

Defining high-risk individuals in a population-based molecular-epidemiological study of lung cancer

ADRIAN CASSIDY¹, JONATHAN P. MYLES², TRIANTAFILLOS LILOGLOU¹,
STEPHEN W. DUFFY^{2,3} and JOHN K. FIELD^{1,4}

¹Roy Castle Lung Cancer Research Programme, University of Liverpool Cancer Research Centre, University of Liverpool, Liverpool L3 9TA; ²The London Department of Epidemiology, Mathematics and Statistics, Wolfson Institute of Preventive Medicine, London EC1M 6BQ; ³Cancer Screening and Aetiology Group, CRUK Clinical Centre at Barts, EC1M 6BQ; ⁴Department of Clinical Dental Sciences, University of Liverpool, Liverpool L3 9TA, UK

Received November 23, 2005; Accepted December 29, 2005

Abstract. Within the framework of the Liverpool Lung Project (LLP), population-based case-control and prospective cohort studies are in progress to identify molecular and epidemiological risk factors and define populations and individuals most at risk of developing lung cancer. This report describes a strategy for selection of a high-risk population and further provides support for the inclusion of occupational and genetic risk factors in future models. Data from the case-control study (256 incident cases and 314 population controls) were analysed to define a high-risk population. Detailed lifestyle and occupational information were collected during structured interviews. Models were constructed using conditional logistic regression and included terms for age, tobacco consumption and previous respiratory disease. Smoking duration was chosen as the most important predictor of lung cancer risk [>50 years (OR 15.65, 95% CI 6.10-40.15)]. However, such a model would preclude younger individuals. Several combinations of previous respiratory disease were also considered, of which a history of bronchitis, emphysema or pneumonia (BEP) was the most significant (OR 1.86, 95% CI 1.28-2.69). A high-risk subset (based on combinations of smoking duration and BEP) was identified, which have a 4.5-fold greater risk of developing lung cancer (OR 4.5, 95% CI 2.33-8.68). Future refinement of the risk model to include individuals occupationally exposed to asbestos and with the p21 genotypes is discussed. There is real potential for environmental and genetic factors to improve on risk prediction and targeting of susceptible individuals beyond the traditional models based only on smoking and age. The

development of a molecular-epidemiological model will inform the development of effective surveillance, early detection and chemoprevention strategies.

Introduction

Lung cancer is the most common cancer in the world, both in terms of incidence (1.35 million new cases, representing 12.4% of all new cancers) and mortality (1.18 million deaths, or 17.6% of the world total) (1). Large geographical differences exist in the levels and trends in both men and women. In men, the incidence rate, including all histologic types, varies up to 35-fold between high and low-risk areas. Males consistently show higher lung cancer incidence than females in all populations, with male:female ratios varying approximately from 1.5 to 20 (2). In recent years a rapid increase in lung cancer incidence has been observed among women in developed countries, contrasting with a levelling off or decrease among men (3,4). Both geographic and gender disparities are mainly due to differences in patterns of tobacco smoking (5).

Detection of lung cancer usually occurs late in the disease when it is beyond effective treatment; consequently, there is a high mortality rate. More than 40% of all lung cancer patients present with metastasised disease at diagnosis (6). In these clinically advanced tumour stages, long-term survival is rarely achieved with conventional cytotoxic agents (7). Individuals at high risk for specific cancers such as non-small cell lung cancer have, apart from smoking cessation, no options to reduce their steadily increasing risk (8). Long-term and heavy smokers who give up smoking do experience a reduced relative risk, but the absolute risk remains high for years after giving up (9). The way forward for improved management and prognosis for individuals at high risk probably lies with early detection of disease. New molecular approaches provide potential hope for early diagnosis and screening of high-risk individuals, determination of prognosis and identification of innovative treatments (10). At the same time, modern imaging technologies afford hope of effective early detection (11). However, this solution can only be viable in health economic terms through accurate stratification of risks and monitoring of those at high risk.

Correspondence to: Professor John K. Field, Roy Castle Lung Cancer Research Programme, The University of Liverpool Cancer Research Centre, The University of Liverpool, 200 London Road, Liverpool L3 9TA, UK
E-mail: j.k.field@liv.ac.uk

Key words: lung cancer, high-risk, smoking, respiratory disease, occupation, genotype

While more than 80% of people who develop lung cancer are current or former smokers, only a minority of smokers develop lung cancer. To better understand the aetiology of lung cancer and to more effectively target high-risk individuals for prevention and screening interventions, it is important to identify factors that influence a smoker's risk of developing lung cancer (12). Several studies have suggested that prior lung diseases, such as asthma, chronic bronchitis, emphysema, pneumonia, tuberculosis, hay fever and impaired pulmonary function may modify lung cancer risk (12-20). The association between occupational exposure and lung cancer has also been investigated. Results on occupations known to entail exposure to lung carcinogens consistently indicate an association, but there is wide variation in the magnitude of the risk estimates, probably because of changes in industrial processes over time and between countries (21). Approximately 14% of lifetime smokers develop lung cancer and 10-15% of all lung cancers occur among non-smokers (22,23). These figures suggest that there are individual differences in susceptibility to lung carcinogens and that these individual differences may be the result of genetic predisposition to lung cancer (24). Results from several studies of familial clustering and segregation analyses are indicative of an inherited component to lung cancer risk (25-31).

Given that the causes of common cancers may have their basis in environmental exposures occurring in the genetically predisposed host, it is essential that we are able to study the interaction between lifestyle factors and susceptibility genes to produce a risk assessment model. The Liverpool Lung Project (LLP) is conducting population-based case-control and prospective cohort studies to identify risk factors and define populations and individuals most at risk of developing lung cancer. The ultimate aim of the LLP will be to develop a model which predicts risk from epidemiological risk factors such as smoking and occupational exposures, from genetic/molecular factors and from changes in phenotypic biological markers. For the last purpose, it is necessary to have serial biological samples within the cohort study. For economic reasons, we propose to take such samples on a very high-risk minority of the cohort study members. In this report, we discuss the options involved in selecting such a group and the issues which arise in the exercise. We describe a strategy for selection of the high-risk subgroup based on interim results from the case-control study with respect to smoking and history of respiratory diseases. In addition, we suggest how the inclusion of occupational and genetic factors might enhance the process of selecting a high-risk population.

Materials and methods

We aim to recruit 800 lung cancer cases and 1,600 population controls to the case-control study. Structured questionnaires collect detailed information on lifestyle factors (active and passive smoking, medical history, family history of cancer, occupation and residence) over the whole life course. Blood and sputum is collected from each individual for DNA extraction and genotyping. In the cohort study, it is planned to recruit 7,500 individuals and follow them up over a 10-year period. By recruiting these individuals from particularly high-risk areas within the Liverpool region, we expect to observe a relatively

Table I. Distribution of select characteristics of lung cancer cases and controls.

Characteristic	Cases		Controls	
	N	%	N	%
Smoking				
Never	19	7.4	89	28.3
Ever	237	92.6	225	71.7
Age (years)				
≤60	66	25.8	81	25.8
60-64	45	17.6	62	19.7
65-69	46	17.9	67	21.3
>70	99	38.7	104	33.1
Mean age (years) (SD ^a)	66.1 (8.7)		64.7 (9.0)	
Bronchitis				
No	152	59.4	211	67.2
Yes	104	40.6	103	32.8
Emphysema				
No	239	93.4	305	97.1
Yes	17	6.6	9	2.9
Asthma				
No	220	85.9	277	88.2
Yes	36	14.1	37	11.8
Pneumonia				
No	190	74.2	268	85.4
Yes	66	25.8	46	14.6

^aSD, standard deviation.

high yield of lung cancers. Detailed lifestyle, residential, and job history data are also collected over the whole life course for cohort study members, and will therefore provide prospective validation of the results of the case-control study. In addition, we plan to take serial biological samples (blood and induced sputum) from a high-risk subgroup of the cohort, comprising of between 20 and 30% of the study population. The Liverpool Lung Project protocol has been approved by Liverpool, South Sefton and St. Helens & Knowsley local research ethics committees (32).

Newly diagnosed cases of primary lung cancer are recruited through specialist NHS chest clinics. Inclusion in the study is dependent upon histological or cytological confirmation. Eligible cases are recruited within three months of diagnosis, as a result, only a minority of cases died before interview and next-of-kin interviews were not required. Population controls were selected from registers of General Practitioners. Given the strong association between lung cancer and age and gender, two controls per case were matched on this basis (+/-2 years). We used an interim analysis of the case-control recruits in

Table II. Effect of various measures of smoking history on risk of lung cancer.

Measure	Case (%)	Control (%)	p-value	OR	95% CI	% AR
Ever smoked						
No	19 (7.4)	89 (28.3)	<0.0001	1.00	-	79
Yes	237 (92.6)	225 (71.7)		6.29	(3.23-12.27)	
Current smoker						
No	162 (63.3)	258 (82.2)	<0.0001	1.00	-	26
Yes	94 (36.7)	56 (17.8)		3.04	(1.91-4.82)	
Duration (years)						
0	17 (6.6)	85 (27.1)	<0.0001	1.00		
1-25	26 (10.1)	93 (29.6)		1.51	(0.62-3.69)	
26-40	66 (25.8)	70 (22.3)		5.88	(2.72-12.71)	
41-50	88 (34.4)	42 (13.3)		9.99	(4.53-22.01)	
>50	59 (23.0)	24 (7.6)		15.65	(6.10-40.15)	
Pack-years						
0	18 (7.0)	86 (27.2)	<0.0001	1.00		
1-15	23 (9.0)	70 (22.1)		1.50	(0.62-3.65)	
16-30	57 (22.3)	64 (20.2)		5.82	(2.64-12.82)	
31-40	56 (21.9)	32 (10.1)		9.46	(4.83-21.70)	
41-50	67 (26.2)	42 (13.3)		10.12	(4.48-22.85)	
>50	35 (13.7)	20 (6.3)		14.50	(5.55-37.88)	
Amount (per day)						
0	18 (7.0)	86 (27.2)	<0.0001	1.00		
1-10	37 (14.4)	61 (19.3)		3.04	(1.37-6.76)	
11-15	39 (15.2)	37 (11.7)		6.43	(2.89-14.26)	
16-20	90 (35.1)	68 (21.5)		7.83	(3.74-16.36)	
21-30	46 (18.0)	38 (12.0)		8.89	(3.82-20.66)	
>30	26 (10.1)	24 (7.6)		8.95	(3.61-22.18)	

2004 to define the high-risk subgroup. The analyses included 256 lung cancer cases and 314 population controls (with either one or two controls per case). The aim was to delineate a high-risk subgroup based on smoking history and on personal history of respiratory diseases.

Smoking history. Lifetime smoking histories were recorded, including age at starting and stopping each smoking epoch, and amount smoked per day during each epoch (a smoking epoch being defined as a period in which the individual's reported type of cigarette and daily consumption remained constant). The risk factors derived from this included: i) ever smoked (never-smoker defined as someone who had smoked less than 100 cigarettes in his or her lifetime); ii) current smoker (yes or no); iii) duration of smoking (years); iv) pack-years (calculated from the number of cigarette packs (of 20) smoked per day and years of smoking); v) amount smoked (average number of cigarettes per day).

Respiratory disease. The various categories of respiratory disease considered in the analyses included: history of: i)

bronchitis or emphysema (BE); ii) bronchitis, emphysema or asthma (BEA); iii) bronchitis, emphysema or pneumonia (BEP); iv) bronchitis, emphysema, asthma or pneumonia (BEAP); v) history of any respiratory disease (ANY).

Statistical analysis. Data were analysed by conditional logistic regression (33). This allows for multivariate effects on risk and takes into account the matched design of the case-control study. For categorical analyses, exposure categories were based on tertiles of the distribution among exposed controls, and individuals never exposed to the risk factor under study comprised the reference category. Conditional logistic regression models were fitted to the data, estimating odds ratios (OR) and 95% confidence intervals (CI). Models included terms for age, tobacco consumption (5 categories) and previous respiratory disease (5 categories).

Results

There was no appreciable difference in age between cases (mean 66.1; SD 8.7) and controls (mean 64.7; SD 9.0) (Table I).

Table III. History of respiratory disease and risk of lung cancer.

Respiratory disease	Case	Control	p-value	OR	95% CI	% AR
BE						
No	145	210	0.04	1.00	-	13
Yes	111	104		1.47	(1.01-2.14)	
BEA						
No	133	202	0.02	1.00	-	17
Yes	123	112		1.56	(1.08-2.24)	
BEP						
No	122	193	0.0001	1.00	-	25
Yes	134	121		1.86	(1.28-2.69)	
BEAP						
No	115	185	0.002	1.00	-	25
Yes	141	129		1.80	(1.25-2.60)	
ANY						
No	109	179	0.002	1.00	-	25
Yes	147	135		1.79	(1.25-2.57)	

BE, bronchitis or emphysema. BEA, bronchitis, emphysema or asthma. BEP, bronchitis, emphysema or pneumonia. BEAP, bronchitis, emphysema, asthma or pneumonia. ANY, history of any respiratory disease.

7.4% of the cases were never smokers compared to 28.3% of the controls. Both cases and controls reported a high lifetime prevalence of bronchitis (case, 40.6%; control, 32.8%) and pneumonia (case, 25.8%; control, 14.6%).

Table II shows the unadjusted association of each measure of smoking history with lung cancer. All measures were significantly associated with increased risk of disease. Smoking duration was chosen as the most important predictor because it was the most significant both in terms of the p-values from a trend test ($p < 0.000$) and of the highest odds ratio between lowest (OR 1.51, 95% CI 0.62-3.69) and high-risk groups (OR 15.56, 95% CI 6.10-40.15) (with fewer categories than pack-years or amount smoked).

Table III shows the unadjusted association of each measure of respiratory disease with lung cancer. The factor BEP shows the most significant effect, the highest odds ratio (OR 1.86, 95% CI 1.28-2.69) with similar attributable risk fraction (AR=25%) as BEAP and ANY. In the last respect, BEP betters BEAP and ANY in specificity since it attains the same AR of 25% with smaller control prevalence. None of these factors were significant when adjusted for smoking duration.

Based on these results, it would seem reasonable to define the high-risk subset in terms of smoking only. Given that the strongest predictor of lung cancer amongst the smoking variables is total duration of smoking, the logical course would be to base our population on smoking duration.

Based on the controls, selecting individuals who have smoked for 40 years will result in a subset of around 20% of the study population. The majority of the cases would be observed in this subset (stratified by smoking duration).

Table IV. History of BEP by case-control status and effect of BEP and smoking duration in individuals aged less than 65 years.

Factor	Case	Control	OR	p-value	95% CI
BEP					
No	58	102	1.97	0.07	(0.96-4.05)
Yes	60	52			
Smoking duration					
Trend	-	-	1.07	<0.001	(1.04-1.07)

BEP, bronchitis, emphysema or pneumonia.

However, in order to have smoked for at least 40 years, an individual will almost certainly be elderly (Table II). Indeed, of the 66 controls with more than 40 years of smoking 71% (47 individuals) are aged 65 or over. To include a number of younger individuals in the high-risk subset, we factored the various combinations of respiratory disease (BE, BEA, BEP, BEAP, ANY) into the regression model. The respiratory disease measure with the highest relative risk (OR 1.86) and lowest p-value ($p = 0.0001$) is bronchitis, emphysema or pneumonia (BEP). Table IV shows BEP by case-control status and the joint effect of BEP and smoking duration in

Table V. Relative risk of lung cancer for prevalent occupational exposures.^a

Exposure	Odds ratio	95% CI
Asbestos	2.08	(1.25-3.45)
Chrysotile	2.15	(1.20-3.59)
Amphibole	1.93	(1.12-3.33)
Sand	1.31	(0.74-2.30)
Silica	1.59	(0.74-3.41)
Concrete	1.49	(0.87-2.56)
Wood	1.21	(0.70-2.08)
Hardwood	1.50	(0.78-2.90)
Softwood	1.17	(0.68-2.03)
Iron fumes	1.56	(0.91-2.66)
Petrol emission	0.92	(0.51-1.67)
Diesel emission	1.14	(0.67-1.94)
Lubricating oil	1.14	(0.73-1.78)
Cutting oils	1.38	(0.81-2.37)
Mineral oils	1.26	(0.65-2.41)

^aAdjusted for smoking and history of bronchitis, emphysema or pneumonia (BEP).

individuals less than 65 years. Including all individuals with BEP would result in the subset being too large. Therefore we combined smoking status and history of BEP. Selecting individuals with a history of BEP and more than 20 years of smoking resulted in 16% of total subgroup aged less than 65 years.

The following criteria were therefore used for the selection of a high-risk subset: i) individuals aged 65 or more with more than 40 years of smoking; ii) individuals aged less than

65 with a history of BEP and more than 20 years of smoking. Overall, this selects a high risk subgroup of 26% of the population. On this basis, this subset will have a 4.5-fold risk of developing lung cancer (OR 4.50, 95% CI 2.33-8.86) ($p < 0.01$). Between 60 and 70% of lung cancer cases would be expected to develop in this group, of which 36% will be less than 65 years old.

Discussion

The above selection strategy would yield a high-risk subset, based on age, smoking duration and history of respiratory disease, of 26% of the population in which 60-70% of the lung cancer cases would arise. A more compact risk group of 21% of the population, with 57% of the cases, could be selected based on smoking duration alone, but this would have confined the age structure of the high-risk subset to the elderly.

Published models of individual lung cancer risk have tended to concentrate on smoking and age (11). However, despite the predominant role of smoking in the causation of lung cancer, other contributing agents also exist. It is therefore worth considering how such models might be improved upon by inclusion of occupational carcinogens and genetic markers.

In this case-control study, detailed occupational information was collected from cases and controls during structured interviews. For each job held, an expert assessed the probability, intensity and duration of exposure to 70 agents or groups of agents on the basis of specialised occupational questionnaires (34). Table V shows the odds ratios associated with workplace exposure to relatively prevalent occupational carcinogens, after adjustment for smoking and history of BEP. Thus, even after taking smoking into account, the most important risk factor for lung cancer, occupational exposures, notably asbestos (OR 2.08, 95% CI: 1.25-3.45), adds further significant information on risk. Changing the high-risk group selection to include all of those who have smoked for at least 25 years,

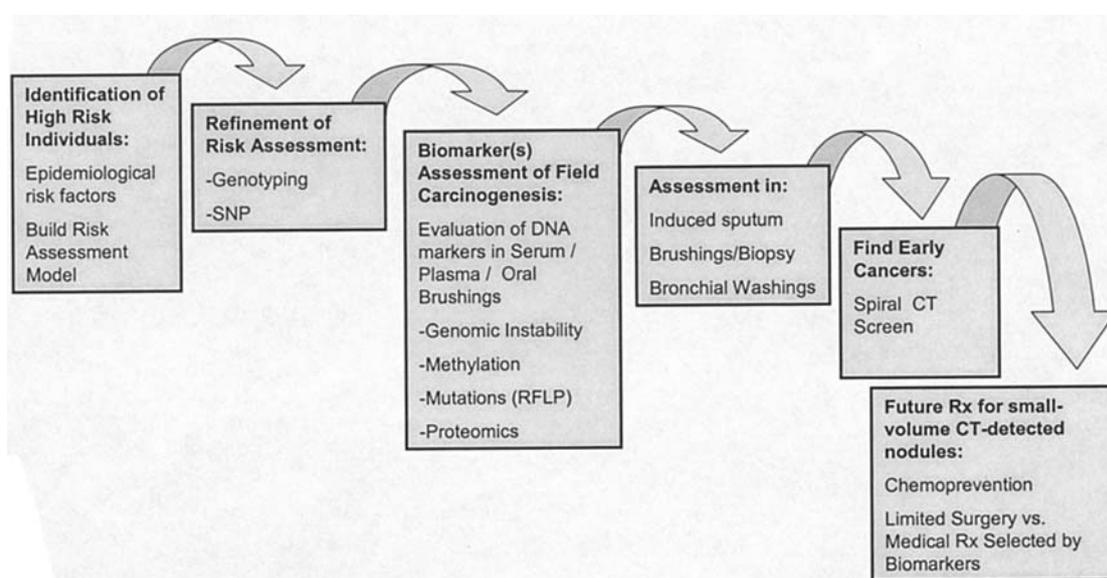


Figure 1. Biomarker-based lung cancer screening journey cascade.

and who have either been exposed to asbestos or have had one of BEP would yield a high risk group of 26% of the population. This group would contain 60-70% of the lung cancer cases. Around 40% of the members of the risk group would be expected to be aged 65 years or less. The considerable effort required to collect detailed job histories and assess occupational exposure would prohibit the use of this methodology for the selection of high-risk individuals within a general population setting. However, since it would only be required in those already satisfying the criteria of 25 years of smoking, it might be practical as part of a risk assessment strategy. Alternatively, the results from expert assessment could be used to develop a general population job-exposure matrix to identify individuals with exposure to occupational carcinogens significantly associated to lung cancer e.g. asbestos.

In a separate analysis of 250 cases and 500 controls, we have estimated the effect on lung cancer risk of the CDKN1A promoter polymorphism rs4135234 (Liloglou, unpublished data). Overall, this suggests a significantly increased risk of lung cancer for individuals with GA or AA genotypes. Furthermore, we have observed a significantly increased risk amongst individuals who have a history of pneumonia or tuberculosis. The CDKN1A gene has a key role in the cell cycle control as its encoded protein p21 inhibits multiple cyclin/cyclin-dependent kinase complexes, namely Cyclin D/CDK4,6, Cyclin E/CDK2 and Cyclin A/CDK2. It is also involved in apoptosis control in p53-dependent and p53-independent manner. Thus, one could hypothesize that the transcriptional activation of CDKN1A following DNA damage is very important in the response of the cell to carcinogen exposure. The G/A polymorphism that has been identified in the promoter region (35) is therefore of interest as it may affect transactivation by p53 due to its high proximity to one of the p53 binding sites in CDKN1A promoter. This suggests that there is scope for genetic effects to further refine risk group selection. As with occupational exposures, genotyping would not be feasible on a population basis, but it might be possible in a small subgroup of long-term smokers, for example. It should be noted that neither the purpose nor the likely outcome of research into genetic risk markers is to identify a population who may smoke without fear of developing lung cancer. The most effective action to reduce the risk of lung cancer is not to start smoking, and in the case of smokers, to stop. The genetic factors identified so far have not had an effect on risk even approaching the magnitude of that of long-term smoking, and no genetic effects have been identified which approach nullifying the substantial risks associated with smoking. The point of such research is to further refine estimates of individual risk, so as to ensure the cost-effectiveness of future early detection and chemoprevention interventions, both in terms of the healthcare burden and the benefit/risk ratio of the individuals concerned.

Currently Spiral CT imaging would appear to be the best clinical modality for identifying very early lung cancer (36). However, the future lies in identifying individuals who are at risk of developing disease prior to its actual inception. Field and colleagues (37) have developed a decision-making paradigm which assimilates this rationale (Fig. 1). In the first instance, high-risk populations would be identified through epidemiological risk factors such as those described in this

report. High-risk individuals would subsequently be asked to provide a blood specimen from which lung cancer susceptibility markers (i.e. SNPs) would be identified. Detection of biomarkers in serum or plasma and bronchial lavage or induced sputum constitutes the third and fourth stage, respectively. The cumulative risk of developing lung cancer at this point would trigger clinical investigations through Spiral CT imaging. Field and colleagues postulate that the combined results of imaging and molecular-pathological investigations will determine the individual's future lung cancer treatment regime. Several molecularly targeted therapies are currently being evaluated to determine whether they might play a preventative role in high-risk individuals (38).

In conclusion, there is real potential for environmental and genetic factors to improve on risk prediction and targeting of susceptible individuals beyond the traditional models based only on smoking and age. Advances in genetics have made studies incorporating information on exposure and genotype feasible and represent the next logical step in understanding the factors determining or predicting the outset of lung cancer. The eventual understanding of the genetic and environmental basis of lung cancers will not only enable us to identify high-risk populations but to develop effective prevention, early detection and chemoprevention strategies.

Acknowledgements

The Liverpool Lung Project is funded by the Roy Castle Foundation. JPM supported by CR-UK.

References

1. Parkin DM, Bray F, Ferlay J and Pisani P: Global cancer statistics, 2002. *CA Cancer J Clin* 55: 74-108, 2005.
2. Parkin DM, Whelan SL, Ferlay J, Teppo L and Thomas D: Cancer incidence in five continents, Vol. VIII. IARC Scientific Publication no. 155. IARC Press, Lyon, 2003.
3. Bray F, Tyczynski JE and Parkin DM: Going up or coming down? The changing phases of the lung cancer epidemic from 1967 to 1999 in the 15 European Union countries. *Eur J Cancer* 40: 96-125, 2004.
4. Tyczynski JE, Bray F, Aareleid T, *et al*: Lung cancer mortality patterns in selected Central, Eastern and Southern European countries. *Int J Cancer* 109: 598-610, 2004.
5. Doll R and Peto R: *The Causes of Cancer*. Oxford University Press, Oxford, 1981.
6. Ginsberg RJ, Vokes EE and Rosenzweig K: Non-small cell lung cancer. In: *Cancer: Principles and Practice of Oncology*. 6th edition. Vol. 1. De Vita VT, Hellmann S and Rosenberg SA (eds). Lippincott Williams & Wilkins, Philadelphia, pp925-983, 2001.
7. Schiller JH, Harrington D, Belani CP, *et al*: Eastern Cooperative Oncology Group. Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. *N Engl J Med* 346: 92-98, 2002.
8. Van Zandwijk N and Hirsch FR: Chemoprevention of lung cancer: current status and future prospects. *Lung Cancer* 42 (Suppl. 1): S71-S79, 2003.
9. Peto R, Darby S, Deo H, Silcocks P, Whitley E and Doll R: Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. *BMJ* 321: 323-329, 2000.
10. Huber RM and Stratakis DF: Molecular oncology - perspectives in lung cancer. *Lung Cancer* 45 (Suppl. 2): S209-S213, 2004.
11. Van Klaveren RJ, De Koning HJ, Mulshine J and Hirsch FR: Lung cancer screening by spiral CT. What is the optimal target population for screening trials? *Lung Cancer* 38: 243-252, 2002.
12. Littman AJ, Thornquist MD, White E, Jackson LA, Goodman GE and Vaughan TL: Prior lung disease and risk of lung cancer in a large prospective study. *Cancer Causes Control* 15: 819-827, 2004.

13. Alavanja MC, Brownson RC, Boice JD Jr and Hock E: Preexisting lung disease and lung cancer among non-smoking women. *Am J Epidemiol* 136: 623-632, 1992.
14. Wu AH, Fontham ET, Reynolds P, *et al*: Previous lung disease and risk of lung cancer among lifetime non-smoking women in the United States. *Am J Epidemiol* 141: 1023-1032, 1995.
15. Mayne ST, Buenconsejo J and Janerich DT: Previous lung disease and risk of lung cancer among men and women non-smokers. *Am J Epidemiol* 149: 13-20, 1999.
16. Osann KE, Lowery JT and Schell MJ: Small cell lung cancer in women: risk associated with smoking, prior respiratory disease, and occupation. *Lung Cancer* 28: 1-10, 2000.
17. Brenner AV, Wang Z, Kleinerman RA, *et al*: Previous pulmonary diseases and risk of lung cancer in Gansu Province, China. *Int J Epidemiol* 30: 118-124, 2001.
18. Talbot-Smith A, Fritschi L, Divitini ML, Mallon DF and Knuiaman MW: Allergy, atopy, and cancer: a prospective study of the 1981 Busselton cohort. *Am J Epidemiol* 157: 606-612, 2003.
19. Santillan AA, Camargo CA Jr and Colditz GA: A meta-analysis of asthma and risk of lung cancer (United States). *Cancer Causes Control* 14: 327-334, 2003.
20. Schabath MB, Delclos GL, Martynowicz MM, *et al*: Opposing effects of emphysema, hay fever, and select genetic variants on lung cancer risk. *Am J Epidemiol* 161: 412-422, 2005.
21. Richiardi L, Boffetta P, Simonato L, *et al*: Occupational risk factors for lung cancer in men and women: a population-based case-control study in Italy. *Cancer Causes Control* 15: 285-294, 2004.
22. Mattson ME, Pollack ES and Cullen JW: What are the odds that smoking will kill you? *Am J Public Health* 77: 425-431, 1987.
23. Crispo A, Brennan P, Jockel KH, *et al*: The cumulative risk of lung cancer among current, ex- and never-smokers in European men. *Br J Cancer* 91: 1280-1286, 2004.
24. Schwartz AG: Genetic predisposition to lung cancer. *Chest* 125 (Suppl. 5): S86-S89, 2004.
25. Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE and Rothschild H: Increased familial risk for lung cancer. *J Natl Cancer Inst* 76: 217-222, 1986.
26. Houlston RS and Peto J: Genetics of common cancers. In: *Inherited Predisposition to Cancer*. Eeles RA, Ponder B, Easton DE and Horwich A (eds). Chapman and Hall, London, pp208-226, 1996.
27. Schwartz AG, Yang P and Swanson GM: Familial risk of lung cancer among non-smokers and their relatives. *Am J Epidemiol* 144: 554-562, 1996.
28. Amos CI, Xu W and Spitz MR: Is there a genetic basis for lung cancer susceptibility? *Rec Res Cancer Res* 151: 3-12, 1999.
29. Mayne ST, Buenconsejo J and Janerich DT: Familial cancer history and lung cancer risk in United States non-smoking men and women. *Cancer Epidemiol Biomarkers Prev* 8: 1065-1069, 1999.
30. Bromen K, Pohlabeln H, Jahn I, Ahrens W and Jockel KH: Aggregation of lung cancer in families: results from a population-based case-control study in Germany. *Am J Epidemiol* 152: 497-505, 2000.
31. Xu H, Spitz MR, Amos CI and Shete S: Complex segregation analysis reveals a multigene model for lung cancer. *Hum Genet* 116: 121-127, 2005.
32. Field JK, Smith DL, Duffy S and Cassidy A: The Liverpool Lung Project Research Protocol. *Int J Oncol* 27: 1633-1645, 2005.
33. Breslow NE and Day NE: *Statistical Methods in Cancer Research*. Vol. 1. IARC Scientific Publications, Lyon, 1980.
34. Siemiatycki J: Risk factors for cancer in the workplace. CRC Press, Boca Raton, pp45-115, 1991.
35. Liloglou T, Risk JM and Field JK: Polymorphisms in the promoter region of the WAF1/CIP1 gene. *Int J Oncol* 9: 559-562, 1996.
36. Mulshine JL and Sullivan DC: Clinical practice. Lung cancer screening. *N Engl J Med* 352: 2714-2720, 2005.
37. Field JK, Beer D, Carbone D, *et al*: A European strategy for developing lung cancer imaging and molecular diagnosis in high risk populations. *Cancer* (In press).
38. Shepherd FA: A targeted approach to reducing lung cancer mortality. *J Clin Oncol* 23: 3173-3174, 2005.