Lung cancer screening: Is it time for a change in policy?

ABSTRACT

Two recent studies of computed tomography (CT) as a screening test for lung cancer have heightened debate about this topic. Although the International Early Lung Cancer Action Program investigators (N Engl J Med 2006; 355:1763–1771) concluded that annual CT screening can detect lung cancer that is curable, Bach et al (JAMA 2007; 297:953–961) concluded that it may not meaningfully reduce the risk of advanced lung cancer or death from lung cancer. We feel that questions remain about the degree of reduction in lung cancer-specific mortality, the potential morbidity caused by screening, the appropriate group to screen, and the cost-effectiveness of screening. These questions warrant further study prior to accepting CT screening as the standard of care. Hopefully, much of this knowledge will be gained when the results of ongoing controlled studies are available.

KEY POINTS

Lead-time bias, length-time bias, and overdiagnosis bias can influence the interpretation of survival results of nonrandomised trials of screening tests, making screening appear to be more beneficial than it really is.

Cohort studies suggest that CT screening can increase the survival rates of patients diagnosed with lung cancer; they cannot comment on the lung cancer-specific mortality rate in the screened population.

As yet, no guidelines from any professional organization recommend in favor of routine CT screening for lung cancer.

THE TWO NEW STUDIES

I-ELCAP: Screening people at risk may be beneficial

In October 2006, the International Early Lung Cancer Action Program (I-ELCAP) investigators reported the results of their large screening study. In brief, over 12 years they screened 31,567 people who were at risk of lung cancer. In total, 484 participants were diagnosed with lung cancer, 85% of which were in clinical stage I. The 10-year survival rate for all those diagnosed with lung cancer was projected to be 80%, and for those with stage I lung...
cancer it was 88%. The investigators concluded that CT screening can detect lung cancer that is curable, and their results support CT screening for lung cancer as a standard of care in people at risk of the disease.

Bach et al: Screening may not reduce deaths from lung cancer
In March 2007, a report by Bach et al. was published comparing the findings from three large CT screening cohorts to predictions of outcome based on validated models. Three times more cases of lung cancer were diagnosed than were expected, and 10 times more lung surgery procedures were performed, but the authors found no decline in the number of advanced cases diagnosed or deaths from lung cancer.

CRITERIA FOR A SUCCESSFUL SCREENING PROGRAM

Most experts believe that a successful screening program should reduce the number of disease-specific deaths in the screened population. The success of a screening program is determined by the disease, the test, and the treatment.

The disease must be of serious consequence to the screened population, and it must be detectable in a preclinical form.

The test must be capable of detecting preclinical disease at a point before it becomes untreatable. It should detect little pseudodisease and cause little morbidity. It needs to be affordable and available.

The treatment must be safe and more effective if the disease is found early in its course.

In the following sections we review some of the issues surrounding lung cancer screening in the context of the criteria for a successful screening program. We discuss how the two recent reports alter our understanding of these issues and conclude with our recommendations for lung cancer screening based on current knowledge.

DOES SCREENING REDUCE LUNG CANCER-SPECIFIC MORTALITY RATES?

In order to comment on a screening program’s ability to reduce lung cancer-specific mortality, a clinical study must screen a group of people and compare them with a similar group of people who are not screened (a control group).

To date, there have been no controlled trials of chest CT screening reported that have been large enough to assess this end point. The vast majority of chest CT screening studies reported have been cohort trials, in which all of the participants were screened. This type of trial can address lung cancer-specific survival (ie, the length of time between the diagnosis of lung cancer and the patient’s death), but not lung cancer-related mortality.

Three types of bias in reporting survival in screening studies
Survival results can be misleading, as they are susceptible to potential biases. The three most commonly discussed biases in this context are lead-time bias, length-time bias, and overdiagnosis bias.

Lead time is the interval between when a disease is detected by the screening test and when the disease would have become known to the patient if he or she had never been screened, eg, through signs and symptoms of the disease. When a lead-time bias exists, the longer survival of the screened group is at least partially the result of finding out about the disease earlier, not living longer.

Length-time bias occurs in diseases in which the length of the preclinical and clinical phases (and thus the aggressiveness of the disease) varies considerably among patients. Screening tends to detect cases with longer preclinical phases. When a length-time bias exists, the screening test is detecting cases of the disease that are less aggressive than usual. Perhaps the cases detected could have been treated as effectively even if diagnosed later.

Overdiagnosis bias is an extension of length-time bias. Overdiagnosis occurs in cases in which the disease was progressing so slowly that it would not have affected the patient’s life in any way.

Are lung cancer screening studies affected by these biases?
The reported large controlled trials of screening with chest radiographs have highlighted the need to consider these biases when...
interpreting survival results. These trials reported higher survival rates in the screened groups but were unable to detect a reduction in disease-specific mortality rates. More cases of early-stage lung cancer were detected in the screened groups, but equal numbers of late-stage cancers were found. Similarly, estimates of lung cancer mortality rates in two of the CT screening cohorts suggested lung cancer mortality rates similar to those in prior screening trials using chest radiographs.

These reports have led to debate about the influence of the above biases on the results. The impact of an overdiagnosis bias has received the most attention. Evidence has been reported that can support either side of the argument about the presence of an overdiagnosis bias.

Studies suggesting that overdiagnosis bias is unlikely to be of major consequence in lung cancer include the following:
- Motohiro et al. reported that, in 802 patients with nonsurgically treated stage I lung cancer, the 5-year survival rate was 16.6% and the 10-year survival rate was 7.4%.
- Sato et al. reported that, in 44 untreated patients with squamous cell carcinoma diagnosed by sputum cytology who had negative chest imaging, the 5-year survival rate was 53% and the 10-year survival rate was 34%.
- Bianchi et al. found that gene profiles of screen-detected cancers were similar to those of cancers detected outside of a screening study.

Studies suggesting that overdiagnosis bias may be of major consequence include the following:
- Read et al. found that moderate to severe comorbidity has a greater impact on survival in patients with early-stage lung cancer than in those with later-stage disease.
- Marcus et al. performed extended follow-up of a controlled trial of chest radiography screening, which suggested that up to 17% of cases of screen-detected cancer could have been overdiagnosed, though some have ascribed these findings to the trial’s design.
- Manser et al. performed an autopsy study that found a small number of incidental lung cancers in people who died of “natural causes.”

This debate has yet to be settled. Overdiagnosis bias is undoubtedly influenced by the population that is screened.

What new information have the recent reports provided?
The I-ELCAP study was a single-arm cohort trial, not a controlled trial. Thus, it cannot answer the question about a reduction in lung cancer-specific mortality rates. Yet it does provide intriguing information about potential survival.

Bach et al. used a validated model to estimate the expected lung cancer mortality rate in the population studied. They removed the results of the prevalence screening (ie, findings on the initial scan) from their primary analysis. They could not find a difference in actual vs estimated mortality. The results from this trial depend on the accuracy of the model that was used to estimate mortality.

Lead-time bias is unavoidable in a screening study that examines survival. In fact, it is necessary that a lead time bias exist if there is any chance of improving the mortality rate in the population screened. The estimated average lead time in a large randomized screening trial using chest radiographs was 16.4 months. The average lead time would be expected to be longer for nodules detected on CT. As most of the cancers detected in the I-ELCAP study were cured, the impact of this bias was probably small.

Length-time bias also may have contributed to the results of the I-ELCAP trial. Seventy-one percent of all cancers detected were adenocarcinomas, which may have been less aggressive than the average cancer. In some populations, including some of the groups included in the trial, small adenocarcinomas have been reported to have higher-than-average cure rates. In addition, most of the cancers were detected on the prevalence screen. Cancers with a longer latent phase tend to be more prevalent at the time of baseline screening. Thus, the survival results reported are more descriptive of a slower-growing subtype of lung cancer.

The I-ELCAP results argue against an overdiagnosis bias in that all eight people with stage I lung cancer who did not undergo treatment died of lung cancer within 5 years of their diagnosis. The number of follow-up
(incidence) screens (< 1 per study subject) and of cases of cancer detected on these follow-up screens (a total of 74) were too low to comment on changes in stage distribution over time.

The study by Bach et al argues in favor of an overdiagnosis bias because it showed an increased number of lung cancer diagnoses and surgeries without an improvement in lung cancer-specific mortality rates.

**Bottom line.** Intriguing and promising survival data have been reported, but serious doubts have been raised about the translation of these survival data into a mortality benefit.

### DOES CT DETECT PRECLINICAL DISEASE BEFORE THE DISEASE BECOMES UNTREATABLE?

Chest CT finds preclinical disease in the form of small lung nodules. One would intuitively think that the smaller the cancerous nodule, the more treatable it should be. However, two reports led to some debate on this issue.

Patz et al described 510 patients with stage IA lung cancer (ie, a tumor smaller than 3 cm without any spread) that was surgically resected. They were unable to find a relationship between tumor size and survival.

Heyneman et al evaluated 620 patients with T1 tumors (ie, < 3 cm) and found no relationship between tumor size and the stage of the cancer at presentation. These reports suggest that finding smaller tumors with CT would not lead to improved outcomes.

Several reports refute these findings. Three reports were generated from data in the national Surveillance, Epidemiology, and End Results registry. The first found that, among 7,620 patients with stage I lung cancer, smaller tumor size at diagnosis led to improved survival. The second found the same trend in 9,191 patients with surgically treated stage I lung cancer. The third evaluated the distribution of disease stage within categories of tumor size in more than 84,000 patients; smaller tumors were more likely to be stage I.

Thus, the bulk of the evidence suggests that CT, which can detect smaller tumors than chest radiography, meets the criterion of detecting preclinical disease.

### What new information have the current reports provided?

In the I-ELCAP study the tumors detected were small. Four hundred twelve (85%) of the 484 tumors were clinical stage I. Pathologic staging (in the 375 patients with clinical stage I cancer who underwent resection) revealed 7% with lymph node metastases and 9% with a second tumor. Thus, 349 (72%) of the 484 patients with cancer were believed to have a solitary stage I cancer after clinical staging with or without pathologic staging. In those with clinical stage I cancer who underwent resection, the median diameter of the tumor was 13 mm if detected at baseline and 9 mm if detected on annual follow-up.

The 10-year survival rate for all 484 participants diagnosed with lung cancer was 80%. The 10-year survival rate of those with pathologic stage I cancer was 94%. However, most of the study participants did not receive long-term follow-up, and lung cancers that were found outside of screening were not reported. Thus, it is unclear if what was reported was representative of the entire population that was screened.

The report from Bach et al highlights the potential for the most advanced lung cancers to escape detection from screening until they have become more difficult to treat.

**Bottom line.** The bulk of the evidence suggests that smaller tumors found by CT screening are more readily cured than larger ones. Questions remain about the ability of CT screening to identify enough lung cancer before it has become advanced.

### DOES CT DETECT LITTLE PSEUDODISEASE AND CAUSE LITTLE MORBIDITY?

Pseudodisease in this context refers to findings on testing that look like the disease in question but do not represent it. Pseudodisease often leads to additional testing to prove it is not the disease in question, with consequent additional expense, risk, and worry.

All of the cohort studies of CT screening have reported a large number of benign nodules being detected. On prevalence screening anywhere from 5% to 51% of study participants have been reported to have at least one nodule. In one study, 73.5% of partici-
pants were found to have a nodule after 5 years of follow-up. The highest rates of nodule detection were in studies that used thinner CT slices. As slice thickness decreases, more nodules will be detected, heightening this problem.

Most of the small nodules detected are not cancer. Intensive follow-up protocols are needed to be sure that cancers are treated expeditiously and that invasive procedures are avoided when the nodules are benign. In the studies reported, for every three cancers resected, one surgical biopsy was performed for benign disease. Positron-emission tomography has been used in evaluating the nodules detected, with variable success.

Benefits in addition to those intended and the potential for harm from the test must be considered. Variable rates of other disease findings (eg, chronic obstructive pulmonary disease, coronary artery calcification, other tumors) have been reported. It is unclear how this will impact the success of the screening program. Smoking cessation efforts have been found to be more successful in those who receive reports of an abnormal scan. The effects of the radiation received are unclear. The age group of the population screened is likely to temper any long-term effects of radiation. One report suggests that 36,000 excess lung cancers would occur from yearly CT screening between the ages of 50 and 75 if 50% of the population at risk were screened. This is highly theoretical and impossible to prove clinically.

What new information have the current reports provided?
The I-ELCAP study defined the initial screening test as being positive if it detected a nodule 5 mm or larger in diameter. By eliminating the smallest nodules from consideration, they minimized the amount of pseudodisease that was reported.

Even so, 13% of all participants had a positive baseline scan. The annual scans were defined as positive if they detected any new nodule, regardless of size. Five percent of all participants who had an annual scan had new nodules. Only 1.3% of study participants (9.7% of those with a nodule) had lung cancer detected by the prevalence scan, while 0.3% (5.1% in those with a new nodule) had lung cancer detected by the follow-up screen. A low rate of biopsy for benign disease was reported (535 biopsies yielding 492 different cancers), and the operative mortality rate during resection of the cancer was remarkably low (0.5%).

Bach et al did not comment on nodule detection. They did report 10 times more lung resections than were expected, suggesting a high potential for increased morbidity.

**Bottom line.** CT screening will uncover many benign nodules that are likely to receive intensive follow-up. The number of biopsies for benign nodules can be minimized. Operative mortality rates are low, but the number of additional surgeries may be large. The long-term risks of the test are unclear.

**IS THE SCREENING TEST AFFORDABLE AND AVAILABLE?**

Several cost-effectiveness studies have been published. They vary in the perspective from which they were written and in the costs used for the calculations. Examples:

- An early report concluded that CT screening would cost $48,000 per life-year gained if 50% of the lung cancers it detected were at a localized stage.
- A study performed from a societal perspective suggested a cost of $116,000 per quality-adjusted life-year (QALY) gained in current smokers, and $2.3 million in former smokers.
- A study performed from a health care perspective suggested that performing a prevalence screen alone would cost $2,500 per year of life saved.
- A report from the perspective of the government as a third-party funder estimated a cost of $105,090 per QALY in current smokers.

The actual cost-effectiveness is sure to vary depending on the algorithm used for follow-up and diagnosis, advances in technology, advancement in treatments, the actual reduction in deaths, the translation of these studies to common practice, and the population screened.

A population with a higher risk of lung cancer would benefit more from the screening.
program, but this does not mean that lower-risk groups would not benefit from being screened. Even nonsmokers have a death rate from lung cancer that is higher than that from most other types of cancer.\textsuperscript{49} It may be difficult to exclude low-risk groups when the program is translated into common practice.

**What new information have the current reports provided?**

The I-ELCAP study did not provide any direct information about cost-effectiveness. If the high survival rate that was reported translates into a large reduction in lung cancer-specific mortality rates, the program is more likely to be cost-effective. People at lower risk (ie, younger people) appeared not to comply as well with returning for annual screening, though the authors did not delve into this trend. This could influence cost-effectiveness.

The report from Bach et al suggests that it is premature to be discussing cost-effectiveness when effectiveness in general has not been proven.

**Bottom line.** There is not enough information to determine the cost-effectiveness of a lung cancer screening program that uses chest CT. Overall effectiveness must be proven first.

**AVAILABLE GUIDELINES**

Four sets of guidelines on lung cancer screening are available.

The American Cancer Society in 2001 suggested that individual patients at risk for lung cancer should be advised of their risk and educated about the current state of early detection. If testing is to occur, it should be in a setting with multidisciplinary specialty groups.\textsuperscript{50}

The Society of Thoracic Radiology stated in 2001 that “it is the consensus of this committee that mass screening for lung cancer with CT is not currently advocated.”\textsuperscript{51}

The American College of Chest Physicians in 2003 said, “we recommend against the use of a single LDCT [low-dose CT] or serial LDCTs to screen for the presence of lung cancer,” giving it a grade of recommendation of I (insufficient evidence for or against its routine use).\textsuperscript{52}

The U.S. Preventive Services Task Force in 2004 also gave lung cancer screening a grade of I. They state, “the USPSTF could not determine the balance between the benefits and harms of screening for lung cancer.”\textsuperscript{53}

All of the guidelines advise that appropriate patients be informed about ongoing lung cancer screening trials. All of these guidelines were produced before the two recent studies were published.

**THE FUTURE**

Two large, randomized, controlled trials of lung cancer screening are well under way.

The National Lung Screening Trial has enrolled 50,000 people at risk and randomized them to receive a chest CT scan or a chest radiograph annually. The trial is expected to be completed in 2010 or 2011.

The Dutch lung cancer screening (NELSON) trial, with 20,000 participants randomized to receive chest CTs or standard of care, is expected to be completed after 2011.

These trials have designs that will allow them to comment more accurately on reduction in lung cancer-specific mortality rates and on cost-effectiveness.

Future advances in our ability to identify populations at risk, in the management of lung nodules,\textsuperscript{54–56} and in tests for lung cancer\textsuperscript{57–61} may improve upon the current screening approaches.

**BOTTOM LINE**

The results of the two recent lung cancer screening studies are intriguing and confusing. They have added to our knowledge and heightened debate about this topic. We feel that questions remain about the degree of reduction in lung cancer-specific mortality rates, the potential morbidity of the screening test and the testing to evaluate the findings, the selection of the appropriate group to screen, and the cost-effectiveness of a lung cancer screening program. These questions warrant further study before we accept lung cancer screening with chest CT as the standard of care. Hopefully, much of this knowledge will be gained when the results of ongoing controlled studies are available.
REFERENCES


