Epidemiology of Lung Cancer

Diagnosis and Management of Lung Cancer, 3rd ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines

Anthony J. Alberg, PhD, MPH; Malcolm V. Brock, MD; Jean G. Ford, MD, MPH, FCCP; Jonathan M. Samet, MD, FCCP; and Simon D. Spivack, MD, MPH

Lung cancer is the leading cause of cancer death in the world. In 2008, >1.6 million people received a new diagnosis of lung cancer, comprising 13% of all new cancer diagnoses, and 1.4 million died of lung cancer, which was 18% of all cancer deaths.2

Many causes of lung cancer have been identified, including active cigarette smoking; exposure to secondhand cigarette smoke (passive smoking); pipe and cigar smoking; occupational exposure to agents such as asbestos, nickel, chromium, and arsenic; exposure to radiation, including radon gas in homes and mines; and exposure to indoor and outdoor air pollution. Despite the identification of this constellation of well-established causal risk factors, the global epidemic of lung cancer is primarily caused by a single factor: cigarette smoking. This dominance of cigarette smoking reflects effective marketing and promotion of an addicting and deadly product by multinational corporations.
Refining the understanding of the etiology and pathogenesis of lung cancer remains a vibrant area of research. As described in this chapter, foci of current research include understanding the root causes of racial and socioeconomic disparities, elucidation of the role of lifestyle factors other than cigarette smoking (eg, diet, physical activity), the risk of indoor and outdoor pollutants, genetic determinants of risk, biomarkers of risk and early detection, and the potential role of infections such as HIV.

An understanding of the epidemiology of lung cancer provides background and contextual information regarding lung cancer that is important for management of guidelines. This article includes no recommendations for individual patients but is included in the American College of Chest Physicians (ACCP) Lung Cancer Guidelines to provide a foundation.

1.0 METHODS

A narrative review of published evidence on the epidemiology of lung cancer was carried out. Key reports that described the occurrence of lung cancer in populations and factors that affect lung cancer risk were identified. This review was accomplished through a combination of approaches that included cataloging searches that included the term “lung cancer” and terms for various exposures that have been studied in relation to lung cancer (eg, “smoking,” “asbestos,” “radiation”). Emphasis was placed on systematic reviews, if available.

The objectives were to provide a summary of the epidemiologic evidence on lung cancer, emphasizing issues that are currently relevant to prevention. The literature is extraordinarily large, and we did not attempt to conduct a comprehensive review and systematic synthesis. Such syntheses have been carried out by expert review groups, including the committees assembled to prepare the US Surgeon General’s reports on smoking and health and committees of other governments and organizations, including the UK Royal College of Physicians and Scientific Committee on Tobacco, the World Health Organization (WHO) International Agency for Research on Cancer (IARC), and the World Cancer Research Fund (WCRF).

The topics covered were agreed on by consensus of the writing committee with initial input from the ACCP Lung Cancer Guidelines Panel. Topics were added as recommended by external reviewers from the ACCP Lung Cancer Guidelines Panel, the Thoracic Oncology NetWork, Guidelines Oversight Committee (formerly known as the Health and Science Policy Committee), and the Board of Regents of the ACCP. All parties agreed to make no attempt to grade the evidence or generate formal guidelines.

2.0 PATTERNS OF OCCURRENCE

The patterns of occurrence of lung cancer with respect to survival, incidence, and mortality rates are reviewed in this section, using the United States as a specific example before going on to consider global variation in rates.

2.1 Survival

The 5-year relative survival rate for lung cancer in the United States for the period of 2001 to 2007 is 16.3%, which is up from 12.3% in 1975 to 1977. The 5-year relative survival rate varies markedly depending on the stage at diagnosis, from 52% to 24% to 4% for local, regional, and distant stage disease, respectively. Stage at diagnosis accounts for the most marked variation in prognosis, but patient characteristics associated with poorer survival also include being older, male, and African American.

2.2 Temporal Trends

Because of the high case fatality rate of lung cancer, incidence and mortality rates are nearly equivalent, and consequently, routinely collected vital statistics provide a long record of the occurrence of lung cancer. We are presently amid an epidemic of lung cancer that began in the 1930s in the United States.

2.3 Sex

As the leading cause of cancer death among women, lung cancer is a major women’s health issue. Historical trends indicate that cigarette smoking prevalence peaked about 2 decades earlier in men than in...
women; thus, the epidemic of lung cancer started later in women. In contrast to men, lung cancer incidence rates in women have not yet begun to decrease consistently, but a recent analysis of 2003 to 2007 data for the first time detected a significant downturn in incidence and mortality rates in US women. Far more men than women still die of lung cancer each year, but the gender gap in lung cancer mortality is steadily narrowing and is expected to close.

2.4 Race and Ethnicity

2.4.1 African Americans: The patterns of occurrence of lung cancer by race and ethnicity make lung cancer a relevant disease for cancer disparities research. Lung cancer incidence rates are similar among African American and white women, but rates are about 47% higher among African American men than among white men. African American men have also experienced a greater mortality from lung cancer, with the largest disparity in rates being 42% greater than for European American men in 1990; the excess decreased to 25% in 2008. Clues to birth cohort changes in disease occurrence may be found in changes in rates in young people. An analysis of trends in lung cancer incidence and mortality rates (per 100,000) from 1992 to 2006 among 20- to 39-year-olds revealed a similar narrowing of the racial gap in this age group, leading to the inference that the drop among African Americans resulted from the striking decrease in smoking prevalence among African American youth since the 1970s. If this inference is correct, continued narrowing of the racial disparity can be anticipated in the coming decades.

This racial disparity may be partially due to a greater susceptibility of African American smokers to smoking-induced lung carcinogenesis, but historical differences in smoking prevalence do not explain all of the higher risks seen in African Americans compared with European Americans. The racial disparity in mortality reflects not only the differences in incidence but also 20% poorer 5-year relative survival among African Americans compared with whites. The poorer lung cancer survival in African Americans remains to be explained, but in multivariate analyses, the racial disparity is diminished by adjustment for receipt of evidence-based therapy and pretreatment health status. Thus, potential contributors to the racial disparity in survival could be treatment-related factors such as later stage at diagnosis and lack of access to, or uptake of, evidence-based stage-specific lung cancer treatment.

Compared with African Americans (72.7) and whites (63.3), age-adjusted incidence rates per 100,000 for the years 2004 to 2008 were significantly lower among American Indians/Alaskan Natives (44.5), Asians/Pacific Islanders (39.0), and Hispanics (32.5). Similar patterns are observed for lung cancer mortality rates.

2.4.2 Asians: Among patients with lung cancer, those of Asian ancestry have consistently been observed to have better survival than whites. The reasons for the more favorable prognosis in Asians are incompletely understood, but one contributory factor is differences in tumor characteristics. For example, in Asians, the prevalence of epidermal growth factor receptor mutations in lung tumors is much higher than in whites, and epidermal growth factor receptor-positive tumors are responsive to treatment with gefitinib. Further delineating the distinct features of the etiology and prognosis of lung cancer in Asians compared with other ethnic groups may lead to novel insights into lung cancer pathogenesis.

2.5 Socioeconomic Status

Increasingly, lung cancer is more likely to occur in poorer and less-educated populations, primarily reflecting the increasing gradient of smoking with socioeconomic indicators that include income, education, and occupation. This pattern, noted decades ago in the United States, has now been observed in many countries worldwide. For example, in Canada, the risk of lung cancer was inversely associated with income, education, and social class, and despite universal health care, lower socioeconomic status was significantly associated with poorer lung cancer survival. In China, a sixfold variation in lung cancer risk was observed between the lowest and highest income categories. In The Netherlands, lung cancer risk was inversely associated with attained education. Lower socioeconomic status has also been associated with later stage at diagnosis for lung cancer as for other cancers. In the United States, studies of lung cancer prognosis that have examined both race/ethnicity and socioeconomic status have shown lower socioeconomic status to be a strong determinant of worse prognosis, whereas racial differences in prognosis tend to diminish when adjusted for socioeconomic status. Socioeconomic status is associated with an unfavorable profile of interacting determinants of lung cancer risk, such as smoking, diet, and exposure to inhaled carcinogens in the workplace and general environment.

2.6 Geographic Patterns

Internationally, lung cancer rates vary markedly across countries: Age-standardized incidence rates vary > 60-fold in both men and women. The geographic distribution is predominantly driven by historical patterns in cigarette smoking prevalence, with
an approximately 20-year lag period from change in smoking pattern to change in incidence, reflecting the slow and multistep process of cancer initiation and progression.\textsuperscript{29}

In men, the highest annual lung cancer incidence rates are in central and eastern Europe and North America (65.7 and 61.2 per 100,000, respectively). In women, the lung cancer incidence rates are highest in North America and northern Europe (35.6 and 21.3 per 100,000, respectively).\textsuperscript{30} For both sexes, the lowest incidence rates are in Africa. These patterns are fluid because lung cancer rates will change commensurate with changes in smoking prevalence.

The situation in China is both unique and of particular concern. Contrary to elsewhere, the high lung cancer mortality rates among Chinese women are not due to a high prevalence of cigarette smoking. Rather, the high rates appear to be a result of exposure to other risk factors that include indoor air pollution from cooking fumes.\textsuperscript{2} Chinese men are a high-risk population of particular concern because of a striking increase in their smoking rates. Per capita cigarette consumption in Chinese men increased from one cigarette per day in 1952, to four in 1972, to 10 in 1992.\textsuperscript{31} As a consequence, the lung cancer incidence rates have already increased and will continue to rise substantially. The increase in lung cancer among Chinese men will have a major impact on the global burden of lung cancer in the 21st century, given the size of this group of smokers.

The tobacco addiction epidemic in China exemplifies a shift in the global burden of lung cancer from high-income western countries to low- and middle-income countries, particularly in Asia. In 2008, newly diagnosed lung cancers in developing countries (884,500) exceeded the number in developed countries (724,300) by 22%.\textsuperscript{2} The trend of the lung cancer burden becoming increasingly concentrated in the developing world is expected to continue for the foreseeable future.

Substantial geographic variation in lung cancer mortality rates is also present within countries. For example, in the United States from 2004 to 2008, the age-adjusted lung cancer incidence rates varied 3.6-fold between the states with the highest (Kentucky, 101 per 100,000) and the lowest (Utah, 28 per 100,000) rates.\textsuperscript{10}

### 3.0 The Etiology of Lung Cancer: Overview

The etiology of lung cancer can be conceptualized as reflecting the joint consequences of the interrelationship between (1) exposure to etiologic agents and (2) individual susceptibility to these agents. Synergistic interactions among risk factors can have substantial consequences for lung cancer risk. Well-known examples include the synergistic effect of cigarette smoking on the lung cancer risk associated with asbestos exposure and radon.\textsuperscript{32}

Given the many known risk factors for lung cancer, a practical question for guiding prevention is the relative contribution of these factors to the overall burden of lung cancer. The population attributable risk approach takes into account the magnitude of the relative risk associated with an exposure along with the likelihood of exposure in the general population.\textsuperscript{33} These attributable risk estimates include joint contributions of risk factors that sometimes have synergistic relationships. For example, the attributable risk estimate for cigarette smoking includes the lung cancer risk attributed to the independent effects of cigarette smoking and further includes the risk of lung cancer from smoking because of its synergistic interactions with factors such as asbestos and radon. For this reason, the total percentage can exceed 100%. Lung cancer has a well-characterized set of important risk factors and established synergistic interactions between risk factors, and these reasons contribute to the attributable risks summing to considerably > 100%. For example, population attributable risk estimates for lung cancer indicate that in the United States, active smoking is responsible for 90% of lung cancer;\textsuperscript{3} and radon is responsible for 15%.\textsuperscript{3}

### 3.1 Environmental and Occupational Agents

#### 3.1.1 Tobacco Smoking

Tobacco smoking: A single etiologic agent—cigarette smoking—is by far the leading cause of lung cancer, accounting for about 80% to 90% of lung cancer cases in the United States and other countries where cigarette smoking is common.\textsuperscript{24} Compared with never smokers, US smokers who have not quit successfully have about a 20-fold increase in lung cancer risk. Few exposures to environmental agents convey such risks for any disease. In general, spatial and temporal trends of lung cancer occurrence closely reflect patterns of smoking, but rates of occurrence lag smoking rates by about 20 years. Prior versions of this review\textsuperscript{35,36} covered smoking and lung cancer extensively, so only a summary of this voluminous literature is provided here. Lung cancer occurs in four major types as classified by light microscopy: adenocarcinoma, squamous cell carcinoma, large cell carcinoma, and small cell carcinoma. All four types are caused by cigarette smoking.\textsuperscript{37} The histologic characteristics of lung cancer in developed countries have changed during the past 50 years. Adenocarcinoma has become more common, whereas squamous cell carcinoma has declined. This shift is notable because adenocarcinoma tends to arise more peripherally and squamous cell carcinoma more centrally.\textsuperscript{38} The most likely explanation for the rise in adenocarcinoma is
the changing cigarette, leading to changes in smoking topography that has included greater depth of inhalation.

Cigar smoking is also an established cause of lung cancer. The lung cancer risks associated with cigar smoking are substantial but less than the risks observed for cigarette smoking because of differences in smoking frequency and depth of inhalation. The same pattern holds true for pipe smoking. Bidi, loose tobacco rolled in a leaf, is the most commonly smoked tobacco product in India; bidi smoking also causes lung cancer. With respect to smoking of nontobacco products, despite the plausibility of marijuana as a risk factor for lung cancer, the evidence to date has not documented an association after adjusting for tobacco smoking.

3.1.1.2 The Changing Cigarette—The composition of cigarettes has evolved considerably since the 1950s. The marketplace has shifted from mainly unfiltered cigarettes to predominantly filtered cigarettes. In the mid-1960s, ventilation holes were added to the filter, which dilute the smoke with air drawn through them. However, smokers can easily block the holes with their fingers, which are left unblocked by the machines used to test cigarettes. Reconstituted tobacco has been used increasingly since the 1960s, there have been changes to the cigarette paper and additives used, and most cigarettes are more ammoniated in the United States. A concomitant shift toward lowered levels of tar and nicotine, as measured by a smoking machine, has occurred. Cigarette tar refers to the condensable residue of cigarette smoke, that is, the total particulate matter of cigarette smoke deposited on the machine’s filter less the moisture and nicotine. Tar is a complex mixture that includes many carcinogens.

Studies show little relation between biomarkers of cigarette smoke and cigarette tar or nicotine yield as measured by Federal Trade Commission (FTC) protocol. These studies have been conducted in both the population context and the controlled laboratory setting. The lack of association of tar and nicotine yields with biomarker levels partially reflects compensatory changes in smoking patterns for smokers switching from higher to lower yield products. The compensation includes blocking the ventilation holes, more frequent and deeper puffs, and an increase in the number of cigarettes smoked.

The gradual reduction in machine-measured tar yield over recent decades would be expected to have reduced smokers’ exposures to carcinogens if the FTC test protocol were predictive of carcinogen doses delivered to the lung. However, substantial evidence indicates that the FTC test method is not informative with regard to lung cancer risk or risks of smoking-caused diseases more generally. For lung cancer and other diseases, three lines of epidemiologic data have been available on changes in products. The first comes from case-control studies that compared the smoking history profiles of persons with lung cancer with those of control subjects. The second comes from cohort studies that tracked the risk of lung cancer over time as the products smoked changed. The third comes from assessment of the temporal changes in age-specific patterns of lung cancer mortality rates compared with changes in cigarette characteristics. These lines of evidence are convergent, and national as well as international groups that have evaluated the evidence concluded that changes in yield over time have not reduced lung cancer risk in smokers; conversely, the cigarette changes may have actually increased lung cancer risk. Under the 2009 Family Smoking Prevention and Tobacco Control Act, cigarette packages will no longer provide machine-measured yields.

3.1.1.3 Menthol—Menthol is a flavoring agent that can be either derived naturally or synthesized in the laboratory. Menthol cigarettes were invented in the 1920s and currently comprise approximately one-third of the US cigarette market. The prevalence of menthol cigarette smoking is by far the highest among African Americans, reflecting patterns of aggressive and targeted marketing that began in the 1960s. Use of menthol cigarettes is increasing among adolescents, an increase that is greater among minorities.

Menthol acts on receptors expressed primarily on sensory nerves in the nose, mouth, and airway to produce a minty taste and aroma. It has cooling, counterirritant, and analgesic properties. These sensory actions of menthol in cigarette smoke have raised concern that it facilitates experimentation and initiation of regular smoking and alters smoking topography in ways that increase doses of exposure to tobacco smoke toxins. These effects of menthol,
combined with the fact that African Americans have disproportionately high lung cancer rates and a very high prevalence of menthol cigarette use, are compatible with the hypothesis that menthol cigarettes may be even more strongly associated with lung cancer risk than nonmenthol cigarettes.

However, two relevant lines of evidence involve comparing smokers of menthol and nonmenthol cigarettes with respect to (1) biomarkers and (2) lung cancer risk. Comparisons of menthol and nonmenthol cigarette smokers with respect to biomarkers of tobacco smoke exposure and dose have not revealed consistent differences, including concentration of nicotine and tobacco-specific nitrosamines.65-69 The results of numerous case-control70-74 and cohort75,76 studies have consistently reinforced the conclusion that menthol and nonmenthol cigarettes are associated with nearly equivalent risks of lung cancer. These lines of evidence were recently summarized by the Tobacco Products Scientific Advisory Committee (TPSAC) of the Food and Drug Administration. Under the 2009 Family Smoking Prevention and Tobacco Control Act, TPSAC was required to develop a report and recommendations to address the impact of menthol in cigarettes on public health, including specifically considering use among children, African Americans, Hispanics, and other racial and ethnic minorities. The TPSAC report concluded that the totality of the evidence did not support the hypothesis that smoking menthol cigarettes was associated with a greater risk of lung cancer than smoking nonmenthol cigarettes.77 Studies subsequent to the TPSAC report provided evidence to support TPSAC’s conclusion; notable among these was a study nested with a large cohort of racially diverse adults in which a lower lung cancer incidence was noted in menthol vs nonmenthol smokers.78

However, menthol cigarettes may still have adverse public health consequences. Several studies suggested that menthol cigarettes are a starter product that may be associated with smoking initiation.59,79 The evidence is conflicting with regard to the effects of menthol on dependence in adult smokers but more clearly points to menthol cigarettes being associated with greater nicotine dependence among adolescents.77 Furthermore, among African Americans, smokers of menthol cigarettes may have a more difficult time successfully quitting smoking than smokers of nonmenthol cigarettes.80-82 The TPSAC concluded on the basis of this evidence that menthol cigarettes are detrimental to public health by increasing the number of smokers and the duration of smoking, resulting in increased smoking prevalence.

3.1.1.4 Secondhand Smoke Exposure—Passive smokers inhale a complex mixture of smoke widely referred to as secondhand smoke. The 2006 US Surgeon General’s report reinforced earlier conclusions that secondhand smoke exposure is a cause of lung cancer among nonsmokers.4 This association holds true regardless of the source of exposure to secondhand smoke, but the most abundant evidence is for nonsmokers who live with a smoker, which is associated with a 20% to 30% increased risk of lung cancer.4 IARC has classified secondhand tobacco smoke exposure as a known human (class A) carcinogen.52 Secondhand smoking is estimated to cause 3,000 lung cancer deaths per year in the United States53 and 21,400 deaths per year globally.54

3.1.1.5 Are Women More Susceptible to Smoking-Induced Lung Cancer? —Results of some studies have suggested a potentially higher risk of smoking-associated lung cancer in women compared with men,85-87 but methodologic issues cloud the interpretation of these studies, particularly due to a lack of focus on the most informative comparisons.88 Furthermore, the evidence from prospective cohort studies fails to support the notion of a sex differential in susceptibility to lung cancer from smoking.89 The equal rates of lung cancer mortality in younger US men and women corresponding to a time of equal smoking prevalence also provides evidence against an important gender difference in susceptibility to smoking-induced lung cancer.90 At present, the evidence does not favor the hypothesis because for a specific degree of smoking history, the relative risk estimates for men and women are very similar.89

Suspicion is naturally cast on a potential hormonal role for any hypothesized gender difference in disease susceptibility. A meta-analysis of two large-scale randomized controlled trials of hormonal therapy with an estrogen plus progestin formulation found a significantly increased risk of lung cancer (relative risk, 1.4; 95% CI, 1.03-1.8), pointing to a potential hormonal role in the etiology of lung cancer.91 Reports published since the meta-analysis further suggested that the lung cancer risk associated with hormonal therapy may be specific to estrogen plus progestin formulations. In the Women’s Health Initiative trial, results were null for estrogen-only formulations.92 In a prospective cohort study, increased lung cancer risk was observed for estrogen plus progestin formulations, with null results for estrogen-only formulations.93

3.1.1.6 Risk Prediction—Models for predicting risk will be needed to guide lung cancer screening with CT scanning or other interventions targeting high-risk smokers. Given the dominance of cigarette smoking as a risk factor for lung cancer in the general population, many analyses have explored the relationships between quantitative measures of smoking and lung cancer risk. These analyses have shown the importance of smoking duration; number of cigarettes smoked; and, for former smokers, the time...
since quitting. A landmark analysis of the data from a cohort study of British physicians showed that duration of smoking and number of cigarettes smoked have quantitatively distinct effects on lung cancer risk and that they should not be combined for estimating lung cancer risk.\textsuperscript{94} Further information on the quantitative relationships of measures of smoking with lung cancer risk can be found in the comprehensive review presented in the 2004 IARC monograph and the 2004 report of the US Surgeon General.\textsuperscript{3,52}

More recently, prediction models have been developed to estimate the probability of lung cancer occurring during specified time intervals. Such models are of particular interest for risk stratification to identify candidates for screening, given the recent demonstration of reduction in lung cancer mortality by low-dose CT scanning in the National Lung Cancer Screening Trial (NLST).\textsuperscript{49} Prediction models parallel the approach taken for breast cancer, wherein a risk factor-based model, the Gail model, has long been available.\textsuperscript{96} This model’s predictions are used for a variety of purposes, including informing patients of their potential risk, guiding screening, and selecting women for clinical trials.

Three prediction models for lung cancer are now available: the Bach model, which is based on data from the CARET (\textbeta{}-Carotene and Retinol Efficacy Trial); the Spitz model, which is based on data from an ongoing case-control study; and the Liverpool Lung Project, which is based on a case-control study in Liverpool, England.\textsuperscript{97-99} The general approach used to develop the models is similar. The epidemiologic data are analyzed to identify risk factors and estimate the associated relative risk. The relative risk estimates are then used to project risk over time. Various statistical approaches are used to identify the most informative variables and to validate the final model. The measure of prediction used generally is the area under the curve (AUC) from receiver operating characteristic (ROC) analysis; the AUC varies from 0 to 1.0, that is, from no predictive value to perfect prediction.

Etzel and Bach\textsuperscript{100} provided a useful comparison of the three models. The AUC values for the models range from about 0.60 to 0.70. D’Amelio and colleagues\textsuperscript{101} compared the performance of the three models in an independent data set, a case-control study carried out in Boston, Massachusetts. The AUC values for the Spitz and Liverpool Lung Project models were 0.69, whereas that for the Bach model was 0.66. Spitz and colleagues\textsuperscript{102} refined the original model in two ways: (1) by adding two markers of DNA repair capacity with a slight gain in AUC\textsuperscript{98} and (2) by developing a model specifically for African Americans, which performed better than the model based on whites, in this racial group.\textsuperscript{102}

Undoubtedly, these models will continue to be refined. Their predictive power may be greatly enhanced if additional genetic markers are identified to determine susceptibility to lung cancer in smokers. For now, they provide a tool for risk stratification but lack the sensitivity and specificity needed for managing individual patients.

3.1.2 Never Smokers: Tobacco smoking causes such a large proportion of all lung cancer cases that there have been few data on the occurrence of lung cancer among never smokers. Global estimates indicate that about 300,000 lung cancer deaths annually are not due to tobacco use.\textsuperscript{103} Even though this estimate represents a minority of the lung cancer burden, the incidence of lung cancer in never smokers ranges from 4.8 to 20.8 per 100,000 among individuals aged 40 to 79 years,\textsuperscript{104} comparable to the rates of myeloma in men and cervical cancer in women.

An analysis of data from 35 cohort and cancer registry studies from around the world revealed the absence of temporal trends of lung cancer in never smokers but indicated that among never smokers, lung cancer death rates are greater in men than in women and greater in African Americans and in Asians living in Asia compared with those of European ancestry.\textsuperscript{105} Among patients with lung cancer who never smoked, the known causes of lung cancer other than cigarette smoking, such as exposure to second-hand smoke, radon, and occupational carcinogens, are present in a substantial\textsuperscript{105} proportion. Conversely, a large fraction of cases were not attributable to these factors.\textsuperscript{106}

3.1.3 Diet and Physical Activity: Lifestyle factors other than cigarette smoking, such as diet and exercise, have been extensively investigated for a potential role in influencing lung cancer risk.

3.1.3.1 Diet—The most thoroughly investigated dietary factors are also those that appear to have the greatest implications for prevention: fruits, vegetables, and specific antioxidant micronutrients that are commonly found in fruits and vegetables. This section relies primarily on evidence ratings from a systematic review of the world’s evidence summarized in the 2007 report of the WCRF.\textsuperscript{107} This rating scale included categories of convincing, probable, and limited-suggestive evidence of an association between a dietary factor and lung cancer.

With a WCRF evidence rating of probable, the evidence points toward greater levels of fruit consumption being inversely associated with lung cancer risk. Similar to fruit, the evidence suggests that vegetable consumption is inversely associated with lung cancer risk, but the results have been less consistent and weaker than for fruit. Hence, the overall evidence
for vegetables was rated as limited-suggestive in the WCRF report.

To better understand the basis of these inverse associations, fruits and vegetables have been grouped into classes and studied in relation to lung cancer risk. For example, cruciferous vegetables have been associated with a reduced risk of lung cancer in a number of studies. The association with cruciferous vegetable intake has persisted even after careful control for cigarette smoking in the design of some studies. The evidence of an inverse association between cruciferous vegetable intake and lung cancer risk has bolstered interest in isothiocyanates as a promising chemopreventive agent. Isothiocyanates, metabolites of the class of phytochemicals known as glucosinolates, could exert anticancer effects by blocking carcinogens through induction of phase 2 detoxification enzymes, such as glutathione S-transferase. Lung cancer risk has also been consistently inversely associated with higher dietary intakes or urinary levels of isothiocyanates, a constituent of cruciferous vegetables.

Fruits and vegetables are the major dietary source of antioxidant micronutrients, such as carotenoids. The example of carotenoids exemplifies the complexities involved in attempting to determine the role of diet in the etiology of lung cancer. Prospective studies of both dietary intake and prediagnostic blood concentrations suggest an inverse association between carotenoids and lung cancer. For example, both dietary intake and circulating concentrations of total carotenoids were associated with a 20% to 30% lower risk of lung cancer in the highest vs lowest exposure categories. On the basis of these data, the WCRF rates foods containing carotenoids as probable protective factors for lung cancer. However, it cannot be determined with certainty whether the inverse association between carotenoids and lung cancer is directly due to carotenoid intake or whether carotenoid intake merely serves as a marker of the intake of other protective substances or healthier dietary habits in general. This uncertainty is heightened by the results of large-scale randomized controlled trials that conclusively demonstrated that high-dose β-carotene consumption is associated with an increased risk of lung cancer in smokers.

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 meta-analyses suggested that alcohol drinking in the highest consumption categories is associated with an increased risk of lung cancer, but a study of alcohol drinking in never smokers was null. Anthropometric measures have also been studied, indicating a tendency for persons with a lower BMI to have increased lung cancer risk relative to heavier people. However, effects of both alcohol drinking and low BMI may be difficult to separate from the comitant effects of smoking because those who smoke more cigarettes per day tend to be leaner and to drink more. At present, when considering the possible relationships between lung cancer and factors such as alcohol drinking and lower BMI, uncontrolled confounding by cigarette smoking cannot be dismissed as a possible explanation.

The overwhelming contribution of cigarette smoking in smokers as a cause of lung cancer poses a challenge to detecting the role that lifestyle factors, such as diet, may play in the etiology of lung cancer. Cigarette smoking is now so closely associated with unhealthy lifestyles in the United States and some other countries that it is difficult to disentangle the dietary factors of interest from the effects of smoking. Cigarette smoke can directly affect circulating concentrations of dietary factors; for example, smokers tend to have lower circulating concentrations of antioxidant micronutrients even after accounting for differences in dietary intake. Additionally, associations between dietary factors and lung cancer risk are likely to be far weaker than the association with active smoking, and in general, diet is measured with much greater error than smoking. Even for a dietary factor such as fruit consumption that is inversely associated with lower lung cancer risk, the highest exposure category is typically associated with at most a halving in the risk of lung cancer. Thus, in interpreting evidence on diet and lung cancer risk, residual confounding cannot be readily set aside as an explanation for the observed associations between dietary factors and lung cancer.

In summary, observational evidence suggests that fruit consumption, and to a lesser extent vegetable consumption, are inversely associated with lung cancer risk. The specific constituents of fruits and vegetables that may confer protection are not known.

3.1.3.2 Physical Activity—A meta-analysis of leisure activity observed that both moderate and high levels of physical activity are associated with a 13% to 30% decrease in lung cancer risk, consistent with the WCRF rating of the evidence as probable that physical activity is inversely associated with lung cancer risk. However, the specific biologic mechanism whereby physical activity could reduce lung cancer risk remains unknown, and as with associations of lung cancer with lifestyle factors that are associated with tobacco dependence, potential residual confounding by cigarette smoking needs to be considered as an alternative explanation.

3.1.4 Environmental Exposures—

3.1.4.1 Occupational Exposures—Investigations of occupational groups who often incur heavy, sustained exposure to workplace agents, particularly in the past

Epidemiology of Lung Cancer
in the United States and other western countries, provide evidence on the carcinogenicity of a number of chemicals and physical agents. Lung cancer is the most common among cancers associated with occupational exposures. Estimates of the proportion of lung cancer caused by occupational exposures through independent or shared causal pathways range widely because of differences between populations in the prevalence of exposure and associated risk, but a figure of around 10% is a reasonable average estimate. Thus, in industrialized nations, the contribution of occupational exposures to the lung cancer burden is small compared with that of cigarette smoking but large compared with most other exposure classes. Cigarette smoking potentiates the effect of some known occupational lung carcinogens.

Lung cancer is causally associated with many workplace exposures. Workers exposed to tar and soot (which contains benzo[a]pyrene), such as coke oven workers, in concentrations exceeding those present in urban air are at increased risk of lung cancer. Occupational exposures to a number of metals, including arsenic, chromium, and nickel, are also causes of lung cancer. These carcinogens can even increase lung cancer risk when the route of exposure is ingestion rather than inhalation; for example, elevated concentrations of arsenic in drinking water are associated with increased lung cancer risk.

For some other workplace agents, the evidence is less clear. The results of numerous case-control and cohort studies are compatible with a weak association between diesel exhaust exposure and lung cancer risk. For example, in a pooled analysis of data from 11 case-control studies, a significant dose-response trend in risk was observed, with an OR of 1.3 (95% CI, 1.2-1.4) between the highest and lowest exposure categories. This is a public health concern because of exposure to diesel exhaust in urban areas and the increasing use of diesel vehicles in some European countries; diesel exhaust particles are known to contain carcinogens.

The question of whether silica dust is a risk factor for lung cancer is controversial. A meta-analysis associated silicosis with a 2.4-fold increase in lung cancer mortality (95% CI, 1.6-3.7) when limited to studies that adjusted for cigarette smoking, this association was weaker but still significant (standardized mortality ratio, 1.6; 95% CI, 1.3-1.9). The evidence on silica exposure, absent consideration of the presence of silicosis, is less clear. In another meta-analysis, a significant overall association between silica dust exposure and lung cancer in people without silicosis (relative risk, 1.2; 95% CI, 1.1-1.3) was observed, but in analyses limited to studies that adjusted for smoking, the overall results were null (relative risk, 1.0; 95% CI, 0.8-1.3). Still another meta-analysis that focused on the dose-response association found that lung cancer risk increased with increased silica exposure. In 1997, the IARC classified crystalline silica as a human carcinogen, however, some still continue to question its carcinogenicity and the role of silica exposure vs that of fibrosis in people with silicosis.

3.1.4.2 Asbestos—Asbestos, a well-established occupational carcinogen, refers to several forms of fibrous, naturally occurring silicate minerals. The association of occupational asbestos exposure and lung cancer is strong, often greater than a fivefold excess risk. Whether asbestos acts as a direct carcinogen or through indirect mechanisms such as chronic inflammation that promotes cancer development remains uncertain. Asbestos and cigarette smoking are both independent causes of lung cancer, but in combination, they act synergistically to markedly increase lung cancer risk.

3.1.4.3 Radiation—Studies of populations exposed to high doses of radiation have found lung cancer to be one of the cancers caused by exposure to ionizing radiation. The following two types of radiation, classified by rate of energy transfer to the tissue, are relevant to lung cancer: (1) low linear energy transfer (LET) radiation (eg, x-rays, γ-rays) and (2) high-LET radiation (eg, neutrons, radon). High-LET radiation produces ionization of relatively higher density in tissues than low-LET radiation, so in equivalent doses, more biologic damage is produced by high-LET than low-LET radiation.

For both types of radiation, epidemiologic evidence comes from cohorts exposed at levels substantially greater than those experienced by the general population. For radon, case-control studies have directly estimated the risk associated with indoor concentrations in homes. Risk assessment methods have been used to estimate the population risks.

3.1.4.3.1 High-LET Radiation: Radon: Radon is an inert gas that is produced naturally from radium in the decay series of uranium. Two of the decay products of radon emit α particles that by virtue of their high energy and mass, can cause damage to the DNA of cells of the respiratory epithelium. Studies of underground miners of uranium and other ores have established exposure to radon daughters as a cause of lung cancer. Cigarette smoking and radon decay products synergistically influence lung cancer risk.

Radon is of broader societal interest because it is a ubiquitous indoor air pollutant, entering buildings in the form of soil-derived gas. Average exposures to indoor radon for the general population are much less than received by occupational groups such as uranium miners, but direct evidence from case-control studies has documented that exposure to
Epidemiology of Lung Cancer

3.1.4.4.1 Air Pollution—

3.1.4.4.1. Atmospheric Air Pollution: Outdoor air can contain a number of hazardous agents, many of which are generated by the combustion of fossil fuels. Carcinogens generated by combustion of fossil fuels include polycyclic aromatic hydrocarbons and metals such as arsenic, nickel, and chromium. In considering respiratory carcinogenesis, the constituents of air pollution will vary by locale and over time depending on the pollution sources.

Particulate matter in urban air, which has multiple sources, has been studied as a potential lung cancer risk factor. A study of six US cities found a nonsignificant 40% increase in risk of lung cancer in urban environments with the highest concentration of fine particles compared with the city with the lowest concentration. The data from the American Cancer Society Cancer Prevention Study II showed that each 10 g/m3 increase in concentration of fine particles carried an increased lung cancer risk of 14%. The evidence continues to coalesce around an association between constituents of ambient air pollution and increased lung cancer mortality, with reports from Japan, China, and New Zealand, documenting increased risks with measures of particulate matter, sulfur dioxide, and nitrogen dioxide. In a study of nitrogen oxides as a measure of air pollution specific to traffic, a significantly elevated risk of lung cancer was noted, but associations have not been observed in all traffic-specific air pollution studies.

3.1.4.4.2 Indoor Air Pollution: Indoor air pollution has large potential health implications for people who spend substantial amounts of time indoors. Indoor air pollution may come from incoming outdoor air or originate indoors, for example, tobacco smoking, soil gases, and combustion products from heating and cooking. In developed countries, two indoor pollutants that most strongly increase lung cancer risk in never smokers are passive smoking and radon. Studies that were strongest for coal users in Asia (OR, 4.9; 95% CI, 3.7-6.5) and weaker, but still statistically significant for wood users in Europe and North America (OR, 1.2; 95% CI, 1.1-1.4). To illustrate the extent of risk that solid fuel use poses, estimates in China indicate that halving solid fuel use by 2033 would avert about 600,000 lung cancer deaths between 2003 and 2033.
susceptibility to developing lung cancer. For example, a family history of lung cancer is strongly associated with increased risk of lung cancer. A meta-analysis of 41 published cohort and case-control studies found that having a positive family history of lung cancer was associated with a 1.7-fold increased risk of lung cancer (95% CI, 1.6-1.9), an association that was only slightly weaker among nonsmokers (OR, 1.4; 95% CI, 1.2-1.7). A positive family history of lung cancer in two or more relatives was associated with substantially greater risk (OR, 3.6; 95% CI, 1.6-8.3). These associations were specific to a family history of lung cancer in that a family history of any cancer or a family history of other smoking-associated cancers was not significantly associated with lung cancer risk. A positive family history of lung cancer is thus a clinically useful risk indicator.

The specific host factors selected for review here are biomarkers of risk and early detection, acquired lung disease, and infections. These areas of active research were considered as needing updated coverage in this review.

4.1 Biomarkers of Risk and Early Detection

Advances in understanding lung carcinogenesis and integration of a transdisciplinary approach hold promise for determining who is at risk for the disease and for elucidating methods of early detection. The ability to identify accurately at the molecular level individuals at high risk for lung cancer would offer tremendous public health benefit.

The lung is a visceral organ; thus, access is limited, posing barriers to the sensitivity of molecular lung cancer early detection modalities. Even sophisticated modalities such as external imaging, endobronchial fiberoptics, and long probes cannot thoroughly examine the entire lung because of the complex branching structures of the bronchial tree, rendering portions of the transforming epithelium unstudied. An alternative is to sample surrogate tissues. This section summarizes lung tissue-based biomarkers before going on to review lung cancer biomarkers derived from the blood, airway, or exhaled breath. Emphasis is placed on less invasive, molecular-oriented approaches to risk assessment and early detection.

4.1.1 Tissue-Based Molecular Assays of Carcinogenesis and Risk: The most definitive approach to detecting lung cancer is by directly examining lung tissue, either conventionally (eg, microscopy, immunohistochemistry) or by newer molecular methods. DNA-based genome-wide searches of tumors have identified somatic copy number aberrations (CNAs) on the basis of single-nucleotide polymorphism (SNP) or comparative genomic hybridization arrays, indicating that these alterations cluster in unknown genes. Premalignant squamous dysplastic lesions harboring CNAs at 3p26.3-p11.1, 3q26.2-29, and 6p25.3-24.3 predicted with 97% accuracy the later occurrence of squamous cell cancer at the same site. The 3q CNA loci most strongly associated with preinvasive nonsmall cell lung cancer (NSCLC) progression identified SOX2 as a lead candidate gene. Early lung cancers detected by CT screening already have a high prevalence of aneuploidy, with areas of common CNAs in 5p and 8p. In small cell carcinoma, CNAs are present in the focal adhesion pathways. Ethnicity-specific differences may affect tumor CNAs. In lung adenocarcinoma, losses in 16p13.13 and 16p13.11 were seen in patients of European ancestry, whereas losses in 19p13.3 and 19p13.11 predominated in those of East Asian ancestry.

4.1.1.1 DNA Methylation Markers—CpG methylation of gene promoter DNA is important in regulating genes, in turn regulating biochemical pathways. Progressive DNA methylation, especially in tumor suppressor genes, plays a central role in lung carcinogenesis. Comprehensive genome-wide searches for methylation-silenced genes, checked for consistency with expression, identified BNC1, MSX1, CCNA1, RASSFIA, p16, ALDH1A1, LOX, and CTSZ as possibly distinguishing malignant from adjacent nonmalignant lung tissue. The candidate genes that had significantly elevated methylation in atypical adenomatous hyperplasia (CDKN2A exon 2, PTPRN2) differed from genes with significant hypermethylation in lung adenocarcinoma in situ (2C35, EYA4, HOXA1, HOXA11, NEUROD1, NEUROD2, and TMED2). In contrast, promoter hypermethylation at CDH13, CDX2, OPCML, RASSF1, SFRP1, and Twist1 and global DNA hypomethylation appear to be present predominantly in invasive cancer. DNA methylation is associated with patient factors such as age, sex, smoking, and alcohol consumption and, thus, is being evaluated as a lung cancer biomarker.

4.1.1.2 RNA Transcriptome in Lung Tumors—Development of coding RNA-based transcriptome biomarkers in lung carcinogenesis uses high-throughput sequencing of complementary DNA to derive genetic differences between tumor vs adjacent nontumor tissue. Technological and procedural limitations have made mRNA findings difficult to reproduce.
Technologies are just now emerging for noncoding RNA signatures. However, there is a noncoding RNA/microRNA signal unique to precursor bronchial lesions for squamous cell carcinoma.

4.1.1.3 Proteomics and Metabolomics—Proteome-wide technologies, such as matrix-assisted laser desorption/ionization mass spectrometry (MS), have yielded validated biomarkers of identifiable proteins with promising discriminative capacity. Metabolomics uses high-throughput assays to generate spectroscopic signatures representative of the metabolic states of a cell; this technology is in its nascent stages. Preliminary data suggest that spectroscopic signals may be able to distinguish lung squamous cell and adenocarcinomas from nonmalignant tissue.

4.1.2 Noninvasive Lung Surrogates of Carcinogenesis and Risk: Noninvasive biomarkers of risk are the optimal population-level approach. Noninvasively obtained tissues in asymptomatic populations include exfoliated cells in sputum or the circulating macromolecules of blood. For these, commonly used techniques are polymerase chain reaction (PCR) assays of DNA and RNA, and MS-based methods. PCR assays have excellent sensitivity. The nucleic acid end points confer greater sensitivity and are considered more biologically proximate end points as opposed to direct measurements of tumor growth, invasion, and metastasis. Proteins and metabolites, direct underpinnings of cancer phenotypes, are also detectable using proteomic and metabolomic assays, respectively.

4.1.2.1 Blood-Based Markers—

4.1.2.1.1 DNA-Based Markers: Germline Sequence Variation: Blood is an attractive surrogate tissue because it is readily available and because metabolic processes presumably pool in the circulation. A study of high-risk family pedigrees reported a susceptibility locus on chromosome arm 6q23-25 but the specific genetic variants responsible for the excess risk have yet to be identified.

Blood-based epidemiologic studies of germline variants in relation to lung cancer risk began by studying individual genetic variants (SNPs) in candidate genes encoding for enzymes in specific biochemical pathways. These were investigated in population-based studies with emphasis on carcinogen metabolism, DNA repair, and inflammation pathways. In general, the results of candidate gene SNP association studies have been inconsistent across populations. More recently, genome-wide association studies have been carried out, and the results of at least four separate studies are consistent in identifying a region on the long arm of chromosome 15 (15q24-25.1). For example, those with at least one variant allele of a specific SNP in this region (rs8034191) in the acetylcholine receptor gene, a variant that was present in approximately one-third of the study population, were at 1.3 times greater risk of lung cancer than those with the wild-type allele.

The quantity of DNA leaching into the blood itself has been proposed as a marker, albeit controversially. Many DNA-based somatic markers in the blood have been tested (summarized in Brennes et al), including oncogene mutations.

DNA adducts of known tobacco carcinogens in blood cells are relevant to lung cancer detection. Although bulky DNA adducts may be weakly associated with lung cancer risk, postlabeling of adducts with 32P does not consistently show a relation to smoking or to lung cancer risk. Compared with DNA adducts, germline polymorphisms in carcinogen metabolism, DNA repair, and related pathway genes may be more predictive blood-based biomarkers of lung cancer.

4.1.2.2 DNA Methylation—Aberrant DNA methylation in blood is detectable but is of unknown origin and has been ineffective as a diagnostic tool. Recently, however, high sensitivity and specificity were reported for methylation-specific PCR in the blood, with evidence of association with the corresponding primary lung tumor. In heavy smokers with abnormal radiographic findings, a four-gene signature (DCC, Kif1a, NISCH, and Rarb) in the blood discriminated between cancer and noncancer with 73% sensitivity and 70% specificity. For small cell carcinoma, a nine-CpG-site DNA methylation signature in peripheral blood leukocytes was found, with an ROC AUC of 86%. In a case-control study of NSCLC, nine genes validated in tissues had a significantly higher frequency of tumor-specific hypermethylation in plasma. A five-gene set (APC, RASSF1A, CDH13, KLK10, and DLEC1) discriminated cases from controls with 84% sensitivity and 74% specificity.

4.1.2.3 RNA-Based Markers—Free circulating RNA assays are inherently challenging for quality template recovery, RNA-specific amplification, and detection of signal above background. RNA expression has been measured in circulating tumor cell and peripheral blood lymphocytes. Blood mononuclear cells are a more robust source than plasma or serum, yielding 76% sensitivity and 82% specificity. MicroRNAs may be more suitable as a blood biomarker given their resistance to degradation and consequent long half-life. In a cohort undergoing lung cancer screening, microRNAs discriminated current case-control status as well as predicted future lung cancer risk, both with ROC > 0.85. Although the source of discriminant signals remains unclear and the small sample size emphasizes the importance of replication, other groups have also reported an informative case-control signature.
4.1.2.4 Proteomic-Based Markers—Proteomics seeks protein patterns that distinguish malignant from premalignant or normal tissue, using any corresponding signals in the blood. Serum from a case-control study using an aptamer-based proteomics approach yielded a protein panel that discriminated NSCLC cases from controls with 90% sensitivity and 83% specificity. Moreover, the same markers had lower sensitivity (58%) but similar specificity in a validation population. A four-protein panel of proteomic-detected markers combined with candidate proteins from known lung cancer pathways had significant diagnostic value. The combination of serum protein biomarkers with PET scans for nodules had 92% specificity but only 26% sensitivity.

4.1.2.4.1 Candidate Proteins: The state of the art is to identify serum proteins, construct panels that distinguish lung cancer cases from controls, and independently validate results in separate populations. Candidate proteins, such as circulating serum carcinoembryonic antigen and cytokeratin 19-fragment (CYFRA 21-1) and lymphocyte antigen 6 complex locus K (LY6K), have discriminated, particularly in combination, between lung cancer cases and controls but not accurately enough to be considered for diagnostic use.

Because autoantibodies with lung tumor-specific epitopes show signals in lung cancer, on-chip protein libraries that detect antitumor antibodies appear promising. For example, a blood panel of p53, c-myc, HER2, NY-ESO-1, CAGE, MUC-1, and GBU4-5 autoantibodies detected by enzyme-linked immunosorbent assay in blood yielded 76% sensitivity and 92% specificity.

4.1.2.5 Metabolomics in Blood and Other Fluids—Metabolomics refers to the use of high-throughput generation of spectroscopic signatures representative of the metabolic state of a cell. Lung cancer-related metabolites in blood, such as S-adenosylmethionine, have been reported as biomarkers of early detection of lung cancer. On a small scale, high-resolution magnetic resonance spectroscopic results in blood were able to separate lung cancer cases from control subjects, and among cases, adenocarcinoma from squamous cell carcinoma. In urine, metabolomics discriminated lung cancer cases from control subjects with 93% sensitivity and 94% specificity.

4.1.2.6 Molecular Phenotypes—Phenotypes for mutagen sensitivity and DNA repair capacity have been developed that presumably integrate a large number of genes and pathways into a cancer-relevant trait. These markers of cytogenetic or DNA damage have been strongly associated with lung cancer risk in case-control studies.

4.1.2.7 Airway-Based Markers—Airway-based markers are attractive in that they are in the same anatomical and embryological compartment as that from which lung cancers arise.

4.1.2.7.1 Airway Specimens for Molecular Analyses Procured by Bronchoscopy: Novel molecular modalities of early diagnosis have been applied to bronchial biopsy specimens and BAL. SHOX2 methylation on bronchial aspirates, for example, from 523 patients separated lung cancer from nonmalignant disorders with 65% sensitivity and 95% specificity. Computer-assisted fluorescence interphase cytogenetics and immunophenotyping also have been used in BAL samples of a nodule.

In the major airways, observed field carcinogenesis patterns of transcriptome-wide gene expression reflect the likelihood that a radiographically detected peripheral lesion far from the bronchoscope is a malignancy. A next-generation sequencing study of the transcriptome also identified noncoding mRNAs differentiating case vs control bronchial epithelia.

4.1.2.7.2 Upper Airway Surrogates for the Lung: Because of field carcinization, the nose may provide easier access to the respiratory epithelium for transcriptional studies. Brush-exfoliated buccal cells are also transcriptionally active and provide another way to study epithelial gene-tobacco interactions. In a case-control study, expression of carcinogen metabolism pathway genes from buccal epithelium was correlated with gene expression in the lung. The value of this approach is currently being assessed in genome-wide studies for both buccal and nasal epithelia.

4.1.2.7.3 Sputum-Based Cytology: Standard sputum cytology is not adequate for lung cancer detection because the sensitivity is only 30% to 50% for proximal, slow growing lesions.

4.1.2.7.4 Sputum for DNA-Based Markers: Sputum-based detection has been studied extensively for genetic and epigenetic alterations. For genetic mutations, case-control evidence suggests that numerous k-ras and p53 mutations in sputum are associated with lung cancer. In lung cancer cases and smoking controls, fluorescence in situ hybridization (FISH)-detected HYAL2 and FHT deletions in sputum were correlated with those seen in tumor tissue. Three loci in sputum were examined for loss of heterozygosity and microsatellite instability within the FHT locus, with 55% sensitivity and 82% specificity. A sputum-based FISH assay may be informative when aimed at specific DNA targets. Compared with those collected > 18 months after lung cancer diagnosis, the sensitivity and specificity were substantially higher for the samples collected within 18 months of lung cancer diagnosis. Sensitivity was the highest for squamous cell cancers. The OR of lung cancer for specimens collected within 18 months before a cancer diagnosis was high (29.9; 95% CI, 9.5-94.1). Whether chromosomal aneusomy by FISH is a marker of lung...
cancer risk or alternatively detects exfoliated cancer cells is unknown.  

4.1.2.7.5 DNA Methylation in Sputum: DNA promoter methylation has been studied in sputum for several genes.  

A nested case-control study of heavy smokers with spirometrically confirmed airflow obstruction suggested that methylation of three or more genes in sputum was associated with a 6.5-fold increased risk of lung cancer within 18 months and that the risk increased with the number of methylated genes.  

In stage III lung cancer, methylation in sputum and tumor was consistent, with sputum methylation giving a 44% to 72% positive predictive value and a >70% negative predictive value.  

Usually, methylated sputum performed better than serum.  

In a study of quantitative methylation-specific PCR, 11 genes were silenced by methylation in lung tumors, adjacent nonmalignant lung tissues, and sputum.  

Three genes, 3-OST-2, DCR-1, and RASSF1A, had the highest levels in tumors and the lowest in adjacent nonmalignant tissues. For sputum, the combination of 3-OST-2, RASSF1A, p16, and APC significantly distinguished cases from controls (ROC AUC, 0.8).  

Double-strand break repair pathway SNPs and repair capacity are also associated with sputum hypermethylation.  

Sputum methylation was lower in those consuming a healthy diet and standard vitamin supplements.  

DNA methylation as a biomarker in this and other matrices has been reviewed.  

4.1.2.7.6 Sputum for RNA-Based Markers: Few RNA-based sputum studies of lung cancer detection credibly identify the amplicon as RNA. Those that showed an ability to amplify several transcripts from sputum lacked important no reverse transcriptase (negative) controls.  

Nevertheless, APRIL (a proliferation-inducing ligand) mRNA amplification from sputum appears to be feasible and yielded 82% sensitivity and 97% specificity.  

Also encouraging, microRNAs are ubiquitous regulatory features targeting mRNAs and have been reported to have a high capacity to discriminate lung cancer cases from controls in sputum.  

The signature of miR-21, miR-486, miR-375, and miR-200b distinguished lung adenocarcinoma cases from controls with 81% sensitivity and 92% specificity; similar results were reported for squamous cell carcinoma case-control discrimination in sputum.  

4.1.2.7.7 Sputum for Protein-Based Markers: Protein expression in sputum has been understudied. A ribonucleoprotein has been suggested as a sputum-based biomarker by immunocytochemistry but with few follow-up data. In a case-control study, sputum telomerase activity yielded 68% sensitivity and 90% specificity for concurrent lung cancer.  

A sputum-based measure of the oncogenic APRIL protein A provided good discrimination between lung cancer cases and controls.  

4.1.2.8 Exhaled Breath for Volatile Small Compounds—The gas phase of exhaled breath can be used to identify individual volatile components (eg, volatile organic compounds [VOC]) or complex volatile mixtures that indicate the presence of lung cancer. Most studies have been case-control studies in which confounders have not been thoroughly considered.  

Whehinger et al collected tidal volume breathing mixed expiratory gas samples prior to diagnostic or therapeutic interventions. A proton transfer MS approach to exhaled gas analysis that avoids preconcentration steps otherwise required for gas chromatography-based techniques was used. Among 17 predominantly early stage lung cancer cases and 170 control subjects, greater than twofold differences of cases vs smoker control subjects were observed for a mass-to-charge ratio of 31 or VOC-31 (tentatively protonated formaldehyde) and VOC-43 (tentatively a protonated fragment of isopropanol). In simulations, lung cancer cases were discriminated from control subjects with 54% sensitivity and 99% specificity.  

In case-control studies of the association between volatile compounds in the gas phase and lung cancer, gas chromatography-coupled MS patterns were significantly associated with lung cancer.  

Variations of lung cancer detection using a commercialized sensor array electronic nose method revealed that the actual volatile components of the unique signal were predominantly volatile hydrocarbons. Corroboration evidence simplifying the sensor array has been reported.  

With the use of gold nanoparticles, a VOC signature was created to distinguish stage III and IV case from control subjects of similar age and smoking history with a >86% accuracy.  

A study on dogs, a species enriched for biosensing capability, trained in study-blinded fashion, reported 99% sensitivity and specificity in distinguishing exhaled breath from lung cancer cases vs healthy control subjects.  

A report that lung cancer cells in vitro evolve unique volatile metabolites in the gas phase above the culture dish lends some credence to the cancer specificity of the evolving exhaled gas detection approach.  

4.1.2.8.1 Exhaled Breath Condensates for Cancer-Related Macromolecules: Large macromolecules indicative of carcinogenesis, such as DNA and proteins, are also detectable, somewhat counterintuitively, in the condensate (aqueous fraction) of exhaled breath, making possible assays for DNA-based markers such as specific p53 gene mutations, microsatellite markers, or methylated DNA.  

Similarly, small polypeptide molecules such as IL-2, tumor necrosis factor-α, and leptin show promise in distinguishing lung cancer cases from controls. In each instance, a higher rate of carcinogenesis-related DNA aberrancy was detected in the lung cancer cases vs controls with about an 80% sensitivity and specificity.
4.1.3 Summary—Innovative approaches are being used to develop biomarkers of lung cancer risk and early detection. Most of the evidence to date has been based on associations in case-control studies. Moving forward, a major question to be tested is whether these case-control associations will translate to meaningful predictive power for risk of lung cancer in prospective studies. In this rapidly advancing field, validation studies for some of the most promising markers are under way, holding promise for future translation to lung cancer prevention and detection efforts.

4.2 Presence of Acquired Lung Disease

Underlying lung disease could increase susceptibility to lung cancer. Acquired lung diseases assume two major forms: (1) airflow obstruction disorders, such as COPD, and (2) fibrotic disorders that restrict lung capacity, such as pneumoconiosis. Clear-cut inferences are elusive because of the complexity of this topic, but associations between both types of acquired lung disease and lung cancer have been noted.

The presence of COPD as well as impaired ventilation lung function are positively associated with the occurrence of lung cancer. However, cigarette smoking is the principal cause of both COPD and lung cancer, making it difficult to discern whether their co-occurrence is due to an etiologic association or, rather, to the shared risk factor of cigarette smoking. A strong association has been documented between airflow obstruction and increased risk of lung cancer, even after controlling for smoking. For example, in a population-based case-control study, lung cancer risk was elevated among individuals with chronic bronchitis (OR, 2.0; 95% CI, 1.5-2.5), emphysema (OR, 1.9; 95% CI, 1.4-2.8), or COPD (OR, 2.5; 95% CI, 2.0-3.1). More persuasive evidence comes from prospective studies of lung cancer incidence. In a nationally representative US cohort, moderate or severe obstructive lung disease was associated with a 2.8-fold (95% CI, 1.8-4.4) increased risk of lung cancer.

COPD may be linked to lung cancer through several mechanisms. In smoking-driven pathways, tobacco smoke exposure promotes inflammatory and mutagenic effects in the lungs that potentially induce both lung carcinogenesis and COPD, the latter through triggering the nuclear factor-κB pathway, a key pathway in COPD-related inflammation. Alternatively, the presence of COPD may indicate that the affected individual has received a greater dose of tobacco carcinogens than the typical unaffected individual. Other proposed mechanisms are independent of smoking. For example, one possible link of COPD with lung cancer is α1-antitrypsin deficiency (α1ATD). The prevalence of α1ATD carriers was observed to be higher among patients with lung cancer, including those who never smoked, than in the general population. In a case-control study, α1ATD carriers had a 1.7-fold increased risk of lung cancer (95% CI, 1.2-2.4), even after adjusting for smoking history and COPD. Regardless of mechanism, the presence of COPD is a clinically useful risk indicator for lung cancer.

In several studies, asthma was inversely associated with lung cancer risk. In contrast, a meta-analysis that rigorously controlled for smoking revealed a positive association between asthma and lung cancer risk, especially nonadenocarcinoma lung cancer. Asthma was also associated with lung cancer mortality in the National Health and Nutrition Examination Survey II Mortality Study. More recently, a case-control study and a cohort study in China observed that asthma was significantly associated with an increased risk of small cell lung cancer. Appropriately designed studies are still needed to establish whether and how asthma might increase the risk of lung cancer. Potential mechanisms proposed for an association between asthma and lung cancer are (1) mucociliary dysfunction leading to accumulation of toxicants, such as lung carcinogens, in the airway; (2) free radical damage to DNA as a result of imbalance between oxidants and antioxidants; and (3) chronic inflammation leading to chronic mitogenesis and an increased likelihood of conversion of endogenous DNA damage into mutations.

Clarifying the possible relationship between pneumoconioses and lung cancer poses particularly vexing challenges. Even for asbestos exposure, which is clearly established as a potent cause of lung cancer, whether lung cancer results from asbestos per se or from asbestosis remains controversial. Asbestos is likely to cause lung cancer through multiple mechanistic pathways. For other mineral fibers, the situation is less clear. For example, determining whether silica exposure or silicosis mediates the increased lung cancer risk in people exposed to silica has proven difficult. The presence of silicosis is associated with an increased risk of lung cancer. Results from a case-control study investigating silica dust exposure in relation to lung cancer observed a doubling in risk among those with the highest vs lowest exposure. In 2009, the IARC concluded that crystalline silica in occupational settings is a lung carcinogen, basing this conclusion primarily on the presence of a dose-response relationship and a pooled analysis of 10 major studies. Understanding the basis of this association will entail isolating the independent effects of silica exposure and lung fibrosis while taking exposure to smoking and other lung carcinogens into account.

The heterogeneity in the evidence between pneumoconioses and lung cancer emphasize that fibrosis...
is not a homogeneous exposure but, rather, depends on the properties of the specific mineral fiber or other environmental agent. In assessing potential harm, the agent’s size, shape, and durability are important considerations along with potential interactions with other exposures, such as cigarette smoking.

In addition to pneumoconioses, two other forms of interstitial lung disease (ILD) have been consistently linked to lung cancer: idiopathic pulmonary fibrosis (IPF) and systemic sclerosis (SSc). The potential relationship between these conditions and lung cancer is controversial because ILD has alternatively been hypothesized to (1) cause lung cancer, (2) be caused by lung cancer, and (3) share common pathogenetic mechanisms with lung cancer.322

4.2.1 Idiopathic Pulmonary Fibrosis: Wide-ranging associations from increased risk to protection have been observed between IPF and lung cancer, which likely is at least partially the result of the use of variable diagnostic criteria.322 Improvements to the international classification system have facilitated investigation of this topic. In autopsy studies, high proportions of lung cancer were seen in those with IPF, but these studies are prone to overestimate the role of IPF as a risk factor for lung cancer because IPF is a histopathologic marker of inflammatory response to toxic exposures commonly seen in lung cancer, including connective tissue disease, chemotherapy, radiotherapy, and surgery. On the other hand, the association between IPF and lung cancer estimated from registry data may be attenuated because of the potential misclassification of smoking status and the lack of histologic confirmation.323

The 10-year cumulative incidence of lung cancer in a series of 103 patients with IPF was 55%.324 In a cohort of patients with and without IPF, IFP was associated with a sevenfold increased lung cancer risk, an association that was robust to adjustment for confounding and was consistent in analyses limited to current smokers.325 These examples of the strikingly high risks of lung cancer observed in patients with IPF indicate that IPF is a clinical risk indicator of lung cancer risk and underscore the importance of further clarifying the etiologic significance of this association.

4.2.2 Systemic Sclerosis: ILD may also occur in the context of SSc, a rheumatologic disorder with a myriad of local and systemic manifestations. ILD, which occurs in most cases of SSc, is the major cause of morbidity and mortality in patients with SSc. Lung cancer is the most frequently reported malignancy in SSc, usually occurring in patients with concurrent SSc and ILD. The most common tumors seen in these patients are bronchioloalveolar cell and adenocarcinoma cell type.326 In studies in which the age-adjusted incidence rate of lung cancer in a registry of patients with SSc is compared with the general population, observed standardized incidence ratios were 1.2 (95% CI, 0.6-2.3),327 2.1 (95% CI, 1.4-3.0),328 4.9 (95% CI, 2.8-8.1),329 and 5.9 (95% CI, 3.0-10.3).330 The significant excess of lung cancer risk in patients with SSc in three of four studies provides evidence of a potential etiologic link between SSc and lung cancer. SSc symptoms could potentially contribute to lung carcinogenesis. For example, inflammation and fibrosis cause repeated cellular injury that could induce genetic damage in lung cells, creating a permissive environment for lung cancer progression.322 Alternatively, the frequent use of immunosuppressive drugs to treat SSc could result in increased lung cancer risk.322

5.0 INFECTIONS

The potential contributory role of infections to the etiology of lung cancer has been a long-standing concern, particularly with respect to TB,331,332 which has been covered in our previous reviews.35,36 More recent questions have centered on infections with human papillomaviruses333,334 and HIV. Because of strong evidence that human papillomavirus is absent in lung tumor tissue,335 HIV infection is the focus of this review.

5.1 HIV

Since the first case report of lung cancer in a 35-year-old patient with HIV in 1984,336 myriad case reports,337-345 patient series,346-351 and case-control studies352-359 have described young smokers with HIV infection presenting mainly with NSCLC of adenocarcinoma histology with advanced stage of disease and shortened survival rates. Data from many industrialized nations indicate that patients with HIV infection and lung cancer are significantly younger than patients with lung cancer in general.340,344,346,349,350

In fact, lung cancer is now the third most frequent neoplasm in individuals with HIV infection, trailing only the AIDS-defining cancers of Kaposi sarcoma and non-Hodgkin’s lymphoma.361-363 Lung cancer is thus the most common and most fatal non-AIDS-associated malignancy in the HIV-infected population, accounting for about 16% of deaths in patients with HIV infection.364,365

Individuals with HIV infection have been repeatedly observed to have a higher lung cancer risk than those without HIV infection,366-375 with relative risk estimates ranging from 2 to 11.366,375 Inherent in these reports was a lack of adjustment for important confounders, such as cigarette smoking, that may affect
the lung cancer risk in the HIV population. This is an important concern because patients with HIV infection and lung cancer are almost exclusively smokers. On average, individuals with HIV infection smoke more than the general population, but the lifetime smoking histories of patients with HIV and lung cancer tend to be less extensive than in patients with lung cancer but not HIV infection because of the younger age of disease onset. A relative risk estimate of 2.5 was derived based on plausible, yet conservative assumptions about tobacco use. Two large cohort studies with internal comparison groups of participants without HIV infection and prospectively collected smoking data corroborate this estimate, with relative risks of 3.6 and 2.6. Thus, the current body of evidence clearly suggests that HIV infection increases lung cancer risk independent of smoking status by a factor of at least 2.5-fold.

Immunosuppressed individuals are at an increased risk for lung cancer, but most patients with HIV infection and lung cancer only have moderate immunosuppression and CD4 counts as well as HIV viral loads are not strongly related to increased lung cancer risk. The average latency between HIV and lung cancer diagnosis is at least 5 years, and there is no convincing evidence that antiretroviral medication increases lung cancer risk.

The widespread use of highly active antiretroviral therapy in the United States since 1996 has prolonged survival among adults with HIV, leading to significant aging of the HIV/AIDS population. From 1990 to 2001, the number of adults with AIDS aged ≥ 50 years increased more than fivefold. Because lung cancer risk increases markedly with age, lung cancer can be expected to become increasingly common as the HIV-infected population ages. The elevated lung cancer risk in the HIV-infected population also has implications for the racial disparity in the occurrence of lung cancer because African Americans comprise a much higher percentage of the HIV-infected population than the general population (46% vs 12%).

Compared with all patients with advanced stage lung cancer, the median survival of patients with concurrent HIV infection is significantly shorter (3-6 vs 10-12 months). Most patients with HIV infection and lung cancer present with advanced stage disease, leaving only 10% to 15% with disease amenable to curative resection. The poor performance status of patients with HIV infection and lung cancer undermines their ability to tolerate surgery, chemotherapy, and radiation therapy so that lung cancer remains untreated in almost one-fourth of these patients. Good performance status, higher CD4 counts, and continuing highly active antiretroviral therapy during cytotoxic chemotherapy confer a survival advantage, but protease inhibitor use has been associated with an unacceptably

| Table 1—Summary of Findings: Key Factors Associated With Risk of Lung Cancer |
|----------------------------------|---------------------------------------------|
| **Factor**                       | **Description**                             |
| **A. Single most important causal determinant of individual and population risk, most valuable indicator of clinical risk** | Active smoking of cigarettes and other tobacco products: Individual risk increases with greater number of cigarettes smoked per day and greater number of years of smoking. Population risk increases with the prevalence of current smokers because population prevalence predicts lung cancer occurrence with a latency period of about 20 y. |
| **B. Other risk factors causally associated with lung cancer** | Secondhand smoke exposure; Ionizing radiation, including radon; Occupational exposures, eg, arsenic, chromium, nickel, asbestos, tar, and soot; Indoor and outdoor air pollution |
| **C. Additional clinical risk indicators** | The risk factors noted above, plus: Older age; Male sex, particularly among those of African American ancestry; Family history of lung cancer; Acquired lung disease, eg, COPD, TB, pneumoconioses, idiopathic pulmonary fibrosis, and systemic sclerosis; Occupational exposures, such as to silica dust; HIV infection |
| **D. Examples of associations with consistent evidence but causal role not presently established** | Fruit and vegetable intake (decreased risk); Physical activity (decreased risk); Marijuana smoking (not associated with risk) |

The evidence for factors listed in these categories is extremely strong to meet epidemiologic criteria for causality.

*The factors listed under clinical risk indicators are all strongly associated with increased risk of lung cancer but are listed in this category either because they are intrinsic characteristics of the patient (age, sex, ethnic ancestry, family history) or are factors with consistent evidence of increased risk that presently falls short of being rated as causal.
The numerous exposures known to cause lung cancer, summarized in Table 1, chart the path to its prevention. Further, as shown in Table 1, these causal agents combined with sociodemographic characteristics, family history, and characteristics such as acquired lung disease or HIV infection provide a suite of clinical risk indicators.

Steps to reduce or eliminate the population’s exposure to the causal agents would be expected to reduce the population’s risk of lung cancer. Preventive strategies can be pursued in the public policy arena or in public health interventions directed at individual behavior. Cigarette smoking provides a useful example to illustrate the multiple levels that can form the basis of preventive strategies. In the legislative and regulatory arena, examples of tobacco control strategies include limiting cigarette advertising, reducing children’s access to cigarettes, and prohibiting smoking in the workplace. A thorough list of policy measures has been set forth in the WHO Framework Convention on Tobacco Control. Litigation against cigarette manufacturers is a productive component of tobacco control strategies, as exemplified by the settlement between US states and the tobacco industry. Behavioral interventions to prevent children and adolescents from starting to smoke cigarettes and behavioral and pharmacologic interventions to promote smoking cessation are individual-level approaches that, if successful, could reduce the occurrence of lung cancer.

In developing lung cancer prevention strategies, certain groups warrant particular attention. Steps need to be taken to reduce the very high lung cancer incidence rates in African American men. Lung cancer is a major women’s health issue. Because of historical cigarette smoking patterns, the epidemic of lung cancer started later in women than men, but in contrast to the situation in men, lung cancer incidence rates in women have not yet begun to decrease consistently. Although lung cancer remains a critical public health problem, the decrease in the overall lung cancer burden that is presently occurring in the United States, as in much of the developed world, reflects the successes of preventive strategies. A critical global priority is to prevent the uptake of cigarette smoking in developing countries where smoking prevalence is still low to prevent the increase in lung cancer rates that consistently follows an increase in smoking prevalence. Another imperative is to identify smokers and ex-smokers at highest risk according to sociodemographic and molecular factors so that prevention and detection efforts can be directed at those at highest risk.

A consideration of the epidemiology of lung cancer consistently reinforces one major theme: The pandemic of lung cancer is a consequence of the tragic and widespread addiction to cigarettes. Curtailing the pandemic of lung cancer will require preventing youths from starting to smoke cigarettes and effectively promoting smoking cessation among dependent smokers.

ACKNOWLEDGMENTS

Author contributions: Dr Alberg had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Dr Alberg: contributed to the content, writing, and editing of the manuscript.

Dr Brock: contributed to the content, writing, and editing of the manuscript.

Dr Ford: contributed to the content, writing, and editing of the manuscript.

Dr Samet: contributed to the content, writing, and editing of the manuscript.

Dr Spizzichino: contributed to the content, writing, and editing of the manuscript.

Financial/nonfinancial disclosures: The authors have reported to CHEST that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Role of Sponsors: The American College of Chest Physicians was solely responsible for the development of these guidelines. The remaining supporters played no role in the development process. External supporting organizations cannot recommend panelists or topics, nor are they allowed prepublication access to the manuscripts and recommendations. Further details on the Conflict of Interest Policy are available online at http://chestnet.org.

Endorsements: This guideline is endorsed by the European Society of Thoracic Surgeons, Oncology Nursing Society, American Association for Bronchology and Interventional Pulmonology, and the Society of Thoracic Surgeons.

REFERENCES


52. Trinidad DR, Pérez-Stable EJ, Messer K, White MM, Pierce JP. Menthol cigarettes and smoking cessation among racial/ethnic groups in the United States. *Addiction*. 2010;105(suppl 1):54-94.


Epidemiology of Lung Cancer


