Genetic Epidemiology 17:253–273 (1999)

Lung Cancer Risk in Families of Nonsmoking Probands: Heterogeneity By Age at Diagnosis

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

[1]Mayo Clinic and Foundation, Rochester, Minnesota
[2]MCP-Hahnemann School of Medicine, Allegheny University of the Health Sciences, Pittsburgh, Pennsylvania
[3]University of Pittsburgh, Pittsburgh, Pennsylvania
[4]Cancer Center and College of Human Medicine, Michigan State University, East Lansing

In an earlier investigation, we did not detect a major genetic component to lung cancer in families of nonsmoking lung cancer probands. However, heterogeneity with respect to familial aggregation, based on probands' age at diagnosis, was evident. We reanalyzed our previously collected data of 257 families, stratified by age at diagnosis of the probands, using complex segregation analysis. We specifically tested the effects of a Mendelian diallelic gene, history of tobacco use, and history of selected chronic lung diseases in families with a proband diagnosed at the age of 60 years or older and in families with a younger proband (i.e., under 60 years of age). Cases were identified from the Metropolitan Detroit Cancer Surveillance System. Information on lung cancer occurrence, smoking history, and chronic respiratory diseases in first-degree relatives was obtained for 210 older probands and for 47 younger probands. In older probands' families, no evidence of a major genetic effect was detected. A history of emphysema and tobacco-smoke exposure were found to be significant risk factors. In younger probands' families, a Mendelian codominant model with significant modifying effects of smoking and chronic bronchitis best explained the observed data. Our results suggest the presence of a high-risk gene contributing to early-onset lung cancer in a population where the probands are nonsmokers. Genet. Epidemiol. 17:253–273, 1999.

Contract grant sponsor: NIH; Contract grant numbers: R29-CA50383, RO1-CA60691; Contract grant sponsor: NIOSH; Contract grant number: RO1-OH02067; Contract grant sponsor: NCI; Contract grant number: N01-CN05225.

*Correspondence to: Ping Yang, M.D., Ph.D., Department of Health Sciences Research, Mayo Clinic, 200 First Street SW, Rochester, MN 55905. E-mail: yang.ping@mayo.edu

Received for publication 26 May 1998; revision accepted 23 December 1998

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

INTRODUCTION

Although 89% of lung cancer incidence can be attributed to cigarette smoking [Shopland, 1996; Schottenfeld, 1996], evidence for genetic predisposition to lung cancer has been reported in several previous studies among both smokers and nonsmokers [Amos et al., 1992; Lynch et al., 1986; Mack et al., 1990; McDuffie, 1991; Osann, 1991; Ooi et al., 1986; Schwartz et al., 1996; Shaw et al., 1991; Tokuhata and Lilienfeld, 1963; Wu et al., 1988]. Based on a Mendelian codominant model from a study of families of 336 probands who were primarily smokers and died of lung cancer, Sellers et al. [1990] predicted that a rare autosomal codominant gene is responsible for one-third of the lung cancer cases under the age of 50 years, one-half of the cases over 60 years, and one-fifth of the cases over 70 years. In a re-analysis of the same data, Gauderman et al. [1997] reported an autosomal dominant gene that confers a 9- to 17-fold increased risk for developing lung cancer.

We investigated the role of genetic factors and the effects of cigarette smoking, passive smoking, and chronic lung diseases in lung cancer risk among relatives of *nonsmoking* probands [Schwartz et al., 1996; Yang et al., 1997a]. In the case-control analysis, Schwartz et al. [1996] revealed a bimodal distribution of lung cancer occurrence in relatives of the lung cancer probands. First-degree relatives of nonsmoking lung cancer cases age 40–59 years had an over sixfold increased risk of lung cancer compared to relatives of controls [Schwartz et al., 1996]. No increased risk was seen in relatives of probands age 60 years or more. Although our previous analysis did not detect a major genetic effect in our total study population [Yang et al., 1997a], these age-specific findings led us to investigate the etiologic heterogeneity between families with a younger age proband (< 60 years) and families with an older age proband (≥ 60 years). In this work, we tested the heterogeneity in Mendelian inheritance patterns of the 257 families by the age at diagnosis of the probands and report the results of two subgroup analyses. We investigated the role of a putative Mendelian susceptibility gene in the context of history of tobacco exposure and chronic pulmonary diseases in families with an older proband and then in families with a younger proband.

MATERIALS AND METHODS

Case Identification

Population-based cases were identified from the Metropolitan Detroit Cancer Surveillance System, a participant in the National Cancer Institute's Surveillance, Epidemiology and End Results program [Ries et al., 1994]. Detailed information on lung cancer occurrence, smoking behavior (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, 185 females) with lung cancer diagnosed between ages 40 and 84, November 1, 1984 through June 30, 1987. 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 [Illis et al., 1987; Schwartz et al., 1996; Swanson et al., 1985].

Data collection was described in detail previously [Schwartz et al., 1996]. In brief, through telephone interview, information on health and smoking history, environmental tobacco smoke exposure, family and occupational history was obtained. Family history included questions about age, gender, race, birth year, residence, health history, cancer status, vital status (if deceased, age, place, and cause of death), smoking history, environmental tobacco smoke exposure history, and usual occupation and industry for each first-degree relative (parents, full siblings, children) and spouse of the cases. For those relatives with cancer and for those relatives who had died of cancer, additional questions were asked pertaining to age and place of the diagnosis of the cancer; age, place, and cause of death. Questionnaire data for 1,820 family members (1,645 first-degree relatives and 175 spouses) of 210 probands were included in the older case families and data for 412 family members (377 first-degree relatives and 35 spouses) of 47 probands were included in the younger case group.

Analysis

To evaluate the role in lung cancer susceptibility of a possible Mendelian diallelic gene, the effects of tobacco exposure, and other chronic lung diseases, we performed complex segregation analyses 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 SAGE program [Statistical Analysis for Genetic Epidemiology (SAGE), 1991].

Heterogeneity Testing

A bimodal age-at-risk distribution was found in the lung cancer probands with regard to a history of lung cancer in at least one first-degree relative [Schwartz et al., 1996]. Excess risk of familial lung cancer was only significant among relatives of nonsmoking probands aged 40–59 years of age (relative risk = 6.1, 95% confidence interval 1.1–33.4) [Schwartz et al., 1996]. Therefore, we chose to partition the total 257 families into 210 families with probands 60 years of age and older and 47 families with probands younger than the age of 60 years. To test whether there was significant heterogeneity between these two groups with respect to the segregation results, a test for heterogeneity, consisting of partitioning the families according to predefined criteria, that is, the age of diagnosis of the lung cancer probands, was used. The heterogeneity chi-square test, as described by Williams and Anderson [1984] and Khoury et al. [1993], was performed by comparing the sum of -2lnL as determined for the subsets with the -2lnL as determined from the pooled data. The degrees of freedom for such a test is determined by the difference in the number of estimated parameters in the pooled data and the sum of the estimated parameters in the two separate subgrouped data [Beaty, 1997].

Two Alternative Regressive Models: REGD and REGTL

REGD incorporates unobserved "type" effects (e.g., genotypes) into the logistic regression, i.e.,

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

where θ is the log-odds of the probability of individual *i* having lung cancer (Y_i), β_j is the log of the odds ratio attributable to a specific "type", such as β_{aa}, β_{ab}, or β_{bb} in the case of a Mendelian diallelic gene with "type" frequencies determined by a frequency parameter *q*, and the ξ_k are *v* logs of the odds ratios attributable to the additional covariates such as tobacco use and previous chronic pulmonary diseases. Equation [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:

$$P(Y_i = 1|\beta_j) = \exp(\theta_i)/[1 + \exp(\theta_i)] \quad [2]$$

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 gender (based on Surveillance, Epidemiology and End Results data), p_i, was included in all models as a fixed covariate shown in equation [3] [Lustbader et al., 1992]. This method of age adjustment is free of assumptions regarding the observed age-of-onset distribution and has been shown to be as robust as the variance component approach by unified mixed models [Lalouel and Morton, 1981; Lustbader et al., 1992]:

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

Under the assumption of Mendelian inheritance, the transmission probabilities (τ_{aa}, τ_{ab}, τ_{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 into the pedigree) and non-founders ($\tau_{aa} = \tau_{ab} = \tau_{bb}$), and can detect an unmeasured etiologic factor (indicated by the parameters *q* and type-specific β_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 multiple logistic regression model [SAS Institute Inc., 1990]: age at diagnosis/interview (as a continuous variable); race (dichotomized as African-American or Caucasian); gender; education (dichotomized as less than high school graduate and high school graduate or more); personal history at least one year before diagnosis/inter-

view of allergies, pneumonia, tuberculosis, chronic obstructive pulmonary disease (COPD 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. Census Bureau classification codes [U.S.Bureau of the Census, 1981] and grouped into 84 occupations and 90 industries. Odds ratio estimates of relative risk were calculated from the regression coefficients as in conventional logistic models [Schlesselman, 1982]. After the initial modeling of all potential lung cancer risk factors that were collected in the interview, only five for older case families and two for younger case families were retained in the subsequent regressive logistic analysis.

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) against an unrestricted general model. The natural logarithm ($\log_e$) of likelihood (lnL) of the unrestricted model, in which all parameters are adjusted to provide a general 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

$$\mathrm{LRT} = -2(\ln L_{\mathrm{specific}} - \ln L_{\mathrm{general}}) \qquad [4]$$

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. In the situation where the LRT is not informative due to indistinguishable test results among multiple hypothesis-bearing models, an alternative LRT method is used in determining the best model. The alternative LRT is to compare the $\ln L_{\mathrm{specific}}$ to the $\ln L_{\mathrm{baseline}}$, which is the sporadic model without any assumptions regarding mode of inheritance or transmission [Lustbader et al., 1992]. When the general model as in [4] was not rejected by more than one specific model or was rejected by multiple specific models by the alternative LRT, the one with the lowest Akaike information criteria (AIC), where

$$\mathrm{AIC} = -2[\ln L + 2k]$$

was considered the best model [Akaike, 1974].

RESULTS

Evidence for Heterogeneity

The heterogeneity χ^2 test, shown in Table I, was based on the partitioning of the 257 families into two separate subgroups of families by age-at-diagnosis of the proband. Under the best fit model for all families combined, i.e., the homogeneous environmental model with five covariates [Yang et al., 1997a], the pattern of lung cancer aggregation for the two subgroups differed significantly, which supports heterogeneity in etiology [Sellers et al., 1992; Schwartz et al., 1996]. Results were the same using either approach of complex segregation analysis, i.e., REGD and REGTL. The heterogeneity χ^2 test from both approaches indicated a significant difference between the younger and the older groups ($P \leq 0.02$) when fitting the best model determined for the two groups combined [Yang et al., 1997a; Williams and Anderson, 1984; Khoury et al., 1993; Beaty, 1997].

Lung Cancer Risk in Families of Nonsmoking Probands Over Age 60

The occurrence of lung cancer for tobacco-exposed and not exposed (active or passive) relatives was 4.4 and 1.2% in males; 2.4% and zero in females, respectively. Tables II and III summarize the distribution of tobacco smoke exposures and personal history of other chronic lung diseases in 1,645 first-degree relatives of the 210 older probands. Compared to the relatives without lung cancer, relatives with lung cancer had higher frequencies of both active and passive cigarette smoke exposures (Table II). For other chronic lung diseases, excluding cancer (Table III), the male relatives with lung cancer reported a personal history of allergy and the female relatives with lung cancer reported a personal history of emphysema and tuberculosis significantly more often than non-affected relatives.

Table IV presents the results of complex segregation analysis to evaluate the role of a Mendelian diallelic gene in the presence of tobacco exposure and chronic lung diseases. Although passive smoking and tuberculosis 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 tuberculosis) in the genetic and/or environmental models (Table IV). Among the six competing models, the sporadic

TABLE I. Test of Heterogeneity in 257 Families of Nonsmoking Lung Cancer Probands Based on the Homogeneous Environmental Model With Five Covariates[a]

Family Type	REGD –2lnL	REGTL –2lnL
All families (n = 257)	270.99	463.21
Probands < 60 years of age (n = 47)	59.89	107.04
Probands ≥ 60 years of age (n = 210)	191.81	324.64
χ^2 test of heterogeneity (df)[b]	19.21 (6)	31.5 (9)
P value	0.02	0.001

[a]In this model, allele frequency q is set to be the same as τ where $\tau_{aa} = \tau_{ab} = \tau_{bb}$. Five covariates include smoking, passive smoking, tuberculosis, allergy, and emphysema.

[b]$\chi^2 = -2\ln L_{(all)} - [(-2\ln L_{(<60)}) + (2\ln L_{(\geq 60)})]$. Degrees of freedom (df) for families < 60 years of age, 3 parameters were at boundary for both REGD and REGTL, and they were not included in the calculated df.

TABLE II. Distribution of Tobacco Smoke Exposures in Relatives of 210 Nonsmoking Probands[a]

Tobacco exposure	No. (%) of relatives without lung cancer[b]			No. (%) of relatives with lung cancer		
	Male	Female	All	Male	Female	All
Cigarette smoking						
Yes	315 (42.9)	167 (21.4)	482 (31.8)	14 (73.7)[c]	4 (100.0)[d]	18 (78.3)
No	420 (57.1)	612 (78.6)	1,032 (68.2)	5 (26.3)	0 (0)	5 (21.7)
Unk	72	36	108	0	0	0
Passive tobacco exposure at work or at home						
Either	442 (69.4)	438 (62.4)	880 (65.7)	15 (93.8)[e]	4 (100.0)	19 (95.0)
Neither	195 (30.6)	264 (37.6)	459 (34.3)	1 (6.3)	0 (0)	1 (5.0)
Unk	170	113	283	3	0	3
Total	807	815	1,622	19	4	23

[a]Unk (unknowns) were not included in percentage calculations for each risk factor.
[b]Three relatives with gender unknown were excluded from this table.
[c]$P < 0.01$ when compared to the relatives without lung cancer.
[d]$P < 0.005$ when compared to the relatives without lung cancer.
[e]$P < 0.05$ when compared to the relatives without lung cancer.

model and all three Mendelian genetic models (dominant, recessive, codominant) were rejected. The two environmental models, one assuming differential risk between founders and their descendants and the other assuming homogeneous risk across all individuals in the family, fit the data equally well with almost identical estimates (models 5 and 6). In the absence of a major susceptibility gene effect, we could still define three risk categories for lung cancer in our study population (Appendix A): A high-risk group (0.1%) of the population whose risk of developing lung cancer was greater than 90% at age 80; a moderate risk group (5.8%) in which the risk varied

TABLE III. Distribution of Chronic Lung Diseases in Relatives of 210 Nonsmoking Probands[a]

Chronic diseases	No. (%) of relatives without lung cancer[b]			No. (%) of relatives with lung cancer		
	Male	Female	All	Male	Female	All
Allergy						
Yes	16 (2.2)	20 (2.7)	36 (2.5)	3 (15.8)[c]	0 (0)	3 (13.0)
No	699 (97.8)	709 (97.3)	1,408 (97.5)	16 (84.8)	4 (100)	20 (87.0)
Unk	92	86	178	0	0	0
Emphysema						
Yes	18 (2.5)	9 (1.2)	27 (2.7)	2 (10.5)	1 (25.0)[d]	3 (13.0)
No	697 (97.5)	724 (98.8)	1,421 (97.3)	17 (89.5)	3 (75.0)	20 (87.0)
Unk	92	82	174	0	0	0
Tuberculosis						
Yes	12 (1.7)	6 (8.0)	18 (1.2)	1 (5.6)	1 (25.0)[d]	2 (10.0)
No	706 (98.3)	727 (99.2)	1,433 (98.8)	17 (94.4)	3 (75.0)	20 (90.0)
Unk	89	82	171	1	0	1
Total	807	815	1,622	19	4	23

[a]Unk (unknowns) were not included in percentage calculation for each risk factor.
[b]Three relatives with gender unknown were excluded from this table.
[c]$P \leq 0.05$ when compared to the relatives without lung cancer.
[d]$P \leq 0.01$ when compared to the relatives without lung cancer.

TABLE IV. Complex Segregation Analysis of 210 Families With Probands ≥ 60 Years of Age at Diagnosis (REGD Class A, SAGE V2.1)[a]

Model	q_a	τ[b]	α_{aa}	α_{ab}	α_{bb}	$\beta_{Smoking}$	β_{ETS}	β_{TB}	$\beta_{Allergy}$	$\beta_{Emphysema}$	–2lnL	AIC	df	Chi-square	*P*-value
1 Sporadic	1.00	[NA]	–3.72	[–3.72][c]	[3.72]	1.72	1.82	1.89	1.91	0.64	216.66	228.66	6	25.89	0.0002
2 Codominant	0.02	[M]	0.13	–3.00	–6.12	2.07	2.43	1.93	2.17	1.13	212.43	228.43	4	21.66	0.0002
3 Dominant	0.00	[M]	1.03	[1.03]	–6.00	1.83	3.85	1.97	2.06	0.73	214.84	230.84	4	24.07	<0.0001
4 Recessive	0.01	[M]	3.29	–5.47	[–5.47]	1.98	3.18	2.02	2.00	0.78	209.36	225.36	4	18.59	0.0010
5 Environmental (τ's equal)	0.03	0.03	5.24	–3.12	–23.00	5.54	3.18	13.38	7.13	4.30	191.55	211.55	2	0.78	0.6771
6 Environmental (no τ)	0.03	[NA]	5.20	–3.14	–23.09	5.56	3.21	13.42	7.16	4.31	191.55	209.55[d]	3	0.78	0.8543
7 General	0.03	0.005 0.040 0.030	5.35	–3.14	–52.92	5.73	3.46	41.76	9.65	5.32	190.77	214.77[e]			

[a]Age-specific ''liability'' parameter, risk adjusted to the general population, is included as a fixed covariate (see equation [3] in Materials and Methods).
[b]Transmission probability includes τ_{aa}, τ_{ab}, and τ_{bb}.
[c][ ] Fixed parameter value based on specified model.
[d]Best model according to Akaike's criteria.
[e]Used as the baseline model for the likelihood ratio test.
[NA], not applicable in the given model; [M], Mendelian transmission probabilities: τ_{aa} = 1.0, τ_{ab}=0.5, τ_{ab}=0; df, degrees of freedom.

from 0.2 to 100% at age 80 depending on the individual's tobacco exposure and history of chronic lung diseases; and a low-risk group (94%). In the moderate-risk group, the lung cancer risk for nonsmokers at the age of 80 years was 0.7% in males and 0.2% in females. For nonsmokers who had emphysema, their lung cancer risk by age 80 was 33% in males and 15% in females; but when exposed to environmental tobacco smoke, their risk increased to 92 and 81%, respectively.

These results demonstrate that in the first-degree relatives of older lung cancer probands, controlling both active and passive cigarette-smoke exposure can dramatically decrease the lung cancer risk. For example, male relatives at age 60 years who have emphysema, which increases lung cancer risk by 53-fold (10.6% vs. 0.2%), can reduce their lung cancer risk by up to 85–89% simply by avoiding active and passive smoking during their earlier life.

Evidence for Mendelian Inheritance of Lung Cancer in Families of Younger Nonsmoking Probands

Ten lung cancer relatives were reported among the 377 first-degree relatives of younger probands (2.7%). The occurrence of lung cancer for smoking and nonsmoking relatives was 4.3 and 1.0% in males; 3.9 and 2.6% in females, respectively. Table V presents lung cancer occurrence in different groups of relatives, e.g., 3.3% in siblings and 4.6% in parents of the probands. No lung cancer was reported among the children of the probands most likely due to their young age (average age of 31 years). After thoroughly examining risk factors for lung cancer using multiple logistic regression [SAS Institute Inc., 1990] and regressive models [Statistical Analysis for Genetic Epidemiology (SAGE), 1991], only tobacco smoking and a history of chronic bronchitis remained statistically significant.

Shown in Table VI are the results of complex segregation analysis. The LRT was not informative in that multiple specific or reduced models cannot be rejected; thus both the alternative LRT and AIC were applied. The final results are consistent from both REGD and REGTL and we presented detailed model specifications and tests from REGD in Table VI. A Mendelian codominant model, with significant modifying effects of smoking and chronic bronchitis (model 4), was judged the best model

TABLE V. Lung Cancer Distribution in First-Degree Relatives of Young Age Nonsmoking Probands[a]

Family members	Sex	Number of relatives	Mean age (in years)	Relatives with lung cancer		
				Number	(%)	Mean age at diagnosis
Probands	Male	17	52.9	17	100	52.9
	Female	30	53.4	30	100	53.4
Parents	Fathers	45	71.5	2	4.4	76.0
	Mothers	42	75.0	2	4.8	82.0
Siblings	Brothers	84	53.0	3	3.6	55.3
	Sisters	66	55.4	2	3.0	55.0
Children	Sons	64	32.0	0	—	—
	Daughters	47	30.1	0	—	—

[a]Twenty-nine individuals were excluded from this table because of missing gender (n = 1) and missing age or age at diagnosis (n = 28).

TABLE VI. Complex Segregation Analysis of 47 Families With Probands < 60 Years of Age (REGD Class A, SAGE V2.1)[a]

Model	q_a	τ[b]	α_{aa}	α_{ab}	α_{bb}	$\beta_{Smoking}$	$\beta_{Bronchitis}$	–2lnL	AIC	df	Chi-square	*P*-value
1 Sporadic	1.000	[NA]	–0.95	[–0.95][c]	[–0.95]	0.99	3.73	73.31	79.31	[d]	—	—
2 Recessive	0.030	[M]	4.83	–1.71	[–1.71]	1.25	4.08	68.51	78.51	2	4.80	0.09
3 Dominant	0.001	[M]	2.24	[2.24]	–2.63	1.71	4.53	70.79	80.79	2	2.52	0.28
4 Codominant	0.004 ± 0.007[f] ±1.18	[M] ±1.56	5.01	[–0.70] ± 1.05	–6.40 ± 2.59	1.44 ± 1.18	4.05 ± 1.56	67.38	77.38[e]	2	5.93	0.05
5 Environmental (τ's equal)	0.010	0.02	3.89	3.77	–3.40	2.10	5.63	68.54	82.54	4	4.77	0.31
6 Environmental (no τ)	0.020	0.02	3.52	3.52	–4.32	2.29	6.43	68.79	78.79	2	4.52	0.10
7 General	0.010	0[g] 0.06 0.02	4.67	3.36	–4.22	2.16	6.43	68.37	84.37	5	4.94	0.55

[a]Age-specific ''liability'' parameter risk adjusted to the general population is included as a fixed covariate (see equation [3] in Materials and Methods).
[b]Transmission probability includes τ_{aa}, τ_{ab}, and τ_{bb}.
[c]Fixed parameter value based on specific model.
[d]Used as the baseline model for likelihood ratio test.
[e]Best model according to Akaike's criteria.
[f]Standard deviation for the estimated parameter.
[g]The parameter was estimated at a boundary. The df was calculated without this parameter.
[NA], not applicable in the given model; [M], Mendelian transmission probabilities: $\tau_{aa} = 1.0$, $\tau_{ab} = 0.5$, $\tau_{ab} = 0$; df, degrees of freedom.

to explain the observed data. The estimated risk allele frequency was 0.004, and, therefore, the frequency of the homozygous risk genotype occurred at a rate of 1.6/100,000 in the study population. Although having this genotype can confer a very high penetrance of early-onset lung cancer (85% in male and 74% in female at the age of 60 years), the attributable risk of lung cancer in the population is negligible (Appendix B). One percent of the study population carried the heterozygous genotype ab. Individuals with this genotype were at a relatively low lung cancer risk, i.e., 7% in males and 4% in females by age 60 if they did not have a history of smoking or chronic bronchitis. However, their risk increased to the same level of the homozygous individuals if they were smokers with chronic bronchitis. The majority of the study population (99%) carrying genotype bb was at minimal risk of lung cancer. In this latter group, the risk of developing lung cancer among smokers with chronic bronchitis by age 80 years was estimated to be 6% in males and 2% in females. Figure 1 illustrates the age-specific lung cancer risks for individuals with a heterozygous genotype, with or without chronic bronchitis, in smokers and nonsmokers. For both females and males, the effect of a history of chronic bronchitis on lung cancer risk is greater than the effects of being a smoker and these effects are additive.

Figure 2 presents the eight pedigrees with multiple lung cancer cases. In four of these families, lung cancer occurred in two generations; and in the other four, lung cancer clustered within sibships. Notably six of the eight pedigrees also had relatives with malignancies other than lung cancer. There were five colon cancer occurrences in four families, liver cancer in one, and skin cancer in another family. One proband (in P1) also had breast cancer and one relative in pedigree (P46) had both lung and brain cancer. We are in the process of verifying the diagnosis of these cancer sites.

To further identify specific pedigrees that contributed to the final model, we derived the family-specific lnL score from both analytic approaches using equation [4] (Fig. 3). On both X (REGD) and Y (REGTL) axes, families with lnL score at or near zero did not contribute to the final model; while families with negative or positive lnL provided evidence against or supportive of the final model, respectively. Obviously, families along or near the diagonal line are consistent in rejecting or approving a given model using either analytic approach. Among our 47 families, families P33, P37, P39 contributed the most to the codominant Mendelian model in both REGD and REGTL. Families P1 and P10 appeared supportive of the model only in REGTL while family P13 provided evidence against the model in REGD.

DISCUSSION

In lung cancer etiology research, the role of genetic factors and environmental determinants, especially tobacco smoking, easily obscure each other. Our study, using 257 population-based families, provided evidence for etiologic heterogeneity in lung cancer risk in families of nonsmoking probands. We were able to tease apart the genetic component and other risk factors by studying the disease in a more homogeneous group of families, i.e., by stratifying the families according to proband's age at diagnosis. The basis for such stratification was suggested by previous studies [Schwartz et al., 1996; Sellers et al., 1990].

In the families of older aged probands, a major genetic factor did not appear to play a significant role. The most important environmental factor is, as expected, to-

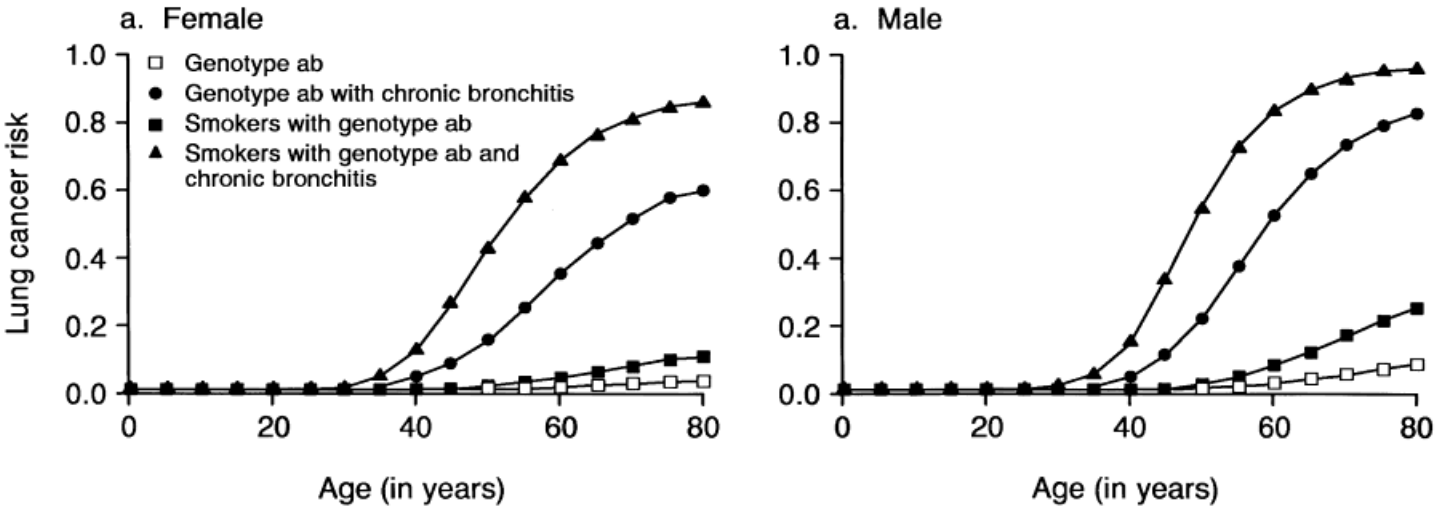

Fig. 1. Estimated lung cancer risk by age for heterozygous individuals (genotype Ab) based on a codominant Mendelian model using REGD class A.

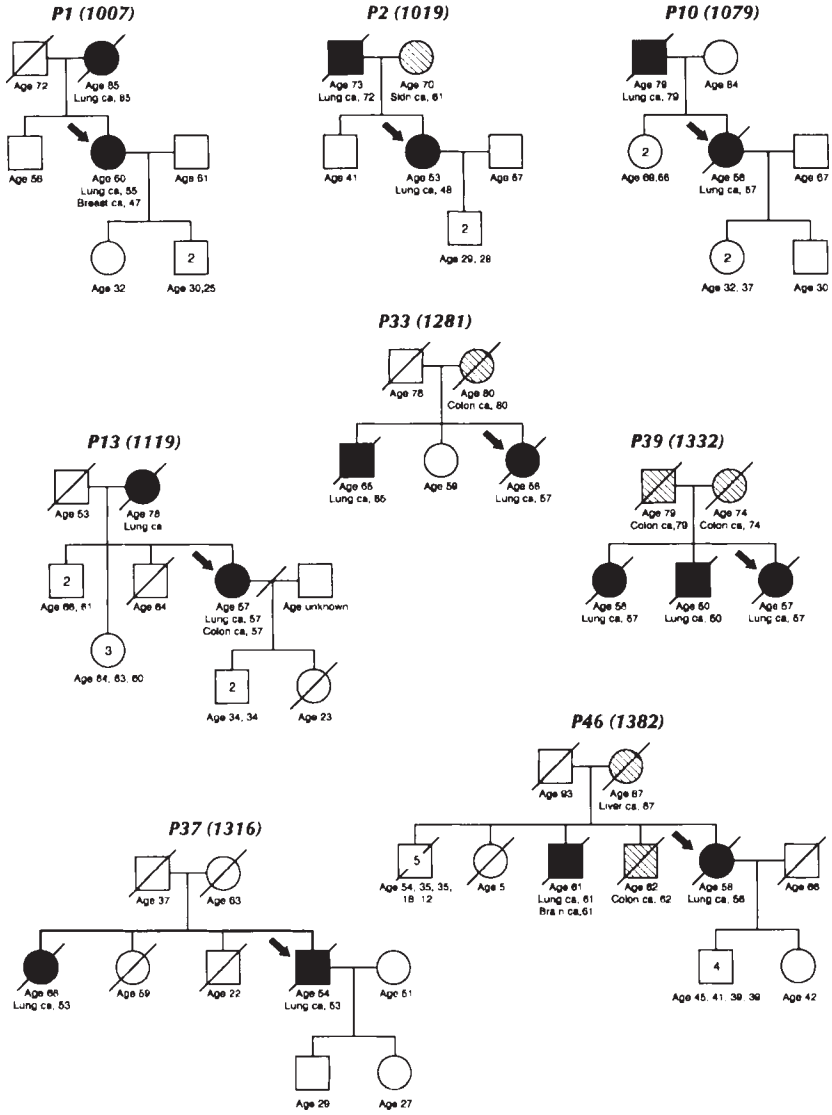

Fig. 2. Multiplex lung cancer pedigrees of nonsmoking probands under the age of 60 years. →, proband; ca., cancer; /, deceased.

bacco smoking. Environmental tobacco exposure and history of other chronic lung diseases also play an important role in lung cancer risk in the families. Since the 210 families in this older group comprised 82% of our original 257 population-based families of nonsmoking probands [Schwartz et al., 1996], it is not surprising that our results are consistent with our earlier report [Yang et al., 1997a], i.e., rejected simple Mendelian inheritance. However, more complicated inheritance, e.g., multiple low-penetrant genes and epistatic effects—a form of gene-gene interaction [Bateson, 1907]—cannot be ruled out. Although we have accounted for the possible cohort difference in tobacco exposure as suggested by Sellers et al. [1992] by stratifying

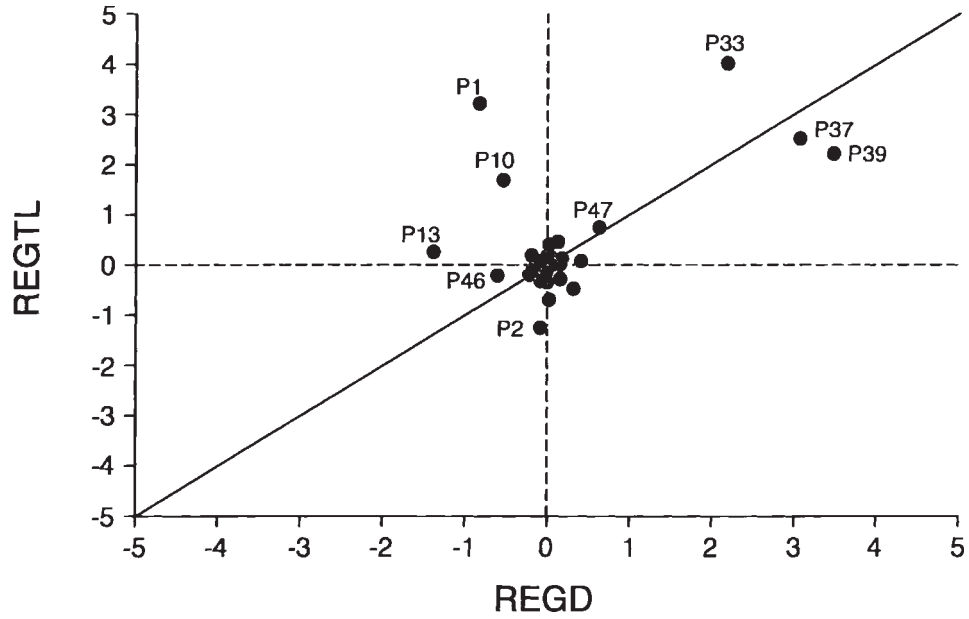

Fig. 3. Pedigree-specific contributions* to a codominant model based on 47 families of nonsmoking lung cancer probands diagnosed under the age of 60 years. *Measured as lnL score for both axes: $-2[\ln(L_{sporadic})-\ln(L_{codominant})]$.

families based on proband's age at diagnosis, there may exist unrecognized risk factors and remaining heterogeneity in these families.

Of particular interest, α_1-antitrypsin deficiency (α_1AD), a common autosomal recessive condition, has been known to cause severe emphysema in adults [Levois and Switzer, 1998] which may partially explain the tendency of familial aggregation of emphysema. Cigarette smoking has a major effect on both the age of onset and the course of pulmonary deterioration in individuals with α_1AD [Cox, 1995]. However, the role of α_1AD disorders, either in homozygous form or in heterozygous form, in relation to lung cancer has not been thoroughly evaluated. Based on our results, family members of nonsmoking probands can be classified into three risk categories. Individuals in the moderate-risk group with emphysema and exposure to environmental tobacco smoke have age-specific lung cancer risks very similar to those in the high-risk group. Carriers of α_1AD may be represented in this risk group [Yang et al., 1997b].

The importance of a major Mendelian genetic factor, cigarette smoking, and history of chronic bronchitis is demonstrated in families of younger aged nonsmoking probands. Our results in several aspects are consistent with the results from another similar study of lung cancer [Sellers et al., 1990, 1992; Gauderman et al., 1997]. The same mode of Mendelian inheritance of a diallelic gene has been detected in our study as in Sellers et al. even though families in their study were ascertained through deceased cases who were mostly smokers [Ooi et al., 1986; Sellers et al., 1987]. The estimated penetrance for high-risk gene carriers (genotype ab) is remarkably similar between the two studies, as shown in Table VII. The discrepancies in penetrance for

TABLE VII. Penetrance Estimations for Genotypes ab and bb From the Current Study and the Study by Sellers et al. [1990]

Genotype	Cigarette smoking		Age in years 50	60	70
Genotype ab	No	Current study[a]	0.003–0.005	0.009–0.018	0.018–0.042
		Sellers et al.	0.001	0.006	0.031
	Yes	Current study	0.013–0.020	0.038–0.072	0.072–0.156
		Sellers et al.[b]	0.011–0.052	0.057–0.164	0.171–0.249
Genotype bb	No	Current study	<0.0001	<0.0001	0.0001
		Sellers et al.	<0.0001	0.0004	0.0024
	Yes	Current study	<0.0001	0.0001–0.0003	0.0003–0.0008
		Sellers et al.	0.0008–0.0044	0.0048–0.0251	0.0277–0.1060

[a]The range between female and male.
[b]The range between ''average'' and ''heavy'' smokers.

non-carriers (genotype bb) could be due to our smaller sample size of the young proband group and, therefore, unstable estimates of the frequencies of smokers. It might also imply that there is a higher proportion of phenocopies in families with smoking probands.

For the three major risk factors from our best-fitting model (genetic predisposition, smoking, and chronic bronchitis), the proportion of all cases due to each factor, i.e., etiologic fraction, is given in Appendix C. Listed in Appendix C are eight strata of lung cancer cases representing various combinations of the three risk factors. Because the homozygous high-risk group (aa) is very rare and less than 0.5% of all lung cancer cases in the study population fall into this high-risk group, this group is omitted. Among female cases (Appendix C.I) in the moderate-risk group (genotype ab), 66% of the patients diagnosed at age 50 in contrast to 43% diagnosed at age 70 could be attributed to the genotype in combination with the other two risk factors. The decline of the genetically determined cases with increasing age is more striking in male patients (Appendix C.II). Reciprocal trends are seen for cases in the low risk group (genotype bb), in both females and males, i.e., the etiologic fraction in tobacco smoking and/or chronic bronchitis related cases increases with age.

As pointed out by Sellers [1996], lung cancer should not be considered solely genetic even though a genetic effect is evident. This is because the gene carriers are very rare in the population and among these carriers, the risk of developing lung cancer (penetrance) is so low for non-smokers (< 5% at age 70). The penetrance increases 4–58 times at age 50 years and 4–8 times at age 70 years for these carriers if they are smokers. Note that within either genotype ab or bb group, smokers with a history of chronic bronchitis account for the majority of the lung cancer cases in the study population. Approximately 60–75% of the cases across all age groups were smokers with chronic bronchitis, 20–30% were nonsmokers with chronic bronchitis, and 1–5% were smokers without chronic bronchitis.

A putative diallelic Mendelian gene consistent with codominant inheritance was suggested in both the current study and the other study [Sellers et al., 1990]. Dominant and recessive inheritance models were rejected. Very likely the high-risk allele is so rare that there is not sufficient power to detect the key difference between the "dominant" and "codominant" mode for the given population size, i.e., the presence

of two distinctive genotypes aa and ab. This observation has been reported in a segregation analysis of primary hepatocellular carcinoma among a native Alaskan population [Yang et al., 1990]. Also possible is that the putative lung cancer susceptibility gene is codominant and the heterozygous individuals are more susceptible than the wild type only when exposed to lung carcinogens.

Most susceptibility genes identified so far for commonly occurring cancers are highly penetrant but rare. Only a small proportion of all cases are attributable to these identified susceptibility genes, that is, they have a small attributable risk to the disease. Common but less penetrant cancer susceptibility genes are difficult to detect because they do not confer substantially increased risks in relatives. This might explain the lack of major genetic effects in our families of older probands. Although by both REGD and REGTL approaches the environmental models fit better than Mendelian models in these families, the homogeneous environmental model (no parent-offspring transmission) was favored by REGD whereas the equal transmission model ($\tau_{aa}=\tau_{ab}=\tau_{bb}$) was favored by the REGTL, as discussed in Yang et al. [1997a].

In this study, familial aggregation of lung and colon cancer is suggested (Fig. 2). A number of other earlier studies have documented a familial component not only to lung cancer risk but also to cancer of other anatomic sites [McDuffie, 1991; Lynch et al., 1986; Wu et al., 1988; Osann, 1991; Shaw et al., 1991; Sellers et al., 1991; Schwartz et al., 1999]. Despite the fact that these reported colon cancers have not all had a pathologically confirmed diagnosis, there might be a common pathway in tumorigenesis between carcinomas of the lung and colon, e.g., DNA repair mechanisms [Weiland et al., 1996; Mizzo et al., 1996; National Cancer Institute Division of Cancer Prevention and Control., 1988].

Recent studies showed that several genetic polymorphisms in the microsomal mixed-function oxidases (cytochrome P450s) and in conjugative drug metabolizing (phase II) enzymes are associated with the risk of lung cancer and other cancers [Nebert et al., 1991; Jacqz et al., 1986; Gaedigk et al., 1991; Miller and Miller, 1983; Lower, 1982]. The P450 enzymes are phase I enzymes, which activate chemical carcinogens producing reactive metabolites capable of inducing or promoting carcinogenesis. Since the balance of phase I and phase II enzyme activity could underlie most carcinogen-related human cancers [Anttila et al., 1995; Nakachi et al., 1991; Seidegård et al., 1986; Romkes-Sparks et al., 1993; Chern et al., 1994, 1995], Yang et al. [1996] investigated whether clusters of various cancers in lung cancer families are associated with combined enzyme genotypes in lung cancer probands. They found that not only was the *mut/mut* genotype of the *NAT2* locus (indicative of a slow phenotype) associated with an increased cancer risk in first-degree relatives of lung cancer patients, but also the patients with combined *GSTμ1 null* and *NAT2 mut/mut* had an over 3-fold increased risk of having cancer-affected first-degree relatives.

A history of certain chronic nonmalignant lung diseases has been reported as a lung cancer risk factor in the literature in the past 10 years. Specifically, COPD among patients with lung cancer has been reported by Cohen et al. [Cohen et al., 1977; Cohen, 1980] and Samet et al. [1986]. These studies [Cohen, 1980; Samet et al., 1986] report a familial tendency to COPD among patients with lung cancer. The available evidence suggests that heritable factors may contribute to the familial aggregation of lung cancer along with other diseases such as COPD. From our current study, an intriguing finding is that emphysema is a potential risk factor for lung cancer only in families of older aged

probands while chronic bronchitis is a potential risk factor for lung cancer only in families of younger aged probands. History of tuberculosis, asthma, and allergy in relation to lung cancer risk has been previously summarized [Schottenfeld, 1996] and discussed [Chern et al., 1995; Yang et al., 1997a].

In conclusion, our results showed no evidence of a major genetic effect in families of older nonsmoking probands whereas a rare autosomal codominant gene is suggested in families of younger probands. In families of older probands, eliminating exposure to cigarette smoking not only can reduce lung cancer risk by over 85% among smokers, but also can reduce the risk by over 60% among passive-smokers. On the other hand, our results from families of young probands demonstrate that the attributable risk (etiologic fraction) from the putative high-risk allele declines with age while the attributable risk from tobacco smoking and chronic bronchitis increases with age. Also, in families with young probands, having a history of chronic bronchitis is a much stronger risk factor than being a cigarette smoker. Cautious interpretation of our results is needed because a relatively small proportion of families are in the young proband group although our probands are all drawn from a well-defined large population [Swanson et al., 1985].

ACKNOWLEDGMENTS

We are grateful to all who contributed to the accomplishment of this work: all the study participants, study manager Mr. Michael Rothrock, study interviewer Ms. Rochelle Steiner, secretary Ms. Sondra Buehler and Ms. Marilyn Goodman, and computer programmer Mr. Blair Powell. Particularly, we acknowledge the essential feedback from Dr. Joan Bailey-Wilson and Dr. Thomas Sellers who served as scientific consultants in the current study. This study was supported by NIH grants R29-CA50383 (A.G.S.) and R01-CA60691 (A.G.S.), N10SH grant R01-OH02067 (G.M.S.), and NCI contract N01-CN05225 (G.M.S.). The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute. The main results reported in this work were obtained by using the SAGE program package, which is supported by US 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 Automatic Control AC 19:716–723.

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.

Anttila S, Luostarinen L, Hirvonen A, Elovaara E, Karjalainen A, Nurminen T, Hayes JD, Vainio H, Ketterer D 1995. Pulmonary expression of glutathione S-transferase M3 in lung cancer patients: Association with GSTM1 polymorphism, smoking, and asbestos exposure. Cancer Res 55:3305–3309.

Bateson W. 1907. The progress of genetics since the rediscovery of Mendel's paper. Progressus Rei Botanicae 1:368.

Beaty TH. 1997. Evolving methods in genetic epidemiology. I. Analysis of genetic and environmental factors in family studies. Epidemiol Rev 19:14–23.

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

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

Chern HD, Lesnick T, Becich MJ, Dickinson A, Persad R, Sibley G, Smith P, Romkes-Sparks M, Branch RA. 1994. Association of low CYP3A4 activity and p53 mutation in aggressive bladder cancer in man. Can J Physiol Pharmacol 72S1:603.

Chern HD, Persad RA, Fleming CM, Bryant JL, Romkes-Sparks M, Sibley GA, Smith PJB, Branch RA. 1995. N-acetyltransferase slow acetylator phenotype increases susceptibility to occupational related aggressive bladder cancer. AACR Conference Proc 36:121.

Cohen BH. 1980. Chronic obstructive pulmonary disease: A challenge in genetic epidemiology. Am J Epidemiol 112:274–288.

Cohen BH, Graves CG, Levy DA, Permutt S, Diamond EL, Kreiss P, Menkes HA, Quaskey S, Tockman MS. 1977. A common familial component in lung cancer and chronic obstructive pulmonary disease. Lancet 2:523–526.

Cox DW. 1995. α_1-antitrypsin deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. The metabolic and molecular bases of inherited disease, 7th ed. New York: The McGraw-Hill Company. p 4125–4158.

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.

Gaedigk A, Blum M, Gaedigk R, Eichelbaum M, Meyer UA. 1991. Deletion of the entire cytochrome P450 CYP2D6 gene as a cause of impaired drug metabolism in poor metabolizers of the debrisoquine/sparteine polymorphisms. Am J Hum Genet 48:943–950.

Gauderman WJ, Morrison JL, Carpenter CL, Thomas DC. 1997. Analysis of gene-smoking interaction in lung cancer. Genet Epidemiol 14:199–214.

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.

Jacqz E, Hall S, Branch RA. 1986. Genetically determined polymorphisms in oxidate drug metabolism. Hepatology 6:1020–1032.

Khoury MJ, Beaty TH, Cohen BH. 1993. Genetic approaches to familial aggregation. I. Analysis of heritability. In: Khoury MJ, Beaty TH, Cohen BH, editors. Fundamentals of genetic epidemiology. New York: Oxford University Press. p 200–232.

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

Levois M, Switzer P. 1998. Differential exposure misclassification in case-control studies of environmental tobacco smoke and lung cancer. J Clin Epidemiol 51:37–54.

Lower GM. 1982. Concepts in causality: Chemically induced human bladder cancer. Cancer 49:1056–1066.

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 Perspect 37:27–35.

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

Miller JA, Miller EC. 1983. Some historical aspects of N-aryl carcinogens and their metabolic activation. Environ Health Perspect 49:3–12.

Mizzo M, Sozzi G, Musso K, Pilotti S, Incarbone M, Pastorino U, Peirotti MA. 1996. Microsatellite alterations in bronchial and sputum specimens of lung cancer patients. Cancer Res 56:2285–2288.

Nakachi K, Imai K, Hayashi S-i, Watanabe J, Kawajiri K. 1991. Genetic susceptibility to squamous cell carcinoma of the lung in relation to cigarette smoking dose. Cancer Res 5119.5177–5180.

National Cancer Institute Division of Cancer Prevention and Control. 1988. Annual Cancer Statistics Review. Bethesda: National Institutes of Health.

Nebert DW, Nelson DR, Coon MJ, Estabrook RW, Feyereisen R, Fujii-Kuriyama Y, Gonzalez FJ, Guengerich FP, Gunsalus IC, Johnson EF, Loper JC, Sato R, Waterman MR, Waxman DJ. 1991. The P450 superfamily: Update on new sequence, gene mapping, and recommended nomenclature. DNA Cell Biol 10:1–14.

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, editors. 1994. SEER Cancer Statistics Review, 1973-1991: Tables and Graphs, National Cancer Institute. NIH Pub. No. 94-2789, Bethesda, MD: National Institutes of Health.

Romkes-Sparks M, Mnuskin A, Chern H-D, Persad R, Branch RA. 1993. Cytochrome P4502D6 mRNA expression in bladder tissue is related to debrisoquine hydroxylation activity in normal subjects and patients with bladder cancer. AACR Conference Proc 34:143.

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.

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, editor. Case-control studies: design, conduct, analysis. New York: Oxford University Press. p 227–290.

Schottenfeld D. 1996. Epidemiology of lung cancer. In: Pass HI, Mitchell JB, Johnson DH, Turrisi AT, editors. Lung cancer principles and practice. Philadelphia: Lippincott-Raven Publishers. p 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.

Schwartz AG, Rothrock M, Yang P, Swanson GM. 1999. Increased cancer risk among relatives of nonsmoking lung cancer cases. Genet Epidemiol (in press).

Seidegård J, Pero RW, Miller DG, Beattie EJ. 1986. A glutathione transferase in human leukocytes as a marker for the susceptibility to lung cancer. Carcinogenesis 7:751–753.

Sellers TA. 1996. Familial predisposition to lung cancer. In: Eeles RA, Ponder BAJ, Easton DF, Horwich A, editors. Genetic predisposition to cancer. London: Chapman & Hall. p 344–353.

Sellers TA, Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H. 1987. Increased familial risk for non-lung cancer among relatives of lung cancer patients. Am J Epidemiol 126:237–246.

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, Potter JD, Folsom AR. 1991. Association of incident lung cancer with family history of female reproductive cancers: The Iowa Women's Health Study. Genet Epidemiol 8:199–208.

Sellers TA, Potter JD, Bailey-Wilson JE, Rich SS, Rothschild H, Elston RC. 1992. Lung cancer detection and prevention: Evidence for an interaction between smoking and genetic predisposition. Cancer Res 52(Suppl):2694s–2697s.

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.

Shopland DR. 1996. Cigarette smoking as a cause of cancer. In Cancer Rates and Risks. National Cancer Institute, NIH Publication No. 967-691. 4th edition. Bethesda: National Institutes of Health. p 67–72.

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

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

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

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

Weiland I, Ammermuller T, Bohm M, Totzeck B, Rajewsky MF. 1996. Microsatellite instability and loss of heterozygosity at the hMLH1 locus on chromosome 3p21 occur in a subset of nonsmall cell lung carcinomas. Oncol Res 8:1–5.

Williams WR, Anderson DE. 1984. Genetic epidemiology of breast cancer; segregation analysis of 200 Danish pedigrees. Genet Epidemiol 1:7–20.

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 48:7279–7284.

Yang P, Buetow KH, Lustbader E, Lanier AP, McMahon BJ, McGlynn KA. 1990. Evidence for a major locus modifying risk of hepatocellular carcinoma. Am J Hum Genet 47(Suppl):A25.

Yang P, Romkes M, Schwartz AG, Rittmeyer L, Adedoyin A, Landreneau R, Mauro K, Branch RA. 1996. Polymorphisms at the NAT2* locus associated with increased cancer risk in first-degree relatives of lung cancer patients. Genet Epidemiol 3:308–309.

Yang P, Schwartz AG, McAllister AE, Aston CE, Swanson GM. 1997a. Genetic analysis of families with nonsmoking lung cancer probands. Genet Epidemiol 14:181–197.

Yang P, Wentzlaff KA, Marks R, Patel A, Allen M, Edell E, Katzmann J, Lesnick T, Deschamps C, Midthun D, Bonner J, Swensen S, Tazelaar H, Krowka MJ. 1997b. Higher rate of alpha-1-antitrypsin (α_1AT) deficiency carriers found in lung cancer patients. Am J Hum Genet 61:.A15.

APPENDIX A. Estimates Lung Cancer Risk (%) Based on an Environmental Model of 210 Families With Nonsmoking Lung Cancer Probands Diagnosed ≥ 60 Years of Age in Metropolitan Detroit, 1984–1987

	Estimated lung cancer risk (%)			
	At age 60 years		At age 70 years	
Risk group (% in study population)	Male	Female	Male	Female
High risk (0.1%)	>80	>70	>90	>80
Moderate risk (5.8%)	0.2	0.1	0.4	0.2
+ emphysema	10.6	5.7	22.1	10.7
+ ETS	3.8	2.0	8.6	3.8
+ smoking	29.3	17.4	47.8	29.5
+ emphysema and ETS	76.6	59.9	87.6	74.8
+ emphysema and smoking	96.9	94.1	98.7	96.9
Low risk (94.1%)	0	0	0	0
+ smoking, emphysema, and allergy	0.01	<0.01	0.02	0.01

ETS, environmental tobacco smoke.

APPENDIX B. Estimated Penetrance Under a Mendelian Codominant Model for 47 Families With Nonsmoking Probands Diagnosed < 60 Years of Age in Metropolitan Detroit, 1984–1987

	Age in years		
Risk group (proportion in the population)	50	60	70
I. Females			
Genotype aa (0.000016)	.4802	.7375	.8477
+ smoking	.7959	.9222	.9592
+ chronic bronchitis	.9815	.9938	.9969
+ both	.9958	.9986	.9993
Genotype ab (.01)	.0031	.0092	.0181
+ smoking	.0128	.0378	.0722
+ chronic bronchitis	.1494	.3482	.5142
+ both	.4258	.6927	.8171
Genotype bb (.99)	0	0	.0001
+ smoking	0	.0001	.0003
+ chronic bronchitis	.0006	.0018	.0035
+ both	.0025	.0075	.0147
II. Males			
Genotype aa (0.000016)	.5966	.8465	.9297
+ smoking	.8619	.9588	.9824
+ chronic bronchitis	.9884	.9969	.9987
+ both	.9974	.9993	.9997

(*continued*)

APPENDIX B. Estimated Penetrance Under a Mendelian Codominant Model for 47 Families With Nonsmoking Probands Diagnosed < 60 Years of Age in Metropolitan Detroit, 1984–1987 (*continued*)

	Age in years		
Risk group (proportion in the population)	50	60	70
Genotype ab (.01)	.0049	.0179	.0419
+ smoking	.0203	.0716	.1560
+ chronic bronchitis	.2195	.5119	.7153
+ both	.5427	.8157	.9138
Genotype bb (.99)	0	.0001	.0001
+ smoking	.0001	.0003	.0006
+ chronic bronchitis	.0009	.0035	.0083
+ both	.0040	.0146	.0343

APPENDIX C. Etiologic Fraction (%) for Lung Cancer Cases in the Study Population Based on a Mendelian Codominant Model[a]

	Age in years		
Risk group	50	60	70
I. Females			
Genotype ab	0.34	0.45	0.55
+ smoking	1.42	1.86	2.20
+ chronic bronchitis	16.58	17.17	15.67
+ both	47.25	34.16	24.90
Subtotal	**65.59**	**53.64**	**43.32**
Genotype bb	0	0	0.30
+ smoking	0	0.49	0.91
+ chronic bronchitis	6.66	8.88	10.67
+ both	27.74	36.98	44.80
Subtotal	**34.40**	**46.35**	**55.77**
II. Males			
Genotype ab	0.38	0.55	0.68
+ smoking	1.58	2.19	2.53
+ chronic bronchitis	17.05	15.67	11.62
+ both	42.15	24.97	14.85
Subtotal	**61.16**	**43.38**	**29.67**
Genotype bb	0	0.31	0.16
+ smoking	0.78	0.92	0.97
+ chronic bronchitis	6.99	10.71	13.48
+ both	31.07	44.69	55.71
Subtotal	**38.84**	**56.63**	**70.32**

[a]Total percentage adds to 100 for females and males at each age, respectively.