Scientific Research

Usage of Cox-Regression Model for Forecasting of Survival Rate in Patients with the Early Stage of Non-Small Cell Lung Cancer

Oleksey P. Kolesnik[1], Anatoliy I. Shevchenko[1], Yuriy E. Lyakh[2], Vitaliy G. Gurianov[2], Pavel A. Alyoshechkin[2]

[1]Department of Oncology, Zaporizhzhya State Medical University, Zaporizhzhya, Ukraine
[2]Department of Medical Biophysics, Medical Informatics and Biostatistics, M. Gorky Donetsk National Medical University, Donetsk, Ukraine
Email: kap_kan@mail.ru, ai_shevchenko@ukr.net, liakh@dsmu.edu.ua, vitaliy.gurianov@dsmu.edu.ua, alyoshechkin-pavel@rambler.ru

Abstract

In the past decades a lot of investigations were focused on searching for more accurate markers of lung cancer progression. Researchers indicate that molecular markers may be useful in forecasting of treatment outcome and overall survival rate in patients with non-small cell lung cancer. The aim of our research was to create a forecasting model in order to identify patients with stage I-II of non-small cell lung cancer and dismal prognosis. Our research covered 254 patients with the early stage of non-small cell lung cancer who underwent a cure from June 2008 till December 2012 in the Department of Thoracic Surgery of Zaporizhzhia Regional Clinical Oncologic Dispensary. Surgery was performed for all patients. Adjuvant chemotherapy was performed for 101 patients. In order to carry out multivariate Cox-regression analysis, STATISTICA 6.0 (StatSoft Inc.) program was used. The most significant from 39 variables were selected (tumor size, histological form of tumor, volume of surgical intervention, volume of conducted lymph node dissection, Ki-67 expression, EGFR expression, E-cadherin expression). We propose the computer system which can forecast survival rate in patients with the early stage of non-small cell lung cancer.

Keywords

Forecasting Model; Survival Rate; Non-Small Cell Lung Cancer

1. Introduction

It is well-known that lung cancer is subdivided into non-small cell lung cancer (NSCLC) which makes approximately 80% - 90% from all malignant lung tumors and small cell lung cancer which makes 10% - 20% respectively. Fundamental type of treatment for patients with the early-stage of NSCLC is surgery. However, 22% - 50% of patients have local recurrences after surgery and 48% - 78% of patients have distant metastases [1]-[5].

In the past decades a lot of investigations were focused on searching for more accurate markers of lung cancer progression. Researchers indicate that molecular markers may be useful in forecasting of treatment outcome and overall survival rate in patients with NSCLC [4] [6]-[9]. Currently large number of markers are offered for patients with NSCLC but none of them shows its high significance in clinical practice [6]. Further research of molecular forecasting markers is needed. By that time their significance in routine practice is not adequate [3].

The aim of our research was to create a forecasting model in order to identify patients with stage I-II of NSCLC and dismal prognosis.

2. Materials and Methods

Our research covered 254 patients with the early stage of NSCLC who underwent a cure from June 2008 to December 2012 in the Department of Thoracic Surgery of Zaporizhzhia Regional Clinical Oncologic Dispensary. Surgery was performed for all patients. Adjuvant chemotherapy was performed for 101 patients.

In order to carry out multivariate Cox-regression analysis STATISTICA 6.0 (StatSoft Inc.) program was used. 39 variables (sex, age, height, weight, body mass index (BMI), BMI (categorized index), localization of the tumor and its clinical form (lung, lobes of the lung, central/peripheral lung cancer), tumor size, tumor size (categorized index), criterion "T", criterion "N", stage of disease, histological and morphological form of tumor differentiation, presence of necrosis in tumor, infiltration of visceral pleura, volume of surgical intervention, volume of conducted lymph node dissection, fact of intraperitoneal delegation of lung blood vessels, conducting of adjuvant radiotherapy and chemotherapy, complaints on weakness, loss of appetite, loss of body weight, cough, shortness of breath, chest pain, hemoptysis, hyperthermia, detection of malignant lung neoplasm during prophylactic examination, Ki-67 expression (categorized index), CD31/CD34 expression, Her2-neu expression, EGFR expression, p53 expression, E-cadherin and pan-cytokeratin expression) were used. The most significant of them were selected. Such selection was performed by means of univariate analysis with retaining of variables that have significance level p < 0.1 [10].

3. Results

All obtained after univariate analysis variables and their characteristics are present in **Table 1**.

Research on collinearity (independence) was performed after ANOVA. The variables which dependent on each other were excluded. Only the most significant markers were left. Reverse selection of variables was used. First model was calculated with all possible predictive variables and then re-calculated after removing of variability with the least significant interconnection, with response variable (survival rate). The process lasted until the model contained only statistically significant variables (**Table 2**).

Survival rate based on these prognostic factors can be calculated for any patient with the early stage of NSCLC.

Let us consider some examples:

Example 1. Patient G., 71-year-old man. Received institutional treatment at Department of Thoracic Surgery of Zaporizhzhia Regional Clinical Oncologic Dispensary between August 26, 2008 and September 17, 2008.

Diagnosis: Central cancer of right lung superior lobe. pT2N0M0. Stage IB.

Concomitant diseases: Coronary heart disease: atherosclerotic cardiosclerosis. Circulatory inefficiency stage II. Obliterating atherosclerosis of the lower extremities.

Clinical examination data: Considered complaints on pain in the right side of the chest, shortness of breath and cough since February 2008. At clinic were performed such examination as: abdominal ultrasound: ultrasound signs of urine acid diathesis; CT-scan of lung: in right lung superior lobe near the root of lung, peribronchially, relatively to proximal bronchus, 49×29 mm neoplasm was marked. Enlarged lymph nodes: paratracheal up to 10 mm, bifurcation up to 13 mm, bronchial (right side) up to 13 mm, bronchopulmonary (right side) up to 13 mm; Fiber-optic bronchoscopy: proximal bronchus with peribronchial tumor growth; Cytological conclusion No7153: moderately differentiated squamous cell carcinoma; Spirography: 1st degree respiratory failure.

Table 1. Variables and their characteristics after univariate analysis.

No	Variable	Min	Max	$(M \pm m)^*$ $(\% \pm m\%)^{**}$	Significance level, p
1	Sex	Men		83.9% ± 2.3%	0.93
		Women		16.1% ± 2.3%	
2	Age, years	34	77	60.7 ± 0.51	0.89
3	Height, cm	150	191	171.4 ± 0.46	0.89
4	Weight, kg	46	115	73.7 ± 0.84	0.56
5	BMI	16.9	38.9	25.1 ± 0.28	0.54
6	BMI (categorized)	<30 mg/m^2		83.9% ± 2.3%	0.33
		≥30 mg/m^2		16.1% ± 2.3%	
7	Localization, lung	Right		49.6% ± 3.1%	0.55
		Left		50.4% ± 3.1%	
8	Localization, lobe of lung	Upper		62.6% ± 3.0%	0.91
		Middle		3.5% ± 1.2%	
		Lower		33.9% ± 3.0%	
9	Localization	Central		39.8% ± 3.1%	0.48
		Peripheral		60.2% ± 3.1%	
10	Tumor size, cm	1.2	14	4.88% ± 1.13	<0.001
11	Tumor size, (categorized)	≥3 cm		78.0% ± 2.6%	0.002
		<3 cm		22.0% ± 2.6%	
12	Criterion "T"	1		12.6% ± 2.1%	0.31
		2		83.5% ± 2.3%	
		3		3.9% ± 1.2%	
13	Criterion "N"	0		71.3% ± 2.8%	0.002
		1		28.7% ± 2.8%	
14	Stage	1		52.4% ± 3.1%	0.002
		2		47.6% ± 3.1%	
15	Histological form	Squamous		58.7% ± 3.1%	0.01
		Non-squamous		41.3% ± 3.1%	
16	Morphological differentiation	High		23.2% ± 2.6%	0.43
		Moderate		50.8% ± 3.1%	
		Low		26.0% ± 2.8%	
17	Presence of necrosis	Absent		55.5% ± 3.1%	0.44
		Present		44.5% ± 3.1%	
18	Infiltration of visceral pleura	Absent		57.5% ± 3.1%	0.48
		Present		42.5% ± 3.1%	
19	Volume of surgical intervention	Lobeectomy		63.0% ± 3.0%	0.002
		Pneumonectomy		37.0% ± 3.0%	
20	Volume of conducted lymph node dissection	Systemic		31.9% ± 2.9%	0.007
		Sampling		68.1% ± 2.9%	
21	Intraperitoneal delegation of lung blood vessels	Absent		87.8% ± 2.1%	0.018
		Present		12.2% ± 2.1%	
22	Adjuvant chemotherapy	Absent		60.2% ± 3.1%	0.91
		Present		39.8% ± 3.1%	

Continued

23	Radiotherapy	Absent	94.5% ± 1.4%	0.33
		Present	5.5% ± 1.4%	
24	Complaint on "weakness"	Absent	67.7% ± 2.9%	0.31
		Present	32.3% ± 2.9%	
25	Complaint on "loss of appetite"	Absent	96.1% ± 1.2%	0.19
		Present	3.9% ± 1.2%	
26	Complaint on "loss of body weight"	Absent	87.0% ± 2.1%	0.02
		Present	13.0% ± 2.1%	
27	Complaint on "presence of cough"	Absent	35.0% ± 3.0%	0.45
		Present	65.0% ± 3.0%	
28	Complaint on "shortness of breath"	Absent	59.1% ± 3.1%	0.52
		Present	40.9% ± 3.1%	
29	Complaint on "chest pain"	Absent	66.5% ± 3.0%	0.33
		Present	33.5% ± 3.0%	
30	Complaint on "hemoptysis"	Absent	90.2% ± 1.9%	0.055
		Present	9.8% ± 1.9%	
31	Complaint on "hyperthermia"	Absent	84.6% ± 2.3%	0.48
		Present	15.4% ± 2.3%	
32	Detection of malignant neoplasm during prophylactic examination	Absent	58.7% ± 3.1%	0.47
		Present	41.3% ± 3.1%	
33	Ki67 expression (categorized index)	≥25%	55.1% ± 3.4%	0.002
		<25%	44.9% ± 3.4%	
34	CD31/CD34 expression	0 212	68.7% ± 2.11	0.25
35	Her2-neu expression	Absent	71.2% ± 2.9%	<0.001
		Present	28.8% ± 2.9	
36	EGFR expression	Absent	51.3% ± 3.5%	<0.001
		Present	48.7% ± 3.5%	
37	P53 expression	≥25%	57.3% ± 3.3%	0.85
		<25%	42.7% ± 3.3%	
38	E-cadherin expression	0,+,++ (low expression)	36.7% ± 3.7%	0.19
		+++, ++++ (high expresion)	63.3% ± 3.7%	
39	pan-cytokeratin expression	Absent	86.7% ± 2.5%	0.12
		Present	13.3% ± 2.5%	

Note: p—connection with survival rate (censored). *M (average); m (standart error); **m% (standart error of proportions).

After examination surgery was performed: right-sided pneumonectomy, lymph node dissection D1. Histological conclusion: No 33959-70: non-keratinizing squamous cell lung carcinoma with foci of necrosis. In bronchus and lymph nodes at the operating region the malignant growth elements were not found (pT2N0M0 G2).

Thus, prognostic factors significance is as follows:

Tumor size 5 cm—5 points;

Histological type of tumor: squamous cell carcinoma—2 points;

Volume of surgical intervention: pneumonectomy—2 points;

Lymph node dissection volume: D1—1 point;

Table 2. Statistically significant variables of forecasting model.

Variable	Coefficient (*β*)	Standard error	Significance level, p	OR	95% CI
Tumor size, cm	0.14	0.06	0.021	1.16	1.06 - 1.76
Histological form	−0.79	0.24	<0.001	0.45	0.04 - 0.29
Volume of surgical intervention	1.16	0.29	<0.001	3.19	3.13 - 24.95
Volume of conducted lymph node dissection	−1.21	0.32	<0.001	0.29	0.04 - 0.47
Ki-67 expression	0.64	0.28	0.020	1.91	1.62 - 10.69
EGFR expression	1.68	0.34	<0.001	5.40	2.03 - 16.25
E-cadherin expression	−1.07	0.29	<0.001	0.34	0.09 - 0.52

Ki-67 expression: 17%—1 point;

EGFR expression: "+"—1 point;

E-cadherin expression: "++"—2 points.

Figure 1 shows forecasting of survival rate in patient G. As it may be seen, survival rate median is not reached, patient's chance to five-year survival rate is more than 60%. Control period of observation—June 01, 2013, patient is alive. Period of observation is 64 months.

Example 2. Patient M., 66-year-old man. Received institutional treatment at Department of Thoracic Surgery of Zaporizhzhia Regional Clinical Oncologic Dispensary between July 14, 2010 and August 02, 2010.

Diagnosis: Peripheral cancer of right lung superior lobe. pT2N0M0. Stage IB.

Clinical examination data: Considered complaints on shortness of breath and cough near 2 months ago. At clinic were performed such examinations as: X-ray chest examination: in rear segment of superior lobe of right lung marked 6×6, 5 cm extensive tuberous neoplasm with path to the root of lung. Also, subpleurally in the right root of the lung determined enlarged bronchial-pulmonic lymph nodes; CT-scan: in S2 of right lung viewed 40×40 mm peripheral neoplasm with luminous contours. Also, in S10 of lower lobe marked 21×26 mm neoplasm. Mediastinal lymph nodes were not enlarged; Abdominal ultrasound: abdomen pathology is not found; Fiber-optic bronchoscopy: without intraluminal pathology;

After examination surgery was performed: right-side upper lobectomy with lymph node dissection D1. Histological conclusion No 22729-34: non-small cell lung carcinoma with necrosis. In bronchus and lymph nodes from operating region, the malignant growth elements were not found (pT2N0M0 G3).

Thus, prognostic factors significance is as follows:

Tumor size 6 cm—6 points;

Histological type of tumor: large cell carcinoma—1 point;

Volume of surgical intervention: upper lobectomy—1 point;

Lymph node dissection volume: D1—1 point;

Ki-67 expression: 100%—2 points;

EGFR expression: "+++"—2 points;

E-cadherin expression: "+"—1 point.

Figure 2 shows forecasting of survival rate in patient M. As it may be seen, survival rate median equals 10 - 11 months; patient's chance to five-year survival rate is almost equals 0%. Indeed patient's survival rate was 11 months and he died from the progression of disease.

4. Discussion

Several studies were devoted to prognostic factors search [4] [7]-[9] [11]. Clinical, morphological, genetic factors have been researched as prognostic factors for patients with NSCLC.

It is noted that the most significant factors are the follows: overall clinical status of the patients, tumor size, vascular invasion, poor differentiation and presence of necrosis in the tumor, distant extrapulmonary metastases, type of conducted chemotherapy, number of white blood cells, level of CYFRA 21–1, Ca 125, Ca 15–3, and Ca 72–4 [2] [4] [7] [8] [10]-[13]. In our research 39 prognostic markers of 254 patients were examined. At the same time independent prognostic factors selected by Cox-regression analysis were: size and histological type of tumor, type of surgical intervention, volume of conducted lymph node dissection.
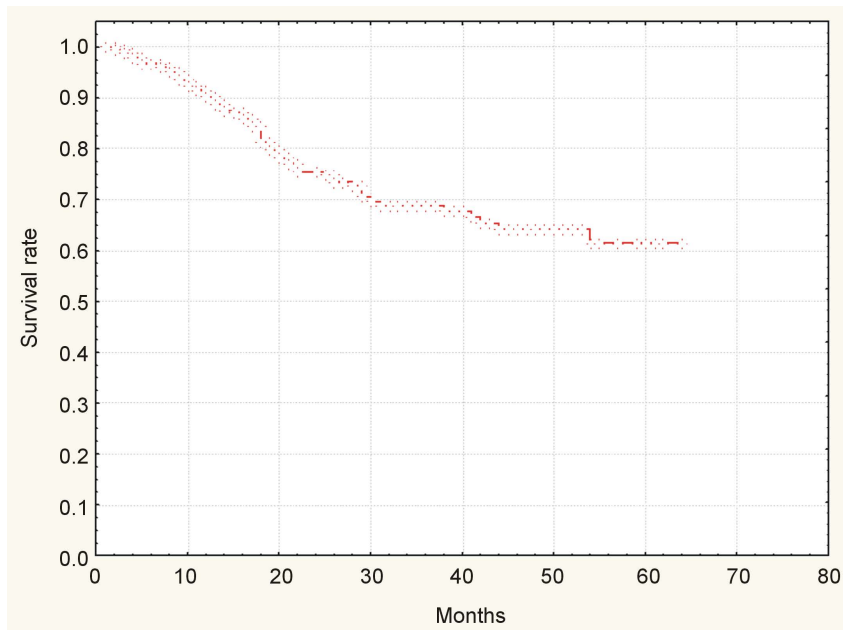
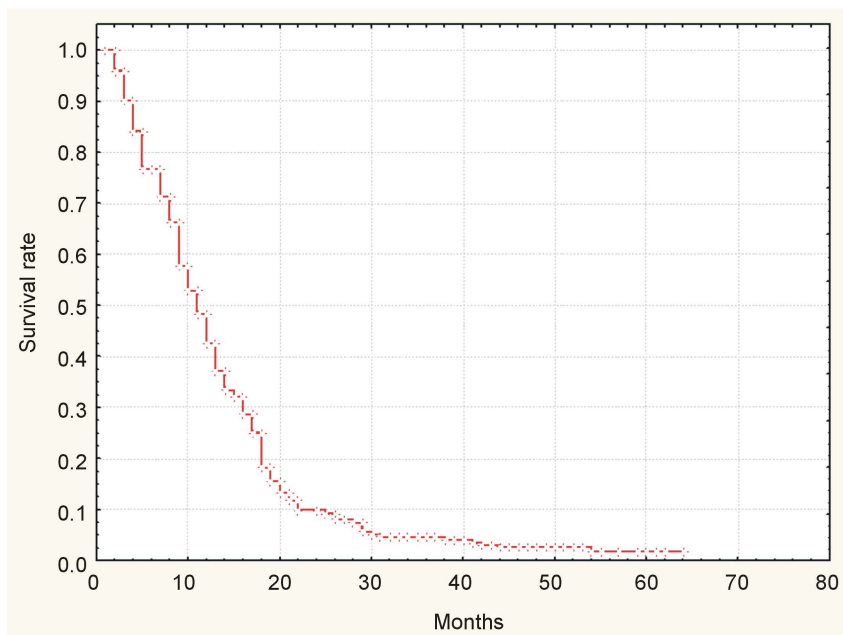**Figure 1.** Patient G. forecasting of survival rate.

Figure 2. Patient M. forecasting of survival rate.

Among the molecular and genetic markers, such factors were prognostically significant: p53, factor 8, erb-b2, Cd-44, retinoblastoma recessive gene, cyclin E, cyclin D1, methylation DAPK, IL-10 expression, p21. Also genetic prognostic markers have been identified—genes Wnt3a, Erbb3, LSK, Rnd3 [2] [5] [6]. We have demonstrated the prognostic role of Ki-67, EGFR and E-cadherin expression. Other factors that were analyzed showed no effect on survival rate in patients with stage I-II of NSCLC.

Several studies of foreign authors state that the usage of single marker for prediction is less accurate than the use of combination of factors [14]-[16]. We have also observed that prediction through multivariate analysis was more accurate than the use of prognostic factor. It is necessary to use forecasting regression model in order to predict survival rate in patients with the early stage of NSCLC. It allows selecting patients with dismal prognosis.

The future researches of conducting adjuvant polychemotherapy in this category of patients are necessary.

Cox-regression model with tumor size, histological type of tumor, volume of surgical intervention, volume of lymph node dissection, Ki-67 expression, EGFR expression and E-cadherin expression, could be useful for forecasting of survival rate in patients with stage I-II of NSCLC.

References

[1] Chen, D.-T., Hsu, Y.-L., William, J.F., *et al.* (2011) Prognostic and Predictive Value of a Malignancy-Risk Gene Signature in Early-Stage Non-Small Cell Lung Cancer. *Journal of the National Cancer Institute*, **103**, 1859-1870. http://dx.doi.org/10.1093/jnci/djr420

[2] D'Amico, T.A., Massey, M., Herndon, J.E., *et al.* (1999) A Biologic Risk Model for Stage I Lung Cancer: Immuno-histochemical Analysis of 408 Patients with the Use of Ten Molecular Markers. *The Journal of Thoracic and Cardiovascular Surgery*, **117**, 736-743. http://dx.doi.org/10.1016/S0022-5223(99)70294-1

[3] Konopa, K. (2010) Do We Have Markers to Select Patients for Adjuvant Therapies of Non-Small-Cell Lung Cancer? *Annals of Oncology*, **7**, vii199-vii202.

[4] Park, S.Y., Lee, H.-S., Jang, H.-J., *et al.* (2011) Tumor Necrosis as a Prognostic Factor for Stage IA Non-Small Cell Lung Cancer. *The Annals of Thoracic Surgery*, **91**, 1668-1673. http://dx.doi.org/10.1016/j.athoracsur.2010.12.028

[5] Zhu, C.-Q., Shih, W., Ling, C.-H. and Tsao, M.-S. (2006) Immunohistochemical Markers of Prognosis in Non-Small Cell Lung Cancer: A Review and Proposal for a Multiphase Approach to Marker Evaluation. *Journal of Clinical Pathology*, **59**, 790-800. http://dx.doi.org/10.1136/jcp.2005.031351

[6] Raz, D.J., Ray, M.R., Kim, J.Y., *et al.* (2008) A Multigene Assay Is Prognostic of Survival in Patients with Early-Stage Lung Adenocarcinoma. *Clinical Cancer Research*, **14**, 5565-5570. http://dx.doi.org/10.1158/1078-0432.CCR-08-0544

[7] Harpole Jr., D.H., Herndon, J.E., Wolfe, W.G., *et al.* (1995) A Prognostic Model of Recurrence and Death in Stage I Non-Small Cell Lung Cancer Utilizing Presentation, Histopathology, and Oncoprotein Expression. *Cancer Research*, **55**, 51-56.

[8] Holdenrieder, S., Nagel, D., Heinemann, V., *et al.* (2008) Predictive and Prognostic Biomarker Models in Advanced Lung Cancer. *Journal of Clinical Oncology*, 26, 19010.

[9] López-Encuentra, A., López-Ríos, F., Conde, E., *et al.* (2011) Composite Anatomical-Clinical-Molecular Prognostic Model in Nonsmall Cell Lung Cancer. *European Respiratory Journal*, **37**, 136-142. http://dx.doi.org/10.1183/09031936.00028610

[10] Lang, T.A., Math, M.S. and Leonov, V.P. (2011) How to Describe the Statistics in Medicine. An Annotated Guide for Authors, Editors and Reviewers. M:, 480.

[11] Lu, C., Soria, J.-C., Tang, X., *et al.* (2004) Prognostic Factors in Resected Stage I Non-Small-Cell Lung Cancer: A Multivariate Analysis of Six Molecular Markers. *Journal of Clinical Oncology*, **22**, 4575-4583. http://dx.doi.org/10.1200/JCO.2004.01.091

[12] Van der Pijl, L.L.R., Birim, O., Van Gameren, M., *et al.* (2010) Validation of a Prognostic Model to Predict Survival after Non-Small-Cell Lung Cancer Surgery. *European Journal Cardio-Thoracic Surgery*, **38**, 615-620. http://dx.doi.org/10.1016/j.ejcts.2010.03.028

[13] Rubio, L., Vera-Sempere, F.J., Lopez-Guerrero, J.A., *et al.* (2005) A Risk Model for Non-small Cell Lung Cancer Using Clinicopathological Variables, Angiogenesis and Oncoprotein Expression. *Anticancer Research*, **25**, 497-504.

[14] Dosaka-Akita, H., Hommura, F., Mishina, T., *et al.* (2001) A Risk-Stratification Model of Non-Small Cell Lung Cancers Using Cyclin E, Ki-67, Andrasp21: Different Roles of G1 Cyclins in Cell Proliferation and Prognosis. *Cancer Research*, **61**, 2500-2504.

[15] Hilbe, W., Dirnhofer, S., Oberwasserlechner, F., *et al.* (2003) Immunohistochemical Typing of Non-Small Cell Lung Cancer on Cryostat sections: Correlation with Clinical Parameters and Prognosis. *Journal of Clinical Pathology*, **56**, 736-741. http://dx.doi.org/10.1136/jcp.56.10.736

[16] Williams, B.A., Sugimura, H., Endo, C., *et al.* (2006) Predicting Postrecurrence Survival among Completely Resected Nonsmall-Cell Lung Cancer Patients. *The Annals of Thoracic Surgery*, **81**, 1021-1027. http://dx.doi.org/10.1016/j.athoracsur.2005.09.020

Abbreviatios

ANOVA—Analysis of Variance
BMI—Body Mass Index
CI—Confidence Interval
CYFRA—Cytokeratin Fragment
DAPK—Death-Associated Protein Kinase
EGFR—Epidermal Growth Factor Receptor
NSCLC—Non-Small Cell Lung Cancer
OR—Odds Ratio