Review

Host Factors in Lung Cancer Risk: A Review of Interdisciplinary Studies

Christopher I. Amos,¹ Neil E. Caporaso, and Ainsley Weston

Division of Biostatistics Research and Informatics, International Agency for Research on Cancer, Lyon, France [C. I. A.], and Genetic Studies Section, National Institute of Arthritis, Musculoskeletal and Skin Diseases [C. I. A.], Environmental Epidemiology Branch [N. E. C.], and Laboratory of Human Carcinogenesis [A. W.], National Cancer Institute, Bethesda, Maryland 20892

Inhalation (CYP1A1) activate many carcinogens in tobacco smoke but have shown inconsistent associations with risk for lung cancer.

Case-control studies that assess the effects of $\text{CYPIID6}$ on lung cancer risk have consistently shown a mildly decreased risk for lung cancer among poor metabolizers. Cell surface markers have shown little relation to risk for lung cancer. Studies involving DNA or hemoglobin adducts, sister chromatid exchange, or oncogene activation only indirectly measure host-specific risk, and these assays have suffered from poor reproducibility and high cost. We describe epidemiological designs to assess specific genetic factors that may alter lung cancer risk.

Introduction

Lung carcinogenesis in humans usually requires exposure to environmental agents including inhalation of tobacco smoke (1), radioactive ores, metals, asbestos, and petrochemicals (2). Studies of host-environment interactions can identify groups of individuals who are at the greatest risk and may help in identifying oncogenic mechanisms. Knowledge of mechanisms that lead to lung cancer may help to identify early stages of carcinogenesis and allow the possibility of better-targeted intervention strategies.

Family studies and case-control designs have been used to assess host factors in lung cancer risk. Family studies collect the same information on all individuals within a specified sampling scheme, according to which families of various sizes and structures may be selected. Case-control studies, on the other hand, are typically designed to collect extensive information from a large set of cases and controls and generally collect only limited information on relatives of the cases and controls. A few case-control studies followed systematic rules for collecting family information and therefore represent both case-control and family studies.

The purpose of this review is to consolidate findings from these two usually distinct approaches (family and epidemiological studies) to facilitate the future design of studies. The design of future studies will depend upon attributes of the studied biomarker as well as characteristics of the assay for the biomarker. We first discuss studies to assess the familiality of lung cancer and familial associations of lung cancer with other diseases; then we discuss associations of specific biomarkers with lung cancer risk.

Familial Clustering of Lung Cancer

Familiality of Lung Cancer. Familiality of lung cancer has been examined in case-control studies, genealogy-based searches for familiality, through segregation analysis, and through twin studies. In establishing familiality, measures of smoking and environmental exposures must be included because these exposures tend to be correlated among family members (3, 4). Failure to allow for correlated environmental exposures may falsely lead to evidence for familiality. For this reason, results from genealogy-based studies of familiality (5, 6) for lung cancer will not be considered here. Studies of familiality have not assessed the effects of passive smoking, although cases were found to have a 1.4-fold increased risk for maternal exposure after allowance for personal smoking measures (7).

Familiality of lung cancer has been examined in case-control studies by comparing the recurrence risk (i.e., the probability that a relative is affected by the same cancer) among first-degree relatives in case families to that observed in control families. Tokuhata and Lilienfeld (3) demonstrated excess lung cancer mortality among the relatives of 270 lung cancer patients compared with 270 age-, race-, sex-, and location-matched controls and their relatives. Smoking was a more important determinant of risk for men, while family history of lung cancer was the more important risk factor in women. Smoking and a family history had synergistic effects. Smoking relatives of cases had a relative risk of 2–2.5 for mortality from lung cancer compared with smoking relatives of controls. Nonsmoking relatives of lung cancer cases were also at higher risk for lung cancer than nonsmoking relatives of controls; and, compared with controls, a significant excess of noncancer respiratory illnesses was observed among case relatives, but this excess was not correlated with smoking status in the relatives.
More recent studies (8, 9) have used retrospective case-control designs to assess effects due to smoking and familial risk for lung cancer. Ooi et al. (8) studied first-degree relatives of 336 deceased lung cancer cases and first-degree relatives of 307 of their spouses in Louisiana. Quantitative measures of smoking behavior and occupational exposures were obtained. After making allowance for smoking and occupational exposures, the case relatives had a relative risk of 2.4 for lung cancer when compared with the control relatives. Familial effects were more strongly observed among the female relatives than among the male relatives. In a similar but smaller study, also conducted in southern Louisiana, Sellers et al. (9) reported a relative risk of 2.5 for lung cancer among siblings of lung cancer cases compared to siblings of the spouse controls. Shaw et al. (10) studied families from south Texas and observed a younger age at onset among probands with relatives who had lung cancer.

Segregation analyses (11) of data collected by Ooi et al. (8) were consistent with Mendelian segregation of a single codominant locus when a model was fitted that allowed for a genetic component affecting the age of onset of lung cancer. Among lung cancer cases younger than age 50, over 70% were predicted to have some genetic predisposition, while over 70% of cases 70 or older were predicted to be attributable solely to smoking. This study did not model specific effects attributable to passive smoking, and effects from occupational exposures were not modeled. Results from this study reflect a rather restrictive set of genetic hypotheses that could be considered, e.g., oligogenic (processes affected by several genes) influences were not specifically considered. The number of lung cancer events observed in twin studies has been too small to draw conclusions regarding the familiality of lung cancer (12). However, Fisher's hypothesis that a common genetic cause underlies the risk for both smoking and lung cancer was disproved by evaluation of monozygotic twins discordant for smoking behavior (13, 14).

Diseases with Altered Risk for Lung Cancer. Families with excess lung cancer of diverse histological types have been reported (15), with alveolar cell carcinoma showing consistent familial aggregation (16–17). Lung cancer may be more common in pedigrees having an excess of breast and ovarian cancers (18), and survivors of familial retinoblastoma may be at increased risk for small-cell lung cancer (19).

COPD is associated with an increased risk for lung cancers (20–22). A causal relationship of either disease with the etiology of the other is unlikely because of their similar age-at-onset distributions. However, oxidizing agents in cigarette smoke may initiate COPD through inflammatory processes and may initiate lung cancer by genotoxic effects mediated by P-450 enzymes (23). Both cross-sectional (24) and longitudinal observations (25) suggest that delayed clearance of inhaled carcinogens among COPD-affected individuals increases the risk for progression to lung cancer. In further support of this hypothesis, increased risk for lung cancer is observed in sarcoidosis and scleroderma (26, 27), which have decreased clearance of inhaled carcinogens. Several studies (3, 24) have noted familial clustering of COPD and other respiratory illnesses and lung cancer. Cohen (24) showed that similar smoking and socioeconomic factors among family members were not sufficient to account for the correlation of COPD and lung cancer in families, suggesting an underlying familial component of risk for both diseases.

Mortality from lung cancer has been reported to be lower among schizophrenic patients (28, 29) in spite of elevated levels of smoking (30). Neuroleptics such as phenothiazines have been reported to inhibit tumor growth in animals (31). Inhibition of P-450 enzyme metabolism might account for the apparent deficit of lung cancer among schizophrenic patients (32).

Genetically Influenced Traits Associated with Variability in Risk for Lung Cancer

Chemical Carcinogenesis. Activation of PAHs and arylamines by cytochrome P-450s (e.g., CYP1A1, CYP1A2, and CYP3A4) leads to the formation of reactive chemical species that can bind covalently to DNA to form carcinogen-DNA adducts (33). Fig. 1 depicts the metabolism of PAHs by cytochrome P-450 enzymes. PAHs such as BP are first oxidized by AHH or CYP1A1 (the structural gene for P-4501A1), resulting in the formation of an arene oxide. This metabolite can be further activated to form a dihydrodiol by the action of an epoxide hydratase (34). Dihydrodiol can then be further metabolized across the olefinic double bond by both cytochrome P-450 (CYP3A4) and other oxidation systems (35), thus forming a diol-epoxide. Diol-epoxides are unstable and rearrange to form highly reactive carbocations. For example, benzo[a]pyrene-7,8-diol,9,10-epoxide forms covalent bonds primarily with the exocyclic amino group of guanine. This reaction has been shown to be responsible for activating mutations in the HRAS-1 protooncogene in several experimental systems (36, 37).

Alternative and competing routes of metabolism can lead to inactivation of carcinogens (e.g., PAHs) through the formation of conjugates (glutathiones, glucuronides, and sulfate esters) or phenols and tetrahydrotetroxins which facilitate excretion. The balance between metabolic activation and metabolic detoxification, as well as the efficiency of DNA repair mechanisms, affects cancer risk in an individual exposed to PAHs.

Carcinogenic aromatic amines, typified by 4-aminobiphenyl, are related to bladder cancer; and the acetylation phenotype, mediated by the polymorphic enzyme N-acetyltransferase, has been related to susceptibility for this tumor. Aromatic amines can be found in tobacco smoke (38, 39), and aromatic amine-DNA adducts have been detected in the lung and bladder (40–42). No association has been noted, however, between the acetylator phenotype and lung cancer (43–44). Analysis of 4-aminobiphenyl hemoglobin adducts from lung cancer cases and controls showed a correlation with tobacco consumption but not with diagnosis of lung cancer (45), arguing against any role for these compounds in lung carcinogenesis.

1 The abbreviations used are: COPD, chronic obstructive pulmonary disease; PAH, poly cyclic aromatic hydrocarbon; BP, benzo[a]pyrene; HRAS-1, Harvey ras protooncogene; AHH, aromatic hydrocarbon hydroxylase; Ah, aromatic hydrocarbon; RFLP, restriction fragment length polymorphism; ECDE, 7-ethoxycoumarin O-deethylase; CYP, structural genes for cytochrome P-450 enzymes; SCE, sister chromatid exchange.
Cytochrome P-450 Enzymes. A variety of P-450 enzymes have been proposed as having a role in lung cancer through the demonstrated role of these enzymes in the activation of carcinogens in tobacco. CYP1A1 catalyzes the oxygenation of PAHs such as BP more efficiently than CYP1A2, while substrates for CYP1A2 enzyme additionally include arylamines and heterocyclic amines (food mutagens) (46). The CYP1A genes are located on chromosome 15q22-qter (47). Levels of both CYP1A genes are inducible and are influenced by the binding of the Ah receptor. Induction follows exposure to aromatic hydrocarbons, but the inducibility varies among individuals. Several other gene products are also inducible by the binding of the Ah receptor to a substrate, including UDP glucuronosyltransferase (48), NAD(P)H:menadione-oxidoreductase (49), and glutathione transferase (50).

Regulation of levels of CYP1A1 occurs via a number of different mechanisms (51). Following the binding of PAHs with the Ah receptor, the complex is transported to the nucleus, where it binds to nuclear chromatin in a temperature-dependent step and induces transcription. An upstream regulatory sequence, an MSP1 site (52, 53) downstream of the coding region, and a product from chromosome 2 (54, 55) may all influence expression in humans. Interaction of the Ah receptor:compound complex with regulatory regions of the CYP1A2 gene does not appear to be as important as it is for regulation of CYP1A1 levels (46), but a large variability in the constitutive mRNA expression of CYP1A2 has been observed in human liver samples (46).

Studies of in vivo levels of CYP1A enzymes as a risk factor for lung and other cancers are confounded by the inducibility of these enzymes. Because individuals often change their smoking around the time of lung cancer diagnosis, case-control studies designed to study the risk for lung cancer associated with the expression of CYP1A enzymes must make some allowance for the state of induction of the study subjects. Petruzelli et al. (56) compared AHH, ECDE, epoxide hydroxylase, glutathione-S-transferase, UDP-glucuronosyltransferase, glutathione, and malondialdehyde contents from nontumorous parenchymal samples from 54 lung cancer patients and 20 patients undergoing resection for nonneoplastic diseases. All enzymatic levels (AHH, ECDE, UDP-glucuronosyltransferase, glutathione S-transferase) were significantly correlated. When smoking was ignored, there were no differences between cancer and noncancer patients. However, lung cancer patients who had smoked within 30 days of surgery had significantly higher AHH levels and ECDE activities than other current smokers. Negative correlations (significant for all but ECDE) were observed between all of the enzyme levels and the number of days that had elapsed since the patients had given up smoking. No significant differences were observed between nonsmoking lung cancer cases and nonsmoking controls. These two observations suggest that the presence of lung cancer itself does not lead to an induction of these enzymes. Quantitative studies of Northern blots of mRNA for CYP1A1 expression were performed on lung tissue resected during surgery for primary lung cancer (57). CYP1A1 expression was absent in nonsmokers (0 of 20) but was detected in 3 of 7 smokers who had abstained less than 1 month and was present in 17 of 19 current smokers.

High levels of AHH inducibility have been shown to be a risk factor for lung cancer in several case-control studies (58); earlier results (59) which indicated that AHH inducibility was predictive of case status, however, could not be duplicated in many studies (60–62). Paigen et al. (60) could not replicate earlier work despite implementing a careful design including cases as well as their healthy offspring and spouses.

Measurement error may have obscured results from some of the earlier studies. Kouri et al. (63) suggested that the ratio of AHH inducibility to cytochrome c activity was a more reliable measure than AHH inducibility alone. Trelle et al. (64), studying 20 healthy subjects, found that the intrapersonal variability in this ratio following exposure to 3-methylcholanthrene was 0.016, while the interindividual variability that can be estimated from their data is 1.63. Among 1589 smokers, 304 ex-smokers and 218 never-smoking individuals, trimodal distributions were observed for each category of smoking status similar to those reported by Kellerman et al. (59), consistent with the observed codominant expression (65).

Isoenzymes of glutathione S-transferase have been identified (66), and the presence of the mu isoenzyme is dominantly inherited. Results from two case-control series conducted by Siedegard (67, 68) and others in Sweden and New York City show that consistently fewer lung cancer patients have the mu isoenzyme than smoking-matched controls. The mu isoenzyme was least frequent among small cell, large cell, and adenocarcinoma patients, but the first two categories included few patients. The odds ratio comparing the presence of the mu isoenzyme in controls to cases is 2.4 for both studies.
combined. Neither of these studies assessed ethnic variation, and in both studies cases and controls were not obtained from the same hospital populations. A subsequent study (69) conducted on a homogeneous population found low levels of glutathione S-transferase to be nonsignificantly protective of lung and other smoking-related cancers, but the power to detect a significant difference at least as large as that found by Siedegard (68) was only 70%.

**CYP2D6 (Debrisoquine) Phenotype.** The ability to metabolize the antihypertensive drug debrisoquine appears to be a risk factor for lung cancer in smokers. Metabolism of debrisoquine along with other drugs including antidepressants, neuroleptics, beta blockers, and codeine depends upon the activity of the CYP2D6 enzyme. A possible mechanism is suggested by the report (70) that CYP2D6 activates the carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butane, a nicotine-derived nitroso ketone from tobacco. Unlike CYP1A enzymes, levels of CYP2D6 are not inducible, but constitutive expression varies widely among individuals. Sample storage (71), circadian (72), and intranidividual variation account for minor degrees of variability in phenotyping, while certain medications, notably quinidine, predictably invalidate phenotyping.

From family and population studies (71, 73–75), the poor metabolizer phenotype can be well differentiated from extensive or intermediate metabolizer phenotypes. Intermediate metabolizers show incomplete dominance, with mean values being that of the extensive metabolizers plus 30% of the difference between that of the extensive and poor metabolizers. Among Caucasians of European descent, the pooled estimate of the poor metabolizer genotype frequency is 6.4% based upon phenotyping of 3280 unrelated, healthy individuals (71), while this frequency is estimated at 2.1% among West Africans. The gene coding for CYP2D6 has been cloned (76), and polymerase chain reaction-based assays now distinguish most poor metabolizers from extensive or intermediate metabolizers (77). A summary (78) of mutant CYP2D6 genes describes three significant mutations. The CYP2D6(B) mutation accounts for about 75% of poor metabolizer alleles and results from a splice site defect in exon 4. CYP2D6(A) accounts for a little more than 5% of defective alleles and results from a deletion in exon 5. The CYP2D6(D) mutation represents a total deletion of the gene and accounts for 10–15% of poor metabolizer alleles. About 10% of poor metabolizers currently remain unrecognized by molecular techniques.

Metabolic status for debrisoquine has been found to be a predictor of risk for lung cancer (Table 1). In a case-control study of 245 cigarette smokers with bronchogenic carcinoma, poor metabolizers were underrepresented (79, 80). Results from subsequent studies (43, 81–84) support these earlier results, although findings from several studies (43, 83, 84) were equivocal. Benitez et al. (83) observed a significant effect only when adenocarcinomas are excluded, a finding which reflects the earlier observation that adenocarcinomas did not share elevated risk based on the debrisoquine metabolic ratio (80). Speirs et al. (85) did not find evidence for an effect of debrisoquine metabolic ratio on risk for lung cancer, but the cases-only approach used precludes compilation or comparison with other studies. Using a Mantel-Haenszel statistic, the odds ratio comparing the relative odds of being a poor metabolizer versus being an extensive or intermediate metabolizer among lung cancer cases versus controls is 2.3 (95% confidence interval = 1.6–3.4; \( P < 0.001 \)). Mild variability among the strata is observed \( (P = 0.01) \), indicating some dissimilarity among the results from the various studies.

**Blood Groups and Serum Markers.** Type O blood group has been associated with decreased risk for lung cancer, but cases and controls in this study were not selected by the same mechanism (43). Although case-control studies to assess the effects of HLA antigens on lung cancer risk have been performed, consistent effects have not been documented, although to date few studies (86) have included Dr or other class II loci. Serum alpha antitrypsin genotypes are not related to lung cancer risk (87) or to abnormal sputum cytologies. However, alpha antitrypsin levels (88) and trypsin inhibitory capacity (87) were significantly higher in lung cancer patients than in controls with normal or abnormal sputum cytologies.

**Oncogene Polymorphisms.** A variable tandem repeat DNA sequence region tightly flanked by MspI restriction sites, which is located 3' with respect to the cellular HRAS-1 structural protooncogene (chromosomal region 11p15.5), accounts for a DNA RFLP. Rare alleles at the HRAS-1 protooncogene locus were found more commonly in lung cancer (89) cases than in controls. No association was found between either risk of lung cancer

---

**Table 1** Summary of findings from case-control studies of lung cancer and debrisoquine phenotypes

<table>
<thead>
<tr>
<th>Investigator</th>
<th>Frequency of control</th>
<th>Frequency of case</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>Criterion to be a PM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EMs*</td>
<td>PMs</td>
<td>EMs</td>
<td>PMs</td>
<td></td>
</tr>
<tr>
<td>Law et al.</td>
<td>95</td>
<td>9</td>
<td>102</td>
<td>2</td>
<td>4.83</td>
</tr>
<tr>
<td>Ayesh et al.</td>
<td>213</td>
<td>21</td>
<td>241</td>
<td>4</td>
<td>5.94</td>
</tr>
<tr>
<td>Caporaso et al. (blacks)</td>
<td>41</td>
<td>1</td>
<td>48</td>
<td>1</td>
<td>1.17</td>
</tr>
<tr>
<td>Caporaso et al. (whites)</td>
<td>39</td>
<td>10</td>
<td>47</td>
<td>0</td>
<td>NE</td>
</tr>
<tr>
<td>Roots et al.</td>
<td>240</td>
<td>30</td>
<td>251</td>
<td>19</td>
<td>1.65</td>
</tr>
<tr>
<td>Duche et al.</td>
<td>234</td>
<td>20</td>
<td>143</td>
<td>10</td>
<td>1.22</td>
</tr>
<tr>
<td>Benitez et al.</td>
<td>133</td>
<td>10</td>
<td>80</td>
<td>4*</td>
<td>1.50</td>
</tr>
<tr>
<td>All studies</td>
<td>995</td>
<td>101</td>
<td>912</td>
<td>40</td>
<td>2.28(^d)</td>
</tr>
</tbody>
</table>

\( ^a \) EMs, extensive metabolizers; PM, poor metabolizers; MR, metabolic ratio; NE, not estimable from the data (infinite).

\( ^b \) Confidence intervals calculated using Cornfield's method.

\( ^c \) Includes three adenocarcinomas.

\( ^d \) Odds ratio calculated with a Mantel-Haenszel statistic, stratifying on study.

---

508 Review: Lung Cancer Genetomics

---

* Downloaded from cebp.aacrjournals.org on October 28, 2017. © 1992 American Association for Cancer Research.
or risk of metastasis and a RFLP defined by an EcoRI site in the second intron of the L-myc protooncogene (90). Both the HRAS-1 and L-myc RFLPs were found to vary significantly with race (89, 90).

**Biomarkers Associated with Carcinogenesis and Progression**

Biomarkers associated with carcinogenesis and tumor progression include oncogene expression, DNA and red blood cell adducts, sister chromatid exchange, and in tumor tissue, presence of microsatellites, alteration in HLA expression, and expression of abnormal proteins such as α-fetoprotein. Generally, the utility of many of these markers in assessing individual risk for tumorigenesis is hampered by the high measurement error of the assays, the expense, technical parameters (i.e., collection, storage, and processing of biological materials under very stringent conditions), and the lack of specificity of the assays. Certain markers, specifically adducts, have the potential to assess the internal dose of specific carcinogens, while others (i.e., SCE) have limited utility in assessing the host-specific risk for lung cancer.

**Adduct Formation and Sister Chromatid Exchange.** Theoretically, the study of DNA adduct formation allows a measure of the internal dose of carcinogen and therefore provides an advantage over traditional exposure instruments such as questionnaires. At least one study has demonstrated a relationship between adduct formation and both a carcinogenic exposure and a genetic susceptibility factor (91). DNA adduct levels in monocytes exposed to BP from 26 lung cancer patients younger than age 46 were more than twice ($P < 0.0025$) that of matched healthy controls (92). Significantly higher adduct levels were observed in monocytes from all smoking lung cancer patients versus all smoking controls and from all nonsmoking cases versus all healthy nonsmokers. Studying 86 first-degree relatives in 15 families, Nowak et al. (93) observed greater inter familial than intrafamilial variability in DNA adduct formation among monocytes exposed to BP, suggesting that host-specific factors affect either adduct formation or the rate of DNA repair.

SCE following exposure of lymphocytes to BP occurred significantly more frequently in lymphocytes from the healthy offspring of lung cancer cases who had four or more relatives with cancer than in lymphocytes from healthy offspring of lung cancer cases who did not have a family history of other cancers (94). SCE was found to occur more frequently in cigarette smokers than among nonsmokers, however (95, 96). Lung cancer patients who were ex-smokers had the same rates of SCE as nonsmokers, but lung cancer patients who smoked had more SCEs than smokers (95, 96).

**Cytogenetic and Molecular Genetic Studies.** Cytogenetic studies are useful in identifying chromosomal regions that are likely to be involved in tumorigenesis. Regions that are typically deleted may contain antioncogene or tumor suppressor loci, while areas that are replicated contain growth factors. Deletion of the short arm of chromosome 3 has been observed in most karyotypes contain growth factors. Deletion of the short arm of tumor suppressor boci, while areas that are typically deleted may contain antioncogene or that are likely to be involved in tumorgenesis. Regions of 17p13 also showed a loss of heterozygosity in 16 of 16 tumor samples heterozygous for RFLPs in this region (102). Loss of heterozygosity for markers on chromosome 17p was found in 8 of 9 squamous cell cancers but in only 2 of 11 adenocarcinomas (105). Similarly, 9 of 21 non-small cell lung cancer tissue samples had loss of heterozygosity for loci in the region 13q12-33 (105). Localization of von Hippel Lindau (106) and renal cell carcinoma susceptibility loci to 3p26 (107, 108), retinoblastoma to 13q (109), and the p53 gene to chromosome 17p (110) provides candidate regions for further evaluation of any families that may appear to segregate lung cancer.

**Prospects for Future Studies**

**Designs for Studies.** Case-control studies may be a useful tool for identifying the environmental and some genetic risk factors important in lung cancer etiology but provide poor estimators of the specific genetic and environmental contributions of the total risk. Despite difficulties in implementing case-control studies which include the study of relatives, the power of this approach to resolve epidemiological versus genetic contributions to disease etiology is high, and this design is often worth the effort. Implementation of designs which include case and control families is hampered by the high mortality from lung cancer, so that affected relatives are likely to be deceased. However, if case definition is restricted to those who developed cancer at an early age, the likelihood is greater that more relatives (and more affected relatives) will be alive and available for study. Generalization of results from this design requires the assumption that individuals affected by early-onset cancer are representative of the general population affected by lung cancer.

Cohort studies, especially if focused on high-risk subjects, can evaluate multiple markers, including P-450 phenotypes, and may also avoid sources of bias from changing behavior patterns. In addition, these studies can more readily assign the proper causal structure underlying disease etiology. Unfortunately, large enough cohorts with sufficient questionnaire data and biological samples are not readily available.

Linkage studies are an extremely powerful tool for identifying the genetic loci involved in disease etiology, although problems encountered in conducting case-control studies include relatives also impede the implementation of linkage studies. A linkage study is more easily accomplished if multiple living affected persons are available in a family. Often, other affected relatives are already deceased once the lung cancer proband has been identified, so that appropriate families are difficult to obtain. In addition, the inheritance pattern for lung cancer does not fit a simple Medelian pattern, further increasing the sample size necessary to identify any specific locus predisposing for this disease.

New methods which address both of these problems are being developed. Genetic linkage studies can be accomplished despite deceased family members when the genotypes of these unavailable persons can be deduced from that of their living relatives. In addition, tissue blocks saved from surgical procedures provide sufficient genetic material for analysis, although nontumorous tissue is necessary. Advances have also been made in modeling inheritance patterns for non-Mendelian phenotypes (111) and in constructing robust linkage tests.
The latter assess evidence for linkage by evaluating the association between affection status and allele sharing for sib and other relative pairs and detect genetic linkage without requiring exact knowledge of the disease process.

**Methodological Considerations for Some Specific Biomarkers.** The lack of reliable and inexpensive measures of metabolic capacity for exogenous carcinogens has hampered the implementation of family studies. However, continued progress may soon reduce the cost and increase the efficacy of these studies. Caffeine metabolites act as an excellent surrogate for 4-aminobiphenyl metabolic capacity and could be used to assay reliably for CYP1A2 activity (114, 115). As yet, however, the relationship between hepatic and liver CYP1A1 enzymatic activity and differential effects of CYP1A1 and CYP1A2 enzymes on caffeine metabolism are unavailable.

Assays for DNA adduct formation are theoretically useful nonspecific measures of susceptibility for the initiation of lung cancer, but some obstacles must be overcome before these assays can be used in family studies. First, the relationship between exposure and subsequent adduct formation in lung tissues and blood or urine must be quantified. Second, measures of adduct formation remain prohibitively expensive for large studies, and reproducibility of any inexpensive assays must be explored prior to implementation. Measures of SCE formation might similarly be studied, but these provide a very nonspecific measure of susceptibility to lung cancer.

Methods for assessing exposures to cigarette smoke and other carcinogens related to lung cancer need further refinement. Passive smoking was shown to elevate lung cancer risk slightly (7, 116-118). In a large study, however, the elevation of risk from passive smoke was conferred primarily by spousal and occupational exposures and did not stem from childhood contacts, although these may be subject to recall errors (118). Passive smoking is thus not likely to provide an adequate explanation for increased familial risk since the familial environmental exposures occur primarily during childhood, although an instrument to measure this exposure must be included in further family studies. Metabolites of nicotine (recovered in blood or urine) reflect recent exposure to tobacco products and can be used to assess the validity of questionnaire data concerning recent exposures (119). However, cotinine measures are affected by individual variation in nicotine metabolism, dietary exposures, and other factors (e.g., medications) and do not provide a measure of tobacco exposures during periods of initiation and early promotion of lung cancer. Methods for collecting proxy information on affected deceased individuals require further standardization.

Studies of genetic factors for lung cancer risk seem difficult to implement because a large number of exposures that increase the risk for lung cancer have been well characterized. However, because the requisite occupational and environmental exposures are well known, delineation of risk into genetic, environmental, and interactions between genetic and environmental factors should be possible.

**Conclusions**

A general schema for incorporating measures of genetic and environmental exposures in lung cancer studies includes both case-control studies as well as family studies. For instance, the case-control studies should be used to evaluate the relationship between measures of metabolism in the lungs and easily available samples (e.g., urine, blood) and to estimate sources of intra-individual variation (such as measurement error and diurnal variation) and interindividual variation (including ethnicity). For non-inducible traits or traits not affected by the disease process, case-control studies can identify biomarkers for lung cancer. Ethnicity must be collected as a covariate in these studies (120). For inducible biomarkers, adequate statistical adjustment for changing behavior may be impossible. Family studies of unaffected relatives of cases and controls would then be used to verify the evidence for an effect of the host factors on the risk for lung cancer.

Finally, evidence that specific loci have a major impact on susceptibility for disease can be assessed by genetic linkage studies. These methods, however, will be expensive to implement because blood samples from many affected individuals will not be available, requiring that many of their offspring be studied. These approaches have nevertheless been useful in studying other diseases with a late age at onset and rapid course (121, 122). Several loci are likely to be involved in the disease process, possibly reducing the power of the studies. Assays of the metabolic status of particular alleles defined by DNA polymorphisms are becoming possible for a few loci such as CYP2D6. As more candidate loci are studied, methods for identifying allelic variation will become simpler and more reliable. Genetic linkage strategies will remain useful because they address the more general hypothesis that a genetic locus situated in a particular region of the genome affects susceptibility to lung cancer.

**References**


Host factors in lung cancer risk: a review of interdisciplinary studies.

C. I. Amos, N. E. Caporaso and A. Weston


Updated version
Access the most recent version of this article at:
http://ceb.p.aacrjournals.org/content/1/6/505

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.