Cytomorphological and histomorphological features of lung adenocarcinoma with epidermal growth factor receptor mutation and anaplastic lymphoma kinase gene rearrangement

NIKOLA GARDIĆ[1,2], ALEKSANDRA LOVRENSKI[1,2], VANESA SEKERUŠ[2,3], SVETLANA KAŠIKOVIĆ LEČIĆ[2,4], MILORAD BIJELOVIĆ[2,5], TANJA LAKIĆ[1,6], ALEKSANDRA ILIĆ[1,6], BOJAN ZARIĆ[2,4] and SOFIJA GLUMAC[7]

[1]Department of Pathology, Faculty of Medicine, University of Novi Sad, Novi Sad 21000, Serbia; [2]Institute for Pulmonary Diseases of Vojvodina, Sremska Kamenica 21204, Serbia; [3]Department of Biochemistry, Faculty of Medicine, University of Novi Sad, Novi Sad 21000, Serbia; [4]Department of Internal Medicine, Faculty of Medicine, University of Novi Sad, Novi Sad 21000, Serbia; [5]Department of Surgery, Faculty of Medicine Foca, University of East Sarajevo, Foča 73300, Bosnia and Herzegovina; [6]Center of Pathology and Histology, University Clinical Center of Vojvodina, Novi Sad 21000, Serbia; [7]Institute of Pathology, Faculty of Medicine, University of Belgrade, Belgrade 11000, Serbia

Received July 15, 2024; Accepted October 16, 2024

DOI: 10.3892/ol.2024.14786

Abstract. Lung cancer is among the lethal and most prevalent oncological diseases globally. It is known that two types of mutations, namely anaplastic lymphoma kinase (ALK) gene rearrangement and epidermal growth factor receptor (EGFR) gene mutation, are responsible for the development of lung adenocarcinoma. The present study aimed to investigate the differences in the frequency of clinical, cytomorphological and histomorphological features of ALK and EGFR-positive lung adenocarcinomas. The present retrospective study comprised 101 patients diagnosed with lung adenocarcinoma. Based on the molecular findings, the patients were categorized into three groups as follows: The ALK-rearranged group (n=28), the EGFR group (n=42) and the negative group (n=31). The clinical features analyzed included sex, age, smoking status and disease stage. The cytomorphological and histomorphological features examined encompassed the following: Cell cluster size, the arrangement of tumor cells, the size of nuclei, nuclear atypia, the visibility of nucleoli, the presence of necrosis, intracytoplasmic vacuoles, signet ring cells, stromal characteristics and the presence of inflammatory infiltrate presence. The results indicated that the female sex was more prevalent in the EGFR group, but statistically significant differences ($P<0.05$) were observed between the EGFR and negative group. A significantly greater percentage of non-smokers was identified in the EGFR group compared with the ALK group ($P<0.01$). The majority of patients with confirmed ALK or EGFR mutations received onco-specific treatment. Focal and abundant necrosis was significantly less common in cytological samples in the EGFR group than in the other groups (21.43 vs. 57.14 and 51.61%, combined, $P<0.01$). No significant differences were observed in other cytomorphological features between the groups. Intracytoplasmic vacuoles, signet ring cells and cells with visible nucleoli were significantly more frequent in histological specimens of the ALK group ($P<0.01$). The predictive model composed of these features or combined with sex and smoking habits exhibited statistically significant differences for mutation status as a criterion ($P<0.01$). Collectively, the findings of the present study confirmed that, in addition to clinical characteristics, certain cytological and histological features of lung adenocarcinoma are associated with the mutational status of the tumor.

Correspondence to: Dr Aleksandra Lovrenski, Department of Pathology, Faculty of Medicine, University of Novi Sad, 4 Hajduk Veljkova, Novi Sad 21000, Serbia
E-mail: aleksandra.lovrenski@mf.uns.ac.rs

Key words: cytology, histopathology, lung adenocarcinoma, epidermal growth factor receptor, anaplastic lymphoma kinase

Introduction

Lung cancer ranks among the most lethal and most prevalent oncological diseases worldwide. According to the latest 'GLOBOCAN 2022' review of global cancer statistics, it stands as the second most commonly diagnosed type of cancer, trailing only breast cancer. In 2022 alone, almost 2.5 million new cases of lung cancer were diagnosed. Notably, lung cancer holds the grim distinction of being the primary cause of mortality among patients afflicted by malignant diseases, claiming the lives of 1.8 million individuals in 2022. Typically, it manifests in individuals aged ≥70, emerging as the leading cause of oncological fatalities among those aged ≥40 (1,2).

The success of lung cancer treatment is dependent upon various factors, including the clinical characteristics of patients, the tumor histological type, the outcomes of predictive biomarker testing, and effective communications between pathologists, radiologists and oncologists. Over the past decade, substantial strides have been made in therapeutic

development, mainly through identifying and utilizing predictive biomarkers (3).

Invasive non-mucinous adenocarcinoma is the most common type of lung cancer. It comprises malignant epithelial cells whose morphology or immunohistochemical phenotype suggest glandular differentiation, and thus it does not meet the criteria for any other type of adenocarcinoma (4).

Anaplastic lymphoma kinase (ALK) rearrangement encompasses a group of gene mutations encoding the transmembrane receptor tyrosine kinase, belonging to the insulin receptor protein superfamily. To date, >20 rearrangement partners of the ALK gene have been identified, with the most prevalent occurring in non-small cell lung cancer (NSCLC): an intra-chromosomal inversion of the short arm of chromosome 2, resulting in the fusion of the 2p21 gene locus of the echinoderm microtubule-associated protein-like protein 4 (EML4) gene and 2p23 ALK genes (5). Among the genomic alterations observed in NSCLC, the ALK rearrangement is a targetable alteration for therapy, providing a therapeutic response that extends patient survival (6,7). While the histological type of the majority of lung cancers with ALK rearrangement is adenocarcinoma, the studies available to date exploring the detailed histomorphological and cytomorphological characteristics of these samples are limited and have yielded contradictory results (8-12).

Epidermal growth factor receptor (EGFR) belongs to the family of tyrosine kinase protein receptors whose mutation leads to the uncontrolled proliferation of malignant cells, their invasion, metastatic spread, the inhibition of apoptosis, as well as tumor angiogenesis, it plays a leading role in tumor carcinogenesis and progression. These somatic mutations mainly target exons 18-21 of the gene encoding part of the tyrosine kinase domain of EGFR. Among the most well-known and frequently occurring mutations are deletions in exon 19 and substitutions in exon 21, particularly the L858R substitution, which collectively account for 80-90% of all EGFR mutations in NSCLC (13). According to certain studies, the histopathological subtype of adenocarcinoma predicts prognosis and mutational status (14-16).

The scarcity of available data in the literature regarding the association between the microscopic morphology of adenocarcinoma and the status of biomarkers available in Serbia underscores the necessity for research in this area. Furthermore, the present study aimed to provide valuable insight into the morphology of primary lung adenocarcinomas, which may aid in the typing of NSCLC, particularly in cases where only cytological smears are available.

Materials and methods

Study design. The present retrospective study analyzed histological and cytological material from the internal tissue bank obtained from patients diagnosed with lung adenocarcinoma between 1st January 2016 and 31st December 2023 at the Institute for Pulmonary Diseases of Vojvodina (Sremska Kamenica, Serbia). The study with research methodology including the use of the external controls for the analysis, was approved by the Institutional Professional and Ethics Committee of the Institute for Pulmonary Diseases of Vojvodina (approval nos. 25-VIII/10 and 24-VII/10). A total of 101 patients were included in the study (mean age, 63.48 and 32 to 84 years old, respectively. The sex distribution of the patients was nearly balanced, with a slight predominance of males (52.48 vs. 47.52%). Patients were divided into the ALK, EGFR and negative groups. The clinical features analyzed included sex, age, smoking status and disease stage. The data were collected from the patients' medical records.

Molecular analyses. A representative paraffin block was selected for the immunohistochemical analysis of ALK rearrangement based on the hematoxylin and eosin (H&E)-stained section. The block was subsequently cut into histological sections that were 4-μm-thick. Appendices removed during appendectomies were used as external control tissue (Fig. S1). Paraffin samples were melted onto the slides overnight at 53˚C. The following day, the slides were labeled and placed in the Benchmark, Ventana Roche machine using an anti-ALK antibody (Rabbit Monoclonal D5F3; ready-to-use; Roche Tissue Diagnostics; cat. no. 790-4794/06679072001) and operated according to the manufacturer's protocol. The antibody for the analysis was incubated for 16 minutes at a temperature of 36˚C.

In all patients included in the study, the qualitative detection and identification of mutations in exons 18, 19, 20 and 21 of the EGFR gene were determined using the real-time PCR Cobas EGFR Mutation Test v2 (cat. no. P/N 07248563190) after DNA isolation with the Cobas DNA Sample Preparation kit. The entire process of amplification, detection, and validation of the samples was conducted using the Cobas 4800 software on the Cobas z 480 analyzer following the manufacturer's protocol (Roche Diagnostics).

All samples for ALK and EGFR testing were selected based on the number of viable tumor cells. The degree of differentiation and the Ki67 proliferative index were not considered in the selection of test samples and, therefore, did not influence the study's results.

Sampling and processing of material. All histological material was obtained through bronchoscopic methods, including bronchial biopsy, transbronchial biopsy and catheter biopsy. Tissues were fixed in 10% neutral formalin for 12-18 h (room temperature), dehydrated in increasing ethanol concentrations (70, 80, 96 and 100%), embedded in paraffin, and cut into 4-μm thick sections using a rotary microtome (Leica Microsystems GmbH). All sections were stained with H&E (Bio-Optica). After rehydration, the sections were stained with hematoxylin for 1 min. The slides were then rinsed in running tap water, followed by differentiation in an acid-alcohol solution to remove excess stain. In the next step, the sections were counterstained with eosin for 1.5 min. The slides were briefly rinsed in water to remove excess eosin and then rehydrated through a graded series of alcohols.

Material for cytopathological analysis was collected via thoracentesis, percutaneous lymph node fine needle aspiration and bronchoscopic methods, including transbronchial fine needle aspiration, brush biopsy and catheter biopsy. Conventional cytological smears were prepared and stained using the May-Grunwald-Giemsa method. The entire

Table I. The method of estimation for cytomorphological features.

Cytomorphological features	Categories		
Size of clusters	≤200 μm	>200 μm, ≤400 μm	>400 μm
Size of nuclei[a]	<3X lymphocyte	3-5xlymphocyte	<5X lymphocyte
Nuclear atypia	Moderate	Severe	
Nucleolar visibility	Visible	Not visible	
Intra-cytoplasmic vacuoles	<20%	≥20%	
Signet ring cells	<5%	≥5%	
Necrosis	Absent	Moderate	Abundant

[a]Nuclear size was estimated according to size of lymphocytes.

Table II. The method of estimation for histomorphological features.

Histomorphological features	Categories		
Predominant arrangement	Lepidic	Papillary/acinar	Solid/micropapillary
Cribriform pattern	No	Yes	
Stroma	Poor	Moderate	Abundant
Size of nuclei[a]	<3X lymphocyte	3-5X lymphocyte	<5X lymphocyte
Nuclear atypia	Moderate	Severe	
Nucleoli visibility	Visible	Not visible	
Intra-cytoplasmic vacuoles	No	Yes	
Signet ring cells	No	Yes	
Inflammatory infiltrate	Poor	Moderate	Abundant
Necrosis	Absent	Moderate	Abundant

[a]Nuclear size was estimated according to size of lymphocytes.

histological and cytological material was evaluated using a light microscope (Leica, DM2500).

Cytomorphological features. Cytomorphological features encompass qualitative characteristics of cellular arrangements observed on cytology smears. The following parameters were analyzed: Size of cell clusters, the size of nuclei, nuclear atypia, visibility of nucleoli, the presence of intracytoplasmic vacuoles, signet ring cells, and necrosis. The method of parameter estimation is outlined in Table I.

Histomorphological features. The analyzed histomorphological characteristics included cell arrangement, the presence of cribriform arrangement, the amount of stroma, nucleus size, degree of nuclear atypia, the visibility of nuclei, the presence of intracytoplasmic vacuoles, the presence of signet ring cells, the presence of inflammatory infiltrate and the presence of necrosis. The method of parameter estimation is detailed in Table II. In the present study, two experienced cytopathologists and a pathology resident evaluated both types of sample cases. Disagreements in estimations were reanalyzed, after which a joint decision was made on the result.

The inflammatory infiltrate was classified into three categories. The infiltrate was considered poor when the tumor was infiltrated with scant inflammatory cells, typically <5% of the tumor area, with sparse lymphocytes and/or neutrophils scattered across the stroma with minimal clustering. It was considered moderate when inflammatory cells comprised 5-30% of the tumor area, with more noticeable clusters of lymphocytes, neutrophils and occasional plasma cells within the stroma and surrounding tumor cells. Abundant inflammatory infiltrate was defined by dense and widespread inflammatory infiltrate, occupying >30% of the tumor area, with prominent clusters of lymphocytes, neutrophils and plasma cells throughout the stroma and within the tumor.

The presence of a partially preserved structure of tumor cells with remains of their outlines in a foci of necrotic tissue was considered a sign of tumor necrosis. Necrosis was assessed as moderate when the sample contained viable primary tumor cells with small foci of necrosis comprising <30% of the sample. Necrosis was classified as abundant when it occupied >30% of the sample or when the sample predominantly consisted of necrotic foci, with viable tumor cells present in only small amounts, complicating the interpretation. The presence of necrotic masses of unrecognizable cells mixed with purulent exudate is considered necrosis of inflammatory/infectious etiology, and such cases are not categorized as tumor necrosis (10,17-20).

The size of the nuclei is expressed as the size of the lymphocyte. Nuclear atypia was scored as moderate or severe nuclear

atypia. The specimen was categorized as moderate nuclear atypia when nuclei were uniform in size and shape, with mild irregularity of the nuclear membrane and homogenous or fine granular chromatin pattern. The specimen was categorized as severe nuclear atypia in cases with varied sizes of nuclei with bizarre shapes and coarse chromatin patterns.

Statistical analysis. Statistical analysis was performed using JASP 0.18.3.0 software (https://jasp-stats.org/). The difference in the frequency of cyto- and histomorphological features relative to the type of mutation was assessed using the Chi-squared test or Fisher's exact test, depending on the conditions met for each analysis. The same statistical method was also used to examine tumor morphology relative to the other clinical parameters. Cyto- and histomorphological features were analyzed as predictors using binary logistic regression, with the mutation status as the criterion. The analysis was performed using the default settings in JASP, applying a logit link function. Predictor variables were entered using the forced entry method. Tolerance values >0.2 and variance inflation factor (VIF) values <5 were used as criteria to confirm the absence of multicollinearity. Model validation was conducted through cross-validation using the k-Nearest Neighbors test. Predictor significance was evaluated using Wald's chi-square tests, and odds ratios (ORs) were calculated with corresponding 95% confidence intervals (CIs). The results are presented tabularly and graphically, with $P<0.05$ considered to indicate a statistically significant difference.

Results

Clinical characteristics. The clinical characteristics of the patients are summarized in Table III. Sex distribution in the ALK group was uniform, while the EGFR group had a higher proportion of female patients (57.14 vs. 42.86%). In the negative group, male patients comprised almost 70% of the cohort. Statistically significant differences in sex distribution were observed between the EGFR and negative groups ($\chi^2=4.439$; $P<0.05$; data not shown). By contrast, no significant differences were noted between the EGFR and ALK group ($\chi^2=0.345$; $P>0.05$; data not shown) or between ALK and negative group ($\chi^2=1.919$; $P>0.05$; data not shown). The mean age of the patients was 64 years, with no statistically significant differences in age observed between the groups. Notably, smoking habits differed significantly between the groups: The patients in the EGFR group were significantly more likely to be non-smokers compared with those in the negative and ALK groups ($P<0.001$ and $P<0.05$, respectively). Furthermore, non-smokers and ex-smokers were significantly more prevalent in the ALK group than in the negative group ($\chi^2=6.679$; $P<0.05$). Conversely, patients in the ALK group were significantly more likely to be smokers compared with those in the EGFR group. There was an approximately equal distribution between the two most common EGFR mutations. The majority of patients had stage IV of the disease. As regards treatment, the majority of patients with confirmed ALK or EGFR mutations received onco-specific treatment. In the group of ALK-positive patients, 82.14% received onco-specific therapy, with alectinib being the most commonly administered, accounting for 80% of these cases. In the EGFR-positive group, 80.96% of patients received onco-specific therapy, with afatinib and gefitinib being the most frequently used drugs. At the time of diagnosis, osimertinib was either unavailable in Serbia or accessible only through clinical trials. Consequently, only 4.76% of patients received osimertinib as a first-line treatment, while an additional 9.52% (4 patients) received it as a second-line treatment following the detection of T790M mutation resistance upon retesting. In the negative group, 61.29% of patients received chemotherapy with or without radiotherapy, while 22.58% received palliative care.

Cytomorphological features. The present study examined the differences in the frequency of cytomorphological features of lung adenocarcinoma between the groups (Table IV). There were no statistically significant differences in the frequency of different sizes of cell clusters, size of the nuclei, nuclear atypia and visibility of nuclei between the groups. Although intracytoplasmic vacuoles and signet ring cells were more frequently present in the ALK group, this difference was not statistically significant ($P>0.05$). The only statistically significant difference between the groups was observed regarding necrosis. Necrosis was significantly prevalent in the ALK and negative group samples than in the EGFR group ($P<0.05$; Fig. 1).

Histomorphological features. The histomorphological features of patients are presented in Table V. Statistically significant differences in the tissue arrangement of adenocarcinoma among the groups were observed ($P<0.01$). In the ALK group, no samples exhibited lepidic or micropapillary arrangements. By contrast, these arrangements were present in 16.67 and 9.51% of the EGFR group, respectively, and in 9.68 and 3.12% of the negative group, respectively. Conversely, papillary arrangement was detected in 21.43% of the ALK-positive adenocarcinomas, while no such arrangement was observed in the EGFR and negative groups.

Visible nucleoli, the presence of intracytoplasmic vacuoles, and the presence of signet ring cells (Fig. 2) were statistically significantly more frequently observed features in the ALK group ($P<0.01$; Figs. 3-5).

No statistically significant differences were found regarding the amount of inflammatory infiltrate, stroma and cribriform cell arrangement between the groups. Additionally, the groups exhibited no statistically significant differences in nuclear size and a degree of nuclear atypia.

A binary logistic regression analysis was conducted to evaluate the association relationship and predictive accuracy of specific microscopic features (cell arrangement, nucleoli visibility, intracytoplasmic vacuoles, signet ring cells and necrosis in smears) in determining the mutational status of lung adenocarcinoma (Tables VI and VII). When compared with the negative group, the collective influence of these variables as predictors of EGFR positivity was statistically significant ($P=0.029$), with pseudo-R^2 values ranging from 0.171 to 0.280. Notably, focal necrosis in smears was associated with a 5.2-fold reduction in the likelihood of EGFR positivity. The presence of signet ring cells led to a 96.5% decrease in the probability of EGFR positivity ($P<0.05$). Incorporating sex and smoking status into the model enhanced its predictive accuracy by 14%, increasing it from 72.60 to 86.30%. The corresponding confusion matrix is presented in Tables VIIIA and IXA. Cross-validation of the model is provided in Tables SI-SIII and Figs. S2 and S3.

Table III. Patient's clinical characteristics.

Clinical characteristics	ALK	EGFR	Negative	Total	Test value	P-value
Total	28	42	31	101		
Sex						
Male	14 (50%)	18 (42.86%)	21 (67.74%)	53 (52.48%)	4.524	0.104
Female	14 (50%)	24 (57.14%)	10 (32.26%)	48 (47.52%)		
Age, years						
Mean ± SD	62.39±10.74	65.52±8.75	61.68±7.28	63.48±9.03	F 1.930	0.151
Median	63	66	62	64		
Minimum	32	46	49	32		
Maximum	83	84	79	84		
Smoking history						
Non-smokers	6 (21.43%)	21 (50%)	2 (6.45%)	29 (28.71%)	30.105	<0.001
Ex-smokers	6 (21.43%)	11 (26.19%)	2 (6.45%)	19 (18.81%)		
Smokers	16 (57.14%)	10 (23.81%)	27 (87.10%)	53 (52.48%)		
Mutation type						
Exon 19	N/A	22 (52.38%)	N/A	N/A	N/A	N/A
Exon 21	N/A	20 (47.62%)	N/A	N/A		
Stage						
IB	0	1 (2.38%)	0	1 (0.99%)	27.635	0.035
IIA	0	1 (2.38%)	1 (3.23%)	2 (1,98%)		
IIB	1 (3.59%)	1 (2.38%)	1 (3.23%)	3 (2.97%)		
IIIA	0	3 (7.14%)	1 (3.23%)	4 (3.96%)		
IIIB	9 (32.14%)	5 (11.91%)	4 (12.90%)	18 (17.82%)		
IIIC	0	3 (7.15%)	1 (3.23%)	4 (3.96%)		
IVA	18 (64.29%)	14 (33.33%)	12 (38.70%)	44 (43.56%)		
IVB	0	14 (33.33%)	11 (35.48%)	25 (24.75%)		
Treatment	Crizotinib	Afatinib	ChT ± RT[b]		N/A	N/A
	3 (10.71%)	16 (38.10%)	19 (61.29%)			
	Alectinib	Gefitinib	Palliative care			
	17 (60.72%)	11 (26.19%)	7 (22,58%)			
	Brigatinib	Erlotinib	Died[a]			
	3 (10.71%)	5 (11.90%)	2 (6,45%)			
	Chemotherapy	Osimertinib	No data			
	1 (3.57%)	2 (4.76%)	3 (9,68%)			
	No data	Palliative care				
	4 (14.29%)	1 (2.38%)				
		Died[a]				
		3 (7.14%)				
		No data				
		4 (9.52%)				

[a]Died during first hospitalization before any specific treatment; [b]ChT ± RT, chemotherapy with or without radiotherapy. ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor.

Similarly, the overall impact of the studied variables as predictors of ALK positivity, relative to the negative group, was statistically significant (P=0.024), with pseudo-R^2 values ranging from 0.235 to 0.370. However, the individual predictors did not reach statistical significance. The classification model demonstrated an overall accuracy of 70%. When sex and smoking status were included, the precision of the model increased by 10%. The confusion matrix is displayed in Tables VIIIB and IXB. Cross-validation of the model is provided in Tables SIV-SVI and Figs. S4 and S5.

Furthermore, the present study assessed the predictive accuracy of the model between the ALK and EGFR groups. The combined effect of the variables as predictors was highly significant (P<0.001), with pseudo-R^2 values ranging

Table IV. Frequency of cytomorphological features of lung adenocarcinoma in smears in relation to mutation status.

Cytomorphological features	ALK (%)	EGFR (%)	Negative (%)	Test	P-value
Size of clusters					
≤200 μm	23 (82.14)	35 (83.33)	23 (74.20)	Fisher's=3.202	0.539
>200 μm, ≤400 μm	3 (10.72)	6 (14.29)	4 (12.90)		
>400 μm	2 (7.14)	1 (2.38)	4 (12.90)		
Size of nuclei					
<3 lymphocyte	4 (14.29)	6 (14.29)	6 (19.35)	χ^2=2.821	0.588
3-5 lymphocyte	15 (53.57)	16 (38.09)	15 (48.39)		
>5 lymphocyte	9 (32.14)	20 (47.62)	10 (32.56)		
Nuclear atypia					
Moderate	11 (39.29)	18 (42.86)	16 (53.33)	χ^2=1.289	0.525
Severe	17 (60.71)	24 (57.14)	14 (46.67)		
Nucleoli visibility					
Visible	23 (82.14)	35 (83.33)	25 (80.65)	χ^2=0.088	0.957
Not visible	5 (17.86)	7 (16.67)	6 (19.35)		
Intracytoplasmic vacuoles					
<20%	14 (50)	27 (64.29)	19 (61.29)	χ^2=1.487	0.475
≥20%	14 (50)	15 (35.71)	12 (38.71)		
Signet ring cells					
<5%	24 (85.71)	38 (90.48)	29 (93.55)	Fisher's=1.053	0.641
≥5%	4 (14.29)	4 (9.52)	2 (6.45)		
Necrosis					
Absent	12 (42.86)	33 (78.57)	15 (48.39)	Fisher's= 11.962	0.015[a]
Moderate	12 (42.86)	6 (14.29)	10 (32.26)		
Abundant	4 (14.28)	3 (7.14)	6 (19.35)		

[a]P<0.05. ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor.

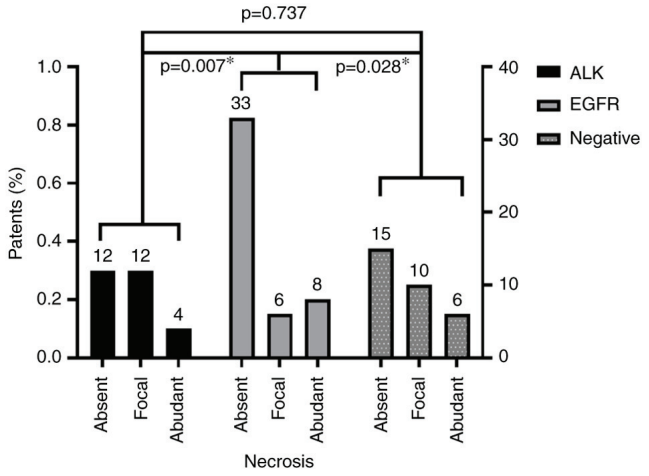

Figure 1. Necrosis in cytological smears in relation to mutation status of the tumor. Smears without necrosis was significantly more often presented in EGFR positive adenocarcinomas than in ALK and negative group. EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma kinase.

from 0.471 to 0.635. Focal necrosis in smears emerged as a significant predictor, reducing the likelihood of EGFR positivity by 87.77%, while the presence of signet ring cells was associated with a 13.75-fold increase in the probability of ALK positivity. The overall accuracy of the model was 84.29%. After incorporating sex and smoking status into the regression model, the accuracy for predicting ALK positivity decreased, whereas the accuracy for EGFR positivity increased, with the overall accuracy remaining unchanged. The classification performance metrics are detailed in Tables VIIIC and IXC. Cross-validation of the model is provided in Tables SVII-SIX and Figs. S6 and S7.

Discussion

World cancer statistics indicate a male predominance in the incidence and prevalence of lung cancer (1). In the present study, there was an equal representation of sex in the ALK group, while females predominated in the EGFR group. Studies on patients with ALK-EML4 gene rearrangement have shown varying results regarding sex distribution (21-24). However, consistent findings across multiple studies confirm that EGFR-positive lung cancer is more common in female patients (23-25). Nonetheless, exceptions, such as a study from Egypt, reported a higher representation of male patients (25). As regards the smoking habits, there is no consensus in patients with ALK-EML4 gene rearrangement regarding the

Table V. Frequency of histomorphological features of lung adenocarcinoma from tissue samples in relation to mutation status.

Histomorphological features	ALK	EGFR	Negative	Test	P-value
Arrangement					
Acinar	17 (60.71%)	26 (61.91%)	21 (67,74%)	Fisher's=19.952	0.003[b]
Lepidic	0	7 (16.67%)	3 (9,68%)		
Micropapillary	0	4 (9.51%)	1 (3,23%)		
Papillary	6 (21.43%)	0	0		
Solid	5 (17.86%)	5 (11.91%)	6 (19,35%)		
Cribriform pattern					
Yes	5 (17.86%)	6 (14.29%)	7 (22,58%)	χ^2=0.838	0.688
No	23 (82.14%)	36 (85.71%)	24 (77,42%)		
Stroma					
Poor	17 (60.71%)	27 (64.29%)	23 (74,19%)	Fisher's=2.635	0.624
Moderate	9 (32.14%)	10 (23.81%)	7 (22,58%)		
Abundant	2 (7.14)	5 (11.90%)	1 (3,23%)		
Size of nuclei					
<3X lymphocyte	7 (25%)	12 (28.57%)	11 (35,48%)	χ^2=1.211	0.876
3-5X lymphocyte	13 (46.43%)	21 (50%)	13 (41,94%)		
>5X lymphocyte	8 (28.57%)	9 (21.43%)	7 (22,58%)		
Nuclear atypia					
Moderate	9 (32.14%)	15 (35.71%)	16 (51,61%)	χ^2=2.786	0.248
Severe	19 (67.86%)	27 (64.29%)	15 (48,39%)		
Nucleoli visibility					
Visible	24 (85.71%)	22 (52.38%)	21 (67,74%)	χ^2=8.399	0.015[a]
Not visible	4 (14.29%)	20 (47.62%)	10 (32,26%)		
Intra-cytoplasmic vacuole					
<20%	3 (10.71%)	17 (40.48%)	8 (25,81%)	χ^2=7.508	0.023[a]
≥20%	25 (89.29%)	25 (59.52%)	23 (74,19%)		
Signet ring cells					
<5%	15 (53.57%)	36 (85.71%)	19 (61,29%)	χ^2=9.511	0.008[b]
≥5%	13 (46.43%)	6 (14.29%)	12 (38,71%)		
Inflammatory infiltrate					
Poor	18 (64.29%)	25 (59.52%)	25 (80,65%)	Fisher's=6.529	0.123
Moderate	9 (32.14%)	16 (38.10%)	4 (12,90%)		
Abundant	1 (3.57%)	1 (2.38%)	2 (6,45%)		
Necrosis					
Absent	26 (92.86%)	42 (100%)	27 (87,10%)	Fisher's=6.465	0.046[a]
Moderate	2 (7.14%)	0	2 (6,45%)		
Abundant	0	0	2 (6,45%)		

[a]P<0.05 and [b]P<0.01. ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor.

association with smoking (17-20). The results of the present study align with some of the literature findings indicating that non-smokers are more commonly associated with EGFR driver mutations (23-26).

The results of the present study confirm that certain cytological and histological features of lung adenocarcinoma are associated with the mutational status of the tumor. This association can aid in selecting appropriate samples for testing and may serve as a valuable supplement to the numerous predictive models currently under development.

Signet ring cells are histologically and cytologically characteristic cells found in tumors of the gastrointestinal system, particularly in the stomach and colon, as well as in ovarian tumors. According to certain studies, tumors containing signet ring cells comprising 10% of the sample represent 7% of adenocarcinomas, while in all lung cancers, they account for ~1.5% (27,28).

A previous study by Japanese authors revealed that specific cytomorphological characteristics, including a pink cytoplasm, vesicular cytoplasm, and smears with predominantly individually

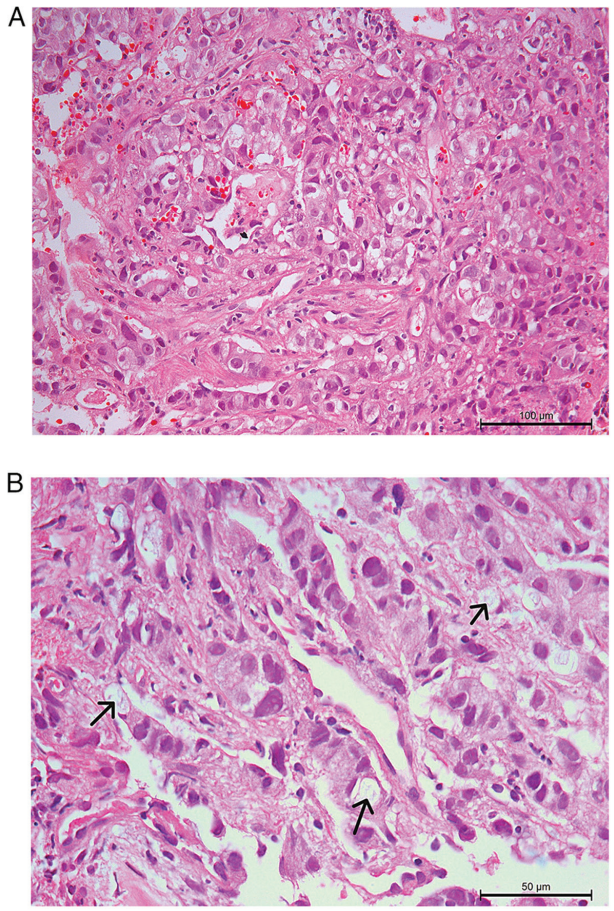

Figure 2. Presence of intra-cytoplasmic vacuoles, visible nucleoli, severe degree of nuclear atypia with nuclear size of 3-5X lymphocyte, with poor stroma, poor inflammatory infiltrate, solid cell arrangement (left or A) and signed ring cells (right or B, arrows) in anaplastic lymphoma kinase positive lung adenocarcinoma.

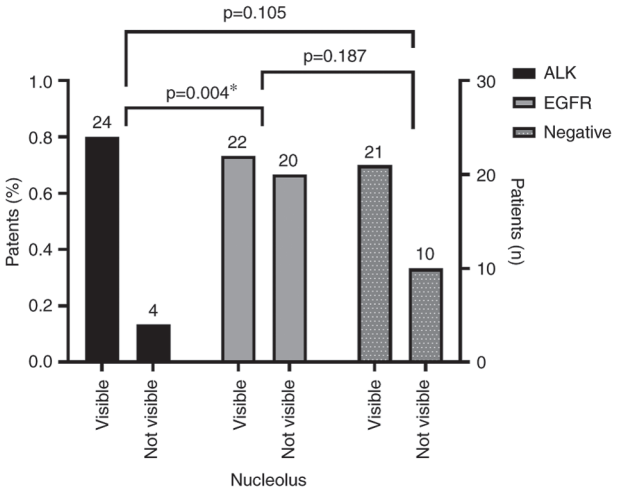

Figure 3. Frequency of visibility of nucleoli in relation to tumor mutation status. EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma kinase.

distributed cells, suggest ALK testing positivity. However, these features cannot replace testing. Nevertheless, in the absence of histological samples for immunohistochemical analysis,

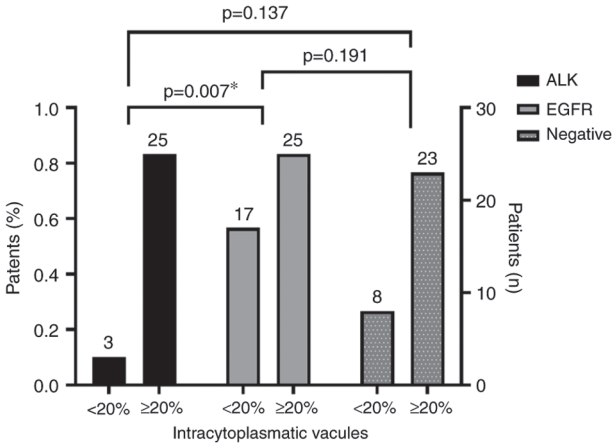

Figure 4. Frequency of intra-cytoplasmic vacuoles (20% of the cells cut off) in relation to tumor mutation status. EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma kinase.

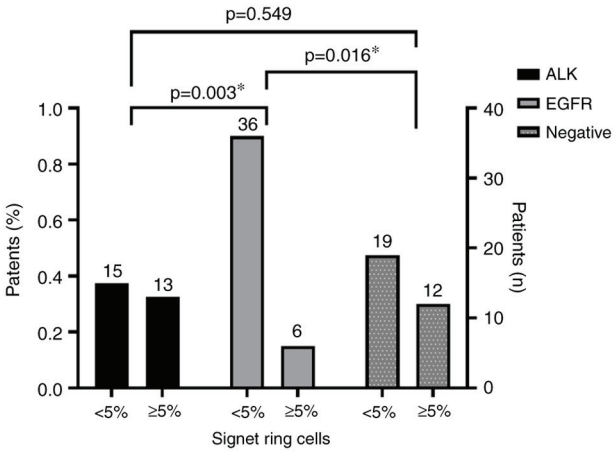

Figure 5. Frequency of signet ring cells (5% of the cells cut off) in relation to tumor mutation status. EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma kinase.

these parameters could aid in predicting ALK positivity (8). Nishino *et al* (10) demonstrated a statistically significant presence of signet ring cells, micropapillary arrangement, and hepatoid cell appearance in ALK-positive adenocarcinomas. They proposed a scoring system with a high sensitivity (88%) and negative predictive value (87%), but low specificity (45%) and positive predictive value (49%) (10). Incorporating useful cyto- and histomorphological parameters into predictive systems could develop a score with higher predictive values than those mentioned. In the present study, intracytoplasmic vacuoles and signet ring cells were statistically significantly more frequent in the ALK group compared with the EGFR and negative groups. Although no studies comparing EGFR and ALK groups were found, studies examining ALK vs. ALK-negative groups identified signet ring cells as a statistically significant parameter of ALK positivity (12). In the present study, the ALK group samples had a higher percentage of signet ring cells, although this difference was not statistically significant. An important consideration in identifying signet ring cells is their mimics, such as vacuolar or fatty degeneration. However, the present study did not address this consideration, representing a limitation (28).

Table VI. Binary logistic regression model for mutation status of lung adenocarcinoma according to micromorphological features (cell arrangement, nucleoli visibility, intra-cytoplasmic vacuoles, signet ring cells, and necrosis in smears).

Model (H1)	df	χ^2	P-value	McFaden	Negelkerke	Tjur	Cox and Snell
EGFR vs. negative	64	17.0223	0.029	0.171	0.280	0.227	0.227
ALK vs. negative	49	19.163	0.024	0.235	0.370	0.250	0.277
EGFR vs. ALK	60	44.361	<0.001	0.471	0.635	0.515	0.469

Table VII. Binary logistic regression model for mutation status of lung adenocarcinoma according to micromorphological features (cell arrangement, nucleoli visibility, intra-cytoplasmic vacuole, signet ring cells, and necrosis in smears), sex and smoking habits.

Model (H1)	df	χ^2	P-value	McFaden	Negelkerke	Tjur	Cox and Snell
EGFR vs. negative	24	52.295	<0.001	0.525	0.687	0.580	0.512
ALK vs. negative	46	29.849	0.003	0.366	0.523	0.416	0.397
EGFR vs. ALK	57	49.719	<0.001	0.528	0.687	0.564	0.509

ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor.

Table VIII. Confusion matrix of classification model for mutation status in lung adenocarcinoma according to micromorphological features.

A, Predicted EGFR and Negative.

Observed	EGFR	Negative	Accuracy (%)
EGFR	31	11	73.81
Negative	9	22	70.97
Accuracy			72.60

B, Predicted ALK and Negative.

Observed	ALK	Negative	Accuracy (%)
ALK	18	10	64.29
Negative	7	24	77.42
Accuracy			71.19

C, Predicted EGFR and ALK.

Observed	EGFR	ALK	Accuracy (%)
EGFR	35	7	83.33
ALK	4	24	85.71
Accuracy			84.29

ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor.

Table IX. Confusion matrix of classification model for mutation status in lung adenocarcinoma according to micromorphological features with sex and smoking habits of the patients.

A, Predicted EGFR and Negative.

Observed	EGFR	Negative	Accuracy (%)
EGFR	38	4	90.47
Negative	6	25	80.65
Accuracy			86.30

B, Predicted ALK and Negative.

Observed	ALK	Negative	Accuracy (%)
ALK	20	8	71.43
Negative	3	28	90.32
Accuracy			81.36

C, Predicted EGFR and ALK

Observed	EGFR	ALK	Accuracy (%)
EGFR	37	5	88.10
ALK	6	22	78.57
Accuracy			84.29

ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor.

In addition to mucinous components, other markers were examined as predictors for ALK positivity. Psammoma bodies and a 'Club cell-like' cytological pattern are statistically significant markers (11). 'Club cell-like' cells exhibit projections

of eosinophilic cytoplasm at the apical compartment lining the surface of papillary cell arrangements (29). In a previous study from Japan (8), the presence of this cytological pattern and papillary tumor growth were predictors of ALK-positive tumors. The results of the present study revealed papillary growth in 21% of ALK-positive tumors, while no such cases were found in the EGFR and negative groups.

Despite the significance of cytomorphological parameters in predicting mutations, data on precise parameters remain insufficient (14-16). Blons *et al* (30) observed an association between the lepidic growth pattern of adenocarcinoma and EGFR status. Sharma *et al* (14) demonstrated a statistically significant presence of acini and single-layer cell bands in EGFR-positive vs. EGFR-negative lung cancer samples. The present study analyzed histological samples dominated by acini formations; however, no statistically significant differences were found compared with the ALK and negative groups. The aforementioned study also suggested an association of EGFR mutations with nuclear atypia and chromatin distribution, noting that mild nuclear atypia was more common in EGFR-positive tumors (14). However, in the present study, there were no statistically significant differences in nuclear atypia between groups, and severe nuclear atypia dominated within each group, contrary to the previous findings (14). The results from an American study revealed that an acinar growth arrangement was significantly more common in EGFR-positive lung adenocarcinomas, with the absence of solid growth serving as a predictor for EGFR negativity, as all solid adenocarcinomas in their sample were EGFR-negative (15). Similarly, in the present study, solid growth arrangement was more common in the ALK and negative groups, although the interpretation of this result should consider the patient-to-group ratio.

The prediction of molecular analysis positivity has been explored through various diagnostic modalities. Song *et al* (31) introduced a deep learning model based on computed tomography and clinicopathological data, successfully predicting ALK positivity with accuracy, sensitivity and specificity of 76.65, 77.44 and 76.32%, respectively. By comparison, the regression model in the present study had an accuracy of 71.19% with only microscopic features. The addition of sex and smoking habits increased the accuracy to 81.36%.

The present study has limitations which should be mentioned. The use of samples obtained by different bronchoscopy sampling techniques in one such limitation. This limitation was unavoidable due to the rarity of the ALK mutation in the population. However, potential differences in the frequency of cyto- and histomorphological characteristics was investigated among samples obtained through different sampling methods. No statistically significant differences were observed, indicating no connection between morphological characteristics and the sampling method. It is worth noting that other studies employing similar methodologies also encountered challenges with different sampling methods (8,10-12).

Another limitation of the present study is the inability to validate the model using external data. Due to the rarity of the mutation, all ALK-mutated adenocarcinomas at the authors' institution with adequate cytological and histological samples were included in the analysis, leaving no additional cases for external verification. However, this limitation is partially mitigated by performing cross-validation using the k-nearest neighbors test, the results of which are provided in Tables SI-SIX.

In conclusion, the present study revealed differences in sex distribution and smoking habits between the groups, alongside statistically significant differences in specific morphological parameters. These findings suggest the potential inclusion of these parameters in future models for predicting the mutational status of NSCLC. Recognizing characteristic patterns in adenocarcinoma samples associated with specific mutational statuses can facilitate the triage of samples for appropriate molecular analyses. However, it is noteworthy that while morphological analyses provide valuable insights, they are not a substitute for molecular testing.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

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

Authors' contributions

NG, AL and SG performed a cytological and histological assessment on the samples, and final review of the manuscript. VS and BZ participated in research design and critically reviewed the intellectual content of the manuscript. AI and TL participated in the analysis and interpretation of the results, while MB and SKL analyzed the research findings, focusing on clinical characteristics and reviewed a discussion based on these insights. All authors read and approved the final version of the manuscript. NG, AL and SG confirm the authenticity of all the raw data.

Ethics approval and consent to participate

The present study was conducted in accordance with the Declaration of Helsinki and was approved (approval nos. 24-VII/10 and 25-VIII/10) by the Professional and Ethical Board of Institute for Pulmonary Diseases of Vojvodina (Sremska Kamenica, Serbia). The Institutional Review Board waived the requirement for informed consent due to the retrospective design of the study with no risk of identity exposure for patients. The present study did not include any minors.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71: 209-249, 2021.
2. Bade BC and Dela Cruz CS: Lung cancer 2020: Epidemiology, etiology, and prevention. Clin Chest Med 41: 1-24, 2020.
3. Šutić M, Vukić A, Baranašić J, Försti A, Džubur F, Samaržija M, Jakopović M, Brčić L and Knežević J: Diagnostic, predictive, and prognostic biomarkers in non-small cell lung cancer (NSCLC) management. J Pers Med 11: 1102, 2021.
4. Nicholson AG, Tsao MS, Beasley MB, Borczuk AC, Brambilla E, Cooper WA, Dacic S, Jain D, Kerr KM, Lantuejoul S, *et al*: The 2021 WHO classification of lung tumors: Impact of advances since 2015. J Thorac Oncol 17: 362-387, 2022.
5. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, Fujiwara S, Watanabe H, Kurashina H, Hatanaka H, *et al*: Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 448: 561-566, 2007.
6. Jiang F, Wang C, Yang P, Sun P and Liu J: Pathological cytomorphologic features and the percentage of ALK FISH-positive cells predict pulmonary adenocarcinoma prognosis: A prospective cohort study. World J Surg Oncol 19: 278, 2021.
7. Hofman P: ALK in non-small cell lung cancer (NSCLC) pathobiology, epidemiology, detection from tumor tissue and algorithm diagnosis in a daily practice. Cancers 9: 107, 2017.
8. Miyata K, Morita S, Dejima H, Seki N, Matsutani N, Mieno M, Kondo F, Soejima Y, Tanaka F and Sawabe M: Cytological markers for predicting ALK-positive pulmonary adenocarcinoma. Diagn Cytopathol 45: 963-970, 2017.
9. Yoshida A, Tsuta K, Nakamura H, Kohno T, Takahashi F, Asamura H, Sekine I, Fukayama M, Shibata T, Furuta K and Tsuda H: Comprehensive histologic analysis of ALK-rearranged lung carcinomas. Am J Surg Pathol 35: 1126-1134, 2011.
10. Nishino M, Klepeis VE, Yeap BY, Bergethon K, Morales-Oyarvide V, Dias-Santagata D, Yagi Y, Mark EJ, Iafrate AJ and Mino-Kenudson M: Histologic and cytomorphologic features of ALK-rearranged lung adenocarcinomas. Mod Pathol 25: 1462-1472, 2012.
11. Pareja F, Crapanzano JP, Mansukhani MM, Bulman WA and Saqi A: Cytomorphological features of ALK-positive lung adenocarcinomas: Psammoma bodies and signet ring cells. Cancer Cytopathol 123: 162-170, 2015.
12. Ha SY, Ahn J, Roh MS, Han J, Lee JJ, Lee B and Yim J: Cytologic features of ALK-positive pulmonary adenocarcinoma. Korean J Pathol 47: 252-257, 2013.
13. Russo A, Franchina T, Riccardi GRR, Picone A, Ferraro G, Zanghi M, Toscano G, Giordano A and Adamo V: A decade of EGFR inhibition in EGFR-mutated non-small cell lung cancer (NSCLC): Old successes and future perspectives. Oncotarget 6: 26814-26825, 2015.
14. Sharma S, Gupta N, Singh N, Chaturvedi R, Behera D and Rajwanshi A: Cytomorphological features as predictors of epidermal growth factor receptor mutation status in lung adenocarcinoma. Cytojournal 15: 11, 2018.
15. Dacic S, Shuai Y, Yousem S, Ohori P and Nikiforova M: Clinicopathological predictors of EGFR/KRAS mutational status in primary lung adenocarcinomas. Mod Pathol 23: 159-168, 2010.
16. Hu H, Pan Y, Li Y, Wang L, Wang R, Zhang Y, Li H, Ye T, Zhang Y, Luo X, *et al*: Oncogenic mutations are associated with histological subtypes but do not have an independent prognostic value in lung adenocarcinoma. Onco Targets Ther 7: 1423-1437, 2014.
17. Kobayashi Y, Yokose T, Kawamura K, Iwasaki S, Murata Y, Onuma S, Hasebe T, Nagai K, Sasaki S and Ochiai A: Cytologic factors associated with prognosis in patients with peripheral adenocarcinoma of the lung measuring 3 cm or less in greatest dimension. Cancer 105: 44-51, 2005.
18. Ravaioli S, Bravaccini S, Tumedei MM, Pironi F, Candoli P and Puccetti M: Easily detectable cytomorphological features to evaluate during ROSE for rapid lung cancer diagnosis: From cytology to histology. Oncotarget 8: 11199-11205, 2017.
19. Radić J, Nikolić I, Kolarov-Bjelobrk I, Vasiljević T, Djurić A, Vidović V and Kožik B: Prognostic and predictive significance of primary tumor localization and HER2 expression in the treatment of patients with KRAS wild-type metastatic colorectal cancer: Single-centre experience from serbia. J Pers Med 14: 879, 2024.
20. Marotti JD, Schwab MC, McNulty NJ, Rigas JR, DeLong PA, Memoli VA, Tsongalis GJ and Padmanabhan V: Cytomorphologic features of advanced lung adenocarcinomas tested for EGFR and KRAS mutations: A retrospective review of 50 cases. Diagn Cytopathol 41: 15-21, 2013.
21. Noronha V, Ramaswamy A, Patil VM, Joshi A, Chougule A, Kane S, Kumar R, Sahu A, Doshi V, Nayak L, *et al*: ALK positive lung cancer: clinical profile, practice and outcomes in a developing country. PLoS One 11: e0160752, 2016.
22. Hou X, Chen H, Liu Y, Gong S, Zhudai M and Shen L: Clinicopathological and computed tomography features of patients with early-stage non-small-cell lung cancer harboring ALK rearrangement. Cancer Imaging 23: 20, 2023.
23. Kang HJ, Lim HJ, Park JS, Cho YJ, Yoon HI, Chung JH, Lee JH and Lee CT: Comparison of clinical characteristics between patients with ALK-positive and EGFR-positive lung adenocarcinoma. Respir Med 108: 388-394, 2014.
24. Liu Q, Huang Q, Yu Z and Wu H: Clinical characteristics of non-small cell lung cancer patients with EGFR mutations and ALK&ROS1 fusions. Clin Respir J 16: 216-225, 2022.
25. Zhou JY, Zheng J, Yu ZF, Xiao WB, Zhao J, Sun K, Wang B, Chen X, Jiang LN, Ding W and Zhou JY: Comparative analysis of clinicoradiologic characteristics of lung adenocarcinomas with ALK rearrangements or EGFR mutations. Eur Radiol 25: 1257-1266, 2015.
26. Mourabiti AY, Sqalli Houssaini M, Benfares A, Bouardi NE, Lamrani MYA, Fatemi HE, Serraj M, Amara B, Qjidaa H, Smahi M, *et al*: Clinical and radiological features associated with EGFR mutation in non-small-cell lung cancer: A study of 149 cases. Egypt J Radiol Nucl Med 54: 171, 2023.
27. Tsuta K, Ishii G, Yoh K, Nitadori J-I, Hasebe T, Nishiwaki Y, Endoh Y, Kodama T, Nagai K and Ochiai A: Primary lung carcinoma with signetring cell carcinoma components: Clinicopathological analysis of 39 cases. Am J Surg Pathol 28: 868-874, 2004.
28. Boland JM, Wampfler JA, Jang JS, Wang X, Erickson-Johnson MR, Oliveira AM, Yang P, Jen J and Yi ES: Pulmonary adenocarcinoma with signet ring cell features: A comprehensive study from 3 distinct patient cohorts. Am J Surg Pathol 38: 1681-1688, 2014
29. Miyata-Morita K, Morita S, Matsutani N, Kondo F, Soejima Y and Sawabe M: Frequent appearance of club cell (Clara cell)-like cells as a histological marker for ALK-positive lung adenocarcinoma. Pathol Int 69: 688-696, 2019.
30. Blons H, Côté JF, Le Corre D, Riquet M, Fabre-Guilevin E, Laurent-Puig P and Danel C: Epidermal growth factor receptor mutation in lung cancer are linked to bronchioloalveolar differentiation. Am J Surg Pathol 30: 1309-1315, 2006.
31. Song Z, Liu T, Shi L, Yu Z, Shen Q, Xu M, Huang Z , Cai Z, Wang W, Xu C and Sun J: The deep learning model combining CT image and clinicopathological information for predicting ALK fusion status and response to ALK-TKI therapy in non-small cell lung cancer patients. Eur J Nucl Med Mol Imaging 48: 361-371, 2021.