**Genetic Epidemiology 14:181–197 (1997)**

# Genetic Analysis of Families With Nonsmoking Lung Cancer Probands

**P. Yang,[1]* A.G. Schwartz,[2] A.E. McAllister,[3] C.E. Aston,[3] and G.M. Swanson[4]**

[1]*Department of Health Sciences Research, Mayo Clinic and Foundation, Rochester, Minnesota*
[2]*Department of Human Genetics, MCP-Hahnemann School of Medicine, Allegheny University of Health Sciences, Pittsburgh, Pennsylvania*
[3]*Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania*
[4]*Michigan State University, East Lansing*

As part of a genetic epidemiologic study of lung cancer among nonsmokers, we investigated the role of genetic predisposition in familial aggregation. Cases were identified from the Metropolitan Detroit Cancer Surveillance System. Information on lung cancer occurrence, smoking habits (active or passive), and chronic respiratory diseases in first-degree relatives was obtained for 257 nonsmoking lung cancer probands (71 males, 186 females) diagnosed at ages 40–84 years. Among the 2,021 first-degree relatives, 24 (2.6%) males and 10 (1.1%) females were reported as having lung cancer. The occurrence of lung cancer among smoking and nonsmoking relatives was 4.5% and 1.1% in males and 2.8% and 0.4% in females, respectively. To evaluate the role of a putative Mendelian gene (one locus, two alleles) in the presence of other risk factors, we performed complex segregation analyses on the data using two different regressive model approaches [Segregation Analysis of a Discrete Trait Under a Class A Regressive Logistic Model, V4.0 (REGD) and Segregation Analysis of a Truncated Trait, V2.0, Model I (REGTL)] as implemented in the Statistical Analysis for Genetic Epidemiology (SAGE) program. Using either approach, an environmental model best explained the observed lung cancer aggregation in families ascertained through nonsmoking probands. Based on our final model, only 0.04% of this population had a very high risk and 4.2% had a moderate risk of lung cancer. The rest of the

Contract grant sponsor: NIH; Contract grant numbers: R29-CA50383, R01-CA60691; Contract grant sponsor: NIOSH; Contract grant number: R01-OH02067; Contract grant sponsor: NCI; Contract grant number: N01-CN05225.

*Correspondence to: Dr. Ping Yang, Department of Health Sciences Research, Mayo Clinic, 200 First Street S.W., Rochester, MN 55905.

Received 10 June 1996; Revised 30 October 1996; Accepted 5 November 1996

population had virtually no risk of lung cancer during their lifetime unless they have multiple risk factors. Among the high-risk individuals without any risk factor under study, the estimated risks at ages 40, 60, and 80 years in males were 16.7%, 83.6%, and 95.4%, and in females were 14.0%, 72.2%, and 88.0%, respectively. Among at-moderate-risk smokers the estimated risks at the same age and gender groups were essentially the same as in the high-risk nonsmokers. Our results suggest that the pattern of lung cancer occurrence in families of nonsmoking lung cancer patients differs from that in families of smoking lung cancer patients. Despite the profound effect of smoking on the risk of lung cancer, other environmental and/or genetic risk factors need to be identified. Genet. Epidemiol. 14:181–197, 1997. 

**Key words: lung cancer; tobacco exposure; chronic lung diseases; familial aggregation**

## INTRODUCTION

Lung cancer has long been deemed an environmentally caused disease with tobacco smoking as the major cause [Levin et al., 1950; Doll and Hill, 1952]. The risk for smokers has been estimated to be more than double in females and over 8 times in males with a strong dose-response relationship [Kahn, 1966; Hammond, 1972; Doll and Peto, 1978]. Not only smokers themselves, but also passive smokers are at a 30–70% increased risk of developing lung cancer [Blot and Fraumeni, 1986; Saracci and Riboli, 1989; Fontham et al., 1991; Brownson et al., 1992; U.S. Environmental Protection Agency, 1992]. Other recognized environmental risk factors include air pollution and occupational exposures to asbestos, radon, mustard gas, polycyclic hydrocarbons, chloromethyl ethers, chromium, and inorganic arsenic [Hammond et al., 1979; Fraumeni and Blot, 1982; Swanson, 1988]. Both smokers and nonsmokers with other chronic lung diseases have also been shown to have an increased lung cancer risk [Osann, 1991; Samet et al., 1986; Schwartz et al., 1996; Wu et al., 1988, 1995]. Estimated risk associated with chronic bronchitis, emphysema, or both is approximately 2-fold [Kuller et al., 1990; Nomura et al., 1991; Skillrud et al., 1986; Tockman et al., 1987; Islam and Schottenfeld, 1994], and the risk associated with impaired lung function has been shown to increase linearly with decreasing forced expiratory volume in 1 sec (FEV1) [Kuller et al., 1990; Islam and Schottenfeld, 1994].

Although tobacco exposure accounts for over 80% of the lung cancer incidence, familial aggregation of this disease has been reported in several studies, all of which indicated an underlying genetic susceptibility [Amos et al., 1992; Lynch et al., 1986; Mack et al., 1990; McDuffie, 1991; Osann, 1991; Ooi et al., 1986; Shaw et al., 1991; Tokuhata and Lilienfeld, 1963; Wu et al., 1988]. The only population study providing strong evidence of Mendelian inheritance in lung cancer was that of Sellers et al. [1990, 1992a]. After adjusting for age, sex, smoking history, and occupational exposures for each relative, a 2.4-fold excess of lung cancer was reported among relatives of lung cancer cases compared to the relatives of spouse controls [Ooi et al., 1986]. The pattern of lung cancer occurrence in these families was consistent with Mendelian codominant inheritance for early age-at-onset of a rare autosomal gene [Sellers et al., 1990, 1992a].

We investigated the role of a putative Mendelian genetic factor and, simulta-

neously, the effects of cigarette smoking, passive smoking, and chronic lung diseases in lung cancer risk among relatives of nonsmoking probands. This study is part of a population-based study of lung cancer in Metropolitan Detroit [Illis et al., 1987; Schwartz et al., 1996; Swanson et al., 1985].

## METHODS

### Case Identification

Population-based cases were identified from the Metropolitan Detroit Cancer Surveillance System (MDCSS), a participant in the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program [Ries et al., 1994]. Detailed information on lung cancer occurrence, smoking habits (active or passive, cigarettes or other type), and history of other lung diseases in first-degree relatives was obtained for 257 nonsmoking probands (71 males, 186 females) with lung cancer diagnosed between ages 40 and 84 years. Eligible cases were newly diagnosed, nonsmoking, African American, and Caucasian lung cancer cases who were identified retrospectively after their participation in the Occupational Cancer Incidence Surveillance Study (OCISS) [Illis et al., 1987; Swanson et al., 1985]. OCISS was developed to monitor occupational cancer occurrence among residents of metropolitan Detroit, functioning in conjunction with the MDCSS. OCISS subjects were identified with cancers of the lung, colon, rectum, urinary bladder, esophagus, eye, liver, salivary gland, stomach, or mesothelioma. For the present study, 401 (14.8%) lung cancers diagnosed among noncigarette smokers from November 1, 1984, through June 30, 1987, were eligible for inclusion from the 5,953 completed OCISS interviews (93.6% of the 6,359 lung cancer cases identified).

### Data Collection

These cases or their proxies reported in the OCISS that they did not smoke cigarettes at any time in their lives. Smoking status was verified in this study. Additional contacts within the family were made if any discrepancies in the reports of smoking status occurred. Eight cases were excluded because of uncertainties in smoking status. For analysis, we further restricted the sample by excluding cigar and pipe smokers, leaving 314 eligible cases. Interviews, designed for the current study and conducted by telephone, were completed for 257 (81.9%) of the eligible cases. Probands or proxies for interview could not be located for 11.5% of the cases and 6.7% refused to participate. Due to the retrospective nature of case identification and the high case fatality associated with lung cancer, 83% of the case family interviews had to be conducted with proxies. Approximately 84% of the proxies completing interviews for cases were either spouses, siblings, offspring, or parents. For 24% of the case subjects, more than one individual was interviewed to obtain complete information.

The telephone interview included questions about health history, smoking history, environmental tobacco smoke exposure, and occupational history. Family history included questions about the age, sex, race, birth year, residence, health history, cancer status, vital status, age at death, place of death, cause of death, smoking history, environmental tobacco smoke exposure history, and usual occupation and industry for each first-degree relative (parents, siblings, children) and spouse of the

cases. For those relatives with cancer and for those relatives who had died, additional questions were asked pertaining to when and where these events occurred. Questionnaire data for 2,219 family members (2,021 first-degree relatives and 198 spouses) of nonsmoking cases were included for analysis.

## Analysis

To evaluate the role of a possible Mendelian diallelic gene, the effects of tobacco exposure and other chronic lung diseases, we performed complex segregation analyses on the data using two different approaches: REGD (Segregation Analysis of a Discrete Trait Under a Class A Regressive Logistic Model, V4.0) and REGTL (Segregation Analysis of a Truncated Trait, V2.0, Model I) implemented in the Statistical Analysis for Genetic Epidemiology (SAGE) [1991] program. REGD incorporates unobserved "type" effects (e.g., genotypes) into the logistic regression, i.e.,

$$\theta_i = \log[P(Y_i = 1) / 1 - P(Y_i = 1)] = \beta_j + \sum_{k=1}^{v} \xi_k x_{ik} \tag{1}$$

where $\theta$ is the log-odds of the probability of individual $i$ having lung cancer ($Y_i$), $\beta_j$ is the specific "type" with a frequency of $q$, such as $\beta_{aa}$, $\beta_{ab}$, or $\beta_{bb}$ in the case of a Mendelian diallelic gene, and the $\xi_k$ are $v$ additional parameters to estimate such as tobacco use and previous chronic pulmonary diseases. Eq. (1) implies that the penetrance, which is the probability of having lung cancer for a given genotype and specified risk factor(s), can be calculated as

$$P(Y_i = 1 \mid \beta_j) = \exp(\theta_i)/[1 + \exp(\theta_i)]. \tag{2}$$

To account for the probability of affection in the general population, the probability of having lung cancer for each individual in a given age stratum and sex (based on SEER data), $p_i$, was included in all models as a fixed covariate shown in Eq. (3). This method of age adjustment is free of assumptions regarding observed age-of-onset distribution and has been shown to be as robust as the variance component approach by unified mixed models [Lustbader et al., 1992; Lalouel and Morton, 1981]:

$$\theta_i = \beta_j + \log[p_i / (1 - p_i)] + \sum_{k=1}^{v} \xi_k x_{ik}. \tag{3}$$

Under the assumption of Mendelian inheritance, the transmission probabilities ($\tau_{aa}$, $\tau_{ab}$, $\tau_{bb}$), which are the probabilities of a parent transmitting the high-risk allele $a$ to a child, are defined as 1, 0.5, and 0, respectively. To specifically test the hypothesis of Mendelian transmission, two environmental models were included: one assumes homogeneous risk among all family members (no transmission or $\tau = q$) and the other allows differential risks between founders (individuals without parents or married-in in the pedigree) and nonfounders ($\tau_{aa} = \tau_{ab} = \tau_{bb}$), and can detect an unmeasured etiologic factor (indicated by the parameters $q$ and type-specific $\beta_j$).

REGTL (Model I) models the putative predisposition as having a type-dependent age-of-onset distribution and a susceptibility parameter (lifetime risk of being

affected) common to all types. Under the dominant model, individuals of types *aa* and *ab*, who carry the susceptibility allele *a*, have the same age-specific probabilities of cancer onset. Individuals with genotype *bb*, who do not carry the susceptibility allele, have relatively smaller age-specific probabilities of lung cancer because of a shift of the age-of-onset distribution such that they have an older mean age-of-onset. Under a recessive model, only individuals of type *aa* have an earlier mean age of onset, the age-of-onset distribution being the same for *ab* and *bb* individuals. Under a model with no dominance effects (codominant model), the three types of individuals have separate mean ages of onset.

Before we modeled major genetic effects, the following covariates were incorporated into the baseline or sporadic model, which is equivalent to an unconditional multivariate logistic regression model [SAS Institute, Inc., 1990]: age at diagnosis/interview (as a continuous variable); race (dichotomized as African American or Caucasian); sex; education (dichotomized as less than high school graduate and high school graduate or more); personal history at least 1 year before diagnosis/interview of allergies, pneumonia, tuberculosis (TB), chronic obstructive pulmonary disease (COPD that includes emphysema and/or chronic bronchitis), and asthma; exposure to environmental tobacco smoke at home and/or at work; usual occupation and usual industry. Occupation and industry were coded using 1980 U.S. Bureau of Census [1981] classification codes and grouped into 84 occupations and 90 industries. Odds ratio (OR) estimates of relative risk were calculated from the regression coefficients the same way as in conventional logistic models [Schlesselman, 1982].

After the initial modeling of all potential lung cancer risk factors which were collected in the interview, only 5 were retained in the subsequent regressive logistic analysis (see Table V): smoking, passive smoking at work and/or at home, allergy, TB, and COPD. Although passive smoking and TB were of borderline significance in the sporadic model, they were statistically significant in our earlier case-control analysis [Schwartz et al., 1996] and their effects were much greater (for passive smoking) or became statistically significant (for TB) in the genetic and/or environmental models (see Table V).

## Hypothesis Testing

Both REGD and REGTL models have been enhanced to allow for, in addition to the transmission of a major gene, multifactorial transmission from one generation to the next [Bonney, 1986; Elston and George, 1989]. For each model fitted, the likelihood was maximized over all unknown parameters, including the regressive coefficients. The likelihood of the data under the above models was conditioned on the likelihood of the phenotype of the proband (assuming one proband per family) to allow for the ascertainment of families via probands [Cannings and Thompson, 1977].

We tested a series of competing models including a sporadic model, Mendelian models (including dominant, recessive, and codominant), and environmental models (including measured and unmeasured risk factors). The natural logarithm ($\log_e$) of likelihood (Ln L) of a baseline model was calculated and compared with the hypothesis-bearing models specified above with one or more pertinent parameters restricted. To test a hypothesis about a specific mode of inheritance, the likelihood ratio test (LRT) statistic

$$LRT = -2(\ln L_{baseline} - Ln L_{specific}) \quad (4)$$

was used, where "specific" indicates the model for a specified hypothesis. The sampling distribution of this statistic is well approximated by a chi-square with *n-k* degrees of freedom where *n* and *k* equal the number of parameters estimated in the baseline model and the specified model, respectively. An alternative LRT method in determining the best model is to compare the $\ln L_{specified}$ to the $\ln L_{unrestricted}$ in which all parameters are adjusted to provide a general unrestricted model. When more than one model was rejected against the baseline model as in Eq. (4), or was not rejected against the unrestricted model by this alternative LRT, the one with the lowest Akaike information criteria (AIC), where

$$AIC = -2[\log(L) + 2k]$$

was considered the best model [Akaike, 1974].

## RESULTS

In a total of 2,219 family members of 257 nonsmoking lung cancer probands, 198 were spouses. Among the 2,021 first-degree relatives, 24 males (18 cigarette smokers, 6 nonsmokers) and 10 females (6 smokers, 3 nonsmokers, 1 unknown smoking status) were reported as having lung cancer. Four male spouses were also reported to have lung cancer. The frequencies of the reported lung cancer occurrence in specific relative groups, which have been described in Schwartz et al. [1996], were 2.1, 2.0, and 1.0 for parents, siblings, and offspring, respectively. Table I summarizes the observed lung cancer frequency in first-degree relatives by 10-year age groups. The male-to-female ratio of 2.4:1 among these relatives is consistent with the population average of 2.5:1 for lung cancer [Schottenfeld, 1996], albeit countered by the 1:2.6 ratio in the probands. Smoking history and mean age at diagnosis of the probands and their affected relatives are shown in Table II. For the probands, who were all nonsmokers by study design, there was no significant difference in

**TABLE I. Age-of-Onset Distribution of Lung Cancer Reported in First-Degree Relatives of 257 Nonsmoking Probands***

| | Frequency of lung cancer | | | | | |
|---|---|---|---|---|---|---|
| | Male relatives | | | Females relatives | | |
| Age group (years) | No. of relatives | No. with lung cancer | % | No. of relatives | No. with lung cancer | % |
| 30–39 | 87 | 1 | 1.1 | 80 | 0 | — |
| 40–49 | 129 | 1 | 0.8 | 126 | 0 | — |
| 50–59 | 178 | 6 | 3.4 | 123 | 4 | 3.3 |
| 60–69 | 178 | 7 | 3.9 | 168 | 2 | 1.2 |
| 70–79 | 162 | 5 | 3.1 | 164 | 1 | 0.6 |
| ≥80 | 115 | 3 | 2.6 | 158 | 1 | 0.6 |
| Unknown | 87 | 1 | 1.1 | 86 | 2 | 2.3 |
| Total | 936 | 24 | 2.6 | 905 | 10 | 1.1 |

*Only relatives 80 years of age are included in this table.

**TABLE II. Mean Age-of-Diagnosis for Lung Cancer in Probands and Relatives by History of Tobacco Exposure**

| | Mean age ± standard deviations (years) | | | |
|---|---|---|---|---|
| | 257 probands | | 34 relatives | |
| Tobacco exposure | Male (N = 71) | Female (N = 186) | Male (N = 24) | Female (N = 10) |
| Cigarette smoking | | | | |
| Yes | NA[a] | NA | 64.7 ± 13.2 | 55.3 ± 6.2* |
| No | 67.5 ± 10.0 | 69.8 ± 9.8 | 59.0 ± 14.2 | 87.7 ± 12.3* |
| Passive smoking[b] | | | | |
| Yes | 66.7 ± 10.4 | 69.5 ± 9.6 | 61.9 ± 10.3 | 64.3 ± 15.2 |
| No | 70.1 ± 8.4 | 70.8 ± 10.9 | 73.0 ± 8.5 | 69.7 ± 25.5 |

[a]NA = not applicable.
[b]Yes = passive smoking either at work or at home; no = neither.
*$P < 0.001$ by a two-tailed t-test of the difference between these two means.

mean age at diagnosis between passive and nonsmokers or between male and female patients. Whereas the same was true for affected male relatives, we observed a striking difference in mean age of onset between female smoking and nonsmoking lung cancer relatives, i.e., 55 and 88 years, respectively ($P < 0.001$). The small number of affected nonsmoking relatives warrants the cautious interpretation of these results.

The distributions of risk factors under evaluation in this study, i.e., history of active or passive tobacco smoke exposure and chronic lung diseases including allergy, TB, emphysema, and chronic bronchitis, are presented in Tables III and IV. Affected relatives were 2–3 times more likely to smoke than unaffected relatives (Table III), i.e., 75% vs. 44% in males and 67% vs. 23% in females, respectively. At-home exposure to tobacco smoke was comparable between males and females, and between affected and unaffected relatives (from 42% to 50% as shown in Table III). A higher percentage (74%) of affected male relatives had at-work exposure than did unaffected male relatives (49%). In Table IV, one can clearly see that more affected relatives had a history of various chronic lung diseases than did unaffected relatives, especially for chronic bronchitis and emphysema.

To account for the above-mentioned risk factors in our search for a putative Mendelian genetic mechanism in lung cancer families, we applied regressive logistic models to perform complex segregation analysis using two different approaches, REGD and REGTL. Using either approach, an environmental model with homogeneous risk across generations best explained the observed data. Two competing hypotheses were rejected by both LRT and AIC information criteria: 1) the occurrence of lung cancer among first-degree relatives is sporadic; and 2) there is a genetic predisposition (i.e., Mendelian recessive, dominant, or codominant locus) accounting for the familial aggregation. Based on the REGD results (Table V), by both LRT and AIC, the best model was an environmental model, in which the risk type was identically distributed among all individuals and the chance of being a high-risk child was independent of the parent's risk type (Table V, model 6). From this model, which is based on family members of nonsmoking probands, a very small subgroup of this population (4/10,000) was predicted to be at a very high risk of developing lung cancer determined by an unmeasured, yet not a simple Mendelian genetic factor.

**TABLE III. Distribution of Cigarette Smoking Exposures in Relatives of 257 Nonsmoking Probands**

| Tobacco exposure | No. (%) of relatives[a] without lung cancer | | | No. (%) of relatives with lung cancer | | |
|---|---|---|---|---|---|---|
| | Male | Female | All | Male | Female | All |
| Cigarette smoking | | | | | | |
| Yes | 404 (43.6) | 216 (22.9) | 620 (33.2) | 18 (75.0)* | 6 (66.7)** | 24 (85.7) |
| No | 522 (56.4) | 726 (77.1) | 1,248 (66.8) | 6 (25.0) | 3 (33.3) | 4 (14.3) |
| Unknown[b] | 80 | 39 | 119 | 0 | 1 | 1 |
| Passive tobacco exposure | | | | | | |
| Passive smoking at work | | | | | | |
| Yes | 377 (49.0) | 196 (23.4) | 573 (35.6) | 14 (73.7)*** | 2 (22.2) | 16 (57.1) |
| No | 393 (51.0) | 643 (76.6) | 1,036 (64.4) | 5 (26.3) | 7 (77.8) | 12 (42.9) |
| Unknown | 236 | 142 | 378 | 5 | 1 | 6 |
| Passive smoking at home | | | | | | |
| Yes | 379 (43.0) | 490 (55.3) | 869 (49.2) | 13 (56.5) | 4 (50.0) | 17 (54.8) |
| No | 502 (57.0) | 396 (44.7) | 898 (50.8) | 10 (43.5) | 4 (50.0) | 14 (45.2) |
| Unknown | 125 | 95 | 220 | 1 | 2 | 3 |
| At work or at home | | | | | | |
| Either | 560 (69.8) | 544 (64.0) | 1,104 (66.8) | 19 (90.5) | 6 (60.0) | 25 (83.3) |
| Neither | 242 (30.2) | 306 (36.0) | 548 (33.2) | 2 (9.5) | 3 (40.0) | 5 (16.6) |
| Unknown | 204 | 131 | 335 | 3 | 1 | 4 |
| Total | 1,006 | 1,029 | 1,987 | 24 | 10 | 34 |

[a]Four relatives with sex unknown were excluded from this table.
[b]Unknowns were not included in percentage calculations for each risk factor.
*$P < 0.001$ when compared to the relatives without lung cancer.
**$P < 0.005$ when compared to the relatives without lung cancer.
***$P < 0.05$ when compared to the relatives without lung cancer.

Another 4.2% of the population was estimated to be at a moderate risk of lung cancer, and the rest of the population had virtually no risk of lung cancer during their lifetime (group 5) unless, as shown in the last row of Table VI, they have multiple risk factors. Table VI also highlights the risks of lung cancer between individuals in different risk groups and between individuals with various risk factors. Among high-risk nonsmokers (group 1), the estimated risks at age 40, 60, and 80 years in males were estimated to be 16.7%, 83.6%, and 95.4%, and in females were 14.0%, 72.2%, and 88.0%, respectively. Although the moderate-risk nonsmokers (group 2) had a very low baseline risk at these age and gender groups, their age- and gender-specific risks are equivalent to those of high-risk nonsmokers if they were either smokers (group 3) or they had a history of COPD and were exposed to environmental tobacco smoke (group 4).

Figure 1a,b illustrates the best model (from REGD) for females and males respectively. For normal nonsmoking individuals who were in the moderate-risk group, the age-specific risk of lung cancer is minimal until late in life, 11% for males and 4% for females at age 80 years. The same gender-specific lung cancer risks would be observed at the age of 40 years if those nonsmokers had a history of TB and would be observed at the age of 35 years if among smokers. Obviously, the risks were tremendously increased for all ages in smokers who had a history of TB. Al-

**TABLE IV. Distribution of Chronic Lung Diseases in Relatives of 257 Nonsmoking Probands**

| Chronic lung diseases | No. (%) of relatives[a] without lung cancer | | | No. (%) of relatives with lung cancer | | |
|---|---|---|---|---|---|---|
| | Male | Female | All | Male | Female | All |
| Allergy | | | | | | |
| Yes | 26 (2.9) | 23 (2.6) | 49 (2.8) | 3 (12.5) | 0 (0) | 3 (9.4) |
| No | 872 (97.1) | 854 (97.4) | 1,726 (97.2) | 21 (87.5) | 8 (100) | 29 (90.6) |
| Unknown[b] | 108 | 106 | 214 | 0 | 2 | 2 |
| Chronic bronchitis | | | | | | |
| Yes | 24 (2.7) | 23 (2.6) | 47 (2.6) | 5 (16.7)* | 1 (12.5) | 5 (15.6) |
| No | 873 (97.3) | 855 (97.4) | 1,728 (97.4) | 20 (83.3) | 7 (87.5) | 27 (84.4) |
| Unknown | 109 | 103 | 212 | 0 | 2 | 2 |
| Emphysema | | | | | | |
| Yes | 24 (2.7) | 11 (1.2) | 54 (2.7) | 3 (12.5)** | 1 (12.5) | 4 (12.5) |
| No | 874 (97.3) | 875 (98.8) | 1,949 (97.3) | 21 (87.5) | 7 (87.5) | 28 (87.5) |
| Unknown | 108 | 95 | 209 | 0 | 2 | 2 |
| Unspecified obstructive disease[c] | | | | | | |
| Yes | 5 (0.6) | 6 (0.7) | 11 (0.6) | 0 (0) | 0 (0) | 0 (0) |
| No | 891 (99.4) | 883 (99.3) | 1,774 (99.4) | 24 (100) | 8 (100) | 32 (100) |
| Unknown | 110 | 92 | 202 | 0 | 2 | 2 |
| TB | | | | | | |
| Yes | 14 (1.6) | 7 (0.8) | 21 (1.2) | 1 (4.3) | 1 (12.5) | 2 (6.5) |
| No | 888 (98.4) | 879 (99.2) | 1,767 (98.8) | 22 (95.7) | 7 (87.5) | 29 (93.5) |
| Unknown | 104 | 95 | 199 | 1 | 2 | 3 |
| Total | 1,006 | 981 | 1,987 | 24 | 10 | 34 |

[a]Four relatives with sex unknown were excluded from this table.
[b]Unknowns were not included in percentage calculations for each risk factor.
[c]COPD was a nonspecific group which could be either or both chronic bronchitis and emphysema. Individuals with a specific diagnosis of chronic bronchitis or emphysema were not counted in this group.
*$P < 0.001$ when compared to the relatives without lung cancer.
**$P < 0.05$ when compared to the relatives without lung cancer.

though TB was a significant risk factor in the final model, its rarity as a disease calls for careful parameter interpretation.

REGTL results were consistent with the REGD results in rejecting all of the Mendelian inheritance models but differed from REGD results in the estimated effects of other chronic lung diseases. The effect of TB on lung cancer risk was not significant in any of the models, while allergy and COPD were significant risk factors using the REGTL approach. The best fit model (by both LRT and AIC) was also environmental and indicated an unmeasured risk factor which could be conferred from parent to offspring generations ($\tau_{aa} = \tau_{ab} = \tau_{bb} = 0.41 \pm 0.11$) yet independent of the parental risk type: high, moderate, or low risk.

## DISCUSSION

This study was carried out to investigate causes of lung cancer in addition to tobacco smoking by focusing on familial aggregation in nonsmoking lung cancer probands. This unique study design not only enabled us to identify risk factors other than smoking among family members of nonsmokers [Schwartz et al., 1996], it also

**TABLE V. Complex Segregation Analysis of 257 Families With a Nonsmoking Proband (REGD)**

| Model | $q_a$ | $\tau$[a] | $\beta_{aa}$ | $\beta_{ab}$ | $\beta_{bb}$ | $\varepsilon_{Smoking}$ | $\varepsilon_{Passive\ smoking}$ | $\varepsilon_{TB}$ | $\varepsilon_{Allergy}$ | $\varepsilon_{COPD}$[b] | –2lnL | AIC | df[c] | Chi-square | *P* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1) Sporadic | NA[d] | NA | –2.51 | =[e] | = | 1.38 | 0.88 | 1.58 | 1.43 | 1.10 | 300.86 | 312.86 | —[f] | | |
| 2) Recessive | 0.02 | M[g] | 4.20 | –3.77 | = | 1.70 | 1.59 | 1.84 | 0.86 | 1.37 | 283.83 | 299.83 | 2 | 17.03 | 0.0002 |
| 3) Dominant | 0.00 | M | 2.19 | = | = –4.13 | 1.58 | 2.03 | 1.79 | 1.54 | 1.33 | 291.38 | 307.38 | 2 | 9.48 | 0.0083 |
| 4) Codominant | 0.02 | M | 1.73 | –1.60 | = –4.93 | 1.69 | 1.31 | 1.57 | 1.60 | 1.42 | 291.99 | 307.99 | 2 | 8.87 | 0.0119 |
| 5) Environmental ($\tau$s equal) | 0.02 | 0.02 | 4.85 | –0.23 | –13.02 | 5.23 | 2.38 | 3.98 | 1.87 | 5.13 | 270.98 | 290.98 | 4 | 29.88 | <0.0001 ($P > 0.08$)[i] |
| 6) Environmental (no $\tau$) | 0.02 | = | 4.93 | –0.20 | –12.91 | 5.17 | 2.34 | 3.98 | 1.87 | 5.11 | 270.99[h] | 288.99[h] | 3 | 29.87 | <0.0001 ($P > 0.18$)[i] |
| 7) General unrestricted | 0.02 | ARB[j] | –129.62 | –60.86 | –249.23 | 137.59 | 64.10 | 3.92 | 1.61 | 47.56 | 266.09 | 290.09 | 6 | 34.77 | <0.0001 |

[a] $\tau$ = Transmission probability.
[b] Chronic obstructive lung diseases including chronic bronchitis, emphysema, or both.
[c] df = degrees of freedom.
[d] NA, not applicable for the given model.
[e] The parameter is set to the same value of the one on the left.
[f] Used as baseline model.
[g] M = Mendelian transmission probabilities: $\tau_{aa} = 1.0$, $\tau_{ab} = 0.5$; $\tau_{bb} = 0$.
[h] Best model according to Akaike's criteria and the LRT.
[i] *P* value by the alternative LRT where the general unrestricted model was used as the baseline.
[j] ARB = arbitrary transmission probabilities: $\tau_{aa} = 0.67$, $\tau_{ab} = 0.00$; $\tau_{bb} = 0.02$.

**TABLE VI. Estimated Lung Cancer Risk Based on the Environmental Model by Complex Segregation Analysis (REGD)**

| | Estimated lung cancer risk (%) | | | | | |
|---|---|---|---|---|---|---|
| | Males at age (in years) | | | Females at age (in years) | | |
| Risk group | 40 | 60 | 80 | 40 | 60 | 80 |
| 1) High-risk[a] nonsmokers | 16.7 | 83.6 | 95.4 | 14.0 | 72.2 | 88.0 |
| 2) Moderate-risk[b] nonsmokers | 0.1 | 2.9 | 11.0 | 0.1 | 1.5 | 4.0 |
| 3) Moderate-risk smokers | 17.3 | 84.1 | 95.6 | 14.4 | 73.0 | 88.4 |
| 4) Moderate-risk passive smokers with COPD | 16.4 | 83.3 | 95.4 | 13.7 | 71.8 | 87.7 |
| 5) Low-risk smokers with COPD and allergy | 0.1 | 0.2 | 0.7 | 0.1 | 0.1 | 0.2 |

[a]4/10,000 in the study population.
[b]4.2% in the study population

engendered the opportunity, for the first time in the literature, to test for a single Mendelian genetic component while simultaneously accounting for known risk factors in lung cancer etiology among families of nonsmoking lung cancer cases. There have been two population-based studies of lung cancer among female nonsmokers [Alavanja et al., 1992; Wu et al., 1995]. The focus of Alavanja et al. [1992] was on the comparison of preexisting lung disease between lung cancer patients and controls. The lack of detailed family information in the study by Wu et al. [1995] precluded any genetic analysis.

Several studies of familial aggregation of primarily smoking lung cancer probands have been reported and all indicated an underlying genetic susceptibility [Amos et al., 1992; Lynch et al., 1986; Mack et al., 1990; McDuffie, 1991; Osann, 1991; Ooi et al., 1986; Shaw et al., 1991; Tokuhata and Lilienfeld, 1963; Wu et al., 1988]. The study by Sellers et al. [1990, 1992a] was the only one to provide strong evidence of Mendelian inheritance in lung cancer. This study was based on families of 337 probands who were primarily smokers and died of lung cancer in a 4-year period among residents of 10 parishes in Louisiana. The pattern of lung cancer occurrence in these families was consistent with Mendelian codominant inheritance for variable age-at-onset of a rare autosomal gene by using the REGTL models [Sellers et al., 1990]. Our results, which rejected all Mendelian models, suggest that the pattern of occurrence of lung cancer in families of nonsmoking lung cancer patients differs from that in families of smoking lung cancer patients.

In an effort to detect etiologic heterogeneity in these families, Sellers et al. [1992a] found that the model of inheritance differed based on age at death of the lung cancer proband. Evidence of codominant Mendelian inheritance was much stronger in families of probands who died of lung cancer before 60 years of age, but such a model of inheritance was not apparent in families of older probands. In addition, family history varied by histologic type of lung cancer in the proband, although the findings were not statistically significant [Sellers et al., 1992b]. Etiologic heterogeneity in our study population, in terms of the histologic type and early vs. late age-at-onset of proband's lung cancer, is currently under investigation.

Our results, from either of the two approaches, REGD or REGTL, in the framework of regressive logistic models, rejected all Mendelian single gene inheritance mechanisms (dominant, codominant, or recessive). In REGD, the two environmental

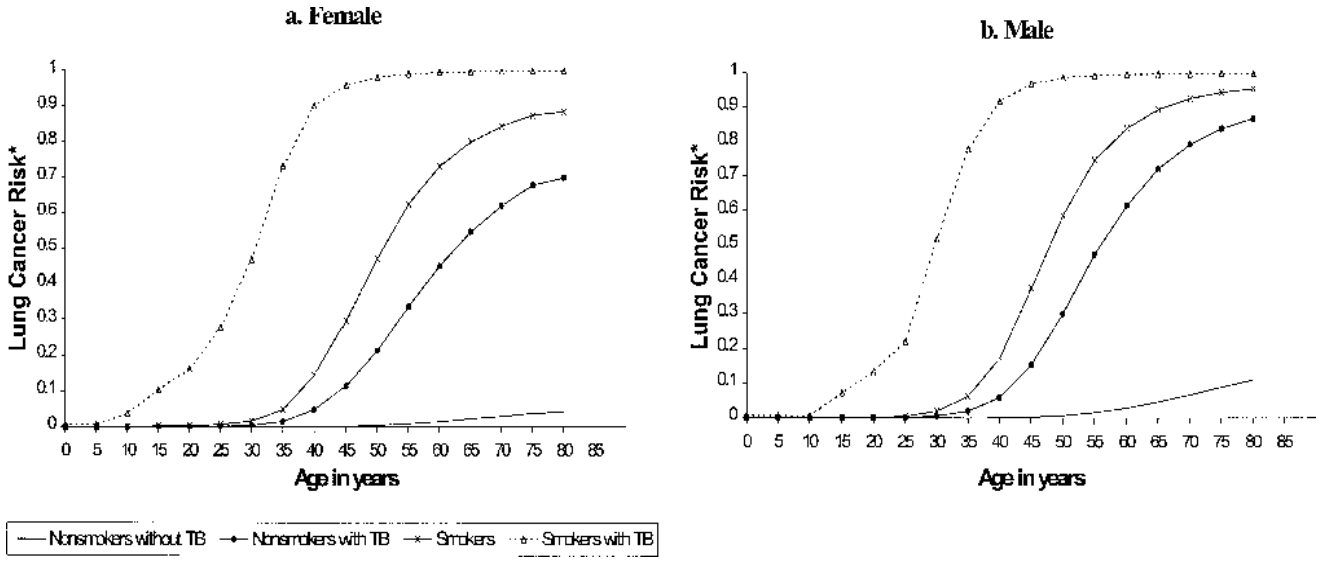


Fig. 1. Lung cancer risk among moderate-risk individuals predicted from the environmental model by complex segregation analysis (REGD). **a:** Female. **b:** Male. COPD = chronic obstructive pulmonary diseases including emphysema and chronic bronchitis. *Adjusted for passive smoking, allergy and COPD.

models were tested against the Mendelian transmission probability, one assuming homogeneous risk among all family members and the other allowing differential risks between individuals with parents and those without parents (founders or married-in individuals) in the pedigree, and converged at the same likelihood point and yielded identical results (Table V, models 5 and 6). In REGTL, the environmental model with differential risks was the best model, indicating the presence of an unmeasured, yet not a simple Mendelian factor, which could be conferred from parent to offspring. This particular environmental model was not presented by Sellers et al. [1990, 1992a]. Although we have modeled our data using two approaches and achieved reasonably consistent results, we have presented REGD models as our main results. Given the observed age-of-onset distribution (Table I), we believe that it is more appropriate to use the REGD models (which are assumption free regarding the age-of-onset distribution) than the REGTL models (which assume the age-of-onset to be well approximated by a logistic distribution).

Susceptibility to lung cancer has been postulated to be due to oncogenes or tumor suppressor genes [Caporaso et al., 1995; Ryberg et al., 1995], and due to individual susceptibility to lung carcinogens (e.g., tobacco) through differential metabolic mechanisms [Seidegard et al., 1990; Nakachi et al., 1991, 1993; Uematsu et al., 1991; Hirvonen, 1995; Kihara et al., 1995]. The interacting and counteracting relationships between genetic polymorphisms for metabolic genes such as *CYP1A1*, *CYP2D6*, *GSTM1*, and *NAT2*, combined with other risk modifiers, may explain the rejection of all Mendelian diallelic models in this study. Further analyses incorporating measured genotypes and various interactions are needed.

Tobacco use and risk of lung cancer have been studied extensively [Levin et al., 1950; Doll and Hill, 1952] and the causal relationship has been supported by a strong dose-response relationship [Kahn, 1966; Hammond, 1972; Doll and Peto, 1978]. Although the interaction between smoking and familial risk of lung cancer was recognized more than 30 years ago [Tokuhata and Lilienfeld, 1963], familial aggregation of smoking does not fully explain the familial aggregation of lung cancer [Osann, 1991]. In the present study, we took the effects of smoking into account by using a dichotomized variable instead of a more quantitative measure because we did not expect to detect different genetic effects at different levels of smoking in our study population. Moreover, using a quantitative measure of smoking such as pack-years would have resulted in the exclusion of approximately 25% of the population for whom pack-years was not available. A similar proportion of missing information regarding pack-years of smoking among relatives of the probands was also reported by Sellers et al. [1990].

Environmental tobacco exposure or passive smoking has also been demonstrated repeatedly to increase lung cancer risk [Blot and Fraumeni, 1986; Saracci and Riboli, 1989; Fontham et al., 1991; Brownson et al., 1992; U.S. Environmental Protection Agency, 1992]. Nonsmoking spouses of smokers have a 30–70% increased risk of developing lung cancer [Blot and Fraumeni, 1986; Saracci and Riboli, 1989; Fontham et al., 1991; U.S. Environmental Protection Agency, 1992]. Because exposure to environmental tobacco smoke in relatives is much harder to measure than active smoking, the only two population studies focusing on nonsmokers either did not collect the actual data at all [Alavanja et al., 1992] or estimated exposures from the probands' smoking status [Wu et al., 1995]. In

this study, we obtained information on passive smoking for each relative at home and at work from interviews with the probands or next-of-kin proxies. The possible confounding effects of such exposures have been adjusted for in all of the models that we have tested.

Previous chronic lung disease has been shown to increase lung cancer risk in both smokers and nonsmokers [Alavanja et al., 1992; Van Der Wal et al., 1966; Osann, 1991; Nomura et al., 1991; Wu et al., 1988, 1995; Samet et al., 1986]. To date, our study is the first in which a detailed history of previous lung diseases including COPD, TB, allergy, asthma, and pneumonia was obtained for all first-degree relatives, and each of these diseases has been evaluated as covariates in the complex segregation analysis. Although only TB was a statistically significant risk factor for lung cancer, the potential effects of allergy and COPD were adjusted for in all models in our analysis (Table IV).

The specific association between TB and lung cancer has been reported by Steinetz [1965], Hinds et al. [1982], and Zheng et al. [1987]. Alavanja et al. [1992] reported a 2-fold increase of lung cancer risk among former smokers who had TB. Ger et al. [1993] found that history of TB was a significant risk factor to squamous and small cell carcinomas of the lung. In our study population, we found a 6-fold increased lung cancer risk in relatives of nonsmokers who had TB when compared with relatives without TB. Chronic inflammation and scarring of the lung and an inherited susceptibility have been suggested as potential mechanisms through which TB may increase lung cancer risk.

Potential hazardous exposures related to occupation and industry were evaluated under the following job titles [based on U.S. Bureau of Census, 1981]: assembly, automaker, chemical, construction, and mechanics. None of these occupations and industries showed any significant effects on lung cancer risk in families of nonsmoking cases, although some significant findings were reported in the related case-control study [Schwartz et al., 1996].

In conclusion, the implications from our results are 2-fold: despite the profound effect of smoking on the risk of lung cancer, other environmental and/or genetic risk factors need to be characterized to identify the small subgroup of the study population (<5%) who are very or moderately susceptible to lung cancer. On the other hand, elimination of tobacco exposure and prevention of other chronic lung diseases (i.e., TB, allergy, and COPD) are evident measures for minimizing lung cancer incidence in the population.

## ACKNOWLEDGMENTS

We feel most grateful to all who contributed to the accomplishment of this work: all the study participants, study manager Mr. Michael Rothrock, interviewer Ms. Rochelle Steiner, and computer programmer Mr. Blair Powell. This study was supported by NIH grants R29-CA50383 (A.G.S.) and R01-CA60691 (A.G.S.), NIOSH grant R01-OH02067 (G.M.S.), and NCI contract N01-CN05225 (G.M.S.). Some of the results reported in this work were obtained by using the SAGE program package which is supported by U.S. Public Health Service Resource grant RR03655 from the Division of Research Resources.

## REFERENCES


Akaike H (1974): A new look at the statistical model identification. IEEE Trans Autom Control AC 19:716–723.

Alavanja MCR, Brownson RC, Boice JD, Hock E (1992): Preexisting lung disease and lung cancer among nonsmoking women. Am J Epidemiol 136:623–632.

Amos CI, Caporaso NE, Weston A (1992): Host factors in lung cancer risk: A review of interdisciplinary studies. Cancer Epidemiol Biomarkers Prev 1:505–513.

Blot WJ, Fraumeni JF (1986): Passive smoking and lung cancer. J Natl Cancer Inst 77:993–1000.

Bonney GE (1986): Regressive logistics models for familial diseases and other binary traits. Biometrics 42:611–625.

Brownson RC, Alavanja MCR, Hock ET, Loy TS (1992): Passive smoking and lung cancer in nonsmoking women. Am J Public Health 82:1525–1530.

Cannings C, Thompson EA (1977): Ascertainment in the sequential sampling of pedigrees. Clin Genet 12:208–212.

Caporaso N, DeBaun MR, Rothman N (1995): Lung cancer and CYP2D6 (the debrisoquine polymorphism): Sources of heterogeneity in the proposed association. Pharmacogenetics 5:S129–S134.

Doll R, Hill AB (1952): A study of the aetiology of carcinoma of the lung. Br Med J 2:1271–1286.

Doll R, Peto R (1978): Cigarette smoking and bronchial carcinoma: Dose and time relationships among regular smokers and lifelong nonsmokers. J Epidemiol Community Health 33(4):303–313.

Elston RC, George VT (1989): Age of onset, age at examination, and other covariates in the analysis of family data. Genet Epidemiol 6:217–220.

Fontham ETH, Correa P, Wu-Williams A, Reynolds P, Greenberg RS, Buffler PA, Chen VW, Boyd P, Alterman T, Austin DF, Liff J, Greenberg DS (1991): Lung cancer in nonsmoking women: A multicenter case-control study. Cancer Epidemiol Biomarkers Prev 1:33–43.

Fraumeni JF, Blot WJ (1982): Lung and pleura. In Schottenfeld D, Fraumeni JF (eds): "Cancer Epidemiology and Prevention." Philadelphia: W.B. Saunders, pp 564–582.

Ger LP, Hsu WL, Chen KT, Chen CJ (1993): Risk factor of lung cancer by histologic category in Taiwan. Anticancer Res 13(5A):1491–1500.

Hammond EC (1972): Smoking habits and air pollution in relation to lung cancer. In Lee DHK (ed): "Environmental Factors in Respiratory Disease." New York: Academic Press, pp 177–198.

Hammond EC, Selikoff IJ, Seidman H (1979): Asbestos exposure, cigarette smoking and death rates. Ann NY Acad Sci 330:473–490.

Hinds MW, Cohen HI, Kolonel LN (1982): Tuberculosis and lung cancer risk in nonsmoking women. Am Rev Respir Dis 125:776–778.

Hirvonen A (1995): Genetic factors in individual responses to environmental exposures. J Occup Environ Med 37:37–43.

Illis WR, Swanson GM, Satariano ER, Schwartz AG (1987): Summary measures of occupational history: A comparison of latest occupation and industry with usual occupation and industry. Am J Public Health 77:1532–1534.

Islam SS, Schottenfeld D (1994): Declining FEV1 and chronic productive cough in cigarette smokers: A 25 year prospective study of lung cancer incidence in Tecumseh, Michigan. Cancer Epidemiol Biomarkers Prev 3:289–298.

Kahn HA (1966): The Dorn Study of smoking and mortality among U.S. veterans: Report on 8½ years of observation. NCI Monogr 19:1–125.

Kihara M, Kihara M, Noda K (1995): Risk of smoking for squamous and small cell carcinomas of the lung modulated by combinations of CYP1A1 and GSTM1 gene polymorphisms in a Japanese population. Carcinogenesis 16:2331–2336.

Kuller LH, Ockene J, Meilahn E, Svendsen KH (1990): Relation of forced expiratory volume in one second (FEV1) to lung cancer mortality in the multiple risk factor intervention trial (MRFIT). Am J Epidemiol 132:265–274.

Lalouel JM, Morton NE (1981): Complex segregation analysis with pointers. Hum Hered 31:312–321.

Levin ML, Goldstein H, Gerhardt PR (1950): Cancer and tobacco smoking: A preliminary report. JAMA 143:336–338.

Lustbader ED, Williams WR, Bondy ML, Strom S, Strong LC (1992): Segregation analysis of cancer in families of childhood soft-tissue-sarcoma patients. Am J Hum Genet 51:344–356.

Lynch HT, Kimberling WJ, Markvicka SE, Biscone KA, Lynch JF, Whorton E, Mailliard J (1986): Genetics and smoking-associated cancers. Cancer 57:1640–1646.

Mack W, Langholz B, Thomas DC (1990): Survival models for familial aggregation of cancer. Environ Health Persp 37:27–35.

McDuffie HH (1991): Clustering of cancer in families of patients with primary lung cancer. J Clin Epidemiol 44:(1)69–76.

Nakachi K, Imai K, Hayashi S, Watanabe J, Kawajiri K (1991): Genetic susceptibility of squamous cell carcinoma of the lung in relation to cigarette smoking dose. Cancer Res 51:5177–5189.

Nakachi K, Imai K, Hayashi S, Kawajiri K (1993): Polymorphisms of the CYP1A1 and glutathione S-transferase genes associated with susceptibility to lung cancer in relation to cigarette dose in the Japanese population. Cancer Res 53:2994–2999.

Nomura A, Stemmermann GN, Chyou P-H, Marcus EB, Buist AS (1991): Prospective study of pulmonary function and lung cancer. Am Rev Respir Dis 144:307–311.

Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H (1986): Increased familial risk for lung cancer. J Natl Cancer Inst 76:217–222.

Osann KE (1991): Lung cancer in women: The importance of smoking, family history of cancer, and medical history of respiratory disease. Cancer Res 51:4893–4897.

Ries LAG, Miller BA, Hankey BF, Kosary CL, Harras A, Edwards BK (eds) (1994): SEER Cancer Statistics Review, 1973–1991: Tables and Graphs, National Cancer Institute. Bethesda MD: NIH Publication No. 94-2789.

Ryberg D, Lindstedt BA, Zienolddiny S, Haugen A (1995): A hereditary genetic marker closely associated with microsatellite instability in lung cancer. Cancer Res 55:3996–3999.

Samet JM, Humble CG, Pathak DR (1986): Personal and family history of respiratory disease and lung cancer risk. Am Rev Respir Dis 134:466–470.

Saracci R, Riboli E (1989): Passive smoking and lung cancer: Current evidence and ongoing studies at the International Agency for Research on Cancer. Mutat Res 222:117–127.

SAS Institute, Inc. (1990): SAS Technical Report P-200, SAS/STAT Software: CALIS and LOGISTIC Procedures, Release 6.04. Cary, NC: SAS Institute, Inc.

Schlesselman JJ (1982): Multivariate analysis. In Schlesselman JJ (ed): "Case-Control Studies: Design, Conduct, Analysis." New York: Oxford University Press, pp 227–290.

Schottenfeld D (1996): Epidemiology of lung cancer. In Pass HI, Mitchell JB, Johnson DH, Turrisi AT (eds): "Lung Cancer Principles and Practice." Philadelphia: Lippincott-Raven, pp 305–321.

Schwartz AG, Yang P, Swanson GM (1996): Familial risk of lung cancer among nonsmokers and their relatives. Am J Epidemiol 144:554–562.

Seidegard J, Pero RW, Markowitz MM, Roush G, Miller DG, Beattie EJ (1990): Isoenzymes of glutathione transferase (class Mu) as a marker for the susceptibility to lung cancer: A follow-up study. Carcinogenesis 11:33–36.

Sellers TA, Bailey-Wilson JE, Elston RC, Wilson AF, Elston GZ, Ooi WL, Rothschild H (1990): Evidence for Mendelian inheritance in the pathogenesis of lung cancer. J Natl Cancer Inst 82:1272–1279.

Sellers TA, Bailey-Wilson JE, Potter JD, Rich SS, Rothschild H, Elston RC (1992a): Effect of cohort differences in smoking prevalence on models of lung cancer susceptibility. Genet Epidemiol 9:261–272.

Sellers TA, Elston RC, Atwood LD, Rothschild H (1992b): Lung cancer histologic type and family history of cancer. Cancer 69:86–91.

Shaw GL, Falk RT, Pickle LW, Mason TJ, Buffler PA (1991): Lung cancer risk associated with cancer in relatives. J Clin Epidemiol 44:429–437.

Skillrud DM, Offord KP, Miller RD (1986): Higher risk of lung cancer in chronic obstructive pulmonary disease. Ann Intern Med 105:503–507.

Statistical Analysis for Genetic Epidemiology (SAGE) SAGE Version 2 (1991): Department of Biometry and Genetics, Louisiana State University Medical Center.

Steinetz R (1965): Pulmonary tuberculosis and carcinoma of the lung: A survey from two population-based disease registers. Am Rev Respir Dis 92:758–766.

Swanson GM (1988): Cancer prevention in the workplace and natural environment: A review of etiology, research design, and methods of risk reduction. Cancer (Suppl) 62:1725–1746.

Swanson GM, Schwartz AG, Brown KL (1985): Population-based occupational cancer incidence surveillance. J Occup Med 27:439–444.

Tockman MS, Anthonisen NR, Wright EC, Donithan MG, the Intermittent Positive Pressure Breathing Trial Group, and Johns Hopkins Lung Project for the Early Detection of Lung Cancer (1987): Airways obstruction and the risk of lung cancer. Ann Intern Med 106:512–518.

Tokuhata GK, Lilienfeld AM (1963): Familial aggregation of lung cancer in humans. J Natl Cancer Inst 30:289–312.

Uematsu F, Kikuchi H, Motomiya M, Abe T, Sagami I, Ohmachi T, Wakui A, Kanamaru R, Watanabe M (1991): Association between restriction fragment polymorphism of the human P450IIE1 gene and susceptibility to lung cancer. Jpn J Cancer Res 82:254–256.

U.S. Bureau of Census (1981): Alphabetic Index of Industries and Occupations, 1980. Washington, DC: U.S. Government Printing Office.

U.S. Environmental Protection Agency (1992): Respiratory Health Effects of Passive Smoking: Lung Cancer and Other Disorders. EPA/600/6-90/006F. Washington, DC: U.S. Environmental Protection Agency.

Van Der Wal AM, Huizinga E, Orie NGM, Sluiter HJ, De Vries K (1966): Cancer and chronic non-specific lung disease. Scand J Respir Dis 47:161–172.

Wu AH, Yu MC, Thomas DC, Pike MC, Henderson BE (1988): Personal and family history of lung disease as risk factors for adenocarcinoma of the lung. Cancer Res 48:7279–7289.

Wu AH, Fontham ETH, Reynolds P, Greenberg RS, Buffler P, Liff J, Boyd P, Henderson BE, Correa P (1995): Previous lung disease and risk of lung cancer among lifetime nonsmoking women in the United States. Am J Epidemiol 141:1023–1032.

Zheng W, Blot WJ, Liao ML, Qang ZX, Levin LI, Zhao JJ, Fraumeni JF Jr, Gao YT (1987): Lung cancer and prior tuberculosis infection in Shanghai. Br J Cancer 56:501–504.