BMJ Open

EGFL6 Expression Predicts Poor Prognosis in Non-Small Cell Lung Cancer

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-021385
Article Type:	Research
Date Submitted by the Author:	27-Dec-2017
Complete List of Authors:	Chang, Chun-Chi ; Changhua Christian Hospital, Division of Chest Medicine, Department of Internal Medicine Sung, Wen-Wei; Chung Shan Medical University, School of Medicine Hsu, Hui-Ting ; Changhua Christian Hospital, Department of Pathology Yeh, Chung-Min ; Changhua Christian Hospital, Department of Pathology Lee, Chien-Hsun ; Changhua Christian Hospital, Department of Pathology Chen, Ya-Ling ; Changhua Christian Hospital, Department of Nursing Liu, Ta-Chih ; Kaohsiung Medical University, Graduate Institute of Clinical Medicine Yeh, Kun-Tu; Changhua Christian Hospital, Department of Pathology
Keywords:	EGF like domain multiple 6, EGFL6, prognosis, non-small cell lung cancer, overall survival

EGFL6 Expression Predicts Poor Prognosis in Non-Small Cell Lung Cancer

Chun-Chi Chang[1,2], Wen-Wei Sung[3,4,5], Hui-Ting Hsu[6], Chung-Min Yeh[6,7], Chien-Hsun Lee[6], Ya-Ling Chen[8], Ta-Chih Liu[1,9,*], Kun-Tu Yeh[3,6,*]

[1]Graduate Institute of Clinical Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

[2]Division of Chest Medicine, Department of Internal Medicine, Changhua Christian Hospital, Changhua, Taiwan

[3]School of Medicine, Chung Shan Medical University, Taichung, Taiwan

[4]Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan

[5]Department of Urology, Chung Shan Medical University Hospital, Taichung, Taiwan

[6]Department of Pathology, Changhua Christian Hospital, Changhua, Taiwan

[7]Department of Medical Technology, Jen-Teh Junior College of Medicine, Nursing and Management, Miaoli, Taiwan

[8]Department of Nursing, Changhua Christian Hospital, Changhua, Taiwan

[9]Division of Hematology and Oncology, Department of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan

* Ta-Chih Liu and Kun-Tu Yeh contributed equally in this work.

Corresponding author and request for reprints:

Kun-Tu Yeh (TEL: 886-4-7238595 ext. 4830, E-mail: 10159@cch.org.tw)

Department of Pathology, Changhua Christian Hospital, Changhua, Taiwan

Running head: EGFL6 predicts poor prognosis in lung cancer

Key words: EGF like domain multiple 6; EGFL6; prognosis; non-small cell lung cancer; overall survival

Abstract

Objective

Non-small cell lung cancer is a common cancer in both genders and has poor clinical outcome. Our aim was to evaluate the role of EGF-like domain multiple 6 (EGFL6) and its prognostic significance in non-small cell lung cancer.

Methods

EGFL6 expression was studied by immunohistochemical staining of specimens from 150 patients with non-small cell lung cancer. The correlation between clinicopathological features and EGFL6 expression was quantitatively analyzed. We used Kaplan-Meier analysis and Cox proportional hazard models to examine the prognostic value of EGFL6 for overall survival.

Results

No significant correlation was found between EGFL6 expression and clinical parameters. However, patients with high EGFL6 expression levels were prone to have poor prognosis with borderline significance. More interestingly, while we grouped the patients according to the age using a medium value, a high EGFL6 expression was significantly associated with poor clinical outcome in young patients. This finding was further confirmed by grouping the patients into three groups according to age. The hazard ratio in patients with high EGFL6 expression was higher in younger patients than in older patients.

Conclusion

High EGFL6 expression may serve as a marker for poor prognosis of non-small cell lung cancer, especially in younger patients.

Strengths and limitations of this study

1. This is the first study evidenced that the EGFL6 expression in the tumor tissue of non-small cell lung cancer might be an independent prognostic marker.
2. The prognostic role of EGFL6 was more significant in young patients.
3. The major limitation of this study is the limited sample size. Further investigation is necessary for clinical application.

Introduction

Non-small cell lung cancer (NSCLC) is a major public health problem worldwide and is a leading cause of cancer death in Taiwan [1,2]. Unlike other types of cancer, which have shown steady increases in survival in recent years, NSCLC continues to have a poor clinical outcome, with a 5-year survival of only 18% [3]. Early detection of NSCLC might improve the clinical outcome; however, no suitable screening tools are available that are both cost effective and the efficient. NSCLC screening via low dose CT scans can provide early detection of lung lesions, such as ground glass opacity lesions [2]. Nevertheless, further intervention might be unnecessary for tissue confirmation and surgical intervention for tumor resection of benign lesions. Therefore, identifying specific biomarkers for selection of patients with a malignant potential of their NSCLC would help in clinical decision making for cancer follow up and the timing of surgical intervention.

The malignant potential of a tumor with metastatic behavior is determined by complex processes, including tumor cell migration, invasion, and angiogenesis to the target site [1,4-6]. The motif features of the epidermal growth factor (EGF) repeat superfamily is conserved, with cysteines and glycines positioned in a domain of 30 to 40 residues [7,8]. EGF-like proteins are characterized by multiple EGF repeats [9]. EGF-like repeat family members are secreted as cell surface molecules, and the binding of EGF-like proteins to their receptors promotes tumor malignancy [8,10-12]. Among these proteins, EGF-like domain 6 (EGFL6) is a secreted protein with involvement in tissue development, promotion of tumor cell migration, and angiogenesis [8,10-14]. A role for EGFL6 in promoting tumor malignancy is indicated in several types of cancer. For example, oral cancer patients show high plasma levels of

EGFL6. Moreover, the plasma EGFL6 level was higher in patients with advanced stage disease than with early stage disease [10]. In ovarian cancer, EGFL6 regulates cell migration and asymmetric division via SHP2 oncoprotein, with concomitant activation of ERK [11].

Accumulating evidence indicates crucial roles for EGFL6 in promoting tumor malignancy. However, the potential for an association between EGFL6 expression and the prognosis of NSCLC remains to be addressed. The aim of the present study was to evaluate the expression of EGFL6 and its clinical significance in NSCLC.

Materials and Methods

Patients

Our study examined 150 samples of NSCLC. Cancers were staged according to the AJCC Cancer Staging Manual (7th edition). Clinicopathological features, including histological type, differentiation, lymph node metastasis, TNM stage, and tumor size, were assessed in this study. Histological diagnosis was confirmed by two pathologists, as described previously [15]. The study was approved by the Institutional Review Board and the Ethics Committee of the Changhua Christian Hospital, Changhua, Taiwan (CCH IRB 170511).

Immunohistochemical Staining and Evaluation of Cytoplasm EGFL6

Immunohistochemistry (IHC) staining was performed at the Department of Surgical Pathology, Changhua Christian Hospital, as previously described [15] [16] using anti-human cytoplasm EGFL6 antibody (EGFL6 antibody, 1:50 dilution; Santa Cruz Biotechnology). Immunoreactivity scores were analyzed by three pathologists, using scores defined as previously described [16]. The pathologists were blinded to the prognostic data of the study. A final agreement was obtained for each score by viewing through a multi-headed microscope (Olympus BX51 10-headed microscope).

Statistical Analysis

The $\chi 2$ test was applied for continuous or discrete data analysis. The associations between cytoplasm EGFL6 expression and patient survival were estimated using univariate analysis and the Kaplan–Meier method and assessed using the log-rank test [1]. Potential confounders were adjusted by Cox regression models of multivariate

analysis, with the cytoplasm EGFL6 expression fitted as indicator variables. All statistical analyses were conducted using SPSS statistical software (version 15.0) (SPSS, Inc., Chicago, IL). All statistical tests were 2-sided, and the values of $p<0.05$ were considered statistically significant.

Results

Cytoplasm EGFL6 is Expressed in the Majority of NSCLC Specimens

We verified the role of EGFL6 in the clinical outcome of NSCLC by recruiting 150 NSCLC patients. EGFL6 expression was detected with IHC staining, as shown in Figure 1. Of the 150 patients, only 6 patients (4%) showed no detectable EGFL6 expression in their cytoplasm. The relationships of the cytoplasm EGFL6 expression according to the clinicopathological characteristics are listed in Table 1. The mean age was 62.1±11.6 years (mean±SD). The EGFL6 expression in the cytoplasm was not significantly associated clinicopathological characteristics, including the gender, grade, age, and TNM stage.

The Prognostic Role of Cytoplasm EGFL6 Expression in NSCLC Patients

We further evaluated the prognostic role of cytoplasm EGFL6 expression in NSCLC patients. Overall survival data were collected and no data were missing among the 150 patients. The mean and median follow-up times after surgery were 5.2 and 3.2 years (range from 0.1 to 11.0 years), respectively. The 5-year survival rate was 42.1%. During the survey, 99 (66.0%) patients died. In the univariate analysis, patients with advanced stage disease, age >63 years old, and male gender had significantly poorer clinical outcomes (Table 2). These factors were still significantly associated with poor prognosis in the multivariate analysis (HR=2.241, 95% CI=1.443–3.481, p<0.001 for stage; HR=1.997, 95% CI=1.303–3.062, p=0.002 for age; HR=1.802, 95% CI=1.180–2.753, p=0.006, Table 1). In the prognostic role of cytoplasm EGFL6 in NSCLC, patients with high cytoplasm EGFL6 expression had lower 5-year survival rate and shorter medium survival compared with those with low cytoplasm EGFL6 expression (52.0% vs 37.0% for 5-year survival; 5.7 years vs 2.5 years for medium

survival, Figure 2A). In the univariate and multivariate analysis, cytoplasm EGFL6 showed borderline significance (HR=1.519, 95% CI=0.980–2.355, p=0.061 for univariate analysis; HR=1.515, 95% CI=0.975–2.354, p=0.064 for multivariate analysis, Table 2).

Significant Prognostic Role of Cytoplasm EGFL6 Expression in Young NSCLC Patients

We examined the potential prognostic role of cytoplasmic EGFL6 in NSCLC patients by analyzing their clinical outcomes according to clinicopathological characteristics. We identified a significant association of cytoplasm EGFL6 in patients with younger age. As show in Figure 2B, patients aged less than 69 years old who had high cytoplasmic EGFL6 expression had lower 5-year survival rate and shorter median survival times when compared with patients with low cytoplasm EGFL6 expression (65.7% vs 40.9% for 5-year survival; 8.4 years vs 2.8 years for median survival, Figure 2B). We confirmed this finding using different cutoff points of age, as the use of median age as a cutoff point resulted in a significantly poor prognosis for patients with high EGFL6 (HR: 2.118, 95% CI: 1.082–4.145, p=0.029, Table 3). Moreover, HR was increased in patients with younger age (HR: 2.894, 95% CI: 1.245–6.726, p=0.014 for age≦59; HR: 2.104, 95% CI: 1.184–3.739, p=0.011 for age ≤ 69, Table 3).

Discussion

In this study, we identified the prognostic role of cytoplasmic EGFL6 in NSCLC especially in patients with younger age. This is the first study to provide clinical evidence of EGFL6 expression in NSCLC. No association was noted between EGFL6 expression and clinical parameters, but the significantly poor clinical outcome of patients with high EGFL6 expression supports the findings of previous reports regarding this expression in other types of cancer [10 12 17].

The role of EGFLF6 was first identified with a non-tumor model in cell division and tissue development [12-14 18]. In a bone remodeling model, EGFL6 induced angiogenesis via a paracrine mechanism to promote the angiogenesis and migration of endothelial cells [12]. Inhibition of phosphorylated ERK in this model decreased the cell migration ability [12]. In a zebrafish model, EGFL6 promoted angiogenesis that depended on the RGD domain and the activation of the Akt and Erk pathways [18]. Loss of EGFL6 decreased the numbers of endothelial cells and vessels, suggesting that EGFL6 acts as a positive regulator of functional vessel formation [18].

Increasing evidence supports a role for EGFL6 in regulating tumor malignancy and shows the potential for EGFL6 to serve as a prognostic marker and therapeutic target. In ovarian cancer, EGFL6 is associated with poor clinical outcome, which is further

explained by its role in promoting cancer cell proliferation and asymmetric division [11]. A model using $ALDH^+$ ovarian cancer cells showed that EGFL6 signaling involves integrin, SHP2, and ERK [11]. Molecular analysis of tumor vascular cells in ovarian cancer via immunohistochemistry-guided laser-capture microdissection and genome-wide transcriptional profiling results also support this result [17]. Oral cancer patients show higher plasma EGFL6 levels and higher tumor EGFL6 mRNA expression [10]. The evidence for the association between plasma EGFL6 and the clinicopathological features in oral cancer patients suggests a potential application for monitoring tumor behavior [10].

In conclusion, our study findings demonstrated that cytoplasm EGFL6 is specifically expressed in NSCLC and that increased expression is associated with poor clinical outcome. The results support the suggestion that cytoplasm EGFL6 can serve as a valuable target for the prediction of tumor malignancy and that it has therapeutic potential, although our findings need to be confirmed by further studies. Additional molecular studies are also needed to provide a more in-depth picture regarding the function of cytoplasm EGFL6 in NSCLC.

Competing interests

The authors declare that they have no competing interests.

Funding:

There is no funding or grant supports this work.

Author contributions statement:

Conception and design: Liu TC, Yeh KT; acquisition of data: Chang CC, Hsu HT, Yeh CM, Lee CH, Chen YL; analysis and interpretation of data: Chang CC, Hsu HT; drafting of the manuscript: Chang CC, Sung WW; critical revision of the manuscript: Liu TC, Yeh KT; statistical analysis: Sung WW; supervision: Liu TC, Yeh KT

Data sharing statement

There was no additional unpublished data.

Reference

1. Sung WW, Wang YC, Lin PL, et al. IL-10 promotes tumor aggressiveness via upregulation of CIP2A transcription in lung adenocarcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research 2013;**19**(15):4092-103.
2. Yang SC, Lai WW, Lin CC, et al. Cost-effectiveness of implementing computed tomography screening for lung cancer in Taiwan. Lung Cancer 2017;**108**:183-91.
3. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA: a cancer journal for clinicians 2017;**67**(1):7-30.
4. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;**144**(5):646-74.
5. Balkwill F. Cancer and the chemokine network. Nature reviews Cancer 2004;**4**(7):540-50.
6. Sung WW, Wang YC, Cheng YW, et al. A polymorphic -844T/C in FasL promoter predicts survival and relapse in non-small cell lung cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 2011;**17**(18):5991-9.
7. Davis CG. The many faces of epidermal growth factor repeats. The New biologist 1990;**2**(5):410-9.
8. Yeung G, Mulero JJ, Berntsen RP, et al. Cloning of a novel epidermal growth factor repeat containing gene EGFL6: expressed in tumor and fetal tissues. Genomics 1999;**62**(2):304-7.
9. Singh AB, Harris RC. Autocrine, paracrine and juxtacrine signaling by EGFR ligands. Cellular signalling 2005;**17**(10):1183-93.
10. Chuang CY, Chen MK, Hsieh MJ, et al. High Level of Plasma EGFL6 Is Associated with Clinicopathological Characteristics in Patients with Oral Squamous Cell Carcinoma. International journal of medical sciences 2017;**14**(5):419-24.
11. Bai S, Ingram P, Chen YC, et al. EGFL6 Regulates the Asymmetric Division, Maintenance, and Metastasis of ALDH+ Ovarian Cancer Cells. Cancer research 2016;**76**(21):6396-409.
12. Chim SM, Qin A, Tickner J, et al. EGFL6 promotes endothelial cell migration and angiogenesis through the activation of extracellular signal-regulated kinase. The Journal of biological chemistry 2011;**286**(25):22035-46.
13. Chim SM, Tickner J, Chow ST, et al. Angiogenic factors in bone local environment. Cytokine & growth factor reviews 2013;**24**(3):297-310.
14. Wang X, Yuan W, Chai R, et al. Egfl6 is involved in zebrafish notochord

development. Fish physiology and biochemistry 2015;**41**(4):961-9.
15. Hwang JC, Sung WW, Tu HP, et al. The Overexpression of FEN1 and RAD54B May Act as Independent Prognostic Factors of Lung Adenocarcinoma. PLoS One 2015;**10**(10):e0139435.
16. Sung WW, Lin YM, Wu PR, et al. High nuclear/cytoplasmic ratio of Cdk1 expression predicts poor prognosis in colorectal cancer patients. BMC cancer 2014;**14**:951.
17. Buckanovich RJ, Sasaroli D, O'Brien-Jenkins A, et al. Tumor vascular proteins as biomarkers in ovarian cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2007;**25**(7):852-61.
18. Wang X, Yuan W, Qi J, et al. The somite-secreted factor Maeg promotes zebrafish embryonic angiogenesis. Oncotarget 2016;**7**(47):77749-63.

Legends

Figure 1. Representative immunostaining of EGFL6 in tissue arrays of NSCLC specimens. EGFL6 expression levels were (A) 0; (B) 1+; (C) 2+.

Figure 2. Kaplan-Meier actuarial analysis of EGFL6 expression with respect to overall survival of patients of (A) all patients, and (B) patients younger than 69 years old.

Table 1. Relationships of EGFL6 expression with clinical parameters in 150 NSCLC patients.

	Cytoplasmic staining of EGFL6			
	low(0,1+)	high(2+)	Total	*p* value
Gender				
F	24(34.3)	46(65.7)	70	0.817
M	26(32.5)	54(67.5)	80	
Grade				
Well	8(36.4)	14(63.6)	22	0.744
Moderate, poor	42(32.8)	86(67.2)	128	
Age				
≦63	27(35.5)	49(64.5)	76	0.564
>63	23(31.1)	51(68.9)	74	
T status				
T1	19(35.2)	35(64.8)	54	0.718
T2,T3,T4	31(32.3)	65(67.7)	96	
Lymph Node Metastasis				
No	26(29.9)	61(70.1)	87	0.292
Yes	24(38.1)	39(61.9)	63	
Stage				
I	19(30.6)	43(69.4)	62	0.558
II,III,IV	31(35.2)	57(64.8)	88	

Table 2. Influence of various parameters on overall survival in NSCLC patients.

	Univariate			Multivariate		
Variable	Hazard Ratio	95% CI	*p*	Hazard Ratio	95% CI	*p*
Expression of EGFL6						
low	1.000			1.000		
high	1.519	0.980-2.355	0.061	1.515	0.975-2.354	0.064
Gender						
female	1.000			1.000		
male	2.184	1.450-3.290	<0.001	1.802	1.180-2.753	0.006
Age						
≦63	1.000			1.000		
>63	1.808	1.214-2.691	0.004	1.997	1.303-3.062	0.002
Stage						
I	1.000			1.000		
II,III,IV	1.871	1.232-2.840	0.003	2.241	1.443-3.481	<0.001

Table 3. Influence of EGFL6 on overall survival in NSCLC patients according to the age.

		Multivariate		
Sub-group	Case number	Hazard Ratio of EGFL6 Expression[1]	95% CI	*p*
Divide via medium age				
≦63	76	2.118	1.082-4.145	0.029
>63	74	1.184	0.661-2.122	0.570
Divide via grouped-age				
≦59	61	2.894	1.245-6.726	0.014
≦69	106	2.104	1.184-3.739	0.011
All	150	1.515	0.975-2.354	0.064

[1]Expression of EGFL6: high vs low

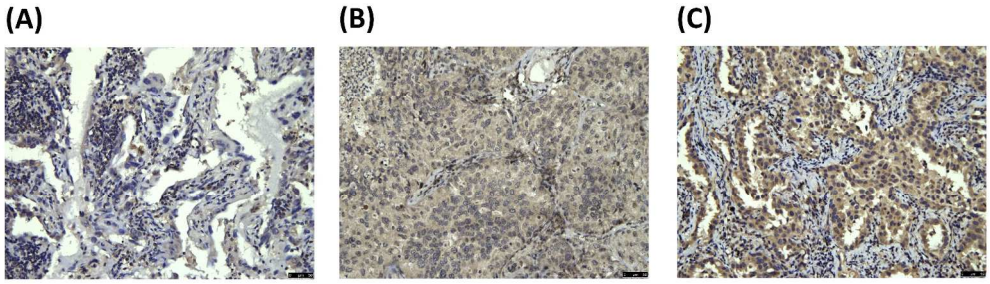


Figure 1. Representative immunostaining of EGFL6 in tissue arrays of NSCLC specimens. EGFL6 expression levels were (A) 0; (B) 1+; (C) 2+.

263x77mm (300 x 300 DPI)

STROBE Statement

Checklist of items that should be included in reports of observational studies

Section/Topic	Item No	Recommendation	Reported on Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	3	State specific objectives, including any prespecified hypotheses	5-6
Methods			
Study design	4	Present key elements of study design early in the paper	7-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-8
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	7-8
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	n/a
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8
Bias	9	Describe any efforts to address potential sources of bias	n/a
Study size	10	Explain how the study size was arrived at	7-8
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	7-8
		(*c*) Explain how missing data were addressed	n/a
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	n/a
		(*e*) Describe any sensitivity analyses	n/a

Section/Topic	Item No	Recommendation	Reported on Page No
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	n/a
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9-10
		(b) Indicate number of participants with missing data for each variable of interest	9-10
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	n/a
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	n/a
		Cross-sectional study—Report numbers of outcome events or summary measures	9-10
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9-10
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12
Generalisability	21	Discuss the generalisability (external validity) of the study results	11-12
Other Information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

**Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.*

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Validation of EGFL6 expression as a prognostic marker in lung adenocarcinoma patients: a retrospective study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-021385.R1
Article Type:	Research
Date Submitted by the Author:	28-Feb-2018
Complete List of Authors:	Chang, Chun-Chi ; Changhua Christian Hospital, Division of Chest Medicine, Department of Internal Medicine Sung, Wen-Wei; Chung Shan Medical University, School of Medicine Hsu, Hui-Ting ; Changhua Christian Hospital, Department of Pathology Yeh, Chung-Min ; Changhua Christian Hospital, Department of Pathology Lee, Chien-Hsun ; Changhua Christian Hospital, Department of Pathology Chen, Ya-Ling ; Changhua Christian Hospital, Department of Nursing Liu, Ta-Chih ; Kaohsiung Medical University, Graduate Institute of Clinical Medicine Yeh, Kun-Tu; Changhua Christian Hospital, Department of Pathology
Primary Subject Heading:	Pathology
Secondary Subject Heading:	Oncology, Pathology
Keywords:	EGF like domain multiple 6, EGFL6, prognosis, non-small cell lung cancer, overall survival, lung adenocarcinoma;

Validation of EGFL6 expression as a prognostic marker in lung adenocarcinoma patients: a retrospective study

Chun-Chi Chang[1,2], Wen-Wei Sung[3,4,5], Hui-Ting Hsu[6], Chung-Min Yeh[6,7], Chien-Hsun Lee[6], Ya-Ling Chen[8], Ta-Chih Liu[1,9,*], Kun-Tu Yeh[3,6,*]

[1]Graduate Institute of Clinical Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

[2]Division of Chest Medicine, Department of Internal Medicine, Changhua Christian Hospital, Changhua, Taiwan

[3]School of Medicine, Chung Shan Medical University, Taichung, Taiwan

[4]Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan

[5]Department of Urology, Chung Shan Medical University Hospital, Taichung, Taiwan

[6]Department of Pathology, Changhua Christian Hospital, Changhua, Taiwan

[7]Department of Medical Technology, Jen-Teh Junior College of Medicine, Nursing and Management, Miaoli, Taiwan

[8]Department of Nursing, Changhua Christian Hospital, Changhua, Taiwan

[9]Division of Hematology and Oncology, Department of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan

* Ta-Chih Liu and Kun-Tu Yeh contributed equally in this work.

Corresponding author and request for reprints:

Kun-Tu Yeh (TEL: 886-4-7238595 ext. 4830, E-mail: 10159@cch.org.tw)

Department of Pathology, Changhua Christian Hospital, Changhua, Taiwan

Running head: EGFL6 predicts poor prognosis in lung cancer

Key words: EGF like domain multiple 6; EGFL6; prognosis; non-small cell lung cancer; lung adenocarcinoma; overall survival

Abstract

Objective

Lung adenocarcinoma belongs to non-small cell lung cancer, a common cancer in both genders, has a poor clinical outcome. Our aim was to evaluate the role of EGF-like domain multiple 6 (EGFL6) and its prognostic significance in lung adenocarcinoma.

Methods

EGFL6 expression was studied by immunohistochemical staining of specimens from 150 patients with lung adenocarcinoma. The correlation between clinicopathological features and EGFL6 expression was quantitatively analyzed. We used Kaplan-Meier analysis and Cox proportional hazard models to examine the prognostic value of EGFL6 in terms of overall survival.

Results

No significant correlation was found between EGFL6 expression and clinical parameters. However, patients with high EGFL6 expression levels showed a tendency toward poor prognosis, with borderline statistical significance. Grouping the patients according to a medium age value revealed a significant association between high EGFL6 expression and poor clinical outcome in young patients. This finding was further confirmed by grouping the patients into three groups according to age. The hazard ratio in patients with high EGFL6 expression was higher in younger patients than in older patients.

Conclusion

High EGFL6 expression may serve as a marker for poor prognosis of lung adenocarcinoma, especially in younger patients.

Strengths and limitations of this study

1. This is a retrospective study using specimens from 150 patients with lung adenocarcinoma.
2. Overall survival but not cancer specific survival was used in this study.
3. This study did not explore the clinical diversity of post-operative chemotherapy and radiotherapy.
4. Considering limited sample size, further study is necessary for clinical application.
5. No information about detail molecular diversity was provided in this analysis.

Introduction

Lung adenocarcinoma belongs to non-small cell lung cancer (NSCLC) is a major public health problem worldwide and NSCLC is a leading cause of cancer death in Taiwan [1 2]. Whereas other types of cancer have shown steady increases in survival in recent years, NSCLC continues to have a poor clinical outcome, with a 5-year survival of only 18% [3]. Early detection of NSCLC might improve the clinical outcome; however, no suitable screening tools are available that are both cost effective and efficient. NSCLC screening via low dose CT scans can provide early detection of lung lesions, such as ground glass opacity lesions [2]. However, it does not discriminate benign lesions that may require no further intervention or surgical intervention. Therefore, the identification of specific biomarkers that indicate the malignant potential of NSCLC lesions would help in clinical decision making for cancer follow up and the timing of surgical intervention.

The malignant potential of a tumor with metastatic behavior is determined by complex processes, including tumor cell migration, invasion, and angiogenesis to the target site [1 4-6]. One group of proteins implicated in tumor malignancy is the epidermal growth factor (EGF) repeat superfamily, whose members have a conserved motif of cysteines and glycines positioned in a domain of 30 to 40 residues [7 8]. These EGF-like proteins are characterized by their multiple EGF repeats [9] and are secreted as cell surface molecules. The binding of EGF-like proteins to their receptors promotes tumor malignancy [8 10-12].

One member of this family, EGF-like domain 6 (EGFL6), is a secreted protein with involvement in tissue development, promotion of tumor cell migration, and angiogenesis [8 10-14]. A role for EGFL6 in promoting tumor malignancy is indicated in

several types of cancer; for example, oral cancer patients show high plasma levels of EGFL6, and the plasma EGFL6 level is higher in patients with advanced stage disease than in patients with early stage disease [10]. In ovarian cancer, EGFL6 regulates cell migration and asymmetric division via the SHP2 oncoprotein, with concomitant activation of ERK [11].

Accumulating evidence indicates crucial roles for EGFL6 in promoting tumor malignancy. However, an association between EGFL6 expression and the prognosis of NSCLC remains to be established. Since there are types of pathology subgroups in NSCLC with different tumor behavior, patients with lung adenocarcinoma were included for investigation. The aim of the present study was to evaluate the expression of EGFL6 and its clinical significance in lung adenocarcinoma.

Materials and Methods

Patients

Our study examined 150 tumor samples from patients with lung adenocarcinoma. Cancers were staged according to the AJCC Cancer Staging Manual (7th edition). The clinicopathological features assessed in this study included histological type, differentiation, lymph node metastasis, TNM stage, and tumor size. Histological diagnosis was confirmed by two pathologists, as described previously [15]. The study was approved by the Institutional Review Board and the Ethics Committee of the Changhua Christian Hospital, Changhua, Taiwan (CCH IRB 170511).

Immunohistochemical Staining and Evaluation of Cytoplasmic EGFL6

Immunohistochemistry (IHC) staining was performed at the Department of Surgical Pathology, Changhua Christian Hospital, as previously described, [15] [16] using anti-human cytoplasmic EGFL6 antibody (EGFL6 antibody, 1:100 dilution; Abcam, ab140079). Immunoreactivity scores were analyzed by three pathologists, using a previously described scoring protocol [16]. The pathologists were blinded to the prognostic data of the study. A final agreement was obtained for each score by having all three evaluators view the specimens simultaneously through a multi-headed microscope (Olympus BX51 10-headed microscope).

Patient and Public Involvement

This study analyzed cancer tissues from de-linked database. Therefore, we did not informe or disseminate the patients about the research question, outcome measures,

and results. Patients did not involve in the study including design, recruitment, and conduct of the study. There was no patient adviser for contributorship statement.

Statistical Analysis

The χ^2 test was applied for continuous or discrete data analysis. The associations between cytoplasmic EGFL6 expression and patient survival were estimated using univariate analysis and the Kaplan–Meier method and further assessed using the log-rank test [1]. Potential confounders were adjusted using Cox regression models of multivariate analysis, with the cytoplasmic EGFL6 expressions fitted as indicator variables. All statistical analyses were conducted using SPSS statistical software (version 15.0; SPSS, Inc., Chicago, IL). All statistical tests were 2-sided, and the values of $p<0.05$ were considered statistically significant.

Results

Cytoplasmic EGFL6 is Expressed in the Majority of Lung Adenocarcinoma Specimens

We verified the role of EGFL6 in the clinical outcome of lung adenocarcinoma by recruiting 150 patients. EGFL6 expression was detected with IHC staining, as shown in Figure 1. Of the 150 patients, only 6 patients (4%) showed no detectable cytoplasmic EGFL6 expression. Table 1 shows the relationships of the cytoplasmic EGFL6 expression according to the clinicopathological characteristics. The mean patient age was 62.1±11.6 years (mean±SD). The cytoplasmic EGFL6 expression was not significantly associated with the clinicopathological characteristics of gender, grade, age, or TNM stage.

The Prognostic Role of Cytoplasmic EGFL6 Expression in Lung Adenocarcinoma Patients

We further evaluated the prognostic role of cytoplasmic EGFL6 expression in lung adenocarcinoma patients. Overall survival data were collected, and no data were missing from any of the 150 patients. The mean and median follow-up times after surgery were 5.2 and 3.2 years (range from 0.1 to 11.0 years), respectively. The 5-year survival rate was 42.1%. During the survey, 99 (66.0%) patients died. In the univariate analysis, patients with advanced stage disease, age >63 years old, and male gender had significantly poorer clinical outcomes (Table 2). These factors were still significantly associated with poor prognosis in the multivariate analysis (HR=2.241, 95% CI=1.443–3.481, p<0.001 for stage; HR=1.997, 95% CI=1.303–3.062, p=0.002 for age; HR=1.802, 95% CI=1.180–2.753, p=0.006, Table 1). A prognostic role for cytoplasmic EGFL6 in lung adenocarcinoma was suggested by the finding that

patients with high cytoplasmic EGFL6 expression had lower 5-year survival rates and shorter medium survival when compared with patients with low cytoplasmic EGFL6 expression (52.0% vs. 37.0% for 5-year survival; 5.7 years vs. 2.5 years for medium survival, Figure 2A). The univariate and multivariate analysis revealed a borderline statistical significance for cytoplasmic EGFL6 (HR=1.519, 95% CI=0.980–2.355, p=0.061 for univariate analysis; HR=1.515, 95% CI=0.975–2.354, p=0.064 for multivariate analysis, Table 2).

Significant Prognostic Role of Cytoplasmic EGFL6 Expression in Young Lung Adenocarcinoma Patients

We examined the potential prognostic role of cytoplasmic EGFL6 in lung adenocarcinoma patients by analyzing their clinical outcomes according to their clinicopathological characteristics. We identified a significant association of cytoplasmic EGFL6 in patients with younger age. As shown in Figure 2B, patients younger than 69 years of age who had high cytoplasmic EGFL6 expression also had a lower 5-year survival rate and shorter median survival times when compared with patients with low cytoplasmic EGFL6 expression (65.7% vs 40.9% for 5-year survival; 8.4 years vs 2.8 years for median survival, Figure 2B). We confirmed this finding using different age cutoff points: the use of the median age as a cutoff point resulted in a significantly poorer prognosis for patients with high EGFL6 (HR: 2.118, 95% CI: 1.082–4.145, p=0.029, Table 3). The HR was also increased in patients of younger age (HR: 2.894, 95% CI: 1.245–6.726, p=0.014 for age $\leqq$ 59; HR: 2.104, 95% CI: 1.184–3.739, p=0.011 for age $\leq$ 69, Table 3).

Discussion

In this study, we identified a prognostic role for cytoplasmic EGFL6 in lung adenocarcinoma, especially in patients of younger age. This is the first study to provide clinical evidence of EGFL6 expression in lung adenocarcinoma. No association was noted between EGFL6 expression and clinical parameters, but the significantly poor clinical outcome of patients with high EGFL6 expression supports the findings of previous reports regarding EGFL6 expression in other types of cancer [10 12 17].

The role of EGFLF6 in cell division and tissue development was first identified in a non-tumor model [12-14 18]. In a bone remodeling model, EGFL6 induced angiogenesis via a paracrine mechanism that promoted angiogenesis and migration of endothelial cells [12]. Inhibition of phosphorylated ERK in this model decreased the ability of the cells to migrate [12]. In a zebrafish model, EGFL6 promoted angiogenesis via a mechanism that depended on the RGD domain and on activation of the Akt and Erk pathways [18]. Loss of EGFL6 decreased the numbers of endothelial cells and vessels, suggesting that EGFL6 acts as a positive regulator of functional vessel formation [18].

Increasing evidence supports a role for EGFL6 in regulating tumor malignancy and shows the potential for EGFL6 to serve as a prognostic marker and therapeutic target.

In ovarian cancer, EGFL6 expression is associated with poor clinical outcome, which is further explained by its role in promoting cancer cell proliferation and asymmetric division [11]. A model using $ALDH^+$ ovarian cancer cells showed that EGFL6 signaling involves integrin, SHP2, and ERK [11]. The results of molecular analysis of ovarian tumor vascular cells obtained with immunohistochemistry-guided laser-capture microdissection and genome-wide transcriptional profiling also supported this result [17]. Oral cancer patients also show high plasma EGFL6 levels and high tumor EGFL6 mRNA expression [10]. The apparent association between plasma EGFL6 and the clinicopathological features in oral cancer patients suggests a potential application for EGFL6 in monitoring tumor behavior [10].

In conclusion, our study findings demonstrated that cytoplasmic EGFL6 is specifically expressed in lung adenocarcinoma, and this increased expression is associated with poor clinical outcome. These results support the suggestion that cytoplasmic EGFL6 can serve as a valuable marker for the prediction of tumor malignancy and that it has therapeutic potential, although our findings need to be confirmed by further studies. Additional molecular studies are also needed to provide a more in-depth picture regarding the function of cytoplasmic EGFL6 in lung adenocarcinoma.

Competing interests

The authors declare that they have no competing interests.

Funding:

There is no funding or grant supports this work.

Author contributions statement:

Conception and design: Liu TC, Yeh KT; acquisition of data: Chang CC, Hsu HT, Yeh CM, Lee CH, Chen YL; analysis and interpretation of data: Chang CC, Hsu HT; drafting of the manuscript: Chang CC, Sung WW; critical revision of the manuscript: Liu TC, Yeh KT; statistical analysis: Sung WW; supervision: Liu TC, Yeh KT

Data sharing statement

There was no additional unpublished data.

Reference

1. Sung WW, Wang YC, Lin PL, et al. IL-10 promotes tumor aggressiveness via upregulation of CIP2A transcription in lung adenocarcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research 2013;**19**(15):4092-103.
2. Yang SC, Lai WW, Lin CC, et al. Cost-effectiveness of implementing computed tomography screening for lung cancer in Taiwan. Lung Cancer 2017;**108**:183-91.
3. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA: a cancer journal for clinicians 2017;**67**(1):7-30.
4. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;**144**(5):646-74.
5. Balkwill F. Cancer and the chemokine network. Nature reviews Cancer 2004;**4**(7):540-50.
6. Sung WW, Wang YC, Cheng YW, et al. A polymorphic -844T/C in FasL promoter predicts survival and relapse in non-small cell lung cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 2011;**17**(18):5991-9.
7. Davis CG. The many faces of epidermal growth factor repeats. The New biologist 1990;**2**(5):410-9.
8. Yeung G, Mulero JJ, Berntsen RP, et al. Cloning of a novel epidermal growth factor repeat containing gene EGFL6: expressed in tumor and fetal tissues. Genomics 1999;**62**(2):304-7.
9. Singh AB, Harris RC. Autocrine, paracrine and juxtacrine signaling by EGFR ligands. Cellular signalling 2005;**17**(10):1183-93.
10. Chuang CY, Chen MK, Hsieh MJ, et al. High Level of Plasma EGFL6 Is Associated with Clinicopathological Characteristics in Patients with Oral Squamous Cell Carcinoma. International journal of medical sciences 2017;**14**(5):419-24.
11. Bai S, Ingram P, Chen YC, et al. EGFL6 Regulates the Asymmetric Division, Maintenance, and Metastasis of ALDH+ Ovarian Cancer Cells. Cancer research 2016;**76**(21):6396-409.
12. Chim SM, Qin A, Tickner J, et al. EGFL6 promotes endothelial cell migration and angiogenesis through the activation of extracellular signal-regulated kinase. The Journal of biological chemistry 2011;**286**(25):22035-46.
13. Chim SM, Tickner J, Chow ST, et al. Angiogenic factors in bone local environment. Cytokine & growth factor reviews 2013;**24**(3):297-310.

14. Wang X, Yuan W, Chai R, et al. Egfl6 is involved in zebrafish notochord development. Fish physiology and biochemistry 2015;**41**(4):961-9.
15. Hwang JC, Sung WW, Tu HP, et al. The Overexpression of FEN1 and RAD54B May Act as Independent Prognostic Factors of Lung Adenocarcinoma. PLoS One 2015;**10**(10):e0139435.
16. Sung WW, Lin YM, Wu PR, et al. High nuclear/cytoplasmic ratio of Cdk1 expression predicts poor prognosis in colorectal cancer patients. BMC cancer 2014;**14**:951.
17. Buckanovich RJ, Sasaroli D, O'Brien-Jenkins A, et al. Tumor vascular proteins as biomarkers in ovarian cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2007;**25**(7):852-61.
18. Wang X, Yuan W, Qi J, et al. The somite-secreted factor Maeg promotes zebrafish embryonic angiogenesis. Oncotarget 2016;**7**(47):77749-63.

Legends

Figure 1. Representative immunostaining of EGFL6 in tissue arrays of lung adenocarcinoma specimens. EGFL6 expression levels were (A) 0; (B) 1+; (C) 2+.

Figure 2. Kaplan-Meier actuarial analysis of EGFL6 expression with respect to overall survival of patients of (A) all patients, and (B) patients younger than 69 years of age.

Table 1. Relationships of EGFL6 expression with clinical parameters in 150 lung adenocarcinoma patients.

	Cytoplasmic staining of EGFL6			
	low(0,1+)	high(2+)	Total	*p* value
Gender				
F	24(34.3)	46(65.7)	70	0.817
M	26(32.5)	54(67.5)	80	
Grade				
Well	8(36.4)	14(63.6)	22	0.744
Moderate, poor	42(32.8)	86(67.2)	128	
Age				
≦63	27(35.5)	49(64.5)	76	0.564
>63	23(31.1)	51(68.9)	74	
T status				
T1	19(35.2)	35(64.8)	54	0.718
T2,T3,T4	31(32.3)	65(67.7)	96	
Lymph Node Metastasis				
No	26(29.9)	61(70.1)	87	0.292
Yes	24(38.1)	39(61.9)	63	
Stage				
I	19(30.6)	43(69.4)	62	0.558
II,III,IV	31(35.2)	57(64.8)	88	

Table 2. Influence of various parameters on overall survival in lung adenocarcinoma patients.

	Univariate			Multivariate		
Variable	Hazard Ratio	95% CI	*p*	Hazard Ratio	95% CI	*p*
Expression of EGFL6						
low	1.000			1.000		
high	1.519	0.980-2.355	0.061	1.515	0.975-2.354	0.064
Gender						
female	1.000			1.000		
male	2.184	1.450-3.290	<0.001	1.802	1.180-2.753	0.006
Age						
≦63	1.000			1.000		
>63	1.808	1.214-2.691	0.004	1.997	1.303-3.062	0.002
Stage						
I	1.000			1.000		
II,III,IV	1.871	1.232-2.840	0.003	2.241	1.443-3.481	<0.001

Table 3. Influence of EGFL6 on overall survival in lung adenocarcinoma patients according to the age.

		Multivariate		
Sub-group	Case number	Hazard Ratio of EGFL6 Expression[1]	95% CI	*p*
Divide via medium age				
≦63	76	2.118	1.082-4.145	0.029
>63	74	1.184	0.661-2.122	0.570
Divide via grouped-age				
≦59	61	2.894	1.245-6.726	0.014
≦69	106	2.104	1.184-3.739	0.011
All	150	1.515	0.975-2.354	0.064

[1]Expression of EGFL6: high vs low

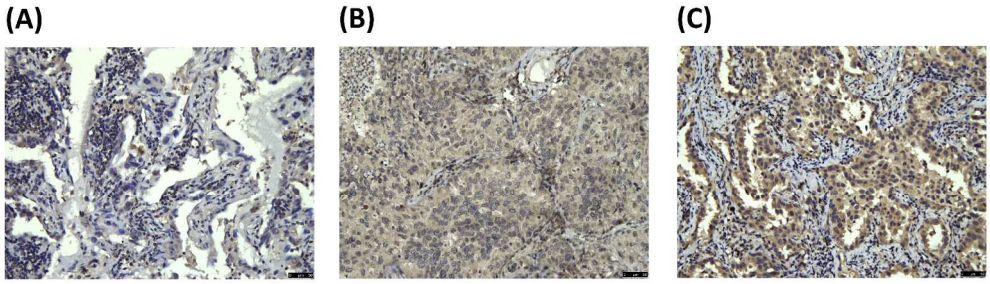


Figure 1. Representative immunostaining of EGFL6 in tissue arrays of NSCLC specimens. EGFL6 expression levels were (A) 0; (B) 1+; (C) 2+.

263x77mm (300 x 300 DPI)

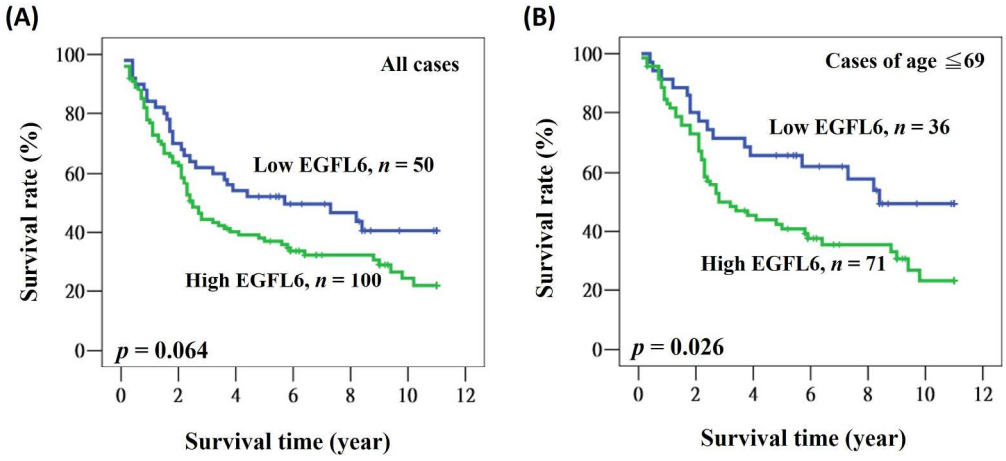


Figure 2. Kaplan-Meier actuarial analysis of EGFL6 expression with respect to overall survival of patients of (A) all patients, and (B) patients younger than 69 years old.

225x103mm (300 x 300 DPI)

STROBE Statement

Checklist of items that should be included in reports of observational studies

Section/Topic	Item No	Recommendation	Reported on Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	3	State specific objectives, including any prespecified hypotheses	5-6
Methods			
Study design	4	Present key elements of study design early in the paper	7-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-8
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	7-8
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	n/a
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8
Bias	9	Describe any efforts to address potential sources of bias	n/a
Study size	10	Explain how the study size was arrived at	7-8
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	7-8
		(*c*) Explain how missing data were addressed	n/a
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	n/a
		(*e*) Describe any sensitivity analyses	n/a

Section/Topic	Item No	Recommendation	Reported on Page No
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	n/a
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9-10
		(b) Indicate number of participants with missing data for each variable of interest	9-10
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	n/a
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	n/a
		Cross-sectional study—Report numbers of outcome events or summary measures	9-10
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9-10
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12
Generalisability	21	Discuss the generalisability (external validity) of the study results	11-12
Other Information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

**Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.*

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Validation of EGFL6 expression as a prognostic marker in lung adenocarcinoma patients in Taiwan: a retrospective study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-021385.R2
Article Type:	Research
Date Submitted by the Author:	12-Apr-2018
Complete List of Authors:	Chang, Chun-Chi ; Changhua Christian Hospital, Division of Chest Medicine, Department of Internal Medicine Sung, Wen-Wei; Chung Shan Medical University, School of Medicine Hsu, Hui-Ting ; Changhua Christian Hospital, Department of Pathology Yeh, Chung-Min ; Changhua Christian Hospital, Department of Pathology Lee, Chien-Hsun ; Changhua Christian Hospital, Department of Pathology Chen, Ya-Ling ; Changhua Christian Hospital, Department of Nursing Liu, Ta-Chih ; Kaohsiung Medical University, Graduate Institute of Clinical Medicine Yeh, Kun-Tu; Changhua Christian Hospital, Department of Pathology
Primary Subject Heading:	Pathology
Secondary Subject Heading:	Oncology, Pathology
Keywords:	EGF like domain multiple 6, EGFL6, prognosis, non-small cell lung cancer, overall survival, lung adenocarcinoma;

Validation of EGFL6 expression as a prognostic marker in lung adenocarcinoma patients in Taiwan: a retrospective study

Chun-Chi Chang[1,2], Wen-Wei Sung[3,4,5], Hui-Ting Hsu[6], Chung-Min Yeh[6,7], Chien-Hsun Lee[6], Ya-Ling Chen[8], Ta-Chih Liu[1,9,*], Kun-Tu Yeh[3,6,*]

[1]Graduate Institute of Clinical Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

[2]Division of Chest Medicine, Department of Internal Medicine, Changhua Christian Hospital, Changhua, Taiwan

[3]School of Medicine, Chung Shan Medical University, Taichung, Taiwan

[4]Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan

[5]Department of Urology, Chung Shan Medical University Hospital, Taichung, Taiwan

[6]Department of Pathology, Changhua Christian Hospital, Changhua, Taiwan

[7]Department of Medical Technology, Jen-Teh Junior College of Medicine, Nursing and Management, Miaoli, Taiwan

[8]Department of Nursing, Changhua Christian Hospital, Changhua, Taiwan

[9]Division of Hematology and Oncology, Department of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan

* Ta-Chih Liu and Kun-Tu Yeh contributed equally in this work.

Corresponding author and request for reprints:

Kun-Tu Yeh (TEL: 886-4-7238595 ext. 4830, E-mail: 10159@cch.org.tw)

Department of Pathology, Changhua Christian Hospital, Changhua, Taiwan

Running head: EGFL6 predicts poor prognosis in lung cancer

Key words: EGF like domain multiple 6; EGFL6; prognosis; non-small cell lung cancer; lung adenocarcinoma; overall survival

Abstract

Objective

Lung adenocarcinoma belongs to non-small cell lung cancer, a common cancer in both genders, has a poor clinical outcome. Our aim was to evaluate the role of EGF-like domain multiple 6 (EGFL6) and its prognostic significance in lung adenocarcinoma.

Methods

EGFL6 expression was studied by immunohistochemical staining of specimens from 150 patients with lung adenocarcinoma. The correlation between clinicopathological features and EGFL6 expression was quantitatively analyzed. We used Kaplan-Meier analysis and Cox proportional hazard models to examine the prognostic value of EGFL6 in terms of overall survival.

Results

No significant correlation was found between EGFL6 expression and clinical parameters. However, patients with high EGFL6 expression levels showed a tendency toward poor prognosis, with borderline statistical significance. Grouping the patients according to a medium age value revealed a significant association between high EGFL6 expression and poor clinical outcome in young patients. This finding was further confirmed by grouping the patients into three groups according to age. The hazard ratio in patients with high EGFL6 expression was higher in younger patients than in older patients.

Conclusion

High EGFL6 expression may serve as a marker for poor prognosis of lung adenocarcinoma, especially in younger patients.

Strengths and limitations of this study

1. This is a retrospective study using specimens from 150 patients with lung adenocarcinoma.
2. Overall survival but not cancer specific survival was used in this study.
3. This study did not explore the clinical diversity of post-operative chemotherapy and radiotherapy.
4. Considering limited sample size, further study is necessary for clinical application.
5. No information about detail molecular diversity was provided in this analysis.

Introduction

Lung adenocarcinoma belongs to non-small cell lung cancer (NSCLC) is a major public health problem worldwide and NSCLC is a leading cause of cancer death in Taiwan [1 2]. Whereas other types of cancer have shown steady increases in survival in recent years, NSCLC continues to have a poor clinical outcome, with a 5-year survival of only 18% [3]. Early detection of NSCLC might improve the clinical outcome; however, no suitable screening tools are available that are both cost effective and efficient. NSCLC screening via low dose CT scans can provide early detection of lung lesions, such as ground glass opacity lesions [2]. However, it does not discriminate benign lesions that may require no further intervention or surgical intervention. Therefore, the identification of specific biomarkers that indicate the malignant potential of NSCLC lesions would help in clinical decision making for cancer follow up and the timing of surgical intervention.

The malignant potential of a tumor with metastatic behavior is determined by complex processes, including tumor cell migration, invasion, and angiogenesis to the target site [1 4-6]. One group of proteins implicated in tumor malignancy is the epidermal growth factor (EGF) repeat superfamily, whose members have a conserved motif of cysteines and glycines positioned in a domain of 30 to 40 residues [7 8]. These EGF-like proteins are characterized by their multiple EGF repeats [9] and are secreted as cell surface molecules. The binding of EGF-like proteins to their receptors promotes tumor malignancy [8 10-12].

One member of this family, EGF-like domain 6 (EGFL6), is a secreted protein with involvement in tissue development, promotion of tumor cell migration, and angiogenesis [8 10-14]. A role for EGFL6 in promoting tumor malignancy is indicated in

several types of cancer; for example, oral cancer patients show high plasma levels of EGFL6, and the plasma EGFL6 level is higher in patients with advanced stage disease than in patients with early stage disease [10]. In ovarian cancer, EGFL6 regulates cell migration and asymmetric division via the SHP2 oncoprotein, with concomitant activation of ERK [11].

Accumulating evidence indicates crucial roles for EGFL6 in promoting tumor malignancy. However, an association between EGFL6 expression and the prognosis of NSCLC remains to be established. Since there are types of pathology subgroups in NSCLC with different tumor behavior, patients with lung adenocarcinoma were included for investigation. The aim of the present study was to evaluate the expression of EGFL6 and its clinical significance in lung adenocarcinoma.

Materials and Methods

Patients

Our study examined 150 tumor samples from patients with lung adenocarcinoma. Cancers were staged according to the AJCC Cancer Staging Manual (7th edition). The clinicopathological features assessed in this study included histological type, differentiation, lymph node metastasis, TNM stage, and tumor size. Histological diagnosis was confirmed by two pathologists, as described previously [15]. Patients with primary lung adenocarcinoma and tissue available in bio-bank were included in this survey. Those with missing data or tissue loss during the staining procedure were excluded from this study to avoid bios from missing data. The study was approved by the Institutional Review Board and the Ethics Committee of the Changhua Christian Hospital, Changhua, Taiwan (CCH IRB 170511).

Immunohistochemical Staining and Evaluation of Cytoplasmic EGFL6

Immunohistochemistry (IHC) staining was performed at the Department of Surgical Pathology, Changhua Christian Hospital, as previously described, [15] [16] using anti-human cytoplasmic EGFL6 antibody (EGFL6 antibody, 1:100 dilution; Abcam, ab140079). Immunoreactivity scores were analyzed by three pathologists, using a previously described scoring protocol [16]. Liver tissue was reported to have EGFL6 expression and served as positive control. IHC assay with a primary antibody in tandem with a specimen that is not exposed to the primary antibody served as negative control (Supplementary Figure 1). The pathologists were blinded to the prognostic data of the study. A final agreement was obtained for each score by having all three evaluators view the specimens simultaneously through a multi-headed microscope (Olympus BX51 10-headed microscope). Immunoreactivity scores were defined as the

cell staining intensity (0-3) multiplied by the percentage of stained cells (0%-100%), leading to scores from 0 to 300 [15 16].

Patient and Public Involvement

This study analyzed cancer tissues from de-linked database. Therefore, we did not informe or disseminate the patients about the research question, outcome measures, and results. Patients did not involve in the study including design, recruitment, and conduct of the study. There was no patient adviser for contributorship statement.

Statistical Analysis

The χ^2 test was applied for continuous or discrete data analysis. The associations between cytoplasmic EGFL6 expression and patient survival were estimated using univariate analysis and the Kaplan–Meier method and further assessed using the log-rank test [1]. Potential confounders including age, gender, and stage were adjusted using Cox regression models of multivariate analysis, with the cytoplasmic EGFL6 expressions fitted as indicator variables. All statistical analyses were conducted using SPSS statistical software (version 15.0; SPSS, Inc., Chicago, IL). All statistical tests were 2-sided, and the values of $p<0.05$ were considered statistically significant.

Results

Cytoplasmic EGFL6 is Expressed in the Majority of Lung Adenocarcinoma Specimens

We verified the role of EGFL6 in the clinical outcome of lung adenocarcinoma by recruiting 150 patients. EGFL6 expression was detected with IHC staining, as shown in Figure 1 (Figure 1A to 1C). Of the 150 patients, only 6 patients (4%) showed no detectable cytoplasmic EGFL6 expression. Table 1 shows the relationships of the cytoplasmic EGFL6 expression according to the clinicopathological characteristics. The mean patient age was 62.1±11.6 years (mean±SD). The cytoplasmic EGFL6 expression was not significantly associated with the clinicopathological characteristics of gender, grade, age, or TNM stage.

The Prognostic Role of Cytoplasmic EGFL6 Expression in Lung Adenocarcinoma Patients

We further evaluated the prognostic role of cytoplasmic EGFL6 expression in lung adenocarcinoma patients. Overall survival data were collected, and no data were missing from any of the 150 patients. The mean and median follow-up times after surgery were 5.2 and 3.2 years (range from 0.1 to 11.0 years), respectively. The 5-year survival rate was 42.1%. During the survey, 99 (66.0%) patients died. In the univariate analysis, patients with advanced stage disease, age >63 years old, and male gender had significantly poorer clinical outcomes (Table 2). These factors were still significantly associated with poor prognosis in the multivariate analysis (HR=2.241, 95% CI=1.443–3.481, p<0.001 for stage; HR=1.997, 95% CI=1.303–3.062, p=0.002 for age; HR=1.802, 95% CI=1.180–2.753, p=0.006, Table 1). A prognostic role for cytoplasmic EGFL6 in lung adenocarcinoma was suggested by the finding that

patients with high cytoplasmic EGFL6 expression had lower 5-year survival rates and shorter medium survival when compared with patients with low cytoplasmic EGFL6 expression (52.0% vs. 37.0% for 5-year survival; 5.7 years vs. 2.5 years for medium survival, Figure 2A). The univariate and multivariate analysis revealed a borderline statistical significance for cytoplasmic EGFL6 (HR=1.519, 95% CI=0.980–2.355, p=0.061 for univariate analysis; HR=1.515, 95% CI=0.975–2.354, p=0.064 for multivariate analysis, Table 2).

Significant Prognostic Role of Cytoplasmic EGFL6 Expression in Young Lung Adenocarcinoma Patients

We examined the potential prognostic role of cytoplasmic EGFL6 in lung adenocarcinoma patients by analyzing their clinical outcomes according to their clinicopathological characteristics. We identified a significant association of cytoplasmic EGFL6 in patients with younger age. As shown in Figure 2B, patients younger than 69 years of age who had high cytoplasmic EGFL6 expression also had a lower 5-year survival rate and shorter median survival times when compared with patients with low cytoplasmic EGFL6 expression (65.7% vs 40.9% for 5-year survival; 8.4 years vs 2.8 years for median survival, Figure 2B). We confirmed this finding using different age cutoff points: the use of the median age as a cutoff point resulted in a significantly poorer prognosis for patients with high EGFL6 (HR: 2.118, 95% CI: 1.082–4.145, p=0.029, Table 3). The HR was also increased in patients of younger age (HR: 2.894, 95% CI: 1.245–6.726, p=0.014 for age $\leqq$ 59; HR: 2.104, 95% CI: 1.184–3.739, p=0.011 for age $\leq$ 69, Table 3).

Discussion

In this study, we identified a prognostic role for cytoplasmic EGFL6 in lung adenocarcinoma, especially in patients of younger age. This is the first study to provide clinical evidence of EGFL6 expression in lung adenocarcinoma. No association was noted between EGFL6 expression and clinical parameters, but the significantly poor clinical outcome of patients with high EGFL6 expression supports the findings of previous reports regarding EGFL6 expression in other types of cancer [10 12 17].

The role of EGFLF6 in cell division and tissue development was first identified in a non-tumor model [12-14 18]. In a bone remodeling model, EGFL6 induced angiogenesis via a paracrine mechanism that promoted angiogenesis and migration of endothelial cells [12]. Inhibition of phosphorylated ERK in this model decreased the ability of the cells to migrate [12]. In a zebrafish model, EGFL6 promoted angiogenesis via a mechanism that depended on the RGD domain and on activation of the Akt and Erk pathways [18]. Loss of EGFL6 decreased the numbers of endothelial cells and vessels, suggesting that EGFL6 acts as a positive regulator of functional vessel formation [18].

Increasing evidence supports a role for EGFL6 in regulating tumor malignancy and shows the potential for EGFL6 to serve as a prognostic marker and therapeutic target.

In ovarian cancer, EGFL6 expression is associated with poor clinical outcome, which is further explained by its role in promoting cancer cell proliferation and asymmetric division [11]. A model using $ALDH^+$ ovarian cancer cells showed that EGFL6 signaling involves integrin, SHP2, and ERK [11]. The results of molecular analysis of ovarian tumor vascular cells obtained with immunohistochemistry-guided laser-capture microdissection and genome-wide transcriptional profiling also supported this result [17]. Oral cancer patients also show high plasma EGFL6 levels and high tumor EGFL6 mRNA expression [10]. The apparent association between plasma EGFL6 and the clinicopathological features in oral cancer patients suggests a potential application for EGFL6 in monitoring tumor behavior [10].

There are some limitations of this study. The use of tissue arrays cannot represent the whole tumor and no duplicated array was investigated in this study. Limited sample size weakens the impact of our finding. Thus, more complete studies with large sample size are still needed in the future. Otherwise, only one clone of antibody was used. The results should be further validated with different antibody clones.

In conclusion, our study findings demonstrated that cytoplasmic EGFL6 is specifically expressed in lung adenocarcinoma, and this increased expression is associated with poor clinical outcome. These results support the suggestion that

cytoplasmic EGFL6 can serve as a valuable marker for the prediction of tumor malignancy and that it has therapeutic potential, although our findings need to be confirmed by further studies. Additional molecular studies are also needed to provide a more in-depth picture regarding the function of cytoplasmic EGFL6 in lung adenocarcinoma.

Competing interests

The authors declare that they have no competing interests.

Funding:

There is no funding or grant supports this work.

Author contributions statement:

Conception and design: Liu TC, Yeh KT; acquisition of data: Chang CC, Hsu HT, Yeh CM, Lee CH, Chen YL; analysis and interpretation of data: Chang CC, Hsu HT; drafting of the manuscript: Chang CC, Sung WW; critical revision of the manuscript: Liu TC, Yeh KT; statistical analysis: Sung WW; supervision: Liu TC, Yeh KT

Data sharing statement

There was no additional unpublished data.

Reference

1. Sung WW, Wang YC, Lin PL, et al. IL-10 promotes tumor aggressiveness via upregulation of CIP2A transcription in lung adenocarcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research 2013;**19**(15):4092-103.
2. Yang SC, Lai WW, Lin CC, et al. Cost-effectiveness of implementing computed tomography screening for lung cancer in Taiwan. Lung Cancer 2017;**108**:183-91.
3. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA: a cancer journal for clinicians 2017;**67**(1):7-30.
4. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;**144**(5):646-74.
5. Balkwill F. Cancer and the chemokine network. Nature reviews Cancer 2004;**4**(7):540-50.
6. Sung WW, Wang YC, Cheng YW, et al. A polymorphic -844T/C in FasL promoter predicts survival and relapse in non-small cell lung cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 2011;**17**(18):5991-9.
7. Davis CG. The many faces of epidermal growth factor repeats. The New biologist 1990;**2**(5):410-9.
8. Yeung G, Mulero JJ, Berntsen RP, et al. Cloning of a novel epidermal growth factor repeat containing gene EGFL6: expressed in tumor and fetal tissues. Genomics 1999;**62**(2):304-7.
9. Singh AB, Harris RC. Autocrine, paracrine and juxtacrine signaling by EGFR ligands. Cellular signalling 2005;**17**(10):1183-93.
10. Chuang CY, Chen MK, Hsieh MJ, et al. High Level of Plasma EGFL6 Is Associated with Clinicopathological Characteristics in Patients with Oral Squamous Cell Carcinoma. International journal of medical sciences 2017;**14**(5):419-24.
11. Bai S, Ingram P, Chen YC, et al. EGFL6 Regulates the Asymmetric Division, Maintenance, and Metastasis of ALDH+ Ovarian Cancer Cells. Cancer research 2016;**76**(21):6396-409.
12. Chim SM, Qin A, Tickner J, et al. EGFL6 promotes endothelial cell migration and angiogenesis through the activation of extracellular signal-regulated kinase. The Journal of biological chemistry 2011;**286**(25):22035-46.
13. Chim SM, Tickner J, Chow ST, et al. Angiogenic factors in bone local environment. Cytokine & growth factor reviews 2013;**24**(3):297-310.

14. Wang X, Yuan W, Chai R, et al. Egfl6 is involved in zebrafish notochord development. Fish physiology and biochemistry 2015;**41**(4):961-9.
15. Hwang JC, Sung WW, Tu HP, et al. The Overexpression of FEN1 and RAD54B May Act as Independent Prognostic Factors of Lung Adenocarcinoma. PLoS One 2015;**10**(10):e0139435.
16. Sung WW, Lin YM, Wu PR, et al. High nuclear/cytoplasmic ratio of Cdk1 expression predicts poor prognosis in colorectal cancer patients. BMC cancer 2014;**14**:951.
17. Buckanovich RJ, Sasaroli D, O'Brien-Jenkins A, et al. Tumor vascular proteins as biomarkers in ovarian cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2007;**25**(7):852-61.
18. Wang X, Yuan W, Qi J, et al. The somite-secreted factor Maeg promotes zebrafish embryonic angiogenesis. Oncotarget 2016;**7**(47):77749-63.

Legends

Figure 1. Representative immunostaining of EGFL6 in tissue arrays of lung adenocarcinoma specimens. EGFL6 expression levels were (A) 0; (B) 1+; (C) 2+.

Figure 2. Kaplan-Meier actuarial analysis of EGFL6 expression with respect to overall survival of patients of (A) all patients, and (B) patients younger than 69 years of age.

Supplementary Figure 1. Representative control immunostaining of EGFL6 in liver tissue. (A) positive and (B) negative control.

Table 1. Relationships of EGFL6 expression with clinical parameters in 150 lung adenocarcinoma patients.

	Cytoplasmic staining of EGFL6			
	low(0,1+)	high(2+)	Total	*p* value
Gender				
F	24(34.3)	46(65.7)	70	0.817
M	26(32.5)	54(67.5)	80	
Grade				
Well	8(36.4)	14(63.6)	22	0.744
Moderate, poor	42(32.8)	86(67.2)	128	
Age				
≦63	27(35.5)	49(64.5)	76	0.564
>63	23(31.1)	51(68.9)	74	
T status				
T1	19(35.2)	35(64.8)	54	0.718
T2,T3,T4	31(32.3)	65(67.7)	96	
Lymph Node Metastasis				
No	26(29.9)	61(70.1)	87	0.292
Yes	24(38.1)	39(61.9)	63	
Stage				
I	19(30.6)	43(69.4)	62	0.558
II,III,IV	31(35.2)	57(64.8)	88	

Table 2. Influence of various parameters on overall survival in lung adenocarcinoma patients.

	Univariate			Multivariate		
Variable	Hazard Ratio	95% CI	*p*	Hazard Ratio	95% CI	*p*
Expression of EGFL6						
low	1.000			1.000		
high	1.519	0.980-2.355	0.061	1.515	0.975-2.354	0.064
Gender						
female	1.000			1.000		
male	2.184	1.450-3.290	<0.001	1.802	1.180-2.753	0.006
Age						
≦63	1.000			1.000		
>63	1.808	1.214-2.691	0.004	1.997	1.303-3.062	0.002
Stage						
I	1.000			1.000		
II,III,IV	1.871	1.232-2.840	0.003	2.241	1.443-3.481	<0.001

Table 3. Influence of EGFL6 on overall survival in lung adenocarcinoma patients according to the age.

		Multivariate		
Sub-group	Case number	Hazard Ratio of EGFL6 Expression[1]	95% CI	*p*
Divide via medium age				
≦63	76	2.118	1.082-4.145	0.029
>63	74	1.184	0.661-2.122	0.570
Divide via grouped-age				
≦59	61	2.894	1.245-6.726	0.014
≦69	106	2.104	1.184-3.739	0.011
All	150	1.515	0.975-2.354	0.064

[1]Expression of EGFL6: high vs low

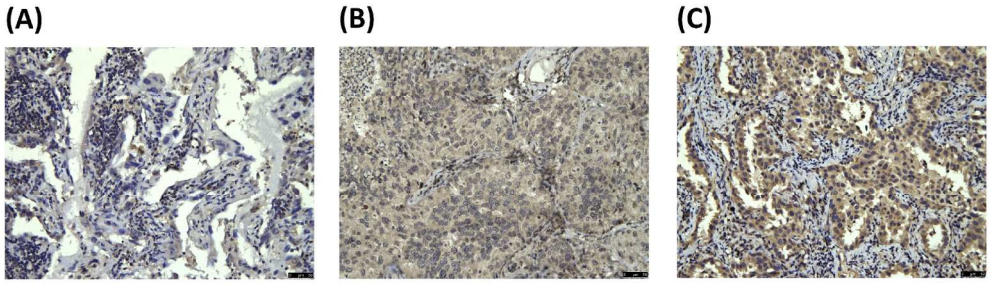


Figure 1. Representative immunostaining of EGFL6 in tissue arrays of NSCLC specimens. EGFL6 expression levels were (A) 0; (B) 1+; (C) 2+.

315x93mm (300 x 300 DPI)

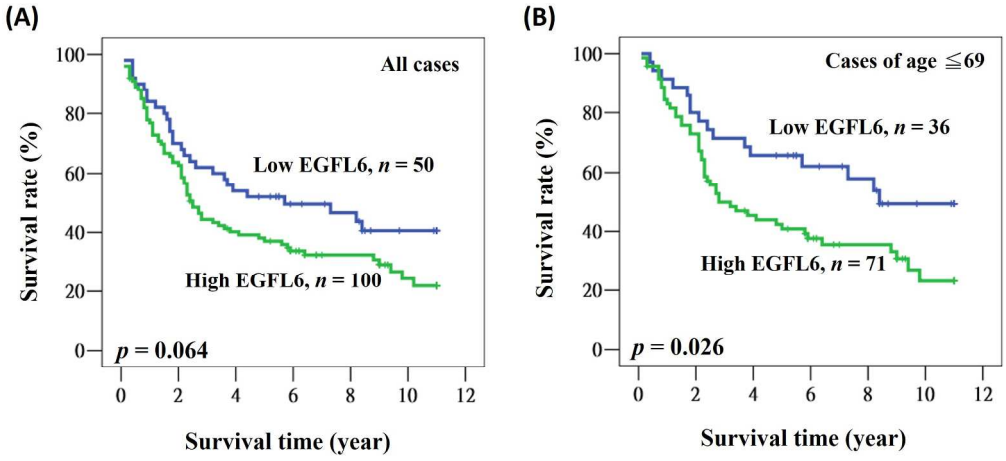
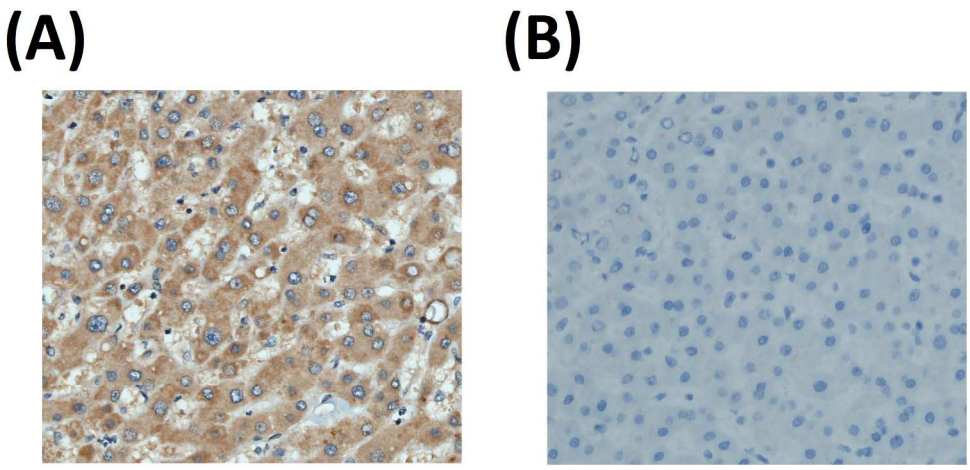


Figure 2. Kaplan-Meier actuarial analysis of EGFL6 expression with respect to overall survival of patients of (A) all patients, and (B) patients younger than 69 years old.

225x103mm (300 x 300 DPI)



Supplementary Figure 1. Representative control immunostaining of EGFL6 in liver tissue. (A) positive and (B) negative control.

237x113mm (300 x 300 DPI)

STROBE Statement

Checklist of items that should be included in reports of observational studies

Section/Topic	Item No	Recommendation	Reported on Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	3	State specific objectives, including any prespecified hypotheses	5-6
Methods			
Study design	4	Present key elements of study design early in the paper	7-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7-8
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	7-8
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	n/a
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8
Bias	9	Describe any efforts to address potential sources of bias	7
Study size	10	Explain how the study size was arrived at	7-8
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	7-8
		(*c*) Explain how missing data were addressed	7
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	n/a
		(*e*) Describe any sensitivity analyses	n/a

Section/Topic	Item No	Recommendation	Reported on Page No
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9-10
		(b) Indicate number of participants with missing data for each variable of interest	9-10
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	n/a
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	n/a
		Cross-sectional study—Report numbers of outcome events or summary measures	9-10
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9-10
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12
Generalisability	21	Discuss the generalisability (external validity) of the study results	n/a
Other Information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

**Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.*

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.