Screening for Lung Cancer*
A Review of the Current Literature

Peter B. Bach, MD; Michael J. Kelley, MD; Ramsey C. Tate, BS; and Douglas C. McCrory, MD, MHS

Study objectives: To review the available data on the early detection of lung cancer, with a focus on three technologies: chest x-ray (CXR), sputum cytology, and low-dose CT (LDCT) scanning.

Design, setting, participants: Review of published clinical studies of early detection technologies. The best available evidence on each topic was selected for analysis. Randomized trials were used to evaluate CXR and sputum cytology. Cohort studies, as well as studies providing evidence regarding rates of overdiagnosis and efficacy of initial treatment, were considered in evaluation of LDCT. Study design and results were summarized in evidence tables. Statistical analyses of combined data were not performed.

Measurement and results: Five randomized trials of CXR with or without sputum cytology have been conducted, each which reports disease-specific mortality as well as other end points. None of these studies provide support for the use of either CXR or sputum cytology for the early detection of lung cancer in asymptomatic individuals. Eight completed and ongoing trials of LDCT were identified. All of these studies report the frequency and stage distribution of lung cancers found during initial (“prevalence”) screening, and several studies also report rates of detection at the time of annual follow-up. No outcome data on survival or treatment are available. A number of studies support the hypothesis of “overdiagnosis”—that some lung cancers detected by LDCT may behave in an indolent manner.

Conclusions: The use of either CXR or sputum cytology for the early detection of lung cancer is not supported by the published evidence. The evidence for LDCT appears promising, in that the technology typically identifies lung cancer at an early stage, although corollary studies suggest that these findings in isolation may be misleading. Further high-quality research is needed to better define the role of LDCT in the evaluation of asymptomatic high-risk individuals.

Key words: evidence-based medicine; lung neoplasms; mass chest x-ray; mass screening; review; sputum; tomography, x-ray computed

Abbreviations: CXR = chest x-ray; ELCAP = Early Lung Cancer Action Project; LDCT = low-dose CT; MSKCC = Memorial Sloan-Kettering Cancer Center; NCI = National Cancer Institute; RCT = randomized controlled trial

Lung cancer accounts for approximately 6% of all deaths in the United States each year, and is the leading cause of cancer death for both men and women. At present, most patients who receive an initial diagnosis of lung cancer have advanced stage disease, making cure with currently available therapies unlikely. In contrast, individuals with early stage disease can achieve cure through surgical resection. Because of this dichotomy in outcome associated with stage at diagnosis, there has been persistent interest in designing and testing methods for early detection of lung cancer. To place these studies in context, we articulate the critical elements of successful screening tests and describe the research methods that can be used to evaluate them. We then focus on lung cancer screening studies with chest x-ray (CXR), sputum cytology, and low-dose CT (LDCT) scanning. For the first two modalities, we review published randomized controlled trials (RCTs). For LDCT, we review published and ongoing studies, discuss the evidence that has fueled the recent debate over efficacy, and discuss future plans of assessment.

Elements of a Beneficial Screening Test
The benefit of a screening test is typically assessed in two domains. First, the screening test must pro-
vide benefit to individuals who have the illness, typically through increasing life expectancy. In order to prolong life expectancy, the test must be capable of detecting the disease at a point in its natural history at which the course can be altered through treatment, typically earlier than is encountered in sporadic practice. Patients survive longer after diagnosis of early stage lung cancer than they do after diagnosis of more advanced stage lung cancer; therefore, it is widely believed that early detection followed by definitive treatment should alter the natural history of the disease and thereby reduce lung cancer mortality.

Second, the screening test should not be dangerous or painful, nor should it have numerous false-positive results that cause anxiety or necessitate invasive or dangerous follow-up tests. From a societal perspective, the screening test should also not be harmful to the large fraction of the population who do not have the disease, either by consuming vast amounts of resources or by directly affecting the capacity of the health-care system to provide for others.

**Research Methods for Evaluating Screening Tests**

Three general methods have been advocated for the evaluation of screening tests: randomized trials of screening, population-based studies of screening, and observational studies of screening in select cohorts. Traditionally, the scientific community ranks RCTs at the top of this hierarchy of study designs. In a randomized trial, individuals are subjected to different intensities of screening, and the outcome that is observed is disease-specific mortality. The efficacy of fecal-occult blood testing for the detection of colon cancer was established in this manner.2–4 As another example, the Mayo Lung Project randomized individuals to receive regularly scheduled screening (intervention arm) or routine care; the primary outcome of interest was lung cancer mortality.5–7 An alternative highly indicative design is the population-based study. In studies of this design, the impact of a broadly implemented screening program is assessed through changes in disease-specific mortality rates within the population. The efficacy of cervical cancer screening was established through multiple studies that documented large declines in the rates of cervical cancer mortality in the 1960s and 1970s in Sweden, Finland, and a number of other countries.5,9 In the third design—observational studies of screening in select cohorts—the efficacy of the screening test is inferred from the frequency with which it can detect early stage cancers. The subsequent impact on disease-specific mortality is then ascertained from historical data on the outcomes of individuals with early stage disease, or through the documentation of survival for the individuals enrolled in the study.

Readers should be aware that all of these study designs have potential weaknesses. Negative results of randomized trials may be viewed with skepticism either because of defects in design (such as inadequate power), or defects in study conduct (such as having excessive crossover between arms). For example, despite seven separate RCTs, the efficacy of mammography for the prevention of death from breast cancer remains uncertain. Some investigators believe that significant design flaws in the majority of mammography studies limit the ability of those studies to measure the impact of mammography. These investigators have pointed out that the best designed studies show no discernible impact on overall mortality,10,11 an assertion that has led an independent panel of US National Cancer Institute (NCI) advisers to revise their information summaries about screening with mammography.12,13 Similarly, population-based studies can be confounded by numerous factors, including the implementation of a screening program itself. For example, there is a substantial amount of evidence suggesting that the act of screening for prostate cancer with the prostate-specific antigen test may artifactually raise incidence and mortality rates from that disease.14,15 A study from Black et al16 suggested that screening interventions in RCTs may also influence the ascertainment of cause of death. Observational studies are subject to particular biases, including those resulting from “lead time” and “length time,” and those resulting from the recruitment and selection of volunteer cohorts.

**Sputum Cytology and CXR**

Sputum cytology and CXR have each been assessed in RCTs. The first study was begun in 1960, in which 55,034 London men were randomized to CXR every 6 months for 3 years, or to CXR at the beginning and the end of the 3-year period.17,18 Follow-up exceeded 99% in both study arms. In this study, a total of 36 additional cases of lung cancer were identified in the frequently screened group (132 cases) than in the group screened only at the study inception and study completion (96 cases). Of the cancers detected during the 3-year study period, 44% were resected in the screened group, while only 29% were resected in the control group. During the 3-year study period, lung cancer mortality was similar: 62 individuals in the frequently screened group died
of lung cancer, compared to 59 individuals in the less frequently screened group. This study, as well as several observational studies, did not support the hypothesis that CXR screening would be associated with a mortality benefit, and widespread screening was not pursued.19,20 Improvements in CXR and sputum cytology technologies led the NCI to readdress this question in the 1970s. The NCI funded the Cooperative Early Lung Cancer Group,21 which conducted three randomized studies of sputum cytologic examination and CXR. During the same time period, a randomized study of CXR and sputum cytology was conducted in Czechoslovakia.22

Two of the NCI-funded RCTs (the Johns Hopkins Lung Project and the Memorial Sloan-Kettering Cancer Center [MSKCC] Lung Cancer Screening Program) randomized subjects to either CXR alone or CXR plus sputum cytology. In neither study were there differences in the number of cancers detected, the fraction of cancers that were “resectable,” or the lung cancer mortality rate (Table 1).23–25 These studies were primarily designed to assess the incremental impact of sputum cytology in a CXR-screened cohort. Thus, the absence of a mortality benefit in these studies may have been due to the insensitivity of sputum cytology as it was implemented.

The NCI-funded Mayo Lung Project focused on the combined impact of CXR and sputum cytology for screening.5–7 This study recruited male smokers > 45 years old between 1971 and 1976 who had at least one pack-per-day cigarette use, an estimated survival of at least 5 years, no evidence of lung cancer on an initial evaluation, and sufficient lung function to tolerate lobectomy. Among the 10,933 subjects initially recruited, 91 prevalent cases of lung cancer were detected on screening studies (0.83%), 51 cases by CXR, 17 cases by sputum cytology alone, and 15 cases by a combination of both tests. Subsequently, individual subjects with negative and satisfactory prevalence screens were randomized to one of two groups. The “screening” group had CXRs and sputum cytology every 4 months for 6 years; the “control” group received a recommendation at the start of the study to receive a yearly CXR and sputum cytologic examination (that being the standard Mayo Clinic recommendation at that time). Subjects were then followed up from 1 to 5.5 years (median, 3 years). Compliance in the screened group was 85% during the first year and fell to 75% by the end of the study. More than one half of the control group also underwent CXR during the study period.

Lung cancer was detected more often in the screened group (206 cases) than in the control group (160 cases; Table 1). Forty-eight percent of the cancers detected in the screened group were localized, and thus amenable to surgical resection with curative intent, while only 32% of the cancers detected in the control group were localized, suggesting that screening may increase the cure rate. However, at the first analysis, there were no substantive differences in lung cancer-specific mortality rates: 3.2 per 1,000 person-years in the screened group and 3.0 per 1,000 person-years in the control group. All-cause mortality was also similar (Table 1). The investigators of the Mayo Lung Project therefore concluded that the combination of CXR and sputum cytology performed every 4 months does not decrease lung cancer mortality compared to the recommendation of annual screening.6 An analysis of the same cohort performed 15 years after study completion yielded a similar finding: no substantive difference in lung cancer mortality rate between the screened group (4.4 per 1,000 person-years) and the control group (3.9 per 1,000 person-years).6,26

The contemporaneous study conducted in Czechoslovakia randomly assigned individuals to CXR and sputum cytology every 6 months or to initial screening followed by screening again after 3 years had elapsed. Subjects in both arms were then screened annually for an additional 3 years. The results of this study at 3 years were similar to the Mayo study; a greater number of cancers had been detected in the intervention group than the control group (39 cases vs 27 cases), including a greater number of cases with early stage disease (20 cases vs 10 cases; Table 1). However, there was effectively no difference between the number of individuals in the screened group and the control group with late-stage cancer (19 patients vs 17 patients), nor was there a difference in the number of patients dying of lung cancer.22,27

**Conclusions: CXR and Sputum Cytology**

The results of the five RCTs suggest that neither CXR nor sputum cytology satisfy the primary criterion of a beneficial screening test. Neither test appears to prolong the life expectancy of an individual with the disease. Whether either test fits the second criterion—that the test is not particularly painful or harmful—was not addressed in sufficient detail in any of these studies to determine.

Readers should appreciate that despite the consistency of the findings in these studies, there is still skepticism about their interpretation. One observa-
<table>
<thead>
<tr>
<th>Study Site</th>
<th>Study Arm</th>
<th>Intervention</th>
<th>Sample Size</th>
<th>Baseline Screen</th>
<th>Repeat Screening</th>
<th>Long Cancer Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cancers, No.</td>
<td>Per 1,000 Persons</td>
<td>Cancer, No.</td>
</tr>
<tr>
<td>London 1960–1964†</td>
<td>All</td>
<td>Intervention CXR every 6 mo for 3 yr</td>
<td>55,034</td>
<td>51</td>
<td>0.9</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control CXR at end of 3 yr</td>
<td>25,311</td>
<td>20</td>
<td>0.8</td>
<td>76</td>
</tr>
<tr>
<td>Mayo 1971–1983†</td>
<td>All</td>
<td>Intervention CXR and sputum cytology every 4 mo for 6 yr</td>
<td>10,933</td>
<td>91</td>
<td>8.3</td>
<td>366</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control Advised to have CXR and sputum cytology annually</td>
<td>4,593</td>
<td>NA</td>
<td>NA</td>
<td>160</td>
</tr>
<tr>
<td>Czechoslovakia 1976–1980†</td>
<td>All</td>
<td>Intervention CXR and sputum cytology every 6 mo for 3 yr</td>
<td>6,364</td>
<td>18</td>
<td>2.8</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control CXR and sputum cytology at end of 3 yr</td>
<td>3,174</td>
<td>NA</td>
<td>NA</td>
<td>27</td>
</tr>
<tr>
<td>MSKCC 1974–1982†</td>
<td>All</td>
<td>Intervention CXR annually and sputum cytology every 4 mo</td>
<td>10,040</td>
<td>53</td>
<td>5.3</td>
<td>235</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control CXR annually</td>
<td>5,072</td>
<td>23</td>
<td>4.5</td>
<td>121</td>
</tr>
<tr>
<td>Johns Hopkins 1973–1982†</td>
<td>All</td>
<td>Intervention CXR annually and sputum cytology every 4 mo</td>
<td>10,386</td>
<td>79</td>
<td>7.6</td>
<td>396†</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control CXR annually</td>
<td>5,226</td>
<td>39</td>
<td>7.5</td>
<td>194</td>
</tr>
</tbody>
</table>

*NA = not available; NR = not reported.
†Randomization occurred subsequent to baseline screening; therefore, the sample sizes of the study arms do not sum to the number of total enrollees.
‡Randomization occurred prior to baseline screening; therefore, the total number of deaths may include individuals found to have lung cancer during the baseline screen.
§Includes 379 cancer detected during screening period and 17 cancers detected after the end of screening.
tion, initially provided by the International Union Against Cancer,\textsuperscript{28} is that many of these studies may be invalid because they did not include a “no-screening” study arm and thus no determination of true efficacy could be made. Other authors have argued that the sample size of these studies was inadequate. Both the Mayo Lung Project and the Czechoslovakian studies were powered to detect a 50% reduction in lung cancer mortality in the screened group compared with the control group.\textsuperscript{27} In contrast, the power to detect a smaller but clinically meaningful effect, such as a 10% reduction in mortality, was much lower (only 0.21 and 0.16, respectively).\textsuperscript{7,27,29}

Adding to the controversy, case-control studies have suggested that there might still be a benefit from CXR screening.\textsuperscript{32} The ongoing NCI-sponsored Prostate, Lung, Colorectal, and Ovarian Trial is a randomized trial designed to further determine the role of CXR screening in a low-risk population of both genders, but the results from this study have not yet been made available.\textsuperscript{33,34}

LDCT Scanning

LDCT scanning is a technique that allows a low-resolution image of the entire thorax to be obtained in a single breath-holding with low radiation exposure. There is enormous enthusiasm for LDCT as a screening test. To date, LDCT has only been evaluated in observational studies of volunteer cohorts (Table 2).

Studies of LDCT Screening

A prevalence screen using a single-spiral CT scan in 5,483 persons aged 40 to 74 years was performed in 1996 in Matsumoto, Japan.\textsuperscript{35} Most of the subjects had undergone yearly CXRs and sputum cytologic screening, and 3,967 subjects underwent CXR concurrently with spiral CT scanning. Smokers also underwent sputum cytology. Among subjects who had abnormalities prompting further evaluation were 59 subjects (1%) with noncancerous but suspicious lung lesions, 84 subjects (2%) with lesions suspicious for lung cancer, and 80 subjects (2%) with indeterminate small lung nodules. Nineteen lung cancers were diagnosed (8.5% of subjects with abnormal findings prompting further evaluation). The initial radiographic appearance of these lesions was suspicious for lung cancer in 14 cases, benign but suspicious in 3 cases, and indeterminate in 2 cases. Eighteen of the 19 cases were surgically confirmed, and 1 case was clinically diagnosed. Pathologic staging in 16 of 19 tumors (84%) was stage I, while the remainder (3 of 19 tumors) were stage IV. Four tumors were ≤ 1 cm, and 14 tumors were between 1 cm and 2 cm. Only 1 of the 17 tumors that were ≤ 2 cm was seen on CXR. One lung cancer that was not seen on CT scan was diagnosed by sputum cytology. Overall, the rate of lung cancer diagnosis was 0.48%, which was significantly higher than the rate of 0.03 to 0.05% in the same area prior to spiral CT screening.

A historical comparison of two screening strategies performed in Japan within the Anti-Lung Cancer Association was updated at the annual meeting of American Society of Clinical Oncology in May 1999.\textsuperscript{36,37} The report compared the outcome of screening with CXR and sputum cytology done from September 1975 to August 1993 (a total of approximately 26,000 screenings) to the same screening strategy plus spiral CT scan done from September 1993 to December 1998. Forty-three patients with primary lung cancer were found prior to CT scanning, compared to 36 patients with primary lung cancer in the CT-scan period. In addition, the percentage of stage IA tumors increased from 42 to 81%, and the 5-year survival improved from 48 to 82%. The results of this nonrandomized, historical comparison suggest that screening with CT scans can increase the ability to diagnose lung cancers at earlier stages but does not constitute strong evidence of a mortality benefit due to the potential for lead time and other biases.

The Early Lung Cancer Action Project (ELCAP) performed yearly low-dose spiral CT scans and CXR in a single-arm study of 1,000 smokers in New York City.\textsuperscript{38} One to 6 noncalcified lung nodules were found at baseline (“prevalence”) in 233 participants; 27 nodules (2.7%) were malignant. Among these 27 malignant nodules, 20 nodules (74%) were not found on standard CXR; in contrast, no malignant nodules were detected by CXR that were not also seen on CT scanning. The majority of these malignancies (23 of 27; 85%) were stage I, and all but one nodule (97%) were resectable. In the second year of screening (“incidence”), seven cancers were detected, and six of these were stage I. For relevance, in the United States, only approximately 20% of sporadically diagnosed lung cancer is stage I.\textsuperscript{39} The findings of these and related studies are summarized in Table 3. All of the studies had a high rate of false-positive results and a preponderance of stage I lesions among found cancers.

Conclusions: LDCT

It is difficult to determine whether LDCT meets either of the two criteria of a beneficial screening
Table 2—LDCT Screening Studies

<table>
<thead>
<tr>
<th>Study Site</th>
<th>Date of Publication</th>
<th>Patients Screened, No.</th>
<th>Abnormal Results, No. (% of Total)</th>
<th>Malignancies Detected, No. (% Total)</th>
<th>Detected Malignancies Stage I, %</th>
<th>Patients Screened, No.</th>
<th>Newly Identified Abnormal Results, No. (% of Total)</th>
<th>Malignancies Detected, No. (% Total)</th>
<th>Detected Malignancies Stage I, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadassah University, Israel</td>
<td>2001</td>
<td>614</td>
<td>NR</td>
<td>4 (0.65)</td>
<td>100</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Cornell University, United States</td>
<td>1999, 2001</td>
<td>1,000</td>
<td>233 (23)</td>
<td>27 (2.70)</td>
<td>81</td>
<td>1,184</td>
<td>63 (5)</td>
<td>7 (0.59)</td>
<td>85</td>
</tr>
<tr>
<td>Mayo Clinic, United States†</td>
<td>2002</td>
<td>1,520</td>
<td>782 (51)</td>
<td>22 (1.4)</td>
<td>59</td>
<td>1,464</td>
<td>191 (13)</td>
<td>3 (0.20)</td>
<td>0</td>
</tr>
<tr>
<td>Moffitt Cancer Center, United States†</td>
<td>2001</td>
<td>493</td>
<td>NR</td>
<td>11 (2.23)</td>
<td>18</td>
<td>91</td>
<td>NR</td>
<td>0 (0)</td>
<td>0</td>
</tr>
<tr>
<td>Muenster University, Germany†</td>
<td>2001</td>
<td>919</td>
<td>NR</td>
<td>17 (1.85)</td>
<td>76</td>
<td>NR</td>
<td>NR</td>
<td>2 (NR)</td>
<td>100</td>
</tr>
<tr>
<td>National Cancer Center Hospital, Japan†</td>
<td>2002</td>
<td>1,611</td>
<td>186 (11.5)</td>
<td>14 (0.87)</td>
<td>77</td>
<td>7,891</td>
<td>721 (9.1)</td>
<td>22 (0.28)</td>
<td>82</td>
</tr>
<tr>
<td>Shinshu University, Japan†</td>
<td>2001</td>
<td>5,483</td>
<td>676 (12)</td>
<td>22 (0.40)</td>
<td>100</td>
<td>8,303</td>
<td>518 (6)</td>
<td>34 (0.41)</td>
<td>86</td>
</tr>
<tr>
<td>Hitachi Health Care Center, Japan†</td>
<td>2001</td>
<td>8,546</td>
<td>NR</td>
<td>35 (0.41)</td>
<td>97</td>
<td>7,434</td>
<td>NR</td>
<td>7 (0.09)</td>
<td>100</td>
</tr>
</tbody>
</table>

*NR = not reported.
†Unpublished proceedings of the 4th International Conference on Screening for Lung Cancer, February 23–25, 2001; New York, NY.
test. Some have argued that the data from the observational studies are so compelling that we can conclude that screening with LDCT will be associated with a reduction in lung cancer mortality (the first criterion).40,41,45 Others have argued that the findings in the observational studies may reflect the impact of biases, rather than a true effect.30,46 Whether LDCT meets the second criterion—that it is not harmful or painful—is also the subject of a contentious debate.

Those arguing that LDCT will reduce lung cancer mortality point to very consistent findings in observational studies: LDCT detects far more lung cancers than CXR, and the vast majority of those detected are early stage. However, those arguing that the finding of early stage disease does not necessarily equate to an improvement in life expectancy raise two concerns. LDCT may be detecting nodules of histologic cancer that are often indolent in their behavior (ie, “overdiagnosis”), so treating these individuals cannot increase their life expectancy. Also, there are questions about whether surgical treatment of screen-detected cancers alters the natural history of the disease. Critics addressing the second criterion of a beneficial screening test have also pointed out that the aggregate harm to screened individuals (and to society) may exceed the benefit conferred to those who have lung cancer detected. In this case, these harms would be due to the high costs of screening and follow-up, and the morbidity and anxiety associated with false-positive results. We review the evidence substantiating each of these claims below.

**Evidence Supporting the Claim That LDCT May Be Associated With Overdiagnosis**

Much of the work on determining whether LDCT may overdiagnose lung cancer has focused on establishing whether indolent lung cancer exists. Some evidence from the randomized trials of CXR and sputum cytology has been advanced in support of the hypothesis that indolent lung cancer can be detected by a screening intervention. To understand this argument, consider that in a randomized trial of screening, if all screen-detected cancers grow, spread, and cause disease, one would expect that an equal number of cancers would be detected in the screened and unscreened groups after a sufficiently long period of follow-up. In the screened group, cancers would be found at earlier, presymptomatic stages while cancers in the nonscreened group would more typically be found at a later, symptomatic stage (stage shift). In contrast, if screening led to the identification of cancers that did not grow, spread, and cause disease, one would expect that an equal number of cancers would be detected in the screened and unscreened groups after a sufficiently long period of follow-up. In the screened group, cancers would be found at earlier, presymptomatic stages while cancers in the nonscreened group would more typically be found at a later, symptomatic stage (stage shift). In contrast, if screening led to the identification of cancers that did not grow, spread, or cause death, these cancers would be detected in the screened group, but would never manifest with symptoms and therefore would be missed in the nonscreened group. The detection of small lesions

<table>
<thead>
<tr>
<th>Study Site</th>
<th>Intervention</th>
<th>Early Stage</th>
<th>Late Stage</th>
<th>Screen Detected</th>
<th>Interval Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass Radiography Service, London</td>
<td>CXR every 6 mo for 3 yr</td>
<td>44*</td>
<td>57*</td>
<td>65</td>
<td>36</td>
</tr>
<tr>
<td>Mayo Clinic, United States</td>
<td>CXR at end of 3 yr</td>
<td>22*</td>
<td>54*</td>
<td>17</td>
<td>59</td>
</tr>
<tr>
<td>Research Institute of Tuberculosis and Respiratory Diseases, Czechoslovakia</td>
<td>CXR and sputum cytology every 6 mo for 3 yr</td>
<td>99</td>
<td>107</td>
<td>90</td>
<td>116†</td>
</tr>
<tr>
<td>Memorial Sloan-Kettering Cancer Center, United States</td>
<td>CXR annually and sputum cytology every 4 mo</td>
<td>50</td>
<td>64</td>
<td>70</td>
<td>44</td>
</tr>
<tr>
<td>Johns Hopkins, United States</td>
<td>CXR annually and sputum cytology every 4 mo</td>
<td>139¶</td>
<td>240</td>
<td>186</td>
<td>193</td>
</tr>
</tbody>
</table>

*Excluding cancers detected by the initial screening.
†Includes 73 cases detected by symptoms and 43 cases detected by screening outside of study protocol.
‡Includes 112 cases detected by symptoms and 48 cases detected by screening outside of study protocol.
§Includes 15 cases detected by symptoms, 1 case detected by screening outside of study protocol, and 3 cases detected by autopsy.
¶Not reported by study arm.
¶¶Calculated value based on reported proportions among screen-detected and interval cases.
that do not grow, spread, or cause death has been referred to as overdiagnosis. It is important to note that a higher percentage of early stage tumors is also found in the overdiagnosis scenario, so that the detection of a higher percentage of early stage tumors in the screened group cannot by itself distinguish stage shift from overdiagnosis.

In both the Mayo and Czechoslovakian randomized trials of CXR, substantially more cancers were detected in the screened than the unscreened group. On extended follow-up, the excess cancers detected by screening appear to have been overdiagnosed. Marcus et al reported on follow-up of all cases diagnosed during the enrollment and study period in the Mayo study (206 in the screened group and 160 in the control group). Nearly all of the 20% excess cancers detected in the screened group were early stage cancers. Failure to detect an equivalent number of early stage cancers in the control group was without apparent ill effect, because the control group experienced no excess number of lung cancer deaths, even when follow-up extended beyond 17 years for all participants. Kubik et al presented long-term follow-up data for the screened and control groups from the Czechoslovakian study. After the third year of this study, both the screened and control groups were entered into a program of annual screening. Despite the fact that substantially more lung cancers were detected in the screened group than in the control group during the first years of the study (year 1 to year 3), the rate of lung cancer detection in the two groups in the latter years (year 4 to year 6), while they were being screened in a similar fashion, was the same. In other words, greater detection of lung cancer in the screened group early in the study did not lead to a compensatory fall in detection rates later on, suggesting that screening was detecting excess lesions that would not have progressed to more advanced disease.

A number of investigators have also looked to data from autopsy series for evidence of indolent lung cancers that could be overdiagnosed by a sensitive screening test. These investigators argue that if indolent lung cancer exists, then numerous people who have died of other causes should have small nodules of lung cancer found during their autopsies. A study based at the Yale-New Haven hospital documented that of all patients who at autopsy had primary lung cancers, 28% were not diagnosed during their lifetime. The overall prevalence of undiagnosed lung cancer among all autopsies was 0.09%. A study from Duke University recently suggested that some lung cancers may not be detected at postmortem examination. Only 19 of 28 nodules (69%) seen on CT scanning within 2 months prior to death were found at autopsy. These findings are provocative but somewhat hard to interpret. Diagnosed lung cancer is relatively rare: the incidence in 50- to 55-year-olds is only 82 per 100,000 (0.08%), and only rises to about 340 per 100,000 (0.34%) in 75- to 79-year-olds (Surveillance, Epidemiology, and End Results Program, 1973 to 1998 Public Use Data). Also, nodules seen by CT scan that were undetected at autopsy may not have been primary lung cancers. As a result, it is difficult to know what rate of discovery would constitute strong evidence for or against the hypothesis that indolent lung cancer is a substantial concern in screening.

If screen-detected cancers are found at rates in excess of those expected based on population incidence statistics, or are discovered in patterns inconsistent with well established risk factors, this would suggest that some of the screen detected cancers are “atypical” (or overdiagnosed). Evidence that LDCT is detecting atypical lung cancer can be seen in two studies. In a large Japanese study of mobile CT scanning, the rate of detection of lung cancer was equivalent in smokers and nonsmokers, strongly contravening the epidemiologic data on sporadically diagnosed cancer. In the New York-based ELCAP study, the number of cancers detected in the annual follow-up LDCT is far less than that detected during the initial scan, even though the size of the detected lesions is similar. This difference in rates of detection is inconsistent with what one would expect if all of the lesions grew at similar rates (ie, all were aggressive malignancies). Instead, that the size of the prevalent lesions detected at initial scan are consistent with those found at subsequent scans but the rates of detection are substantially higher suggests that a meaningful proportion of cancers detected during the initial scan would have behaved in an indolent manner.

Evidence Supporting a Limited Impact of Surgical Resection on Outcome of Screen-Detected Disease

The intervention universally advocated for early stage lung cancer is surgical resection. Studies in which survival is assessed within a standardized population of individuals with early stage disease strongly support the efficacy of this intervention. However, there are no RCTs in which surgical resection has been compared to no treatment, and such a trial will probably never be done. Some indirect evidence raises doubts about the benefit of surgery in screen-detected disease. In the Mayo RCT of CXR, for example, more patients in the screened group had early stage lung cancer and underwent surgery, but there was no statistically
significant difference in lung cancer mortality (in fact, lung cancer mortality rates were marginally higher in the screened group). In other words, increased rates of surgery did not confer a benefit to the group as a whole. Similar results were seen in the RCT of CXR conducted in Czechoslovakia.

EVIDENCE THAT LDCT MAY RESULT IN AGGREGATE HARM TO SCREENED INDIVIDUALS

Distinct from the issues of lung cancer mortality, there are a number of concerns about the potential harms that could be caused by LDCT if implemented as a screening tool. First among these is the identification of large numbers of individuals with abnormalities on LDCT that do not require treatment, but may require serial monitoring and biopsies. These individuals likely experience anxiety when an abnormality is found during their screening test for lung cancer. The likelihood of such events appears high. In the ELCAP study, 23% of participants had at least one noncalcified nodule, and only 2.9% had diagnosed lung cancer, although it should be noted that all of the individuals who were identified as requiring biopsies because their lesions were highly suspicious did in fact have histologically confirmed lung cancer. In a parallel study conducted at the Mayo Clinic in Rochester, MN, 51% had at least one noncalcified nodule on CT scan, and only 1.4% had diagnosed lung cancer. In other words, the apparent false-positive to true-positive ratio is probably in the range of 10 to 30. A corollary source of harm may result from some histologically proven lung cancers being in fact overdiagnosed. Unnecessary treatments may include thoracotomy, lung resection, chemotherapy, and radiation.

The cost of mass screening with LDCT is unknown. A large potential population of high-risk subjects, combined with a high rate of findings that necessitate follow-up CT scanning, subspecialist consultation, and other testing, all endorse the expectation of a high total cost. An additional source of excess cost will come from the fact that costs of care for individuals with early stage lung cancer exceed costs of care for individuals with more advanced stage disease. Whether the LDCT will be cost-effective, meaning that the incremental improvement in survival per dollar spent is reasonable, is a related question that cannot be answered until effectiveness is established. The few published cost-effectiveness analyses on this topic rely on assumptions about effectiveness that are unproven.

SUMMARY

Sputum cytology screening has been evaluated in the context of RCTs and does not appear to meet the criteria of a beneficial screening test. It appears to detect only a minority of all lung cancers, and does not appear to reduce lung cancer mortality, although it has been primarily assessed in combination with other screening evaluations. More sensitive analyses of sputum for abnormalities suggestive of malignancy are being investigated currently. CXR screening every 4 months or 6 months in combination with sputum cytology has been assessed in two RCTs, and CXR alone every 6 months as a single modality was assessed in a third RCT. In none of these studies was the control arm a no-screening intervention, and the power of these studies to detect small reductions in mortality was limited. Yet, in none of these studies was CXR associated with a reduction in lung cancer mortality. CXR is being further evaluated in the nearly completed Prostate, Lung, Colorectal, and Ovarian Trial.

LDCT has only been evaluated in the context of observational studies with volunteer cohorts. The vast majority of cancers detected by LDCT scanning in these studies are stage I at the time of discovery. This suggests that the preponderance of them might be cured through surgical intervention. However, evidence from numerous sources casts doubt on the extent to which these screen-detected lesions represent aggressive malignancies, and therefore the extent to which surgical treatment will alter outcome. Readers should note that the finding of larger numbers of early stage cancers also characterizes all of the RCTs of sputum cytology and CXR, and in none of these studies was there an alteration in lung cancer mortality. The impact of LDCT on lung cancer mortality to date is consequently unknown.

REFERENCES

lung nodules at autopsy: implications for lung cancer screening and overdiagnosis bias. Lung Cancer 2001; 33:11–16
53 Cancer rates and risks. Bethesda, MD: National Cancer Institute, Cancer Statistics Branch, 1996