**EGFR gene mutation and association with carcinoembryonic antigen status in non-small cell lung cancer — a cross-sectional study**

Lam Van Ngo¹, Toan Khanh Nguyen¹, Hung Duc Tran², Uyen Tu Thi Nguyen¹, Hien Thu Thi Tran¹, Thanh Van Le¹

¹Department of Medical Oncology, Nghe An Oncology Hospital, Nghe An, Vietnam
²Department of Anatomic Pathology, Nghe An Oncology Hospital, Nghe An, Vietnam

**Abstract**

**Introduction.** Epidermal growth factor receptor (EGFR) gene mutations are an important aspect in the diagnosis and treatment of non-small cell lung cancer (NSCLC). Some studies suggest that serum CEA levels may serve as a predictor of the outcome of EGFR mutations. Therefore, we conducted a study to determine the prevalence of EGFR gene mutations and evaluate the prognostic value of serum CEA levels in predicting the frequency of EGFR gene mutations in NSCLC.

**Material and methods.** From January 2018 to December 2022, a cross-sectional study was conducted on 384 NSCLC patients at the Nghe An Oncology Hospital, in Vietnam. EGFR mutations were analyzed using the real-time PCR method to determine the sensitivity and specificity of CEA values in predicting EGFR mutation traits by ROC curve analysis, and the association was assessed using univariate and multivariate logistic regression analyses.

**Results.** The EGFR gene mutation rate in NSCLC is 41.9%. Among patients with genetic mutations, 50.9% had the del exon 19 mutation, 34.8% had the L858R exon 21 mutation, 3.7% had rare exon 18 mutations, 5.6% had dual mutations, and 5.0% had exon 20 insertion mutations. The CEA cutoff value was determined to be 8.95. The sensitivity and specificity of CEA were 76.4% and 47.5%, respectively, and the area under the curve (AUC) was 0.627 (95% CI 0.571–0.683; p < 0.01). The EGFR gene mutation was found to be closely associated with the CEA subgroup ≥ 8.95 ng/mL (OR 2.54; 95% CI 1.57–4.13).

**Conclusions.** This study shows a high incidence of EGFR mutations in NSCLC and suggests that CEA can aid in predicting the likelihood of these mutations.

**Keywords:** EGFR mutation, CEA, non-small cell lung cancer, Vietnam

---

**Introduction**

In the past, chemotherapy was the main treatment modality for advanced-stage non-small cell lung cancer (NSCLC). However, significant advancements have been made, and epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) have become the first-line treatment for advanced-stage NSCLC with positive EGFR mutations [1, 2]. Indeed, the superior efficacy of targeted therapies in NSCLC patients with EGFR mutations compared to chemotherapy has been demonstrated in clinical trials [3–6]. Furthermore, in a phase III clinical trial, osimertinib was found to have superior advantages over erlotinib and gefitinib [7]. A study in Vietnam evaluated the excellent efficacy, favorable response rates, and manageable adverse effects of Afatinib in NSCLC patients with EGFR mutations [8].

The frequency of EGFR gene mutations in NSCLC varies among different geographical regions. A 2015 meta-analysis reported the highest mutation rate in the Asia-Pacific region at 47%, followed by South
America at 36%, North America at 22%, Europe at 15%, and the lowest rate in Oceania at 12%. The frequency of EGFR mutations is higher in females and non-smokers [9]. When analyzing the mutation profiles of NSCLC patients in Vietnam, EGFR mutations had the highest frequency at 32.3% [10]. Another study conducted at the Bach Mai Hospital, Vietnam, reported a mutation rate of 40.7% [11].

An epidemiological study in Asian countries reported an overall EGFR mutation rate of 51.4%. The mutation frequency was higher in females compared to males, lower in Indian patients compared to other regions, highest in the never-smoking group at 60.7%, and decreased with increasing tobacco consumption. A multivariate logistic regression analysis found a correlation between ethnic group, smoking status, and mutation frequency. Specifically, within the Vietnamese subgroup, the EGFR gene mutation rate in this study was 64.2% [12]. Furthermore, another meta-analysis reported an EGFR gene mutation rate of 32.3%, with a higher rate in females, non-smokers, and adenocarcinoma patients [13].

Several studies in China have shown that elevated carcinoembryonic antigen (CEA) levels are associated with a higher frequency of EGFR mutations [14–16]. This finding may serve as a novel prognostic indicator for the response to Tyrosine kinase inhibitors in cases with high CEA levels.

A different research study discovered that lower SUVmax values and higher CEA levels were linked to mutations in the EGFR gene, which could be a factor in predicting the response to EGFR TKI medications. The rate of EGFR gene mutations tends to increase as the CEA value increases. Analysis of the receiver operating characteristic (ROC) curve established a CEA cutoff of 9.6 ng/mL with sensitivity of 67.0%, specificity of 68.1%, and an area under the curve (AUC) of 0.632 [95% confidence interval (CI) 0.546–0.719] [17].

Epidermal growth factor receptor TKI drugs have become increasingly prevalent in Vietnam in recent years. Therefore, doctors must understand the importance of EGFR mutations in clinical practice. Apart from well-established factors such as sex, histopathology, and smoking status that can predict the presence of EGFR gene mutations, knowing the association between serum CEA levels and EGFR gene mutations can be helpful. This understanding can aid in the diagnosis and treatment of patients. Thus, this study aimed to determine the rate of EGFR gene mutations and the prognostic value of serum CEA levels in predicting this mutation in patients with NSCLC.

### Material and methods

#### Study location and patients

This study was conducted at the Nghe An Oncology Hospital, a specialized cancer center in the North Central region of Vietnam. With a capacity of 1300 beds, it is the largest hospital specializing in cancer treatment in the North Central region. The study protocol was approved by the hospital’s ethics committee. Before performing the EGFR gene test, the doctor explained CEA and EGFR in lung cancer and conducted the test only after obtaining the patient’s consent, especially for squamous carcinoma. The diagnosis of EGFR gene mutations was performed using the real-time polymerase chain reaction (PCR) method.

The study included patients aged ≥ 18 years who were diagnosed with NSCLC based on histopathology and underwent EGFR testing using the real-time PCR method. Patients with small cell histology or samples that did not meet the criteria for mutation testing were excluded.

#### Study design and sample size

A cross-sectional study was conducted on NSCLC patients from January 2018 to December 2022. The sample size was calculated using the World Health Organization’s estimation formula for a single proportion with absolute precision [18]. We used a significance level of 0.05, absolute precision of 0.05, and an estimated EGFR mutation rate of 51.4% (based on the previous PIONEER study) [12]. A total of 384 NSCLC patients were included in the analysis.

#### Variables

The outcome variable of this study was positive or negative EGFR gene mutation status in NSCLC patients. Additionally, we determined specific EGFR mutation subtypes (Del exon 19, L858R exon 21, exon 20 insertions, point mutations in exon 18, or other dual mutations). EGFR mutations were determined using the real-time PCR method.

Independent patient-related variables included age group (< 60 and ≥ 60), sex (male and female), smoking status (yes or no), histological type (non-squamous or squamous), and CEA subgroup (with subgroups above and below cutoff).

#### Statistical analysis

Quantitative variables were presented as means and standard deviations. Categorical variables were presented as frequencies and percentages. The chi-square test was used for categorical variables to compare differences between groups to determine the sensitivity and specificity of CEA values in predicting EGFR mutation traits by ROC curve analysis. We evaluated the association between CEA and EGFR gene mutations using univariate and multivariate logistic regression analyses.

Statistical analysis was performed using SPSS 25.0. A p-value < 0.05 was considered statistically significant.
**Results**

From January 2018 to December 2022, a total of 384 patients with NSCLC were recruited to the study. Table 1 presents the characteristics of the study population. The mean age was 64.2 ± 10.4 years (range: 26–91), with the majority of patients being over 60 years old (66.4%). Female patients were the majority, and non-squamous histology accounted for 94.0% of cases. A higher proportion of patients had CEA values ≥ 8.95 ng/mL (62.5%), and 61.5% of cases had a history of smoking.

Figure 1 presents the EGFR gene mutation rate and the rate of mutation subgroups. The rate of EGFR mutations is 41.9%. The most common mutations happen in exon 19 deletion (50.9%) and exon 21 L858R (34.8%). Mutations in exon 20 insertion, double mutations, and point mutations at exon 18 are less common, with rates of 5.0%, 5.6%, and 3.7%, respectively.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Number (%)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>161 (41.9)</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Age [years]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 60</td>
<td>108 (42.4)</td>
<td>0.81</td>
</tr>
<tr>
<td>&lt; 60</td>
<td>53 (41.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>62 (26.4)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Female</td>
<td>99 (66.4)</td>
<td></td>
</tr>
<tr>
<td><strong>Smoking</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>63 (26.7)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>No</td>
<td>98 (66.2)</td>
<td></td>
</tr>
<tr>
<td><strong>CEA [ng/mL]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 8.95</td>
<td>38 (26.4)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>≥ 8.95</td>
<td>123 (51.2)</td>
<td></td>
</tr>
</tbody>
</table>

CEA — carcinoembryonic antigen, SD — standard deviation

Table 2 presents the mutation rate of EGFR in patients with NSCLC according to various factors. The overall the EGFR mutation rate is 41.9%; there is a statistically significant difference in the mutation rate of the EGFR gene between males and females (26.4% vs. 66.4%, p < 0.01); smoking status (26.7% vs. 66.2%, p < 0.01); and the rate of EGFR gene mutations in the CEA ≥ 8.95 ng/mL subgroup is higher than the CEA < 8.95 ng/mL subgroup (51.2% vs. 26.4%), with a statistically significant difference p < 0.01.

Figure 2 presents an ROC curve to select a cut-off value for the CEA level, which is used to iden-
Table 3. Association between carcinoembryonic antigen (CEA) status and the EGFR gene mutation in non-small cell lung cancer: univariate and multivariate logistic regression analysis

<table>
<thead>
<tr>
<th>Variables</th>
<th>Rate of EGFR mutation</th>
<th>OR crude (95% CI)</th>
<th>p value</th>
<th>OR adjustment (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEA [ng/mL]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 8.95</td>
<td>1</td>
<td>1</td>
<td>&lt; 0.01</td>
<td>1</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>≥ 8.95</td>
<td>2.93 (1.87–4.59) &lt; 0.01</td>
<td>2.54 (1.57–4.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI — confidence interval; OR — odds ratio

Discussions

This is the first study to determine the mutation rate of the EGFR gene and its association with CEA status in non-small cell cancer patients in Vietnam. We found that the mutation rate of EGFR in NSCLC is 41.9%, with the highest rate being exon 19 deletion mutation at 50.9%. Furthermore, we determined the prognostic value of serum CEA levels in predicting the frequency of EGFR gene mutations. This understanding will provide clinicians with information crucial for decision-making in treatment interventions and may lead to further study on this issue.

In 2015, a systematic review and global map by ethnicity showed that the rate of EGFR gene mutations varies across different regions, with a higher rate often observed in Asia compared to Europe, Africa, and the Americas [9]. The mutation frequency of EGFR in our study is consistent with studies on the Asian population [9, 19–21]. A study at the Bach Mai Hospital, Vietnam, also reported an EGFR gene mutation rate of 40.7%, and the most common mutation was deletion exon 19 [11]. The data in this study indicate that in Vietnam the mutation rate of EGFR is high in the patient population with NSCLC. Although this rate is lower than the mutation rate of EGFR in the Vietnamese patient subgroup in the PIONEER study, there may have been different criteria for sample selection [12]. Our hospital is a leading hospital in the field of oncology in the central region of Vietnam, and the use of Food and Drug Administration (FDA)-approved standard testing methods, along with a large number of patients and sufficient time, gives us confidence that these results are representative of the NSCLC population in central Vietnam.

Our study found a higher incidence of EGFR gene mutations in non-smoking patients (66.2%), women (66.4%), and those with high CEA levels (51.2%). Similar results have been reported in previous studies, with factors such as female sex, non-smoking, and adenocarcinoma being predictors of a higher likelihood of having EGFR mutation [12, 22].

Currently, serum CEA is widely used to diagnose NSCLC. A study in Japan suggested that CEA levels can predict response to EGFR TKI [23]. Data in our study found different rates of EGFR gene alterations among CEA subgroup values, in which the alteration rate was higher in the CEA group with values ≥ 8.95 ng/mL at 51.2% compared to 26.4% in the CEA < 8.95 ng/mL group, and this was highly statistically significant (p < 0.01). The relationship between EGFR gene mutation status and CEA value subgroups is consistent through univariate and multivariate logistic regression analysis. Some previous studies have also reported that high CEA values may be a prognostic factor for a higher likelihood of EGFR gene mutations [17, 24, 25]. A study on a Chinese patient population reported that EGFR gene mutations and CEA are related when performing multivariate logistic regression analysis [15].

In our study, a cutoff value of 8.95 was determined and ROC analysis of CEA levels indicated sensitivity of 76.4%, specificity of 47.5%, and an AUC of 0.627 (95% CI 0.571–0.683).

Table 3 presents the association between CEA status and the EGFR gene mutation in NSCLC. The adjusted model finds the association between CEA and EGFR gene mutations according to the following factors: age, sex, histopathology, and smoking status. The mutation status of the EGFR gene is associated with the CEA value ≥ 8.95 ng/mL group [odds ratio (OR) = 2.54; 95% CI 1.57–4.13; p < 0.01].

Current results are consistent through univariate and multivariate logistic regression analysis 

The results of this study and some previous studies show that serum CEA levels are significantly related to EGFR gene mutations in NSCLC. They may be one of the factors that can predict the patient’s ability to respond to EGFR TKI drugs, guiding clinicians during diagnosis and treatment.

Limitations of our study include it being conducted at a single hospital unit; the results can be used as a reference for other hospitals, but the data will have limited applicability when extrapolating to the entire community of patients with NSCLC. Additionally, this was a descriptive study, so there will be many limitations in determining causal relationships.

Conclusions

The NSCLC population exhibits a high rate of EGFR gene mutations. CEA is a significant predictor of EGFR gene mutation outcomes. Further research is required to reach a definitive conclusion on this matter.

Article Information and Declarations

Data availability statement

The patient was treated at Nghe An Oncology Hospital in Vietnam. The diagnosis and treatment procedures were conducted according to the guidelines laid out by the Vietnamese Ministry of Health. The collection of data adhered to ethical principles and was solely intended to improve the patient’s health outcomes and protect their well-being.

Ethics statement

The study proposal was approved by the Ethics Committee of the Nghe An Oncology Hospital. All patients were provided with information about the medical procedure and consented to it.

Author contributions

L.V.N.: developed the idea and study design, analyzed and interpreted the data, and wrote the article, contributed to data collection; T.K.N., H.D.T.: developed the idea and study design, analyzed and interpreted the data, and wrote the article; U.T.T.N., H.T.T.T., T.V.L.: contributed to data collection. All authors read and approved the final article.

Funding

There was no funding for this study.

Acknowledgments

The authors would like to thank the patients who agreed to participate in the study. This article represents the personal views of the authors and does not represent the views of other organizations with which the authors are affiliated.

Conflict of interest

Authors declare no conflict interest.

Supplementary material

None.


