Low Lung Function and Incident Lung Cancer in the United States

Data From the First National Health and Nutrition Examination Survey Follow-up

David M. Mannino, MD; Samuel M. Aguayo, MD; Thomas L. Petty, MD; Stephen C. Redd, MD

Background: Obstructive lung disease and lung cancer are tobacco-related diseases that can remain clinically silent until late in the disease process. We sought to define the risk for incident lung cancer among a national cohort of US adults with and without obstructive lung disease.

Methods: We studied participants in the First National Health and Nutrition Examination Survey, who had up to 22 years of follow-up. We classified subjects as having moderate or severe obstructive lung disease at baseline if the ratio of forced expiratory volume in 1 second (FEV1) to forced vital capacity (FVC) was less than 70% and the FEV1 was less than 80% of the predicted value. We also determined incident cases of lung cancer during the follow-up period.

Results: A total of 113 lung cancers occurred in the 5402 adults in the cohort. In the proportional hazards model adjusted for covariates of age, sex, race, education, smoking status, and duration and intensity of smoking, the presence of moderate or severe obstructive lung disease was associated with a higher risk for incident lung cancer (hazard ratio, 2.8; 95% confidence interval, 1.8-4.4).

Conclusions: The presence of moderate or severe obstructive lung disease is a significant predictor of incident lung cancer in long-term follow-up. This finding may be useful clinically and in studies evaluating the utility of new tools for the early detection of lung cancer.

Arch Intern Med. 2003;163:1475-1480

Lung cancer is the leading cause of cancer-related mortality in the United States. In 2000, an estimated 156,900 deaths were attributed to lung cancer.1 Tobacco smoking is the most important risk factor for lung cancer, although exposures to other agents, such as radon and asbestos, are also important factors.2 A diagnosis of obstructive lung disease has been found to be associated with a higher rate of lung cancer in several studies,3-5 although none of these have been representative of the US adult population.

The diagnosis of obstructive lung disease has traditionally depended on the presence of symptoms such as chronic cough or sputum production.6 New international guidelines for the diagnosis of obstructive lung disease depend almost exclusively on measured lung function to diagnose and classify disease.7

We applied spirometric criteria for the diagnosis of obstructive lung disease to the cohort of 5402 subjects who had pulmonary function measurements obtained as part of the First National Health and Nutrition Examination Survey (NHANES I).8,9 We searched the NHANES I follow-up database for incident cases of lung cancer in the up to 22 years of follow-up and determined the significant predictors of lung cancer in this cohort.

Methods: The NHANES I was conducted by the National Center for Health Statistics from 1971 to 1975.6,9 This was a survey of a probability sample of the civilian, noninstitutionalized population of the United States. Follow-up surveys of the adult participants (aged 25-74 years old) in the NHANES I occurred in 1982 to 1984, 1986, 1987, and 1992.10,11 Data collected on participants included hospitalization records, vital status, and, for decedents, death certificates. Through 1992, 96% of the original cohort had been successfully traced, and death certificates were available for 98% of the 4604 documented decedents.11

Questionnaire data

Participants in the NHANES I completed an extensive questionnaire that included the demographic data of age, race, sex, and education level. We classified subjects as having no greater than 12 or at least 13 years of education. A nationally representative subset of participants...
completed a cardiorespiratory module that included a series of questions about the presence of respiratory symptoms and the diagnosis of respiratory disease. The pulmonary symptoms included in the analysis (to define an atypical subset of the population to calculate equations for lung function) were cough (defined as a positive response to “Have you ever had a cough first thing in the morning in the winter?” or “Have you ever had a cough first thing in the morning in the summer?”), sputum (defined as a positive response to “Have you ever had any phlegm from your chest first thing in the morning in the winter?” or “Have you ever had any phlegm from your chest first thing in the morning in the summer?”), and wheeze (defined as a positive response to “Have you ever had wheezy or whistling sounds in your chest?”). Participants were also asked if they had physician-diagnosed chronic bronchitis (nonallergic), emphysema, or asthma.

Complete smoking histories were obtained on all participants. We considered participants to be current smokers if they reported the use of cigarettes, cigars, or pipes at the time of the survey. We considered participants to be former smokers if they reported any previous use of cigarettes (≥100), cigars (≥20), or pipes (≥3 packages of tobacco), but no current use. Long-term intensity of use data were available only for cigarettes. We calculated pack-years of cigarette use by multiplying the average number of cigarettes smoked daily by the number of years smoked and dividing the product by 20. Former cigarette smokers reported how long it had been since they smoked cigarettes fairly regularly.

PULMONARY FUNCTION DATA

Subjects who participated in the cardiorespiratory module performed spirometry using a spirometer (Model 800; Ohio Medical Instrument Company, Inc, Cincinnati, Ohio). The procedures used have been documented previously. Subjects were excluded from this analysis if they did not perform spirometry or had results that were not reliable. Values used in this analysis included forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1), and the FEV1/FVC ratio. We determined predicted values of FEV1 and FVC by performing linear regression (stratified by sex and using age and height as predictors) on a subgroup of participants who were white, had never smoked, and did not report respiratory symptoms or physician-diagnosed lung disease. The results from these regression models were applied to the data from all participants to obtain predicted values of FEV1 and FVC. We used an adjustment factor of 0.88 to estimate predicted values for African American participants. We classified the subjects as having moderate or severe obstructive lung disease if the FEV1/FVC ratio was less than 70% and the FEV1 was less than 80% of the predicted value; as having mild obstructive lung disease if the FEV1/FVC ratio was less than 70% and the FEV1 was at least 80% of the predicted value; as having mild obstructive lung disease if the FEV1/FVC ratio was less than 70% and the FEV1 was at least 80% of the predicted value; as having mild obstructive lung disease if the FEV1/FVC ratio was less than 70% and the FEV1 was at least 80% of the predicted value; as having moderate obstructive lung disease if the FEV1/FVC ratio was less than 70% and the FEV1 was at least 80% of the predicted value; and as having restrictive disease if the FEV1/FVC ratio was at least 70% and the FVC was less than 80% of the predicted value.

INCIDENT CASES OF LUNG CANCER

Participants who reported cancer at the baseline interview were excluded from the analysis. Lung cancer cases were defined as any hospital discharge diagnosis coded 162 using the International Classification of Diseases, Ninth Revision, Clinical Modification, or a death certificate containing this code.

RESULTS

A total of 14407 adults aged 25 to 74 years participated in the NHANES I. From this sample, 6913 participated in the cardiorespiratory survey and examination. We excluded 1371 subjects who did not undergo pulmonary function testing or had results that were not reliable, and 140 subjects who admitted to a history of cancer at the baseline evaluation, resulting in 5402 subjects in the final analytic cohort. Subjects excluded because of missing pulmonary function data were more likely to be older than 60 years (37.8% vs 22.8%; P<.001) and to be of nonwhite race (21.8% vs 11.8%; P<.001) than subjects included in the final cohort. Current smokers had a mean smoke exposure of 33.1 pack-years (SE, 0.7 pack-years) and 27.9 cigarettes per day (SE, 0.4 cigarette per day), whereas former smokers had a mean smoke exposure of 33.3 pack-years (SE, 1.0 pack-years) and 25.4 cigarettes per day (SE, 0.6 cigarette per day) and had quit smoking 8.6 years (SE, 0.3 years) before the initial survey.

Median duration of follow-up of the cohort was 17.9 years (interquartile range [IQR], 15.4-19.9 years). During the follow-up period, 113 cases of lung cancer (2.1% of cohort) (Table 1) were identified, including 82 identified from hospitalization records and 31 from death certificates. Lung cancer developed during follow-up in a similar proportion of subjects who did not participate in the cardiorespiratory survey (2.0%), but it developed in a higher proportion of subjects without pulmonary function testing (3.8%). Of the 113 lung cancer cases, 85 occurred in current smokers, 18 in former smokers, and 10 in never smokers. Among current smokers, lung cancer developed in 3.7% in the follow-up period, compared with 1.7% of former smokers and 0.5% of those who never smoked (Table 2). Mild and moderate obstructive disease and restrictive disease at the baseline evaluation were associated with a higher proportion of
lung cancer development among current and former smokers (Table 2).

The median time to lung cancer diagnosis (for the 82 cases detected on the basis of hospitalization records) was 9.8 years (IQR, 5.3-15.5 years). The median time to lung cancer diagnosis varied by level of lung function. Among subjects with no obstructive lung disease, this time was 12.4 years (IQR, 8.5-16.6 years); among subjects with restrictive lung disease, 10.0 years (IQR, 6.3-15.3 years); among subjects with mild obstructive lung disease, 9.5 years (IQR, 6.7-17.8 years); and among subjects with moderate or severe obstructive lung disease, 5.3 years (IQR, 2.6-9.8 years). Similar results were obtained when the end point was lung cancer diagnosis or death due to lung cancer among current and former smokers (Figure).

The median age at lung cancer diagnosis (for the 82 cases detected on the basis of hospitalization records) was 69.6 years (IQR, 63.1-74.4 years). The median age at lung cancer diagnosis varied by level of lung function. Among subjects with no obstructive lung disease, this age was 68.2 years (IQR, 63.1-74.4 years); among subjects with moderate or severe obstructive lung disease, 71.1 years (IQR, 67.1-74.4 years).

Moderate or severe obstructive lung disease was a significant predictor of incident lung cancer in logistic regression models that controlled for age, smoking status, sex, education level, pack-years of cigarette smoking, and years since quitting smoking (odds ratio [OR], 2.4; 95% confidence interval [CI], 1.5-3.8). Similar results were obtained in a model restricted to current and former smokers (OR, 2.6; 95% CI, 1.6-4.2).

In the multivariate proportional hazards analysis (Table 3), moderate or severe obstructive lung disease was a significant hazard for incident lung cancer in the follow-up period among the entire cohort (hazard ratio [HR], 2.8; 95% CI, 1.8-4.4) and when limited to current or former smokers (OR, 3.0; 95% CI, 1.9-4.8).

**COMMENT**

In this analysis of a nationally representative cohort of the US population with up to 22 years of follow-up, we found that the presence of moderate or severe obstructive lung disease, particularly moderate or severe obstructive lung disease, was a significant risk factor for incident lung cancer. Further studies are needed to confirm these findings and to explore the potential mechanisms underlying the association between obstructive lung disease and lung cancer risk.
The development of lung cancer during the follow-up period.3 Another US study that included controls (all among current smokers) during the 10-year follow-up period.4 The National Lung Health Education Program, recommends spirometric testing for all smokers (current and former smokers, stratified by degree of lung function impairment. From the National Health and Nutrition Examination Survey, 1971 through 1975, and follow-up through 1992. COPD indicates chronic obstructive pulmonary disease.

International studies have shown similar results. A Danish study of 13946 participants followed up for 10 years found increased mortality risks due to lung cancer in subjects with an FEV₁ of less than 40% (HR, 3.9 [95% CI, 2.2-7.2]) or of 40% to 79% (HR, 2.1 [95% CI, 1.3-5.4]) when compared with subjects with an FEV₁ of at least 80%.3 In a Scottish study of 15411 subjects followed up for 15 years, men and women with low lung function (FEV₁ of <73% and <75%, respectively) had a higher risk for lung cancer mortality (HR, 2.5 and 4.4, respectively) compared with subjects in the highest quintile of lung function.18

One difference between our study and the previous studies was our separate evaluations of mild obstructive lung disease and restrictive lung disease. There was a suggestion in the univariate models that lung cancer is more likely to develop in both of these groups (Table 3), although these findings did not remain significant in the multivariate models.

We used objective criteria for obstructive lung disease to define the population with moderate or severe airway obstruction.7 We know from other studies that 63% of the adult US population with this degree of airway obstruction has never received a diagnosis of any lung disease,19 and that dependence on clinical signs or symptoms will not detect most of these individuals.20,21 To address this gap, a new national health care initiative, the National Lung Health Education Program, recommends spirometric testing for all smokers (current and former).

### Table 2. Participants Included in Analysis Stratified by Smoking Status and Degree of Pulmonary Function Impairment

<table>
<thead>
<tr>
<th>Pulmonary Function Impairment</th>
<th>No. of Participants</th>
<th>No. (%) of Lung Cancer Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current smokers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal lung function</td>
<td>1552</td>
<td>33 (2.1)</td>
</tr>
<tr>
<td>Restrictive lung disease†</td>
<td>222</td>
<td>10 (4.5)</td>
</tr>
<tr>
<td>Mild COPD‡</td>
<td>207</td>
<td>11 (5.3)</td>
</tr>
<tr>
<td>Moderate/severe COPD§</td>
<td>288</td>
<td>31 (10.8)</td>
</tr>
<tr>
<td>Total</td>
<td>2269</td>
<td>85 (3.7)</td>
</tr>
<tr>
<td><strong>Former smokers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal lung function</td>
<td>807</td>
<td>6 (0.7)</td>
</tr>
<tr>
<td>Restrictive lung disease†</td>
<td>78</td>
<td>2 (2.6)</td>
</tr>
<tr>
<td>Mild COPD‡</td>
<td>89</td>
<td>3 (3.4)</td>
</tr>
<tr>
<td>Moderate/severe COPD§</td>
<td>105</td>
<td>7 (6.7)</td>
</tr>
<tr>
<td>Total</td>
<td>1079</td>
<td>18 (1.7)</td>
</tr>
<tr>
<td><strong>Never smokers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal lung function</td>
<td>1643</td>
<td>7 (0.4)</td>
</tr>
<tr>
<td>Restrictive lung disease†</td>
<td>201</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Mild COPD‡</td>
<td>127</td>
<td>2 (1.6)</td>
</tr>
<tr>
<td>Moderate/severe COPD§</td>
<td>93</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>2054</td>
<td>10 (0.5)</td>
</tr>
<tr>
<td><strong>All subjects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal lung function</td>
<td>4002</td>
<td>46 (1.2)</td>
</tr>
<tr>
<td>Restrictive lung disease†</td>
<td>501</td>
<td>13 (2.6)</td>
</tr>
<tr>
<td>Mild COPD‡</td>
<td>423</td>
<td>16 (3.8)</td>
</tr>
<tr>
<td>Moderate/severe COPD§</td>
<td>476</td>
<td>38 (8.0)</td>
</tr>
<tr>
<td>Total</td>
<td>5402</td>
<td>113 (2.1)</td>
</tr>
</tbody>
</table>

Abbreviations: COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity.

†Defined as FEV₁/FVC of ≥70% and FVC of <80%.
‡Defined as FEV₁/FVC of <70% and FEV₁ of ≥80%.
§Defined as FEV₁/FVC of <70% and FEV₁ of <80%.

former) aged 45 years or older and anyone with respiratory symptoms.22

A surprising finding in our analysis was that the proportion of former smokers with moderate obstructive disease in whom lung cancer developed was almost as high as the proportion of current smokers (6.7% vs 10.8%). Many lung cancers are being diagnosed in former smokers because of the increasing numbers of former smokers in the population and their greater age (relative to current smokers).23,24 Most studies suggest that even after 3 to 5 years of smoking abstinence, the risk for cancer is similar to that of current smokers.23,25 In our univariate models, former smokers who stopped smoking at least 10 years before the survey had an HR of 3.7 for development of lung cancer, compared with those who never smoked (Table 3). In our analysis, we only determined smoking status at baseline, although a proportion of these subjects likely stopped smoking during the follow-up period, which would result in an even higher proportion of lung cancer in former smokers.26

Another finding in our analysis was that although subjects without obstructive lung disease had a longer interval from the time of the survey to the time of lung cancer diagnosis (12.4 vs 5.3 years for subjects with obstructive lung disease), they received a diagnosis of lung cancer at a slightly younger age (68.2 vs 71.1 years). This difference raises the possibility that the factors related to lung cancer in the presence of obstructive lung disease differ from those in its absence.

Lung cancer remains one of the most deadly malignancies in the United States, with an estimated 164100 new cases and 156900 deaths in 2000.1 Screening for lung cancer is currently not recommended, although the basis for this policy consists of studies performed more than 20 years ago.1,17,28 Modern imaging techniques involving computed tomography scanning with low-dosage radiation may be able to alter this recommendation.29,30 The findings of our study point to the potential utility of pulmonary function testing among current and former smokers to identify a subgroup of patients in whom lung cancer is more likely to develop and who are more likely to benefit from screening or case detection.31-34

There are certain limitations to our analysis. Pulmonary function data were obtained only at the baseline evaluation, so we could not determine in which subjects obstructive lung disease may have developed in the follow-up period. Hospitalizations and deaths associated with lung cancer had no independent confirmation of the malignancy. Some subjects may have died with occult lung cancer present. Despite these potential limitations, our overall findings of risk factors for lung cancer are consistent with what is known about the epidemiology of lung cancer, suggesting any misclassification did not significantly alter our results.

### CONCLUSIONS

In this nationally representative cohort of US adults, the presence of moderate or severe obstructive lung disease was independently associated with an increased risk for incident lung cancer during up to 22 years of follow-up. This finding may be useful in better defining popula-

---

### Table 3. Proportional Hazards Model for Incident Lung Cancer

<table>
<thead>
<tr>
<th>Lung function</th>
<th>Univariate Models</th>
<th>Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal lung function</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Restrictive lung disease†</td>
<td>2.5 (1.4-4.6)</td>
<td>1.5 (0.8-2.8)</td>
</tr>
<tr>
<td>Mild COPD‡</td>
<td>3.6 (2.0-6.3)</td>
<td>1.4 (0.8-2.6)</td>
</tr>
<tr>
<td>Moderate/severe COPD§</td>
<td>8.8 (5.8-13.5)</td>
<td>2.8 (1.8-4.4)</td>
</tr>
</tbody>
</table>

Age, y
25-29 1.0 1.0
40-49 8.0 (2.8-23.3) 5.8 (2.0-16.6)
50-59 15.3 (5.5-42.6) 9.0 (3.2-25.4)
60-69 22.4 (8.1-62.3) 16.1 (5.8-44.3)
70-74 42.4 (14.6-123.4) 40.2 (13.6-119.0)

Race
White 0.9 (0.5-1.7) 1.0 (0.6-1.8)
Nonwhite 1.0 1.0

Sex
Male 3.2 (2.1-4.8) 1.7 (1.1-2.6)
Female 1.0 1.0

Education, y
≤12 4.0 (2.1-7.7) 2.5 (1.3-4.9)
≥13 1.0 1.0

Smoking status
Current smoker 8.4 (4.4-16.0) 3.5 (1.0-12.4)
Former smoker 3.6 (1.7-7.8) 1.1 (0.4-3.3)
Never smoker 1.0 1.0

Pack-years
Never smoker 1.0 1.0
<30 2.9 (1.4-5.9) 1.0
30 to <60 9.0 (4.5-18.1) 1.4 (0.9-2.4)
≥60 19.9 (10.0-39.6) 2.2 (1.3-3.8)

Years since regularly smoked
Never smoker 1.0 1.0
0 7.4 (3.9-14.3) 1.6 (0.6-4.2)
<1 to <10 6.3 (2.9-13.5) 2.0 (0.8-5.1)
≥10 3.7 (1.3-10.0) 1.0

Product of smoking status and years since regularly smoked
Never smoker 1.0 1.0
<30 2.9 (1.4-5.9) 1.0
30 to <60 9.0 (4.5-18.1) 1.4 (0.9-2.4)
≥60 19.9 (10.0-39.6) 2.2 (1.3-3.8)

Yearly average pack-years
Never smoker 1.0 1.0
<30 2.9 (1.4-5.9) 1.0
30 to <60 9.0 (4.5-18.1) 1.4 (0.9-2.4)
≥60 19.9 (10.0-39.6) 2.2 (1.3-3.8)

Abbreviations: CI, confidence interval; COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; HR, hazards ratio.

†Defined as FEV1/FVC of ≥70% and FEV1 of <80%.
‡Defined as FEV1/FVC of <70% and FEV1 of ≥80%.
§Defined as FEV1/FVC of <70% and FEV1 of ≥80%.

### REFERENCES


Accepted for publication September 16, 2002.

Corresponding author: David M. Mannino, MD, National Center for Environmental Health, Centers for Disease Control and Prevention, 1600 Clifton Rd, Mail Stop E-17, Atlanta, GA 30333 (e-mail: dmannino@cdc.gov).


