ERJ Express. Published on January 20, 2012 as doi: 10.1183/09031936.00097311

Prognostic value of *TP53*, *KRAS* and *EGFR* mutations in Non-Small Cell Lung Cancer: EUELC cohort

Chiara Scoccianti[1,2], Aurélien Vesin[3], Ghislaine Martel[1], Magali Olivier[1], Elisabeth Brambilla[3], Jean-François Timsit[3], the EUELC Consortium[4], Luca Tavecchio[5], Christian Brambilla[4], John K. Field[6] and Pierre Hainaut[1,7]

1: International Agency for Research on Cancer, Lyon, France; 2: School of Public Health, Imperial College, London, UK; 3: Institut Albert Bonniot, Université Joseph Fourier, INSERM U823, Grenoble, France; 5: Department of Experimental Oncology, Istituto Nazionale Milan, Italy; 6: Field JK: Roy Castle Lung Cancer Research Programme, The University of Liverpool Cancer Research Centre, Institute of Translational Medicine, The University of Liverpool, UK; 7: International Centre for Research and Training (CIPE),Sao Paulo, Brazil;

4: EUELC Consortium: John K Field[1,] Christian Brambilla[2], Yves Martinet[3], Erik Thunnissen[4], Peter Snijders[5], Gabriella Sozzi[6], Angela Risch[7], Stuart Elborn[8], Luis M. Montuenga[9], Ken O'Byrne[10], David J. Harrison[11], Jacek Niklinski[12]

[1]The University of Liverpool Cancer Research Centre, Institute of Translational Medicine, The University of Liverpool, United Kingdom; [2]INSERM, Grenoble, France; [3]Centeral Hospitalier Universitaire de Nancy, France; [4]Canisius Wilhelmina Ziekenhuis, Nijmegen, The Netherlands; [5]University Hospital Vrije Universiteit, Amsterdam, The Netherlands; [6]Department of Experimental Oncology, Istituto Nazionale Milan, Italy; [7]German Cancer Research Centre, Heidelberg, Germany; [8]Belfast City Hospital, Belfast, United Kingdom; [9]University of Navarra, Pamplona, Spain; [10]St James Hospital, Dublin, Ireland; [11]University of Edinburgh, Edinburgh, United Kingdom; [12]Medical Academy of Bialystok, Bialystok, Poland

Correspondence to:

Pierre Hainaut,

hainaut@iarc.fr; pierre.hainaut@cipe.accamargo.org.br

Tel: +33 4 72 73 85 32 Fax +33 4 72 73 83 22

Keywords

EGFR, *KRAS*, mutations, non small cell lung cancer, prognosis, *TP53*.

Abstract

Non-Small Cell Lung cancer samples from the European Early Lung Cancer biobank were analysed to assess the prognostic significance of mutations in *TP53*, *KRAS* and *EGFR* genes.

The series included 11 never smokers, 86 former smokers, 152 current smokers and 1 patient without smoking status informed. There were 110 Squamous Cell Carcinomas (SCC), 133 ADenoCarcinoma (ADC) and 7 Large Cell Carcinoma or mixed histologies. Expression of p53 protein was analysed by immunohistochemistry. DNA was extracted from frozen tumour tissues.

TP53 mutations were detected in 48.8% of cases and were more frequent among SCC than ADC ($p<0.0001$). *TP53* mutation status was not associated with prognosis. G to T transversions, known to be associated with smoking, were marginally more common among patients who developed a second-primary lung cancer or recurrence/metastasis (Progressive Disease). *EGFR* mutations were almost exclusively found in never smoking women ($p=0.0067$). *KRAS* mutations were detected in 18.5% of cases, mainly ADC ($p<0.0001$), and showed a tendency for association with Progressive Disease status.

These results suggest that mutations are good markers of different aetiologies and histopathological forms of lung cancers but have little prognostic value, with the exception of *KRAS* mutation which may have a prognostic value in ADC.

Introduction

The development of Non-Small Cell Lung cancers (NSCLC) is accompanied by multiple genetic and epigenetic alterations, with some differences according to aetiology and histological type [1, 2]. A survey of 139 NSCLC cell lines has identified a panel of frequently mutated genes which may be useful for NSCLC stratification on the basis of activating mutations. These genes include *KRAS, EGFR, ALK, MET, PDGFR, ROS, ERBB2, BRAF PI3K and MEK1* [3]. The most commonly mutated of these genes are *EGFR* and *KRAS* (the latter mostly in ADC, and being mutually exclusive). *EGFR* encodes a transmembrane receptor for Epidermal Growth Factor and related ligands, which contains an intracellular tyrosine kinase domain. Mutations are found almost exclusively in lung cancers of never-smokers and cluster in domains of the kinase that constitutively activate its activity and signal transduction. *KRAS* encodes a GTP/GDP exchange factor acting as downstream effector of *EGFR* signalling that mediates the activation of growth promoting signalling cascades of kinases. Mutations mostly fall at codon 12, located in the GTP binding pocket and preventing its hydrolysis.

In addition to these activating mutations, inactivating mutations in *TP53* are detected in the majority of NSCLC. *TP53* encodes an all-round tumour suppressor transcription factor, p53, which mediates multiple anti-proliferative effects in response to a variety of stresses, including in particular DNA damage. Most known mutations fall within the DNA-binding domain and inactivate the

suppressor by preventing DNA binding and transactivation. There is evidence that *TP53* or *KRAS* transversion mutations in NSCLC of smokers occur prevalently at G bases and are commonly the sites of adduct formation by metabolites of polycyclic aromatic hydrocarbons, one of the main family of tobacco carcinogens [4-6]. These observations suggest that at least some of these mutations may occur as the consequence of exposure to tobacco smoke and precede the development of cancer, therefore having an impact on molecular and biological patterns of lung carcinogenesis. However, the impact of these mutations on clinical prognosis remains a matter of debate.

The purpose of this study was to investigate the prognosis impact of mutations in *TP53*, *KRAS* or *EGFR* in resected, early stage NSCLC and to evaluate their use as biomarkers of disease progression. We took advantage of the EUELC project and biobank [7, 8] to select a group of patients with good quality frozen tissues. EUECL patients were recruited in 12 centres in 8 European countries and were followed on a 6-month basis after surgery. We show that *TP53* and *EGFR* mutations, although common in these cancers, have limited if any prognostic value, whereas *KRAS* mutations could be associated with Progressive Disease status.

Materials and methods

Study subjects and tumours

The European Early Lung Cancer (EUELC) project is a collaboration involving twelve Centres in France, Germany, southern Ireland, Italy, the Netherlands, Poland, Spain and the UK. This study has recruited 762 patients with surgically resected primary lung cancers who were considered at very high risk of developing Second Primary Lung Cancers (SPLC) and/or metastasis in relation to occupational or lifestyle risk factors. Among those, 739 were evaluated for disease progression and followed-up at 6-month intervals for up to 48 months (median: 29 months). All patients filled a lifestyle and medical questionnaire at each follow-up visit. Patients with a history of a completely resected primary lung or head & neck cancer who developed a SPLC, a recurrence or metastasis, or who died of the disease were grouped as Progressive Disease (PD). Patients who were alive and asymptomatic for the disease and who were not undergoing treatment by chemotherapy and/or radiotherapy at the time of the last follow-up were classified as Disease-Free (DF). Data on smoking and occupation were collected using a standardised lifestyle questionnaire. Instructions for interviewing and coding were developed and training to research interviewers was carried out in each centre. All questionnaires were translated to ensure consistency across EU partners.

A total of 306 samples were available for p53 protein detection by immunohistochemistry

(Figure 1). 273 frozen tissues were found to be suitable for DNA extraction and mutation analysis. 250 patients with follow-up status informed were finally selected for statistical analysis of *TP53* and *KRAS*. The series included 11 never smokers (<100 cigarettes smoked in a lifetime), 86 former smokers (smoking cessation ≥2 years before diagnosis), 152 current smokers (still smoking, or <2 years since cessation) and 1 patient without smoking status informed. There were 110 SCC, 133 ADC and 7 recorded as "other histologies" (large cell carcinoma or mixed histologies). *EGFR* mutations were analysed in 130 ADC based on earlier reports that this gene is rarely mutated in other types of NSCLC than ADC and in smokers [9].

Mutation analysis

DNA previously extracted from frozen tissue was received and analysed for *TP53* (exons 4-10 including flanking splice sites) mutations by pre-screening with denaturing high-pressure liquid chromatography (dHPLC) followed by a second PCR and bi-directional automated sequencing as described elsewhere [10]. Specimens with matched dHPLC and sequencing results were considered as containing a mutation. *KRAS* mutations at codon 12 were analysed by mutant-enriched PCR as described elsewhere [10], allowing enrichment of the mutant sequence, and sequenced.

EGFR mutations were detected using PCR-based direct sequencing of the four exons of the TK domain (exons 18–21) using primers and annealing conditions as described elsewhere [11].

Immunohistochemistry for p53 was performed as detailed previously [12] using the Ventana automated immunostainer (Ventana Corp., Tucson, AZ) with specified procedures and reagents. Percentage of stained tumour cells was evaluated on a scale from 0 to 4 (0: absent; 1: <10%; 2: 10% to 50%; 3: 50% to 90%; 4: >90%). Intensity of staining was assessed on a scale from 0 (absent) to 3 (marked). The results for percentage and intensity were summed up to generate a composite score as follows: sum of 0, no staining (score 0); sum of 1 to 3, slight staining (score 1); sum of 4 to 5, moderate staining (score 2); and sum of 6 to 7, marked staining (score 3).

Statistical analysis

Mantel-Haenszel chi square was used to test the association between clinical parameters and biomarkers but also between biomarkers. Fine & Gray (F&G) model was used to measure association between clinical variables and biomarkers with cancer progression. The model takes into account the presence of competing risks which in our study are patients who died from causes other than lung cancer. Hazard Ratios (HR) in the F&G model can be interpreted like Relative Risks. Bootstrap was performed to obtain non-parametric confidence intervals for risk estimates. According to the distribution of follow-up duration we censored the analysis at 48 months. Each biomarker was assessed one at a time in a multivariate model adjusted with the clinical variables significantly associated to the disease progression risk in the univariate analysis. Cumulative incidence

plots were performed to illustrate the risk of disease progression through time according to the mutation status of the genes.

Standard survival analysis was performed using the Cox proportional hazard model to assess association between overall death, lung cancer specific death and biomarkers. Adjustment on clinical parameters associated to death was done. All the analyses were stratified by centre. All statistical analyses were performed using SAS 9.1.3 (SAS Institute, NC).

Results

Patients, mutation prevalence and associations with individual and pathological parameters

Selected characteristics of patients are shown in Table 1 and mutation prevalence is shown in Table 2. A total of 48.4% *TP53* mutations, including 5 silent mutations, were detected. *KRAS* mutations at codon 12 and *EGFR* mutations were detected in 18.5% and in 13.1% of samples, respectively. Eighteen patients had mutations in two genes (Table 2), including 11 patients with both *TP53* and *KRAS* mutations and 6 patients with both *TP53* and *EGFR* mutations. One patient with *KRAS* mutation also had a *EGFR* silent mutation in exon 21 (codon 836, CGC>CGT Arg>Arg). No patient had mutations in the 3 genes.

[INSERT TABLES 1 and 2]

The patterns of mutations in *TP53* are shown in Supplementary Figure 1. *TP53* mutations and the codon distribution were in agreement with the known smoking-patterns [4], with about 33% of G:C to T:A transversions and hotspots at codons 157 and 158. Mutations in *EGFR* were spread among the 4 exons tested (4% in exon 18, 3% in exon 19, 3% in exon 20 and 5% in exon 21) and were all previously reported in the COSMIC mutation database [13].Table 3 shows the associations between mutations and selected pathological or individual

variables. *TP53* mutations were less frequent in ADC (39.7%) than SCC (57%); p<0.0001 (Table 3a). *KRAS* mutations were preferentially found in ADC (89.1%) than SCC (10.9%); p<0.0001 (Table 3b). None of these mutations were associated with either T or N status of TNM classification of tumours. *TP53* mutations were marginally more common in subjects who reported a past history of pulmonary illness or a familial history of lung cancer, but these associations were not statistically significant (Supplementary Tables 1b and 1c: p=0.1505 and p=0.1620, respectively). Neither smoking status nor history of asbestos exposure were associated with *TP53* or *KRAS* mutation status. No significant association was found between *TP53* mutation and smoking duration, age at smoking initiation, consumption in pack-years, time since quitting smoking or cigarette type (data not shown). Among *EGFR* mutations, 23.5% were found among never smoking women (p = 0.007) (Table 3c).

[INSERT TABLES 3a to 3c]

Association between *TP53* mutations and p53 expression

Missense *TP53* mutations may lead to nuclear accumulation of mutant p53 protein. Information on both mutation status and p53 IHC was available for 230 patients. There was a strong correlation between mutation status and p53 IHC (p<0.0001, Supplementary Table 3). Among tumours with mutations, 62% were highly positive for p53 protein. Among tumours with wild-type *TP53*, however, 25% had widespread, high expression of p53 across the tumour, suggesting that

p53 may be widely expressed in a subset of lung cancers without missense mutations in exons 4 to 9.

Prognostic significance of mutations

There were 26.4% PD and 73.6% DF. The following parameters were significantly associated with PD status (data not shown): T status of TNM (T1 versus T2 or more, p< 0.0001), N status of TNM (N0 versus N1, N2 or NX, p< 0.0001). *TP53*, *KRAS* or *EGFR* mutation status, however, were not associated with prognosis (Figure 2). No prognostic value was found when mutations were grouped into different categories according to their predicted effects on p53 protein structure or function [14]. G to T transversions were marginally more common among PD patients than DF (Table 4) but this effect was not statistically significant (adjusted HR: 1.49 [0.66-3.36], p=0.13).

[INSERT TABLE 4]

Likewise, p53 IHC positive status was not associated with prognosis (p=NS). Since there were important disparities in the recruitment of patients among countries and centres, we repeated these analyses on the largest homogenous subgroup, comprising the 103 patients from the French centres (Nancy and Grenoble). Again, in this subgroup, neither *KRAS* nor *TP53* mutations had prognostic value (results not shown). However, patients with tumours containing both mutations had a marginally significantly higher risk of developing a PD (adjusted HR: 3.30 [1.08-10.0], p=0.036).

TP53 mutations in relation with *TP53* polymorphisms

The *TP53* gene is highly polymorphic and there is evidence that mutations may occur at different rates on different *TP53* alleles. We analysed the distribution of 3 common polymorphisms located within a 312 bp region of the *TP53* gene encoding the N-terminus of p53, in relation with *TP53* mutation status. These three polymorphisms are PIN2 (G to C, intron 2, rs.1642785), PIN3 (16bp duplication, intron 3, rs.17878362) and PEX4 (non-silent G to C, codon 72, R to P; rs.1042522). Results (Table 5) show that there was a tendency for more mutations to occur in subjects who were carriers of two PEX4 C alleles encoding P at codon 72 (85.7%, as compared to 43.9 and 46.6 in G-C heterozygotes and G-G homozygote respectively; p=0.05). The two other polymorphisms did not appear to be associated with significant differences in mutation prevalence (data not shown).

[INSERT TABLE 5]

Discussion

Many studies have investigated the prognostic value of *TP53* or *KRAS* mutations in lung cancer. There is evidence that both the pattern and frequency of mutations vary according to risk factors such as tobacco smoke. However, it remains unclear whether mutations are associated with an increased risk of disease progression and of unfavourable outcome. Here we have used the setup of a large European collaborative study, EUELC, to assess the prognostic value of *TP53* and *KRAS* mutations in a series of 250 NSCLC cases with detailed follow-up information. We have analysed the relationships between *TP53* mutations and several common *TP53* polymorphisms. Finally, we have assessed *EGFR* mutations in 130 adenocarcinomas, as mutations in this gene have been reported to be rare in other lung cancers histologies [15, 16].

Results show that *TP53* mutations were present in 48.4% and *KRAS* mutations in 18.5% of the cases. For both genes, the codon distribution showed a high proportion of G to T transversions in agreement with the well-documented prevalence of this mutation type in lung cancers of smokers. We also observed differences between the two main histological forms of NSCLC, SCC and ADC. *TP53* mutations were detected in 57% of SCC, versus 39.7% in ADC. In contrast, *KRAS* mutations were detected in 89.1% of ADC versus 10.9% of SCC. As shown in other case series, *KRAS* mutations tended to be more common in lung cancers

of ever than former or never-smokers (52.2%, 45.7% and 2.2%, respectively). Among clinical and etiological factors, only histology was statistically associated with mutation prevalence, while never-smoking status was significantly associated with *EGFR* mutation (p=0.0067). One tumour contained both *EGFR* and *KRAS* mutation, an extremely rare occurrence according to the literature. Interestingly, the *EGFR* mutation in this tumour was a silent one (codon 836 CGC>CGT Arg>Arg) and was thus not supposed to lead to tyrosine kinase activation.

In the present case series, mutation of none of the three genes analysed to carry a significant prognostic value in the cohort as a whole or in specific histological subgroups. Given the multi-centric character of the study and the possibility of a bias due to recruitment centre, we performed a separate analysis on the largest and most homogeneous subgroup which revealed a borderline effect in patients carrying both *TP53* and *KRAS* mutations (HR=3.26 [1.07-9.90], p=0.038). The value of these analyses is constrained by the relatively small sample size and it will be important to verify this interpretation in larger cohorts.

Similar to our results, a study on Japanese patients with surgically resected ADC did not identify any prognostic implication for *TP53* or *KRAS* mutations [17]. The authors detected a significant association between *EGFR* mutation and longer survival while none of the gene mutations appeared to be an independent prognosis marker. Of note, in this Japanese series 49% of the patients had *EGFR*

mutations, a much higher rate than in the present Caucasian series (13.1%). It is well documented that mutations in *EGFR* are associated with never-smoking status, female gender and Asian ethnicity [15-17]. The relatively low prevalence of *EGFR* mutations in our series may reflect the characteristics of the patients recruited in EUELC, i.e. Caucasian, 84% males and 95.2% ever smokers. Given these characteristics the *EGFR* mutation showed a higher than expected rate, and it was not restricted to NSCLC of never-smokers since detected in about 10% of former (5/48) or current (8/72) smokers.

Based on these results, the conservative conclusion is that mutation status does not predict short-term outcomes in completely resected lung cancers and given the overall poor prognosis of lung cancer over a period of 5 to 8 years, it remains to be determined whether it may be a prognostic factor for longer-term outcomes.

From a biological viewpoint, *TP53* and *KRAS* mutations may represent very early events in lung carcinogenesis, occurring before tumour onset as the result of genetic damage by tobacco components. Although these mutations do participate in launching bronchial cells towards transformation and progression, it is likely that the tumour behaviour may be dictated by specific, additional events occurring after initiation by tobacco carcinogens. We found that tumours carrying both *TP53* and *KRAS* mutations might have a worse prognosis and this underline a possible higher exposure to tobacco carcinogens or a particular susceptibility to

their mutagenic effects. These patients may have increased risk of acquiring additional mutations which, in turn, may be responsible of their poorer prognosis. Thus, presence of both *TP53* and *KRAS* mutations in the same lesion may act as a marker to identify a small group of tumours that are genetically unstable and prone to the accumulation of mutations which may accelerate disease progression and/or escape from therapy. Further studies are needed to identify the targets of such genetic instability in NSCLC. Candidate markers may involve genes with activating mutations, making it possible to treat these cancers using selective pharmacological inhibitors [3], and epigenetic changes in DNA methylation patterns and in microRNA expression which may distinguish different NSCLC subgroups [18].

Our data on *TP53* polymorphisms show that *TP53* mutations tend to occur at different rates on different *TP53* alleles. Although the group of patients was small, patients with two PEX4 C alleles tended to have more frequently a *TP53* mutation than patients with at least one G allele. This suggests that the *TP53* C allele may be intrinsically more “mutable” than the G allele, perhaps as a result of subtle differences in the functional properties of p53 proteins. Experimental studies have identified such functional differences, including a greater ability to induce apoptosis for 72P than for 72A [19]. This observation is in agreement with results from Mechanic et al. [20] who found that common genetic variation in *TP53* could modulate lung cancer pathways, as suggested by the association

of *TP53* codon 72 polymorphism with lung cancer in African Americans and with somatic *TP53* mutation frequency in lung tumours. Thus, in future studies, it may be important to take into account both *TP53* mutation and *TP53* haplotypes in assessing the prognostic and predictive significance of *TP53* gene status in lung cancer.

Acknowledgements

This study was supported in part by the PNES program on Lung Cancers of the French National Cancer Institute (INCa). The European Early Lung Cancer study was supported by a Framework V grant from the European Union (QLG1-CT-2002-01735).

References

1. Jedrychowski W, Becher H, Wahrendorf J, *et al*. Effect of tobacco smoking on various histological types of lung cancer. *J Cancer Res Clin Oncol* 1992; 118: 276–82.
2. Pfeifer GP, Denissenko MF, Olivier M, *et al*. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* 2002; 21: 7435–51.
3. Sharma SV, Haber DA, Settleman J. Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. Nat Rev Cancer 2010; 10(4):241-53.
4. Denissenko MF, Pao A, Tang MS *et al*. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspot in P53. *Science* 1996; 274: 430–32.
5. Hainaut P, Pfeifer GP. Patterns of p53 G>T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. *Carcinogenesis* 2001; 22: 367–74.
6. Hussain SP, Amstad P, Raja K, *et al*. Mutability of p53 hotspot codons to benzo(a)pyrene diol epoxide (BPDE) and the frequency of p53 mutations in non-tumorous human lung. *Cancer Res* 2001; 61: 6350–5.
7. Cassidy A, Balsan J, Vesin A, *et al*. Cancer diagnosis in first-degree relatives and non-small cell lung cancer risk: Results from a multi-centre case-control study in Europe. *Eur J Cancer* 2009; 45(17): 3047-53.
8. Field JK, Liloglou T, Niaz A, *et al*. EUELC project: a multi-centre, multipurpose study to investigate early stage NSCLC, and to establish a biobank for ongoing collaboration. *ERJ* 2009; 34(6): 1477-86.
9. Mounawar M, Mukeria A, Le Calvez F, *et al*. Patterns of EGFR, HER2, TP53, and KRAS Mutations of p14arf Expression in Non–Small Cell Lung Cancers in Relation to Smoking History. *Cancer Res* 2007; 67(12): 5667-72.

10. Le Calvez F, Mukeria A, Hunt JD, *et al*. TP53 and KRAS Mutation Load and Types in Lung Cancers in Relation to Tobacco Smoke: Distinct Patterns in Never, Former, and Current Smokers. *Cancer Res* 2005; 65(12): 5076-83.
11. Pao W, Miller V, Zakowski M, *et al*. EGF receptor gene mutations are common in lung cancers from ‘‘never smokers’’ and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA* 2004; 101: 13306–11.
12. Burke L, Flieder DB, Guinee DG, *et al*. Prognostic implications of molecular and immunohistochemical profiles of the Rb and p53 cell cycle regulatory pathways in primary non-small cell lung carcinoma. *Clin Cancer Res* 2005;11(1):232-41.
13. COSMIC Database. www.sanger.ac.uk/genetics/CGP/cosmic/. Last Modified Mar 23 2011.
14. Petitjean A, Mathe E, Kato S, *et al*. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum Mutat* 2007; 28(6): 622-9.
15. Shigematsu H, Lin L, Takahashi T, *et al*. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *JNCI* 2005; 97: 339–46.
16. Yatabe1 Y and Mitsudomi T. Epidermal growth factor receptor mutations in lung cancers. *Pathology International* 2007; 57: 233–44.
17. Kosaka T, Yatabe Y, Onozato R, *et al*. Prognostic implication of EGFR, KRAS, and TP53 gene mutations in a large cohort of Japanese patients with surgically treated lung adenocarcinoma. J Thorac Oncol 2009; 4(1):22-9.
18. Voortman J, Goto A, Mendiboure J, *et al*. MicroRNA expression and clinical outcomes in patients treated with adjuvant chemotherapy after complete resection of non-small cell lung carcinoma. Cancer Res 2010; 70(21):8288-98.

19. Dumont P, Leu JI, Della Pietra AC, *et al*. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat Genet* 2003; 33(3): 357-65.
20. Mechanic LE, Bowman ED, Welsh JA, *et al*. Common genetic variation in TP53 is associated with lung cancer risk and prognosis in African Americans and somatic mutations in lung tumors. *CEBP* 2007; 16(2): 214-22.

Table 1: Characteristics of selected patients included for statistical analysis

Variable	Items	n
Gender	Man	210
	Woman	40
Age	< 60	89
	[60-65[	82
	[65-70[	30
	≥ 70	49
Education level[$]	No/Primary Level	181
	High education	59
	Missing	10
Histology	ADC	133
	SCC	110
	Others	7
Asbestos exposure	None	191
	Yes	57
	Missing	2
pT	T1	76
	T2	150
	T3	15
	T4	8

	Missing	1
pN	N0	173
	N1	65
	N2	2
	NX	9
	Missing	1
Past pulmonary illness	No	110
	Yes	138
	Missing	2
Smoking status£	Current smoker	152
	Former smoker	86
	Never smoker	11
	Missing	1
Total		250

$ Education level was missing for 10 subjects

£ Former smokers: patients who quitted smoking at least 2 years before interview. Current smokers: patients who were smokers in the last 2 years before interview

Table 2: Single and multiple mutation prevalence in EUELC patients

Gene (n)	Status	n (%)	n (%)
TP53 (250)	Wild-type	129 (51.6)	
	Mutant (exons 4-9)	121 (48.4)	
KRAS (249)	Wild-type	203 (81.5)	
	Mutant (codon 12)	46 (18.5)	46 (100)
	TP53 Wild-type		35 (76.1)
	TP53 Mutant		11 (23.9)
EGFR (130)[$]	Wild-type	113 (86.9)	
	Mutant	17 (13.1)	17 (100)
	TP53 Wild-type		11 (64.7)
	TP53 Mutant		6 (35.3)
	KRAS Wild-type		16 (95)
	KRAS Mutant		1 (5)

n= number of analysed samples

$ The group of cases analysed includes only ADC.

Table 3: Associations between mutations, patient's variables and clinical parameters

3a: *TP53*

Variable (missing)	Items	Wild-type (n = 129)		Mutated (n = 121)		P[$]
		n	%	n	%	
Histology	ADC	85	65.9	48	39.7	**<.0001**
	SCC/others	44	31.8	73	57	
pT (1)	T1	39	30.5	37	30.6	0.9810
	T2	76	59.4	74	61.2	
	T3	8	6.3	7	5.8	
	T4	5	3.9	3	2.5	
pN (1)	N0	89	69.5	84	69.4	0.3696
	N1	30	23.4	35	28.9	
	N2	2	1.6	0	0	
	Nx	7	5.5	2	1.7	
Smoking status (1) *	Current smoker	77	60.2	75	62	0.8287
	Former smoker	44	34.4	42	34.7	
	Never smoker	7	5.5	4	3.3	
Asbestos exposure (2)	None	100	78.7	91	75.2	0.8007
	Yes	27	21.3	30	24.8	

3b: *KRAS*

Variable (missing)	Items	Wild-type (n = 203)		Mutated (n = 46)		P[$]
		n	%	n	%	
Histology	ADC	93	45.8	41	89.1	**<.0001**
	SCC/others	110	54.2	5	10.9	
pT (1)	T1	68	33.7	9	19.6	0.10
	T2, T3, T4	134	66.3	37	80.4	
pN (1)	N0	140	69.3	33	71.7	0.60
	N1, N2, Nx	62	30.7	13	28.3	
Smoking status (1) *	Current smoker	128	63.4	24	52.2	0.08
	Former smoker	64	31.7	21	45.7	
	Never smoker	10	5	1	2.2	
Asbestos exposure [£]	None	154	76.2	36	80	0.82
	Yes	48	23.8	9	20	

3c: *EGFR*

Variable (missing)	Items	Wild-type (n = 113)		Mutated (n = 17)		P[$]
		n	%	n	%	
Gender	Man	92	81.4	11	64.7	0.35

	Woman	21	18.6	6	35.3	
Smoking status*	Current smoker	64	56.6	8	47.1	0.11
	Former smoker	43	38.1	5	29.4	
	Never smoker	6	5.3	4	23.5	
Gender / Smoking status	Others	109	96.5	13	76.5	**0.0067**
	Never smoking women	4	3.5	4	23.5	
Pack-years (1) [£]	≤ 40	61	54.5	13	76.5	0.15
	> 40	51	45.5	4	23.5	

$ Mantel Haenszel test controlling for centre

* Former smokers: patients who quitted smoking at least 2 years before interview. Current smokers: patients who were smokers in the last 2 years before interview

Table 4: Associations between biomarkers and disease progression

Variable	Items	DF		PD		HR	P*	Adj. HR	P$
		n	%	n	%	(95% CI)		(95% CI)	
TP53 status	Wild-type	70	49.3	59	54.6	1		1	
							0.45		0.64
	Mutated	72	50.7	49	45.4	0.86 (0.59 – 1.27)		0.91 (0.62 – 1.40)	
Type	0 – Others	124	88.6	84	80.8	1		1	
							0.19		0.14
	1 – All G > T	16	11.4	20	19.2	1.4 (0.9 – 2.3)		1.46 (0.89 – 2.41)	
KRAS status	Wild-type	118	84.3	85	78	1		1	
							0.26		0.46
	Mutated	22	15.7	24	22	1.30 (0.82 – 2.06)		1.19 (0.75 – 1.90)	
KRAS / *TP53* status	Otherwise	138	97.9	102	93.6	1		1	
							0.07		0.21
	Both Mutated	3	2.1	7	6.4	2.08 (0.95 – 4.57)		1.67 (0.74 – 3.77)	
p53 haplotype	GNA-CDP	33	23.6	23	21.1	1.07 (0.64 – 1.76)		1.16 (0.69 – 1.96)	
	GNA-CNP	17	12.1	15	13.8	1.15 (0.63 – 2.07)		1.52 (0.63 – 2.11)	
							0.95		0.93
	OTHERS	65	46.4	49	45	1.15 (0.69 – 1.92)		1.11 (0.66 – 1.89)	
	GNA-GNA	25	17.9	22	20.2	1		1	
EGFR£ status	Wild-type	62	87.3	51	86.4	1		1	
							0.48		0.68
	Mutated	9	12.7	8	13.6	1.31 (0.62 – 2.80)		0.97 (0.67 – 1.38)	

* F&G model with centre stratification

$ F&G model with centre stratification adjusted on pT and pN

£ The group of cases analysed includes 130 ADC

Table 5: Associations between *TP53* mutation and polymorphisms

Variable	Items	*TP53* status				P[$]
		WT	%	MT	%	
PIN2	CC	7	5.5	15	12.8	
	GC	58	45.3	44	37.6	0.21
	GG	63	49.2	58	49.6	
PIN3	DD	4	3.1	8	6.8	
	ND	43	33.6	29	24.8	0.16
	NN	81	63.3	80	68.4	
PEX4	CC	2	1.6	12	10.3	
	CG	55	43.0	43	36.8	**0.05**
	GG	71	55.5	62	53.0	

$ Mantel Haenszel Chi square controlling for centre

Figure 1: Flow chart of samples selection for mutational analysis

The initial number of tumour samples qualified for the study is indicated and the number of samples analysed for *TP53*, *EGFR* and *KRAS* is given.

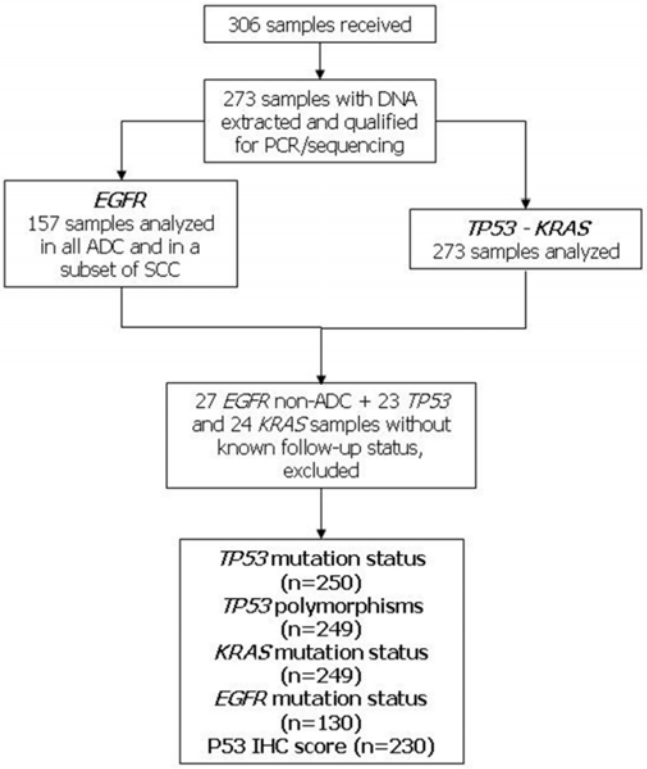
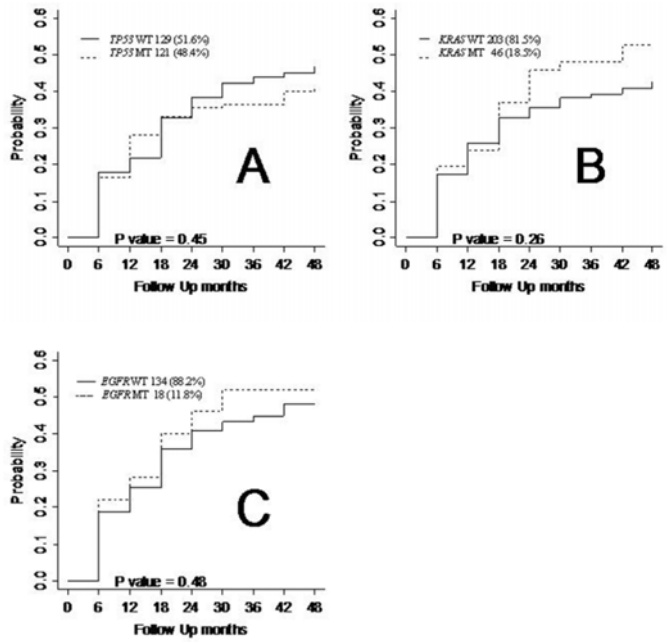


Figure 2: Cumulative incidence plots of the Progressive Disease risk for *TP53*, *KRAS* and *EGFR* mutations

Proportion of subjects with Progressive Disease detected during follow-up, of maximum 48 months, after complete resection of the primary tumour. Numbers of cases with and without mutations and percentage, are given. *P* values from univariate F&G model. Panel A: *TP53,* Panel B: *KRAS*, Panel C; *EGFR*.


TP53 WT 129 (51.6%)
TP53 MT 121 (48.4%)
Probability
A
P value = 0.45
Follow Up months
KRAS WT 203 (81.5%)
KRAS MT 46 (18.5%)
Probability
B
P value = 0.26
Follow Up months
EGFR WT 134 (88.2%)
EGFR MT 18 (11.8%)
Probability
C
P value = 0.48
Follow Up months