RESEARCH ARTICLE

Prediction of the epidermal growth factor receptor gene mutations in lung adenocarcinoma based on CT imaging

Xianghua Wang*

Zhoukou Vocational and Technical College, Zhoukou, Henan, China

Received: November 1, 2023; **accepted:** December 23, 2023.

The epidermal growth factor receptor (EGFR) is a receptor protein located on the cell membrane that regulates cellular behavior. Due to this characteristic, EGFR can serve as a critical target for cancer diagnosis and treatment. Timely detection of EGFR mutations in patients provides valuable opportunities for disease management and treatment. Given the existing research gap in utilizing EGFR for diagnosing lung cancer, this study explored the application of EGFR in this context, offering valuable therapeutic opportunities for alleviating and treating patient conditions, thereby potentially improving their health to a certain extent. Computer Tomography (CT) scans were employed to detect lesions in patients, and regions of interest were manually defined. Subsequently, feature selection was performed based on the Lasso regression method. Image features with predictive capabilities were chosen, and an EGFR personalized prediction model was established using these features, enabling the prediction of EGFR gene mutations. The parts involving model establishment and computations were implemented on the Linux platform using the C++ programming language. 70 patients enrolled were categorized as 2 groups with 50 patients as training set and 20 patients in the experimental validation set. The results showed that the personalized EGFR gene mutation prediction model displayed an area under the curve (AUC) of 0.894 in the training set and 0.889 in the validation set. In addition, its sensitivity and specificity in the training set were 0.67 and 0.86, respectively, while the values of these two indicators in the validation set were 0.94 and 0.58, respectively. The results indicated the effectiveness of this model in predicting EGFR gene mutations in lung adenocarcinoma patients. This study suggested that CT imaging detection before the deterioration of a patient's condition and predicting disease progression provided a new approach for diagnosing EGFR mutations in lung adenocarcinoma patients.

Keywords: computed tomography; epidermal growth factor receptor; lung adenocarcinoma; prediction model; gene mutation; imaging.

***Corresponding author:** Xianghua Wang, Zhoukou Vocational and Technical College, Zhoukou 466000, Henan, China. Email: 18638052805@163.com.

Introduction

Lung cancer, as one of the most prevalent and deadliest cancers, poses a threaten risk to human health and survival. Therefore, reducing the mortality rate of lung cancer has become a global challenge [1]. Among all recorded lung cancer cases, lung adenocarcinoma accounts over 85% of them and has become a major focus of research [2]. Lung adenocarcinoma is characterized by rapid metastasis, high malignancy, and high recurrence rates. Although there are treatment methods available, a considerable number of patients present with severe symptoms and seek medical attention at an advanced stage, missing the optimal treatment window. Therefore, early detection of lung adenocarcinoma in its early stages is crucial

in reducing the mortality rate. Among the various treatment methods for lung adenocarcinoma, the epidermal growth factor receptor (EGFR) has emerged as an effective clinical therapeutic target. EGFR acts as a crucial element in cell growth, differentiation, and proliferation. When epidermal growth factor (EGF) or other related ligands bind to EGFR, it activates the tyrosine kinase activity within EGFR, triggering a series of signaling pathways involved in cellular physiology and pathology. EGFR acts as a crucial element in many types of cancers, making it a significant target for tumor treatment. Inhibitors targeting EGFR have been widely used in clinical therapy to suppress tumor cell growth and metastasis.

For cancer treatment research, there is a considerable amount of literature available. Liu *et al*. summarized the important role of EGFR tyrosine kinase mutation-mediated overexpression and activation in cell proliferation and epithelial tumor development. They believed that existing inhibitors targeting the intracellular tyrosine kinase (TK) domain of EGFR could be used in clinical settings but might lead to the development of acquired resistance [3]. Wang *et al*. evaluated the potential application of a series of novel 5-trifluoromethylpyrimidine derivatives in cancer treatment and tested their anti-tumor activity against various kinases *in vitro*. The results showed that some of these compounds exhibited promising activity. Among them, one specific compound demonstrated high anti-tumor efficacy against the four human cancer cell lines used in the experiments [4]. Hong and Chen reported a case of advanced lung adenocarcinoma and focused on the development of multiple resistance mechanisms during targeted therapy using various biopsy methods. In the treatment of this study, healthcare professionals used EGFR tyrosine kinase inhibitors to treat EGFR-mutated lung adenocarcinoma and documented the impact of acquired resistance on treatment response and survival rate [5]. Peng and Tan, building upon the issue of resistance, summarized the existing knowledge on how insulin-like growth factor (IGF) signaling affected cell resistance to specific drugs. The IGF receptor family acted as a crucial element in cellular signal transduction. They also reviewed current drugs targeting the IGF signaling pathway and their research progress, including clinical trials and preclinical studies [6]. Bhatia summarized the therapeutic achievements targeting the EGFR in neck plus head squamous cell carcinoma and noted that objective progress had been made in recent years with drugs targeting EGFR and programmed cell death. However, treatment response remained challenging for most patients due to primary or acquired resistance [7]. Analyzing the recent literature in the field revealed that most studies focused on the treatment of lung adenocarcinoma and other cancers, as well as the study of patient resistance to EGFR-related drugs. However, for the treatment of major diseases such as cancer, early detection and diagnosis are equally important. Timely diagnosis can effectively alleviate patient suffering, optimize prognosis, improve quality of life, and extend life expectancy. Therefore, it is necessary to conduct research in the diagnosis of lung adenocarcinoma to fill the existing research gaps in this area.

Computed Tomography (CT) adopts X-rays to generate detailed cross-sectional images of the internal human body with computer processing. CT emits multiple X-ray beams and captures X-ray images of multiple slices through scans at different angles. The computer processes and reconstructs these images to create high-resolution 3D images that can clearly display organs, tissues, and abnormalities within the body. This technology is widely used in the diagnosis of various major diseases. Mukherjee *et al*. explored the utility of quantitative CT radiological features of the pancreas in identifying patients who would develop pancreatic cancer within the subsequent 3 to 36 months. The results demonstrated the reliable overall efficacy of this approach [8], which indicated the application value of predictive diagnosis based on CT pathological features. This study, based on the matured CT technology, focused on feature extraction of the lesion area

in patients to establish an EGFR prediction model for predicting EGFR mutations in lung cancer. The results of this study would provide a reference for the EGFR gene mutation in lung adenocarcinoma.

Materials and Methods

Experiment subjects

A total of 89 patients diagnosed with lung adenocarcinoma from January 2018 to August 2020 in Zhoukou People's Hospital, Zhoukou, Henan Province, China were involved in this study. All patients were screened in accordance with the excluding and inclusion criteria of the study and only 70 patients who met the criteria were selected as the research subjects [9, 10] with 42 males and 28 females and the age ranged from 38 to 76 years old. 51 out of 70 patients had a history of smoking. The inclusion criteria were (1) patients who had undergone chest CT examination before lung cancer resection surgery at the designated hospital, and the imaging results were available for review; (2) pathological examination at the designated hospital confirmed the diagnosis of lung adenocarcinoma; (3) EGFR gene mutation results were available in the patient's medical records [11-13], while the excluding criteria were (1) patients with tumors located in the hilum of the lung or with lung collapse; (2) patients who did not undergo lung cancer resection surgery to remove the tumor; (3) patients who had received chemotherapy, radiotherapy, or other treatment methods; (4) the interval between the surgery and chest CT examination was less than one month [14, 15]. All participants involved in this study were informed of relevant information and signed informed consent forms by themselves. In addition, the study was approved by the Ethics Committee of Zhoukou People's Hospital (Approval number: 170627).

CT image segmentation and feature extraction

The CT image data were obtained through uCT960 instrument (Lianying Company, Shanghai, China). The accurate segmentation of the images was the essential for subsequent feature extraction. The segmentation methods for the regions of interest (ROI) in the images could be divided into manual segmentation and automatic segmentation. In this study, the image segmentation was done by using ITK-SNAP software (http://www.itksnap.org) for segment 3D medical images, which allowed manual adjustments of contrast, brightness, and other parameters, enabling more accurate segmentation. In order to achieve more precise image segmentation, this study compared manual segmentation with automatic segmentation to determine which method would be used for this study. An improved watershed algorithm for automatic segmentation of lung images was applied in this study and was implemented on the Linux platform using C++ programming language.

Wavelet features were extracted through wavelet transform and could capture local details and frequency information of images or signals. Image intensity features referred to the grayscale or intensity values of pixels in an image. By performing grayscale statistical analysis on the image, features such as average grayscale, variance, and energy of different regions could be extracted. Image intensity features could reflect information such as brightness, contrast, and texture in the image, playing an important role in tasks. The combined use of wavelet features and image intensity features could provide a more comprehensive description and analysis of three-dimensional image features with wavelet features providing local details and frequency information and image intensity features providing global brightness and texture information. In addition to these two types of features, shape features and texture features were also analyzed. There were four feature categories of abstract concepts in three-dimensional image analysis including shape, image intensity, texture, and wavelet. Through high-dimensional imaging features, lung tumors could be quantified, enabling more scientific, comprehensive, and quantitative reflection of the conditions within lung tumors.

Selection of imaging features and establishment of labels

In order to establish a predictive model for EGFR gene mutations, it is necessary to obtain imaging features with predictive capabilities. Therefore, it is necessary to verify whether the imaging features have excellent predictive capabilities for gene mutations. The 70 patients enrolled in this study were divided to 2 groups with 50 patients as training set and 20 patients as the experimental validation set. As the existence of the small sample size and the high feature dimensionality, there was inevitably feature redundancy. Therefore, the study used the Lasso regression method to select the best-performing features for further research. In Lasso regression, the following expression can be fitted:

$$y = \sum_{k=1}^{N} w_k x_k = W^T x \quad (1)$$

Where *x* was the samples. *N* was the dimensionality of the features. *W* was the parameters to be fitted. Lasso regression added an L_1 regularization term to the least squares algorithm, which could be expressed as follows:

$$J_2(w) = \arg\min_w \left\{ J(w) + \lambda \sum_{k}^{N_1} |w_k| \right\} \quad (2)$$

where

$$J(w) = \arg\min_w \frac{1}{n} \sum_{i=1}^{N} (y_i - w_0 - \sum_{j=1}^{N_i} x_{ij} w_j)^2 \quad (3)$$

In equation (2), the following condition was satisfied:

$$\sum_{j}^{N_1} |w| \le c \quad (4)$$

Equation (4) was a constraint term that restricted Lasso within a norm, limiting the range of variable selection. The performance of the EGFR gene mutation prediction model might change as the number of features changes. The performance of the model fluctuated significantly with the change in the number of features. In addition, the variation of feature coefficients during the feature selection process was also an essential research item. As the range of norm constraints approached the objective function, the Lasso parameters concentrated on the coordinate axis. When the parameter point intersected with the objective function, the optimal solution could be obtained. Through the Lasso algorithm, 7 imaging features with modeling capabilities and 3 clinical variables were finally selected.

Construction of personalized prediction model for EGFR

A logistic function was hired for establishing a predictive model for EGFR mutations. The logistic function was similar in shape to the sigmoid function and could optimize and improve the classification and features in simple linear regression methods. Let the research object included N feature variables as shown in equation (5):

$$x = (x_1, x_2, \mathrm{K}, x_N), x \in R, y \in \{0,1\} \quad (5)$$

where x was the samples. Subsequently, the probability of the event occurring was defined as *P* as shown in equation (6).

$$P = P(y = 1|x) \quad (6)$$

Based on the definition of P, the logistic function could be mathematically expressed as equation (7).

$$P = \frac{1}{1 + e^{-(\alpha + \beta_1 x_1 + \beta_2 x_2 + \mathrm{L} + \beta_N x_N)}} \quad (7)$$

where $P = P(y = 1|x)$ could also be represented as $P(y|x)$. When applied to the study of binary diseases, it could be used to represent the ratio of the probability of patients

not having the disease to the probability of having the disease as follows.

$$\ln(\frac{P(y|x)}{1-P(y|x)}) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \mathrm{L} + \beta_N x_N \quad (8)$$

The output value of the completed personalized prediction model for EGFR mutations was defined as the imaging score, and its model was expressed using the following equation.

$$\begin{cases} Rad\text{-}score = 0.9318379 + A + B + C + D + E \\ A = IIF.range * 0.6567268 - IIF.Skewness * 1.0960980 \\ B = -W_{LLH}F.IIF.mean\_absulute\_deviation * 1.0136071 \\ C = W_{LHH}F.IIF.median * 2.1692973 + W_{LLH}F.IIF.mean * 0.5761592 \\ D = -W_{LLH}F.GLCM.variance * 1.7123142 + GLRLM.HGLRE \\ * 1.0256924 \\ E = -gender * 0.2748692 - smoking\_status * 0.3309877 \\ -histologic\_subtype * 0.39116817 \end{cases} \quad (9)$$

where *IIF.range*, *IIF.Skewness*, $W_{LLH}F.IIF_{mean\_absolute\_debiation}$, *GLRLM.HGLTE*, $W_{LLH}F.IIF.median$, $W_{LLH}F.IIF.mean$, and $W_{LLH}F.GLCM.variance$ were the 7 imaging features selected in this study with predictive capabilities, while gender, smoking status, and histologic subtype were the 3 significant clinical variables selected in the study.

Statistic analyzing

SPSS 19.0 statistical software (IBM, Armonk, New York, USA) was employed in this study. All obtained data were tested for normal distribution. Technical data were represented using frequencies, and results were represented using medians. Chi-square tests and logistic regression analysis were performed. The dependent variable in the study was whether EGFR gene mutations occurred, and a logistic equation was introduced for multivariate analysis. $P < 0.05$ was reckoned as statistically significant.

Results

To verify whether the selected imaging features have predictive capabilities, a feature model was constructed and evaluated *via* receiver operating characteristic (ROC) curve and the area under the curve (AUC). The results demonstrated the ROC curve of the imaging labels for training set (Figure 1a) and the ROC curve of the imaging labels in the validation set (Figure 1b) with the AUC values of 0.7643 and 0.7824, respectively, which indicated that the imaging labels had strong predictive capabilities and robustness in both sets.

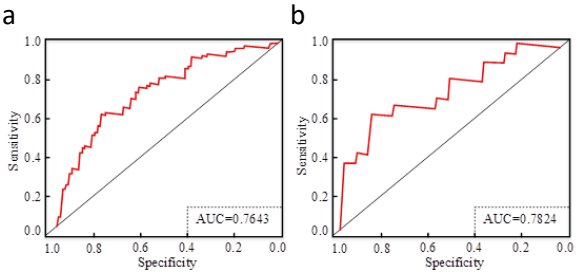

Figure 1. ROC curves of imaging labels. a. training set. b. validation set.

The univariate analysis of the personalized prediction model for EGFR was shown in Figure 2. The results showed that, for both sets, the mentioned indicators of individual features were not very high, indicating that imaging features alone did not have accurate predictive capabilities.

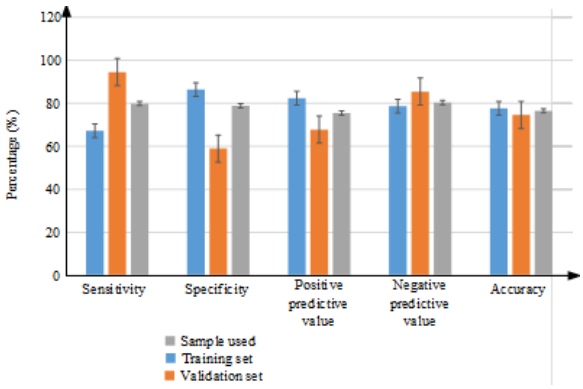

Figure 2. Indicators analysis of the model.

After accurately recording the values of each indicator, the evaluation indicator values of the personalized prediction model under different

Table 1. Evaluation indicator values of the model under different conditions.

	Sensitivity	Specificity	PPV	NPV	Accuracy
Training set	0.67	0.86	0.82	0.78	0.77
Validation set	0.94	0.58	0.67	0.85	0.74
Sample used	0.79	0.78	0.75	0.8	0.76

conditions were obtained (Table 1). The positive predictive value (PPV), which refers to the actual proportion of illness in cases where all test results are positive, and the negative predictive value (NPV), which refers to the proportion of individuals who have not actually developed the disease despite all negative test results, were positive predictive and negative predictive values, respectively. The model sensitivity and specificity for training set were 0.67 and 0.86, while they were 0.94 and 0.58 in the validation set. In terms of accuracy, the model had an accuracy of 0.77 for training and 0.74 for validation. The overall accuracy of the samples used was 0.76, which was relatively close to the model's performance. The multiple features were combined, and their predictive performance was validated using ROC curves.

The performance comparison of the clinical variable model, imaging model, and personalized prediction model in the training set was shown in Figure 3. The ROC curve was the main criterion for judgment. The model based on imaging labels had an AUC value of 0.753, while the clinical variable prediction model had an AUC value of 0.784. The combined imaging labels and clinical variables model had an AUC value of 0.894. The ROC curves of the models in the validation set showed that the model based on imaging labels had an AUC value of 0.784, while the clinical variable prediction model had a much lower AUC value of 0.528 (Figure 4). The combined imaging labels and clinical variables model demonstrated an AUC value of 0.889. The results indicated that the comprehensive prediction model after combining the imaging label model and clinical variable prediction model could evolve the model behavior in the training and validation sets and had a high predictive ability for EGFR gene mutations.

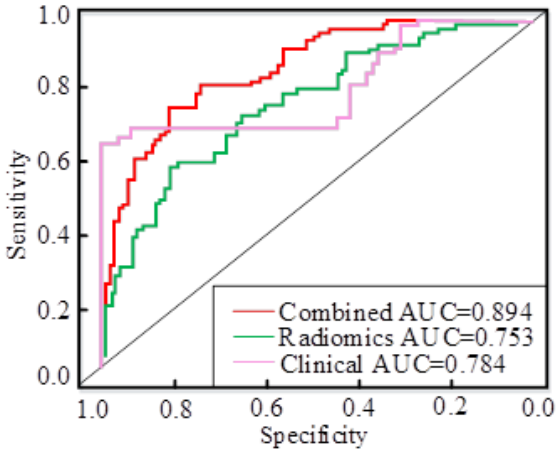

Figure 3. Performance comparison of imaging labels and clinical variable models in the personalized prediction model in the training set.

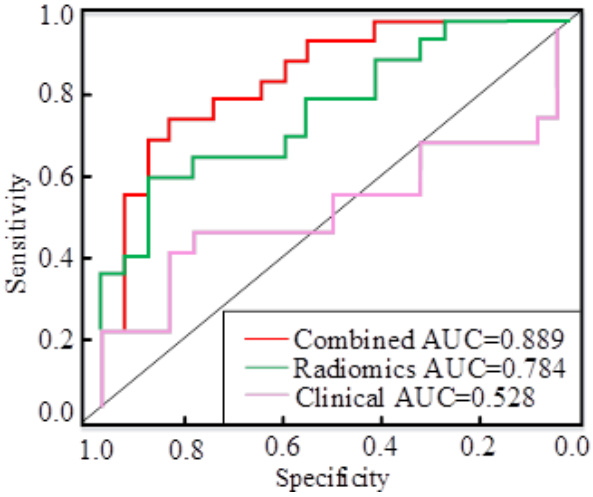

Figure 4. Performance comparison of imaging labels and clinical variable models in the personalized prediction model in the validation set.

Discussion

Lung cancer is one of the most common and deadliest cancers worldwide, posing a significant threat to human health. Therefore, reducing the mortality rate of lung cancer is a global challenge. Lung adenocarcinoma is characterized by rapid metastasis, high malignancy, and high recurrence rate. Despite the available methods to combat this disease, a considerable proportion of patients are diagnosed at an advanced stage when severe symptoms appear, missing the optimal treatment window [16-18]. Among the factors influencing tumor progression, the EGFR is one of the important factors. EGFR is a type of receptor on the cell membrane surface that acts as a crucial element in cell differentiation, growth, and other vital activities due to its tyrosine kinase activity. Similarly, EGFR is also expressed in lesions, and in tumor tissues, EGFR can promote the occurrence, development, and metastasis of tumors, making it a significant factor in clinical medicine. With the increasing incidence of lung cancer worldwide, the mutation status detection for EGFR gene has become a major way to predict and treat lung adenocarcinoma patients. At the same time, numerous clinical studies have confirmed the importance of imaging techniques in predicting and detecting the condition of lung cancer patients. With the maturity of CT imaging technology in recent years, the features in lung cancer imaging have been gradually recognized, and some researchers have attempted to apply CT technology in the detecting of EGFR gene mutations in lung cancer. Some studies suggested that when the tumor diameter was less than 3 cm and the solid component in ground-glass nodules (GGN) was more than 50%, there was a high possibility of EGFR mutation [19]. However, the investigation conducted in this study found that the early screening work for lung cancer in China had become more mature, and an increasing number of early-stage tumor patients had undergone tumor surgery. Therefore, most patients have tumors smaller than 2 cm, making this study statistically insignificant. Furthermore, researchers have found that a history of non-smoking is an independent risk factor for EGFR mutations, indicating that smoking history is a factor associated with the occurrence of lung adenocarcinoma, and its influence is not regulated or interfered by other factors. According to this conclusion, smoking history still has a significant impact on the disease after controlling for other potential confounding factors [20]. However, in this study, the conclusion drawn from data analysis was contrary to the aforementioned viewpoint. Smoking status of patients, as one of the three clinical variables, showed that the EGFR mutation rate in smokers was slightly lower than that in non-smokers, but the difference had no statistically significant, which might be cases of patients concealing or falsely reporting their smoking status in this study, leading to distortion and inaccuracy in the data analysis. Conclusions derived from false or concealed data might also be erroneous. Therefore, further exploring is necessary for investigating the true status of smoking history as a clinical variable, and to confirm it based on higher quality and more reliable data. In future studies, further investigation and validation of the authenticity of data from study subjects should be conducted to improve research accuracy.

The innovation of this study lies in starting from clinical practical problems and classifying the characteristics of lung adenocarcinoma tumors through training. Multiple aspects such as training accuracy, testing accuracy, AUC, specificity, sensitivity, and ROC curve were compared during the training. AUC and ROC are commonly hired indicators for classification models' performance evaluating. The ROC curve can show the trade-off relationship between the sensitivity and specificity of the model at different classification thresholds. An AUC with higher value indicates that the studied object has better performance. AUC and ROC can be hired for performance comparison for different models with a higher AUC value indicating better classification ability of the model. Additionally, AUC can also be used to select the optimal

classification threshold to balance the sensitivity and specificity of the model. By plotting the ROC curve and calculating the AUC value, the classification ability of the model and the optimal classification threshold can be evaluated. Based on the model constructed using imaging labels in this study, the AUC value in the validation set was 0.784, while the AUC value of the clinical variable prediction model was only 0.528. However, the combined model of both imaging labels and clinical variables had an AUC value of 0.889, which indicated that the comprehensive prediction model after combining the imaging label model and clinical variable prediction model could improve the performance of the model in the training set and had a high predictive ability for EGFR gene mutations. Overall, this design improved the predictive accuracy of the research model and enhances the feasibility of this study.

Through the exploration of this study, it is believed that most of the current research focuses on a single type of cancer. This approach has the advantage of making the research more targeted. However, as the number of known cancer types increases and the harm to humans continues to intensify, research on a single type of cancer may not meet future development needs. Therefore, if multiple types of cancer can be analyzed together and common patterns can be identified, it may provide support for the research of a group of cancers. Secondly, based on the rapid computer technology growth, deep learning has achieved significant achievements. Therefore, it is believed that deep learning can be optimized and its integration with the medical field can be strengthened. By linking deep learning with traditional imaging techniques and modeling the mixed features of imaging, the predictive accuracy of the model can be improved. Finally, efforts should be made to strengthen clinical research and improve the stability and reliability of its results. Various machine learning methods can be used for research, such as random forest, clustering analysis, or support vector machines. By combining the analysis of various models, more accurate and persuasive conclusions can be obtained, providing a more rigorous approach to solving clinical problems. However, some limitations do exist. Firstly, it is a retrospective study, so there may be subjective biases in the selection of experimental subjects. Secondly, because the research data is from January 2018 to August 2020, the time for survival follow-up activities is not sufficient, and this will be supplemented in future research. Additionally, since EGFR gene mutations mostly occur in lung adenocarcinoma, this study only focused on adenocarcinoma and did not consider other pathological types. Other pathological types of lung cancer may have different molecular characteristics and clinical manifestations, so neglecting the research on other pathological types may result in an incomplete understanding of the overall disease mechanism and treatment strategies. This may limit a comprehensive and in-depth understanding of lung cancer. Therefore, in future studies, analysis of other pathological types can be conducted. Finally, this study did not investigate the relationship between EGFR mutation status and CT characteristics. It is possible that CT is not only related to EGFR mutations but also associated with their mutation status. Therefore, further research is needed to study the EGFR mutation status, timing, CT characteristics, and other factors.

References

1. Nie G, Wei X, Ye J. 2022. miR-1298 derived from bone marrow mesenchymal stem cells (BMSCs) inhibits non-small cell lung cancer aggressiveness by obstructing chemokine receptor 4 (CXCR4)-induced epithelial-mesenchymal transition (EMT) process. J Biomater Tissue Eng. 12(6):1194-1201.
2. Ayati A, Moghimi S, Toolabi M, Foroumadi A. 2021. Pyrimidine-based EGFR TK inhibitors in targeted cancer therapy. Eur J Med Chem. 221:113523.
3. Liu B, Duenas D, Zheng L, Reckamp K, Shen BH. 2022. Genomic instability as a major mechanism for acquired resistance to EGFR tyrosine kinase inhibitors in cancer. Protein Cell. 13(2):82-89.
4. Wang Q, Rao N, Liu L, Yan L, Le Y. 2022. Development of 5-trifluoromethylpyrimidine derivatives as dual inhibitors of EGFR and SRC for cancer therapy. Heterocycles. 104(3):556-572.
5. Hong E, Chen XE, Mao J, Zhou JJ, Chen L, Xu JY, *et al*. 2022. Sequential occurrence of T790M mutation and small cell lung

cancer transformation in EGFR-positive lung adenocarcinoma: A case report. World J Clin Cases. 10(9):2836-2843.
6. Peng Y, Tan J. 2023. The Relationship between IGF pathway and acquired resistance to tyrosine kinase inhibitors in cancer therapy. Front Biosci (Landmark Ed). 28(8):163.
7. Bhatia A. 2022. Targeting epidermal growth factor receptor in head and neck cancer. The Cancer Journal. 28(5):331-338.
8. Rosenthal MH, Schawkat K. 2023. Beyond the AJR: CT radiomic features of the pancreas predict development of pancreatic cancer. AJR Am J Roentgenol. 220(5):763.
9. Bennett P, Finall A, Medeiros F, Gerrard G, Taniere P. 2022. Re: Inadequacy of PCR genotyping in advanced non-small cell lung cancer: EGFR L747_A755delinsSS Exon 19 deletion is not detected by the real-time PCR IdyllaTM EGFR mutation test but is detected by ctDNA NGS and responds to osimertinib. Eur J Cancer. 174:315-317.
10. Widiandani T, Purwanto BT, Siswandono. 2022. The potency of 4-nitrobenzoyl-3-allylthiourea as an agent of breast cancer with EGFR/HER2: in silico and in vitro study. RASĀYAN J Chem. 3(15):2083-2088.
11. Kodama H, Kenmotsu H, Kawabata T, Notsu A, Yabe M, Nishioka N, *et al*. 2021. Impact of angiogenesis inhibitor eligibility on the prognosis of patients with non-small cell lung cancer harboring EGFR mutation. Cancer Med. 10(21):7503-7513.
12. Park YL, Kim HP, Ock CY, Min DW, Kang JK, Lim YJ, *et al*. 2022. EMT-mediated regulation of CXCL1/5 for resistance to anti-EGFR therapy in colorectal cancer. Oncogene. 41(14):2026-2038.
13. Kim DS, Ji W, Kim DH, Choi YJ, Im K, Lee CW, *et al*. 2022. Generation of genetically engineered mice for lung cancer with mutant EGFR. Biochem Biophys Res Commun. 632:85-91.
14. Hondelink LM, Ernst SM, Atmodimedjo P, Cohen D, Wolf JL, Dingemans AC, *et al*. 2023. Prevalence, clinical and molecular characteristics of early stage EGFR-mutated lung cancer in a real-life West-European cohort: Implications for adjuvant therapy. Eur J Cancer. 181:53-61.
15. Elkrief A, Yu H. 2022. Combination regimens in EGFR-mutated lung cancer: can we get ORIENT-ed? Lancet Oncol. 23(9):1113-1114.
16. Herrera-Juarez M, Serrano-Gomez C, Bote-De-Cabo H, Paz-Ares L. 2023. Targeted therapy for lung cancer: Beyond EGFR and ALK. Cancer. 129(12):1803-1820.
17. Lin CY, Huang KY, Lin YC, Yang SC, Chung WC, Chang YL, *et al*. 2021. Vorinostat combined with brigatinib overcomes acquired resistance in EGFR-C797S-mutated lung cancer. Cancer Lett. 508:76-91.
18. Takahashi S, Noro R, Seike M, Zeng C, Matsumoto M, Yoshikawa A, *et al*. 2021. Long non-coding RNA CRNDE is involved in resistance to EGFR tyrosine kinase inhibitor in EGFR-mutant lung cancer via eIF4A3/MUC1/EGFR signaling. Int J Mol Sci. 22(8):4005.
19. Fatima I, Barman S, Uppada JP, Chauhan S, Dhawan P. 2021. MASTL regulates EGFR signaling to impact pancreatic cancer progression. Oncogene. 40(38):1-14.
20. Kim S, Jeon JS, Choi YJ, Baek GH, Kim SK, Kang KW. 2022. Heterogeneity of glutamine metabolism in acquired-EGFR-TKI-resistant lung cancer. Life Sci. 291:120274.