



OPEN ACCESS

EDITED BY

Jiansong Ji,
Lishui Central Hospital, China

REVIEWED BY

Fu Xiong,
Huazhong University of Science and
Technology, China
Chunli Kong,
Lishui Central Hospital, China

*CORRESPONDENCE

Lin Wu
wulin-calf@yeah.net

[†]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 03 July 2022

ACCEPTED 25 July 2022

PUBLISHED 16 August 2022

CITATION

Jiang M, Yang P, Li J, Peng W, Pu X,
Chen B, Li J, Wang J and Wu L (2022)
Computed tomography-based
radiomics quantification predicts
epidermal growth factor receptor
mutation status and efficacy of first-
line targeted therapy in lung
adenocarcinoma.
Front. Oncol. 12:985284.
doi: 10.3389/fonc.2022.985284

COPYRIGHT

© 2022 Jiang, Yang, Li, Peng, Pu, Chen,
Li, Wang and Wu. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

Computed tomography-based radiomics quantification predicts epidermal growth factor receptor mutation status and efficacy of first-line targeted therapy in lung adenocarcinoma

Meilin Jiang^{1†}, Pei Yang^{2,3†}, Jing Li^{4†}, Wenyong Peng⁵,
Xingxiang Pu¹, Bolin Chen¹, Jia Li¹, Jingyi Wang¹ and Lin Wu^{1*}

¹The Second Department of Thoracic Oncology, Hunan Cancer Hospital, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China, ²The General Surgery Department of Xiangya Hospital Affiliated to Central South University, Changsha, China,

³The National Clinical Research Center for Geriatric Disorders of Xiangya Hospital Affiliated to Central South University, Changsha, China, ⁴Medical Oncology, Hunan Cancer Hospital, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China, ⁵The Second Department of Oncology, Yunnan Cancer Hospital, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Center, Kunming, China

Background: Biomarkers that predict the efficacy of first-line tyrosine kinase inhibitors (TKIs) are pivotal in epidermal growth factor receptor (EGFR) mutant advanced lung adenocarcinoma. Imaging-based biomarkers have attracted much attention in anticancer therapy. This study aims to use the machine learning method to distinguish EGFR mutation status and further explores the predictive role of EGFR mutation-related radiomics features in response to first-line TKIs.

Methods: We retrospectively analyzed pretreatment CT images and clinical information from a cohort of lung adenocarcinomas. We entered the top-ranked features into a support vector machine (SVM) classifier to establish a radiomics signature that predicted EGFR mutation status. Furthermore, we identified the best response-related features based on EGFR mutant-related features in first-line TKI therapy patients. Then we test and validate the predictive effect of the best response-related features for progression-free survival (PFS).

Results: Six hundred ninety-two patients were enrolled in building radiomics signatures. The 13 top-ranked features were input into an SVM classifier to establish the radiomics signature of the training cohort ($n = 514$), and the predictive score of the radiomics signature was assessed on an independent validation group with 178 patients and obtained an area under the curve (AUC) of 74.13%, an F1 score of 68.29%, a specificity of 79.55%, an accuracy of 70.79%,

and a sensitivity of 62.22%. More importantly, the skewness-Low (≤ 0.882) or 10th percentile-Low group (≤ 21.132) had a superior partial response (PR) rate than the skewness-High or 10th percentile-High group ($p < 0.01$). Higher skewness (hazard ratio (HR) = 1.722, $p = 0.001$) was also found to be significantly associated with worse PFS.

Conclusions: The radiomics signature can be used to predict EGFR mutation status. Skewness may contribute to the stratification of disease progression in lung cancer patients treated with first-line TKIs.

KEYWORDS

lung adenocarcinoma, computed tomography, radiomic response biomarker, epidermal growth factor receptor mutation status, machine learning

Introduction

Lung cancer is the most prevalent cancer worldwide, causing the highest cancer-related death rate of all malignancies (1). Adenocarcinoma comprises 80% of non-small cell lung cancer (NSCLC), and epidermal growth factor receptor (EGFR) mutations mostly appear in this subtype (2, 3). With the discovery and development of tyrosine kinase inhibitors (TKIs), the clinical treatment strategy for advanced activating EGFR mutation lung adenocarcinoma has evolved into a personalized approach (4, 5). Based on the National Comprehensive Cancer Network (NCCN) and Chinese Society of Clinical Oncology (CSCO) guidelines, EGFR TKIs have been approved as first-line standard therapy for driver mutation-positive metastatic adenocarcinoma based on studies that have shown better survival than chemotherapy (3, 6–9).

Nowadays, the individual diagnosis and treatment of EGFR-mutant lung adenocarcinoma depend on invasive biopsy testing. However, low DNA quality and testing methods can limit the reliability of results and sequencing applications (10–14). Furthermore, the EGFR mutation result was only determined by a part of tumor tissue, ignoring the heterogeneity of the entire tumor, which might be the reason for the inconsistent treatment outcome. When patients preliminarily elect for EGFR-TKI therapy only based on EGFR mutation, their response will not last long and varies so markedly after treatment (4, 9, 15). In sum, it is crucial and urgent to use the whole picture of the tumor to predict the potential resistance or the likelihood of rapid progression comprehensively before patients receive EGFR TKIs.

Radiomics is a non-invasive and high-throughput image assessment approach based on medical imaging (16, 17). A correlation between radiomics features and underlying intertumoral heterogeneity of lung cancer has been observed

(18–26). Furthermore, molecular images have been used to identify patients with different therapeutic outcomes of EGFR-TKI therapy (27–30). Tian et al. built a signature to discriminate lung cancer patients with rapid and slow progression to EGFR-TKI therapy using the least absolute shrinkage and selection operator (LASSO) Cox regression model based on two-direction imaging data. Cook et al. found the association between features and survival by Cox regression analyses. However, compared to the predictive model that was made of an ‘unknown process’, oncologists tend to identify some specific image features and link them to the medical explanation.

Hence, our study aimed to locate some specific image features that were highly related to the survival outcome and could be linked to clinical practice. We proposed a radiomics signature based on all three computed tomography (CT) image dimensions for predicting EGFR mutation status. We further explored in-depth the relevance between EGFR mutation-related features and risk stratification of progression-free survival (PFS) in EGFR mutant advanced adenocarcinoma.

Materials and methods

Patients

The institutional research board of Hunan Cancer Hospital (Changsha, China) approved this retrospective study. A total of 1,219 lung adenocarcinoma patients at Hunan Cancer Hospital were initially collected between July 2013 and September 2019. Patients were included in this research based on the following inclusion criteria: 1) pathologically confirmed primary pulmonary adenocarcinoma in our institute, 2) there are measurable target lesions under the Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST v 1.1), 3) next-

generation sequencing-proven EGFR mutational status by tumor tissue sample, and 4) available patient characteristic clinical data. Finally, 692 patients were included in our study. Furthermore, clinical data were collected, including therapy protocol, response evaluation, and follow-up material. In the process of building the predictive radiomics signature, patients

confirmed between 1 July 2013 and 1 May 2018 were enrolled in a training cohort, and those confirmed between 1 June 2018 and 1 September 2019 were enrolled in a test cohort (Figure 1).

Response assessment routinely took place 4–6 weeks after treatment completion by diagnostic CT scans and laboratory tests according to the RECIST v 1.1. PFS is the study endpoint

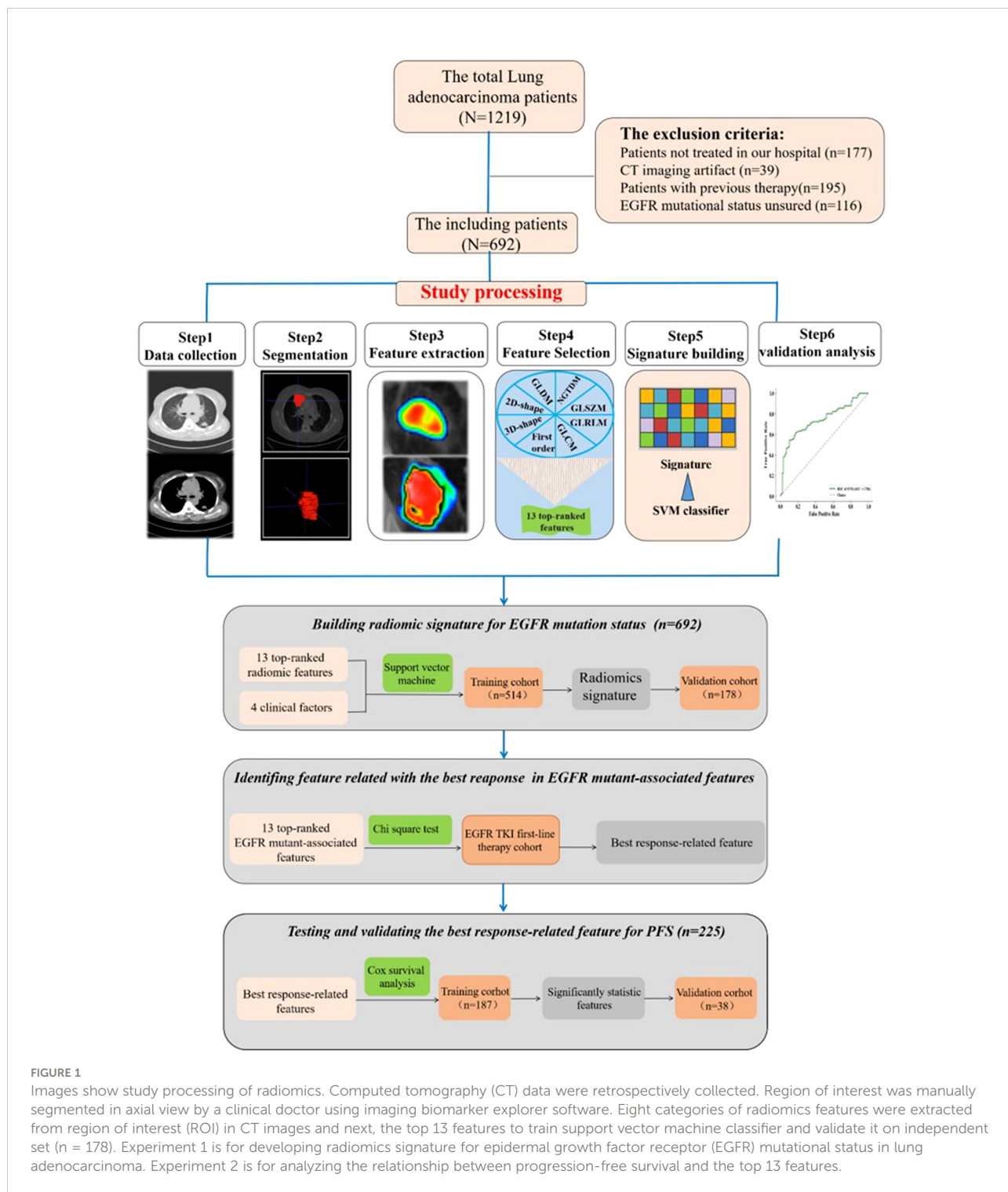


FIGURE 1

Images show study processing of radiomics. Computed tomography (CT) data were retrospectively collected. Region of interest was manually segmented in axial view by a clinical doctor using imaging biomarker explorer software. Eight categories of radiomics features were extracted from region of interest (ROI) in CT images and next, the top 13 features to train support vector machine classifier and validate it on independent set (n = 178). Experiment 1 is for developing radiomics signature for epidermal growth factor receptor (EGFR) mutational status in lung adenocarcinoma. Experiment 2 is for analyzing the relationship between progression-free survival and the top 13 features.

considered the time from the initiation of therapy to confirmation of progression or death.

CT scanning protocol

All thoracic CT examinations were performed at Hunan Cancer Hospital. CT images of all patients were acquired on CT-on-rails (Brilliance CT 16, Hunan Tumor Hospital, Changsha) with the following parameters: a 5.0-mm slice thickness reconstruction, 313-mA tube current, and 120-kV peak voltage.

Tumor imaging segmentation and feature extraction

In this study, all nodules were identified by a radiologist with more than 10 years of experience, and the clinician manually annotated the regions of interest (ROIs) on axial view piece by piece using imaging biomarker explorer (IBEX) software (31, 32). In the end, each ROI of the subject was reviewed by a radiologist. Imaging features were extracted by the PyRadiomics toolbox (33), which is an open-source python software package. To mine rich radiomics, each original image was processed by eight image filters. 1) Wavelet filter: yields eight decompositions per level (all possible combinations of applying either a high- or low-pass filter in each of the three dimensions). 2) Laplacian of Gaussian filter: edge enhancement filter, emphasizes areas of gray-level change, where sigma defines how coarse the emphasized texture should be. A low sigma emphasis on fine textures (change over a short distance), where a high sigma value emphasizes coarse textures (gray-level change over a large distance). 3) Square: takes the square of the image intensities and linearly scales them back to the original range. 4) SquareRoot: takes the square root of the absolute image intensities and scales them back to the original range. 5) Logarithm: takes the logarithm of the absolute intensity + 1. 6) Exponential: takes the exponential, where filtered intensity is $e^{(\text{absolute intensity})}$. 7) Gradient: returns the magnitude of the local gradient. 8) Local Binary Pattern: computes the Local Binary Pattern in a by-slice operation (two-dimensional (2D)) and three-dimensional (3D) using spherical harmonics (34). Then, the features were quantified by the following eight categories of imaging features: 1) first-order statistics with 19 features, 2) 3D shape-based with 16 features, 3) 2D shape-based with 10 features, 4) gray-level co-occurrence matrix (GLCM) with 24 features, 5) gray-level run length matrix (GLRLM) with 16 features, 6) gray-level size zone matrix (GLSZM) with 16 features, 7) neighboring gray-tone difference matrix (NGTDM) with five features, and 8) gray-level dependence matrix (GLDM) with 14 features. In the end,

2,153 quantitative radiological features from each ROI were obtained.

Feature selection and signature building

The Mann–Whitney statistical test (13) was first conducted to distinguish the redundant features. Each feature with a p -value >0.05 was redundant and eliminated. After redundant features were removed, the residual parameters were normalized by the z-score method, which is widely used in machine learning. Then, the feature where the variance is equal to zero was removed again. To further decrease the dimension, the minimum redundancy maximum relevance (mRMR) method was used to determine the most remarkable radiomics features.

Finally, the top-ranked radiomics features were entered into a support vector machine (SVM) classifier to establish a radiomics signature that predicts EGFR mutation status. The parameters of the classifier were optimized by a grid searching technology on the training cohort using 10-fold cross-validation. The radiomics signature with the best accuracy was confirmed. Previous studies have shown that clinical features are associated with the outcome of lymph node metastasis (35). In this study, we found a radiomics signature based on the top-ranked features and then added critical clinical features to explore the predictive score of EGFR mutation status.

Evaluation of radiomics signature

The performance of the radiomics signature in predicting EGFR mutation status was estimated by the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. In addition, accuracy, sensitivity, specificity, and an F1 score were also used to measure the signature.

Statistical analysis

Statistical analysis was performed with SPSS version 22. The independent-samples t-test was used to evaluate the difference in median age between the EGFR-positive and EGFR-negative groups. The chi-square test was used for statistical analysis of gender, tumor stage, smoking history, family history, and tumor position. In the EGFR mutational advanced patients, the cutoff points of statistically significant features were defined by the AUC value of the ROC curve. Survival analysis included patients with disease progression treated with first-line EGFR TKIs. Based on the cutoff points, the chi-square test was used to identify the relationship between radiomics features and the best response. Cox regression analysis was used to explore the predictive

capability of the best response-related features for PFS. Parameters with a *p*-value <0.1 in univariate analysis were selected in multivariate Cox proportional hazards regression analysis. The results were presented as hazard ratio (HR) and 95% CI. The reported statistical remarkable levels were all two-sided, and *p*-values <0.05 were significant.

Results

Patient characteristics

The clinicopathologic features of patients are shown in Table 1. In all patients, 355 patients with confirmed EGFR-positive type were enrolled, while 337 patients were EGFR wild type. Most patients were diagnosed with inoperable stage

III or IV disease (677/692, 97.8%), and 50.4% of 692 patients were former or active smokers (Table 1). Patient characteristics including age, gender, and smoking history were demonstrated to be different between EGFR-positive and EGFR-negative type cohorts, which is consistent with a previous clinical study (Table 1).

Two hundred twenty-five patients with EGFR mutation who experienced disease progression following first-line TKI therapy were included in the efficacy analysis presented in Table 2. The median follow-up time was 1 year (range, 0.7–37.7 months). In the training and validation cohorts (187 and 38 cases, respectively), the results showed no significant difference in PFS (median PFS: training cohort, 12 months; validation cohort, 11.8 months; Mann–Whitney, *p* = 0.304). Moreover, there were also no significant differences (*p* > 0.05) regarding age, gender, smoking history, family history, tumor stage, and position between the two cohorts (Table 2).

TABLE 1 Clinical characteristics of all patients included in the study.

Factors	Training cohort		<i>p</i> -Value	Validation cohort		<i>p</i> -Value
	EGFR-wild	EGFR-mutant		EGFR-wild	EGFR-mutant	
Subject (<i>N</i>)	514			178		
Age (years)	57 ± 9	55 ± 4	<0.001 ^b	57 ± 11	59 ± 9	<0.001 ^b
Gender			<0.001 ^b			<0.001 ^b
Male	171	136		76	40	
Female	78	129	12	50		
Smoking history			<0.001 ^b			<0.001 ^b
Yes	158	95		67	29	
No	91	170	21	61		
Family history			0.417			0.565
Yes	31	26		11	15	
No	218	239	77	75		
TNM stage ^a			0.717			0.076
I	2	2		0	1	
II	5	3	2	0		
III	42	38	11	4		
IV	200	222	75	85		
Tumor position			0.853			0.116
RUL	81	87		18	27	
RML	20	27	12	16		
RLL	45	52	14	16		
LUL	66	63	17	18		
LLL	37	36	27	13		
EGFR mutation type	0	265		0	90	
Wild type	249	0		88	0	
Exon 19 deletion	0	167	0	58		
Exon 21 insertion	0	89	0	31		
Other types	0	9	0	1		

EGFR, epidermal growth factor receptor; RUL, right upper lung; RML, right middle lung; RLL, right lower lung; LUL, left upper lung; LLL, left lower lung.

^aBased on American Joint Committee on Cancer (AJCC) 8th edition.

^bOnly statistically significant (*p* < 0.05) results are reported for analysis.

TABLE 2 Clinical characteristics of patients included in treatment response analysis.

Factors	Training cohort N (%)	Validation cohort N (%)	<i>p</i> -Value
Subject(N)	187 (100)	38 (100)	
Age(years)			0.114
Median	55	57	
Range	29–80	37–75	
Gender			0.707
Male	80 (42.8)	15 (39.5)	
Female	107 (57.2)	23 (60.5)	
Smoking history			0.894
Yes	57 (30.5)	12 (31.6)	
No	130 (69.5)	26 (68.4)	
Family history			0.469
Yes	19 (10.2)	6 (15.8)	
No	168 (89.8)	32 (84.2)	
TNM stage ^a			0.083
III	19 (10.2)	0 (0)	
IV	168 (89.8)	38 (100)	
Tumor position			0.466
RUL	57 (30.5)	9 (23.7)	
RML	17 (9.1)	9 (23.7)	
RLL	33 (17.7)	9 (23.7)	
LUL	53 (28.3)	7 (18.4)	
LLL	27 (14.4)	4 (10.5)	
EGFR-TKI therapy			0.718
Gefitinib	66	13	
Erlotinib	62	15	
Icotinib	59	10	
Median PFS (months)	12	11.8	0.304

PFS, progression-free survival; RUL, right upper lung; RML, right middle lung; RLL, right lower lung; LUL, left upper lung; LLL, left lower lung; EGFR, epidermal growth factor receptor; TKI, tyrosine kinase inhibitor.

^aBased on American Joint Committee on Cancer (AJCC) 8th edition.

Building and validating the predictive radiomics signature for epidermal growth factor receptor mutation status

The feature with a *p*-value >0.05 was excluded using the Mann–Whitney statistical test. Thus, the number of radiomics features was reduced from 2,153 to 1,545. Then, 13 normalized features with variance equal to zero were removed. The residual 1,532 features were sorted using mRMR algorithm to pick the 13 top-ranked features (34), including six (skewness, minimum, kurtosis, variance, minimum, and 10th percentile) in the Firstorder features that describe the distribution of voxel intensities within the image region defined by the mask through commonly used and basic metrics, one (SumSquares) in the gray-level co-occurrence matrix features that describe the second-order joint probability function of an image region constrained by the mask and is defined, three

(SizeZoneNonUniformity, HighGrayLevelZoneEmphasis, and ZoneVariance) in the gray-level size zone matrix features that quantify the number of connected voxels sharing the same gray-level intensity in an image, and three (LargeDependenceHighGrayLevelEmphasis, LargeDependenceHighGrayLevelEmphasis, and DependenceEntropy) in the gray-level dependence matrix features that quantify the number of connected voxels within distance, which are dependent on the center voxel in an image. Then, the top-ranked features and four clinical features (age, gender, smoking, and tumor family history) were input into the SVM classifier to establish a radiomics signature that predicts EGFR mutation status in the training group (*n* = 514). The predictive score of the radiomics signature was assessed on an independent validation group with 178 patients and obtained an AUC of 74.13%, an F1 score of 68.29%, a specificity of 79.55%, an accuracy of 70.79%, and a sensitivity of 62.22% (Figure 2).

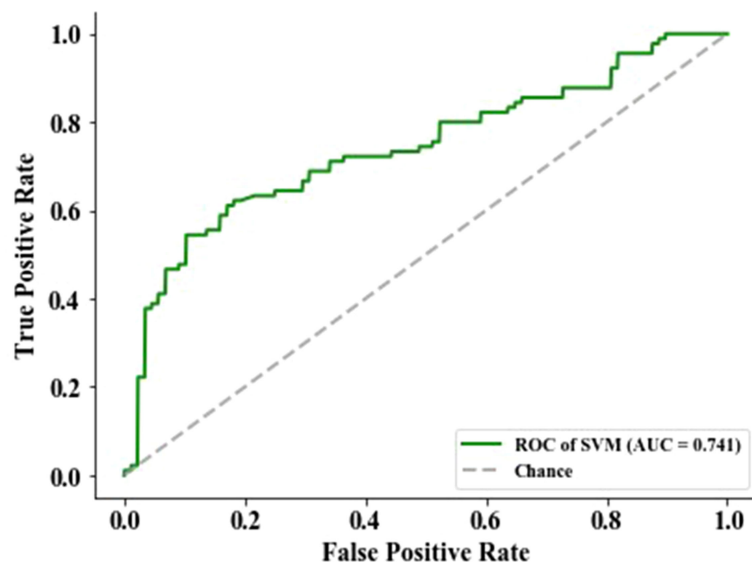


FIGURE 2

The performance of epidermal growth factor receptor (EGFR) status-related radiomics signature was evaluated by the area under the curve (AUC) of the receiver operating characteristic (ROC) curve.

Identification of the best response-related features based on 13 epidermal growth factor receptor mutant-associated features in epidermal growth factor receptor tyrosine kinase inhibitor therapy patients

To identify the imaging biomarkers candidates for the best response of EGFR TKI first-line treatment, the 13 top-rank radiomics features associated with EGFR mutation were further used to explore by logistic analysis. The two features that significantly negatively correlated with the best response were skewness ($p = 0.004$) and 10th percentile ($p = 0.002$) (Table S1). Skewness and 10th percentile were divided into two groups based on predicting the best response. ROC curve was applied to confirm the optimal cutoff points of significant features. For skewness and 10th percentile, the AUC values were 0.832 ($p = 0.004$, Youden's index = 0.614) and 0.653 ($p = 0.002$, Youden's index = 0.289), respectively. The best cutoff points of skewness and 10th percentile, as confirmed by the AUC value, were 0.882 and 21.132, respectively.

Among the results from skewness, the skewness-L (≤ 0.882) group had a superior partial response (PR) rate than had the skewness-H (> 0.882) group (89/117, 76.1% vs 27/108, 25.0%, HR = 9.536, 95% CI: 5.189–17.52, $p < 0.0001$) (Figure 3, Table S2). For the 10th percentile, the SD/PD rate was inferior in the 10th percentile-H group (> 21.132) than in the 10th percentile-L group (≤ 21.132) (76/130, 58.5% vs 33/95, 34.7%, HR = 2.644, 95% CI:

1.529–4.574, $p = 0.0005$) (Figure 3, Table S2). In conclusion, we suggest that skewness and 10th percentile may be better predictive markers for differentiating response to first-line EGFR TKIs.

Testing the correlation between the best response-associated features and progression-free survival

To explore whether advanced lung cancer patients with a good curative outcome can be distinct, we tested the defined cutoff points of the skewness of first-orders (≤ 0.882 versus > 0.882) and the 10th percentile of first-orders (≤ 21.132 versus > 21.132) in the training cohort ($n = 187$). Univariate analysis revealed that the skewness > 0.882 ($p = 0.001$) and the 10th percentile > 21.132 ($p = 0.015$) before treatment were associated with a significantly worse PFS. We then carried out a multivariate Cox proportional regression analysis containing these covariates to ensure independent factors. The relationship between two features and PFS was obvious in multivariate analysis; for the skewness, HR = 1.722, 95% CI: 1.261–2.352, $p = 0.001$ (Figure 4A, Table 3); for the 10th percentile, HR = 1.466, 95% CI: 1.085–1.981, $p = 0.013$ (Figure 4B, Table 3). Therefore, the skewness and 10th percentile of first-order features at baseline level could be used to predict the efficacy in EGFR-mutant advanced lung adenocarcinoma following standard first-line EGFR-TKI therapy.

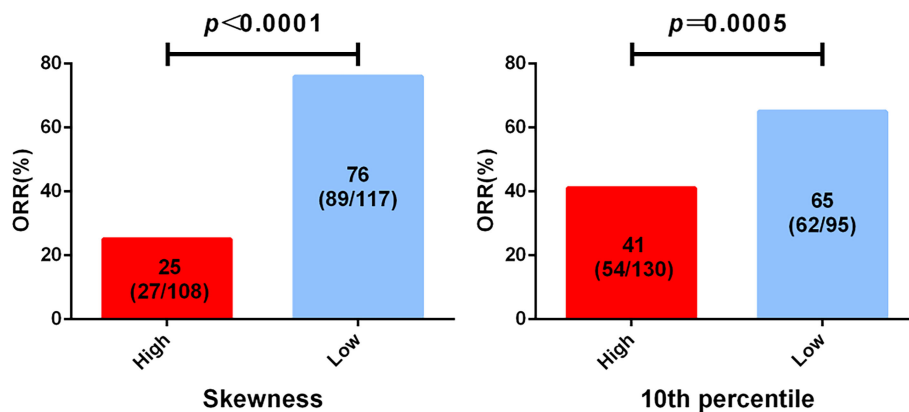


FIGURE 3

Analysis of epidermal growth factor receptor (EGFR) mutation-associated features from computed tomography (CT) imaging before treatment and the best clinical response to tyrosine kinase inhibitor (TKI) first-line therapy. All patients were divided into two groups according to the cutoff of skewness and 10th percentile.

Validation of the predictive effect of skewness and 10th percentile features for progression-free survival

Next, we intended to validate the clinical effect of the skewness and 10th percentile of first-order features in the validation cohort ($n = 38$). Here, we used the aforesaid cutoff value in the training group: high skewness value (>0.882) versus low (≤ 0.882) and high 10th percentile value (>21.132) versus low (≤ 21.132). The correlation between the skewness of first-order and PFS is consistent with the training cohort (Figure 4C, Table 3). Probably due to the limited number of samples, we did not observe a statistical difference between the 10th percentile of first-order before treatment and PFS in the validation cohort (Table 3). However, the skewness of first-order is an effective biomarker that could better indicate the response of first-line EGFR-TKI therapy.

Discussion

Therapeutic opportunities for EGFR mutant lung adenocarcinoma patients have radically changed because of the application of EGFR-TKI therapy. The response varies markedly, and more objective markers are needed to identify patients best suited for certain targeted therapies (4, 9, 15). EGFR mutation types are a well-studied biomarker of response to TKI therapy (6, 7, 36–41). Next-generation sequencing of tissue samples is the standard technique of EGFR status detection. Nevertheless, a biopsy is an invasive procedure of locating tissue that ignores organizational heterogeneity of tumor and microenvironment where distinct bioactive molecules can drive tumor development and progression.

Radiomics is a non-invasive technology that collects routine clinical medical images to assess the tumor phenotype (16, 17). Previous studies used clinical and radiomics models to predict EGFR mutation status (13, 14, 21–26). The radiomics features combined with the clinical factors had a greater prediction effect (13, 14, 22). For example, Aerts et al. found that homogeneity, inverse variance, sum entropy, short-run emphasis, maximum diameter, and tumor volume radiomics features had important roles in discriminating EGFR mutant status in lung adenocarcinoma (14). These features belong to GLCM, GLRLM, and shape features that reveal that EGFR mutation is more likely to be heterogeneous. Similarly, Ye et al., in a single group association study of lung adenocarcinomas, showed that CT imaging characteristics including bubble-like lucency and homogeneous enhancement were remarkably independent predictive factors for EGFR-activating mutation (22). The deep learning model also revealed that the deep learning features such as circle or arch shapes and horizontal and diagonal edges had a significant correlation between high-dimensional CT image characteristics and EGFR genotype (26). Based on previous research progress, we carried out new research for further exploration in this study, and we confirmed a radiomics signature by the SVM classifier combined with four clinical factors to forecast EGFR status in advanced lung adenocarcinoma using preoperative three-dimensional CT images. The radiomics signature showed strong predictive performance in the test group (AUC, 0.7413; specificity, 79.55%; accuracy, 70.79%). We found 13 radiological features that were remarkably associated with EGFR mutations; the first-order category had six features such as skewness, minimum, kurtosis, variance, and 10th percentile, which describes the distribution of voxel intensities within the image region defined by the mask through commonly used and basic

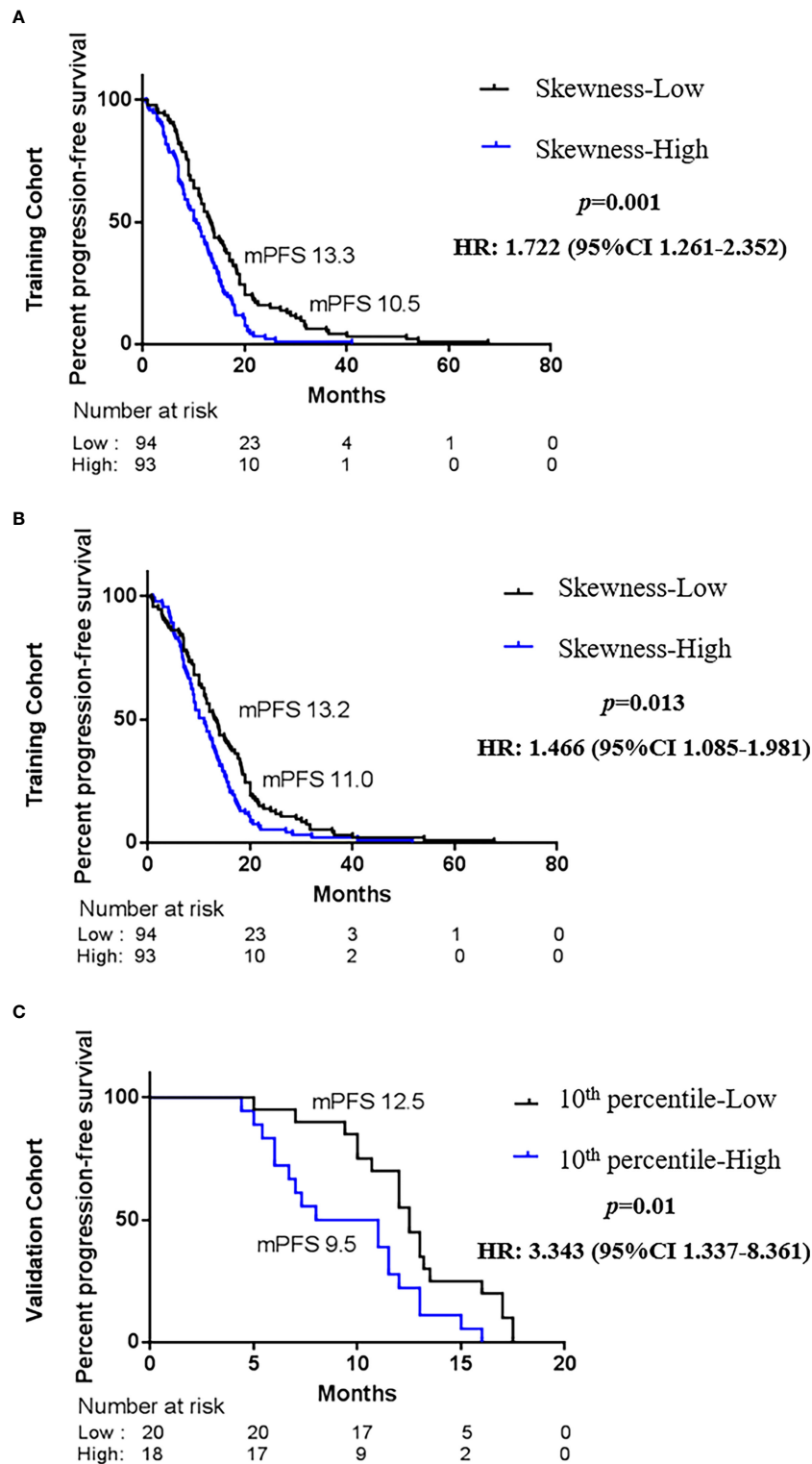


FIGURE 4

Kaplan–Maier survival curves of progression-free survival under biomarker-defined subgroups. In the tyrosine kinase inhibitor (TKI) therapy training cohort, (A) stratification by the skewness of first-order category (low ≤ 0.882 versus high > 0.882); (B) stratification by the 10th percentile of first-order category (low ≤ 21.132 versus high > 21.132). In the TKI therapy validation cohort, (C) stratification by the skewness of first-order category (low ≤ 0.882 versus high > 0.882). p -Values are calculated with multivariate Cox models adjusted by age, gender, smoking history, family history, TNM stage, and tumor position.

TABLE 3 Multivariate analysis of the categorization of two features and PFS.

Factors	Categorization	Training cohort (N=187)		Validation cohort (N=38)	
		HR (95% CI)	p-Value	HR (95% CI)	p-Value
Age	Continuous	0.992 (0.976–1.009)	0.368	1.015 (0.976–1.055)	0.461
Gender	Female vs male	1.181 (0.751–1.856)	0.471	0.719 (0.154–3.360)	0.675
Smoking history	No vs yes	0.994 (0.609–1.621)	0.98	0.822 (0.180–3.744)	0.8
Family history	No vs yes	1.246 (0.764–2.031)	0.378	1.019 (0.357–2.907)	0.972
TNM stage ^a	III vs IV	0.854 (0.517–1.413)	0.54	0	0
Tumor position	RUL vs RML vs RLL vs LUL vs LLL	0.980 (0.887–1.083)	0.695	0.879 (0.671–1.153)	0.352
Skewness	≤Cutoff1 ^c vs >Cutoff1 ^c	1.722 (1.261–2.352)	0.001 ^b	3.343 (1.337–8.361)	0.01 ^b
Age	Male vs female	0.995 (0.979–1.012)	0.579	0.998 (0.961–1.036)	0.905
Gender	Continuous	1.154 (0.740–1.800)	0.528	1.694 (0.425–6.748)	0.455
Smoking history	0 or 1 vs 2	0.870 (0.534–1.416)	0.575	0.437 (0.109–1.751)	0.243
Family history	No vs yes	1.128 (0.694–1.833)	0.628	0.638 (0.237–1.718)	0.374
TNM stage ^a	First vs second	0.682 (0.418–1.111)	0.124	0	0
Tumor position	RUL vs RML vs RLL vs LUL vs LLL	0.996 (0.902–1.101)	0.943	1.013 (0.798–1.285)	0.919
10th percentile	≤Cutoff2 ^c vs >Cutoff2 ^c	1.466 (1.085–1.981)	0.013 ^b	1.122 (0.492–2.561)	0.784

PFS, progression-free survival; EGFR, epidermal growth factor receptor; TKI, tyrosine kinase inhibitor; RUL, right upper lung; RML, right middle lung; RLL, right lower lung; LUL, left upper lung; LLL, left lower lung; HR, hazard ratio; CI, confidence interval.

^aBased on American Joint Committee on Cancer (AJCC) 8th edition.

^bOnly statistically significant ($p < 0.05$) results are reported for analysis.

^cCutoff1 = 0.882; Cutoff2 = 21.132.

metrics (34). The GLSZM and GLDM categories had three features each. The GLSZM is defined as the gray-level zone quantization of connected voxel numbers that share the same gray-level intensity in an image, while GLDM quantifies gray-level dependencies defined as the number of connected voxels within a distance that are dependent on the center voxel in an image. Together, the representation of these features indicates tumor heterogeneous related to EGFR mutation phenotype, which provides an alternative non-invasive way easily to forecast EGFR status in routine CT diagnosis and supply a good supplement to biopsy in real clinical practice.

Radiomics markers that can predict the efficacy of first-line EGFR-TKI therapy are now more needed; we also have found two CT features for progression risk stratification to first-line EGFR-TKI remedy in advanced lung adenocarcinoma. The skewness and 10th percentile of first-order features included in ROI preprocessed by gradient and exponential filter, respectively, were significantly negatively correlated with the best response and PFS of EGFR TKI therapy. Few studies used radiomics to explore the response of targeted therapy. For example, Jie Tian et al. extracted features from two-directional CT images and used the LASSO Cox regression analysis to select 12 CT features for discriminating between patients with rapid and slow progression to EGFR TKI therapy. The 12 CT features are part of GLCM, GLRLM, and first-order features (30). However, our study retrospectively obtained more comprehensive data of the tumors in 3D from CT images and was the first to further explore the relationship between EGFR mutation-related features and the response of EGFR TKI

therapy in advanced lung adenocarcinoma. To date, several studies have found the first-order features with response and prognosis of EGFR TKI therapy, including energy, standard deviation, uniformity, and entropy (26, 30), but we are the first to find the predictive value of other first-order features for the best response and PFS. The skewness feature assesses the asymmetry of the distribution of values about the mean value, while the 10th percentile represents the first 10% proportion of voxels with positive order of susceptibility. The two features indicated the whole tumor heterogeneity and the asymmetry distribution of tumor parenchyma, which corresponds to the inhomogeneity of gross findings in CT images checked by the radiologist. It could explain why radiomics characteristics reveal treatment outcomes, while further work is needed to explore the potential mechanisms of the above features and predict the efficacy of lung cancer.

There are several limitations that could not be ignored. First, this was a retrospective study and CT images were acquired with 5-mm slice thicknesses, which is indeed used in hospitals. Although we may ignore some tumor information, our results are certainly closer to practice. Second, given that this was a single-center study, the study lacks an external validation group of patients, which is a key component of radiological analyses, and required validation in a larger patient population study. Third, our investigation only concentrated on EGFR mutation status. The interrelationship among radiomics features, EGFR, and other driver mutations (i.e., ROS-1, ALK, and c-Met) is unknown but could be studied in future research. Nonetheless, the study still had significant positive results. Further assessment

of two indicators could be contained together with other predictive biomarkers in the evaluation of lung and other solid tumor patients who are candidates for treatment efficacy.

Data availability statement

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

Ethics statement

Ethical review and approval was not required for the study of human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

Conception and design: LW, MJ, PY, and JiL. Acquisition of data (acquired and managed patients): MJ, PY, JiL, WP, XP, BC, JaL, and JW. Analysis and interpretation of data: MJ and PY. Writing, review, and/or revision of the manuscript: MJ, PY, and JiL. LW designed and supervised the study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Hunan Provincial Natural Science Foundation of China (No. 2021JJ30430), Wu Jieping

Medical Foundation (No. 320.6750.19088-11) and the Research Foundation of Chinese Society of Clinical Oncology (No. Y-2019Genecast-024 and No. Y-HR2019-0464) to LW. This work was supported by the National Cancer Center (NCC2017A17) to PY.

Acknowledgments

We thank all patients, family, nurses, and doctors for their contribution to this study.

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.985284/full#supplementary-material>

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics. *CA Cancer J Clin* (2018) 68:7–30. doi: 10.3322/caac.21442
2. Alamgeer M, Ganju V, Watkins DN. Novel therapeutic targets in non-small cell lung cancer. *Curr Opin Pharmacol* (2013) 13:394–401. doi: 10.1016/j.coph.2013.03.010
3. Savas P, Hughes B, Solomon B. Targeted therapy in lung cancer: IPASS and beyond, keeping abreast of the explosion of targeted therapies for lung cancer. *J Thorac Dis* (2013) 5 Suppl 5:S579–92. doi: 10.3978/j.issn.2072-1439.2013.08.52
4. Chan BA, Hughes BG. Targeted therapy for non-small cell lung cancer: current standards and the promise of the future. *Transl Lung Cancer Res* (2015) 4 (1):36–54. doi: 10.3978/j.issn.2218-6751.2014.05.01
5. Gadgeel SM. Personalized therapy of non-small cell lung cancer (NSCLC). *Adv Exp Med Biol* (2016) 890:203–22. doi: 10.1007/978-3-319-24932-2_11
6. 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 (10):947–57. doi: 10.1056/NEJMoa0810699
7. Gao G, Ren S, Li A, Xu J, Xu Q, Su C, et al. Epidermal growth factor receptor-tyrosine kinase inhibitor therapy is effective as first-line treatment of advanced non-small-cell lung cancer with mutated EGFR: A meta-analysis from six phase III randomized controlled trials. *Int J Cancer* (2012) 131(5):E822–9. doi: 10.1002/ijc.27396
8. Li T, Kung HJ, Mack PC, Gandara DR. Genotyping and genomic profiling of non-small-cell lung cancer: implications for current and future therapies. *J Clin Oncol* (2013) 31(8):039–1049. doi: 10.1200/JCO.2012.45.3753
9. Chong CR, Janne PA. The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nat Med* (2013) 19(11):1389–400. doi: 10.1038/nm.3388
10. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. *Nature* (2013) 501(7467):355–64. doi: 10.1038/nature12627
11. Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res* (2012) 72(19):4875–82. doi: 10.1158/0008-5472.CAN-12-2217
12. Taniguchi K, Okami J, Kodama K, Higashiyama M, Kato K. Intratumor heterogeneity of epidermal growth factor receptor mutations in lung cancer and its correlation to the response to gefitinib. *Cancer Sci* (2008) 99(5):929–35. doi: 10.1111/j.1349-7006.2008.00782.x

13. Girard N, Sima CS, Jackman DM, Sequist LV, Chen H, Yang JC, et al. Nomogram to predict the presence of EGFR activating mutation in lung adenocarcinoma. *Eur Respir J* (2012) 39(2):366–72. doi: 10.1183/09031936.00010111
14. 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
15. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res* (2013) 19(8):2240–7. doi: 10.1158/1078-0432.CCR-12-2246
16. 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
17. 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
18. Gevaert O, Xu J, Hoang CD, Leung AN, Xu Y, Quon A, et al. Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data-methods and preliminary results. *Radiology* (2012) 264:387–96. doi: 10.1148/radiol.12111607
19. Aerts HJ WL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P. Decoding tumour phenotype by non-invasive imaging using a quantitative radiomics approach. *Nat Commun* (2014) 5:4006. doi: 10.1038/ncomms5006
20. Zhou M, Leung A, Echegaray S, Gentles A, Shrager JB, Jensen KC, et al. Non-small cell lung cancer radiomics map identifies relationships between molecular and imaging phenotypes with prognostic implications. *Radiology* (2018) 286:307–15. doi: 10.1148/radiol.2017161845
21. Sugano M, Shimizu K, Nakano T, Kakegawa S, Miyamae Y, Kaira K, et al. Correlation between computed tomography findings and epidermal growth factor receptor and KRAS gene mutations in patients with pulmonary adenocarcinoma. *Oncol Rep* (2011) 26(5):1205–11. doi: 10.3892/or.2011.1412
22. 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
23. 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
24. Zhou JY, Zheng J, Yu ZF, Xiao WB, Zhao J, Sun K. 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
25. Rizzo S, Petrella F, Buscarino V, De Maria F, 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(1):32–42. doi: 10.1007/s00330-015-3814-0
26. 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
27. Dingemans AC, de Langen AJ, van den Boogaart V, Marcus JT, Backes WH, Scholtens HTGM, et al. First-line erlotinib and bevacizumab in patients with locally advanced and/or metastatic non-small-cell lung cancer: A phase II study including molecular imaging. *Ann Oncol* (2011) 22:559–66. doi: 10.1093/annonc/mdq391
28. Nishino M, Dahlberg SE, Cardarella S, Jackman DM, Rabin MS, Hatabu H. Tumor volume decrease at 8 weeks is associated with longer survival in EGFR-mutant advanced non-small-cell lung cancer patients treated with EGFR TKI. *J Thorac Oncol* (2013) 8:1059–68. doi: 10.1097/JTO.0b013e318294c909
29. Dai D, Li XF, Wang J, Liu JJ, Zhu YJ, Zhang Y. Predictive efficacy of 11 c-PD153035 PET imaging for EGFR-tyrosine kinase inhibitor sensitivity in non-small cell lung cancer patients. *Int J Cancer* (2016) 138:1003–12. doi: 10.1002/ijc.29832
30. Song J, Shi J, Dong D, Fang M, Zhong W, Wang K, et al. A new approach to predict progression-free survival in stage IV EGFR-mutant NSCLC patients with EGFR-TKI therapy. *Clin Cancer Res* (2018) 24(15):3583–92. doi: 10.1158/1078-0432.CCR-17-2507
31. Perez NP, Guevara Lopez MA, Silva A, Ramos I. Improving the Mann-Whitney statistical test for feature selection: an approach in breast cancer diagnosis on mammography. *Artif Intell Med* (2015) 63(1):19–31. doi: 10.1016/j.artmed.2014.12.004
32. Zhang L, Fried DV, Fave XJ, Hunter LA, Yang J, Court LE. IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. *Med Phys* (2015) 42(3):1341–53. doi: 10.1118/1.4908210
33. 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
34. 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(2):328–38. doi: 10.1148/radiol.2020191145
35. Wang X, Zhao X, Li Q, Xia W, Peng Z, Zhang R, et al. Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT. *Eur Radiol* (2019) 29(11):6049–58. doi: 10.1007/s00330-019-06084-0
36. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, Tan EH, Hirsh V, Thongprasert S, et al. Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* (2005) 353(2):123–32. doi: 10.1056/NEJMoa050753
37. Sequist LV, Yang JCH, Yamamoto N, O'Byrne K, Hirsh V, Mok T. Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. *J Clin Oncol* (2013) 31(27):3327–34. doi: 10.1200/JCO.2012.44.2806
38. Mok TS, Wu YL, Ahn MJ, Garassino MC, Kim HR, Ramalingam SS. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. *N Engl J Med* (2017) 376(7):629–40. doi: 10.1056/NEJMoa1612674
39. Zhang Y, Sun Y, Pan Y, Li C, Shen L, Li Y, et al. Frequency of driver mutations in lung adenocarcinoma from female never-smokers varies with histologic subtypes and age at diagnosis. *Clin Cancer Res* (2012) 18(7):1947–53. doi: 10.1158/1078-0432.CCR-11-2511
40. Song Z, Zhu H, Guo Z, Wu W, Sun W, Zhang Y. Correlation of EGFR mutation and predominant histologic subtype according to the new lung adenocarcinoma classification in Chinese patients. *Med Oncol* (2013) 30(3):645. doi: 10.1007/s12032-013-0645-1
41. Jia Y, Yun CH, Park E, Ercan D, Manuia M, Juarez J. Overcoming EGFR (T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors. *Nature* (2016) 534:129–32. doi: 10.1038/nature17960