Lung cancer screening in 2008: A review and update

Emily J. Brawner, S. Patrick Nana-Sinkam, James R. Jett

Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, The Ohio State University, DHLRI 473, West 12th Avenue, Room 201, Columbus, OH 43210, USA

Division of Pulmonary Medicine and Medical Oncology, Mayo Clinic, 200 First Street, S.W., Rochester, MN, 55905, USA

Summary
This article discusses the strengths and weaknesses of using sputum cytology, plain chest radiograph and computerized tomography (CT) as screening modalities for lung cancer and provides recommendations for screening.

Introduction
Lung cancer is the second most common new cancer diagnosis with an estimated 213,000 new cases in the United States in 2007. More importantly, lung cancer is the leading cause of cancer death for both men and women, accounting for a staggering 29% (164,840) of all cancer deaths in the United States. There are two major categories of lung cancer: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). The vast majority (85–90%) of lung cancers are NSCLCs. The 5-year relative survival rate for NSCLC is 49.5% when the disease is still localized and only 2.1% for metastatic disease.1 Despite efforts at early detection and new treatment modalities, the overall 5-year survival rate remains only 15%.1 Many speculate that the poor survival may be attributable to the indolent nature, thus resulting in the symptomatic patient presenting late in the course of the disease. Therefore, screening programs targeted to high-risk individuals may offer earlier detection, thus increasing potentially curative treatment options and improving patient survival.

Historical perspective
To date, there have been several studies focused on early detection of lung cancer. Sputum cytology and chest X-ray have been proposed as modalities for lung cancer screening.2–6 In the 1970s, the National Cancer Institute (NCI) sponsored randomized controlled trials (RCTs) at Johns Hopkins, Memorial Sloan-Kettering Cancer Center and the Mayo Clinic.1–6 The objectives of these trials were to determine if the addition of sputum cytology to chest X-ray interval examinations could lead to early detection of lung cancer and if screening would reduce lung cancer mortality. Approximately 31,000 men at high risk for lung cancer, defined by an active or former history of smoking one pack per day or more, were randomized to chest X-ray with sputum cytology or chest X-ray alone. The baseline screen, or “prevalence” screen, was followed by interval screens over 5 years. Participants were then followed for an additional 5 years thereafter. The prevalence screen (or baseline screen) detected lung cancer in 0.5–0.8% of the
participants. Chest X-ray screening detected almost twice as many cases of lung cancer compared to cytology alone. Early stage lung cancer detected with cytology alone were predominantly centrally located squamous cell carcinomas while those detected with chest X-ray alone were peripherally located adenocarcinomas. The dual approach of chest X-ray with sputum analysis, however, detected only 23% (7 out of 30) stage I lung cancer, as defined by the American Joint Committee of Cancer (AJCC). The study demonstrated that screening with chest X-ray and sputum cytology resulted in the detection of early stage cancers suitable for resection and therefore, translated to increased survival. However, there was no difference in the number of deaths in the screened population compared to the control group. Therefore, screening with chest radiographs and/or sputum cytology has not been recommended.

Over the last few years, the focus has shifted to the utility of screening with low-dose computed tomography (CT). Several studies have demonstrated that helical CT detects more cases of early stage lung cancer than chest X-ray and/or sputum cytology. In one study, investigators detected 23 out of 25 cases of lung cancer by CT alone compared to two cases detected by sputum cytology alone (n = 1520). However, the detection of early stage lung cancer was also accompanied by a high rate of detecting benign nodules. In another study, the proportion of lung cancer detected by low-dose CT surpassed that of chest X-ray and sputum cytology for the baseline screen as well as for the repeat screen.

Despite the ability to detect early stage disease, it is unclear if screening studies will ultimately affect the overall outcome of the disease. A retrospective review of patients with pathologically proven NSCLC ≤ 3 cm in size found that regardless of tumor size (<1 cm, 1–2 cm or 2–3 cm), the stage distribution was comparable in all the patient groups (i.e. the proportion of stage I lung cancer to stage IV lung cancer was similar in all groups). Therefore, finding small tumors by CT screening may not result in a stage shift, or identification of lung cancer in the earlier stage of the disease. A true stage shift should increase detection of early stage disease with a concomitant decreased detection of late stage disease. Unfortunately, this is yet to be demonstrated. Therefore, substantial debate continues to surround studies on CT screening for lung cancer, exemplified more recently by two studies.

**Recent studies**

From 1993 to 2005, the International Early Lung Cancer Action Program (I-ELCAP) enrolled participants to determine the utility of screening by spiral CT scanning. Participants were asymptomatic men and women 40 years of age and older deemed to be at risk for lung cancer because of either a history of cigarette smoking, exposure to secondhand smoke or occupational exposure (to asbestos, beryllium, uranium or radon). A total of 31,567 participants underwent baseline CT screening between 1993 and 2005. Of those, 27,456 participants underwent a repeat annual CT screening approximately 1 year after the baseline screen. Investigators diagnosed lung cancer in 484 participants: 405 cases were detected during the baseline screen while 74 cases were detected during the annual screen. The remaining five cases were "interim" diagnoses, in which the baseline or annual screen was negative, but further diagnostic work-up was pursued because of suggestive symptoms or incidental findings. Approximately 85% (412 of 484) were found to have clinical stage I lung cancer. Investigators followed all cases of lung cancer for a median of 40 months (1–123 months). Deaths from lung cancer, including deaths resulting from treatment of lung cancer, were recorded and Kaplan–Meier curves were developed to estimate the 10-year lung cancer-specific survival rates. The estimated 10-year survival rate for all participants diagnosed with lung cancer, regardless of tumor stage and treatment, was 80%. Participants with stage I lung cancer had an estimated 10-year survival rate of 88%, and 92% if they underwent surgical resection. These findings were a substantial improvement compared to the 5-year survival rate of 60–80% quoted by several studies for patients with pathological stage I disease undergoing surgical resection. The authors concluded that annual spiral CT screening of at-risk patients can detect lung cancer that is curable. However, there were several limitations. The study was a non-randomized, observational study and the 10-year survival rate was extrapolated from a shorter duration of follow-up (median of 40 months).

Recently, investigators conducted a study in which they recruited 3246 asymptomatic current or former smokers for screening with low-dose CT through three academic centers, two in the United States and one in Italy. Participants underwent a baseline CT scan followed by three to four subsequent annual CT scans with a median follow-up of 3.9 years. The number of lung cancer events detected in the cohort was compared to prediction models developed to estimate the risk of being diagnosed with lung cancer or the risk of dying from lung cancer. These models have been described and validated in previous studies. The models included individuals who are at high risk based on their age (50–80 years of age), smoking history (average of 10–60 cigarettes per day over 25–60 years) and if former smokers, how long they had quit (within the past 20 years).

The authors demonstrated a threefold increase in the number of lung cancer diagnosed using low-dose CT screening, and a 10-fold increase in the likelihood of undergoing thoracic surgery for a lung lesion, as compared to the validated prediction model. However, screening did not reduce the number of advanced lung cancer diagnoses or lung cancer deaths. They concluded that patients should not undergo screening for lung cancer with low-dose CT outside of clinical investigations until further conclusive data are obtained. The authors acknowledged, however, that the study was limited by its smaller sample size and that perhaps, a longer duration of screening and/or follow-up may have detected a benefit of screening. The study sample size could not exclude a potential lung cancer mortality reduction of 30% from screening.

**Limitations of screening**

Studies on lung cancer screening have been met with varying levels of criticism that are primarily based on the perceived inherent biases in many studies. In order for screening to be effective, the target disease should be a significant health problem (resulting in substantial morbidity and mortality),
should be detectable in those who are at risk before symptoms occur, and earlier detection should result in improved treatment outcomes. These requisites are certainly applicable to lung cancer. The screening test should detect little "pseudodisease", or findings that are similar to the target disease but do not represent it, should be cost-effective and cause little morbidity.\textsuperscript{19,20}

There are many aspects of screening as outlined by Black.\textsuperscript{21} The natural history of disease includes preclinical and clinical phases. Within the preclinical phase is the detectable preclinical phase (DPCP) of the disease, wherein a screening test begins to detect disease before signs or symptoms arise. The critical point in the disease occurs when therapy would be less effective. In the case of lung cancer, this would be the advanced stages of the disease when it is no longer localized and therefore carries a poorer prognosis. Thus, if the DPCP occurs after that critical point, then screening will not be applicable. An effective screening modality would detect a target disease before it is clinically apparent so that the development of advanced disease is prevented.

There are certain biases that may affect screening tests. The interval between diagnosis in an asymptomatic individual and the development of signs and symptoms of disease is known as the lead time. Lead time bias refers to the relatively longer survival observed by earlier diagnosis despite no change in death from lung cancer (see Figure 1). Length time bias occurs when the length of the preclinical stage of disease is longer than its clinical stage (e.g. identifying a slow growing tumor). In other words, the disease outcome may be the same if it was diagnosed later as long as it was within the lengthened preclinical stage (see Figure 2). Overdiagnosis bias refers to the diagnosis of disease that otherwise would not have caused a patient’s demise, such as less aggressive or indolent cancers, even if it was not detected initially (see Figure 3). Overdiagnosed lung cancers were previously defined by Yankelevitz et al.\textsuperscript{22} as those having a volume doubling time (VDT) of 400 days. Lindell et al.\textsuperscript{23} conducted a retrospective review of lung cancers detected in high-risk individuals who underwent annual chest CT screening for 5 years. "High-risk" individuals were defined as males or females aged 50 years or older, with a smoking history of at least 20 pack years, no prior history of cancer, no need for supplemental oxygen and a life expectancy of at least 5 years. VDT was also calculated (based on a modified Schwartz equation) for cancers that

---

**Figure 1**  
*Lead time bias.* Screening detects disease earlier but the time of death remains unchanged.

**Figure 2**  
*Length time bias.* Screening detects less aggressive disease due to its longer preclinical phase.
were imaged at more than one CT examination. Out of 1520 participants, 61 primary lung cancers were identified in 59 patients, most of which were non-bronchioloalveolar cell adenocarcinomas. VDT was calculated for 48 cancers with a mean value of 518 days $\pm$ 1094 (median, 166 days; range 10–5810 days). The authors found that 27% ($n = 13$) of these cancers were calculated to have a VDT longer than 400 days. Recalculated for stage I lung cancer only, 25% ($n = 10$) would be considered overdiagnosed. Interestingly, the mean tumor VDT was longer in women (688 days for women and 234 days for men) and in almost every tumor histologic subtype, the mean VDT was longer in women than in men. Thus, the authors concluded that overdiagnosis may be substantial in lung cancer screening, especially in women.

Lead time, length and overdiagnosis bias are the most significant and widely appreciated biases that make survival a less than ideal measure of screening benefit. The disease-specific mortality, or the number of deaths in screened participants attributed to lung cancer, is a better measure of screening benefit than disease-specific survival since it escapes these biases. Much debate on the impact of overdiagnosis bias on the screening tests for lung cancer remains.  

Recalculated for stage I lung cancer only, 25% ($n = 10$) would be considered overdiagnosed. Interestingly, the mean tumor VDT was longer in women (688 days for women and 234 days for men) and in almost every tumor histologic subtype, the mean VDT was longer in women than in men. Thus, the authors concluded that overdiagnosis may be substantial in lung cancer screening, especially in women.

Our current knowledge of the risks of cancer from radiation exposure come from the follow-up of atomic bomb survivors.  

Risks from radiation exposure have also been cause for concern. The effective dose, or the absorbed energy non-homogenously distributed within an organ in CT, is expressed in sieverts (Sv). The National Council on Radiation Protection and Measurements (NCRP) quoted an average natural (non-medical) radiation exposure per individual in the United States of about 3 mSv/year. Man-made sources of radiation contribute another 0.6 mSv; the majority of which comes from medical X-rays (0.4 mSv).  

The effective dose equivalent with chest radiography ranges from 0.06 to 0.25 mSv, 3–27 mSv for conventional chest CT and 0.3–0.55 mSv for low-dose CT. Thus, CT screening which can include interval follow-up CT studies within a year can exceed the average natural radiation exposure per individual. What becomes of the cancer risk from radiation exposure? This concern was brought to light more recently with a publication in the New England Journal of Medicine. Our current knowledge of the risks of cancer from radiation exposure come from the follow-up of atomic bomb survivors.

### Risks of screening

Ideally, a screening test should detect little “pseudodisease” and cause little morbidity. “Pseudodisease” refers to findings that are similar to the target disease but do not represent it, such as benign pulmonary nodules encountered in screening for lung cancer. Pseudodisease may lead to additional testing and surgical intervention accompanied by further risk, expense and psychologic burden.

Crestanello et al. reviewed the number of thoracic surgeries performed on 1520 participants screened with CT from 1999 to 2002. Participants were at least 50 years of age with a history of 20 pack years or greater. A total of 3130 indeterminate pulmonary nodules were detected in 73% of the participants ($n = 1112$) and 3.6% ($n = 55$) underwent 60 thoracic operations. Lung cancer was diagnosed in 81.9% ($n = 45$) while benign disease was diagnosed in 18.1% ($n = 10$). The operative mortality was 1.7%; however, none of the deaths occurred in those who underwent surgery for benign disease. Nevertheless, this should be considered in the context of the actual proportion of lung cancer identified with screening, which only comprises 1.5% (45 of 3130) of the indeterminate pulmonary nodules detected.

The financial and economic implications of CT screening are also of major concern. Estimates have ranged from $2500 to as much as $2.3 million per life year gained. Mahadevia et al. used a computer simulation program to compare annual CT screening to no CT screening of 100,000 hypothetical patients. Adjusted for known biases and with a projected stage shift of 50%, they estimated a cost of $116,300 per quality-adjusted life year (QALY) gained for current smokers and over $2.3 million per QALY gained for former smokers. Chirikos et al. estimated the expected economic value of screening by combining the projected cost-effectiveness ratios and life expectancy (based on the stage of lung cancer at diagnosis) of two hypothetical cohorts. The authors determined that screening with CT costs approximately $48,000 per life year gained if 50% of lung cancer was detected at a localized stage. This cost increased as smaller proportions of early stage lung cancer were detected with CT screening. The lowest estimate of cost-effectiveness analysis was published by Wisnivesky et al. The authors incorporated data from the Early Lung Cancer Action Project (ELCAP) into a decision analysis model comparing CT screening in high-risk participants ($\geq 60$ years of age and $\geq 10$ pack year smoking history) to observation without screening. They found that a single baseline low-dose CT screening cost $2500 per year of life saved and that screening was expected to increase survival by 0.1 year at an incremental cost of about $230. However, the cost of the baseline low-dose CT screen exceeded $50,000 per year of life saved if the likelihood of overdiagnosis bias was $\geq 50%$. Risks from radiation exposure have also been cause for concern. The effective dose, or the absorbed energy non-homogenously distributed within an organ in CT, is expressed in sieverts (Sv). The National Council on Radiation Protection and Measurements (NCRP) quoted an average natural (non-medical) radiation exposure per individual in the United States of about 3 mSv/year. Man-made sources of radiation contribute another 0.6 mSv; the majority of which comes from medical X-rays (0.4 mSv). The effective dose equivalent with chest radiography ranges from 0.06 to 0.25 mSv, 3–27 mSv for conventional chest CT and 0.3–0.55 mSv for low-dose CT. Thus, CT screening which can include interval follow-up CT studies within a year can exceed the average natural radiation exposure per individual. What becomes of the cancer risk from radiation exposure? This concern was brought to light more recently with a publication in the New England Journal of Medicine.
in Japan, spanning over 50 years. The authors note that
the survivors who were exposed to even lower radiation
doses (mean dose 40 mSv) had an increased overall risk of
cancer. This implies that two or three conventional CT scans
in an adult could potentially equal or surpass 40 mSv. The
authors, however, derived this conclusion chiefly from CT
studies obtained for diagnosis rather than for screening
purposes. Nevertheless, the concern remains.

Complementary modalities

The utility of F-18-fluorodeoxyglucose positron emission
tomography (FDG-PET) as an adjunct to spiral CT was
evaluated most recently by Bastarikka et al. F-18-
fluorodeoxyglucose is a radioactive tracer that accumulates
in lesions that have a high metabolism, such as rapidly
growing tumors. A total of 911 asymptomatic current or
former smokers were enrolled, the majority (76%) of which
were men who had a median smoking history of 30 pack
years. All participants received a baseline spiral CT screen.
Negative baseline studies were followed with a repeat
screening CT after 12 months. Non-calcified nodules
≥10 mm or smaller growing nodules (7 mm) found on
baseline spiral CT screen were evaluated with FDG-PET.
Nodules that were positive on FDG-PET, defined by visual
analysis rather than standardized uptake value (SUV), were
assessed with percutaneous fine needle aspiration (FNA) or
an intraoperative biopsy. Nodules that were negative on
FDG-PET were followed with a repeat CT 3 months later.
Eleven NSCLC and one SCLC were diagnosed with the
baseline screen (prevalence rate of 1.32%). Two NSCLCs
were discovered on the annual screen (incidence rate of
0.47%). The sensitivity and specificity for diagnosing
malignancy were 69% and 91%, respectively. The positive
predictive value and negative predictive value for FDG-PET
were 90% and 71%, respectively. For nodules found to be
negative on FDG-PET followed by a 3 month repeat CT, the
negative predictive value was 100%. Lung cancer was
diagnosed 3 months earlier in 9 of 13 patients using FDG-
PET compared to short-term follow-up CT. Six PET-positive
nodules were followed by percutaneous FNA, which yielded
diagnosis of lung cancer in four of the six cases. Thus, the
addition of FDG-PET to CT may streamline the work-up of
nodules detected on screening but may not completely
eliminate unnecessary invasive procedures. Further investiga-
tions on FDG-PET have yet to be undertaken with particular
attention to its cost, radiation risk and possibly,
enhanced detection of "pseudodisease".

Future/ongoing studies

To date, studies investigating the effectiveness of CT
scanning in screening have largely been observational.
However, there are several current randomized trials that
may clarify the potential role for CT scans in lung cancer
screening. The National Lung Screening Trial (NLST) is a
multi-center RCT sponsored by the NCI, comparing the
effectiveness of helical CT versus chest x-ray in screening
individuals at high risk for lung cancer. The primary
objective of the study is to determine if CT screening can
reduce the lung cancer mortality in the screened popula-
tion. The study was launched in September 2002 and
enrollment closed in February 2004. Over 50,000 current
or former smokers aged 55–74, without prior cancer history
within the past 5 years, were enrolled. Participants were
randomized into the low-dose CT arm or chest X-ray arm.
Baseline screening is followed by repeat screening annually
for 2 years. Test results are mailed to the participant and
their physician, who determines if follow-up tests or
consultations are needed. Participants are followed over a
course of 8 years using annual health questionnaires. Some
centers also collect blood, urine and/or sputum for future
tlung cancer biomarker studies. Several additional questions
being addressed by the NLST include all cause mortality,
stage of lung cancer at diagnosis, the cost-effectiveness of
screening, what follow-up tests are done for a positive result
and how screening affects the quality of life and smoking
behavior of screened individuals. Results are expected to be

Another multi-center RCT, the NELSON Trial, was
launched in 2003 in the Netherlands and Belgium. The
purpose of the study is to compare lung cancer mortality in
high-risk patients screened with chest CT to those who
receive no screening at all. Participants are enrolled
via population-based recruitment rather than volunteer-
based recruitment. Initial recruitment is based on a
questionnaire focused on general health, alcohol consump-
tion, physical exercise, cancer history, a family history of
lung cancer, body weight and height, education and their
opinion on screening programs in general. As of October
2005, over 15,000 current and former smokers aged 50–75
years have been enrolled and roughly 4000 participants
recruited from the Danish trial will be pooled into the
analysis. It is estimated that approximately 17,300–27,900
participants will be required to demonstrate a 20–25% reduction in lung cancer mortality 10 years after randomiza-
tion, with a power of 80%. This represents the only trial in
which the control arm does not undergo any form of
screening.

Summary

Currently, there are at least four available guidelines on
lung cancer screening (see Table 1). Authoritative
bodies either do not advocate lung cancer screening or if
screening is to be undertaken, they recommend that it
should be within the context of a well-designed trial. The US
Preventative Task Force states that there is insufficient
evidence for or against screening.

In order for screening to be truly effective, it must
demonstrate a decrease in the clinical stage of the disease
and more importantly, death attributable to the disease.
A percentage of pulmonary nodules detected by screening
will ultimately prove to be benign, indolent or clinically
insignificant tumors. Therefore, it is important to consider
what population of patients to screen. It is unclear what
defines "high risk". Should screening only encompass those
deemed high risk? Prior studies such as the NCI sponsored
investigation in the 1970s recruited only males. Although
recent studies have now included females, the defined
smoking history that would place one at high risk varied
among different studies. In addition, screened participants
Lung cancer screening in 2008: Review and update

were mostly current or former smokers recruited on a voluntary basis and therefore, would not represent the overall population demographics that a screening program might include. Increased understanding of the preclinical phase of disease and the implementation of complementary imaging modalities such as PET scan, autofluorescent bronchoscopy and genomic platforms as biomarkers of disease may assist in further classifying the high-risk individual.

Another question pertains to the follow-up of patients testing positive on screening. The Lung Screening Study, which paved the way for the NLST, demonstrated that the number of participants who returned for repeat screening, or compliance rate, was less for those who tested positive than those who tested negative. Therefore, how are we to assure that adequate follow-up is performed for these patients, especially if mass screening is undertaken?

Thus far, efforts at lung cancer screening using low-dose CT have resulted in the detection of early stage disease but not an obvious improvement in the disease-specific mortality rate. There are many limitations to screening that include inherent biases such as lead time bias, length time bias and overdiagnosis bias. In addition, the enhanced detection of early stage disease by CT is also accompanied by detection of benign nodules that eventually may lead to further testing and/or surgical interventions that carry risk, morbidity and mortality. Hence, the jury is still out and the hope is that the ongoing RCTs will shed light on the utility and potential benefit of lung cancer screening.

Main points

- Screening with low-dose CT detects more early stage lung cancer than chest X-ray or sputum analysis. However, a concomitant decrease in late stage disease or decreased mortality has not been observed despite improved identification of early stage disease.
- Lung cancer screening studies that measure survival are influenced by multiple biases such as lead time bias, length time bias and overdiagnosis bias.
- The enhanced detection of nodules can also result in the discovery of "pseudodisease" which may be accompanied by risks such as unnecessary testing and surgeries, cost and psychological burden.
- Lung cancer screening remains a highly debated topic. Previous data have been garnered from cohort studies. The medical community is certainly awaiting the results of ongoing RCTs that may provide the final verdict on lung cancer screening.

Useful links

American Cancer Society http://www.cancer.org/docroot/

Appendix. Questions for lung cancer screening paper:

1. When considering lung cancer, which of the following statements is true?
   a. Lung cancer is the leading cause of cancer deaths in men but not in women.
   b. Only about 10 to 15% of lung cancers are small cell cancers.
   c. Early detection has already improved 1 and 5 years lung cancer survival.
   d. Most patients are symptomatic early in their episode of lung cancer.

2. Studies of screening for lung cancer have included the following modalities and results:
   a. Routine chest X-rays increased identification of resectable tumors and reduced lung cancer mortality.
   b. The addition of screening sputum cytology to screening chest X-ray further reduced lung cancer mortality.
   c. Helical CT screening can increase lung cancer detection rates.
   d. Helical CT screening is very specific for lung cancer. E.g. identifies few benign lesions.

3. Debate surrounding CT screening for lung cancer focuses on the following clinical issues:
   a. The size of the primary tumor (especially non-small cell) is directly related to the likelihood of metastases at detection.
   b. Identifying small tumors by CT clearly results in identification of lung cancer at an earlier Stage.
   c. Some of the early evidence that appeared to be very convincing for the advantage of CT screening was done in observational rather than randomized control trial studies.
   d. More recent data also demonstrates discovery of more resectable tumors and leads to reduced mortality at 1 and 5 years.
4. To consider using screening to improve disease outcomes the following should be true (check all that are correct)
   a. The disease must result in great expense for individuals with the condition, especially for rare diseases.
   b. The condition needs to be identifiable prior to the onset of clinical symptoms that would be expected to lead to evaluation and diagnosis.
   c. Earlier diagnosis should improve outcomes.
   d. Screening tests must be inexpensive and very specific.

5. When considering clinically important cancers, which of the following should be true (check all that are correct)
   a. The disease must result in great expense for individuals with the condition, especially for rare diseases.
   b. The condition needs to be identifiable prior to the onset of clinical symptoms that would be expected to lead to evaluation and diagnosis.
   c. Earlier diagnosis should improve outcomes.
   d. Screening tests must be inexpensive and very specific.

References