

# The NAD(P)H:quinone Oxidoreductase 1 Gene Polymorphism and Lung Cancer: Differential Susceptibility Based on Smoking Behavior<sup>1</sup>

Li Lian Xu, John C. Wain, David P. Miller,  
Sally W. Thurston, Li Su, Thomas J. Lynch, and  
David C. Christiani<sup>2</sup>

Departments of Environmental Health, Occupational Health Program [L. L. X., D. P. M., L. S., D. C. C.] and Biostatistics [S. W. T.], Harvard School of Public Health, Boston, Massachusetts 02115; and Thoracic Surgery Unit, Departments of Surgery [J. C. W.] and Hematology-Oncology [T. J. L.], and Pulmonary and Critical Care Unit, Department of Medicine [D. C. C.], Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114

## Abstract

**We conducted a hospital-based case-control study of 814 lung cancer patients and 1123 controls to examine the association of the NAD(P)H:quinone oxidoreductase 1 (*NQO1*) gene polymorphism with lung cancer susceptibility. Using PCR-RFLP genotyping assay techniques, we analyzed DNA samples to detect the variant forms of the *NQO1* gene in exon 6 on chromosome 16q. We examined the relationship between lung cancer odds and *NQO1* genotypes after adjusting for age, gender, and smoking behavior using generalized additive modeling. We found no overall association between *NQO1* genotypes and lung cancer susceptibility, regardless of age, gender, family history of cancer, or histological cell type. However, our data demonstrated that in both former and current smokers, there was an association between *NQO1* genotypes and lung cancer susceptibility that was dependent upon cigarette smoking duration and smoking intensity. For both current and former smokers, smoking intensity was more important in predicting cancer risk than smoking duration for all of the genotypes. Among former smokers, individuals with the *T/T* genotype were predicted to have a greater cancer risk than those with the *C/C* genotype for smoking durations up to 37 years. The predicted cancer risk for former smokers with the *C/T* versus *T/T* genotype depended on both smoking intensity and smoking duration. Our results support the concept that differential susceptibility to lung cancer is a function of both an inheritable trait in *NQO1* metabolism and individual smoking characteristics.**

Received 9/22/00; revised 12/8/00; accepted 1/19/01.

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.

<sup>1</sup> Supported by NIH Grants CA74386, ES/CA06409, and ES00002. Dr. Xu was supported by training Grant ES T3207069.

<sup>2</sup> To whom requests for reprints should be addressed, at Occupational Health Program, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115. Phone: (617) 432-3323; Fax: (617) 432-3441; E-mail: dchris@hohp.harvard.edu.

## Introduction

Extensive epidemiological data clearly establish cigarette smoking as the major cause of lung cancer (1). Although up to 90% of lung cancer in the United States is attributed to cigarette smoking, only 10% of smokers develop a bronchogenic carcinoma (2). This observation implies that host factors may influence individual susceptibility to tobacco smoke. The pleiotropic genes that control oxidative metabolism (phase I) and conjugation of reactive intermediates (phase II) are examples of such host factors. Procarcinogens in tobacco smoke, such as the polycyclic aromatic hydrocarbons and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, require metabolic activation to exert their carcinogenic effects. These metabolic processes are regulated by pleiotropic genes via complex enzymatic mechanisms. The variant alleles that encode the same type of proteins with different activities or expression patterns will affect individual susceptibility to cancer.

*NQO1*<sup>3</sup> is an important flavoprotein that catalyzes the two-electron reduction of carcinogenic quinoid compounds into their reduced form, such as hydroquinones (3). The *NQO1* gene is located on chromosome 16q22. The polymorphic variant is a C to T point mutation at position 609 of exon 6 of the *NQO1* cDNA that encodes for a proline to serine substitution at position 187 in the amino acid sequence of the protein. These three genotypes of *NQO1* are known to be associated with different enzyme activities: *C/C* is the homozygous wild type with normal activity; *C/T* is heterozygous with mild activity; and *T/T* is the homozygous variant with only 2–4% of the enzyme activity as the wild type (4–8).

There have been several studies examining the relationship between the *NQO1* genetic polymorphism and lung cancer risk, but the conclusions have been contradictory. Rosvold *et al.* (9) reported in 1995 that the rare, null activity allele was observed to be approximately twice as common in Caucasian cases than in controls. More recent studies (10–12) have shown that this type of variant is inversely associated with lung cancer risk. The specific relationship between the *NQO1* genotype and lung cancer susceptibility remains unclear.

In the present study, we conducted a large scale (1937 subjects) hospital-based case-control study to evaluate the relevance of the three *NQO1* polymorphisms (*C/C*, *C/T*, and *T/T*) with respect to lung cancer susceptibility. Generalized additive modeling (13) was used to adjust for potential important confounding variables and to examine the relationship between the odds of lung cancer and each covariate.

<sup>3</sup> The abbreviations used are: *NQO1*, NAD(P)H:quinone oxidoreductase 1; MGH, Massachusetts General Hospital; GAM, generalized additive models; OR, odds ratio; CI, confidence interval.

## Materials and Methods

**Study Population.** This study was approved by the Human Subjects Committees of both MGH and the Harvard School of Public Health. There were 814 primary lung cancer cases and 1123 controls recruited for the study. Lung cancer cases included newly diagnosed lung cancer patients that were admitted to the Massachusetts General Hospital Thoracic Surgery, Oncology and Pulmonary Services for surgery, chemotherapy, or radiation therapy between December 1992 and December 1999. Diagnosis of primary lung cancer was confirmed through a review of each patient's pathology report by an MGH lung pathologist.

Controls were recruited first among the friends and non-blood related family members of the cases with no specific matching characteristics. If related family members were not available, controls were recruited from friends and family members of patients either receiving thoracic surgery for conditions other than lung cancer or receiving chemotherapy or radiation treatment for a condition other than lung cancer.

A detailed interviewer-administered questionnaire was completed for each case and control by a trained interviewer. The standardized American Thoracic Society respiratory questionnaire was modified to include detailed occupational and environmental exposure history (14, 15). The questionnaire included information on the average number of cigarettes smoked daily and the number of years the subjects had been smoking. For ex-smokers, the time elapsed since quitting was recorded. If the participant was not able to fill out the questionnaire at the time of recruitment, a prestamped envelope was provided to allow the questionnaires to be filled out at home and returned by mail. If there was missing data, participants were contacted by phone to obtain a complete set of information. In addition, follow-up information was obtained from patient abstracts at the MGH cancer registry.

This selection process provided both the control group and cases with natural balance, with no overmatching for potential confounding factors such as smoking status, ethnicity, race, socioeconomic status, and age (15).

***NQO1* *HinfI* Polymorphism.** Blood samples were obtained from each of the cases and controls via venipuncture and shipped to the Molecular Epidemiology Laboratory of the Harvard University Occupational Health Program. Within 24 h, DNA was extracted from the sample using the Puregene DNA isolation kit (Gentra Systems, Minneapolis, MN). Genotyping was performed by investigators who were blinded to the subjects' case or control status. This procedure was modified slightly from that of Traver *et al.* (16). A PCR-RFLP analysis was used to characterize the wild type and variant of the codon 187 of the *NQO1* alleles. Using a Perkin-Elmer Corp. 9600 thermocycler, PCR products were generated using 100–300 ng of genomic DNA as a template. The sense primer (5'TCCT-CAGAGTGGCATTCTGC-3') and antisense primer (5'TCTC-CTCATCCTGTACCTCT-3') amplified a 230-bp oligonucleotide including the *NQO1* 609 C to T polymorphism (the last seven bases of exon 5 and the first 204 bases of exon 6). This C→T transition creates a new *HinfI* restriction site to distinguish the variant from the wild-type allele. The PCR reaction mixture contained a final concentration of 50 μM of each deoxynucleotide triphosphate, 3.5 mM MgCl<sub>2</sub>, 0.4 μM of each primer, and 1.5 units of Taq polymerase (Perkin-Elmer Corp). After an initial DNA denaturation at 94°C for 4 min, the cycling patterns used included: two cycles at 94°C for 15 s, 69°C for 15 s, and 72°C for 30 s, followed by two cycles at 94°C for 15 s, 67°C for 15 s, and 72°C for 30 s. Final amplification included

31 cycles at 94°C for 30 s, 62°C for 30 s, and 72°C for 60 s. The final extension was 72°C for 5 min. The PCR products were then digested with 5 units of restriction enzyme *HinfI* (New England Biolabs, Beverly, MA) at 37°C for 3 h. The genotype pattern for each PCR product was detected by electrophoresis on a 2% agarose gel (Sigma Chemical Co., St. Louis, MO) containing 0.5 μg/ml of ethidium bromide. Complete digestion of the 230-bp PCR product produced a cut fragment of 195 bp for the wild-type allele and 151 bp for the variant allele.

**Statistical Analysis.** Contingency tables were created to compare cases and controls for demographic variables and genotype prevalence using  $\chi^2$  and *t*-tests. GAMs were then used to examine the relationship between the logarithm of the odds of lung cancer and each continuous covariate (13). A GAM extends the generalized linear models framework, such as logistic regression, by allowing the relationship between the outcome and each covariate to be an unspecified but smooth function. GAM plots of the log odds of lung cancer *versus* each covariate, in a model that adjusts for the other covariate, were created in S-plus (17). The plots were examined for departures from linearity. If such departures were found, the plots were further examined to see if the shape of the relationship suggested that a parametric transformation of the covariate would be linearly related to the odds of lung cancer. A second GAM was then fit that used the transformed covariate. The plot created from the second GAM was then examined to see if the transformed variable was linearly related to the log odds of cancer. Logistic regression, using transformed covariates as needed, was used to model the relationship between the log odds of the lung cancer and *NQO1* genotypes, controlling for age, sex, and smoking behavior.

Pack-years is a composite quantitative measure of smoking behavior. In our analysis, we elected to decompose pack-years into their individual components of cigarette smoking duration (measured in years) and the average number of cigarettes smoked/day (smoking intensity) to better delineate any differential effects of these variables. For former smokers, we collected data on the number of years since quitting smoking. Because the number of years since quitting smoking is undefined for lifelong nonsmokers, our GAM and logistic regression analyses excluded those subjects. A lack of fit test, as described in Hosmer and Lemeshow (18) was performed to summarize the goodness-of-fit for each logistic regression model.

## Results

A total of 814 lung cancer cases and 1123 controls were used in this study, of which 97% were Caucasians. Table 1 summarizes the distribution of demographic variables for both cases and controls. Compared with controls, cases were somewhat younger and had a higher proportion of males. A greater proportion of cases had a heavy cigarette smoking history (defined as more than 30 years) when compared with controls. There is a well-documented strong ethnic variation in the prevalence of the *NQO1* polymorphism; the frequency of the *NQO1* *T/T* genotype ranges from approximately 3% in Caucasians to 22% in Chinese populations (10, 11, 19). Because Caucasians comprised more than 97% of our data, the analyses were restricted to Caucasians.

The frequency distribution of the *NQO1* genotypes of *T/T*, *C/T*, and *C/C* were 2.8%, 32%, and 65.2% in cases and 3.7%, 31%, and 65.3% in controls, respectively. The control genotype distribution was consistent with Hardy-Weinberg equilibrium ( $P > 0.05$ ;  $\chi^2$  goodness of fit). These frequencies were similar to those of other published studies (9, 11). We did not find any

Table 1 Frequency (%) of the *NQO1* gene homozygous variant among lung cancer patients and controls defined by demographic characteristics

	Case (%)				Control (%)			
	<i>n</i>	TT <sup>a</sup>	CT <sup>b</sup>	CC <sup>c</sup>	<i>n</i>	TT <sup>a</sup>	CT <sup>b</sup>	CC <sup>c</sup>
Total	814	2.8	31.9	65.2	1123	3.7	31.0	65.3
Gender								
Male	445	2.9	30.1	67.0	506	4.4	27.5	68.2
Female	369	2.7	34.2	63.1	617	3.2	33.9	62.9
Race								
Caucasian	780	2.7	31.5	65.8	1096	3.7	31.1	65.2
Other	34	5.9	41.2	52.9	27	3.7	22.2	74.1
Age (yr)								
<55	139	2.2	32.4	65.5	423	4.0	28.1	67.9
55–64	191	3.1	34.0	62.8	256	1.6	32.0	66.4
>64	484	2.9	31.0	66.1	444	4.7	33.1	62.2
Smoking Status								
Never	50	0.0	36.0	64.0	396	4.6	33.6	61.9
Former	432	3.0	31.0	66.0	515	2.7	28.7	68.5
Current	332	3.0	32.2	64.8	211	4.7	31.8	63.5
Cigarettes/day								
1–5	25	4.0	4.0	92.0	79	1.3	25.3	73.4
>5–20	367	3.0	31.1	65.9	395	3.5	32.2	64.3
>20–40	381	3.4	36.5	60.1	200	3.0	28.0	69.0
>40	119	2.5	28.6	68.9	47	4.3	21.3	74.4
Years since cessation (ex-smokers)								
1–4	102	2.9	31.4	65.7	55	1.8	32.7	65.5
>4–14	158	2.5	29.8	67.7	153	2.0	24.2	73.8
>14–29	128	2.3	31.3	66.4	195	4.1	30.8	65.1
>29	44	6.8	34.1	59.1	112	1.8	29.5	68.7
Yr smoked								
≤10	16	12.5	37.5	50.0	104	1.9	24.0	74.1
>10–20	42	2.4	31.0	66.6	139	2.9	25.2	71.9
>20–30	115	1.7	29.6	68.7	192	2.1	29.2	68.8
>30–40	228	1.8	31.6	66.6	158	3.8	33.5	62.7
>40–50	204	3.4	28.4	68.1	94	5.3	34.0	60.7
>50–60	126	3.2	36.5	60.3	30	6.7	30.0	63.3
>60	19	5.3	42.1	52.6	4	0.0	75.0	25.0
Family history of cancer								
No	251	2.8	33.9	63.3	436	4.4	30.5	65.1
Yes	508	2.6	31.1	66.3	631	3.5	31.2	65.3

<sup>a</sup> Homozygous variant.<sup>b</sup> Heterozygous.<sup>c</sup> Homozygous wild-type.

statistically significant differences for the genotype distribution between cases and controls as a whole or when analyzed by subgroups (gender, age, and family history of cancer). When lung cancer patients were stratified into different histological cell types, we found a slightly increased frequency of the *TT* genotype in the squamous cell lung cancer (4.2%) compared with adenocarcinoma (2.6%) and the other subtypes (1.9%), such as mixed cell, large cell, and small cell lung cancer, although this difference was not statistically significant ( $P > 0.05$ ; Table 2).

The GAM plots showed a linear association between lung cancer risk and cigarettes/day when this variable was square-root transformed, and they showed linear associations between lung cancer risk and each of the other smoking variables. Therefore, the logistic regression analysis was performed using the square-root transformed cigarettes/day.

We examined logistic regression models containing gene-environment interactions between *NQO1* genotypes and different smoking variables (to allow for the possibility that the effect of smoking behavior on the risk of lung cancer differed among the different genotypes) and interactions between smoking variables and smoking status (former or current). The relationship between the risk of lung cancer

Table 2 Distribution of the *NQO1* polymorphism among lung cancer patients and controls by histological cell types

Histological cell type	<i>n</i>	<i>NQO1</i> polymorphism (%)		
		TT <sup>a</sup>	CT <sup>b</sup>	CC <sup>c</sup>
Controls	1123	3.7	31	65
Cancer patients <sup>d</sup>	814	2.9	32	65
Adenocarcinoma	390	2.6	30	68
Squamous carcinoma	217	4.2	35	61
Other <sup>e</sup>	207	1.9	33	65

<sup>a</sup> Homozygous variant<sup>b</sup> Heterozygous.<sup>c</sup> Homozygous wild type.<sup>d</sup> 154 individuals had missing histological cell type data.<sup>e</sup> Other refers to large cell, small cell, and mixed lung cancer histological subtypes.

and cigarettes/day was statistically significantly different in current *versus* former smokers ( $P < 0.005$ ). We then decided to analyze former and current smokers in separate models to avoid multiple two-way interactions and three-way interactions between smoking status, smoking variables, and *NQO1* genotypes. All of the ORs were adjusted for age, gender,

Table 3 OR of lung cancer risk for the *NQO1* *T/T* genotype compared with *C/C* using logistic regression models stratified by smoking duration in former smokers<sup>a,b</sup>

Yr of smoking	5 cigarettes/day		20 cigarettes/day		40 cigarettes/day	
	OR	95% CI	OR	95% CI	OR	95% CI
1	4.51	0.60–33.8	4.38	0.68–28.0	4.27	0.48–38.0
10	3.11	0.57–16.9	3.02	0.74–12.2	2.94	0.51–16.8
20	2.05	0.46–9.12	1.99	0.74–5.38	1.94	0.51–7.41
30	1.36	0.30–6.17	1.32	0.56–3.09	1.28	0.42–3.94
40	0.90	0.15–5.19	0.87	0.29–2.58	0.85	0.26–2.81
50	0.59	0.07–5.04	0.57	0.12–2.66	0.56	0.12–2.58
60	0.39	0.03–5.31	0.38	0.05–2.97	0.37	0.05–2.69
70	0.26	0.01–5.85	0.25	0.02–3.43	0.24	0.02–2.99

<sup>a</sup> *T/T* is homozygous variant; *C/C* is homozygous wild-type.

<sup>b</sup> Adjusted for age, sex, years smoked, the average number of cigarettes smoked/day, years since smoking cessation, the *NQO1* genotype (with three alleles: *C/C*, *C/T*, and *T/T*), and the interactions between *NQO1* genotype and both the number of cigarettes smoked/day and years smoked. Note that this model considers the square-root of the number of cigarettes smoked/day, and years of smoking as continuous variables.

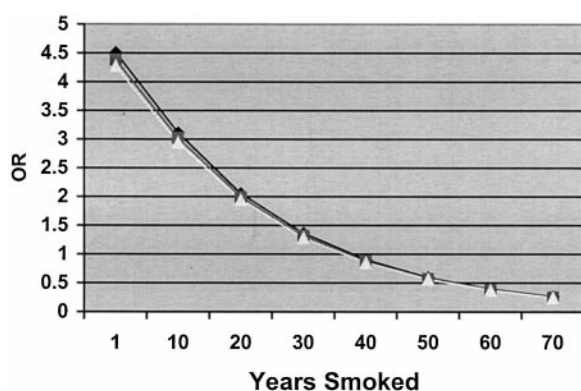


Fig. 1. OR of lung cancer risk for *NQO1* *T/T* versus *C/C* using logistic regression models stratified by smoking duration in former smokers. Model incorporates age, gender, years smoked, average number of cigarettes smoked/day, years since smoking cessation, the *NQO1* genotype (with three alleles: *C/C*, *C/T*, and *T/T*), and the interactions between *NQO1* genotype and both the number of cigarettes smoked/day and years smoked. Note that this model considers the square root of the number of cigarettes smoked/day and years of smoking as continuous variables.  $\blacklozenge$ , light smoker, five cigarettes/day;  $\blacksquare$ , mild smoker, 20 cigarettes/day;  $\blacktriangle$ , heavy smoker, 40 cigarettes/day.

smoking duration, square-root transformed cigarettes smoked/day, and years since quitting smoking.

For former smokers, the interactions between genotype and cigarettes smoked/day and between genotype and smoking duration were both important ( $P = 0.006$  for both pairs of interactions together). After adjusting for age, gender, and years since quitting smoking, the predicted risk of cancer increased with smoking intensity for all of the three genotypes. Smoking duration was not an important predictor of cancer risk, except for the *C/C* genotype for which cancer risk was higher with longer duration. Except for long-term smokers of greater than 37 years, the overall risk of cancer was greater for the *T/T* genotype than the *C/C* genotype. The effect of smoking duration on cancer risk was slightly smaller for the *T/T* genotype compared with the *C/C* genotype, whereas the effect of number of cigarettes smoked/day was about the same for these two genotypes (Table 3; Fig. 1). Comparing *T/T* with *C/C*, for an average of 20 cigarettes smoked/day, the OR for 10 years of smoking was 3.02 (95% CI, 0.74–12.23), whereas for 50 years of smoking the OR was 0.57 (95% CI, 0.12–2.66). In the same model, when the *C/T* genotype was compared with the *C/C* genotype, the increased risk of cancer from a longer smoking

duration was significantly less for the *C/T* genotype ( $P = 0.03$ ), whereas the increased cancer risk from a greater daily smoking intensity was significantly larger for the *C/T* genotype ( $P = 0.002$ ), as compared with the *C/C* genotype (Table 4; Fig. 2). For five cigarettes smoked/day, the OR for 10 years of smoking was 0.65 (95% CI, 0.28–1.50), and for 50 years of smoking the OR was 0.20 (95% CI, 0.08–0.50). In contrast, for 40 cigarettes smoked/day, the OR for 10 years of smoking was 2.99 (95% CI, 1.42–6.28), and the OR for 50 years of smoking was 0.91 (95% CI, 0.48–1.71).

For current smokers, the risk of cancer for all of the genotypes increased with smoking intensity ( $P < 0.0001$ ) but not with smoking duration. There was a slight indication that current smokers with the *T/T* genotype had a smaller cancer risk than current smokers with *C/T* or *C/C* genotypes; the OR for *T/T* versus *C/C* was 0.38 (95% CI, 0.19–1.00).

Although there were no statistically significant interactions between genotype and smoking duration or intensity, the direction of these associations was similar to ex-smokers. In models that incorporated the interaction terms comparing *T/T* with *C/C* for an average of 20 cigarettes/day, the OR for 10 years of smoking was 4.09 (95% CI, 0.11–151), and for 50 years of smoking the OR was 0.21 (95% CI, 0.05–0.84). In the same model, when the *C/T* genotype was compared with the *C/C* genotype, for five cigarettes smoked/day, the OR for 10 years of smoking was 0.60 (95% CI, 0.12–3.02), and for 50 years of smoking, the OR was 0.45 (95% CI, 0.15–1.35). In contrast, for 40 cigarettes smoked/day, the OR for 10 years of smoking was 2.26 (95% CI, 0.61–8.35), and the OR for 50 years of smoking was 1.70 (95% CI, 0.68–4.25).

All of the models reported in the study showed no lack of fit using the Hosmer and Lemeshow test (18).

## Discussion

Metabolism is an essential part of the chemical carcinogenic process. In humans, a significant proportion of xenobiotic metabolizing enzymes are polymorphic. Different phenotypes may arise from the existence of multiple alleles at important loci within genes encoding chemical metabolizing enzymes; this may explain the varying susceptibility of individuals to the mutagenic and carcinogenic effects of environmental chemicals.

*NQO1* has attracted considerable attention because of its ability to detoxify a number of natural and synthetic compounds (such as quinones) and, conversely, to activate certain anticancer agents (20, 21). Quinones (e.g., benzo( $\alpha$ )-pyrene-3,6-quinone) are found in all burnt organic material, including



Table 4 OR of lung cancer risk for the *NQO1* *C/T* genotype compared with *C/C* using logistic regression models stratified by smoking duration in former smokers<sup>a,b</sup>

Yr of smoking	5 cigarettes/day		20 cigarettes/day		40 cigarettes/day	
	OR	95% CI	OR	95% CI	OR	95% CI
1	0.85	0.32–2.27	1.95	0.85–4.052	3.91	1.52–10.0
10	0.65	0.28–1.50	1.50	0.80–2.80	2.99	1.42–6.28
20	0.48	0.23–1.00	1.11	0.72–1.71	2.22	1.28–3.86
30	0.36	0.17–0.73	0.82	0.59–1.16	1.65	1.05–2.58
40	0.26	0.12–0.58	0.61	0.40–0.94	1.22	0.75–1.98
50	0.20	0.08–0.50	0.45	0.24–0.85	0.91	0.48–1.71
60	0.15	0.05–0.45	0.34	0.14–0.79	0.67	0.29–1.57
70	0.11	0.03–0.41	0.25	0.08–0.75	0.50	0.17–1.47

<sup>a</sup> *C/T* is heterozygous variant, *C/C* is homozygous wild-type.

<sup>b</sup> Adjusted for age, sex, years smoked, the average number of cigarettes smoked/day, years since smoking cessation, the *NQO1* genotype (with three alleles: *C/C*, *C/T*, and *T/T*), and the interactions between *NQO1* genotype and both the number of cigarettes smoked/day and years smoked. Note that this model considers the square root of the number of cigarettes smoked/day, and years of smoking as continuous variables.

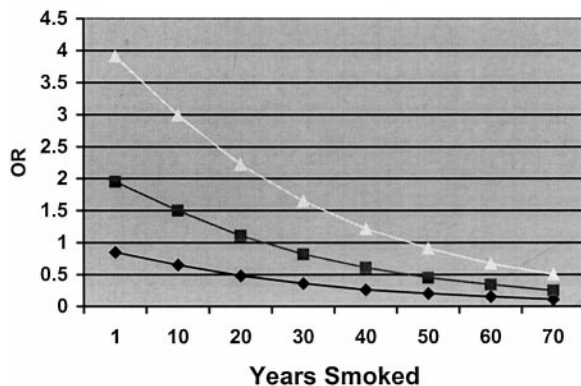


Fig. 2. OR of lung cancer risk for *NQO1* *C/T* versus *C/C* using logistic regression models stratified by smoking duration in former smokers. Model incorporates age, gender, years smoked, average number of cigarettes smoked/day, years since smoking cessation, the *NQO1* genotype (with three alleles: *C/C*, *C/T*, and *T/T*), and the interactions between *NQO1* genotype and both the number of cigarettes smoked/day and years smoked. Note that this model considers the square root of the number of cigarettes smoked/day and years of smoking as continuous variables. ◆, light smoker, five cigarettes/day; ■, mild smoker, 20 cigarettes/day; △, heavy smoker, 40 cigarettes/day.

automobile exhaust, cigarette smoke, and urban air particulates (22). The pathways of xenobiotic metabolism have been classified as either phase I (referred to as activation) or phase II (referred to as detoxification). One-electron reduction of quinones by NAD(P)H-cytochrome p450 reductase (phase I) results in the formation of mutagenic semiquinones. Semiquinones enter the redox cycling pathway in the presence of molecular oxygen and generate reactive oxygen species that stimulate oxidative stress and DNA/membrane damage. Cytotoxicity, mutagenicity, and possibly carcinogenicity may ensue. In contrast, *NQO1* (phase II) competes with the one-electron-reducing enzymes and catalyzes a two-electron reductive metabolic conversion of the quinones to hydroquinones, leading to their detoxification (23). *NQO1* activity is also known to catalyze the activation of some procarcinogens, such as nitrosamines and certain anticancer agents such as mitomycin C and 2,5-diaziridinyl-3-(hydroxymethyl)-6-methyl-1,4-benzoquinone (24).

*NQO1* is a highly inducible enzyme. The procarcinogens in tobacco smoke, such as polycyclic aromatic hydrocarbons, are examples of such inducers (21). High levels of *NQO1* gene expression have been observed in liver, lung, colon, and breast

tumors when compared with normal tissues of the same origin. In addition to established tumors, *NQO1* gene expression is also increased in developing tumors, thereby indicating a role in cellular defense during tumorigenesis (25, 26).

Because there was a potential role for the *NQO1* enzyme in the metabolism of procarcinogens in tobacco smoke, we explored a possible association between the *NQO1* gene <sup>609</sup>C → T mutation, which results in decreased *NQO1* enzyme activity, and lung cancer susceptibility. We used a generalized additive modeling technique, which allows the data to inform us about the nature of the relationship between the outcome and each continuous covariate, without making an assumption of linearity.

This is the largest study of *NQO1* genotype and lung cancer susceptibility to date. Our data did not demonstrate a significant overall association between these two primary factors, regardless of histological type, age, or gender. We did find a gene-environment interaction between *NQO1* genotype and smoking; the association between *NQO1* genotype and lung cancer risk differed depending on cigarette smoking status (current versus former smoker), smoking duration, and smoking intensity (as defined by the number of cigarettes smoked daily). Among former smokers who smoked 20 cigarettes/day, the *T/T* genotype had a greater predicted risk of cancer than the *C/C* genotype for smoking durations up to 37 years. The point of equal predicted risk for the *T/T* as compared with the *C/C* genotype was similar for other smoking intensities. The cancer risk for the *C/T* as compared with the *C/C* genotype depended on smoking time and intensity. Specifically, among those who smoked 20 cigarettes/day, the predicted cancer risk was greater for the *C/T* genotype for individuals who smoked 23.5 years or less, whereas among those who smoked 40 cigarettes/day, the predicted cancer risk was greater for the *C/T* genotype for individuals who smoked for 47 years or less. Among those who only smoked five cigarettes/day, those with the *C/T* genotype had a lower predicted cancer risk than those with the *C/C* genotype for all years of smoking. These differential interactions between genotypes and smoking patterns were not examined in previous reports (9–12) and may explain the heterogeneity of published results.

Tables 3 and 4 further demonstrate the predominant effects of gene versus environment at different levels of smoking. At very low levels of smoking (five cigarettes/day and for less than 10 years of smoking) and for heavy smokers (40 cigarettes/day and for at least 40 years of smoking), genotype does not appear to play a major role in cancer risk modification (ORs at these levels cross unity in both Tables in these smoking ranges).

In separate secondary analyses (data not shown), light former smokers (defined as individuals who smoked 35 pack-years or less) and individuals who had quit smoking for at least 10 years were primarily responsible for the gene-smoking interactions; *e.g.*, in the comparison of *T/T* with *C/C* for the years smoked and cigarettes smoked/day combinations given in Table 3, the OR ranged from 0.95 and 1.21 for heavy former smokers and from 0.03 and 7.42 for light former smokers. All of these results conform to the current theories that genotype susceptibility is most important in those individuals whose exposure to smoking is moderate or remote, where the overwhelming factor of current or heavy carcinogen exposure plays less of a role in carcinogenesis.

In former smokers with a lower *NQO1* enzyme activity, a longer smoking duration appears to decrease the risk of lung cancer compared with individuals with normal *NQO1* enzyme activity (Figs. 1 and 2). This was because a longer smoking duration had little impact on the *C/T* and *T/T* genotypes in former smokers but increased cancer risk in individuals with the *C/C* genotype. This lack of effect of smoking duration by the *C/T* and *T/T* genotypes may be explained by the bioactivation functions of the *NQO1* enzyme. *NQO1* appears to have phase I in addition to phase II function, depending on the substrate; *e.g.*, 2,5-diaziridinyl-3-(hydroxymethyl)-6-methyl-1,4-benzoquinone and 2,5-diaziridinyl-3,6-dimethyl-1,4-benzoquinone are compounds that are bioactivated by *NQO1*. These bioactivated forms selectively attack cells that contain elevated *NQO1* enzyme activity (23). Because the carcinogens contained in cigarette smoke may contain similar benzoquinone compounds, subjects with lower *NQO1* activity (*T/T* and *C/T* genotype) may be protected in the long term against the harmful effects of these compounds when compared with patients with normal *NQO1* activity (*C/C* genotype).

In current smokers, the effect of smoking intensity on lung cancer risk was so strong ( $P < 0.001$ ) that any modest differential effect of smoking duration on lung cancer risk by the variant genotypes may be masked. The use of spouse and friend controls may have further reduced the main effect of smoking duration on lung cancer risk but would not change the direction of the association nor significantly alter the interaction of genotype and smoking duration, unless there was an association between the smoking variables (such as duration) and *NQO1* genotype. There is no biological basis to suspect this. Although we did not see statistically significant interactions among smoking duration, smoking intensity, and genotype in modeling lung cancer risk in current smokers, the overall direction of risk in current smokers was similar to ex-smokers. These differences between current smokers and ex-smokers may reflect the predominance of continued enzyme induction of *NQO1* in the presence of tobacco carcinogen activation by *NQO1* in the *C/C* genotype among current smokers.

Because our data set found that both the *C/T* and *T/T* genotype produced a higher risk of lung cancer compared with the wild-type genotype in those who smoked more intensely over a shorter period of time in former smokers (Table 4), we hypothesize that the phase II activity of *NQO1* is most important in the early years after an individual starts to smoke, whereas the phase I effects predominate in longer term smokers. Previous studies have invoked both phase I (10–12) and phase II *NQO1* activity (9) to explain their results, but our study suggests that an individual's lung cancer susceptibility is dependent on variable smoking behaviors that lead to differences in the predominance of phase I or phase II activity.

Although this study underscores the complexity of *NQO1* enzymatic function, it raises a number of questions. How does

phase I and phase II *NQO1* activity interact to determine an overall risk susceptibility of lung cancer in a particular environmental setting for individuals? Does race play a role in determining whether phase I or phase II functions of *NQO1* dominate in any one individual? What is the relationship between *NQO1* and other phase I (*e.g.*, cytochrome P450) and phase II (*e.g.*, glutathione *S*-transferase) enzymes and with the other oxidoreductases such as NADPH:P-450 reductase and myeloperoxidase? *NQO2* is a polymorphic gene that encodes an enzyme with similar activity to *NQO1*. Is *NQO2* also important in determining lung cancer risk? These questions require further exploration.

In summary, our results of over 1900 individuals did not support an overall association between *NQO1* genotype and lung cancer risk. However, we found that different smoking behaviors greatly modified lung cancer susceptibility. From these gene-environment analyses, we theorize that an individual's overall susceptibility to lung cancer is a consequence of both an inheritable trait in *NQO1* metabolism and a behavioral trait in smoking characteristics. Smoking intensity was much more important in predicting cancer risk than smoking duration. For a given increase in smoking intensity, the predicted risk of cancer increases more rapidly for current smokers than for former smokers. Future studies using larger sample sizes should focus on these behavioral traits in relation to *NQO1* genotypes.

### Acknowledgments

We thank Linda Lineback and Barbara Bean for patient recruitment; Lucy-Ann Principe-Hasan and Nick Weidemann for data entry and management; D. K. Kim for technical assistance, and Stephanie Shih, Wei Zhou, and Geoffrey Liu for manuscript preparation.

### References

- Blot, W. J., and Fraumeni, J. F., Jr. Cancers of the lung and pleura. In: D. Schottenfeld and J. Fraumeni, Jr. (eds.), *Cancer Epidemiology and Prevention*, pp. 637–665. New York, NY: Oxford University Press, 1996.
- United States Department of Health and Human Services, The Health Benefits of Smoking Cessation, DHHS publication no. 90-8416, 1990.
- Lind, C., Cadenas, E., Hochstein, P., and Ernster, L. DT-diaphorase: purification, properties, and function. *Methods Enzymol.*, 186: 287–301, 1990.
- Jaiswal, A. K., McBride, O. W., Adesnik, M., and Nebert, D. W. Human dioxin-inducible cytosolic NAD(P)H:menadiol oxidoreductase: cDNA sequence and localization of gene to chromosome 16. *J. Biol. Chem.*, 263: 13572–13578, 1988.
- Jaiswal, A. K. Human NAD(P)H quinone oxidoreductase (*NQO1*) gene structure and induction by dioxin. *Biochemistry*, 30: 10647–10653, 1991.
- Traver, R. D., Horikosh, T., Danenberg, K. D., Stadlbauer, T. H., Danenberg, P. V., Ross, D., and Gibson, N. W. NAD(P)H:quinone oxidoreductase gene expression in human colon carcinoma cells: characterization of a mutation which modulates DT-diaphorase activity and mitomycin sensitivity. *Cancer Res.*, 52: 797–802, 1992.
- Ross, D., Traver, R. D., Siegel, D., Kuehl, B. L., Misra, V., and Rauth, A. M. A polymorphism in NAD(P)H:quinone oxidoreductase (*NQO1*): relationship of a homozygous mutation at position 609 of the *NQO1* cDNA to *NQO1* activity. *Br. J. Cancer*, 74: 995–996, 1996.
- Siegel, D., McGuinness, S. M., Winski, S. L., and Ross, D. Genotype-phenotype relationships in studies of a polymorphism in NAD(P)H:quinone oxidoreductase 1. *Pharmacogenetics*, 9: 113–121, 1999.
- Rosvold, E. A., McGlynn, K. A., Lustbader, E. D., and Buetow, K. H. Identification of an NAD(P)H:quinone oxidoreductase polymorphism and its association with lung cancer and smoking. *Pharmacogenetics*, 5: 199–206, 1995.
- Wiencke, J. K., Spitz, M. R., McMillan, A., and Kelsey, K. T. Lung cancer in Mexican-Americans and African-Americans is associated with the wild-type genotype of the NAD(P)H:quinone oxidoreductase polymorphism. *Cancer Epidemiol. Biomark. Prev.*, 6: 87–92, 1997.
- Chen, H., Lum, A., Seifried, A., Wilkens, L. R., and Le Marchand, L. Association of the NAD(P)H:quinone oxidoreductase 609 C → T polymorphism with a decreased lung cancer risk. *Cancer Res.*, 59: 3045–3048, 1999.

12. Lin, P., Wang, H. J., Lee, H., Lee, H. S., Wang, S. L., Hsueh, Y. M., Tsai, K. J., and Chen, C. Y. NAD(P)H: quinone oxidoreductase polymorphism and lung cancer in Taiwan. *J. Toxicol. Environ. Health*, *58*: 187–197, 1999.
13. Hastie, T. J., and Tibshirani, R. J. (eds.). *Generalized Additive Models*. New York: Chapman and Hall, Inc., 1990.
14. Ferris, B. G. Epidemiology standardized project. *Am. Rev. Respir. Dis.*, *118*: 1–120, 1978.
15. Xu, X., Kelsey, K. T., Wiencke, J. K., Wain, J. C., and Christiani, D. C. Cytochrome P450 CYP1A1 MspI polymorphism and lung cancer susceptibility. *Cancer Epidemiol. Biomark. Prev.*, *5*: 687–692, 1996.
16. Traver, R. D., Siegel, D., Beall, H. D., Phillips, R. M., Gibson, N. W., Franklin, W. A., and Ross, D. Characterization of a polymorphism in NAD(P)H: quinone oxidoreductase (DT-diaphorase). *Br. J. Cancer*, *75*: 69–75, 1997.
17. Becker, R. A., Chambers, J. M., and Wilks, A. R. (eds.). *The New S Language*. Pacific Grove, California: Wadsworth, 1988.
18. Hosmer, D. W., and Lemeshow, S. (eds.). *Applied Logistic Regression*, pp. 140–156. New York: John Wiley & Sons, Inc., 1989.
19. Kelsey, K. T., Ross, D., Traver, R. D., Christiani, D. C., Zuo, Z. F., Spitz, M. R., Wang, M., Xu, X., Lee, B. K., Schwartz, B. S., and Wiencke, J. K. Ethnic variation in the prevalence of a common NAD(P)H quinone oxidoreductase polymorphism and its implications for anticancer chemotherapy. *Br. J. Cancer*, *76*: 852–854, 1997.
20. Ernster, L. In: K. Yagi (ed.), *Pathophysiology of Lipid Peroxides and Related Free Radicals*, pp.149–168. Basel: S. Karger AG, 1998.
21. Ross, D. Quinone reductases. In: F. P. Guengerich (ed.), *Comprehensive Toxicology*, Vol. 3, pp. 179–197. New York: Pergamon, 1997.
22. Monks, T. J., Hanzlik, R. P., Cohen, G. M., Ross, D., and Graham, D. G. Quinone chemistry and toxicity. *Toxicol. Appl. Pharmacol.*, *112*: 2–16, 1992.
23. Venugopal, R. Gene expression of D-T diaphorase in cancer cells. In: H. J. Forman and E. Cadenas (eds.), *Oxidative Street and Signal Transduction*. New York: Chapman and Hall, Inc., 1997.
24. Winski, S. L., Hargreaves, R. H., Butler, J., and Ross, D. A new screening system for NAD(P)H:quinone oxidoreductase (NQO1)-directed antitumor quinones: identification of a new aziridinylbenzoquinone, RH1, as a NQO1-directed antitumor agent. *Clin. Cancer Res.*, *4*: 3083–3088, 1998.
25. Belinsky, M., and Jaiswal, A. K. NAD(P)H: quinone oxidoreductase 1 (DT-diaphorase) expression in normal and tumor tissues. *Cancer Metastasis Rev.*, *12*: 103–117, 1993.
26. Schlager, J. J., and Powis, G. Cytosolic NAD(P)H oxidoreductase in human normal and tumor tissues: effects of cigarette smoking and alcohol. *Int. J. Cancer*, *45*: 403–409, 1990.