Neoplasma 2023; 70(3): 443–450

doi:10.4149/neo_2023_230305N115

Prognostic implications of the *EGFR* polymorphism rs763317 and clinical variables among young Chinese lung cancer population

Zheng-Xing LI[1,#], Fan WANG[2,#], Zhen-Yu SUN[2,#], Chang XU[3], Zi-Xiu ZOU[4], Zong-Xu ZHAO[5], Hai-Jian WANG[4], Shi-Cheng GUO[6], Li JIN[4,7], Qiang LI[8], Jiu-Cun WANG[4,7], Fan XIA[9,*], Xiao-Qiang YUE[10,*], Jun-Jie WU[11,12,*]

[1]Department of Surgery, Navy Military Medical University Affiliated to Changhai Hospital, Shanghai, China; [2]Company of Basic Medical Science, Navy Military Medical University, Shanghai, China; [3]Clinical College of Xiangnan University, Chenzhou, China; [4]State Key Laboratory of Genetic Engineering, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China; [5]Hainan Medical College, Haikou, China; [6]Department of Medical Genetics, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, Wisconsin, United States; [7]Human Phenome Institute, Fudan University, Shanghai, China; [8]Department of Respiratory and Critical Care Medicine, Shanghai East Hospital, Tongji University, Shanghai, China; [9]Department of Respiratory Diseases, Navy 905 Hospital, Shanghai, China; [10]Department of Traditional Chinese Medicine, Changzheng Hospital, Naval Medical University (Second Military Medical University), Shanghai, China; [11]Department of Pulmonary Medicine, Zhongshan Hospital, Fudan University, Shanghai, China; [12]Department of Pulmonary and Critical Care Medicine, Shanghai Geriatric Medical Center, Shanghai, China

**Correspondence: wjjxcc@126.com; xqyue@smmu.edu.cn; xfs2000@hotmail.com*
#Contributed equally to this work.

Received March 5, 2023 / Accepted June 9, 2023

The 5-year survival rate for patients with lung cancer, the world's second most frequent malignant tumor, is less than 20%, and its prognosis cannot be clearly predicted. Our aim was to analyze the epidermal growth factor receptor (EGFR) rs763317 (G>A) single nucleotide polymorphism and its association with prognosis in Chinese Han lung cancer patients. 839 patients with primary lung cancer were recruited, and genomic DNA was extracted and genotyped by SNPscan. Kaplan-Meier technique and multivariate Cox proportional hazards model were used to analyze the association between prognosis and *EGFR* polymorphism rs763317. A significant association after stratification by age, significantly increased lung cancer risk was associated with the AA homozygous genotype of rs763317 (adjusted hazard ratio = 2.53, 95% CI: 1.31–4.88, p=0.005), and conferred a poor survival for lung cancer patients (MST: median survival time: 13.6 months) compared with GG genotype (MST: 41.5 months), and in the recessive model AA genotype (AA vs. GG + GA; adjusted hazard ratio = 2.57, 95% CI: 1.34–4.93, p=0.004) who were young (<60 years) had a significantly increased risk of death. The *EGFR* polymorphism rs763617 might serve as a significant genetic marker for predicting the prognosis of lung cancer.

Key words: lung cancer; EGFR; rs763317; polymorphism; prognosis

Lung cancer is the second most prevalent malignant tumor worldwide; the most common cancers are lung cancers in China, which is the main cause of cancer mortality globally. Moreover, the 5-year survival rate for patients is less than 20%, and the onset and poor prognosis of cancer are getting young [1–4]. The patient's age is both a risk factor and an independent prognostic variable for developing lung cancer [5, 6]. The impact of lung cancer prognosis is a complex biological process related to many variables, such as age. Numerous studies found [7–10] that genetic variation is also a significant factor affecting the prognosis of patients who have lung cancer. These results indicate that genetic factors and clinical variables influence cancer patients' risk and survival prognosis [11–15]. Identification of precise prognostic genetic markers can increase the efficacy of therapies and survival time.

The epidermal growth factor (EGF) receptor, the first receptor protein tyrosine kinase described, is a member of the ErbB family, regulates several signaling pathways that include cell proliferation, and is one of the most potent oncogenes that are commonly altered in cancers [16, 17]. *EGFR* genetic polymorphism results in aberrant *EGFR* trafficking in situations such as *EGFR* amplification or overexpression, resulting in kinase activation, which leads to enhanced signaling and the formation of tumors [18–20]. Receptor endocytosis is changed in oncogenic *EGFR* polymorphism

and large genomic rearrangements (seen in ovarian and lung malignancies), which adds to enhanced signaling characteristics and is an important indicator for disease recurrence or shorter patient survival [21]. Meanwhile, EGFR signaling is frequently altered in glioblastoma and lung cancer due to gene amplification and/or protein overexpression, mutations, or in-frame deletions [22].

Previous studies have found interethnic differences in the distribution of *EGFR* polymorphisms and mutations [23], cancer development [24, 25], and patient prognosis [26]. Therefore, to analyze the association of rs763317 polymorphism with the prognosis of Chinese Han patients with lung cancer, the authors used newly labeled SNPs in the *EGFR* of lung cancer patients to undertake a retrospective investigation of the alterations in genotype structure and frequency distribution. Interestingly, the study data observe that rs763317 G>A, a standard marker in *EGFR* intron 1, significantly conferred a poor prognosis for young patients with lung cancer.

Patients and methods

Patient demographics and data collection. Between January 2009 and November 2019, 888 Han Chinese patients with primary lung cancer were recruited, consisting of 536 patients from Changhai Hospital connected with the Naval Military Medical University (Second Military Medical University) and 352 patients from Taizhou Institute of Health Sciences, Fudan University. The inclusion criteria were as follows: patients must have been histopathologically diagnosed with primary lung cancer and have no history of malignant diseases of any organs. Clinical information was received from the patient's medical records, and clinical information was gathered by telephone follow-up. The Fudan University School of Life Sciences Ethics Committee authorized this research, and the subjects gave informed consent to gathering epidemiological survey data and blood samples.

Table 1. Clinical variables and survival prognosis of Chinese Han patients with lung cancer.

Stratification	N (%)	MST (month)	Log rank p-value
Total	839	36.73	
Gender			0.01
Male	610 (72.7)	34.27	
Female	229 (27.3)	40.17	
Age			0.003
≥60	524 (62.5)	33.20	
<60	315 (37.5)	40.87	
Smoking status			<0.001
Yes	582 (69.4)	33.90	
No	237 (28.2)	41.03	
Unknown	20 (2.4)	67	
Family cancer history			0.462
Yes	302 (36)	33.63	
No	537 (64)	38.03	
Histological type			0.211
ADC	367 (43.7)	38.80	
SCC	282 (33.6)	33.63	
SCLC	72 (8.6)	33.90	
Others*	118 (14.1)	36.20	
TNM staging			<0.001
Stage I+II	154 (18.4)	113.93	
Stage III+IV	625 (74.5)	29.40	
Unknown	60 (7.1)	66.43	

*Notes: other carcinomas include adenosquamous carcinoma, large cell carcinoma, carcinosarcoma, and mucoepidermoid carcinoma; Abbreviations: MST-median survival time; CI-confidence interval

SNP selection and genotyping. Before starting treatment, each patient contributed 5 ml of blood, and the genomic DNA used for PCR optimization was extracted from blood samples at a constant temperature of 37 °C from primary lung cancer patients using the QIAamp Blood Mini Kit (Qiagen, Hilden, Germany, 51106) [27]. For genotyping, a 248-plex SNPscan TM kit (catalog number G0104; Genesky Biotechnologies, Shanghai, China) was used [28–30]. Genotyping quality was determined using thorough protocols, and the genotyping detection rate was over 95%. Internal positive control samples were examined for deviations from Hardy-Weinberg equilibrium (HWE), and duplicate samples were genotyped. The laboratory personnel who performed the genotyping analysis were not informed of the patient's clinical information.

Statistical analyses. Pearson's chi-square test evaluated the HWE. Overall survival (OS) was calculated between the time patient information was collected and the time of the final follow-up or death from any cause. After univariate or multivariate Cox regression analysis, the age- and sex-adjusted hazard ratio (HR) and 95% confidence interval (CI) were calculated. Cox regression analyses the SNPs' allele, genotype, and dominant and invisible genetic models, and stratified analysis was performed for age, sex, smoking status, family tumor history, histological type of lung cancer, and TNM stage. The $p<0.01$ value was judged statistically significant. Version 3.6.2 of R (Vienna, Austria) was used for all statistical analyses.

Results

Patient condition and clinical variables. In this research, forty-nine patients did not meet the inclusion criteria and were thus eliminated, and 839 patients were recruited for analysis. The subjects were all Han people of the same ethnic group in China. There were 668 (79.6%) deaths, 103 (12.3%) survived for more than 5 years, 68 (8.1%) failed follow-up. 610 (72.7%) males, and 229 (27.3%) females; 315 (37.5%) <60 years old, 524 (62.5%) ≥60 years old; 582 (69.4%) smokers, 237 (28.2%) non-smokers; 302 (36.0%) patients with a history of malignant tumor; 367 patients (43.7%) of adenocarcinoma and 282 patients (33.6%) of squamous cell carcinoma. 72 patients (8.6%) had small cell lung cancer, whereas 118 patients (14.1%) had other types of cancer. 154 patients

(18.4%) with stages I and II, whereas stages III and IV were diagnosed in 625 (74.5%) patients (Table 1).

The association between patient clinical variables and prognosis. According to Table 1, MST of all patients was 36.73 months, and the MST for male patients was substantially lower than for female patients (34.27 vs. 40.17 months, p=0.01); MST in patients younger than 60 years of age was substantially higher than that in patients older than 60 years of age (40.87 vs. 33.20 months, p=0.003); MST of smoking patients was substantially higher than that of non-smoking patients (41.03 vs. 33.90 months, p<0.001); MST was shorter in patients with a family history of cancer than in patients without a family history of malignancy (33.63 vs. 38.03 months, p=0.462); the association between patients of histological type and prognosis showed no significant difference (p=0.211). Patients with stage I+II lung cancer had a substantially higher MST than those with stage III+IV lung cancer (113.93 vs. 29.40 months, p=0.001). There was no statistically significant difference across hospitals in the association between patient clinical variables and lung cancer outcomes.

Prognostic analysis. According to the log-rank test, MST varied substantially with age (p<0.01), and the results showed that in young (<60 years) patients, the MST of patients with low expression of the AA genotype (mutated genotype) was 13.60 months, respectively, while the MST of patients with high expression of GG genotype (non-mutated genotype) was 41.50 months, patients with the GG genotype had a significantly longer MST than those with the AA genotype, consistent with the Cox regression analysis. Furthermore, the MST of patients with low expression of AA genotype (AA vs. GG + GA) was 13.60 months in the recessive model, the MST in patients with high expression recessive model GG + GA lung cancer was 42.30 months, and the MST also showed a difference that was statistically significant.

Association of *EGFR* polymorphism with lung cancer prognosis. At the *EGFR* rs763317 locus, 527 GG genotypes, 267 GA genotypes, and 38 AA genotypes were detected. The percentage of genotype detection was 99.16% (Table 3). The genotype frequency of the *EGFR* rs763317 was the same as that of the HWE (p=0.561), demonstrating that the research sample was genetically balanced and that the survey data were reliable. In the death group, the frequencies of genotype GG and AA were 79.32% (418/527) and 76.32% (29/38), accordingly, and observe no substantial difference between the allele frequencies of the two genotypes (p>0.01, Table 3).

Association of *EGFR* polymorphism with prognosis by stratification of clinical variables in patients with lung cancer. Interestingly, we observed a greater mortality risk in younger (<60 years) lung cancer patients with AA genotypes compared with GG genotypes (adjust hazard ratio for AA=2.53 95%CI: 1.31–4.88, p=0.005, shown in Table 3), and conferred a poor survival for lung cancer patients (MST: 13.60 months) compared with GG genotype (MST: 41.50 months), and a statistically significant difference in MST was observed between those with the AA

Table 2. Association between *EGFR* rs763317 polymorphism in allele mode and prognosis in Chinese Han patients with lung cancer.

Stratification	Death/Survive		Polymorphism frequency %(A)	HR (95% CI)	p-value	HR[a] (95% CI)	p-value[a]
	G (ref)	A					
Total	1051/270	273/70	20.61	1.04 (0.91–1.19)	0.581	1.03 (0.90–1.18)	0.674
Gender							
Male	782/170	210/44	21.06	1.12 (0.96–1.31)	0.146	1.12 (0.96–1.31)	0.137
Female	269/100	63/26	19.43	0.84 (0.64–1.11)	0.227	0.80 (0.60–1.05)	0.104
Age (year)							
≥60	687/134	183/38	21.21	0.99 (0.84–1.17)	0.896	1.00 (0.85–1.17)	0.966
<60	364/136	90/32	19.61	1.12 (0.89–1.41)	0.350	1.11 (0.88–1.40)	0.359
Smoking status							
Yes	764/145	198/43	20.96	1.02 (0.82–1.27)	0.833	1.01 (0.81–1.27)	0.895
No	266/113	72/23	20.04	1.16 (0.89–1.50)	0.273	1.14 (0.87–1.48)	0.339
Family cancer history							
Yes	389/84	99/24	20.64	1.09 (0.91–1.30)	0.868	1.08 (0.90–1.29)	0.423
No	662/186	174/46	20.60	1.05 (0.89–1.25)	0.536	1.06 (0.89–1.25)	0.503
Histological type							
ADC	434/145	114/35	20.47	1.09 (0.88–1.34)	0.425	1.05 (0.85–1.29)	0.648
SCC	383/69	93/17	19.57	1.09 (0.87–1.37)	0.436	1.10 (0.87–1.38)	0.418
NSCLC	964/249	244/65	20.30	1.06 (0.92–1.22)	0.438	1.05 (0.91–1.20)	0.533
SCLC	87/21	29/5	23.94	0.87 (0.57–1.33)	0.519	0.90 (0.59–1.38)	0.625
TNM staging							
I+II	847/131	223/39	21.94	0.96 (0.83–1.11)	0.485	1.14 (0.76–1.71)	0.512
III+IV	578/98	494/72	45.57	1.02 (0.90–1.15)	0.565	0.95 (0.82–1.10)	0.520

Notes: [a]adjusted by age, gender; Abbreviations: CI-confidence interval; HR-hazard ratio; ref-reference

Table 3. Association between *EGFR* rs763317 polymorphism in genotype model and prognosis in Chinese Han patients with lung cancer.

Stratification	Death/Survive			Polymorphism frequency % (A)	G/A vs. G/G		A/A vs. G/G	
	G/G (ref)	G/A	A/A		HR[a] (95% Cl)	p-value[a]	HR[a] (95% Cl)	p-value[a]
Total	418/109	215/52	29/9	10.82	1.04 (0.89–1.23)	0.602	1.01 (0.69–1.48)	0.939
Gender								
Male	309/67	164/36	23/4	10.45	1.11 (0.92–1.35)	0.267	1.29 (0.84–1.98)	0.241
Female	109/42	51/16	6/5	11.79	0.86 (0.61–1.20)	0.371	0.53 (0.23–1.22)	0.135
Age (year)								
≥60	271/55	145/24	19/7	9.60	1.10 (0.90–1.35)	0.355	0.79 (0.49–1.26)	0.315
<60	147/54	70/28	10/2	12.86	0.95 (0.71–1.26)	0.711	2.53 (1.31–4.88)	0.005
Smoking status								
Yes	302/56	160/33	19/5	9.91	1.00 (0.82–1.21)	0.991	1.00 (0.62–1.59)	0.986
No	107/49	52/15	10/4	12.24	1.25 (0.89–1.75)	0.191	1.04 (0.54–2.00)	0.902
Family cancer history								
Yes	156/31	77/22	11/1	11.41	0.90 (0.68–1.18)	0.445	1.59 (0.85–2.97)	0.147
No	262/78	138/30	18/8	10.49	1.15 (0.94–1.42)	0.177	0.90 (0.55–1.46)	0.664
Histological type								
ADC	177/61	80/23	17/6	12.64	1.09 (0.84–1.42)	0.513	1.01 (0.61–1.67)	0.970
SCC	152/27	79/15	7/1	8.19	1.14 (0.87–1.50)	0.346	1.06 (0.50–2.28)	0.874
NSCLC	386/101	192/47	26/9	10.78	1.10 (0.93–1.31)	0.274	0.95 (0.64–1.42)	0.803
SCLC	32/8	23/5	3/0	11.27	0.63 (0.36–1.10)	0.103	3.65 (1.03–12.99)	0.045
TNM staging								
I+II	48/51	26/23	2/3	18.30	1.30 (0.80–2.10)	0.293	0.83 (0.20–3.45)	0.793
III+IV	338/52	171/27	26/6	9.12	0.96 (0.80–1.15)	0.656	0.90 (0.60–1.34)	0.599

Notes: [a]adjusted by age, gender; Abbreviations: CI-confidence interval; HR-hazard ratio; ref-reference

Table 4. Association between *EGFR* polymorphism in dominant model and prognosis in Chinese Han patients with lung cancer.

Stratification	Death/Survive		Polymorphism frequency % (A)	HR (95% CI)	p-value[a]	HR[a] (95% CI)	p-value[a]
	G/G (ref)	G/A+A/A					
Total	418/109	244/61	36.66	1.04 (0.89–1.22)	0.596	1.00 (0.69–1.45)	0.997
Gender							
Male	309/67	187/40	37.65	1.11 (0.93–1.34)	0.245	1.13 (0.94–1.36)	0.180
Female	109/42	57/21	34.06	0.87 (0.63–1.19)	0.379	0.81 (0.58–1.12)	0.197
Age (year)							
≥60	271/55	164/31	37.42	1.04 (0.86–1.27)	0.682	1.05 (0.86–1.28)	0.606
<60	147/54	80/30	35.37	1.04 (0.79–1.36)	0.796	1.03 (0.79–1.36)	0.814
Smoking status							
Yes	302/56	179/38	37.74	1.01 (0.84–1.22)	0.921	1.00 (0.83–1.21)	0.988
No	107/49	62/19	34.18	1.22 (0.89–1.68)	0.207	1.21 (0.88–1.66)	0.235
Family cancer history							
Yes	156/31	88/23	37.25	0.95 (0.73–1.23)	0.685	0.95 (0.73–1.23)	0.696
No	262/78	156/38	36.33	1.11 (0.91–1.35)	0.317	1.12 (0.92–1.37)	0.273
Histological type							
ADC	177/61	97/29	34.62	1.11 (0.87–1.42)	0.412	1.08 (0.84–1.38)	0.558
SCC	152/27	86/16	36.30	1.13 (0.86–1.47)	0.385	1.13 (0.87–1.48)	0.355
NSCLC	386/101	218/56	36.01	1.09 (0.92–1.29)	0.325	1.08 (0.91–1.28)	0.357
SCLC	32/8	26/5	43.66	0.67 (0.40–1.14)	0.141	0.71 (0.41–1.23)	0.220
TNM staging							
I+II	48/51	28/26	35.29	1.23 (0.77–1.98)	0.384	1.25 (0.78–2.00)	0.360
III+IV	338/52	197/33	37.10	0.95 (0.80–1.14)	0.581	0.95 (0.80–1.13)	0.574

Notes: [a]adjusted by age, gender; Abbreviations: CI-confidence interval; HR-hazard ratio; ref- reference

and GG genotypes, as shown by the Kaplan-Meier curves (log-rank p=0.00667; Figure 1A). Compared with the GG + GA recessive model, patients with AA genotype younger (<60 years) lung cancer had a higher risk of death (adjust hazard ratio for recessive model AA (p=2.57 CI: 1.34–4.93, p=0.004, Table 5) and a poorer prognosis of death (AA vs. GG + GA; MST: 13.60 vs. 42.30 months), and the Kaplan-Meier curves demonstrated that MST was shorter in young patients with the AA genotype compared to those with the GG+GA genotype (log-rank p=0.00399; Figure 1B). Additionally, there was no statistical significance between allele patterns and prognosis (p>0.01, Table 2). In the dominant model, the dominant genotype GG + GA was not significantly associated with the prognosis compared to genotype AA (p>0.01, Table 4).

Discussion

This research investigated the association between the prognosis and *EGFR* polymorphism rs763317 by taking different variables such as exposure to smoke, age, and sex among the Chinese Han population with lung cancer as the research object. The authors observed that rs763317 G>A, an SNP in the *EGFR* intron 1, was clearly associated with poor prognosis in young (<60 years) lung cancer patients.

Recent research has shown that *EGFR* gene polymorphisms have significance in the growth of malignant cancers

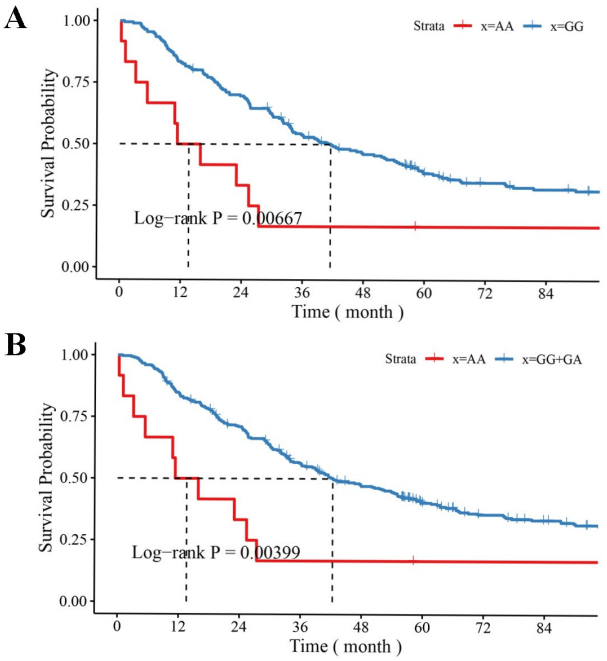

Figure 1. The Kaplan-Meier survival curve analysis of the *EGFR* gene rs763316 SNP and youth (<60 years) survival probability of lung cancer patients. A) Association of survival probability between AA genotype (mutant) lung cancer patients and GG genotype (non-mutant) lung cancer patients; B) Association of survival probability between the recessive model AA genotype lung cancer patients and the recessive model GG + GA genotype.

Table 5. Association between *EGFR* polymorphisms in recessive model and prognosis in Chinese Han patients with lung cancer.

Stratification	Death/Survive		Polymorphism frequency %(A)	HR (95% CI)	p-value[a]	HR[a] (95% CI)	p-value[a]
	G/G + G/A (ref)	A/A					
Total	633/161	29/9	4.57	1.35 (0.89–2.06)	0.159	1.25 (0.82–1.90)	0.308
Gender							
Male	473/103	23/4	4.48	1.35 (0.89–2.06)	0.159	1.25 (0.82–1.90)	0.308
Female	160/58	6/5	4.80	0.61 (0.27–1.38)	0.236	0.56 (0.25–1.28)	0.169
Age(year)							
≥60	416/79	19/7	4.99	0.76 (0.48–1.20)	0.241	0.76 (0.48–1.21)	0.244
<60	217/82	10/2	3.86	2.47 (1.31–4.66)	0.005	2.57 (1.34–4.93)	0.004
Smoking status							
Yes	462/89	19/5	4.17	1.09 (0.69–1.73)	0.701	1.00 (0.63–1.58)	0.988
No	159/64	10/4	5.91	1.03 (0.54–1.95)	0.934	0.97 (0.51–1.85)	0.931
Family cancer history							
Yes	233/53	11/1	4.03	1.73 (0.95–3.18)	0.075	1.65 (0.89–3.06)	0.113
No	400/108	18/8	4.87	0.87 (0.54–1.39)	0.560	0.85 (0.53–1.37)	0.515
Histological type							
ADC	257/84	17/6	6.32	1.07 (0.65–1.75)	0.794	0.98 (0.60–1.61)	0.940
SCC	231/42	7/1	2.85	1.04 (0.49–2.21)	0.919	1.02 (0.48–2.16)	0.965
NSCLC	578/148	26/9	4.60	0.97 (0.65–1.44)	0.876	0.92 (0.62–1.37)	0.680
SCLC	55/13	3/0	4.23	5.89 (1.73–20.04)	0.004	4.29 (1.22–15.08)	0.023
TNM staging							
I+II	74/74	2/3	3.27	0.88 (0.22–3.62)	0.863	0.75 (0.18–3.12)	0.696
III+IV	509/79	26/6	5.16	0.94 (0.63–1.40)	0.761	0.91 (0.61–1.36)	0.646

Notes: [a]adjusted by age, gender; Abbreviations: CI-confidence interval; HR-hazard ratio; ref-reference

and affect the therapeutic effect [31]. In Jou's [32] control research of 730 patients with lung cancer and 730 subjects without cancer, rs763317 polymorphism in *EGFR* intron 1 was shown to be significantly related to elevated lung cancer in the Taiwanese population, especially non-smoking female adenocarcinoma patients carrying the A allele. Another study [33] found that *EGFR* was overexpressed in over 90% of cervical cancer patients by immunohistochemistry (IHC) to detect the expression level of *EGFR* before treatment and the role of ionizing radiation (IR)-induced expression changes, and this overexpression was associated with poor outcomes for cervical cancer patients. Furthermore, *EGFR* gene polymorphism caused the occurrence of lung cancer in significant interethnic differences, which is more frequent in Asia, meanwhile, compared to other types of lung cancer, it was shown that the frequency of *EGFR* mutations is also greater in women, non-smokers, and adenocarcinoma patients [34]. The author's study observed that young (<60 years) lung cancer patients with the rs763317 AA genotype had a poor prognosis.

The rs763317 G>A polymorphism is located in intron 1, 6.9 kilobases downstream of the polymorphism region of the dinucleotide CA repeat (near the second enhancer) in the *EGFR* gene [35]. Analyses of the Mitogen-Activated Protein Kinase/Extracellular-signal-Regulated Kinase (MAPK/ERK) signaling system [36] demonstrated that the MAPK/ERK pathway was the center of frequently altered genes throughout the progression of multiple primary lung cancers, including mutated downstream genes and altered upstream activators. The most enriched alterations in the early stages, such as atypical adenomatous hyperplasia, were in MAP2K1 and BRAF, and in the later stages of adenocarcinoma, BRAF and *EGFR* alterations were the most enriched. This suggested that the MAPK/ERK pathway aberrations may be the critical factors in the multiple primary lung cancers tumorigeneses. Studies [37, 38] have shown that localized amplification of *EGFR* intron 1 site, usually confined to short alleles, is negatively correlated with *EGFR* expression levels and that the carcinogenesis and poor prognosis of patients caused by *EGFR* protein overexpression are less attributable to gene amplification and more attributable to *EGFR* gene polymorphism. All of these data demonstrate the significance of *EGFR* in the genesis of cancer prognosis.

In this research, *EGFR* rs763317 G>A was associated with poor lung cancer outcomes in young patients (<60 years). Different ages and regions are significant differences according to the frequency of occurrence of *EGFR* mutations and have been shown for the adult population in a meta-analysis, and its mutation rate in China is significantly higher than in other regions and has an important effect on the prognosis of patients [39]. For example, in an analysis [40] of the expression of *EGFR* in mice during the skin development process, strong *EGFR* expression in all age groups except the small embryonic day (E12.5) group, showed moderate expression. In another research on the prediction of *EGFR* gene polymorphism in the prognosis of patients with locally progressed pharyngeal squamous cell carcinoma after concurrent chemoradiotherapy, the *EGFR* R521K genetic polymorphism was found to have a substantially increased probability of death in young patients (<50 years) after age analysis [41]. Genetic alterations within the driver gene and increased mutation rate are major are important factors for poor survival prognosis in young (<60 years) lung cancer patients. In the study of Nagashima [42] on 12 young patients with lung cancer, 9 out of 12 patients (75%) had *EGFR* mutation and EML4-ALK fusion gene driver oncogene. Additionally, Arnold's [43, 44] study on the association between age and target gene mutations in NSCLC patients found that the occurrence of gene mutations in young (<50 years) lung cancer patients was 59% greater than that in elderly patients in a retrospective controlled study of 2,237 lung cancer patients. As a result, we hypothesized that the age difference in *EGFR* SNP affecting lung cancer patients' survival prognosis was caused by different *EGFR* driver gene mutations and an increase in the mutation rate at different ages.

Previous screening methods for *EGFR* variants with diagnostic and prognostic potential were diverse, mainly involving polymerase chain reaction (PCR) and sequencing methods, and dominated applications such as genome-wide sequencing [45]. Genome-wide sequencing may overlook crucial genes due to the size and complexity of the data. The SNPscan technology is a multi-gene mutation screening method that improves the multiplex ligation-dependent probe amplification technology. This technique utilizes four different fluorescent dyes and lengthens ligations thereby eliminating the need for long ligation probe synthesis and increasing the number of SNPs simultaneously interrogated to over 100 loci in one reaction [46]. Therefore, our study used SNPscan technology to directly sequence the important locus of rs763317 in the candidate *EGFR* gene and confirmed that the rs763317 polymorphism is associated with the prognosis of young lung cancer patients, improving the sensitivity and accuracy of the analysis.

The study has some advantages and limitations. We performed SNP detection on known tags in *EGFR* in all patients to avoid using genome-wide association studies that miss an important gene due to the complexity and volume of data. Unfortunately, there are several limitations to this research. First, the evaluation of the gene model is limited by the correspondingly small sample size, leading to possible bias in the analysis results. Secondly, cell biology has not verified changes in SNP rs763317 in *EGFR* expression levels in patients. Therefore, different cell biology or biochemical experiments based on patients' blood should be conducted to verify this conclusion.

In conclusion, the results showed that among age-related *EGFR* gene polymorphisms, rs763317 polymorphisms were associated with poor prognosis in patients with lung cancer. Younger carriers of the *EGFR* polymorphism (rs763317, AA) had a worse prognosis than carriers of the *EGFR* gene

(no polymorphism GG; or recessive model GG+GA), and these results have substantial therapeutic implications for predicting the prognosis of lung cancer patients by adding new predictive biological markers.

Acknowledgments: Supported in part through the computational resources and staff expertise provided by Scientific Computing at the Fudan University in Shanghai. The work is currently receiving five grants from the National Natural Science Foundation of China (Grant NO. 81372236), the Shanghai Postdoctoral Research Foundation in 2021 (Class A) (Grant NO. 12R21411500), the National Natural Science Foundation of China (Grant NO. 81972822 and NO. 82272863), Shanghai Science and Technology Commission project (Grant NO. 19401971700). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

[1] XIA C, DONG X, LI H, CAO M, SUN D et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. Chin Med J (Engl) 2022; 135: 584–590. https://doi.org/10.1097/CM9.0000000000002108

[2] SUNG H, FERLAY J, SIEGEL RL, LAVERSANNE M, SOERJOMATARAM I et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021; 71: 209–249. https://doi.org/10.3322/caac.21660

[3] SIEGEL RL, MILLER KD, FUCHS HE, JEMAL A. Cancer Statistics, 2021. CA Cancer J Clin 2021; 71: 7–33. https://doi.org/10.3322/caac.21654

[4] WU F, WANG L, ZHOU C. Lung cancer in China: current and prospect. Curr Opin Oncol 2021; 33: 40–46. https://doi.org/10.1097/CCO.0000000000000703

[5] YOUNG RP, HOPKINS RJ, HAY BA, EPTON MJ, MILLS GD et al. Lung cancer susceptibility model based on age, family history and genetic variants. PLoS One 2009; 4: e5302. https://doi.org/10.1371/journal.pone.0005302

[6] LUBIN JH, CAPORASO NE. Cigarette smoking and lung cancer: modeling total exposure and intensity. Cancer Epidemiol Biomarkers Prev 2006; 15: 517–523. https://doi.org/10.1158/1055-9965.EPI-05-0863

[7] LUO J, MARTUCCI VL, QUANDT Z, GROHA S, MURRAY MH et al. Immunotherapy-Mediated Thyroid Dysfunction: Genetic Risk and Impact on Outcomes with PD-1 Blockade in Non-Small Cell Lung Cancer. Clin Cancer Res 2021; 27: 5131–5140. https://doi.org/10.1158/1078-0432.CCR-21-0921

[8] ZHOU YK, LI XP, YIN JY, ZOU T, WANG Z et al. Association of variations in platinum resistance-related genes and prognosis in lung cancer patients. J Cancer 2020; 11: 4343–4351. https://doi.org/10.7150/jca.44410

[9] YEO MK, CHOI SY, SEONG IO, SUH KS, KIM JM et al. Association of PD-L1 expression and PD-L1 gene polymorphism with poor prognosis in lung adenocarcinoma and squamous cell carcinoma. Hum Pathol 2017; 68: 103–111. https://doi.org/10.1016/j.humpath.2017.08.016

[10] SHIN KM, JUNG DK, HONG MJ, KANG HJ, LEE WK et al. The pri-let-7a-2 rs1143770C>T is associated with prognosis of surgically resected non-small cell lung cancer. Gene 2016; 577: 148–152. https://doi.org/10.1016/j.gene.2015.11.036

[11] CANCER GENOME ATLAS RESEARCH N. Author Correction: Comprehensive molecular profiling of lung adenocarcinoma. Nature 2018; 559: E12. https://doi.org/10.1038/s41586-018-0228-6

[12] ZHANG H, CHEN B, ZOU Z, FENG J, LI Y et al. Associations Between CAMKK1 Polymorphism rs7214723 and the Prognosis of Patients With Lung Cancer. Front Oncol 2021; 11: 757484. https://doi.org/10.3389/fonc.2021.757484

[13] MEHTA A, DOBERSCH S, ROMERO-OLMEDO AJ, BARRETO G. Epigenetics in lung cancer diagnosis and therapy. Cancer Metastasis Rev 2015; 34: 229–241. https://doi.org/10.1007/s10555-015-9563-3

[14] DURUISSEAUX M, ESTELLER M. Lung cancer epigenetics: From knowledge to applications. Semin Cancer Biol 2018; 51: 116–128. https://doi.org/10.1016/j.semcancer.2017.09.005

[15] QIAN Y, XIE L, LI L, FENG T, ZHU T et al. Association between sex hormones regulation-related SNP rs12233719 and lung cancer risk among never-smoking Chinese women. Cancer Med 2021; 10: 1880–1888. https://doi.org/10.1002/cam4.3772

[16] THOMAS R, WEIHUA Z. Rethink of EGFR in Cancer With Its Kinase Independent Function on Board. Front Oncol 2019; 9: 800. https://doi.org/10.3389/fonc.2019.00800

[17] BASHIR NA, RAGAB ES, KHABOUR OF, KHASSAWNEH BY, ALFAQIH MA et al. The Association between Epidermal Growth Factor Receptor (EGFR) Gene Polymorphisms and Lung Cancer Risk. Biomolecules 2018; 8: 53. https://doi.org/10.3390/biom8030053

[18] SIGISMUND S, AVANZATO D, LANZETTI L. Emerging functions of the EGFR in cancer. Mol Oncol 2018; 12: 3–20. https://doi.org/10.1002/1878-0261.12155

[19] CHUNG I, AKITA R, VANDLEN R, TOOMRE D, SCHLESSINGER J et al. Spatial control of EGF receptor activation by reversible dimerization on living cells. Nature 2010; 464: 783–787. https://doi.org/10.1038/nature08827

[20] WILSON KJ, GILMORE JL, FOLEY J, LEMMON MA, RIESE DJ 2ND. Functional selectivity of EGF family peptide growth factors: implications for cancer. Pharmacol Ther 2009; 122: 1–8. https://doi.org/10.1016/j.pharmthera.2008.11.008

[21] YARDEN Y, PINES G. The ERBB network: at last, cancer therapy meets systems biology. Nat Rev Cancer 2012; 12: 553–563. http://doi.org/10.1038/nrc3309

[22] ROSKOSKI R JR. The ErbB/HER family of protein-tyrosine kinases and cancer. Pharmacol Res 2014; 79: 34–74. https://doi.org/10.1016/j.phrs.2013.11.002

[23] SHIGEMATSU H, LIN L, TAKAHASHI T, NOMURA M, SUZUKI M et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005; 97: 339–346. https://doi.org/10.1093/jnci/dji055

[24] ALANAZI IO, SHAIK JP, PARINE NR, AZZAM NA, ALHARBI O et al. Association of HER1 and HER2 Gene Variants in the Predisposition of Colorectal Cancer. J Oncol 2021; 2021: 6180337. https://doi.org/10.1155/2021/6180337

[25] HAN L, LEE CK, PANG H, CHAN HT, LO IL et al. Genetic predisposition to lung adenocarcinoma among never-smoking Chinese with different epidermal growth factor receptor mutation status. https://doi.org/10.1016/j.lungcan.2017.10.012

[26] OBRADOVIĆ J, DJORDJEVIĆ N, TOŠIC N, MRDJANOVIĆ J, STANKOVIĆ B et al. Frequencies of EGFR single nucleotide polymorphisms in non-small cell lung cancer patients and healthy individuals in the Republic of Serbia: a preliminary study, https://doi.org/10.1007/s13277-016-4930-4

[27] JURISIC V, OBRADOVIC J, TOSIC N, PAVLOVIC S, KULIC M et al. Effects of DMSO, glycerol, betaine and their combinations in detecting single nucleotide polymorphisms of epidermal growth factor receptor (EGFR) gene promoter sequence in non-small-cell lung cancer (NSCLC) patients, J Pharm Biomed Anal 2016; 128: 275–279. https://doi.org/10.1016/j.jpba.2016.05.010

[28] ZHANG YH, XU Q, ZHAO Z, WU J, LIU WX et al. Polymorphism rs7214723 in CAMKK1 and lung cancer risk in Chinese population. Tumour Biol 2013; 34: 3147–3152. https://doi.org/10.1007/s13277-013-0883-z

[29] TANG LL, CHEN FY, WANG H, HU XL, DAI X et al. Haplotype analysis of eight genes of the monoubiquitinated FANCD2-DNA damage-repair pathway in breast cancer patients. Cancer Epidemiol 2013; 37: 311–317. https://doi.org/10.1016/j.canep.2012.12.010

[30] YIN J, WANG L, SHI Y, SHAO A, TANG W et al. Interleukin 17A rs4711998 A>G polymorphism was associated with a decreased risk of esophageal cancer in a Chinese population. Dis Esophagus 2014; 27: 87–92. https://doi.org/10.1111/dote.12045

[31] MENDELSOHN J, BASELGA J. Status of epidermal growth factor receptor antagonists in the biology and treatment of cancer, J Clin Oncol 2003; 21: 2787–2799. https://doi.org/10.1200/JCO.2003.01.504

[32] JOU YS, LO YL, HSIAO CF, CHANG GC, TSAI YH et al. Association of an EGFR intron 1 SNP with never-smoking female lung adenocarcinoma patients. Lung Cancer 2009; 64: 251–256. https://doi.org/10.1016/j.lungcan.2008.09.014

[33] CERCIELLO F, RIESTERER O, SHERWEIF M, ODERMATT B, CIERNIK IF. Is EGFR a moving target during radiotherapy of carcinoma of the uterine cervix? Gynecol Oncol 2007; 106: 394–399. https://doi.org/10.1016/j.ygyno.2007.04.019

[34] JURISIC V, OBRADOVIC J, PAVLOVIC S, DJORDJEVIC N. Epidermal Growth Factor Receptor Gene in Non-Small-Cell Lung Cancer: The Importance of Promoter Polymorphism Investigation. Anal Cell Pathol (Amst) 2018; 2018: 6192187. https://doi.org/10.1155/2018/6192187

[35] MARINOVIC S, VUKOVIC K, SKRTIC A, POLJAK M, PETEK S et al. Epidermal growth factor receptor intron 1 polymorphism and microsatellite instability in sporadic colorectal cancer. Oncol Lett 2021; 21: 131. https://doi.org/10.3892/ol.2020.12392

[36] LIANG N, BING Z, WANG Y, LIU X, GUO C et al. Clinical implications of EGFR-associated MAPK/ERK pathway in multiple primary lung cancer. Clin Transl Med 2022; 12: e847. https://doi.org/10.1002/ctm2.847

[37] NOMURA M, SHIGEMATSU H, LI L, SUZUKI M, TAKAHASHI T et al. Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. PLoS Med 2007; 4: e125. https://doi.org/10.1371/journal.pmed.0040125

[38] BUERGER H, PACKEISEN J, BOECKER A, TIDOW N, KERSTING C et al. Allelic length of a CA dinucleotide repeat in the egfr gene correlates with the frequency of amplifications of this sequence--first results of an inter-ethnic breast cancer study. J Pathol 2004; 203: 545–550. https://doi.org/10.1002/path.1542

[39] JURISIC V, VUKOVIC V, OBRADOVIC J, GULYAEVA LF, KUSHLINSKII NE et al. EGFR Polymorphism and Survival of NSCLC Patients Treated with TKIs: A Systematic Review and Meta-Analysis. J Oncol 2020; 2020: 1973241. https://doi.org/10.1155/2020/1973241

[40] FALODAH FA, AL-KARIM S. Immuno- and gene expression analysis of EGFR and Nestin during mice skin development. Tissue Cell 2016; 48: 274–281. https://doi.org/10.1016/j.tice.2016.02.004

[41] SU NW, LEU YS, LEE JC, LIU CJ, CHENG CY et al. EGF and EGFR genetic polymorphisms predict prognosis in locally advanced pharyngolaryngeal squamous cell carcinoma patients receiving postoperative concurrent chemoradiotherapy. Onco Targets Ther 2014; 7: 2197–2204. https://doi.org/10.2147/OTT.S70188

[42] NAGASHIMA O, OHASHI R, YOSHIOKA Y, INAGAKI A, TAJIMA M et al. High prevalence of gene abnormalities in young patients with lung cancer. J Thorac Dis 2013; 5: 27–30. https://doi.org/10.3978/j.issn.2072-1439.2012.12.02

[43] ARNOLD BN, THOMAS DC, ROSEN JE, SALAZAR MC, BLASBERG JD et al. Lung Cancer in the Very Young: Treatment and Survival in the National Cancer Data Base. J Thorac Oncol 2016; 11: 1121–1131. https://doi.org/10.1016/j.jtho.2016.03.023

[44] HAN X, MA S. [Current Situation of Clinical Feature and Gene Phenotype of Young Adult Lung Cancer]. Zhongguo Fei Ai Za Zhi 2020; 23: 388–392. https://doi.org/10.3779/j.issn.1009-3419.2020.101.21

[45] OBRADOVIC J, TODOSIJEVIC J, JURISIC V. Application of the conventional and novel methods in testing EGFR variants for NSCLC patients in the last 10 years through different regions: a systematic review. Mol Biol Rep 2021; 48: 3593–3604. https://doi.org/10.1007/s11033-021-06379-w

[46] LI H, WANG B, LIU D, WANG T, LI Q et al. SNPscan as a high-performance screening tool for mutation hotspots of hearing loss-associated genes. Genomics 2015; 106: 83–87. https://doi.org/10.1016/j.ygeno.2015.05.004