Lung Cancer Screening with Sputum Cytologic Examination, Chest Radiography, and Computed Tomography: An Update for the U.S. Preventive Services Task Force

Linda L. Humphrey, MD, MPH; Steven Teutsch, MD, MPH; and Mark Johnson, MD, MPH

Background: Lung cancer is the leading cause of cancer-related death in the United States and worldwide. No major professional organizations, including the U.S. Preventive Services Task Force (USPSTF), currently recommend screening for lung cancer.

Purpose: To examine the evidence evaluating screening for lung cancer with chest radiography, sputum cytologic examination, and low-dose computed tomography (CT) to aid the USPSTF in updating its recommendation on lung cancer screening.

Data Sources: MEDLINE, the Cochrane Library, reviews, editorials, and experts.

Study Selection: Studies that evaluated mass screening programs for lung cancer involving the tests of interest were selected. All studies were reviewed, but only studies with control groups were rated in quality since these would most directly influence the USPSTF screening recommendation.

Data Extraction: Data were abstracted to data collection forms. Studies were graded according to criteria developed by the USPSTF.

Data Synthesis: None of the 6 randomized trials of screening for lung cancer with chest radiography alone or in combination with sputum cytologic examination showed benefit among those screened. All studies were limited because some level of screening occurred in the control population. Five case-control studies from Japan suggested benefit to both high- and low-risk men and women. All studies were limited by potential healthy screenee bias. Six cohort studies showed that when CT was used to screen for lung cancer, lung cancer was diagnosed at an earlier stage than in usual clinical care. However, these studies did not have control groups, making mortality evaluation difficult. In addition, the studies demonstrated a high rate of false-positive findings.

Conclusions: Current data do not support screening for lung cancer with any method. These data, however, are also insufficient to conclude that screening does not work, particularly in women. Two randomized trials of screening with chest radiography or low-dose CT are currently under way and will better inform lung cancer screening decisions.


For author affiliations, see end of text.

See related article on pp 738-739.
believed (primarily on the basis of indirect evidence) (22–28) that early surgical resection is associated with better outcomes. Therefore, the current standard of practice is to resect most non–small-cell lung cancer without evidence of metastatic spread. For many of these reasons, screening for and treating early lung cancer is intuitively appealing.

**Methods**

This review discusses studies of chest radiography, sputum cytologic examination, and low-dose computed tomography (CT) for lung cancer screening and focuses on the outcomes of screening in populations. We reviewed the MEDLINE and Cochrane databases from their inception through January 2003 using the search terms lung neoplasms, lung cancer, and any screening. The search strategy is detailed in Appendix Table 1 (available at www.annals.org). To ensure complete ascertainment, we reviewed the bibliographies of reviews, editorials, book chapters, and letters discussing lung cancer screening, as well as a recent Cochrane review and analysis (29). We sought studies evaluating screening in the general population, as well as in high-risk populations, and included observational studies and clinical trials. Observational studies with control groups and controlled trials evaluating disease-specific mortality were evaluated for quality according to criteria created by the USPSTF (30) (Appendix, available at www.annals.org). For the purposes of this review, high-risk persons are those who currently smoke or have ever smoked and low-risk persons are those who have never smoked. To rate each of the studies, we reviewed all original articles discussing the study’s methods or findings. We also used studies of the various screening methods to estimate the screening test characteristics of chest radiography and low-dose CT. Finally, we used data from screening studies (when available), as well as clinical series, to evaluate the harms associated with screening and treatment. For completeness, all studies are described in the tables; however, only studies rated as fair or better quality are described in the text.

Methodologic issues relevant to understanding screening studies include lead-time bias (when the time of diagnosis is advanced by screening but the time of death is unchanged), length bias (bias toward detecting less aggressive tumors in a periodically screened sample) (31), and volunteer bias (a type of selection bias in which volunteers are compared with nonvolunteers) (32). Overdiagnosis occurs when cancer that would never have been important during an individual’s lifetime is diagnosed and treated. These biases can be eliminated only in randomized, controlled trials that include death as an outcome. Therefore, public health guidelines and this review place the most emphasis on information from randomized, controlled trials.

This research was funded by the Agency for Healthcare Research and Quality. Agency staff and USPSTF members reviewed interim analyses and the final report.

Before preparation of this manuscript, the full report was reviewed by 17 content experts in lung cancer screening and was revised accordingly.

**Data Synthesis**

In our searches, we identified 809 citations and abstracts; 149 full-text papers were reviewed. Of these, 1 randomized trial of chest radiography in conjunction with a multiphasic screening program (33, 34) and 5 randomized, controlled trials of chest radiography, sputum cytologic examination, or both as screening for lung cancer (35–40) were reviewed. In addition, 6 case–control studies (41–46), 1 nonrandomized controlled trial (47), and 4 older cohort studies (48–52) were reviewed (Appendix Table 2, available at www.annals.org). We also reviewed 6 recent cohort studies of lung cancer screening with CT (53–62).

**Lung Cancer Screening with Chest Radiography with or without Sputum Cytologic Examination**

**Controlled Trials**

The methods and quality of the 6 randomized, controlled trials and the single nonrandomized controlled trial of lung cancer screening (33–40, 47, 63–85) are shown in Tables 1 and 2. The Figure shows the relative risks and 95% CIs of these randomized trials. In the 1960s, the Northwest London Mass Radiography Service conducted a cluster randomized trial of chest radiography screening in approximately 55,000 men older than 40 years of age (35, 36). In this trial, 29,723 male factory workers from 75 randomly identified firms were offered chest radiography every 6 months and were compared with 25,300 controls from other factories who were offered screening at baseline and at 3 years. After 3 years, the annual mortality rate from lung cancer was 0.7 per 1000 person-years in the intervention group and 0.8 per 1000 person-years in the control group, not a statistically significant difference.

The National Cancer Institute sponsored 3 randomized, controlled trials of lung cancer screening in male smokers in the United States in the 1970s (37–39, 63, 64, 68, 73–75, 80). The Memorial Sloan-Kettering Study (37, 63–67) and the Johns Hopkins Study (38, 68–72) were identical in design and were conducted to evaluate the incremental benefit of adding sputum cytologic examination to annual chest radiography. Of the 20,427 male smokers (≥20 pack-years of smoking) age 45 years or older who volunteered for these 2 studies, 10,234 were randomly assigned to a dual-screening group that was offered screening with chest radiography annually and sputum cytologic examination every 4 months for 5 years; 10,233 were assigned to a chest radiography group that was offered annual screening for 5 years. Each group was followed for 5 to 8 years.

In the Memorial Sloan-Kettering Study, baseline screening identified 30 (6.0 per 1000) malignant tumors in the dual-screening group and 23 (4.6 per 1000) in the chest radiography group (63). After prevalence screening,
114 subsequent (incident) cases of lung cancer were identified in the dual-screening group and 121 were identified in the annual radiography group during the screening period. Thirty-three and 32 cases, respectively, were diagnosed in the 2 years following screening. When the incidence and prevalence tumors are combined, 144 cases of lung cancer were detected in each group during the study period, 206 incident cases of lung cancer were identified in the dual-screening group and 160 were identified in the control group. After 20 years of follow-up, lung cancer death rates were 4.4 (95% CI, 3.9 to 4.9) and 3.9 (CI, 3.5 to 4.4) per 1000 person-years in the dual-screening and control groups, respectively (80).

The Mayo Lung Project was the first individually randomized, controlled trial to specifically evaluate the role of chest radiography in lung cancer screening. It was also the most influential in determining current public health policy. Although it is rated as fair quality by USPSTF criteria, the study has several limitations. First, prevalence screening involved 10933 male smokers age 45 years or older (39, 73–83). All participants underwent prevalence screening with sputum cytologic examinations and chest radiography, and 91 cases of cancer were identified (prevalence, 0.83%) (39, 73, 75). After prevalence screening, 4618 men were randomly assigned to a study group screened with chest radiography and pooled 3-day sputum cytologic examination every 4 months for 6 years; 4593 were assigned to a control group advised to have annual multiphasic health checkup, including CXR. Control group: 5557 received usual care and were screened over 5–8 y.

The mortality rate was 2.7 per 1000 person-years in both the chest radiography and dual-screening groups. In the Johns Hopkins Study, prevalence screening identified 39 malignant tumors in the dual-screening group and 40 in the chest radiography group (38, 71). After 8 years of follow-up, 194 incident cases of cancer were identified in the dual-screening group and 202 were identified in the chest radiography group. The mortality rates were 3.4 per 1000 person-years in the dual-screening group and 3.8 per 1000 person-years in the control group (not statistically significant differences) and were similar to community lung cancer mortality rates at the time (71, 72).

The first trial to evaluate the value of intense screening with chest radiography was the Mayo Lung Project, which...
detected 91 cases of lung cancer (0.83%). Thus, there was no completely unscreened control group. Also, these cases were followed separately and were not evaluated in the randomized comparison. Thus, any effect they had on mortality could not be determined. Second, nearly half of the controls obtained annual chest radiography during the study, and one third of the malignant tumors in the control group were discovered by screening chest radiography; 73% of controls received chest radiography during the study’s last 2 years. Third, adherence was 75% in the intervention group, reducing the study’s power (73).

The incidence of lung cancer in the Mayo Lung Project was approximately 22% higher in the intervention group than in the control group (73). Marcus and Prorok (81) evaluated the possibility of nonrandom distribution of lung cancer risk factors and found that distribution did not vary significantly between the intervention and control groups. Although little detailed information is provided, review of the Mayo Lung Project publications reveals evidence showing that not all patients were asymptomatic (39, 73). This could alter the findings of the screening study if patients with symptoms were disproportionately enrolled in the intervention group. However, there is no evidence to support this. The radiation exposure associated with chest radiography in the Mayo Lung Project is generally considered insufficient to increase lung cancer incidence (86). Finally, another possibility is that the higher incidence of lung cancer in the screened sample may represent the diagnosis of insignificant disease, that is, overdiagnosis.

**Case-Control Studies**

Five fair-quality case-control studies were conducted in Japan between 1992 and 2001 (42–46) (Table 3). Lung cancer was fatal in all participants (high-risk men and low- or unknown-risk women). All case-patients were matched to controls by age, sex, and health insurance status. Some studies included adjustment for geographic region, number of previous health examinations, or both, and all accounted for smoking by matching or statistical adjustment. For screening with chest radiography, with or without sputum cytologic examination within 1 year of diagnosis, the odds

<table>
<thead>
<tr>
<th>Prevalence</th>
<th>Patients with Incident Lung Cancer</th>
<th>Advanced Tumors (Stages III and IV)</th>
<th>Nonresectable Tumors</th>
<th>Mortality Rate per 1000 Person-Years†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>Intervention group: 31 (0.10)</td>
<td>Intervention group: 101</td>
<td>Intervention group: NR</td>
<td>Intervention group: 56</td>
<td>3-y follow-up</td>
</tr>
<tr>
<td>Control group: 20 (0.08)</td>
<td>Control group: 76</td>
<td>Control group: NR</td>
<td>Control group: 71</td>
<td>Intervention group: 0.7 Control group: 0.8</td>
</tr>
<tr>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>16-y follow-up</td>
</tr>
<tr>
<td>Dual-screening group: 30 (0.6)</td>
<td>Dual-screening group: 146</td>
<td>Dual-screening group: 64 (1.2) (incidence)</td>
<td>Dual-screening group: 49</td>
<td>5- to 8-y follow-up</td>
</tr>
<tr>
<td>CXR group: 23 (0.46)</td>
<td>CXR group: 155</td>
<td>CXR group: 63 (1.2) (incidence)</td>
<td>CXR group: 47</td>
<td>Dual-screening group: 2.7</td>
</tr>
<tr>
<td>Dual-screening group: 39 (0.75)</td>
<td>Dual-screening group: 194</td>
<td>NR</td>
<td>Dual-screening group: 53</td>
<td>5- to 8-y followup</td>
</tr>
<tr>
<td>CXR group: 40 (0.78)</td>
<td>CXR group: 202</td>
<td>CXR group: 56</td>
<td>CXR group: 56</td>
<td>Dual-screening group: 3.4</td>
</tr>
<tr>
<td>91 (0.83)</td>
<td>Dual-screening group: 206</td>
<td>Dual-screening group: 107 (2.3)</td>
<td>Dual-screening group: 32</td>
<td>20-y follow-up</td>
</tr>
<tr>
<td>Usual care group: 160</td>
<td>Usual care group: 109 (2.4)</td>
<td>Usual care group: 19</td>
<td>Usual care group: 3.9</td>
<td>Intervention group: 4.4</td>
</tr>
<tr>
<td>19 (0.30)</td>
<td>Dual-screening group: 108</td>
<td>Dual-screening group: 53 (1.7)</td>
<td>Dual-screening group: 77</td>
<td>15-y follow-up</td>
</tr>
<tr>
<td>Control group: 82</td>
<td>Control group: 46 (1.4)</td>
<td>Control group: 46</td>
<td>CXR group: 77</td>
<td>Dual screening group: 7.8 Control group: 6.8</td>
</tr>
<tr>
<td>Intervention group: 54</td>
<td>Intervention group: 320</td>
<td>NR</td>
<td>Intervention group: 72</td>
<td>10-y follow-up</td>
</tr>
<tr>
<td>Control group: 68</td>
<td>Control group: 599</td>
<td>Control group: 81</td>
<td>Control group: 0.8</td>
<td>Intervention group: 0.6</td>
</tr>
</tbody>
</table>

**Table 1—Continued**

www.annals.org
groups, have evaluated CT screening for lung cancer. The medically who had a median of 45 pack-years of smoking and no symptomatic volunteers (46% women) age 60 years or older Lung Cancer Action Project (54) involved 1000 asymptomatic persons with suspicious nodules, and diagnosis was confirmed, of which 26 were resectable and 23 were stage I (54). Four other cases of lung cancer were also diagnosed on the basis of non-nodule CT abnormalities. Approximately 1184 subsequent annual examinations resulted in further evaluation (usually high-resolution CT) in 40 persons (4%); biopsies in 9 persons; and lung cancer diagnoses in 9 persons (7.2 per 1000), 6 of which were stage IA (55). No mortality data are yet available on this cohort.

Lung Cancer Screening with Low-Dose CT

Several recent cohort studies, all without control groups, have evaluated CT screening for lung cancer. The details of these studies are shown in Table 4. The Early Lung Cancer Action Project (54) involved 1000 asymptomatic volunteers (46% women) age 60 years or older who had a median of 45 pack-years of smoking and no previous malignant disease. Participants were evaluated as medically fit for surgery and underwent chest radiography and CT. Baseline chest radiography identified 68 individuals with suspicious nodules, and diagnosis was confirmed by CT in 33. Seven patients had malignant nodules, all of which were resectable. Baseline CT identified 233 persons with nodules. After follow-up of 30 recommended biopsies, 27 malignant tumors were identified, of which 26 were resectable and 23 were stage I (54). Four other cases of lung cancer were also diagnosed on the basis of non-nodule CT abnormalities. Approximately 1184 subsequent annual examinations resulted in further evaluation (usually high-resolution CT) in 40 persons (4%); biopsies in 9 persons; and lung cancer diagnoses in 9 persons (7.2 per

Table 2. Methods and Quality of Controlled Trials of Lung Cancer Screening*

<table>
<thead>
<tr>
<th>Study (Reference)</th>
<th>Assembly of Comparable Groups: Randomization/Allocation Concealment</th>
<th>Maintenance of Comparable Groups</th>
<th>Outcomes Assessment: Validity of Method, Masking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northwest London Mass Radiography Service Study (35, 36)</td>
<td>Cluster randomization by random number; examiners not clearly blind; comparable in age structure and smoking habits; no apparent occupational exposures</td>
<td>99% follow-up</td>
<td>Cause of death determined from hospital records and General Register’s office; blinding not described</td>
</tr>
<tr>
<td>Kaiser Permanente Study (33, 34)</td>
<td>Randomization by patient record numbers with concealed code; more chronic lung disease in intervention group (8.9% vs. 7.5%)</td>
<td>Poor follow-up</td>
<td>Ill blind review of death</td>
</tr>
<tr>
<td>Memorial Sloan-Kettering Study (37, 63–67)</td>
<td>Computer-generated randomization (not described); similar all-cause mortality</td>
<td>Formal protocol/algorithm for follow-up; 55 lost to follow-up</td>
<td>All deaths reviewed by statisticians, clinicians, and pathologists who were blind to study group</td>
</tr>
<tr>
<td>Johns Hopkins Study (38, 68–72)</td>
<td>Computer-generated randomization (not described); allocation concealment unclear; fairly comparable when evaluated by age, smoking history, nontobacco carcinogen exposure</td>
<td>Formal algorithm for follow-up; 1.3% lost to follow-up</td>
<td>All deaths reviewed by statisticians, clinicians, and pathologists who were blind to study group</td>
</tr>
<tr>
<td>Mayo Lung Project (39, 74–76, 80)</td>
<td>Randomization method not described; allocation concealment unclear; similar distribution for age, smoking, exposure to nontobacco carcinogens, and pulmonary disease</td>
<td>Adequate; good follow-up of all participants in both groups</td>
<td>All deaths reviewed by statisticians, clinicians, and pathologists who were blind to study group</td>
</tr>
<tr>
<td>Czech Study (40, 84, 85)</td>
<td>Randomization stratified by age, smoking history, socioeconomic status, residence, occupational exposure; allocation concealment unclear; no differences observed in these characteristics; all-cause mortality rates, smoking-related deaths higher in intervention group</td>
<td>Not well reported</td>
<td>Cause of death ascertained from death certificates; autopsy performed in one third of patients; blind review not described</td>
</tr>
<tr>
<td>Wilde (47)</td>
<td>Nonrandomized; similar community distribution of smoking habits and economic structure; similar all-cause mortality rates; sample age not described</td>
<td>Adequate description; more dropouts in control group</td>
<td>Blinding not described; nonsystematic ascertainment of cause of death</td>
</tr>
</tbody>
</table>

* CXR = chest radiography; MHC = multiphasic health checkups; NR = not reported.
Among this cohort, 2916 annual incidence screening tests identified 336 individuals (12%) with new nodules, and 10 new diagnoses of lung cancer (6.7 per 1000) were made with CT alone. There were 2 cases of interval cancer and 2 cases of cancer diagnosed with sputum cytologic examination only. Of the 40 persons with malignant tumors, 36 were non–small-cell lung cancer; 31 (86%) were resected for cure. Eight patients had surgery for benign disease.

Finally, a German study (53) involving 817 asymptomatic volunteers age 40 years or older with at least 20 pack-years of smoking was conducted between November 1995 and July 1999. Baseline CT identified 350 persons with nodules. Of these, 269 underwent high-resolution CT, and nodules were ultimately identified in 29 persons. Thirteen of these 29 had biopsies; malignant disease was diagnosed in 10, and 1 case of interval cancer was also diagnosed. After an average of 2.7 years of follow-up, 6 patients are alive without evidence of recurrence.

Lung Cancer Screening among Women

Lung cancer is the leading cause of cancer-related death among women in the United States, and most cases are attributed to smoking (2). In addition, women have substantial exposure to passive smoking, which is thought to cause a significant proportion of lung cancer in non-smoking women (18). Although controversial, some studies suggest that for any level of smoking, women are at higher risk for lung cancer than men (4, 87, 88). For unknown reasons, women also tend to develop adenocarcinoma of the lung disproportionately to men (17, 88, 89), and adenocarcinoma is also found more commonly among nonsmokers (17). This cell type tends to occur peripherally (89, 90) and may be more apt to be detected with chest radiography, CT, or both than other cell types. Consequently, radiologic imaging and screening for lung cancer may perform differently among women. Unfortunately, no randomized trials of lung cancer screening have included women. The only data evaluating screening among women and including controls come from 4 Japanese case–control studies involving primarily nonsmoking women (passive smoking was not assessed) (43–46). These studies, which are summarized in Table 5, showed that lung cancer mortality odds ratios or relative risks for screening conducted within 12 months of lung cancer diagnosis ranged from 0.39 to 0.61; 2 studies found statistically significant differences. However, interpretation of these studies is limited.
by the screening biases discussed in this review. Five studies of CT have included women, but mortality data are not yet available. Randomized trials of lung cancer screening with chest radiography, low-dose CT, or both involving women are under way.

**DISCUSSION**

The personal and public health importance of lung cancer in the United States and worldwide is enormous, and even a small benefit associated with screening could save many lives. However, the outcomes of screening, as shown in this report, are mixed. Some lower-quality evidence evaluating chest radiography with or without sputum cytologic examination (case–control studies) has shown benefit, and higher-quality evidence (randomized, controlled trials) has not. Studies show that lung cancer can be diagnosed at an earlier stage with CT screening than in usual clinical practice, but little is known about patient

---

**Table 3. Case–Control Studies of Lung Cancer Screening with Chest Radiography and Lung Cancer Mortality Rates**

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Setting</th>
<th>Case-Patients with Fatal Lung Cancer</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ebeling and Nischan, 1987 (41)</td>
<td>Berlin</td>
<td>130 men age &lt;70 y</td>
<td>204 patients from community center</td>
</tr>
<tr>
<td>Okamoto et al., 1999 (42)</td>
<td>Japan</td>
<td>158 men and 35 women age 40–74 y</td>
<td>579*</td>
</tr>
<tr>
<td>Sobue, 2000 (43)</td>
<td>Japan</td>
<td>208 high-risk men, 65 low-risk women</td>
<td>1269*</td>
</tr>
<tr>
<td>Sagawa et al., 2001 (44)</td>
<td>Japan</td>
<td>258 smoking and nonsmoking men and 70 nonsmoking women age &gt;39 y</td>
<td>1886*</td>
</tr>
<tr>
<td>Tsukada et al., 2001 (45)</td>
<td>Japan</td>
<td>149 high-risk men and 25 non–high-risk (nonsmoking) women age &gt;40 y</td>
<td>801*</td>
</tr>
<tr>
<td>Nishii et al., 2001 (46)</td>
<td>Japan</td>
<td>412 men and women age 40–79 y</td>
<td>3490*</td>
</tr>
</tbody>
</table>

* All matched by age, sex, and location.
† Received a poor score because selection of controls was potentially biased.
‡ Received a poor score for not controlling for smoking.
§ High-risk individuals were also screened with sputum cytologic examination.
Excluding screening at <12 mo.

---

**Figure.** Mortality in randomized, controlled trials of lung cancer screening with chest radiography with or without sputum cytologic examination.
outcomes. Unfortunately, none of the existing randomized trials answer the question faced by clinicians: Should patients be screened for lung cancer at all? The case-control studies from Japan give some support to screening for lung cancer with chest radiography. Although case-control studies are not considered the gold standard in evaluating screening efficacy and effectiveness, several authors believe they can be a useful and efficient way to evaluate a screening method (31, 91, 92). However, it is very difficult to overcome the possibility of volunteer or healthy screenee bias in case-control studies, even well-conducted ones. This might bias such studies toward ben-

### Table 4. Outcomes of Lung Cancer Screening with Low-Dose Computed Tomography*

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Screening Interval</th>
<th>Screening Type</th>
<th>Screening Tests Performed</th>
<th>Positive Test Results</th>
<th>Referral Biopsy</th>
<th>Study Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diederich et al., 2002 (53)</td>
<td>12 mo</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>350 (43)</td>
<td>269 29 13 1 (1)</td>
<td>11 (1.3) (1 interval)</td>
<td>Poor†‡</td>
</tr>
<tr>
<td>Henschke et al., 1999, 2001 (54, 55)</td>
<td>6-18</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>1000</td>
<td>237 (24) 104 27 0</td>
<td>31 (3.1)</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>12 mo</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>1184</td>
<td>40 9 NR 0</td>
<td>9 (0.9)</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>12 mo</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>1000</td>
<td>68 (6.8) 33 NR 0</td>
<td>7 (0.7)</td>
<td>66</td>
</tr>
<tr>
<td>Nawa et al., 2002 (56)</td>
<td>12</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>7956</td>
<td>541 64 NR NR 36 (0.5)</td>
<td>4 (0.1)</td>
<td>86</td>
</tr>
<tr>
<td>Sone et al., 2001 (58)</td>
<td>12</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>5568</td>
<td>148 7 NR NR</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Sobue et al., 2002 (59)</td>
<td>6</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>8301†</td>
<td>279 (5.1) 297 NR NR 37 (0.6)</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>24 mo</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>1611</td>
<td>186 (11.5) 25 21 0</td>
<td>13 (0.8)</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>24 mo</td>
<td>Baseline LDCT Incidence LDCT</td>
<td>7891</td>
<td>721 57 35 1 (0)</td>
<td>19 (0.2)</td>
<td>79</td>
</tr>
<tr>
<td>Swensen et al., 2002, 2003 (61, 62)</td>
<td>12</td>
<td>Baseline LDCT Combined data Incidence LDCT</td>
<td>1520</td>
<td>782 (51.4) NR NR NR</td>
<td>27 (1.8)‡</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>24 mo</td>
<td>Baseline LDCT Combined data Incidence LDCT</td>
<td>2916</td>
<td>336 NR NR NR</td>
<td>11 (0.7) (+2 interval)‡</td>
<td>66</td>
</tr>
</tbody>
</table>

* All data are presented by individual except incidence, which refers to screening tests performed. CXR = chest radiography; HRCT = high-resolution computed tomography; LDCT = low-dose computed tomography; NR = not reported.
† Percentage of lung cancer for incidence = cases of lung cancer identified with incidence screening/cases of lung cancer in cohort minus cases of prevalence cancer.
‡ 1 case of malignant disease diagnosed with sputum cytologic examination only.
Lung Cancer Screening Studies Including Women*

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Study Type</th>
<th>Setting</th>
<th>Description of Sample</th>
<th>Intervention</th>
<th>Odds Ratio or Relative Risk of Lung Cancer Death (95% CI) or Number of Malignant Tumors Identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobue, 2000 (43)</td>
<td>Case-control</td>
<td>Japan</td>
<td>65 low-risk patients</td>
<td>CXR with or without sputum cytologic examination</td>
<td>0.42 (0.20–0.87) for screening &lt;12 mo</td>
</tr>
<tr>
<td>Sagawa et al., 2001 (44)</td>
<td>Case-control</td>
<td>Japan</td>
<td>70 low-risk patients age &gt;39 y</td>
<td>CXR with or without sputum cytologic examination</td>
<td>0.57 (0.30–1.11) for screening &lt;12 mo</td>
</tr>
<tr>
<td>Tsukada et al., 2001 (45)</td>
<td>Case-control</td>
<td>Japan</td>
<td>25 low-risk patients age &gt;40 y</td>
<td>CXR with or without sputum cytologic examination</td>
<td>0.61 (0.23–1.68) for screening &lt;12 mo</td>
</tr>
<tr>
<td>Nishi et al., 2001 (46)</td>
<td>Case-control</td>
<td>Japan</td>
<td>412 mixed-risk patients age 40–79 y</td>
<td>CXR</td>
<td>0.39 (0.24–0.64) for screening &lt;12 mo</td>
</tr>
<tr>
<td>Henschke et al., 1999, 2001 (54, 55)</td>
<td>Cohort</td>
<td>United States</td>
<td>460 high-risk participants</td>
<td>Baseline LDCT and repeated LDCT</td>
<td>NR by sex</td>
</tr>
<tr>
<td>Sone et al., 2001 (58)</td>
<td>Cohort</td>
<td>Japan</td>
<td>2512 participants</td>
<td>Baseline LDCT</td>
<td>11 malignant tumors identified</td>
</tr>
<tr>
<td>Diederich et al., 2002 (53)</td>
<td>Cohort</td>
<td>Germany</td>
<td>1816 participants</td>
<td>Repeated LDCT</td>
<td>4 malignant tumors identified</td>
</tr>
<tr>
<td>Nawa et al., 2002 (56)</td>
<td>Cohort</td>
<td>Japan</td>
<td>1367 participants (4.3% current or former smokers)</td>
<td>Baseline LDCT</td>
<td>12 cases of malignant disease identified, all in nonsmokers</td>
</tr>
<tr>
<td>Swensen et al., 2002, 2003 (61, 62)</td>
<td>Cohort</td>
<td>United States</td>
<td>735 participants</td>
<td>Annual repeated LDCT</td>
<td>0</td>
</tr>
</tbody>
</table>

* CXR = chest radiography; LDCT = low-dose computed tomography; NR = not reported.

The hope of benefit from lung cancer screening is high. However, the implications of screening, especially in the absence of proven benefit, are also great. Evaluating harm or potential harm associated with screening for lung cancer is difficult. One approach to this issue is to evaluate the 4 possible outcomes of screening: false-positive, false-negative, true-positive, and true-negative findings. The best data about outcomes from chest radiography screening come from the recent CT studies, since data from the chest radiography trials precede the use of CT for evaluation of radiography abnormalities, and more patients had previous thoracotomy or biopsy than would have in current clinical practice. Table 4 shows positive chest radiography rates and the diagnostic outcomes associated with chest radiography from the CT studies. Most abnormalities on chest radiography are resolved or screening results are found to be false-positive when evaluated by CT (54, 59). For radiographs identified as suspicious for cancer in the National Cancer Institute studies, the positive predictive value ranged from 41% to 60% (29).

In the CT studies, the false-positive rate was the number of patients who required further evaluation after CT but did not have cancer. When this criterion was used, the false-positive rates in the CT studies ranged from 5% to 50% in prevalence screening and 3% to 12% in incidence screening; most abnormalities were resolved with high-resolution CT. Among the CT studies reporting referral rates, 4.8% to 14.5% of patients undergoing high-resolution CT were referred for biopsy, and most (63% to 90%) then received a diagnosis of cancer (Table 4). For comparison, in U.S. and European clinical practices, approximately half of patients undergoing surgical biopsy of indeterminate nodules subsequently receive a benign diagnosis (61, 94). In the current practice setting, positron emission tomography is commonly used as a noninvasive means of discriminating between malignant and nonmalignant lesions (95) and may reduce the rate of invasive procedures performed to evaluate indeterminate nodules.

Persons with false-positive results can experience high anxiety and concern, and those pursuing further evaluation experience associated cost and risk. Although the false-positive rate is high in the lung cancer screening studies, false-positive results on a lung cancer screening study (either chest radiography or CT) may have a different effect on patients than false-positive results on other types of cancer screening tests. Patients who smoke potentially have some control over their subsequent risk and may be able to more effectively modify their high-risk behavior. Data from the Early Lung Cancer Action Project suggest that CT scan results, in combination with smoking cessation counseling, improved smoking cessation rates among all participants.
An important and controversial issue in lung cancer screening is the question of overdiagnosis and consequent overtreatment. The relatively high prevalence of unrecognized lung cancer in several studies suggests that there is a significant preclinical pool of lung cancer in high-risk populations (38, 54, 97). Whether all of these tumors would eventually present clinically is uncertain. Supporting overdiagnosis are data from the Mayo Lung Project showing increased rates of early tumors in the screened group compared with the control group without a change in the number of advanced tumors or subsequent mortality rates. These findings suggest diagnosis of a pool of indolent tumors (98). Although the higher lung cancer mortality rate among the intervention group in the Mayo Lung Project was not statistically significant, a major concern is that the increase in mortality might not be due to chance and may be a consequence of screening (that is, more persons in the screened group were evaluated and treated, which, with treatment-associated risk, resulted in a true increase in mortality rates). Alternatively, an increase in lung cancer mortality rates among screened individuals may be a consequence of misclassification of cause of death or “sticky-diagnosis bias” (98), meaning that in the absence of autopsy data, there is a propensity to label any diagnosed malignant condition as the cause of death regardless of its clinical course. This results in bias against screening in evaluations of disease-specific mortality (99). Black and colleagues (100) noted that the excess lung cancer mortality, particularly death from metastatic adenocarcinoma, observed among the screened group in the Mayo Lung Project was probably at least partially a consequence of this type of differential misclassification.

Arguments against an important role for overdiagnosis in lung cancer are based on autopsy studies showing low rates (0.8%) of unrecognized lung cancer (101). Whether autopsy data are generalizable to living persons is questionable, particularly given selection biases for autopsy. Further data against overdiagnosis come from 2 natural history studies of screening- and symptom-detected unresected stage I non–small-cell lung cancer, which showed that almost all patients with lung cancer die of the disease over 5 to 10 years (25, 26). Whether a strong case for overdiagnosis should be made on the basis of current data is uncertain. However, it is possible that with an increasingly sensitive detection tool, such as CT, overdiagnosis may occur. The issue of overdiagnosis is particularly relevant to the harm associated with lung resection for cancer, which involves significant associated mortality and morbidity. More data are needed to definitively evaluate this issue.

Another potential harm of screening is false-negative findings and possible false reassurance. In current practice, the best estimate of the rate of false-negative results on chest radiography comes from the CT studies, where false-negative rates as high as 75% were shown (54, 59). Clinical series of chest radiography suggest that retrospective identification of lung cancer ranges from 12% to 90% (102, 103). While CT is considered the gold standard for evaluating nodules, it has also been shown to yield false-negative results (62). The potential for false reassurance with CT certainly exists, particularly if those screened believe that they are undergoing a definitive examination.

The rate of biopsy-associated complications was not described in the CT studies. The morbidity and mortality associated with thoracotomy for positive test results (true or false) are also difficult to evaluate. Studies of symptomatic patients suggest that morbidity and mortality are directly related to the amount of lung tissue removed. Overall, mortality rates ranged from 1.3% to 11.6% and morbidity rates ranged from 8.8% to 44% among several series reviewed. Rates are lower among patients undergoing smaller resections, those with fewer comorbid conditions, and those treated at centers with greater surgical volume (28, 47, 104–110). Complication rates from studies of symptomatic patients are likely to be greater than complication rates among asymptomatic individuals in screening programs directed at those judged healthy enough for surgery.

Currently, most patients in the United States are not screened for lung cancer (111). However, because conclusions about lung cancer screening have been based on limited data and no trials have compared screening with no screening or screening among women, the issue is being reevaluated. Routine annual chest radiography is being compared with usual care in the Prostate, Lung, Ovarian, and Colorectal Cancer Trial, which involves more than 100 000 men and women age 55 to 74 years (112, 113). Data from this study should be available in 2010. The National Lung Screening Trial will compare routine screening CT with chest radiography in high-risk men and women age 55 to 74 years (114).

New technologies may also contribute to the early detection of and possibly screening for lung cancer. Some currently being investigated include immunocytochemical analysis of sputum with monoclonal antibodies (115), identification of genetic mutations (116), abnormal DNA methylation (117, 118), abnormal patterns of immunostaining, and other molecular changes (119–122). Several other potential targets in sputum, bronchial fluid, and expired air may have a role in early lung cancer detection and are currently being investigated (123, 124).

In summary, studies evaluating chest radiography screening for lung cancer have had mixed findings. Stronger evidence from 30-year-old trials suggest no benefit among male smokers and possible overdiagnosis, and weaker study designs suggest benefit to men and women. There are important methodologic limitations to all of these studies. The studies of CT have demonstrated that...
l lung cancer can be diagnosed at a significantly earlier stage with CT screening than in current clinical practice. However, whether this will translate to a mortality benefit is unclear. In addition, even if CT is shown to be effective, the issue of cost-effectiveness remains (125). Critical information will come from the current randomized, controlled trials of screening CT. Given the uncertainty associated with chest radiography screening, it is unfortunate that the National Lung Screening Trial does not include nonscreened control groups. However, data on chest radiography screening will be available from the Prostate, Lung, Ovarian, and Colorectal Cancer Trial in the next 5 to 8 years. In the meantime, other approaches for evaluation of screening should be considered, such as rigorously conducted case-control studies of chest radiography, screening CT, or both. We hope that new methods of screening for lung cancer will be developed and refined. Even a small decrease in lung cancer mortality from screening would save thousands of lives each year.

From Oregon Health & Science University Evidence-based Practice Center and Portland Veterans Affairs Medical Center, Portland, Oregon; Merck & Co., West Point, Pennsylvania; and University of Medicine and Dentistry of New Jersey Medical School, Newark, New Jersey.

Disclaimer: The authors of this article are responsible for its contents, including any clinical or treatment recommendations. No statement in this article should be construed as an official position of the Agency for Healthcare Research and Quality, the U.S. Department of Health and Human Services, or Merck & Co., Inc.

Acknowledgments: The authors thank Kathryn Krages, AMLS, MA; Susan Wingenfeld; and Kim Peterson, MS, for their help in preparation of the full evidence report and the manuscript. They also thank Mark Helfand, MD, MPH; William Holden, MD; John McAnulty, MD; and James Reuler, MD, for their helpful reviews of the manuscript.

Grant Support: This study was conducted by the Oregon Health & Science University Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality (contract no. 290-97-0018, Task Order 2, Rockville, Maryland).

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Reprints are available from the Agency for Healthcare Research and Quality Web site (www.preventiveservices.ahrq.gov) and through the Agency for Healthcare Research and Quality Publications Clearinghouse (telephone, 800-358-9295).

Current author addresses are available at www.annals.org.

References
7. Acknowledgments: The authors thank Kathryn Krages, AMLS, MA; Susan Wingenfeld; and Kim Peterson, MS, for their help in preparation of the manuscript.
27. Snijder RJ, Bruntel de la Riviere A, Elbers HJ, van den Bosch JM. Survival in resected stage I lung cancer with residual tumor at the bronchial resection margin.


APPENDIX. U.S. PREVENTIVE SERVICES TASK FORCE
QUALITY RATING CRITERIA

For randomized, controlled trials, the criteria are as follows.
1. Initial assembly of comparable groups: adequate randomization, including first concealment and whether potential confounders were distributed equally among groups.
2. Maintenance of comparable groups (includes attrition, crossovers, adherence, contamination).
3. Levels of follow-up: differential loss between groups; overall loss to follow-up.
4. Measurements: equal, reliable, and valid, including masking of outcome assessment.
5. Clear definition of interventions.
6. Important outcomes considered.
7. Analysis: intention to treat.

Definition of ratings are as follows, based on these criteria.
A Good study meets all criteria: comparable groups are assembled initially and maintained throughout the study; follow-up is at least 80%; reliable and valid measurement instruments are applied equally to the groups; interventions are clearly defined; important outcomes are considered; and appropriate attention is paid to confounders in the analysis. In addition, for randomized, controlled trials, intention-to-treat analysis is used.

In a Fair study, comparable groups are generally assembled initially, but some question remains whether some (although not major) differences occurred in follow-up; measurement instruments are acceptable (although not the best) and are generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention-to-treat analysis is done for randomized, controlled trials.

In a Poor study, groups assembled initially are not close to being comparable or maintained throughout the study; measurement instruments are unreliable or invalid or not applied at all equally among groups; outcome assessment is not masked; and key confounders are given little or no attention. For randomized, controlled trials, intention-to-treat analysis is not performed.

For case–control studies, the criteria are as follows.
1. Accurate ascertainment of case-patients.
2. Nonbiased selection of case-patients and controls with exclusion criteria applied equally to case-patients and controls; response rate is at least 80%; diagnostic procedures and measurements are accurate and are applied equally to case-patients and controls; and appropriate attention is paid to confounding variables.

In a Fair study, there is appropriate ascertainment of case-patients and controls and exclusion criteria is applied equally to case-patients and controls; response rate is at least 80%; diagnostic procedures and measurements are accurate and are applied equally to case-patients and controls; and appropriate attention is paid to confounding variables.

In a Poor study, there is major selection or diagnostic work-up bias; response rates are less than 50%; or no attention is paid to confounding variables.

Current Author Addresses: Dr. Humphrey: Oregon Health & Science University, Mailcode BICC, 3181 SW Sam Jackson Park Road, Portland, OR 97239-3098.
Dr. Teutsch: Merck & Co., Inc., Sumneytown Pike and Broad Street, PO Box 4, West Point, PA 19437.
Dr. Johnson: Department of Family Medicine, University of Medicine and Dentistry–New Jersey Medical School, 185 South Orange Avenue, MSB-B648, Newark, NJ 07103.
### Appendix Table 2. Cohort Studies of Lung Cancer Screening with Chest Radiography*

<table>
<thead>
<tr>
<th>Study, Year (Reference)</th>
<th>Study Sample</th>
<th>Intervention</th>
<th>Malignant Disease</th>
<th>Resectable Tumors</th>
<th>Survival Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Philadelphia Neoplasm Research Project, 1951 (48, 49)</td>
<td>6136 men age ≥45 y</td>
<td>Photofluorography and questionnaires every 6 mo for 10 y</td>
<td>Prevalence: 84 (1.37)</td>
<td>35</td>
<td>8 (5 y)</td>
</tr>
<tr>
<td>Tokyo Metropolitan Government Study, 1953 (51)</td>
<td>1 871 374 men and women, all ages</td>
<td>Intermittent CXR over 26 y (sputum cytologic examination in some)</td>
<td>Incidence: 121 193 (0.01)</td>
<td>56</td>
<td>44 (5 y for resectable tumors) (usual survival at that time, 20)</td>
</tr>
<tr>
<td>Veterans Administration Trial, 1958 (50)</td>
<td>141 607 men; median age, 62.8 y</td>
<td>CXR and sputum cytologic examination</td>
<td>73 (0.052)</td>
<td>36</td>
<td>17 (32 mo)</td>
</tr>
<tr>
<td>South London Cancer Study, 1959 (52)</td>
<td>67 400 men age ≥45 y</td>
<td>CXR every 6 mo</td>
<td>234 (0.35)</td>
<td>56</td>
<td>18% (4 y) (usual survival at that time, 9)</td>
</tr>
</tbody>
</table>

* CXR = chest radiography.