

Incorporation of a Genetic Factor into an Epidemiologic Model for Prediction of Individual Risk of Lung Cancer: The Liverpool Lung Project

Olaide Y. Raji¹, Olorunsola F. Agbaje², Stephen W. Duffy³, Adrian Cassidy¹, and John K. Field¹

Abstract

The Liverpool Lung Project (LLP) has previously developed a risk model for prediction of 5-year absolute risk of lung cancer based on five epidemiologic risk factors. SEZ6L, a Met430Ile polymorphic variant found on 22q12.2 region, has been previously linked with an increased risk of lung cancer in a case-control population. In this article, we quantify the improvement in risk prediction with addition of SEZ6L to the LLP risk model. Data from 388 LLP subjects genotyped for SEZ6L single-nucleotide polymorphism (SNP) were combined with epidemiologic risk factors. Multivariable conditional logistic regression was used to predict 5-year absolute risk of lung cancer with and without this SNP. The improvement in the model associated with the SEZ6L SNP was assessed through pairwise comparison of the area under the receiver operating characteristic curve and the net reclassification improvements (NRI). The extended model showed better calibration compared with the baseline model. There was a statistically significant modest increase in the area under the receiver operating characteristic curve when SEZ6L was added into the baseline model. The NRI also revealed a statistically significant improvement of around 12% for the extended model; this improvement was better for subjects classified into the two intermediate-risk categories by the baseline model (NRI, 27%). Our results suggest that the addition of SEZ6L improved the performance of the LLP risk model, particularly for subjects whose initial absolute risks were unable to discriminate into "low-risk" or "high-risk" group. This work shows an approach to incorporate genetic biomarkers in risk models for predicting an individual's lung cancer risk. *Cancer Prev Res*; 3(5); 664–9. ©2010 AACR.

Introduction

Interest in methods of assessing individual risk of developing diseases has continued to grow over the years, partly due to their usefulness in selection of high-risk individuals that would benefit in prevention or screening programs (1–3). Within the Liverpool Lung Project (LLP), we have previously developed and validated a predictive model for 5-year absolute risk of developing lung cancer for an individual with a specific combination of epidemiologic risk factors including smoking duration, previous diagnosis of

pneumonia, prior diagnosis of malignant tumor, occupational exposure to asbestos, and family history of lung cancer (4). Recent advancement in genetic epidemiology leading to identification of genetic and molecular variants affecting the risk of disease means that genetic markers such as single-nucleotide polymorphisms (SNP) can be added to risk models for improved prediction of future risk of disease (5–7).

DNA pooling and high-throughput sequencing recently conducted as part of the Sequenom-Genefinder project has identified a candidate region in chromosome 22q12.2 that contained SNPs showing significant differences between lung cancer case-control subjects (8). This region overlies the seizure 6-like (*SEZ6L*) gene, with polymorphic Met430Ile (SNP rs663048) identified in the region as a top candidate variant modulating the risk of lung cancer. The SEZ6L marker SNP was further validated in individual genotyping in two independent data sets: the LLP and the M.D. Anderson Cancer Study populations (8). The two studies independently found a significant effect of the variant with a combined 3-fold risk of lung cancer for the homozygotes mutant allele compared with the wild-type.

This study incorporates the SEZ6L SNP into the LLP risk model for predicting 5-year absolute risk of developing

Authors' Affiliations: ¹Roy Castle Lung Cancer Research Programme, School of Cancer Studies, The University of Liverpool Cancer Research Centre, Liverpool, United Kingdom; ²Division of Cancer Studies, Cancer Epidemiology Unit, King's College London, School of Medicine, Academic Oncology, Guy's Hospital; and ³Cancer Research UK Centre for EMS, Queen Mary University of London, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, London, United Kingdom

Corresponding Author: John K. Field, Liverpool Cancer Research Centre, Roy Castle Building, University of Liverpool, 200 London Road, Liverpool L3 9TA, United Kingdom. Phone: 44-151-794-8900; Fax: 44-151-794-8989; E-mail: J.K.Field@liv.ac.uk

doi: 10.1158/1940-6207.CAPR-09-0141

©2010 American Association for Cancer Research.

Table 1. Distribution of subjects' epidemiologic characteristics by SEZ6L SNP genotype

Characteristics	SEZ6L genotype			P
	GG freq (%)	TG freq (%)	TT freq (%)	
Subject status				
Case	104 (52)	76 (38)	20 (10)	0.02
Control	111 (59)	71 (38)	6 (3)	
Smoking duration (y)				
Never	43 (61.4)	25 (35.7)	2 (2.9)	0.81
<20	35 (54.7)	24 (37.5)	5 (7.8)	
20-39	58 (54.7)	39 (36.8)	9 (8.5)	
40+	79 (53.3)	59 (39.9)	10 (6.8)	
History of pneumonia				
Yes	34 (50.0)	30 (44.1)	4 (5.9)	0.53*
No	181 (56.6)	117 (36.6)	22 (6.9)	
Previous cancer diagnosis				
Yes	14 (53.9)	10 (38.5)	2 (7.7)	0.95*
No	201 (55.5)	137 (37.9)	24 (6.6)	
Occupation asbestos exposure				
Yes	73 (54.9)	48 (36.1)	12 (9.0)	0.41
No	142 (55.7)	99 (38.8)	14 (5.5)	
Family history of lung cancer				
No	165 (55.4)	109 (36.6)	24 (8.0)	0.16*
Early onset	13 (44.8)	16 (55.2)	0 (0.0)	
Late onset	37 (60.6)	22 (36.1)	2 (3.3)	

*P value for Fisher's exact test; no statistically significant interaction of SEZ6L with any of the risk factors in the LLP risk model.

lung cancer and assesses the predictive ability and improved accuracy of the extended model.

Patients and Methods

Data collected as part of the LLP case-control study were used in this study. The detailed LLP recruitment procedure and study protocol have been previously described elsewhere (4, 9). Briefly, incident cases of histologically or cytologically confirmed lung cancer, ages between 20 and 80 years, were included. Lung cancer included any of topographical subcategories of code C34 of the International Classification of Disease for Oncology 9th revision. Two population controls per case, matched on year of birth (± 2 years) and gender, were selected from registers of general practitioners in Liverpool area. All participants were resident in the Liverpool area and provided written informed consent.

The current analysis is based on 200 histologically and cytologically confirmed lung cancer cases and 188 age (± 2 years) and sex frequency-matched controls. This group of subjects were among those used for developing the original LLP model and were genotyped as part of the Sequenom-Genefinder project. A standardized questionnaire was used to collect clinical and lifestyle data including the five epidemiologic risk factors included in the LLP risk model (4). These were smoking duration (years), occupational exposure to asbestos, family history of lung

cancer including age at onset in the affected relatives, prior diagnosis of malignant diseases, and prior diagnosis of pneumonia. The study protocol was approved by the Liverpool Research Ethic committee.

SNP marker and genotyping

The Sequenom-Genefinder pilot study aimed to identify polymorphic susceptibility variants for lung cancer and genotyped and tested 83,715 SNP, located primarily in gene-based regions. A subset of the LLP case-control study subjects were genotyped as part of the validation study. The genotype was determined by extracting the genomic DNA from blood peripheral leukocytes using the Qiagen DNA Blood Mini Kit according to the manufacturer's instructions. Details of the genotyping process are provided elsewhere (8).

Statistical analysis

Pearson's χ^2 test was used to assess the relationship between the SNP marker and each of the epidemiologic risk factors included in the previously published LLP risk model. This test of association was done for all subjects together and for cases and controls separately.

The risk data were analyzed by conditional logistic regression. We refitted the original risk model for subjects with both genetic and epidemiologic data ($n = 388$) and compared the estimated coefficients to that of the original model ($n = 1,275$). We then fitted the model, including

the SEZ6L SNP, to estimate the effect of the latter, adjusting for the SNP.

The previously published multivariable conditional logistic regression model was used to generate estimates of predicted 5-year absolute risk of lung cancer with and without genotype information. The baseline risk (α —the constant term in the regression model) for the prediction of 5-year absolute risk using the model with genetic information was recalculated. The procedure for calculating the baseline α from age- and sex-specific lung cancer incidence rates for the Liverpool area has earlier been described (4); the only difference is that the linear combination of the β coefficients in the probability model now includes information on the SNP genotype.

The area under the receiver-operating characteristic curve was calculated as a measure of discriminatory ability of the models and to formally compare the models with and without the SEZ6L SNP (10). We also calculated the net reclassification improvement (NRI) to quantify the improvement in the LLP risk model containing the SEZ6L. NRI cross-classifies subjects' predicted risks from the original and extended models, assesses the proportions of subjects reclassified into new risk categories (case and control separately), and quantifies the correct movement in categories—upward for cases and downward for controls (11). Simple measures such as the true positive fraction (sensitivity) and false positive fraction (1 – specificity) were computed for the risk thresholds classifying the subjects

into high-, intermediate-, and low-risk groups. Improvement in model calibration was also assessed by comparing the closeness of the predicted risks to the observed risks (goodness of fit) for each model using the Hosmer-Lemeshow's χ^2 statistic (12, 13) and the Aikake information criteria (14).

Results

Subjects' characteristics

The distribution of the five epidemiologic and lifestyle characteristics of the 388 case-control subjects used in this study by SEZ6L genotype is presented in Table 1. The majority of subjects (55%) had SEZ6L genotype GG, 38% had heterozygote genotype TG, and approximately 7% had homozygote mutant genotype TT. The distribution by case-control status revealed a statistically significant increase in the proportion of cases with homozygote mutant genotype (10%) compared with controls (3%). There was no statistically significant relationship between any of the five epidemiologic factors and the SNP for either the combined or the separate (data not shown) analysis of the case and control subjects.

Multivariable risk model

Table 2 gives the odds ratios (OR) and 95% confidence intervals (95% CI) for the multivariate conditional logistic regression using different subsets of LLP subjects. The

Table 2. Summary of multivariable conditional logistic regression results for lung cancer risk prediction

Variables	LLP subjects with both EPI variables and SEZ6L SNP (n = 388)		Original LLP risk model using LLP case-control subjects (n = 1,275)	Model with SEZ6L only using all subjects genotyped (n = 463)
	Model with EPI variables only	Model with EPI + SEZ6L variables		
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Smoking duration (y)				
1-19	1.84 (0.79-4.24)	1.69 (0.72-3.94)	2.16 (1.22-3.82)	
20-39	5.32 (2.48-11.41)	5.08 (2.34-11.01)	4.28 (2.63-6.96)	
40+	13.72 (6.54-28.76)	13.37 (6.33-28.22)	12.39 (7.50-20.46)	
Pneumonia (Yes)	1.25 (0.68-2.31)	1.24 (0.67-2.33)	1.82 (1.26-2.63)	
Asbestos exposure (Yes)	2.35 (1.35-4.08)	2.32 (1.32-4.06)	1.89 (1.35-2.63)	
Previous tumor (Yes)	1.62 (0.61-4.32)	1.58 (0.58-4.29)	1.97 (1.23-3.15)	
Family history of cancer				
Early onset (<60 y)	2.16 (0.82-5.67)	2.34 (0.88-6.20)	2.02 (1.18-3.46)	
Late onset (>60 y)	0.87 (0.46-1.67)	0.93 (0.49-1.79)	1.19 (0.80-1.78)	
SEZ6L genotype marker				
TG		1.17 (0.72-1.92)		1.19 (0.78-1.82)
TT		4.51 (1.56-13.02)		3.79 (1.45-9.86)
Goodness-of-fit statistic				
Hosmer-Lemeshow (P)	6.24 (0.62)	5.54 (0.70)		
AIC	418.12	413.38		

Abbreviations: EPI, epidemiological; AIC, Aikake information criteria.

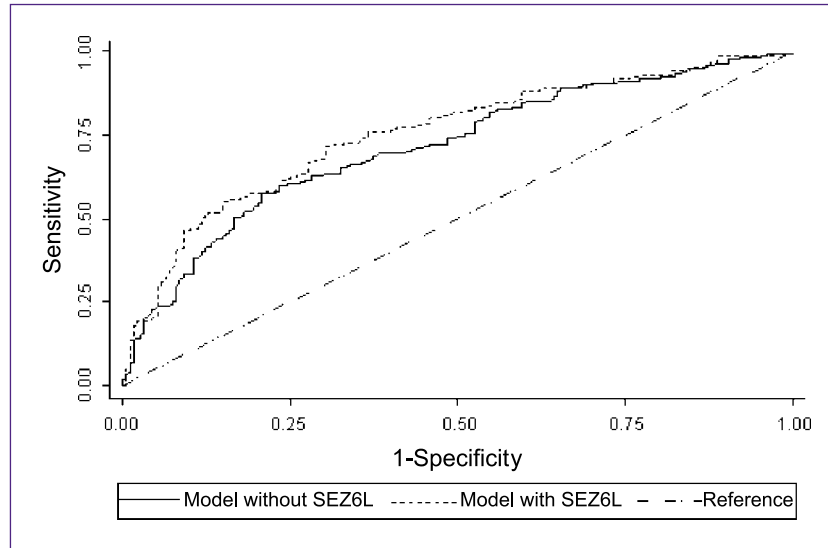


Fig. 1. Receiver-operating characteristic (ROC) curves with and without the SEZ6L SNP in the risk prediction model.

estimated ORs and 95% CIs for models with and without SEZ6L seemed comparable, suggesting no serious confounding effect of SEZ6L on the relationship between each of the original epidemiologic risk factors and lung cancer risk. The parameter estimates for the model with reduced subjects were similar to that of the original LLP risk model, but with lesser precision. We therefore retained coefficients of the original model and incorporated the adjusted estimate for the SEZ6L SNP for the expanded model, but with the α 's recalculated as previously mentioned above.

Predictive ability of the expanded risk model

There was an improvement in the calibration when SEZ6L was incorporated into the risk model, as shown by the reduced Hosmer-Lemeshow χ^2 statistic of 5.54 ($P = 0.70$) compared with the 6.24 ($P = 0.62$) observed for the model without SEZ6L (Table 2). This finding was supported by the reduction in the Aikake information criteria value from 418.12 for the model without SEZ6L to 413.38 for the model with SEZ6L, suggesting improved fit for the extended model.

Figure 1 shows the receiver operating characteristic curve for the models with and without the SEZ6L SNP. A significant ($P = 0.01$) increase of about 4% in area under the receiver-operating characteristic curve was observed (from 0.72 to 0.75) for the model extended with SEZ6L SNP.

The sensitivity and specificity corresponding to the three different threshold classifying the subjects into low-risk (<0.91), intermediate-risk (0.91 to <5.12), and high-risk (>5.12) groups are shown in Table 3. The threshold values were defined from the predicted 5-year absolute risks for the original LLP control samples ($n = 1,272$), assuming the risk distribution in this group to be similar to that of the general Liverpool population. The upper threshold (5.12) corresponds to the value for the top 20% of predicted absolute risks in the population; individuals whose 5-year predicted absolute risk is or above this value are designated as "high-risk" group. The lower threshold value of 0.91 corresponds to the bottom 40% of absolute risks in the control population and demarcates the "low-risk" group. This definition of high-risk and low-risk groups was used in a cardiovascular diseases study (15). The

Table 3. Sensitivity and specificity for models with and without SEZ6L at specified risk thresholds

Risk threshold	LLP risk model using all subjects ($n = 1,275$)		Refitted LLP model using subject with EPI + SNP variables ($n = 388$)			
	Se (Sp) (%)		Sensitivity (%)		Specificity (%)	
	No SEZ6L	With SEZ6L	No SEZ6L	With SEZ6L	No SEZ6L	With SEZ6L
0.91	86.1 (39.5)	83.0	84.5	43.1	42.0	
2.50	67.8 (64.2)	65.0	67.0	67.6	72.3	
5.12	49.9 (79.8)	50.5	50.0	83.5	88.3	

Abbreviations: Se, sensitivity; Sp, specificity.

Table 4. Reclassification (number and percentage) of predicted risks for cases and controls using models with and without the SEZ6L genotype

Model without SEZ6L	Model with SEZ6L				% Correct/total reclassification	% Net gain
	<0.91%	0.91-<2.5%	2.5-<5.12%	>5.12%		
Cases					13.5/24.5	2.5
<0.91%	27 (79.4)	7 (20.6)	—	—	20.6/20.6	
0.91-<2.5%	4 (10.5)	23 (60.5)	9 (23.7)	2 (5.3)	29.0/39.5	
2.5-<5.12%	—	6 (21.4)	13 (46.5)	9 (32.1)	32.1/53.6	
>5.12%	—	—	12 (12.0)	88 (88.0)	0/12.0	
Controls					14.4/19.7	9.1
<0.91%	78 (95.1)	4 (4.9)	—	—	0/4.9	
0.91-<2.5%	3 (6.7)	39 (86.6)	3 (6.7)	—	6.7/13.3	
2.5-<5.12%	—	12 (40.0)	15 (50.0)	3 (10.0)	40.0/50.0	
>5.12%	—	—	12 (38.7)	19 (61.3)	38.7/38.7	

observed values for the two measures (sensitivity and specificity) were comparable for the two models in the two extreme groups (low risk and high risk); this means that knowledge of the SNP genotype had little or no effect for subjects already classified to be of low or high risk. In contrast, the model with SEZ6L seemed to discriminate better in the intermediate group and is particularly more specific among this group of subjects.

Table 4 gives the joint distribution of the classified predicted risks from each model. Overall, approximately one quarter of cases (49 of 200) and about 20% of controls (37 of 188) had their predicted risks reclassified into other risk groups when SEZ6L was incorporated into the risk predictive model. This reclassification showed improvement (upward shift) in approximately 14% of cases and became worse (downward shift) for 11%, resulting in a net gain of about 3%. The net gain was higher for controls (9%) with overall improvement in risks (downward shift) for 14% and worst performance (upward shift) for only 5% of controls. These figures thus correspond to a statistically significant NRI of approximately 12% ($P = 0.03$). Concentrating only on those subjects whose initial risks are in the two intermediate groups (cases, 66; controls, 75), 20 of 30 reclassified cases and 15 of 21 reclassified controls had improved risks, resulting in a better NRI of 27%.

Discussion

We have evaluated the contribution of the SEZ6L gene, a genetic marker variant identified at 22q12.2 region, to prediction of individual absolute risk of developing lung cancer within a 5-year period using the recently developed LLP risk model. Our results suggest a modest increase in the overall predictive ability of the model with the SEZ6L SNP genotype. A greater effect was confined to subjects whose initial risks from the baseline model were in the medium-risk category. For these subjects, a NRI of around 27% was recorded, demonstrating the usefulness of the SNP in risk prediction.

The results presented in this article are in line with recent suggestions on taking predictive risk models of lung cancer to the next level where mediating genetic biological markers are included to allow a more accurate prediction of interindividual lung cancer risks. This expansion of risk factors in predictive risk models beyond the traditional epidemiologic data has been elegantly pursued in cardiovascular diseases, wherein biological assay data and some genetic variants had been added into risk score models for prediction of absolute risk of cardiovascular disease (15, 16). Similar efforts in cancer studies include addition of mammographic density data to the Gail model for breast cancer (17), inclusion of assay data for DNA repair capacity and mutagen sensitivity into the Spitz lung cancer risk models for former and current smokers (18), and recent combination of a panel of low-risk SNPs with demographic data to form a simple algorithm for lung cancer risk prediction (7).

It is known that a single gene would normally convey a small risk and confer a small cumulative addition to prediction (19). However, the combined effect of a number of genetic risk factors may be substantial, and indeed, the addition of this single SNP added more to the area under the curve than, for example, history of a previous non-lung malignancy, one of the LLP risk model parameters. The internal improvement in accuracy is encouraging, but the next step is to apply the combined LLP including the SEZ6L SNP on independent data sets for external validation.

We noted that the SEZ6L SNP variant (rs663048) was not among the top-ranked genes reported in a genome-wide association study identified as lung cancer susceptibility genes (20), nor was it statistically significant in a recent meta-analysis of lung genome-wide association study SNPs (21). However, biological evidence supports the importance of the SEZ6L gene in lung cancer oncogenesis. Frequent allelic losses in non-small-cell lung carcinomas have been reported on chromosomes 22q, a region where the SEZ6L gene resides, indicating the presence of tumor

suppressor gene(s) at the location (22). Also, the *SEZ6L* gene is a Met430Ile amino acid substitution that has been predicted to be functional by both SIFT and PolyPhen, suggesting a protein disturbing functional role (8). *SEZ6L* has also been found to be highly hypermethylated in primary colorectal tumor (23) and gastric carcinoma (24), linking the mutation to the progression of other neoplasms. Also, it should be noted that we observed the effect in three independent studies, two case-control studies with individual genotyping, and one DNA pooling study.

Risk prediction models may potentially play a significant role in future control of lung cancer, as prevention and screening could be targeted at those at increased risk of developing the disease thereby improving diagnosis, treatment, and survival (2). In addition, the usefulness of risk prediction models in the design of public screening program or clinical trial for diagnostic or treatment procedure has recently been shown (25). These necessitate accurate prediction of risk. Recent development in molecular and genetic epidemiology has thus encouraged the inclusion of validated susceptibility genes into existing risk models to improve their predictive ability. Apart from accurate prediction of absolute risks, the expanded risk

models may be reserved for individuals with intermediate risk, where the decision for classification into high or low risk group is ambiguous.

Our results show in principle how a genetic risk marker can be incorporated into an epidemiologic risk prediction model. In the example here, inclusion of the *SEZ6L* SNP shows a significant improvement in risk prediction. This work provides "proof of concept" in taking lung cancer risk prediction models to the next level.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

Roy Castle Lung Cancer Foundation, United Kingdom.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 07/01/2009; revised 10/23/2009; accepted 11/30/2009; published OnlineFirst 04/27/2010.

References

- Cassidy A, Myles JP, Liloglou T, Duffy SW, Field JK. Defining high-risk individuals in a population-based molecular-epidemiological study of lung cancer. *Int J Oncol* 2006;28:1295–301.
- Cassidy A, Duffy SW, Myles JP, Liloglou T, Field JK. Lung cancer risk prediction: a tool for early detection. *Int J Cancer* 2007;120:1–6.
- Field JK, Duffy SW. Lung cancer screening: the way forward. *Br J Cancer* 2008;99:557–62.
- Cassidy A, Myles JP, van Tongeren M, et al. The LLP risk model: an individual risk prediction model for lung cancer. *Br J Cancer* 2008;98:270–6.
- Field JK. Lung cancer risk models come of age. *Cancer Prev Res* 2008;1:226–8.
- Pepe MS, Janes HE. Gauging the performance of SNPs, biomarkers, and clinical factors for predicting risk of breast cancer. *J Natl Cancer Inst* 2008;100:978–9.
- Young RP, Hopkins RJ, Hay BA, et al. Lung cancer susceptibility model based on age, family history and genetic variants. *PLoS One* 2009;4:e5302.
- Gorlov IP, Meyer P, Liloglou T, et al. Seizure 6-like (*SEZ6L*) gene and risk for lung cancer. *Cancer Res* 2007;67:8406–11.
- Field JK, Smith DL, Duffy S, Cassidy A. The Liverpool Lung Project research protocol. *Int J Oncol* 2005;27:1633–45.
- Bamber D. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *J Math Psychol* 1975;12:387–415.
- Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157–72.
- Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 1997;16:965–80.
- Hosmer DW, Lemeshow S. A goodness-of-fit test for the multiple logistic regression model. *Commun Stat* 1980;9:1043–69.
- Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 1974;19:716–23.
- Wang TJ, Gona P, Larson MG, et al. Multiple biomarkers for the prediction of first major cardiovascular events and death. *N Engl J Med* 2006;355:2631–9.
- Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann Intern Med* 2009;150:65–72.
- Chen J, Pee D, Ayyagari R, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *J Natl Cancer Inst* 2006;98:1215–26.
- Spitz MR, Etzel CJ, Dong Q, et al. An expanded risk prediction model for lung cancer. *Cancer Prev Res* 2008;1:250–4.
- Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med* 2008;358:2796–803.
- Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008;452:633–7.
- Landi MT, Chatterjee N, Kai Y, et al. A comprehensive genome-wide association study of lung cancer. Proceedings of the 100th American Association of Cancer Research Annual Meeting. Denver (CO), Philadelphia (PA): AACR; 2009, LB139.
- Nishioka M, Kohno T, Takahashi M, et al. Identification of a 428-kb homozygously deleted region disrupting the *SEZ6L* gene at 22q12.1 in a lung cancer cell line. *Oncogene* 2000;19:6251–60.
- Suzuki H, Gabrielson E, Chen W, et al. A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat Genet* 2002;31:141–9.
- Kang GH, Lee S, Cho NY, et al. DNA methylation profiles of gastric carcinoma characterized by quantitative DNA methylation analysis. *Lab Invest* 2008;88:161–70.
- Duffy SW, Raji OY, Agbaje OF, Cassidy A, Field JK. Use of risk models in planning research and service programmes in CT screening for lung cancer. *Expert Rev Anticancer Ther* 2009;9:1467–72.