

Association of lung cancer with ALK and EGFR and its targeted therapy

Siyu Ji^{1,†}, Shangxuan Li^{2,†}, Yifei Song^{3,*}, Shouye Zhu^{4,†}

¹ College of Pharmacy, Shenyang Pharmaceutical University, China, 110016

² Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, M13 9PT

³ College of Chemistry and Molecular Sciences, Wuhan University, China, 430072,

⁴ Department of Biology Science, Xi'an Jiaotong-liverpool University, 215123

* Corresponding Author Email: 2019302030048@whu.edu.cn

† This author contributed equally.

Abstract. As one of the diseases that are extremely harmful to human health, cancer is difficult to treat and has a high mortality rate. And Lung cancer is the most common disease in the world and the type of cancer with the highest mortality rate. In order to better understand and monitor the incidence and mortality of cancer and trends as well as the prevention and treatment of lung cancer. In terms of modern medical technology, the treatment of lung cancer is mainly divided into local treatment and systemic treatment, the main method of local treatment is surgery and radiotherapy, and systematic treatment is a drug-based treatment mode, but because lung cancer cells may still remain in the body after treatment. Over time, these cells may divide and grow. How to achieve thorough treatment of lung cancer is a difficult problem to be overcome.

Keywords: Lung Cancer, ALK, EGFR, Targeted Therapy.

1. Introduction

The bronchial mucosa and lung glands produce the malignant tumor known as lung cancer, which is also the most dangerous kind of cancer to humans due to the tumor cells' escalating morbidity and mortality. Many countries have observed considerable increases in lung cancer mortality and prevalence over the previous 50 years. Additionally, according to the national tumor registry, there were around 787,000 new instances of lung cancer in China in 2015, ranking as the top of the incidence of malignant tumors in China. The most common kind of occupational cancer is lung cancer. An estimated 10% of lung cancer patients have a history of exposure to different carcinogens at work or in the environment. By-products of aluminum goods, arsenic, asbestos, dichloromethether, chromium compounds, coke ovens, mustard gas, nickel-containing impurities, and vinyl chloride have all been demonstrated to increase the risk of lung cancer [1]. Additionally, tobacco contains more than 3,000 chemicals, of which multi-chain aromatic hydrocarbon compounds (like benzopyrene) and nitrosamines have substantial carcinogenic activity, therefore smoking is currently regarded as the most significant risk factor for lung cancer. Numerous ways exist by which nitrosamines and multi-stranded aromatic hydrocarbons can damage the DNA of bronchial epithelial cells and inactivate tumor suppressor genes (such as p53, FHIT gene, etc.).

Former scientists have discovered the genes that cause lung cancer, namely the ALK gene and the EGFR gene, Figuring out these is indispensable for lung cancer therapy and research. The ALK gene is mainly expressed during the embryonic stage. Its role is to promote the proliferation of nerve cells, which plays a role in the development of the brain and peripheral nervous system, and the ALK is switched off after the nervous system is perfected, so under normal circumstances, the ALK in adults is dormant and does not express proteins. The forms of ALK mutation could be fusion of genes with other genes, which triggers point mutations, and so on, eventually causes tumor mutation. ALK fusion mutation is a common driver gene (and a more potent oncogenic driver gene) in NSCLC (non-small cell lung cancer).

The EGFR gene instructs cells to build a specific class of protein known as a receptor, which is found on the cell surface. The recognition and binding of other substances by receptors often has a specific impact on the cell. One type of receptor known as a receptor tyrosine kinase, which helps control cell growth, division, survival, and death, is the epidermal growth factor receptor. Its main forms of mutation include the pair deletion of exon 19, frameshift mutation of exon 20, and missense mutation of L858R exon 21.

Cancer cells are mutated cells of the human body, with the characteristics of infinite proliferation, and they grow uncontrollably. During growth, cancer cells rob the body of nutrients to increase the rate at which they divide and multiply. In addition, cancer cells also disguise themselves as normal cells, tricking the immune system and avoiding white blood cell attacks. As cancer cells grow and proliferate, it invades blood vessels, causing damage to the body's systems and metastasis. Even if the malignancy is surgically removed, it is difficult to ensure that the cancer cells will not recur [2]. In addition, precision therapy is the general direction of the treatment of cancer, targeted drugs, immunotherapy belong to the category of precision treatment, the bottleneck is the inevitable drug resistance, the research direction of targeted drugs has changed, the emergence of allosteric inhibitors should provide new ideas for solving drug resistance, in general, the future treatment of cancer will be more accurate and effective. The review focused on how the EGFR and ALK genes affect the development of lung cancer and what roles these two genes play in cancer, and later includes targeted therapies for both genes and drug use.

2. The gene can induce lung cancer—ALK

According to reliable data, lung cancer is the disease that causes the greatest morbidity in men, and non-small cell lung cancer accounts for 80%-85% of cases (NSCLC). Radiotherapy and surgery are the main traditional lung cancer treatment options. However, more and more instances demonstrate that patients have a low chance of undergoing surgery [3]. The anaplastic lymphoma kinase (*ALK*) gene alterations discovered at this time have revolutionized the study of lung cancer. Meanwhile, by examining the relationship between the *ALK* gene and lung cancer, some scientists have also given patients with the disease a great deal of hope. We will go through how *ALK* results in lung cancer below [4].

2.1. The basic description of the ALK gene and its mechanism of action

The *ALK* gene is frequently expressed in NSCLC, so it is considered a special marker to detect the carcinogenic principle of lung cancer. The following is about discussing the basic summary of the *ALK* gene.

First of all, through research, *ALK* gene is situated on chromosome 2's short arm. The tyrosine monomer of the *ALK* gene transmembrane, for which the *ALK* protein is encoded by the *ALK* gene. It does not express itself after birth but is physiologically expressed in the neural system throughout the embryonic stage. *ALK* has an extracellular domain, a transmembrane segment and a cytoplasmic receptor kinase section.^[1] The fact is that if the *ALK* gene is expressed at an improper period, it will result in excessive cell proliferation and lead to death, which will eventually result in lung cancer.

The expression of the *ALK* gene is to synthesize the *ALK* protein. Two protein monosomes will form a dimer of *ALK*. Then, after the dimer binds to some ligands, it is phosphorylated.

2.2. The EML4-ALK fusion gene can cause lung cancer.

The N-terminal, basic region, the hydrophobic *EML4* protein (HELP) domain, and the variable tryptophan-aspartic acid (WD) repeats make up Echinoderm microtubule-associated protein-like 4 (*EML4*). The trimerization domain (TD) is a coiled-coil structure near the N-terminal of the protein.^[3] The N-terminal coiled-coil domain's strong ties to the *ALK* dimer are the most significant feature. According to studies, the basic region is likely the crucial region where the *EML4-ALK* fusion protein is drastically reduced (by roughly 84%).

The gene was inserted into the short arm of chromosome 2 and the *EML4-ALK* fusion formed when the 3' end of the *ALK* gene. This caused a conformational change, activated intracellular tyrosine kinase phosphorylation, activated numerous cells, and caused lung cancer via signal pathways.

According to molecular screening, the majority of lung cancer cases- more than 80%- are adenocarcinomas. Even though smoking is the primary cause of lung cancer, 25% of patients have little to no smoking history. We can better understand the occurrence of lung cancer by using genetic analyses.

EML4-ALK fusion genes are frequently observed in NSCLC. The fusion gene for the *EML4* gene activates downstream signal pathways such as PL3K/mToR, JAK/STAT, MAPK, and others, producing unchecked cell proliferation and death, which can have serious carcinogenic implications. Before knowing how the *EML4-ALK* fusion gene results in NSCLC, it is important to understand some genes that are connected to cancer. Cancer-related genes can exist in cells under certain normal conditions, and their normal expression is essential for many aspects of life, including cellular growth, tissue regeneration, and individual development. When these genes change or turn into oncogenes, they have cancer-causing effects. Proto-oncogenes are these genes that have the capacity to lead to cell malignancies. Then because some protooncogenes can undergo mutations that will likely result in cancer as a result of the action of physical, chemical, and biological carcinogens (HPV, HBV, and others), it is clear from this that *ALK* is a proto-oncogene that goes through *EML4-ALK* fusion mutation.

Rearrangement (*ALK-R*), amplification (*ALK-A*), and point mutation are the three different forms of *ALK* gene alterations. If the *ALK* gene has merged with other genes, a fusion oncogene that is highly expressed in cancer will result. After extensive research and effort, it was found that NSCLC contains the *ALK-R*. *EML4* and *ALK*'s gene sequences are oriented in opposition on the short arm of chromosome 2, therefore one of them needs to be joined. Consequently, *EML4-ALK* is made up of *EML4*'s amino-terminal region linked to *ALK*'s kinase domain. When some Japanese patients found that their NSCLC had an *ALK* rearrangement with *EML4*, which resulted in the fusion oncogene *EML4-ALK*, it is reported that the *ALK* mutations were initially identified in NSCLC in 2007. The oncogenic *ALK* tyrosine kinase was produced as a result of *ALK* rearrangement and may activate a variety of downstream signaling pathways, increasing cell proliferation and survival [5].

The *EML4-ALK* fusion protein, which is created when the amino-terminal of *EML4* and the kinase domain of *ALK* are combined, is produced by the fusion gene of *ALK* and *EML4* and has tumorigenic activity. The constitutive dimerization of the amino terminal of the *EML4* on the kinase domain of *ALK* leads to abnormal activation of downstream signal pathways such as AKT, STAT3, etc. The mutation of the fusion gene is an important form of oncogene mutation for the *ALK* gene. At present, the mutated *ALK* protein is involved in a wide range of downstream signal pathways, such as the RAS-MAPK pathway, the PI3K-AKT pathway, and so many others. These signal pathways are directed toward cell proliferation, apoptosis resistance, and eventually lung cancer induction. [4]

The fluorescence can be used in situ hybridization (FISH) to detect whether there is an *EML4-ALK* fusion in the body. This kind of fusion gene mostly occurs in young non-smoking patients, which is an important risk factor for lung cancer and an important risk for targeted therapy of NSCLC. The advent of targeted drugs has prolonged the survival time of advanced NSCLC patients with gene mutations.

3. EGFR

Similarly, another gene called *EGFR* (epidermal growth factor receptor) is belonging to the human epidermal growth factor receptor (*Her*) including *HER1* (*erbB1*, *EGFR*), *HER2* (*erbB2*, *NEU*), *HER3* (*erbB3*), and *HER4* (*erbB4*). The receptors of the *Her* family are important players in the development of cancer mutations and are also the key breakthroughs for people who want to treat cancer. In the

meantime, the *EGFR* gene is closely related to lung cancer in this family, and its mutation and activation ability can lead to the generation of lung cancer tumors.

3.1. The definition and expression of the EGFR gene.

EGFR is a transmembrane glycoprotein that penetrates the cell membrane and has a molecular weight of 170KDa. From Her is a receptor family, it is easy to know that *EGFR* is also a receptor, a tyrosine kinase type receptor. *EGFR* is divided into three regions, namely, the extracellular ligand-binding region full of glycosylation on the extracellular domain, the hydrophobic transmembrane region, and the intracellular kinase region, which has protein kinase activity in the near-membrane region, and a region possessing a carboxyl tail in the intramembrane segment. As a receptor, it has a great binding ability. Unlike Her3 in the same family, which only has two ligands that bind to it, *EGFR* even has a total of 7 suitable binding ligands, which are *EFG* (epidermal growth factor), *TGF- α* (transforming growth factor alpha), neuregulin 2b, amphiregulin, epigen, betacellulin, heparin-binding EGF, epiregulin respectively and each ligand that binds to it will have different effects in different fields. *EGFR* first exists in a monomeric form without tyrosine kinase activity. When in the monomeric form, it seeks to bind to the corresponding ligand. After binding to a suitable ligand, the ligand will lead to the extracellular ligand domain. After a strong conformational change because of the ligand, at which time the part in the extracellular region exposes the dimerization arm for binding, two ligand-*EGFR* complex monomers that have been bound to the ligand will bind to each other to form kinase-active Back-to-back dimers, and the ligands will be located on opposite sides. At this time, the phosphokinase activity of the dimer will play a role, *EGFR* itself will be phosphorylated at first, and tyrosine kinase activity will be enhanced after autophosphorylation. The proteins in the downstream signaling pathways are also phosphorylated, stimulated by the reaction between the kinase-active carboxyl terminus of the tail and the amino terminus of the receptor, thereby activating the downstream *RAS*, *MAPK*, *Src*, *STAT3/5*, *PLC γ* , *PKC*, *PI3-Kinase* and many other signaling pathways enable *EGFR* to have its most important ability to promote cell proliferation, differentiation, and migration. Finally, in the absence of ligands, *EGFR* monomers on the cell membrane surface will be continuously endocytosed into cells, part of which will be degraded, and the other part will be brought to the surface of the cell membrane again [5].

3.2. The reasons EGFR gene cause lung cancer

The variation of *EGFR* accounts for 10%-40% of the causes of lung cancer. The reasons include the overexpression of *EGFR*, which leads to two *EGFR* monomers without any ligand that will automatically combine into a dimer and the dimers will have continuous self-activation, and the *EGFR* gene mutation, which results in the ongoing production of tyrosine kinase [6].

Even 40% to 80% of the causes of non-small-cell lung cancer are due to overexpression. During the meantime, structural variations of the *EGFR* gene are responsible for 30%–40% of non-small cell lung malignancies in the yellow race. and 10%-20% in the white race. At present, the research on *EGFR* structural mutations is increasingly in-depth. The tyrosine kinase active site of *EGFR* is located on exons 18-24 and has the function of controlling activating mutations and drug resistance mutations, and exon 19 of the mutation of *EGFR* will change the direction of the entire function, making it develop in an opposite direction, and it has become one of the most classic cases of *EGFR* structural variation. The variation of exon 19 is specifically called exon 19 inframe deletions, and deletion of gene segments is the method of gene mutation [7]. In the meanwhile, the Leu858Arg point mutation in exon 21 is also one of the main mutation methods. This gene mutation method makes the 858th amino acid residue wrong after transcription and translation, and the leucine is translated into arginine by mistake. Both mutations result in *EGFR* directly having sustained tyrosine kinase activity without ligand binding. The former accounts for 45% and 40% of *EGFR* structural variants, respectively. These mutations have been the subject of numerous research, and the prognosis for individuals with this subtype of non-small-cell lung cancer is high-quality. Nevertheless, the other one only accounts for 0.3%-3.4% of the *EGFR* mutations that have the non-small cell lung cancer prognosis resulting

in a minor probability mutation relatively difficult and poor [8]. This is the insertion mutation of exon 20. This type of mutation is the addition and insertion of a gene, which disrupts a certain coding sequence. The most famous example is the Thr790Met mutation, a transcriptional translation error caused by a coding error that eventually leads to threonine altering to methionine at position 790, thus the protein's properties are changed. Additionally, it is not as classic as deletion exon 19 and points exon 21. It is generally considered to be a heterogeneous variant that causes conceptual changes around *EGFR*, most of which are in a C-terminal of the C-helix loop. Most seriously, while the first two structural variants can be improved by treatment with *EGFR* tyrosinase inhibitors, exon 20 insertion mutations will cause patients resistant to tyrosinase inhibitors, and conventional treatment methods such as chemotherapy and immunotherapy have no significant effect. At present, scientists are studying new targeted drugs and new treatment directions to solve the non-small-cell lung cancer problem of the *EGFR* gene's exon 20 mutations. [9]

4. Targeted therapy for NSCLC

4.1. ALK inhibitors

Take crizotinib as an example of *ALK* inhibitors, a small molecule ATP-mimicking compound which is ATP-competitive. Crizotinib can inhibit the growth and survival of tumor cells by inhibiting *C-Met* kinase, disrupting the signaling pathway of *C-Met*, and then inhibiting the *ALK* fusion gene. Crizotinib has good efficacy and minimal adverse reactions.

After receiving 1-2 years of treatment, some individuals tend to acquiring medication resistance. Secondary drug resistance mutations and driver gene conversion are the primary mechanisms of drug resistance. The main secondary resistance mutations are L1196M, G1269A, C1156Y and other mutations. Some patients also have secondary mutations in *ALK* kinase region for a second time. Crizotinib inhibits tumor growth mainly by blocking *ALK* and its downstream signaling pathways. Tumor cells can activate other pathways through *EGFR* mutation or phosphorylation, *KRAS* mutation and *C-Kit* amplification, which can replace the dependence of tumor cells on *ALK* and its downstream signaling [10].

4.2. EGFR inhibitors

Currently, those used clinically are the first three generations of TKI. The first generation of TKI is small molecules of quinazoline compounds, including gefitinib, erlotinib, icotinib etc. Although it has a significant curative impact in the initial stages of treating advanced NSCLC, for the vast majority of patients receiving targeted therapy for about 10 months, the tumor showed a greater resistance due to the occurrence of T790M mutation. The second generation of TKI, mainly represented by afatinib and dacomitinib, has molecular structures similar to gefitinib and erlotinib. In order to overcome the resistance of T790M mutation, acrylamide functional groups capable of Michael addition to C797 are added. The third generation of TKI, mainly represented by osimertinib and furmonertinib, is an *EGFR* kinase inhibitor especially developed for patients with T790M mutation. It can irreversibly bind to *EGFR*, T790M, L858R and exon 19 deletion mutants. In order to solve the toxicity of the second generation of drugs, the U-shaped conformation is adopted in T790M [11].

4.2.1. The first generation of TKI

Taking erlotinib as an example, erlotinib competitively binds to the ATP site of the tyrosinase domain to block the downstream signaling of *EGFR*, thus achieving the inhibitory effect. After about 10-14 months of treatment, most patients would develop resistance mutations to erlotinib. More than half of patients develop resistance due to the T790M mutation, which causes resistance by obstructing binding to TKI with stereoscopic interference or by increasing affinity for ATP. The main adverse effects of erlotinib are rash, gastrointestinal dysfunction, and abnormal liver function. The incidence of adverse reactions is low and most of them are tolerable [12].

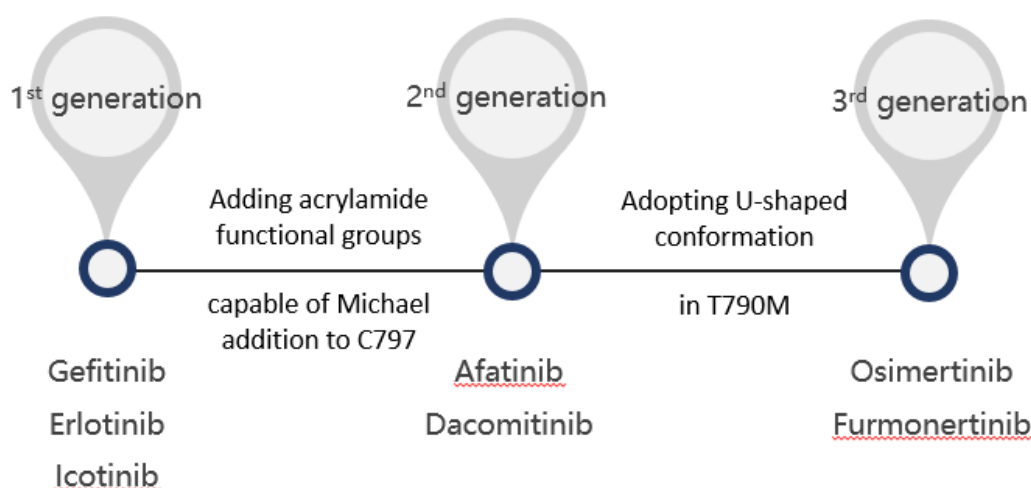


Figure 1. Iteration of three generations of TKI

4.2.2. The second generation of TKI

Afatinib, for example, covalently binds to C805 on *HER2*, C803 on *ErbB4*, and C797 on *EGFR*, which irreversibly inhibits the receptor's tyrosine kinase function, causing autophosphorylation of intracellular tyrosine residues, thereby blocking *EGFR-HER2* signaling and acting as an inhibitory agent. Moreover, afatinib still has a good effect on patients with erlotinib and gefitinib resistance.

After about 9-13 months of treatment, most patients develop resistance mutations to afatinib. Afatinib combined with cetuximab can inhibit the T790M mutation. However, due to severe adverse reactions and insufficient doses, the drug resistance problem cannot be solved, and some patients were forced to stop taking the drug because of unacceptable rash and diarrhea [13].

4.2.3. The third generation of TKI

Osimertinib, for instance, is a monoaniline pyrimidine compound that covalently binds to the ATP binding site at position C797 in the *EGFR* catalytic domain, through unsaturated acrylamide groups, which inhibits *EGFR* phosphorylation and downstream signaling substrates *AKT* and *ERK*, thus playing an anticancer role.

The most common mechanisms of osimertinib resistance are C797S/G and L792H mutations. It can also result from the disappearance of T790M mutation and *EGFR* amplification. In addition, many receptor tyrosine kinases share a portion of the same downstream signaling pathway as *EGFR*, so the activation of other receptor tyrosine kinases can transactivate the *EGFR* pathway, leading to osimertinib resistance in tumors.

The mechanism of action of osimertinib is selective inhibition of T790M mutation, which does not target wild-type *EGFR*. Therefore, osimertinib is well tolerated, and the incidence of rash, diarrhea and other adverse effects is lower than that of the first or second-generation TKIs [14].

5. Conclusion

The treatment of lung cancer has made great strides in the development of recent years. Lung cancer has not been completely conquered. But the treatment of lung cancer has developed as diversely as possible. Whether it is surgery, chemotherapy, targeted therapy, or the most cutting-edge immunotherapy, these treatments have their irreplaceable significance.

References

- [1] Thorac Cancer. 2018 Apr; 9(4): 423–430. Published online 2018 Feb 28. doi: 10.1111/1759-7714.12613
- [2] J Clin Oncol. 2018 Apr 20; 36(12): 1199–1206. Published online 2018 Jan 26. doi: 10.1200/JCO.2017.76.2294

- [3] Eichenmuller B, Everley P, Palange J, Lepley D, Suprenant KA. The human EMAP-like protein-70 (ELP70) is a microtubule destabilizer that localizes to the mitotic apparatus. *J Biol Chem*. 2002 Jan 11;277(2):1301-9. doi: 10.1074/jbc.M106628200. Epub 2001 Nov 2. PMID: 11694528.
- [4] Clin Cancer Res. Author manuscript; available in PMC 2011 Jan 24. Published in final edited form as: *Clin Cancer Res*. 2008 Jul 1; 14(13): 4275–4283. doi: 10.1158/1078-0432.CCR-08-0168
- [5] Roskoski R Jr. The ErbB/HER family of protein-tyrosine kinases and cancer. *Pharmacol Res*. 2014 Jan;79:34-74. DOI: 10.1016/j.phrs.2013.11.002. Epub 2013 Nov 20. PMID: 24269963.
- [6] Harrison PT, Vyse S, Huang PH. Rare epidermal growth factor receptor (EGFR) mutations in non-small cell lung cancer. *Semin Cancer Biol*. 2020 Apr;61:167-179. DOI: 10.1016/j.semcancer.2019.09.015. Epub 2019 Sep 25. PMID: 31562956; PMCID: PMC7083237.
- [7] Matsui T, Tanizawa Y, Enatsu S. [Exon 19 Deletion and Exon 21 L858R Point Mutation in EGFR Mutation–Positive Non–Small Cell Lung Cancer]. *Gan To Kagaku Ryoho*. 2021 May;48(5):673-676. Japanese. PMID: 34006711.
- [8] Li AM, Boichard A, Felip E, Kurzrock R. New therapeutic approaches to overcoming resistant EGFR exon 20 alterations. *Crit Rev Oncol Hematol*. 2020 Jul;151:102990. DOI: 10.1016/j.critrevonc.2020.102990. Epub 2020 May 20. PMID: 32485428; PMCID: PMC7416421.
- [9] Yang X, Zhao J. [Advances in Treatment of Non-small Cell Lung Cancer Harboring EGFR Exon 20 Insertion Mutations]. *Zhongguo Fei Ai Za Zhi*. 2022 May 20;25(5):337-350. Chinese. DOI: 10.3779/j.issn.1009-3419.2022.103.01. PMID: 35599009; PMCID: PMC9127758.
- [10] Tao JIANG, Caicun ZHOU. Treatment of Patients with ALK Gene Rearranged Non-small Cell Lung Cancer after Resistance to Crizotinib [J]. *Chinese Journal of Lung Cancer*,2015(2):69-74. DOI:10.3779/j.issn.1009-3419.2015.02.03.
- [11] Aiping HOU. Research progress of targeted medicine for non-small cell lung cancer withr epidermal growth factor receptor mutation [J]. *Contemporary Medicine*,2022,28(15):181-186. DOI:10.3969/j.issn.1009-4393.2022.15.063.
- [12] Yan ZHAO, Ru ZHANG. The Mechanism and Therapy Perspective of Erlotinib in Patients with Non-small Cell Lung Cancer [J]. *Chinese Journal of Pharmacovigilance*,2010,7(10):617-619. DOI:10.3969/j.issn.1672-8629.2010.10.014.
- [13] Haiqiong LI. Research Progress of Afatinib for Treating Non-Small Cell Lung Cancer [J]. *China Pharmaceuticals*,2017,26(19):5-8. DOI:10.3969/j.issn.1006-4931.2017.19.002.
- [14] Jun CHEN, Jiali HUANG, Zhixiang CHENG. Progress of Ositinib for Treating Advanced Non-Small Cell Lung Cancer [J]. *Journal of Southeast University(Medical Science Edition)*,2020,39(1):85-89. DOI:10.3969/j.issn.1671-6264.2020.01.019.