ABSTRACT

In the United States, the 20th century witnessed the emergence of a lung cancer epidemic that peaked and began to decline by the century’s end, a decline that continues today. However, lung cancer continues to be an unabating pandemic. In research carried out over the last half of the 20th century, many factors were causally associated with lung cancer and studies were implemented to identify determinants of susceptibility to these factors. Cigarette smoking was identified as the single most predominant cause of the lung cancer epidemic, but other causes were found, including workplace agents (eg, asbestos, arsenic, chromium, nickel, and radon) and other environmental factors (passive smoking, indoor radon, and air pollution). Contemporary epidemiologic research on lung cancer now focuses on a new set of issues, primarily related to susceptibility to the well-identified causal factors, particularly smoking, and on the consequences of changes in tobacco products for risks to smokers. Diet and the possibility of reducing risk through chemoprevention remain a focus of research emphasis through experimental and observational approaches. Questions have also been raised about possible differences in susceptibility to lung cancer by sex and race.

Population patterns in smoking prevalence will continue to be the most powerful predictor of the future occurrence of lung cancer. Evaluation of recent US patterns in smoking prevalence indicates that for the next approximately 10 to 15 years, lung cancer rates will decrease, but will then level off starting in approximately 2030. Unless further reductions in the prevalence of cigarette smoking are achieved over the next decade, lung cancer will remain as an all too common, but avoidable, disease.

VOLUME 23 • NUMBER 14 • MAY 10 2005

INTRODUCTION

While writing this manuscript on the epidemiology of lung cancer in 2004, we faced the task of identifying what findings are new and what research challenges can be anticipated. After all, the invitation to write on the topic came 40 years after the release of the landmark 1964 report from the US Surgeon General on smoking and health,1 a report that concluded that smoking causes lung cancer in men, and in the same year as the 2004 report of the Surgeon General,2 which once again covered lung cancer.

The context for this article is set by more than half a century of epidemiologic research on lung cancer. While some epidemiologic studies on lung cancer were published before 1950, most notably in Nazi Germany,3 the year 1950 is widely considered as the start of a rising wave of epidemiologic research on lung cancer. In 1950, the results of five case-control studies associating smoking with lung cancer were published, including the often-cited papers of Wynder and Graham,4 Levin et al,5 and Doll and Hill.6 By 1954, the retrospective findings of the case-control studies were confirmed with the initial findings from Doll and Hill’s prospective cohort study of British physicians.7 During the 1950s, findings of additional case-control and cohort studies were reported, leading to the first syntheses of the evidence and the conclusion that smoking caused lung cancer.8-10 The 1964 report by the Surgeon General’s committee was definitive and readily withstood counterattack by tobacco industry critics.
Over the four decades since the 1964 Surgeon General’s report, epidemiologic research on lung cancer has been conducted with rising frequency. While cigarette smoking has always remained as a central theme, other causes have been evaluated and substantial research has been directed at environmental and genetic factors that might modify the risk of smoking. The first major epidemiologic studies on passive smoking and lung cancer were reported in 1981 and subsequently more than 50 studies have reported on this topic. For several decades, researchers have addressed genetic determinants of lung cancer risk, and much of the ongoing epidemiologic research is based in the hybrid approach of molecular epidemiology, which joins laboratory and population research methods.

Trends in the population prevalence of cigarette smoking strongly predict lung cancer incidence and mortality rates, which closely parallel incidence because of the still tragically high case fatality of those with lung cancer. In the United States, the prevalence of cigarette smoking in males declined consistently from the release of the 1964 Surgeon General’s report until approximately 1990, at which point the prevalence leveled at approximately 25%. The smoking prevalence in US females also leveled off at approximately the same time and at the same level as in men (Fig 1). The historic trends in US smoking prevalence thus provide an explanation for past trends in lung cancer rates and for the current rates, while also providing a sufficiently adequate predictor of future occurrence. The patterns of smoking prevalence to the present indicate that the lung cancer mortality rates will continue to decrease until approximately 2020, assuming an approximate 30-year lag between the population smoking patterns and subsequent lung cancer incidence, and then remain constant, reflecting the current, relative stability of smoking prevalence. By 2030, the lung cancer cases will no longer have a predominance of men, but include equal numbers of men and women, corresponding to the present pattern of smoking. This prediction is a strong rationale for continuing strong tobacco control efforts; without sustained continued cessation and reduction of initiation, lung cancer rates will not fall in the future.

Also contributing to the current burden of lung cancer are workplace exposures of earlier decades, particularly the widespread exposure to asbestos in the 1940s through the 1960s. Even though the contribution to overall rates is relatively small compared with cigarette smoking, the improved control of workplace exposures to lung carcinogens represent a notable achievement and will contribute to the downward trend in future rates. For example, the incidence rates of mesothelioma among men have begun to decline, consistent with controls of exposure to asbestos in the workplace.

Globally, lung cancer shows marked regional variation, with age-standardized mortality rates varying over 25-fold in both men and women. Lung cancer tends to be most common in developed countries, particularly in North America and Europe, and less common in developing countries, particularly in Africa and South America. Within countries, lung cancer rates among men invariably outpace those seen in women, reflecting sex differences in historic patterns in the prevalence of cigarette smoking. In Chinese women, the lung cancer mortality rates are incongruous with the low prevalence of smoking, with notably high rates of adenocarcinoma.

In this article, we pose questions concerning some of the most active areas of current epidemiologic research on lung cancer. We then address the critical role that molecular epidemiology will play in future research. Our approach highlights those areas where significant questions remain to be answered and contributions can be anticipated. For recent reviews of the topic in general, we recommend the 2003 review by Alberg and Samet and the 2004 reports of the International Agency for Research on Cancer of the WHO and of the US Surgeon General. We begin by considering the changing histopathologic pattern in lung cancer diagnoses.

**TOPICAL ISSUES OF THE PRESENT**

**Why Are the Population Patterns of Lung Cancer Histopathology Changing?**

The four major histologic types of lung cancer—squamous cell carcinoma, adenocarcinoma, large-cell carcinoma, and small-cell undifferentiated carcinoma—together account for over 90% of lung cancer cases in the United States. Smoking causes each of the major histologic types, although the trend in risk with number of cigarettes smoked varies across the types, being steepest for small-cell undifferentiated carcinoma. Among the few suggestive links of histologic type with occupational agents are small cell lung cancer in excess in workers exposed to chloromethyl ethers and in underground miners exposed to radon progeny.

Notable shifts have taken place in the incidence rates of lung cancer by histologic type. In the initial decades of the smoking-caused epidemic of lung cancer, squamous cell carcinoma was the most frequent type of lung cancer observed among smokers, and small cell carcinoma was the next most frequent. After steadily increasing occurrence during the period 1973 to 1987, adenocarcinoma supplanted squamous cell carcinoma as the most frequent form of lung cancer.

The hypothesis that changing patterns of diagnosis and classification of lung cancers could have led to these changes over time has largely been set aside, as the rise in adenocarcinoma antedated diagnostic innovations
such as the fiberoptic bronchoscope and thin-needle aspiration, and the development of improved stains for mucin, the hallmark of adenocarcinoma. Hypotheses concerning the shift in histopathology have focused on the potential role of changes in the characteristics of cigarettes and consequent changes in the inhaled doses of carcinogens. Puff volume has likely increased over recent decades, with the possibility that patterns have changed to enhance deposition of tobacco smoke in the peripheral airways and alveoli of the lung. Nitrate levels in tobacco smoke have also increased, which enhance the combustion of tobacco smoke. More complete combustion decreases the concentrations of polycyclic aromatic hydrocarbons, but the increased production of nitrogen oxides contributes to increased formation of tobacco-specific nitrosamines. An increase in dose of the potent tobacco-specific nitrosamine, NNK, has been postulated as one factor leading to the increase in adenocarcinoma. NNK induces lung carcinomas, predominantly adenomas and adenocarcinomas, in mice, regardless of route of administration.

Few studies can provide data to test these hypotheses because of the need for longitudinal observation of lung cancer risk in relation to the characteristics of the cigarettes smoked over time. Thun et al compared risks for lung cancers of the different histologic types among participants in the American Cancer Society’s Cancer Prevention Study (CPS) I and CPS II with the markedly rising risks associated with smoking for adenocarcinoma of the lung in both men and women over the approximately 20 years separating the two studies, suggesting that changes in smoking behavior and cigarette design are contributing to the rise in the occurrence of adenocarcinoma.

**Does Lung Cancer Risk Differ by Race and Ethnicity?**

Addressing disparities in the population burden of cancer is an important element of the nation’s health agenda, and lung cancer is a disease with large disparities by race/ethnicity and socioeconomic status. The patterns of occurrence of lung cancer by race and ethnicity make lung cancer a relevant disease for those concerned with the health of minorities. Of particular note is that whereas lung cancer incidence rates are similar among black and white American women, lung cancer occurrence is much higher among black American men than among white American men. The United States currently has two sources of cancer surveillance data, the National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) Program and the Centers for Disease Control and Prevention’s National Program of Cancer Registries (NPCR). In the third report on nationwide cancer incidence in 2001, which incorporated SEER and NPCR data, the percentage difference in lung cancer incidence rates between black and white men was 26% (109.0 to 86.8 per 100,000 per year) in the NPCR/SEER data and 36% in the SEER data alone for the period 1997 to 2001 (104.1 to 76.6 per 100,000 per year). Thus, both data sets highlight a striking racial/ethnic disparity in lung cancer incidence rates.

The higher mortality rates of lung cancer in black Americans compared to white Americans can be attributed not only to higher incidence rates, but also to the poorer survival of lung cancer patients who are black compared with those who are white. The 5-year relative survival rate was 14.3% lower in black Americans compared with white Americans during the period 1995-2001.
Lung cancer mortality rates among Hispanics, Native Americans, and Asians/Pacific Islanders are significantly lower than rates among blacks and non-Hispanic whites. Nevertheless, lung cancer poses a considerable public health burden among these groups.

The basis for the higher lung cancer rates in black American males compared with white American males is not understood. This disparity is not readily explained on the basis of historic smoking patterns. The hypothesis has been proposed that menthol cigarettes may increase the risk of lung cancer even more than non-menthol cigarettes; if so, the greater prevalence of menthol cigarette use in black Americans compared to white Americans (69% to 22%) may contribute to this racial disparity. Menthol cigarettes could be associated with greater risk, for example, if smokers of menthol cigarettes inhaled more deeply than smokers of non-menthol cigarettes because of the local anesthetic effects of menthol. The evidence to date suggests that menthol cigarette smokers do not have a greater risk of lung cancer than non-menthol smokers, regardless of racial/ethnic group. Priority should be placed on testing new hypotheses to help better understand the persistent excess in lung cancer rates in blacks compared to other racial/ethnic groups so that this disparity can be eliminated.

Why Is Lung Cancer More Common Among Those of Lower Socioeconomic Status?

Lung cancer is more likely to occur in the poor and less educated, a pattern that is observed in many countries worldwide. For example, in Canada, the risk of lung cancer in both sexes was inversely associated with income, education, and social class, even after adjustment for cigarette smoking. In China, those classified as low income had a six-fold increased risk of lung cancer compared with those in the high income category. In the Netherlands, the risk of lung cancer was inversely associated with attained education, an association that was not attributable to occupational exposures. Lower socioeconomic status has also been associated with later stage at diagnosis.

Socioeconomic status is associated with a constellation of interacting determinants of lung cancer risk, such as smoking, diet, and exposures to inhaled carcinogens in the workplace and in the general environment. Lower socioeconomic status is uniformly associated with an unfavorable profile for all of these factors. Advancing our understanding of the complex linkages between components of socioeconomic status and lung cancer risk is essential to effectively addressing this social class disparity and reducing lung cancer rates in the poorer segments of society. The pace of publications on this important topic does not seem commensurate with its importance, particularly given the strong, negative association of smoking prevalence with indicators of socioeconomic status.

Are Women More Susceptible to Smoking-Induced Lung Cancer?

In the 1964 Surgeon’s General report, the conclusion that smoking caused lung cancer was limited to men because not enough evidence was yet available to make a conclusive determination of causality in women. Since that time, smoking was conclusively found to be a cause of lung cancer in women.

Results of some case-control studies have suggested a potentially higher risk of smoking-associated lung cancer in women compared with men, but the interpretation of these studies is complicated by the reliance on relative, rather than absolute, comparisons with data from case-control studies. Furthermore, the evidence from prospective cohort studies fails to support a sex differential in susceptibility to lung cancer from smoking. The equal rates of lung cancer mortality in younger US men and women corresponding to a time of equal smoking prevalence also provide evidence against a major sex difference in susceptibility to smoking-induced lung cancer.

At present, the evidence weighs against greater risks in women, particularly as studies that have compared the relative risk estimates for men and women with similar smoking histories find similar associations.

Does Diet Influence Lung Cancer Risk?

Diet has been of interest as a potential determinant of the risk of lung cancer, particularly in smokers, spurred initially by the pioneering epidemiologic work of Bjelke and the original vitamin A and beta-carotene hypotheses. Bjelke and subsequent researchers originally focused on vitamin A because of its role in cellular differentiation and the promise of the initial observational findings, but this line of inquiry was subsequently expanded to include antioxidant micronutrients, with an emphasis on beta-carotene. The more general hypothesis has been advanced that antioxidant micronutrients may protect against oxidative damage to DNA and thereby protect against cancer. The results of the chemoprevention trials, which failed to show that beta-carotene (and alpha-tocopherol and retinol) protects against lung cancer in high-risk groups, have not slowed the pace of publications exploring the link between diet and lung cancer. Some of the more recent research has enriched the evidence base for understanding the associations between lung cancer risk and exposures that were commonly studied in earlier studies, such as intake of fruits and vegetables and micronutrients. The results of case-control and prospective cohort studies have tended to show that individuals with high dietary intake of fruits or vegetables have a lower risk of lung cancer than those with low fruit or vegetable intake. Evidence from cohort studies published since 2000 has tended to
reinforce this pattern of associations.\textsuperscript{54-58} In the European Prospective Investigation into Cancer and Nutrition study,\textsuperscript{59} and in a pooled analysis of cohort studies,\textsuperscript{60} the protective association was stronger for fruit than for vegetable consumption.

New lines of inquiry have also emerged, such as studies of phytochemicals such as flavonoids and isothiocyanates. Phytochemicals are low molecular weight molecules produced by plants. Of the many classes of phytochemicals, those studied in relation to lung cancer include phytosterogens, flavonoids, and glucosinoids. The tumor promoting effects of steroid hormones can be blocked by phytosterogens. Soya beans are a primary source of a specific class of phytosterogens known as isoflavonoids. The relatively few studies to date of isoflavonoids in relation to lung cancer have not provided evidence of a link.\textsuperscript{61} Flavonoids exhibit potent antioxidant activity. Flavonoid intake has been at least weakly associated with lung cancer in some of the preliminary studies of this topic.\textsuperscript{61,62} Isothiocyanates are metabolites of the class of phytochemicals known as glucosinolates. Isothiocyanates could exert anticancer effects by blocking carcinogens via induction of phase II detoxification enzymes, such as glutathione S-transferase. Cruciferous vegetables contain high concentrations of glucosinolates, and hence consumption leads to higher endogenous isothiocyanate concentrations. As with cruciferous vegetables,\textsuperscript{63} lung cancer risk is also consistently lower with higher intakes or urinary levels of isothiocyanates.\textsuperscript{63-65} When isothiocyanates have been studied in combination with GSTM1, the decreased risk of lung cancer associated with isothiocyanates has been especially pronounced in persons with the GSTM1 null genotype.\textsuperscript{63-65} This provides an example of a potential gene-diet interaction that may be relevant to lung carcinogenesis.

Studies of fruits, vegetables, and micronutrients have been the centerpiece of studies of diet and lung cancer, but a wide range of dietary and anthropometric factors have been investigated. For example, the results of a meta-analysis showed alcohol drinking in the highest consumption categories was associated with increased risk of lung cancer.\textsuperscript{66} Anthropometric measures have also been studied, indicating a tendency for persons with lower body mass index to have increased lung cancer risk relative to heavier persons.\textsuperscript{67,68} However, both alcohol drinking and low body mass index may be difficult to separate from the concomitant effects of smoking. At present, when considering the possible relationships between lung cancer and factors such as alcohol drinking and lower body mass index, cigarette smoking cannot be dismissed as a possible explanation.

Molecular Epidemiology: The Future?

Much current and future epidemiologic research is based in—and will continue to be based in—advancing knowledge of the molecular and cellular basis of respiratory carcinogenesis. The foundation for this line of research is the combination of sound epidemiologic design, coupled with collection of biologic materials, often including tissue specimens and blood for serum, RBCs, and WBCs. To measure exposure to environmental factors, questionnaires are often quite sophisticated. While a smoking history is readily obtained, valid instruments for diet and occupation are extensive. Additionally, substantial sample sizes are needed to have sufficient statistical power, particularly for studies of gene-environment interactions.

Lung cancer is unique among human solid cancers in that a single environmental factor—tobacco smoke—is believed to promote sequential changes in target cells that lead to carcinogenesis. Recently, progress has been made in elucidating changes involved in the transformation of a normal bronchial epithelium to frank malignancy. The sequence of morphologic changes associated with lung tumor progression is well established only for one major histologic type, squamous cell carcinoma.\textsuperscript{59} Under this model, based on observations in underground uranium miners, the sequence of events flows from normal bronchial mucosa to basal cell hyperplasia to squamous metaplasia to squamous dysplasia to carcinoma in situ and on to squamous cell carcinoma. Not until 1999 was the first preinvasive entity, atypical adenomatous hyperplasia, added for the most common histologic subtype, pulmonary adenocarcinoma.\textsuperscript{70}

Since the morphologic pathway for squamous cell carcinoma was described more than 30 years ago, the prevailing model of carcinogenesis has favored the hypothesis that underling these sequential histologic changes is a progression of genetic and epigenetic abnormalities that eventually lead to the loss of normal control mechanisms of cellular growth. Multiple genetic and epigenetic hits accumulate over time, affecting factors such as oncogenes, tumor suppressor genes, growth factors, and DNA repair genes, resulting in constitutive growth stimulation. The step-wise progression of lung carcinogenesis has yet to be as elegantly described as in colon cancer, but recent molecular evidence is refining understanding of preneoplastic morphologic categories.

Distinctions in molecular markers have even been identified when the bronchial mucosal biopsies of cigarette smokers and nonsmokers were histologically indistinguishable. For example, genetic mutations, such as allelic loss of one or more chromosomal regions, were observed much more frequently in the histologically nonmalignant bronchial epithelium of smokers with lung cancer than in never-smokers.\textsuperscript{71} Studies of epigenetic abnormalities also provide supportive evidence. Smokers without lung cancer had a 35% frequency of p16 promoter hypermethylation in their sputum, but methylation was not detected in the bronchial epithelium of never-smokers.\textsuperscript{72}
These new molecular data suggest revising the preinvasive paradigm of lung cancer to include an additional nonmorphologic, but distinct, subcategory for smoke-damaged, histologically normal bronchial mucosa as an intermediate step in the progression from normal bronchial mucosa to basal cell hyperplasia. Adding this new step to the pathway from non–smoke damaged bronchial mucosa to squamous cell lung cancer, we will briefly describe the very early molecular changes involved in lung cancer progression with a view to examining how risk factors such as tobacco exposure may relate to these intermediate markers. These intermediate stages would be useful outcomes for epidemiologic studies. Table 1 briefly outlines some of the molecular changes observed in the sequential development of early squamous cell lung cancer.

Early molecular changes. One of the earliest, most consistent abnormalities in lung cancer is the loss or inactivation of genetic material on the short arm of chromosome 3p. In lung cancer patients with 3p changes detectable in tumor, greater than 30% of nonmalignant bronchial epithelium also had 3p changes. Both the size of the 3p deletions and the frequency of their occurrence progressively increased with the severity of the histopathologic changes, to 42% for basal cell hyperplasia and squamous metaplasia, 81% for squamous dysplasia, and finally to 100% for carcinoma in situ and histologically proven squamous cell carcinoma.

This 3p region is targeted early for inactivation since it is thought to be the location for multiple tumor suppressor genes, including RARβ (retinoic acid receptor beta) on 3p24, RASSF1A on 3p21.3, and FHIT (fragile histidine triad) gene (3p14.2). Eliminating large portions of chromosomal material on 3p is not the sole way that critical tumor suppressor genes in this region can be inactivated. Perhaps as frequent an event is the epigenetic silencing of tumor suppressor genes by DNA promoter hypermethylation.

When these early epigenetic and somatic changes have been assessed in relation to smoking, the evidence suggests that smoking produces an increase of epigenetic and genetic alterations over time with increasing alterations with greater amount smoked. RARβ was methylated more frequently in people with a smoking history than in never-smokers; moreover, the frequency of RARβ methylation progressed with the severity of the clinical stage of the lung cancer. The frequency of RASSF1A methylation is associated with duration of smoking history, as in two studies RASSF1A hypermethylation was most frequent in lung cancer patients who began smoking as teenagers. Early genetic alterations, including the somatic deletions on the short arm of 3p, have also been shown to be strongly associated with smoking history.

The tumor suppressor gene, p16, is another frequent target for inactivation in lung cancer by epigenetic silencing or genetic deletions. Since p16 is an inhibitor of cyclin D kinases 2, 4, and 6, its lack of expression allows unregulated phosphorylation of the Rb protein so that the Rb protein is incapable of preventing cell cycle progression and uncontrolled cell division. As with genetic deletions, p16 promoter hypermethylation appears to be an early alteration observed in the precursor lesions of squamous cell cancer and was more prevalent with each successive stage of the progression model: basal cell hyperplasia (17%), squamous metaplasia (24%), and carcinoma in situ (50%). In another study, p16 methylation was detected in the sputum of 4% of heavy smokers without cancer compared to 32% of patients with lung cancer. Lung cancer from smokers with less than 30 pack-year history had significantly less frequent p16 and RASSF1A methylation than lung cancers from smokers with greater than 30 pack-year history, with a dose-response relationship reported.

Other molecular changes. Mutation of the H-, K-, and N-ras family of oncogenes is a common event in human cancer, particularly lung cancer. Among the potential mutations of the Ras family of oncogenes, it is the K-ras gene that is most often mutated, with point mutations occurring predominately at codon 12 and, occasionally, at codons 13 or 61. These mutations result in amino acid substitutions that preserve the activated conformation of K-ras leading to constitutively functioning signal transduction and uncontrolled cell proliferation. K-ras mutation is particularly common in patients with extensive smoking histories, with prevalences of approximately 30% versus 5% in pulmonary adenocarcinomas from ever-smokers compared to nonsmokers. A carcinogen in cigarette smoke, benzo[a]pyrene, is known to induce a

| Table 1. Molecular Targets in the Progression Pathway to Squamous Cell Carcinoma of the Lung |
|----------------------------------|-------------------------------------------------|
| Bronchial mucosa with molecular abnormalities | Allelic loss of genetic material from four chromosomal regions-3p |
| | Epigenetic changes in p16 |
| | Microsatellite alterations |
| | Small telomeric deletions |
| Basal cell hyperplasia | Telomerase dysregulation |
| | MYC overexpression |
| | Epigenetic changes in p16 |
| Squamous metaplasia | 8p allelic loss |
| | Epigenetic changes in p16 |
| Squamous dysplasia | 17p loss |
| | p53 immunostaining |
| | p53 mutations |
| | Aneuploidy |
| | Epigenetic changes in p16 |
| Carcinoma in situ | Deletions contiguous to 3p |
| | 5q APC loss |
| | K-ras mutations |
| Squamous cancer | Epigenetic abnormalities in p16, APC, CDH13, DAPK, MGMT, TIMP3, RARβ, FHIT, RASSF1A |
| | Progressive increase in epigenetic aberrations of p16, APC, CDH13, with increasing cancer stage |

Information downloaded from jco.ascopubs.org and provided by Harvard Libraries on January 6, 2010 from 128.103.149.52. Copyright © 2005 by the American Society of Clinical Oncology. All rights reserved.
G-to-T transition involving codon 12 of the \textit{K-ras} gene.\textsuperscript{90} Destro et al\textsuperscript{86} frequently recognized \textit{K-ras} point mutations in patients who had stage I lung tumors, but unlike their \textit{p16} methylation data, \textit{K-ras} mutations were not detected in the sputa of chronic smokers without lung cancer.

\textit{p53} is the most frequently mutated gene in human cancer, and lung cancer is no exception. \textit{p53} is abnormal in approximately 50\% of non–small-cell lung cancers and 90\% of small-cell lung tumors.\textsuperscript{91,92} The \textit{p53} gene encodes for a 53-kd nuclear phosphoprotein that directs the cells to arrest either at the G1/S or the G2/mitosis phase if the cell suffers DNA damage. Alternatively, \textit{p53} induces the cell towards an apoptotic death.\textsuperscript{93,94} Cells with mutant \textit{p53} lack these checkpoints, and damaged DNA becomes incorporated into the genome of the daughter cells, facilitating malignant transformation. In non–small-cell lung cancer, most \textit{p53} mutations involve G-to-A transversions, whereas in small-cell lung cancer, G-to-T transitions are relatively common.\textsuperscript{95} As seen with \textit{K-ras} mutations, these alterations are known consequences of benzo[a]pyrene, a carcinogen in cigarette smoke. Similar to epigenetic changes, tobacco smoking is associated with a dose-response increase in \textit{p53} somatic mutations in lung cancer.\textsuperscript{96}

Even in nonsmokers, genetic changes may affect patients’ responses to treatment. In Asia, where there is a high incidence of non–small-cell adenocarcinoma lung cancer in nonsmoking women, clinical trials of gefitinib (a tyrosine kinase inhibitor) indicate more frequent and more complete responses to the drug than observed in largely European populations.\textsuperscript{97} An activating mutation of the epidermal growth factor receptor that sensitizes nonsmoking female patients with adenocarcinoma to gefitinib therapy has recently been described.\textsuperscript{98,99} These activating mutations are found in whites, but appear to have a higher prevalence in Asian women, and this may explain the dramatic response rates to therapy seen in Asia.\textsuperscript{98,100}

As the genetic and epigenetic alterations in the carcinogenic pathway are elucidated, an important line of future inquiry will be to conduct carefully designed epidemiologic studies to characterize the relationships of known lung cancer risk factors to critical changes at different steps in the pathway to lung carcinogenesis. This will not only lead to further understanding of the specific mechanisms through which lung carcinogens cause lung cancer, but an epidemiologically based understanding of the sequence of these changes will also be essential for translating the findings of this research into prevention strategies.

\textbf{What Accounts for Differences in Host Susceptibility to Lung Cancer?}

An extensive set of causal risk factors for lung cancer has been identified, including smoking—an exceptionally strong causal factor. A patient’s history of exposure to established lung carcinogens provides a clinically meaningful risk factor profile. Clearly diagnostic probabilities depend strongly on smoking and occupational risk factors. Yet despite this constellation of causes, it remains difficult to identify susceptible persons. For example, even though cigarette smoking is associated with a greatly increased risk for lung cancer, relative to never-smokers, the absolute lifetime risk of a smoker developing lung cancer is approximately 10\% to 20\%.\textsuperscript{101} Those smokers at greatest risk have yet to be identified. The genetic variation between individuals, measured by either genotype or phenotype, has been the focus in the hunt to identify the determinants of susceptibility to tobacco-induced lung cancer using both case-control and cohort designs. Emphasis has been placed on metabolic pathways related to carcinogen detoxification and activation, as well as DNA repair.

The metabolism of toxic agents, including carcinogens, generally proceeds through two phases.\textsuperscript{102} In phase I, unreactive, nonpolar compounds are converted, usually by oxidative reactions, to highly reactive intermediates. These intermediates are then able to form complexes with conjugating molecules in phase II conjugation reactions, which are usually less reactive and more easily excreted. However, the intermediate metabolite may react with other cellular components, such as DNA, before conjugation occurs. This binding to DNA may be the first step in the initiation of carcinogenesis.\textsuperscript{102}

Many carcinogenic compounds in tobacco smoke (eg, polycyclic aromatic hydrocarbons) undergo metabolic activation by phase I enzymes of the cytochrome p450 system to form reactive intermediates that bind to DNA and cause genetic injury. Several of these enzymes have been investigated with regard to lung cancer risk, including \textit{CYP1A1} and \textit{CYP2D6}. For the \textit{CYP1A1} gene, the current evidence suggests that two specific polymorphisms, the \textit{Msp1} polymorphism\textsuperscript{103} and a polymorphism in exon 7,\textsuperscript{104} are associated with increased risks of lung cancer.

Both metabolic phenotype and genotype have been examined for \textit{CYP2D6}. This enzyme determines the phenotype for debrisoquine metabolism, which was studied extensively in the past as a phenotypic risk factor for lung cancer.\textsuperscript{105} The initial case-control studies found that fast metabolizers had a greater lung cancer risk, consistent with the hypothesized role of rate of metabolism in determining lung cancer risk,\textsuperscript{106} although a subsequent and larger study found no association.\textsuperscript{107} More recent studies that have assessed genotype have generated inconsistent results\textsuperscript{108} suggesting the possibility that the \textit{CYP2D6} fast metabolizer genotype may be weakly associated with increased risk of lung cancer.\textsuperscript{109}

Glutathione S-transferase is a phase II enzyme that detoxifies reactive metabolites of polycyclic aromatic hydrocarbons. There are at least four genetically distinct classes
of the glutathione S-transferases: $\mu$, $\alpha$, $\pi$, and $\theta$. The risk estimates from a meta-analysis indicate that individuals with the $GSTM1$ null genotype have higher risk of lung cancer than those with the $GSTM1$ present genotype, but a pooled analysis of data from 21 case-control studies did not indicate that this susceptibility was stronger among cigarette smokers than among nonsmokers.\textsuperscript{110} The importance of interactions between genes is highlighted by the joint assessment of the $CYP1A1$ Ile462Val and $GSTM1$ null polymorphisms in nonsmokers, which indicated the combination of the two variant genotypes was associated with a greater than four-fold increased likelihood of lung cancer compared with the combination of the two nonvariant genotypes.\textsuperscript{111}

With rapidly developing technologic advances for high-throughput genotyping, studies of multiple genetic markers can be anticipated. Experience to date suggests that few genes with powerful effects on susceptibility will be identified. The expanded genetic information will enable assessment within and across candidate genetic pathways, rather than candidate genes. Substantial research has been directed at DNA repair and susceptibility in lung cancer and other tumors.\textsuperscript{112,113} Persons with specific rare recessive traits (eg, xeroderma pigmentosa) have long been known to be at increased risk for cancer. DNA repair capacity has now been examined as a specific risk factor for lung cancer, with the underlying hypothesis that lesser capacity would lead to greater lung cancer risk from the multiple DNA-damaging components of tobacco smoke. While much research remains to be done to clarify the association between variation in DNA repair capacity and lung cancer risk, the evidence to date suggests this is a promising lead.\textsuperscript{114} There are a variety of phenotypic assays for susceptibility to DNA damage. Individuals with a less proficient DNA repair capacity phenotype, as measured by a nonspecific mutagen sensitivity assay, have been shown to have an increased risk of lung cancer in some studies.\textsuperscript{115-117} Studies of DNA repair genes have been conducted, including studies of $XPA$, $XPD$, x-ray repair complementation group 1 and 3 ($XRCC1$ and $XRCC3$), excision repair complementation group 1 ($ERCC1$), and $hoGG1$. One of the most extensively studied DNA repair genes is the nucleotide excision repair gene $XPD$\textsuperscript{118-121}; the evidence to date has not yet revealed a consistent pattern of associations for any specific polymorphism.

**CONCLUSION: ENDING THE GLOBAL EPIDEMIC**

Lung cancer accounts for 12% of all cancers diagnosed worldwide, making it the most common malignancy, other than nonmelanoma skin cancer.\textsuperscript{122} The high frequency of diagnoses, combined with the poor survival for lung cancer, makes lung cancer far and away the leading cause of cancer death in the world.

Developed countries have much higher rates of lung cancer than developing countries, but due to developing countries representing a greater proportion of the world’s population, the absolute numbers of lung cancer diagnoses is almost equally divided between developed and developing nations.\textsuperscript{122} The global burden of lung cancer will undergo drastic shifts in the coming decades, unfortunately from the developed to the developing world. As this occurs, lung cancer mortality rates are projected to worsen from the world’s 10th to the world’s fifth leading cause of death.\textsuperscript{123} Some scenarios for future lung cancer occurrence offer foreboding predictions. China, which now has 350 million smokers, due primarily to the current high smoking prevalence in males, will sustain an extraordinary number of cases in the future. As the world’s most populous nation, the ubiquitousness of smoking among Chinese men leads to a staggering future burden predicted for China, which will fuel drastic increases in the global burden of lung cancer.\textsuperscript{124,125}

Nonetheless, there is a solid scientific foundation for prevention. Experience in the United States and elsewhere shows that tobacco use can be controlled. Research in progress should lead to markers for identifying susceptible people and for early detection. The lung cancer epidemic can be ended.

**Authors’ Disclosures of Potential Conflicts of Interest**

The authors indicated no potential conflicts of interest.
32. National Cancer Institute: Cancer rates and risks. National Cancer Institute, Bethesda, MD, 2004

www.jco.org

The Future of Lung Cancer Epidemiology


