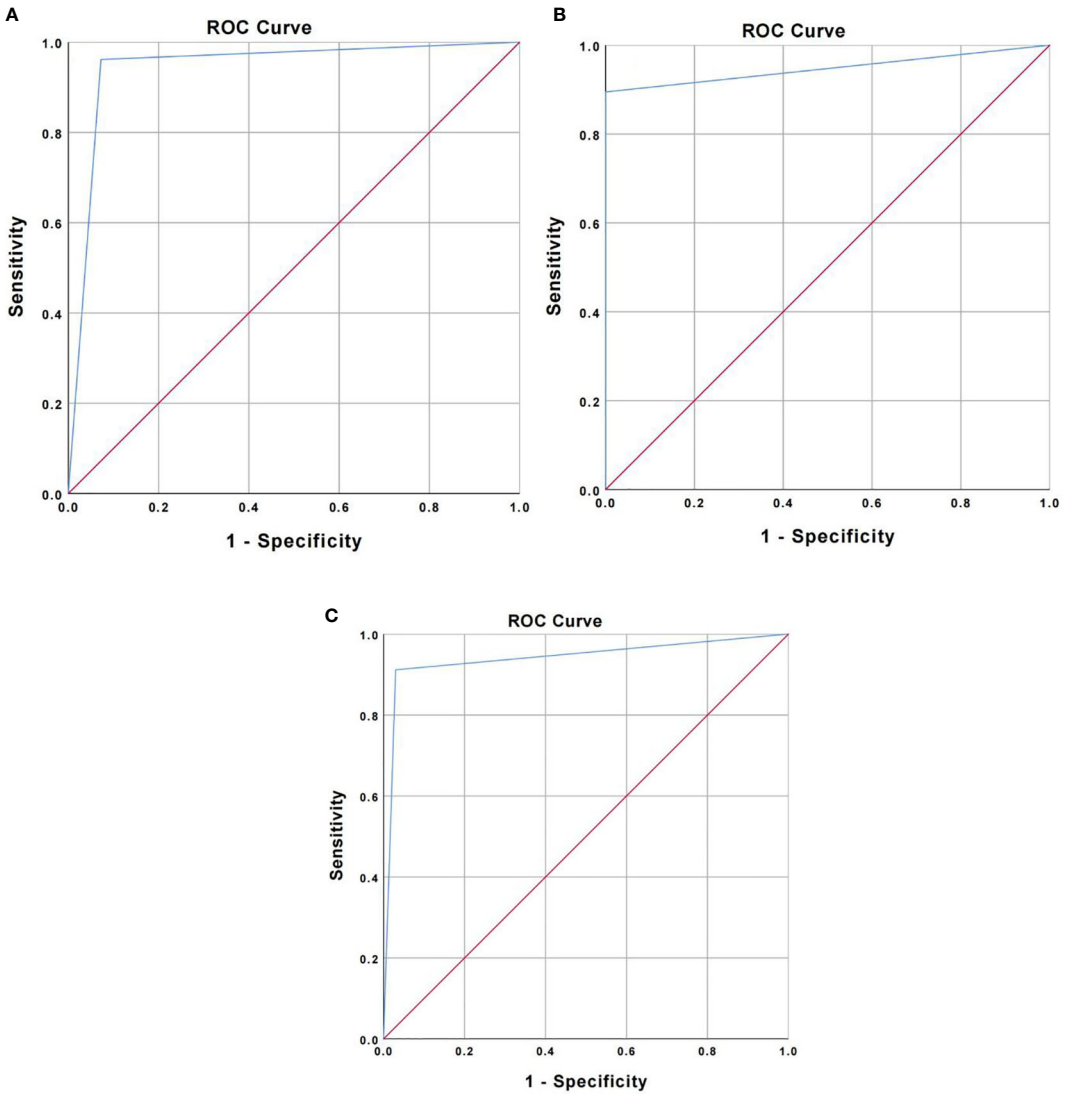


FIGURE 4
(A) ROC curve for verifying the accuracy of major OCT image features of squamous cell carcinoma. **(B)** ROC curve for verifying the accuracy of major OCT image features of adenocarcinoma. **(C)** ROC curve for verifying the accuracy of main OCT image features of small cell lung cancer.

operation, there might be adverse reactions caused by the catheter touching the bronchial wall or the lesions, so scholars have listed OCT examination as one of the "minimally invasive" examinations (19). The adverse events were consistent with the above study, and there were no adverse reactions directly related to OCT examination. Only several cases of mild bleeding or cough were noticed. The adverse reactions were improved and did not affect the continuation of OCT examination after short observation or local administration of a small amount of hemostatic drugs. In this study, the analysts for OCT image features consisted of a professional OCT image analyst and a trained respiratory physician, pathologist, and thoracic surgeon, which were of great significance for judging the histological features of lung cancer, evaluating OCT image features and imaging artifacts, and assessing whether OCT imaging can be evaluated in real time during the examination. In addition, although the real-time scanning location of OCT catheter was close to that of biopsy, it was not completely consistent *in vivo*, so the OCT images could not match the images of tissue pathological sections under microscope, and the one-by-one comparative analysis was not available; therefore, the clinical value of OCT images was evaluated based on the results of histopathology. The OCT examination takes only a few minutes, and the tissue

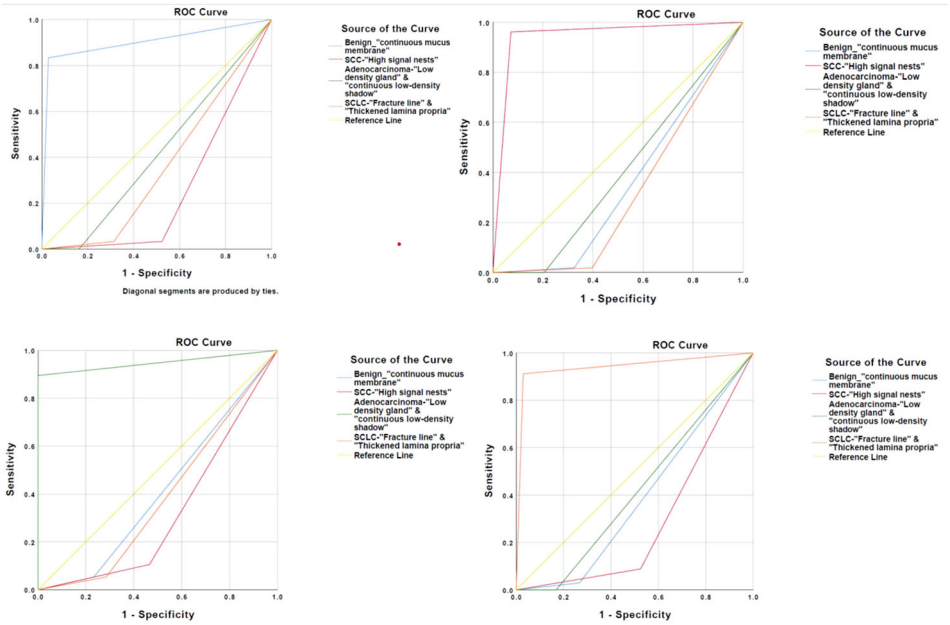


FIGURE 5
Four main OCT image characteristics have a high degree of differentiation to predict the corresponding disease.

microstructure of bronchus such as mucous membrane, submucosa, cartilage, and adventitia were imaged with high resolution without contact with the bronchus in our study; these data are consistent with the findings in other literatures (20, 21). It is worth noting that it is recommended to increase local topical anesthesia for the target lesions and its adjacent bronchi prior to OCT examination, which is useful to increase the tolerance of patients, and reduce the possibility of OCT catheter touching the bronchial wall and lesions, so as to prevent cough stimulation from affecting the imaging effect and destroying the lesions.

The traditional method of respiratory endoscopy for distinguishing benign bronchial lesions from malignant ones is to identify with the naked eyes through the endoscopic images formed in the process of WLB or AFB. WLB has significantly increased the image resolution and improved the ability to distinguish the nature of bronchial lesions by using electronic bronchoscope instead of fiberoptic bronchoscope. However, it is still difficult for experienced bronchoscopic operators to identify subtle mucosal lesions (22, 23). Fluorescence bronchoscopy has been proved to be highly sensitive in differentiating benign and malignant endobronchial lesions, but it has a high false positive rate due to bronchial mucosal inflammation and bleeding. Low specificity is another problem for fluorescence bronchoscopy (24, 25), which is further confirmed in our research results. Former studies (12, 13, 26) have confirmed that OCT can distinguish normal bronchial tissue from tumor lesions with the principle that the mucous layer and submucosa of bronchial wall thickened in different stages of cancer; however, no study

TABLE 2 The logistic regression analysis of the general clinical features and main OCT image characteristics in the differentiation of benign and malignant bronchial lesions and histopathological classification of lung cancer.

OCT	B	S.E.	Wald	p	OR (95% CI)
Distinguish between benign and malignant	5.136	0.764	45.232	<0.001	170 (38.058, 759.374)
SCC-"High signal nests"	5.771	0.836	47.6	<0.001	320.833 (62.272, 1652.987)
Adenocarcinoma-"Low density gland" & "continuous low-density shadow"	4.642	1.06	19.162	<0.001	103.765 (12.983, 829.312)
SCLC-"Fracture line" & "Thickened lamina propria"	5.822	0.842	47.795	<0.001	337.556 (64.796, 1758.491)

had summarized and proposed the main OCT image features in distinguishing benign bronchial lesions from malignant ones. In our study, for the first time, we found that whether the mucosal layer is edematous and the normal structural layer is destroyed are the main OCT image features in distinguishing between benign and malignant bronchial lesions. The accuracy of the main OCT image features in distinguishing benign and malignant bronchial lesions was 94.1%, and the sensitivity and specificity were 97.1 and 83.3%, respectively, both of which were significantly higher than those of AFB. Therefore, comparing with AFB, OCT has more clinical value in distinguishing benign and malignant bronchial lesions. It is worth noting that the histopathological results of malignant lesions were invasive cancer in our study, whereas OCT images of precancerous lesions and carcinoma *in situ* could also show the integrity of normal tissue structure, which was reported in early studies. However, there was a significant difference between benign and malignant bronchial lesions in OCT images in our findings. Therefore, the differentiation between bronchial benign lesions and bronchial precancerous lesions or carcinoma *in situ* still needs more research and further exploration.

OCT has been used to assist the diagnosis and treatment of lung cancer with the development of interventional diagnosis and treatment of respiratory diseases. Many studies (14, 15, 27) have preliminarily reported that the accuracy of main OCT image features for distinguishing adenocarcinoma, squamous cell carcinoma, and poorly differentiated lung cancer *in vitro* was more than 82.6%. However, whether the OCT image features of tissue specimens *in vitro* can reflect the actual condition still needs to be further confirmed, because tissue degeneration and internal blood flow changes will occur in the excised specimens, and the effects of cough, spontaneous breathing, heartbeat, and secretions on the imaging results will not be truly reflected in the examination process. In our study, the clinical value of OCT in real-time diagnosis of lung cancer was evaluated for the first time, and the main OCT image features for predicting adenocarcinoma, squamous cell carcinoma, and small-cell lung cancer were proposed and verified. There are three main OCT image features with squamous cell carcinoma. The sensitivity of any image features in predicting squamous cell carcinoma was 100%, but the specificity was only about 31.2%. The specificity was 100% by combining the three image features to predict squamous cell carcinoma, but the sensitivity was very low. However, the main image features (2) showed high clinical value with the accuracy of 95.6%, and the sensitivity and specificity were 96.2 and 93.8%, respectively. The accuracy and specificity of the main OCT image features of squamous cell carcinoma, which was proposed by previous studies *in vitro*, were 82.6 and 87% (14), respectively. Therefore, the main OCT image features (1), (2), or (3) can be used as the preliminary screening image for the diagnosis of squamous cell carcinoma, and the final image features should be the main OCT image features (2) or the combination of features (1), (2), and (3). It is recommended to choose the combination of each (1) and (2) as the main OCT image features in predicting adenocarcinoma and small-cell lung cancer. The accuracy of the main OCT image features of adenocarcinoma was 94.3%; the sensitivity and specificity were 89.5 and 100%, respectively, which were significantly higher than those in the previous studies *in vitro* (80.3 and 88.6%). The main OCT image features of small-cell lung cancer were proposed by the study for the first time and initially showed high accuracy, sensitivity, and specificity. Therefore, compared with previous studies, the main OCT image features in the study can be used to predict the histological classification of lung cancer significantly. In addition, the sensitivity and specificity have been comparable with the results of studying esophageal and cardiovascular diseases for which the application of OCT is more mature (28–30) and the accuracy results by verification had significant statistical results ($P<0.05$). In addition, the study also found that the accuracy, sensitivity, and specificity of OCT in predicting the pathological classification of lung cancer were higher than those of WLB, so the clinical value should be affirmed. While the OCT images cannot be compared with the histopathological images one by one, therefore, the analysis of OCT images is affected by subjective factors to a certain extent, and the related research data need to be confirmed by more studies. The latest study (31) reports that the OCT catheter is integrated into the 19G puncture needle, which design is useful to highly match the location of OCT examination and aspiration tissue, and the OCT images are more matched with histopathological images, which is expected to obtain more OCT image features with reality and reliability. However, currently, the product is tested using animals; we are looking forward to the relevant findings for clinical research after the product is put on the market.

To sum up, OCT is very useful to clinicians in differentiating benign and malignant bronchial lesions and for the histological classification of lung cancer, especially for patients who are unable to perform tissue biopsy or cannot obtain accurate histological pathological results after biopsy. Meanwhile, it can also provide real-time imaging of the lesions during bronchoscopy and guide the location of bronchoscopy biopsy to improve the positive rate of biopsy and the diagnosis rate of lung cancer. We can even choose direct operation after evaluating the lesion by OCT examination for patients who was diagnosed as lung cancer clinically and have the chance for operation, and there is no need for bronchoscopic biopsy so that the lung cancer metastasis caused by biopsy can be avoided. In addition, studies (32, 33) have suggested that OCT is also helpful in the treatment of lung cancer including auxiliary airway stent implantation, APC, cryosurgery, and other interventional therapy and to explore the relationship between OCT image

features and gene mutations in patients with lung cancer who cannot be biopsied because of high risk of secondary biopsy after targeted drug resistance. However, the findings in the studies were still preliminary. It is expected that more research findings will form a sufficient basis to support its clinical application. It is worth noting that artificial intelligence has been widely used in the diagnosis and treatment of lung cancer in recent years, and OCT will be combined with artificial intelligence in the diagnosis and treatment of lung cancer, but there are still few OCT image features that were found so far, which is not enough for mechanical learning and deep learning. However, the development of follow-up research and the establishment of the database will carry out mechanical learning with massive image features and establish a basic model according to the learning results, and deep learning combined with basic clinical features is conducive to the automation, standardization, and individualization of OCT in the diagnosis and treatment of lung cancer.

Although the findings of the study are quite encouraging, the accuracy, sensitivity, and specificity are not sufficient to support OCT as a complete substitute for tissue biopsy. A study (34) reported that the histopathological results of specimens obtained by FNA were consistent with the pathological results of resection specimens, including 96.2% for adenocarcinoma and 84.7% for squamous cell carcinoma. There are differences between the accuracy of the main OCT image features of lung adenocarcinoma in our study and the above findings, which might due to some important factors that could affect the OCT process. For example, we mainly analyzed the OCT image features of the most common primary lung cancer, but the histological classifications of lung cancer diverses and the OCT image features of some rare lung cancer and metastatic lung tumors still need to be further explored. It has been reported that only 32% of poorly differentiated cancers can be accurately classified by histopathological method (34–36). With the rapid development of targeted therapy in the field of lung cancer treatment in recent years, it may be necessary to perform multiple tissue biopsies for gene detection to evaluate the targeted drug resistance of lung cancer and guide the next step of treatment (37–39). This situation makes other non-invasive tissue biopsy techniques including OCT difficult to completely replace tissue biopsy in a short time. In addition, the resolution of OCT cannot fully reach the level of the microscope, and adjustable factors such as lens selection, focal length adjustment, and high power lens field of view of the microscope cannot be achieved by OCT temporarily. Some factors affecting OCT imaging include spontaneous respiration, heartbeat, blood, airway secretions, calcification, fibrosis, and so on; more advanced technologies need to be developed for data acquisition and automatic image processing to improve the authenticity of airway images and standardize the clinical application of OCT images (40, 41).

Conclusions

In summary, the safety and effectiveness of OCT were confirmed by our study; however, the current findings are not sufficient to support OCT as a complete substitute for tissue biopsy. Nevertheless, optical signals have stronger ability to penetrate tissues with the further development of OCT technology, and the images will have higher resolution and better matching with histopathological images after reducing the influencing factors. It is expected to carry out larger sample size research in the future including early screening of lung cancer and diagnosis of peripheral lung cancer to find more OCT image features of clinical significance. Big data integrate the clinical data of patients with OCT image features and combine with the artificial intelligence, which will play an important role in the clinical application of OCT for the diagnosis and treatment of lung cancer.

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 local institutional review board approved this study (Ethics No. 2018-232-01). This study is registered at Chinese Clinical Trial registry with identifier ChiCTR1900021466. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

LC conceived this study and drafted the manuscript. QZ and ZY collected the clinical data, interpreted the results, and wrote the manuscript. ZL, WZ, HY, MZ, YX, YJ, MG, CZ participated in data collection and critical revision of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by Research on the artificial intelligence-aided diagnosis and clinical decision-making system for ground glass nodules of lung (2020-1-5011), and the Research on the construction of big data platform for diagnosis and treatment with respiratory endoscopy (2019MBD-052).

Acknowledgments

We thank all our colleagues at the Interventional Diagnosis and Treatment Center for Lung Cancer and Respiratory Diseases of the Chinese PLA General Hospital for making this study possible.

Conflict of interest

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

Publisher's note

References

1. Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet.* (2003) 362:847–52. doi: 10.1016/S0140-6736(03)14338-3

2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 0:3–31. doi: 10.3322/caac.21492

3. Zheng S, Sun K, Hao J. Analysis on the prevalence of malignant tumors in China in 2015. *Chin J Oncol* (2019) 41:1. doi: 10.3760/cma.j.issn.0253-3766.2019.01.008

4. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, et al. Diagnosis of lung cancer in small biopsies and cytology: implications of the 2011 international association for the study of lung Cancer/American thoracic Society/European respiratory society classification. *Arch Pathol Lab Med* (2013) 137:668–84. doi: 10.5858/arpa.2012-0263-RA

5. Wang Memoli JS, Nietert PJ, Silvestri GA. Meta-analysis of guided bronchoscopy for the evaluation of the pulmonary nodule. *Chest* (2012) 142:385–93. doi: 10.1378/chest.11-1764

6. Maxwell AW, Klein JS, Dantey K, Mount SL, Butnor KJ, Leiman G. CT-guided transthoracic needle aspiration biopsy of subsolid lung lesions. *J Vasc Interv Radiol* (2014) 25:340–6. doi: 10.1016/j.jvir.2013.11.037

7. Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a metaanalysis. *JAMA* (2001) 285:914–24. doi: 10.1001/jama.285.7.914

8. Keating J, Singhal S. Novel methods of intraoperative localization and margin assessment of pulmonary nodules. *Semin Thorac Cardiovasc Surg* (2016) 28:127–36. doi: 10.1053/j.semtcvs.2016.01.006

9. Hyodo T, Kanazawa S, Dendo S, Kobayashi K, Hayashi H, Kouno Y, et al. Intrapulmonary lymph nodes: thin-section CT findings, pathological findings, and CT differential diagnosis from pulmonary metastatic nodules. *Acta Med Okayama* (2004) 58:235–40. doi: 10.18926/AMO/32102

10. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the fleischner society 2017. *Radiology* (2017) 284(1):228–43. doi: 10.1148/radiol.2017161659

11. Pitris C, Brezinski ME, Bouma BE, Tearney GJ, Southern JF, Fujimoto JG. High resolution imaging of the upper respiratory tract with optical coherence tomography. *Am J Respir Crit Care Med* (1998) 157:1640–1644. doi: 10.1164/ajrccm.157.5.9707075

12. Lam S, Standish B, Baldwin C, McWilliams A, leRiche J, Gazdar A, et al. *In vivo* optical coherence tomography imaging of preinvasive bronchiM lesions. *Clin Cancer Res* (2008) 14:2006–11. doi: 10.1158/1078-0432.CCR-07-4418

13. Michel RG, Kinasewitz GT, Fung K-M, Keddissi JI. Optical coherence tomography as an adjunct to flexible bronchoscopy in the diagnosis of lung cancer. *Chest* (2010) 138(4):984–8. doi: 10.1378/chest.10-0753

14. Hariri LP, Mino-Kenudson M, Lanuti M, Miller AJ, Mark EJ, Suter MJ, et al. Diagnosing lung carcinomas with optical coherence tomography. *Ann Am Thorac Soc* (2015) 12(2):193–201. doi: 10.1513/AnnalsATS.201408-370OC

15. Shostak E, Hariri LP, Cheng GZ, Adams DC, Suter MJ, et al. Needle-based optical coherence tomography to guide transbronchial lymph node biopsy. *J Bronchology Interv Pulmonol* (2018) 25(3):189–97. doi: 10.1097/LBR.0000000000000491

16. 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

17. Travis WD, Brambilla E, Mu ller-Hermelink HK, Harris CC. *Pathology and genetics: tumours of the lung, pleura, thymus and heart*. Lyon, France: IARC (2004).

18. Ridgway JM, Su J, Wright R, Guo S, Kim DC, Barretto R, et al. Optical coherence tomography of the newborn airway. *Ann Otol Rhinol Laryngol* (2008) 117(5):327–34.

19. Williamson JP, Armstrong JJ, McLaughlin RA, Noble PB, West AR, Becker S, et al. Measuring airway dimensions during bronchoscopy using anatomical optical coherence tomography. *Eur Respir J* (2010) 35(1):34–41. doi: 10.1183/09031936.00041809

20. Han S, El-Abbadi NH, Hanna N, Mahmood U, Mina-Araghi R, Jung WG, et al. Evaluation of tracheal imaging by optical coherence tomography. *Respiration.* (2005) 72(5):537–41. doi: 10.1159/000087680

21. Jung W, Zhang J, Mina-Araghi R, Hanna N, Brenner M, Nelson JS, et al. Feasibility study of normal and septic tracheal imaging using optical coherence tomography. *Lasers Surg Med* (2004) 35(2):121–7. doi: 10.1002/lsm.20072

22. Kato H, Imaizumi T, Aizawa K, Iwabuchi H, Yamamoto H, Ikeda N, et al. Photodynamic diagnosis in respiratory tract malignancy using an excimer dye laser system. *J Photochem Photobiol* (1990) 6:189–96. doi: 10.1016/1011-1344(90)85089-F

23. Woolner LB, Fontana RS, Cortese DA, Sanderson DR, Bernatz PE, Payne WS, et al. Roentgenographically occult lung cancer pathologic findings and frequency of multicentric during a 10-year period. *Mayo Clin Proc* (1983) 59:435–66. doi: 10.1016/s0025-6196(12)60434-0

24. Ikeda N, Honda H, Katsumi T, Okunaka T, Furukawa K, Tsuchida T, et al. Early detection of bronchial lesions using lung imaging fluorescence endoscope. *Diagn Ther Endoscopy* (1999) 5:85–90. doi: 10.1155/DTE.5.85

25. Kakihana M, Kim K, Okunaka T, Furukawa K, Hirano T, Konaka C, et al. Early detection of bronchial lesions using system of autofluorescence endoscopy (SAFE) 1000. *Diagn Ther Endoscopy* (1999) 5:99–104. doi: 10.1155/DTE.5.99

26. Tsuboi M, Hayashi A, Ikeda N, Honda H, Kato Y, Ichinose S, et al. Optical coherence tomography in the diagnosis of bronchial lesions. *Lung Cancer* (2005) 49:387–94. doi: 10.1016/j.lungcan.2005.04.007

27. McLaughlin RA, Scolaro L, Robbins P, Hamza S, Saunders C, et al. Imaging of human lymph nodes using optical coherence tomography: Potential for staging cancer. *Cancer Res* (2010) 70(7):2579–84. doi: 10.1158/0008-5472.CAN-09-4062

28. Yabushita H, Bouma BE, Houser SL, Aretz HT, Jang IK, Schlendorf KH, et al. Characterization of human atherosclerosis by optical coherence tomography. *Circulation* (2002) 106:1640–5. doi: 10.1161/01.CIR.0000029927.92825.F6

29. Evans JA, Bouma BE, Bressner J, Shishkov M, Lauwers GY, Mino-Kenudson M, et al. Identifying intestinal metaplasia at the squamocolumnar junction by using optical coherence tomography. *Gastrointest Endosc* (2007) 65:50–6. doi: 10.1016/j.gie.2006.04.027

30. Evans JA, Poneros JM, Bouma BE, Bressner J, Halpern EF, Shishkov M, et al. Optical coherence tomography to identify intramucosal carcinoma and high-grade dysplasia in barrett's esophagus. *Clin Gastroenterol Hepatol* (2006) 4:38–43. doi: 10.1016/S1542-3565(05)00746-9

31. Li J, Quirk BC, Noble PB, Kirk RW, Sampson DD, McLaughlin RA, et al. Flexible needle with integrated optical coherence tomography probe for imaging during transbronchial tissue aspiration. *J Biomed Optics* (2017) 22(10):1–5. doi: 10.1117/1.JBO.22.10.106002

32. Wijmans L, d'Hooghe JNS, Bonta PI, Annema JT. Optical coherence tomography and confocal laser endomicroscopy in pulmonary diseases. *Curr Opin Pulm Med* (2017) 23(3):275–83. doi: 10.1097/MCP.0000000000000375

33. Hariri LP, Villiger M, Applegate MB, Mino-Kenudson M, Mark EJ, Bouma BE, et al. Seeing beyond the bronchoscope to increase the diagnostic yield of bronchoscopic biopsy. *Am J Respir Crit Care Med* (2013) 187(2):125–9. doi: 10.1164/rccm.201208-1483OE

34. Da Cunha Santos G, Lai SW, Saieg MA, Geddie WR, Pintilie M, Tsao M-S, et al. Cyto-histologic agreement in pathologic subtyping of non small cell lung carcinoma: review of 602 fine needle aspirates with follow-up surgical specimens over a nine year period and analysis of factors underlying failure to subtype. *Lung Cancer* (2012) 77:501–6. doi: 10.1016/j.lungcan.2012.05.091

35. Wu Y, Xi J, Huo L. Robust high-resolution fine OCT needle for side-viewing interstitial tissue imaging. *IEEE J Sel Top Quantum Electron* (2010) 16:863–9. doi: 10.1109/JSTQE.2009.2035362

36. Mazzone P, Jain P, Arroliga AC, Matthay RA. Bronchoscopy and needle biopsy techniques for diagnosis and staging of lung cancer. *Clin Chest Med* (2002) 23:137–58. doi: 10.1016/S0272-5231(03)00065-0

37. Mukhopadhyay S, Katzenstein AL. Subclassification of non–small cell lung carcinomas lacking morphologic differentiation on biopsy specimens: utility of an immunohistochemical panel containing TTF-1, napsin a, p63, and CK5/6. *Am J Surg Pathol* (2011) 35:15–25. doi: 10.1097/PAS.0b013e3182036d05

38. Cooper WA, O'Toole S, Boyer M, Horvath L, Mahar A. What's new in non–small cell lung cancer for pathologists: the importance of accurate subtyping, EGFR mutations and ALK rearrangements. *Pathology* (2011) 43:103–115. doi: 10.1097/PAT.0b013e328342629d

39. Travis WD, Brambilla E, Riely GJ. New pathologic classification of lung cancer: relevance for clinical practice and clinical trials. *J Clin Oncol* (2013) 31:992–1001. doi: 10.1200/JCO.2012.46.9270

40. Quirk BC, McLaughlin RA, Curatolo A, Kirk RW, Noble PB, Sampson DD, et al. *In situ* imaging of lung alveoli with an optical coherence tomography needle probe. *J BioMed Opt* (2011) 16:036009. doi: 10.1117/1.3556719

41. Lorenser D, Yang X, Kirk RW, Quirk BC, McLaughlin RA, Sampson DD. Ultrathin side-viewing needle probe for optical coherence tomography. *Opt Lett* (2011) 36:3894–6. doi: 10.1364/OL.36.003894

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact