

Interplay between mutagen sensitivity and epidemiological factors in modulating lung cancer risk

Xifeng Wu*, Jie Lin, Carol J. Etzel, Qiong Dong, Olga Y. Gorlova, Qing Zhang, Christopher I. Amos and Margaret R. Spitz

Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Houston, TX

Few studies have assessed mutagen sensitivity and lung cancer (LC) risk associations in the context of multiple epidemiological risk factors. We evaluated mutagen sensitivity as a susceptibility marker and explored the interplay of the genetic marker and multiple epidemiologic risk factors in modulating LC risk. This largest case–control study included 977 newly diagnosed LC patients and 977 controls, matched by age, gender, ethnicity and smoking status. Cases exhibited significantly higher mutagen sensitivity than controls in bleomycin (0.76 vs. 0.62 breaks/cell, $p < 0.001$) and benzo[a]pyrene diol epoxide (BDPE) assays (0.70 vs. 0.61 breaks/cell, $p < 0.001$). Mutagen sensitivity also exhibited dose–response relationship with LC risk in quartile analysis (p for trend < 0.001). In smokers, history of emphysema, absence of hay fever history, LC in first-degree relatives, bleomycin sensitivity, and high pack-year were identified to be the top 5 significant risk factors in a stepwise logistic regression model (odds ratios (ORs) of 2.69, 1.91, 1.84, 1.73 and 1.67, respectively). Analyses of joint effects of risk factors showed that compared to the reference group, subjects with no exposure to any of the aforementioned risk factors, the ORs for exposure to 1, 2, 3, 4, 5 or more of the risk factors were 1.56, 2.39, 3.75, 6.74 and 17.39, respectively (p for trend < 0.001). This study strongly supports mutagen sensitivity as a predisposition factor for LC and demonstrates the importance of assessing multiple risk factors to comprehensively assess LC risk. This new integrative approach should facilitate identification of high-risk subgroups and has important implications in LC prevention.

© 2007 Wiley-Liss, Inc.

Key words: mutagen sensitivity; epidemiologic risk factors; lung cancer risk

Lung cancer (LC) is the leading cause of cancer-related death in the United States,¹ although it is estimated that $>80\%$ of LC can be attributed to cigarette smoking,² a variety of other exposures also contribute to LC risk, including occupational exposure to asbestos, dusts (wood dust, metal, concrete, fiberglass, etc.), polycyclic hydrocarbons, inorganic arsenic, diesel, exhaust, fumes.³ Previous studies have also suggested a link between family history of LC and the risk of smoking-related cancer and LC among first-degree relatives.^{4–6} Chronic respiratory diseases such as emphysema, hay fever, pneumonia, asthma and bronchitis have all been implicated in modifying LC risk.^{7–9} In never smokers, exposure to environmental tobacco smoke, or second-hand smoke, has been shown to confer increased risk of LC.^{10,11}

The fact that only a small fraction of smokers develops LC suggests that host susceptibility may be involved in LC risk.^{12–15} Host susceptibility, such as interindividual variation in DNA repair capacity to environmental carcinogens found in tobacco smoke, may predispose susceptible individuals to increased risk of LC. Hsu *et al.*¹⁶ developed a mutagen sensitivity assay that quantifies the frequency of chromatid breaks induced by mutagens in cultured lymphocytes *in vitro* as an integrated biomarker of mutagen sensitivity and an indirect measure of DNA repair capacity. Hsu *et al.*^{12,16} hypothesized that there may be a spectrum of mutagen sensitivity within the general population. Patients with chromosome breakage syndromes, such as ataxia telangiectasia and xeroderma pigmentosum are located at the extreme end of the spectrum. These patients exhibit high rates of spontaneous chro-

mosome breaks, increased susceptibility to induction of breaks by mutagens and increased cancer risk. Thus, the mutagen sensitivity assay identifies host DNA repair capacity and genomic stability. The mutagen sensitivity phenotype has been shown to increase risks of upper aerodigestive tract cancer,^{17,18} head and neck cancer,^{19–21} oral premalignant lesions²² and LC.^{13–15,23,24}

To date, there have been few studies that have comprehensively evaluated the association between LC risk in the context of multiple risk factors, including markers of genetic susceptibility and a panel of epidemiological risk factors. The interplay between genetic susceptibility and potential epidemiological covariates in predicting LC risk has not yet been addressed. Understanding how these components contribute to LC risk should effectively facilitate identification of high-risk subgroups, and therefore has important implications in LC prevention.

In this large case–control study of LC, we aimed to corroborate previous pilot data linking mutagen sensitivity and LC risk and further explored the interplay between mutagen sensitivity and a range of established and putative epidemiological risk factors in modulating LC risk.

Material and methods

Study population and epidemiologic data

The recruitment of cases and controls has been described previously.⁶ Briefly, LC patients were recruited at The University of Texas M.D. Anderson Cancer Center. They were all newly diagnosed and histologically confirmed LC enrolled before initiation of chemo- or radiation-therapy. The control subjects without a prior history of cancer (except nonmelanoma skin cancer) were recruited for participation from the Kelsey-Seybold clinics, the largest private multispecialty physician group in the Houston metropolitan area. The clinics consist of more than 23 clinic locations and over 300 physicians. The potential controls were identified by reviewing short survey forms distributed to patients visiting the Kelsey-Seybold clinics. The majority of control subjects are healthy individuals coming to the clinics for annual health check-ups.²⁵ There were no age, gender, ethnicity or stage restrictions on recruitment. Control subjects were frequency matched to the cases by age (± 5 years), gender, ethnicity and smoking status (never, former and current). The study design to recruit cases and controls has been proven to be valid and efficient for large-scale molecular epidemiological studies for cancer.²⁵ The response rates for both cases and controls have been 75%.⁹ All study participants signed a written informed consent detailing the research methods. Trained M.D. Anderson staff interviewers administered an epidemiologic

Grant sponsor: NCI; Grant numbers: CA 55769, CA 70907; Grant sponsor: DOD; Grant number: DAMD17-02-1-0706; Grant sponsor: Flight Attendant Medical Research Institute.

*Correspondence to: Department of Epidemiology, Unit 1340, The University of Texas M.D. Anderson Cancer Center, 1155 Hermann Pressler Blvd., Houston, TX 77030, USA.

Fax: 713-792-4657. E-mail: xwu@mdanderson.org

Received 11 August 2006; Accepted after revision 15 December 2006

DOI 10.1002/ijc.22588

Published online 8 February 2007 in Wiley InterScience (www.interscience.wiley.com).

questionnaire to study participants. Data collected included demographic characteristics, tobacco use history, occupation, exposure to a list of predefined exposures at work or from hobbies, prior medical history, and any family history of cancer (cancer in first-degree relatives). Collection of medical history has been previously described.⁹ Briefly, participants were asked whether they had ever been diagnosed by a physician with emphysema, hay fever, asthma, bronchitis and/or pneumonia. Methods applied to assess occupational exposure have also been described previously.^{14,26} An individual who had never smoked or had smoked less than 100 cigarettes in his or her lifetime was defined as a never smoker. An individual who had smoked at least 100 cigarettes in his or her lifetime but had quit more than 12 months before LC diagnosis (for cases) or before the interview (for controls) were coded as a former smoker. Current smokers were those who were currently smoking or had quit less than 12 months before diagnosis (for cases) or before the interview (for controls). Exposure to environment smoke, or second-hand smoke, was defined as having been exposed to someone else's cigarette smoke at home or at work on a regular basis, *i.e.* daily or weekly exposure.¹¹ Following the interview, a 40-mL blood sample was drawn into coded heparinized tubes for laboratory analyses. The research has been approved by the Institutional Review Boards (IRB) of the M.D. Anderson Cancer Center and Kelsey Seybold Clinics.

Mutagen sensitivity assay

As previously described,²⁴ 2 parallel blood cultures were set up when the blood was received. For the BPDE assay, on the 3rd day of incubation, the cultures were treated with 2 μ M BPDE for 24 hr. For the bleomycin assay, on the 4th day of incubation, the cultures were treated with 0.03 U/mL bleomycin for 5 hr. To arrest the cells at metaphase, 0.2 μ g/mL Colcemid was added to the culture 1 hr before harvest. Cells were then fixed and dropped onto clean microscopic slides, air-dried and stained with 4% giemsa solution. The number of breaks in 50 metaphases per sample were counted and expressed as the average number of breaks/cell.²⁷ Only frank chromatid breaks or exchanges were recorded; chromatid gaps or attenuated regions were disregarded. Each simple chromatid break was scored as 1 break. All measurements and counts were done blind with respect to case-control status.

Statistical analysis

Distributions in demographic variables (including gender, ethnicity and smoking status) between cases and controls were evaluated by the χ^2 test. Differences between cases and controls in age and self-reported pack years were tested using the Student's *t*-test. When the data distribution significantly deviated from normal, the Wilcoxon rank sum test was performed. Odds ratios (ORs) and 95% confidence intervals (CI) were calculated as estimates of relative risk. To calculate the ORs associated with mutagen sensitivity, the median number of breaks per cell (*b/c*) among the controls was used as the cutoff value. Breaks per cell values greater than this cutoff point were considered as the mutagen sensitive phenotype. To assess the dose-response relationships, we divided the *b/c* values among the controls in quartiles. Because data on mutagen sensitivity were not available for all 977 pairs of study subjects, unconditional logistic regression was used to assess the risk. We also performed conditional logistic regression with paired subjects. Because the results from conditional logistic regression and unconditional logistic regression were comparable, we only present results from the unconditional logistic regression in which a full set of mutagen sensitivity data was used. Possible confounding effects were controlled in a multivariate model. Because models adjusting for multiple confounding effects did not yield appreciable change in ORs (less than 10%), results of main effects of mutagen sensitivity were only adjusted by age, gender, ethnicity and smoking status. A trend test was performed to test for a linear trend in the ORs.

To assess LC risk in the context of multiple risk factors, we applied stepwise logistic regression to select important predictors among a panel of potential risk factors. In addition to logistic regression, we applied the recursive partitioning technique²⁸ to further discriminate low- and high-risk subgroups. The recursive partitioning method used for this analysis was derived from the methodology of Classification and Regression Tree (CART).^{28,29} A tree-based model is created by recursively partitioning the data and identifies effect modifications between variables that are less visible by traditional logistic regression. The classification tree divides the study sample into a number of homogeneous subgroups based on risk factors. Complete homogeneity of terminal nodes in tree modeling is ideal, meaning the split by certain variables creates terminal nodes containing only cases or controls. Thus, the numerical objective of partitioning is to make the contents of the terminal nodes as homogeneous as possible.

Multifactor dimensionality reduction (MDR)^{30–32} was carried out to assess interactions among multiple risk factors. In MDR, a set of *N* factors was selected and the *N* selected factors and their possible multifactor classes are represented in *N* dimensional space. The ratio for the number of cases to the number of controls was calculated within each multifactor class. Each multifactor class in *N* dimensional space was then labeled as "high risk" or as "low risk" based on whether the case to control ratio has exceeded a threshold of 1.0. Thus, a new variable incorporating information from several risk factors is defined and crossvalidation and permutation tests were used to assess the capability of this new variable to predict outcome risk. The crossvalidation consistency is defined as the percentage of the same combination of risk factors selected as the best model among crossvalidation data sets. The average testing accuracy is calculated as an average of the prediction accuracies across each of the crossvalidation subsets. In this study, we used 100-fold crossvalidation and 10,000-fold permutation testing. All statistical tests were at the significance level of 0.05.

Results

Mutagen sensitivity as a risk factor for LC

The demographic characteristics of the cases and controls are summarized in Table I. The study included 977 LC cases and 977 controls matched to cases by age (± 5 years), gender, ethnicity and smoking status. By study design, gender, ethnicity and smoking status were perfectly matched ($p = 1.00$ for all) and the mean ages of the cases and controls were 62.0 and 61.2 years, respectively ($p = 0.08$). In smokers, the self-reported median pack years were significantly higher in the cases than in the controls (45 for cases and 36 for controls, $p < 0.001$). Cases also self-reported to start smoking at an earlier age than controls (median age: 17 vs. 18, $p = 0.002$), to smoke a longer duration (median duration of smoking: 36 vs. 32, $p < 0.001$), to smoke more cigarettes per day (median cigarettes per day: 24 vs. 20, $p < 0.001$) and to stop smoking for a shorter period than the controls (median years of quitting: 11 vs. 14, $p < 0.001$) (Table I). There were also significant case-control differences in a panel of epidemiological factors, including family history of any cancer (cancer in first-degree relatives), family history of LC (LC in first-degree relatives), prior history of emphysema and hay fever, previous exposure to asbestos and wood dust (all $p < 0.01$) (Table I). The mean number of bleomycin-induced breaks per cell was significantly higher in cases than in controls (0.76 vs. 0.62, $p < 0.001$). Likewise, the mean number of BPDE-induced breaks per cell was significantly higher in cases as compared to the controls (0.70 vs. 0.61, $p = 0.003$) (Table I). Histology types of LC are listed in Table I. Adenocarcinoma comprised 47.6% of all the cases, followed by squamous cell carcinoma (24.1%).

In controls, there was an increasing trend in bleomycin sensitivity with age (p for trend = 0.002) (Table II). This was also evident in the cases, but the trend was not significant. Control current smokers exhibited significantly lower numbers of bleomycin-

TABLE I – SELECTED HOST CHARACTERISTICS BY CASE-CONTROL STATUS

Variables	Cases (N = 977)	Controls (N = 977)	p-value
Age, mean (SD)	62.0 (9.8)	61.2 (9.6)	0.08 ¹
Sex, n (%)			
Male	524 (53.6)	524 (53.6)	1.00 ²
Female	453 (46.4)	453 (46.4)	
Ethnicity			
Caucasians	803 (82.2)	803 (82.2)	1.00 ²
African Americans	111 (11.4)	111 (11.4)	
Hispanics	63 (6.5)	63 (6.5)	
Smoking status, N (%)			
Never	139 (14.2)	139 (14.2)	1.00 ²
Former	496 (50.8)	496 (50.8)	
Current	342 (35.0)	342 (35.0)	
Pack-year, median (range) ³	45 (0.26–256)	36 (0–189)	<0.001 ⁴
Age at start, median (range) ³	17 (5–58)	18 (6–54)	0.002 ⁴
Smoking duration, median (range) ³	36 (2–68)	32 (1–67)	<0.001 ⁴
Years of cessation, median (range) ⁵	11 (1–54)	14 (1–53)	<0.001 ⁴
Cigarettes/day, median (range) ³	24 (1–80)	20 (0–80)	<0.001 ⁴
Family history of cancer, N (%)			
Yes	665 (68.4)	603 (61.8)	0.002 ²
No	308 (31.7)	373 (38.2)	
Family history of lung cancer, N (%)			
Yes	199 (20.5)	129 (13.2)	<0.001 ²
No	774 (79.6)	847 (86.8)	
Emphysema			
Yes	162 (17.0)	63 (6.5)	<0.001 ²
No	790 (83.0)	914 (93.6)	
Hay fever			
Yes	140 (14.9)	227 (23.3)	<0.001 ²
No	803 (85.2)	749 (76.7)	
Asbestos exposure			
Yes	364 (38.9)	290 (29.9)	<0.001 ²
No	571 (61.1)	680 (70.1)	
Wood dust exposure			
Yes	122 (13.2)	78 (8.2)	<0.001 ²
No	801 (86.8)	875 (91.8)	
Mutagen sensitivity, mean (SD), N			
Bleomycin (b/c)	0.76 (0.41), 933	0.62 (0.33), 960	<0.001 ¹
BPDE (b/c)	0.70 (0.35), 434	0.61 (0.37), 509	0.0003 ¹
Histology			
Adenocarcinoma	459 (47.6)		
Squamous	232 (24.1)		
Large cell	46 (4.8)		
Small cell	52 (5.4)		
Non-small cell	115 (11.9)		
Other	60 (6.2)		

¹Student's *t*-test. ² χ^2 test. ³Ever smokers only. ⁴Wilcoxon rank sum test. ⁵Former smokers only.

induced mean breaks per cell than did never or former smokers (0.55 vs. 0.64, $p = 0.03$; 0.55 vs. 0.66, $p < 0.001$) (Table II). There was no such trend among the cases. Smoking variables, including pack year of smoking, age at starting, duration of smoking, years of quitting, and numbers of cigarettes per day, did not significantly modify bleomycin sensitivity profiles (Table II). Interestingly, controls reporting a history of emphysema showed higher sensitivity than those who had no emphysema history (0.72 vs. 0.61, $p = 0.02$). Cases reporting a history of hay fever exhibited lower sensitivity than cases without a hay fever history (0.65 vs. 0.72, $p = 0.003$) (Table II).

Mean BPDE-induced breaks were highest in never smokers and lowest in current smokers for both cases and controls, although the differences were significant only in the controls (0.54 vs. 0.82, $p < 0.001$) (Table II). Compared to never smokers, control heavy smokers (pack year ≥ 36) showed significantly lower mean number of BPDE-induced breaks per cell (0.62 vs. 0.82, $p = 0.005$). Control light smokers (pack year < 36) also showed significantly lower mean breaks per cell than never smokers (0.58 vs. 0.82, $p < 0.001$).

Other smoking variables showed the same trend as pack year in modifying BPDE sensitivity profiles, but the effects were not significant (Table II). Moreover, controls, but not cases with a posi-

tive family history of LC, showed higher number of breaks than those who had no family history (0.74 vs. 0.59, $p = 0.003$). Histology and tumor stage did not influence mutagen sensitivity in the cases (data not shown).

To estimate LC risk associated with mutagen sensitivity, we dichotomized the mean number of breaks per cell based on the median value in controls. We observed a 1.63-fold increased risk (95% CI: 1.36–1.97) for those showing bleomycin sensitivity and a 1.85-fold increased risk (95% CI: 1.42–2.42) for those with BPDE sensitivity (Table III). In quartile analysis, using the lowest quartile (subjects showing the least sensitivity to bleomycin challenge) as the reference group, the ORs for the 2nd, 3rd and 4th quartiles were 1.28 (95% CI: 0.97–1.70), 1.35 (95% CI: 1.04–1.77), and 2.39 (95% CI: 1.84–3.09), respectively (p for trend < 0.001) (Table III). Similar trends were observed for BPDE sensitivity and LC risk (Table III) when using the first quartile as the reference group (2nd quartile: OR = 1.89, 95% CI: 1.24–2.87; 3rd quartile: OR = 2.58, 95% CI: 1.72–3.88; 4th quartile: OR = 2.81, 95% CI: 1.87–4.22; p for trend < 0.001) (Table III).

We next assessed the joint effect of mutagen sensitivity (bleomycin sensitivity) with other epidemiological risk factors (Table IV). In this and subsequent analyses, we used bleomycin sensitivity as the mutagen sensitivity variable since BPDE data were only

TABLE II – MUTAGEN SENSITIVITY BY HOST CHARACTERISTICS IN CASES AND CONTROLS

	Case			Control			
	<i>N</i>	Mean (SD)	<i>p</i> -value	<i>N</i>	Mean (SD)	<i>p</i> -value	
Bleomycin							
Overall	933	0.76 (0.41)		960	0.62 (0.33)		
Sex							
Male	507	0.77 (0.40)		516	0.63 (0.35)		
Female	426	0.74 (0.42)	0.45	444	0.61 (0.31)	0.22	
Age							
<50	108	0.68 (0.36)		115	0.53 (0.30)		
50–60	246	0.75 (0.45)		261	0.59 (0.33)		
61–70	360	0.77 (0.40)		397	0.65 (0.34)		
71–80	199	0.77 (0.40)		164	0.66 (0.32)		
≥80	20	0.83 (0.47)	0.23	23	0.68 (0.22)	0.002	
Smoking status							
Never	133	0.73 (0.38)	Ref.	137	0.64 (0.30)	Ref.	
Former	476	0.77 (0.41)	0.65	490	0.66 (0.36)	0.73	Ref.
Current	324	0.74 (0.42)	0.99	333	0.55 (0.29)	0.03	<0.001
Pack-year ¹							
Never	133	0.73 (0.38)	Ref.	137	0.64 (0.30)	Ref.	
<36	571	0.77 (0.42)	0.77	624	0.62 (0.34)	0.93	Ref.
≥36	229	0.75 (0.41)	0.98	199	0.62 (0.32)	0.97	1.00
Age at start ¹							
≥18	369	0.76 (0.43)		426	0.62 (0.33)		
<18	431	0.76 (0.41)	0.84	397	0.62 (0.35)	0.94	
Smoking duration ¹							
<32	297	0.76 (0.41)		392	0.63 (0.35)		
≥32	503	0.76 (0.42)	0.80	431	0.61 (0.32)	0.58	
Cigarettes/day ¹							
<20	140	0.74 (0.40)		221	0.62 (0.32)		
≥20	660	0.76 (0.42)	0.50	602	0.62 (0.34)	0.99	
Years of cessation ²							
<14	270	0.76 (0.43)		235	0.65 (0.34)		
≥14	206	0.79 (0.38)	0.45	255	0.67 (0.37)	0.47	
Family history of cancer							
No	295	0.77 (0.41)		369	0.61 (0.32)		
Yes	634	0.75 (0.41)	0.46	590	0.63 (0.34)	0.40	
Family history of lung cancer							
No	740	0.76 (0.40)		830	0.62 (0.33)		
Yes	189	0.75 (0.45)	0.95	129	0.63 (0.36)	0.62	
Emphysema							
No	755	0.76 (0.41)		900	0.61 (0.32)		
Yes	155	0.74 (0.41)	0.66	60	0.72 (0.44)	0.02	
Hay fever							
No	768	0.77 (0.42)		736	0.62 (0.32)		
Yes	134	0.65 (0.33)	0.003	223	0.63 (0.35)	0.58	
Wood dust exposure							
No	768	0.75 (0.41)		859	0.62 (0.33)		
Yes	116	0.78 (0.43)	0.55	77	0.60 (0.31)	0.60	
Asbestos exposure							
No	547	0.74 (0.40)		670	0.61 (0.31)		
Yes	347	0.77 (0.40)	0.21	283	0.64 (0.38)	0.39	
BPDE							
Overall	434	0.70 (0.35)		509	0.61 (0.37)		
Sex							
Male	228	0.69 (0.35)		275	0.59 (0.34)		
Female	206	0.71 (0.35)	0.63	234	0.65 (0.41)	0.06	
Age							
<50	44	0.64 (0.37)		67	0.73 (0.49)		
50–60	115	0.69 (0.33)		149	0.62 (0.40)		
61–70	189	0.72 (0.37)		201	0.56 (0.32)		
71–80	82	0.68 (0.31)		81	0.61 (0.30)		
≥80	4	0.70 (0.20)	0.56	11	0.82 (0.36)	0.006	
Smoking status							
Never	43	0.73 (0.31)	Ref.	45	0.82 (0.46)	Ref.	
Former	222	0.71 (0.35)	0.99	271	0.64 (0.40)	0.005	Ref.
Current	169	0.68 (0.35)	0.82	193	0.54 (0.27)	<0.001	0.013
Pack-year							
Never	43	0.73 (0.31)	Ref.	45	0.82 (0.46)	Ref.	
<36	280	0.69 (0.34)	0.92	337	0.58 (0.35)	<0.001	Ref.
≥36	111	0.70 (0.37)	0.98	127	0.62 (0.37)	0.005	0.71
Age at start ¹							
≥18	170	0.70 (0.38)		235	0.57 (0.37)		
<18	221	0.70 (0.33)	0.97	229	0.62 (0.34)	0.19	
Smoking duration ¹							
<32	133	0.69 (0.39)		205	0.61 (0.40)		
>32	258	0.70 (0.33)	0.86	259	0.58 (0.32)	0.38	

TABLE II – MUTAGEN SENSITIVITY BY HOST CHARACTERISTICS IN CASES AND CONTROLS (CONTINUED)

	Case			Control		
	N	Mean (SD)	p-value	N	Mean (SD)	p-value
Cigarettes/day ¹						
<20	66	0.66 (0.32)	0.36	95	0.63 (0.37)	0.32
≥20	325	0.70 (0.36)		369	0.59 (0.35)	
Years of cessation ²						
<14	130	0.71 (0.36)	0.90	149	0.65 (0.43)	0.42
≥14	92	0.71 (0.35)		122	0.61 (0.36)	
Family history of cancer						
No	126	0.72 (0.37)	0.48	193	0.60 (0.38)	0.68
Yes	305	0.69 (0.34)		315	0.62 (0.37)	
Family history of lung cancer						
No	345	0.70 (0.34)	0.55	438	0.59 (0.35)	0.003
Yes	86	0.68 (0.37)		70	0.74 (0.45)	
Emphysema						
No	340	0.70 (0.35)	0.25	471	0.62 (0.38)	0.08
Yes	86	0.66 (0.27)		38	0.51 (0.29)	
Hay fever						
No	365	0.70 (0.33)	0.07	401	0.62 (0.38)	0.32
Yes	58	0.62 (0.34)		108	0.58 (0.32)	
Wood dust exposure						
No	363	0.68 (0.33)	0.09	472	0.62 (0.37)	0.35
Yes	55	0.76 (0.37)		36	0.56 (0.35)	
Asbestos exposure						
No	242	0.70 (0.34)	0.81	354	0.64 (0.39)	0.06
Yes	173	0.69 (0.33)		153	0.57 (0.32)	

Age-dichotomized by control median.

¹Dichotomized by control median (ever smokers only).–²Dichotomized by control median (former smokers only).

available in half of the study participants. Compared to subjects with none of the risk factors, mutagen sensitive subjects were at 2.88-fold increased risk (95% CI: 1.22–6.84); risk increased to 4.63-fold (95% CI: 1.95, 10.95) for mutagen sensitive subjects who also reported heavy pack-years (Table IV). Adding history of emphysema resulted in an OR of 6.51 (95% CI: 1.76, 24.06). The risk for mutagen sensitive subjects who reported a history of emphysema, but no hay fever, and heavy pack-years (pack-years > 36) was 11.65 (95% CI: 4.27, 31.75). The highest at-risk group included subjects who were mutagen sensitive, had a history of emphysema but not hay fever, were heavy smokers, had a positive family history and reported exposures to asbestos and wood dust. There were 17 cases but no controls in this group (Table IV).

LC risk is modified by multiple risk factors

A stepwise backward elimination procedure was performed to identify significant predictors of LC. Again, in this and subsequent analyses, we used bleomycin sensitivity as a mutagen sensitivity variable as BPDE data were only available in half of the study participants.

In ever smokers (Table V), the model started with all matching variables, mutagen sensitivity (*i.e.* bleomycin sensitivity), history of emphysema, family history of LC, asbestos exposure, exposure to wood dust and pack-years smoked (categorized by tertiles: <26, 26–47, ≥47). Mutagen sensitivity remained a significant predictor of LC risk with an OR of 1.73 (95% CI: 1.40–2.15). History of emphysema (OR: 2.69; 95% CI: 1.92–3.78), family history of LC (OR: 1.84; 95% CI: 1.39–2.44), higher pack-years smoked (OR: 1.67; 95% CI: 1.33–2.08), wood dust exposure (OR: 1.52; 95% CI: 1.07–1.73), asbestos exposure (OR: 1.38; 95% CI: 1.09–1.73), all remained in the model as significant predictors of LC risk and were all associated with increased LC risk. History of hay fever remained in the model and was protective against LC with an OR of 0.52 (95% CI: 0.39–0.70) (the OR for absence of hay fever was 1.91, 95% CI: 1.43–2.54) (Table V).

In never smokers, the model started with the full set of variables above except that pack-years was substituted for second-hand smoke exposure and emphysema was not included as none of the never smokers had a prior history of emphysema. Only second-

TABLE III – MUTAGEN SENSITIVITY AND LUNG CANCER RISK

	Cases N (%)	Controls N (%)	Multivariate OR*
Bleomycin (breaks/cell)			
Median			
<0.56	333 (35.7)	459 (47.8)	1.00 (ref.)
≥0.56	600 (64.3)	501 (52.2)	1.63 (1.36–1.97)
Quartile			
<0.38	161 (17.3)	250 (26.0)	1.00 (ref.)
0.38–0.56	172 (18.4)	209 (21.8)	1.28 (0.97–1.70)
0.56–0.78	232 (24.9)	263 (27.4)	1.35 (1.04–1.77)
≥0.78	368 (39.4)	238 (24.8)	2.39 (1.84–3.09)
p for trend			<0.001
BPDE (breaks/cell)			
Median			
<0.54	153 (35.3)	256 (50.3)	1.00 (ref.)
≥0.54	281 (64.8)	253 (49.7)	1.85 (1.42–2.42)
Quartile			
<0.36	52 (12.0)	126 (24.8)	1.00 (ref.)
0.36–0.54	101 (23.3)	130 (25.5)	1.89 (1.24–2.87)
0.54–0.80	137 (31.6)	127 (25.0)	2.58 (1.72–3.88)
≥0.80	144 (33.2)	126 (24.8)	2.81 (1.87–4.22)
p for trend			<0.001

*Adjusted by age, gender, ethnicity and smoking status.

hand smoke exposure remained in the model with an OR of 2.15 (95% CI: 1.02–4.52) (Table V).

CART analysis resulted in 6 terminal nodes (Fig. 1). Again, we used bleomycin sensitivity as the mutagen sensitivity variable. Consistent with stepwise regression analysis, the significant splitting variables were history of emphysema, mutagen sensitivity, LC in first-degree relatives, history of hay fever and pack-years smoked. As shown in Figure 1, terminal nodes 1, 2 and 4 defined groups with low risk of LC. Compared to the reference group, node 1 (ever smokers who had no history of emphysema, who exhibited no mutagen sensitivity, who had no family history of LC), the ORs for node 2 (ever smokers who had no history of emphysema, who exhibited no mutagen sensitivity, who had a family history of LC and who smoked 36 or less pack years) and node 4 (ever smokers who had no history of emphysema, who were mutagen sensitive and who had a prior history of hay fever) were 1.0

TABLE IV – JOINT EFFECTS OF MUTAGEN SENSITIVITY WITH SELECTED EPIDEMIOLOGICAL RISK FACTORS IN LUNG CANCER RISK IN SMOKERS

Risk factors ¹	Cases	Controls	Multivariate OR*
None	7 (1.44)	26 (5.9)	1.00 (ref.)
MS	158 (32.4)	203 (46.0)	2.88 (1.22–6.84)
MS, PY	239 (49.1)	189 (42.9)	4.63 (1.95–10.95)
MS, ESM, PY	11 (2.3)	6 (1.4)	6.51 (1.77–24.06)
MS, ESM, NHF, PY	55 (11.3)	17 (3.9)	11.65 (4.27–31.75)
MS, ESM, NHF, PY, FH, WD, ASBS	17 (3.5)	0 (0.0)	N/A
<i>p</i> -trend			<0.001

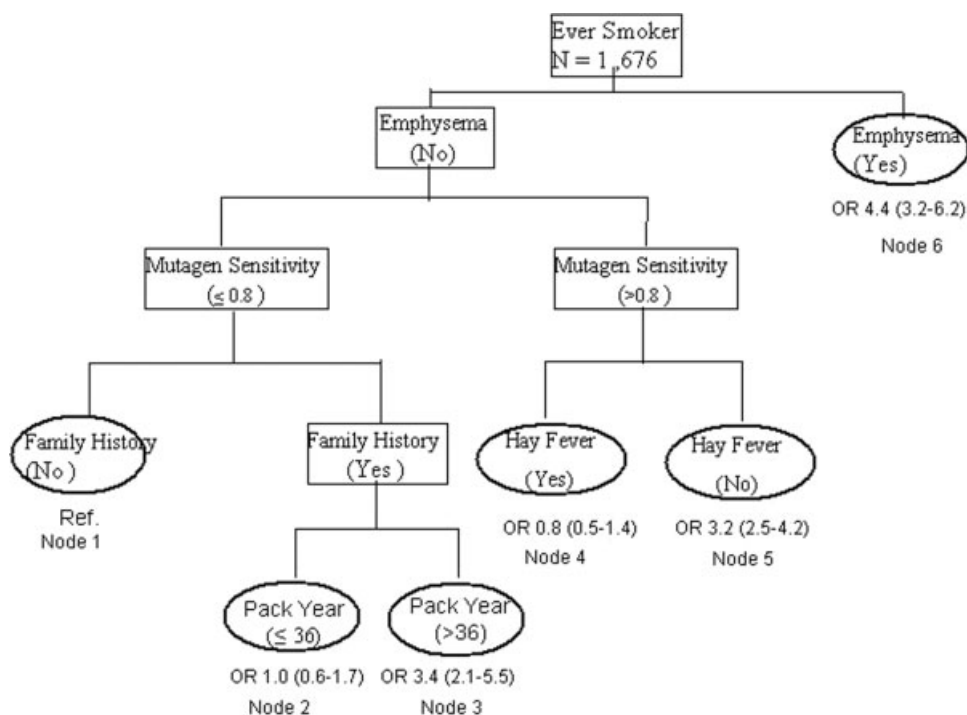
*Adjusted by matching variables.

¹MS, mutagen sensitivity (bleomycin sensitivity); ESM, emphysema; NHF, no hay fever; PY, pack year (dichotomized by median in controls, i.e. 36); FH, family history of lung cancer; WD, wood dust exposure; ASBS, asbestos exposure.

TABLE V – MULTIVARIATE RISK ASSESSMENT MODEL USING STEPWISE LOGISTIC REGRESSION

Variables	Odds ratio (95% CI)*	<i>p</i> -Value
Ever smokers		
Emphysema	2.69 (1.92–3.78)	<0.001
No hay fever	1.91 (1.43–2.54)	<0.001
Family history of lung cancer	1.84 (1.39–2.44)	<0.001
Mutagen sensitivity	1.73 (1.40–2.15)	<0.001
Pack-year > 47	1.67 (1.33–2.08)	<0.001
Wood dust exposure	1.52 (1.07–2.15)	0.02
Asbestos exposure	1.38 (1.09–1.73)	0.006
Never smokers		
Second-hand smoke	2.15 (1.02–4.52)	0.04

*Obtained by stepwise logistic regression starting with matching variables (age, gender and ethnicity) and all 7 risk factors in the model: mutagen sensitivity, history of emphysema, history of hay fever, family history of lung cancer, exposure to wood dust, asbestos, pack-year smoked (categorized by tertiles in controls: <26, 26–47, >47) or second-hand smoke for never smokers.

**FIGURE 1 – Recursive partitioning of risk factors in ever smokers.**

(95 % CI: 0.6–1.67) and 0.8 (95% CI: 0.5–1.4), respectively. Terminal nodes 3 (ever smokers who had no history of emphysema, who exhibited no mutagen sensitivity, who had a family history of LC and who smoked more than 36 pack-years) and 5 (ever smokers who had no history of emphysema, who were mutagen sensitive and who had no prior history of hay fever) characterized medium risk groups with moderate increased risk with ORs of 3.4

(95 % CI: 2.1–5.5), and 3.2 (95 % CI: 2.5–4.2), respectively. The highest at-risk group was defined by terminal node 6, ever smokers who had a history of emphysema (OR: 4.4; 95% CI: 3.2–6.2).

The results of the MDR analysis for each risk combination are presented in Table VI. The model including pack-years of smoking was the best single factor model for predicting LC risk with average testing accuracy of 0.5728 and cross validation consis-

TABLE VI – MULTIFACTOR DIMENSIONALITY REDUCTION (MDR) MODEL RESULTS

Number of risk factors	Best model*	Average crossvalidation consistency	Average testing accuracy	Permutation test <i>p</i> -value
1	PY	100/100	0.5728	<0.001
2	MS, HF	59.5/100	0.5218	0.653
3	MS, HF, ESM	100/100	0.6071	<0.001
4	MS, ESM, FH, PY	100/100	0.6196	<0.001
5	MS, ESM, FH, PY, HF	100/100	0.6317	<0.001
6	MS, ESM, FH, PY, HF, ASBS	100/100	0.6192	<0.001
7	MS, ESM, FH, PY, HF, ASBS, WD	100/100	0.6144	<0.001

*Variable abbreviations are the same as Table 4.

TABLE VII – JOINT EFFECTS OF RISK FACTORS IN LUNG CANCER RISK IN EVER SMOKERS

No. of risk factors ¹	Cases	Controls	Multivariate OR*
0	7 (1.0)	26 (3.3)	1.00 (ref.)
1	67 (9.1)	157 (19.7)	1.56 (0.64–3.78)
2	159 (21.6)	248 (31.2)	2.39 (1.01–5.65)
3	229 (31.1)	240 (30.1)	3.75 (1.59–8.84)
4	167 (22.7)	101 (12.7)	6.74 (2.81–16.21)
≥5	108 (14.7)	26 (3.3)	17.39 (6.75–44.80)
<i>p</i> -trend			<0.001

*Adjusted by age, gender and ethnicity.

¹Risk factors include bleomycin sensitivity, asbestos exposure, wood dust exposure, history of emphysema, no history of hay fever, family history of lung cancer, pack year.

tency of 100%. Among two-factor models, the combination of mutagen sensitivity and history of hay fever was the best model, but the cross validation consistency was low (59.5%) and the permutation test was not significant. When 3 factors were considered, the combination of mutagen sensitivity, history of hay fever and history of emphysema was the best model with 100% cross validation consistency and an average testing accuracy of 0.6071. When risk factors were considered 4 at a time, the model containing mutagen sensitivity, emphysema, family history of LC and pack-years was the best with 100% cross validation consistency and an average testing accuracy of 0.6196. The average test accuracy further increased to 0.6217 in a five-factor model including mutagen sensitivity, emphysema history, family history of LC, pack-years and history of hay fever (cross validation consistency reached 100%). Considering 6 and 7-factor models did not improve testing accuracy. Consistent with the CART analysis above, which showed the low and high-risk groups as defined by the interaction between 5 risk factors (emphysema history, mutagen sensitivity, family history of LC, history of hay fever and pack-year), the MDR analysis identified a model including the same set of 5 variables to be the best model to predict LC risk.

Finally, we assessed LC risk in the context of the multiple risk factors identified as significant in multivariate analyses. In ever smokers, these risk factors were mutagen sensitivity, history of emphysema, no prior history of hay fever, exposure to wood dust, asbestos, family history of LC and pack-years smoked. Compared to the reference group, subjects who had none of these factors, the ORs in the presence of 1, 2, 3, 4, 5 and more of the risk factors were 1.56 (95% CI: 0.64–3.78), 2.39 (95% CI: 1.01–5.65), 3.75 (95% CI: 1.59–8.84), 6.74 (95% CI: 2.81–16.21) and 17.39 (95% CI: 6.75–44.80) (Table VII), respectively. A significant trend in OR was apparent with increasing number of risk factors (*p* for trend <0.001) (Table VII). This analysis was not performed for never smokers as only exposure to second-hand smoke was identified to be significant.

Discussion

Our study is the largest to comprehensively evaluate LC risk as determined by markers of genetic susceptibility and a panel of epidemiological risk factors. Our results corroborate previous findings

of mutagen sensitivity as a predisposition factor for LC^{13–15,23,24} and also suggest a substantial interplay between mutagen sensitivity and other established LC risk factors in modulating risk.

Previous studies have prospectively shown that mutagen sensitivity at baseline in head-and-neck cancer is a predictor of the development of secondary malignant tumors³³ and multiple primary malignancies.³⁴ We and others have also demonstrated that mutagen sensitivity remained a significant predictor of cancer risk after adjusting for confounding effects of age, sex, smoking habits or occupational exposure to carcinogens.^{13,14,17,23} Wu *et al.*²² reported increased risk of developing oral premalignant lesions in those exhibiting elevated sensitivity to bleomycin and BPDE. In that study, none of the subjects had a cancer diagnosis. Moreover, prospective epidemiological analyses showed that chromosomal aberration frequency predicts subsequent cancer risk in healthy subjects.^{35–37} Recently, we validated mutagen sensitivity as a valid genetic marker by showing high heritability of mutagen sensitivity in a twin study.³⁸ Our current data showed that in smokers, mutagen sensitivity remained a significant predictor in the risk assessment model after stepwise selection among a panel of potential risk factors. In CART analysis, mutagen sensitivity was identified to be a significant risk factor (*i.e.* created a significant split) and the interaction between mutagen sensitivity and other factors defined high vs. low risk subgroups. In MDR analysis, the best 2, 3, 4 and 5-factor models all included mutagen sensitivity. The agreement in results obtained from different approaches strongly supports the utility of mutagen sensitivity as a valid biomarker for LC risk.

The mutagen sensitivity assay has been proven to have high repeatability and reliability. In the study of Cloos *et al.*,³⁹ the mean interindividual variation (0.35 breaks per cell) exceeded mean intraindividual variation (0.08 breaks per cell). We previously performed the assays in duplicated samples (samples from the same subjects) to assess measurement error in mutagen sensitivity assays.³⁸ The correlations between duplicated samples were high with the correlation coefficients for bleomycin and BPDE assays of 0.70 (*p* < 0.001; *N* = 25) and 0.75 (*p* < 0.001; *N* = 63), respectively, with an average coefficient of variation for bleomycin and BPDE assays of 19.7% and 16.4%, respectively. In a multicenter, case–control mutagen sensitivity analysis, Cloos *et al.*²¹ found that there were no differences across institutions in the distribution of bleomycin sensitivity in patients and controls.

In both bleomycin and BPDE assays, our results suggested that controls who were current smokers had the lowest chromosome breaks per cell as compared to never or former smokers. Also, compared to light smokers, heavy smokers in controls exhibited the lowest BPDE-induced chromosome breaks per cell. The observation that current smokers exhibit lower levels of chromosome breaks may suggest a more proficient DNA repair capacity because DNA damage caused by tobacco carcinogens may stimulate DNA repair.^{23,40} The observation of increasing bleomycin sensitivity with age is consistent with previous investigations,^{15,22} supporting the notion of increased susceptibility to environmental carcinogens in older people since both immune function and DNA repair efficiency decrease with age. However, the pattern was not clear for BPDE sensitivity. A smaller sample size of BPDE data may be part of the explanation. One interesting observation of this

study is that history of chronic respiratory diseases also modifies the mutagen sensitivity profile. Emphysema is a chronic inflammatory condition that results in fixed narrowing of small airways and alveolar wall destruction.⁴¹ Inflammatory cells release various growth factors, promote angiogenesis, activate stroma and generate free radicals, cellular processes that have been demonstrated to cause DNA damage. Thus, patients with emphysema, an inflammatory disease, may have accumulated more DNA repair errors than healthy people, and consequently exhibit increased mutagen sensitivity. It is also interesting to note that subjects with emphysema (all ever smokers) still exhibited higher bleomycin sensitivity than those without emphysema (all never smokers), suggesting that the detrimental effects of emphysema or genetic defects underlying emphysema on host bleomycin sensitivity is probably substantial such that even control smokers who are proficient in DNA repair capacity due to stimulation of DNA repair by carcinogens still exhibited high mutagen sensitivity in the presence of emphysema. We also observed low mutagen sensitivity in LC cases who reported a prior history of hay fever. We speculate that either the genetic factors underlying hay fever or the anti-inflammatory agents used to treat hay fever may have protected cells from DNA damage caused by the inflammatory reaction. Because of the small sample size of exposed subjects, further investigation is needed to confirm the findings.

Our data did not support an association between mutagen sensitivity and LC risk in never smokers. The sample size of never smokers in our study may not be large enough to detect an effect and further study with a larger sample size of never smokers is needed to explore the putative association.

This study underscores the importance to incorporate chronic inflammatory lung diseases to assess LC risk. In the CART analysis (Fig. 1), prior history of emphysema produced the first significant split that defined subgroups at contrasting risk of LC. It is interesting to note that the subgroup with a positive history of emphysema did not split further, suggesting minimal modulating effects from other factors in the presence of emphysema. There has been compelling evidence supporting the hypothesis that chronic inflammation contributes to cancer development and that inflammation plays an important role in malignant initiation and progression. However, the complex mechanism by which inflammation leads to cancer initiation is not yet fully understood and the specific cellular and molecular pathways that lead to inflammatory responses to particular microbes with malignant transformation vary by target organ and microbe.⁴² Cellular reactions to inflammation include releasing of cytokines, chemokines and free radicals, which contribute to malignant transformation by peroxidizing lipids, inducing genetic mutations that may cause direct DNA damage and inhibition of DNA repair. The impaired DNA

repair machinery in inflammatory cells may be most susceptible to carcinogens and consequently cancer development.

In this study, in addition to the traditional logistic regression, we also applied 2 innovative statistical approaches (CART and MDR) to identify higher-order interactions. Logistic regression is a parametric statistical approach to model the relationship between discrete risk factors and discrete phenotypes (*e.g.*, cancer). However, because the number of possible interaction terms grows exponentially as the number of risk factors increases, logistic regression is limited in its ability to deal with multiple higher-order interactions. CART and MDR are 2 commonly used non-parametric statistical methods to better model higher-order interactions. CART uses a binary recursive partitioning method to generate a tree-structured node with binary splits. All the daughter nodes are more homogeneous (significantly more cancer cases or more controls) than the original node (roughly half cases and half controls), resulting in significantly increased or reduced risks of LC for each terminal node compared to the original node. MDR models high order interactions by reducing the degree of freedom necessary for modeling, *i.e.*, by defining a single variable that incorporates information from several risk factors that can be divided into high-risk and low-risk combinations. This new variable can be evaluated for its ability to classify and predict disease status using cross validation and permutation testing. In our analysis, both methods resulted in similar joint effects among these risk factors and categorized individuals into subgroups with varying degrees of risk for LC. Compared with single risk factors from logistic regression analysis, higher-order interactions yielded higher power of risk prediction.

LC offers an excellent system to comprehensively study the interplay between genetic predisposition and environmental risk factors. In this study, we demonstrated that mutagen sensitivity remains as a significant predisposition factor of LC in the presence of a panel of epidemiological risk factors. Of particular interest is the joint effect of mutagen sensitivity and these factors. Specifically, in CART analysis (Fig. 1), emphysema, mutagen sensitivity, family history of LC, smoking, and history of hay fever jointly defined high and low risk subgroups, suggesting a substantial interplay between mutagen sensitivity and established epidemiological risk factors in modulating LC risk. To date, we know of no published studies that have included such a broad range of risk factors in combination with mutagen sensitivity to estimate LC risk. This study underscores the importance of assessing multiple factors including relevant genetic markers, history of inflammation conditions and deleterious exposures to comprehensively assess LC risk. This new integrative approach also should effectively facilitate identification of high-risk subgroups and therefore has important implications in LC prevention.

References

1. Jemal A, Tiwari RC, Murray T, Ghafoor A, Samuels A, Ward E, Feuer EJ, Thun MJ, American Cancer Society. Cancer statistics. *CA Cancer J Clin* 2004;54:8–29.
2. Peto R, Lopez AD, Boreham J, Thun M, Heath C, Doll R. Mortality from smoking worldwide. *Br Med Bull* 1996;52:12–21.
3. Alberg AJ, Brock MV, Samet JM. Epidemiology of lung cancer: looking to the future. *J Clin Oncol* 2005;23:3175–85.
4. Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H. Increased familial risk for lung cancer. *J Natl Cancer Inst* 1986;76:217–22.
5. Schwartz AG, Yang P, Swanson GM. Familial aggregation of lung cancer among nonsmokers and their relatives. *Am J Epidemiol* 1996;144:554–62.
6. Etzel CJ, Amos CI, Spitz MR. Risk for smoking-related cancer among relatives of lung cancer patients. *Cancer Res* 2003;63:8531–35.
7. Skillrud DM, Offord KP, Miller RD. Higher risk of lung cancer in Chronic obstructive pulmonary disease: a prospective, matched, controlled study. *Ann Int Med* 1986;105:503–7.
8. Mayne ST, Buenconsejo J, Janerich DT. Previous lung disease and risk of lung cancer among men and women nonsmokers. *Am J Epidemiol* 1999;149:13–20.
9. Schabath MB, Delclos GL, Martynowicz MM, Greisinger AJ, Lu C, Wu X, Spitz MR. Opposing effects of emphysema, hay fever, and selected genetic variants on lung cancer risk. *Am J Epidemiol* 2005;161:412–22.
10. Hu J, Mao Y, Dryer D, White K. Risk factors for lung cancer among Canadian women who have never smoked. *Cancer Detect Prev* 2002;26:129–38.
11. Gorlova OY, Zhang Y, Schabath MB, Lei L, Zhang Q, Amos CI, Spitz MR. Never smokers and lung cancer risk: a case-control study of epidemiological factors. *Int J Cancer* 2006;118:1798–804.
12. Hsu TC, Spitz MR, Schantz SP. Mutagen sensitivity: a biologic marker of cancer susceptibility. *Cancer Epidemiol Biomarkers Prev* 1991;1:83–9.
13. Spitz MR, Hsu TC, Wu XF, Fueger JJ, Amos CI, Roth JA. Mutagen sensitivity as a biologic marker of wood-dust exposure, mutagen sensitivity, and lung-cancer risk. *Cancer Epidemiol Biomarkers Prev* 1995;4:99–103.
14. Wu X, Delclos GL, Annegers FJ, Bondy ML, Honn SE, Henry B, Hsu TC, Spitz MR. A case-control study of wood-dust exposure, mutagen sensitivity, and lung-cancer risk. *Cancer Epidemiol Biomarkers Prev* 1995;4:583–8.
15. Zheng YL, Lofferedo CA, Yu Z, Jones RT, Krasna MJ, Alberg AJ, Yung R, Perlmutter D, Enewold L, Harris CC, Shields PG. Bleomy-

- cin-induced chromosome breaks as a risk marker for lung cancer: a case-control study with population and hospital controls. *Carcinogenesis* 2003;24:169-274.
16. Hsu TC, Johnston DA, Cherry LM, Ramkissoon D, Schantz SP, Jessup JM, Winn RJ, Shirley L, Furlong C. Sensitivity to genotoxic effects of bleomycin in humans: possible relationship to environmental carcinogenesis. *Int J Cancer* 1989;43:403-9.
17. Spitz MR, Fueger JJ, Beddingfield NA, Annegers JF, Hsu TC, Newell GR, Schantz SP. Chromosome sensitivity to bleomycin-induced mutagenesis, an independent risk factor for upper aerodigestive tract cancers. *Cancer Res* 1989;49:4626-8.
18. Spitz MR, Fueger JJ, Halabi S, Schantz SP, Sample D, Hsu TC. Mutagen sensitivity in upper aerodigestive tract cancer: a case-control analysis. *Cancer Epidemiol Biomarkers Prev* 1993;2:329-33.
19. Schantz SP, Hsu TC, Ainslie N, Moser RP. Young adults with head and neck cancer express increased susceptibility to mutagen-induced chromosome damage. *JAMA* 1989;262:3313-15.
20. Cloos J, Braakhuis BJ, Steen I, Copper MP, de Vries N, Nauta JJ, Snow GB. Increased mutagen sensitivity in head-and-neck squamous-cell carcinoma patients, particularly those with multiple primary tumors. *Int J Cancer* 1994;56:816-19.
21. Cloos J, Spitz MR, Schantz SP, Hsu TC, Zhang ZF, Tobi H, Braakhuis BJ, Snow GB. Genetic susceptibility to head and neck squamous cell carcinoma. *J Natl Cancer Inst* 1996;88:530-5.
22. Wu X, Lippman SM, Lee JJ, Zhu Y, Wei QY, Thomas M, Hong WK, Spitz MR. Chromosome instability in lymphocytes: a potential indicator of predisposition to oral premalignant lesions. *Cancer Res* 2002;62:2813-18.
23. Strom SS, Wu X, Sigurdson AJ, Hsu TC, Fueger JJ, Lopez J, Tee PG, Spitz MR. Lung cancer, smoking patterns, and mutagen sensitivity in Mexican-Americans. *J Natl Cancer Inst (Monograph)* 1995;18:29-33.
24. Wu X, Gu J, Amos CI, Jiang H, Hong WK, Spitz MR. A parallel study of in vitro sensitivity to benzo[a]pyrene diol epoxide and bleomycin in lung cancer cases and controls. *Cancer* 1998;83:1118-27.
25. Hudmon KS, Honn SE, Jiang H, Chamberlain RM, Xiang W, Ferry G, Gosbee W, Hong WK, Spitz MR. Identifying and recruiting healthy control subjects from a managed care organization: a methodology for molecular epidemiological case-control studies of cancer. *Cancer Epidemiol Biomarkers Prev* 1997;6:565-71.
26. Schabath MB, Spitz MR, Delclos GL, Gunn GB, Whitehead LW, Wu X. Association between asbestos exposure, cigarette smoking, myeloperoxidase (MPO) genotypes, and lung cancer risk. *Am J Ind Med* 2002;42:29-37.
27. Lee JJ, Trizna Z, Hsu TC, Spitz MR, Hong WK. A statistical analysis of the reliability and classification error in application of the mutagen sensitivity assay. *Cancer Epidemiol Biomarkers Prev* 1996;5:191-7.
28. Zhang H, Singer B. Recursive partitioning in the health sciences. New York: Springer, 1999.
29. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. California: Wadsworth, 1984.
30. Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, Moore JH. Multifactor dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet* 2001;69:138-47.
31. Hahn LW, Ritchie MD, Moore JH. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. *Bioinformatics* 2003;19:376-82.
32. Moore JH. Computational analysis of gene-gene interactions using multifactor dimensionality reduction. *Expert Rev Mol Diagn* 2004;4:795-803.
33. Spitz MR, Hoque A, Trizna Z, Schantz SP, Amos CI, King TM, Bondy ML, Hong WK, Hsu TC. Mutagen sensitivity as a risk factor for second malignant tumors following malignancies of the upper aerodigestive tract. *J Natl Cancer Inst* 1994;86:1681-4.
34. Schantz SP, Spitz MR, Hsu TC. Mutagen sensitivity in patients with head and neck cancers: a biologic marker for risk of multiple primary malignancies. *J Natl Cancer Inst* 1990;82:1773-5.
35. Bonassi S, Hagmar L, Stromberg U, Montagud AH, Tinnerberg H, Forni A, Heikkila P, Wanders S, Wilhardt P, Hansteen IL, Knudsen LE, Norppa H. Chromosomal aberrations in lymphocytes predict human cancer independently of exposure to carcinogens. *Cancer Res* 2000;60:1619-25.
36. Hagmar L, Brogger A, Hansteen IL, Heim S, Hogstedt B, Knudsen L, Lambert B, Linnainmaa K, Mitelman F, Nordenson I. Cancer risk in humans predicted by increased levels of chromosome aberrations in lymphocytes: Nordic study group on the health risk of chromosome damage. *Cancer Res* 1994;54:2919-22.
37. Hagmar L, Bonassi S, Stromberg U, Brogger A, Knudsen LE, Norppa H, Reuterwall C. Chromosomal aberrations in lymphocytes predict human cancer: a report from European study group on cytogenetic biomarkers and health (ESCH). *Cancer Res* 1998;58:4117-21.
38. Wu X, Spitz MR, Amos CI, Lin J, Shao L, Gu J, de Andrade M, Benowitz NL, Shields PG, Swan GE. Mutagen sensitivity has high heritability: evidence from a twin study. *Cancer Res* 2006;66:5993-6.
39. Cloos J, Steen I, Joenje H, Ko JY, De Vries N, Vander Sterre MLY, Nauta JJP, Snow GB, Braakhuis BJM. Association between bleomycin genotoxicity and nonconstitutional risk factors for head and neck cancer. *Cancer Lett* 1993;74:161-5.
40. Wang Y, Spitz MR, Zhu Y, Dong Q, Shete S, Wu X. From genotype to phenotype: correlating XRCC1 polymorphisms with mutagen sensitivity. *DNA Repair* 2003;2:901-8.
41. Barnes PJ, Shapiro SD, Pauwels RA. Chronic obstructive pulmonary disease: molecular and cellular mechanisms. *Eur Respir J* 2003;22:672-88.
42. Moss SF, Blaser MJ. Mechanisms of disease: inflammation and the origin of cancer. *Nat Clin Pract Oncol* 2005;2:90-7.