Computed Tomography Screening for Lung Cancer

Review of Screening Principles and Update on Current Status

William C. Black, MD

Department of Radiology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire.

Center for the Evaluative Clinical Sciences, Department of Community and Family Medicine, Dartmouth Medical School, Hanover, New Hampshire.

Screening for lung cancer with low-dose computed tomography (CT) is controversial. In favor of screening, lung cancer is the leading cause of cancer death in the United States, and those at greatest risk are identified readily on the basis of age and smoking history. In addition, it is well established that CT is far more sensitive than chest radiography in detecting lung cancer when it is small and asymptomatic. Furthermore, very high rates of survival were reported recently for screen-detected lung cancers in a large, multinational, single-arm observational study. However, a reduction in lung cancer mortality has not been demonstrated to date, and a recent longitudinal study with a simulated control group suggested little or no mortality reduction. In addition, there are important harms from CT screening, including false-positive test results and overdiagnosis. Furthermore, healthcare resources are finite. Therefore, even if the benefits do outweigh the harms, the cost-effectiveness of CT screening for lung cancer still will need to be considered in the context of competing healthcare alternatives. The objectives of this article were 3-fold: 1) to review the basic principles of screening and study designs related to cancer screening, 2) to summarize the results of the observational and analytical studies of CT screening that have been reported to date, and 3) to describe the design of the 2 ongoing, randomized controlled trials of CT screening and what may be learned from these studies in the near future. ***Cancer* 2007;110:2370–84.** © *2007 American Cancer Society.*

KEYWORDS: lung neoplasms, tomography, x-ray computed, mass screening, research design, selection bias, randomized controlled trials, cost effectiveness.

Lung cancer is the leading cause of cancer death in the United States, killing more individuals than cancers of the colon, breast, and prostate combined. In 2007, there will be an estimated 213,380 new diagnoses of lung cancer and 160,390 deaths from this disease.[1] The high ratio of mortality to incidence in lung cancer is a reflection of its poor prognosis. Despite modest improvements in treatment during the last few decades, the overall 5-year survival rate for lung cancer remains only approximately 16%.

Tobacco smoke is the single most important etiologic factor in the development of lung cancer. It is estimated that 90% of all lung cancers are attributed to smoking.[2] Smoking has been associated most strongly with squamous cell and small cell carcinoma[3] but also has been associated with adenocarcinoma, including the bronchioloalveolar subtype.[4] It has been estimated that approximately 10% of lung cancer deaths are attributable to various occupational exposures, including asbestos, environmental tobacco smoke, radon progeny, and arsenic.[5] Heredity also plays a role,

Dr. Black is a site principal investigator and a member of the Executive Committee for the National Lung Screening Trial.

Address for reprints: William C. Black, MD, Department of Radiology, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756; Fax: (603) 650-5455; E-mail: william.black@hitchcock.org

Received May 1, 2007; revision received May 22, 2007; accepted May 29, 2007.

DOI 10.1002/cncr.23059
Published online 16 October 2007 in Wiley InterScience (www.interscience.wiley.com).

because several polymorphisms have been associated with an increased risk of the disease.[6] The role of diet is not clear. Although the consumption of fruits and vegetables rich in β carotene has been associated with a lower risk of lung cancer,[7] 2 randomized controlled trials (RCTs) of β carotene supplementation in heavy smokers have demonstrated a higher risk of lung cancer in individuals who were randomized to receive the supplement than individuals who were randomized to receive the placebo.[8,9]

Despite initial enthusiasm for screening with chest radiography, several RCTs during the 1970s failed to demonstrate a reduction in lung cancer mortality.[10] However, those trials have been criticized on several grounds; and, because of their limited power alone, they do not exclude the possibility of a lung cancer mortality reduction in the range from 10% to 20%. Chest radiography currently is being evaluated in the ongoing Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial,[11] which is powered to detect a mortality reduction[10] as small as 10%.

Regardless of whether it is demonstrated ultimately that chest radiographic screening is effective, there are several reasons to consider screening for lung cancer with low-dose computed tomography (CT). First, CT is far more sensitive than chest radiography. In the Early Lung Cancer Action Project (ELCAP), CT detected almost 6 times as many stage I lung cancers as chest radiography, and most of those tumors measured $\leq$1 cm in greatest dimension.[12] In the Mayo Clinic study, CT screening detected all lung cancers that measured $\geq$8 mm.[13] Second, CT is noninvasive and can be performed in a few seconds during a single breath hold. Currently, the radiation dose for a single examination is approximately 3.3 mSv,[14] only slightly more than the United States average annual effective dose equivalent per person from natural sources,[15] and there is potential to lower the radiation dose from screening CT further by an order of magnitude.[16] Third, the population at greatest risk for lung cancer can be identified readily on the basis of age and smoking history.

However, there also are reasons for healthy skepticism about CT screening for lung cancer. First, its effectiveness is unknown. Although high rates of survival have been reported for screen-detected lung cancers in single-arm observational studies, especially the international ELCAP (I-ELCAP) study,[17] a reduction in lung cancer mortality has not been demonstrated to date. In the previous RCTs of screening with chest radiography, earlier detection of lung cancer did not translate into a reduction in mortality. Second, against any benefit that ultimately may be demonstrated, there are harms that must be weighed, including false-positive screening results and overdiagnosis.[18,19] Radiation exposure also may become a nonnegligible issue with repeated screening over time.[20] Third, there are costs to consider, including not only the costs of the CT screening examinations but also the costs of the follow-up diagnostic tests in patients who screen positive and the costs of treatment in patients who are diagnosed with lung cancer or have other conditions detected by screening. Therefore, even if the benefits do outweigh the harms, the cost-effectiveness of CT screening for lung cancer still will need to be considered in the context of competing healthcare alternatives.[21]

In this article, I provide a review of the basic principles of screening and study designs related to screening. In addition, the results of the observational and analytic studies that have been reported to date are summarized, and their limitations are discussed. After summarizing those results, I describe the design of the 2 ongoing RCTs, the National Lung Screening Trial (NLST) and the Dutch-Belgian Randomized Lung Cancer Multi-slice CT Screening (NELSON) trial, and what we expect to learn from those studies. Finally, I review the latest guidelines pertaining to CT screening for lung cancer.

Screening Principles

The American College of Radiology Task Force on Screening Technologies has adopted the following definition of screening[22]: Screening can be defined as the systematic testing of individuals who are *asymptomatic* with respect to some target disease. The purpose of screening is to prevent, interrupt, or delay the development of advanced disease in the subset of individuals who have a *preclinical* form of the target disease through early detection.

Natural History of Disease

In the analysis of screening, disease is modeled as a dynamic process that evolves over time. The natural history of disease can be represented by a time line with certain key events (Fig. 1).[23,24]

Preclinical phase

The preclinical phase begins with the onset of disease, which may occur long before the onset of signs and symptoms. For example, according to the exponential model of tumor growth[25] (which is probably a gross oversimplification but still is useful for illustration), lung cancer starts as a single cancer cell about 10 microns in diameter and, on average, must grow for about 17 years (35 doublings) to attain a

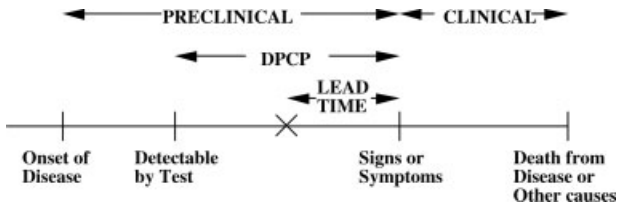

FIGURE 1. Natural history of disease. The preclinical phase of disease begins with its onset and ends when it produces signs or symptoms. The clinical phase begins at the end of the preclinical phase and ends with death. The detectable preclinical phase (DPCP) is subset of the preclinical phase that begins when disease becomes detectable by test. Detection (X) during DPCP advances the time of diagnosis by the duration of lead time. Reprinted with permission from the *American Journal of Roentgenology* (see Black and Welch, 1997[23]).

diameter of 3 cm, assuming a constant doubling time of 180 days.[26,27]

Clinical phase

In the absence of screening, the disease may progress and produce signs or symptoms, which would mark the beginning of the clinical phase. Eventually, the disease may cause some adverse outcome, such as death. For example, starting as a nodule 3 cm in greatest dimension, an average lung cancer must grow for about 2.5 years (5 doublings) to cause death. Death from other causes could occur anywhere along the time line and interrupt the disease process.

Detectable preclinical phase and lead time

An important component of the preclinical phase is the detectable preclinical phase (DPCP) of disease, which begins when the disease becomes detectable by the screening test. For an average lung cancer, which is detectable by CT screening at 5 mm and becomes symptomatic at 3 cm, the DPCP would be about 4 years (8 doublings). The time interval between when a preclinical case is screen-detected and when it would have produced signs or symptoms is known as the *lead time*. The amount of lead time gained by screening depends on the length of the DPCP and the frequency of testing. For example, the average lead time gained from continuous screening would be equal to the average length of the DPCP, whereas the average lead time gained from a single screening would be equal to 50% of the average length of the DPCP.[28]

Critical point

For many diseases, there is a critical point in time beyond which therapy is less effective. For most cancers, the critical point occurs when the primary tumor metastasizes. For screening to be effective, the critical point must occur within the DPCP. If the critical point occurs before the DPCP, then screening is ineffective. It is believed that some lung cancers metastasize when they are only about 1 mm in greatest dimension, at the time that angiogenesis begins to take place.[29,30] If the critical point occurs after the DPCP, then screening is unnecessary.

Overdiagnosis

A key concept in screening is overdiagnosis,[18] ie, the diagnosis of a condition that would not have become significant clinically had it not been detected by screening. Conceptually, 2 forms of overdiagnosis have been distinguished[24]: Type I pertains to the detection of preclinical disease that does not progress or that actually regresses. For example, a pathologist may overcall a case of benign atypical adenomatous hyperplasia as malignant bronchioloalveolar carcinoma (BAC), a distinction on which pathologists often disagree.[31] Type II overdiagnosis pertains to the detection of preclinical disease that progresses, but not rapidly enough to produce any signs or symptoms, before the individual dies from competing causes. Type II overdiagnosis is most prevalent among individuals with slow-growing tumors and short life expectancies because of age or comorbidity. For example, 13 of 48 screen-detected lung cancers in the Mayo Clinic study[32] had volumetric doubling times longer than 400 days. Given the median greatest dimension of 12 mm, such a lung cancer would not be expected to cause death for >10 years, during which time death from another cause would be likely to occur in those who are screened for lung cancer because of age and smoking history.[33,34]

Overdiagnosis is not recognizable in a living individual because, by definition, it cannot be confirmed until the patient dies from other causes. Furthermore, overdiagnosed individuals usually are treated, and the lack of subsequent signs and symptoms usually is attributed to cure from treatment. Consequently, overdiagnosis in cancer screening has received relatively little attention until recently.

Test Performance

Sensitivity

Sensitivity is defined as the ratio of true-positive to true-positive plus false-negative test results. In the evaluation of screening tests, true-positive results usually refer to screen-detected cases, whereas false-negative results refer to interval cases, ie, those that surface clinically after a negative screen during some specified period of time.[35] Sensitivity will be overestimated if the number of true-positive results is over-

estimated because of overdiagnosis or if the number of interval cases is underestimated because of incomplete clinical follow-up of individuals who screen negative. Another approach to measuring sensitivity is to review previous screening results of screen-detected cases and count as false-negative those cases for which the cancer can be observed retrospectively. Because this approach includes as false-negative results those patients who still are asymptomatic, it results in a lower estimate of sensitivity than the approach using interval cases. With either approach, sensitivity varies from 0% to 100% as a cancer grows from a single cell to a large mass, and it often is helpful to report sensitivity according to size or anatomic extent.[36]

Prevalence, incidence, and prevalence ratio

The proportion of the screened population that has disease detected at the first, or prevalence, screen is known as the *detected prevalence*, which is a composite measure of the prevalence of preclinical disease and the sensitivity of screening. The proportion of the screened population that has disease detected at the subsequent, or incidence, screens is the *detected incidence*. The *prevalence ratio*, which is the ratio of the detected prevalence to the detected incidence, is another measure of sensitivity and is approximately equal to the number of screening intervals that comprise the length of the DPCP. For example, if the prevalence ratio in an annual program of screening is 3.0, then the length of the DPCP is 3 years. However, if overdiagnosis occurs during the prevalence and incidence screens, then the prevalence ratio and the length of the DPCP will be underestimated. When overdiagnosis is substantial, the prevalence ratio may approach 1.0.

Stage distribution

The stage distribution of detected cancer at the prevalence screen is a function of the level of surveillance of the participants before screening, the test sensitivity, and overdiagnosis. The stage distribution of detected cancer at the incidence screen(s) depends less on the level of surveillance of the participants before screening than the stage distribution at the prevalence screen because of the effects of the intervening prevalence screen.

All of the measures of test performance described above—sensitivity, prevalence, prevalence ratio, and stage distribution—are functions of lead time and are related to the potential for early detection of the target disease. However, as pointed out by Cole and Morrison[37] nearly 30 years ago, earlier diagnosis is a "double-edged sword," and measures of sensitivity should not be used as the sole indicators of screening effectiveness.

Specificity

Specificity is defined as the ratio of true-negative to true-negative plus false-positive test results. The specificity of a test is equal to 1 minus its false-positive rate, which is a measure of 1 of the 2 major harms of screening (the other major harm is overdiagnosis). Similar to sensitivity, specificity is a function of the positivity threshold. For example, the specificity of screening CT for lung cancer with regard to all detectable, noncalcified nodules is only about 50%.[13] Most of these patients will only require a follow-up noncontrast CT in 6 to 12 months.[38] However, the specificity with regard to noncalcified nodules that measure $\geq$10 mm is >90%,[12] but most of those patients will require a more invasive evaluation. The specificity usually is higher (and the false-positive rate usually is lower) during the incidence screens than during the prevalence screen, because stability over time allows false-positive results at the prevalence screen to be reclassified correctly as negative.

Positive predictive value

The positive predictive value is defined as the ratio of true-positive to true-positive plus false-positive test results. Because the prevalence of preclinical disease usually is very low (<5%), the positive predictive value usually also is low. For example, if the prevalence of preclinical lung cancer is 1% and the sensitivity and specificity are 95%, then the positive predictive value is only 16% ($1 \times 95/[1 \times 95 + 99 \times 5]$).

Primary Outcomes of Screening

Disease-specific mortality

The purpose of screening, as stated above, is to prevent or delay the development of advanced disease and its adverse effects. Therefore, when considering the adverse event of death, disease-specific mortality is the most appropriate outcome measure in the evaluation of screening effectiveness.[39] The disease-specific mortality rate in a population is the ratio of the number of deaths from the target disease to the number of person-years of observation. Regardless of the study design, effectiveness usually is expressed in terms of the relative risk reduction, which is equal to the disease-specific mortality in the control group minus the disease-specific mortality in the screened group divided by the disease-specific mortality in the control group. Deaths resulting from positive (true-positive and false-positive) screening results and treatment of the target disease usually are counted as disease-specific deaths.

Absolute versus relative risk reduction

The effectiveness of screening usually is reported as a relative risk reduction, as discussed above. However, this metric used alone can be misleading, because it conveys no information about an individual's baseline risk. It is well recognized that a more appropriate measure of effectiveness for all types of interventions is the absolute risk reduction,[40] which is the product of risk and relative risk reduction. For example, suppose a screen-eligible individual has a 1% probability of dying from some disease over the next 20 years. If the relative risk reduction from screening (or some other intervention) is 50%, then the absolute risk reduction is 0.5%. The reciprocal of the absolute risk reduction is the number needed to screen to prevent 1 death or adverse event. In the example provided above, this number is 200 (1/.005). It was demonstrated in a telephone survey that respondents were much more likely to accept screening when its effectiveness was presented in terms of the relative risk reduction rather than the number needed to screen.[41]

All-cause mortality

Although the most widely accepted endpoint in studies of cancer screening is disease-specific mortality, the validity of this endpoint rests on the assumption that the cause of death can be determined accurately. Misclassification in the cause of death can result in either over or underestimation of screening effectiveness, depending on how the misclassification occurs.[42] A complementary endpoint is all-cause mortality, which depends only on the accurate determination of deaths and when they occur.

Although a statistically significant effect on all-cause mortality rarely is demonstrable with screening (because the target disease usually is responsible for only a small proportion of all deaths), it is useful to examine all-cause mortality along with disease-specific mortality for 3 reasons. First, examination of all-cause mortality may reveal major deficiencies in a study, such as flaws in the randomization or ascertainment of vital status. Second, examination of all-cause mortality helps ensure that a major harm or benefit of screening is not being missed. Third, examination of all-cause mortality puts the magnitude of expected benefits from screening into an appropriate perspective for decision making.

Survival

Survival is the most common measure used in the evaluation of treatment effectiveness, especially for those with a diagnosis of cancer. Although survival is an appropriate measure for the evaluation of treatment, this measure is inappropriate for the evaluation of screening for 2 reasons. First, in screening, the vast majority of participants never develop clinical manifestations of the target disease. Second, in screening, diagnosis occurs after the intervention. Therefore, to the extent that screening advances the time of diagnosis, survival is a biased measure of screening effectiveness.[23,43] Three distinct biases affect the comparison of survival in screen-detected versus clinically detected cases of disease.

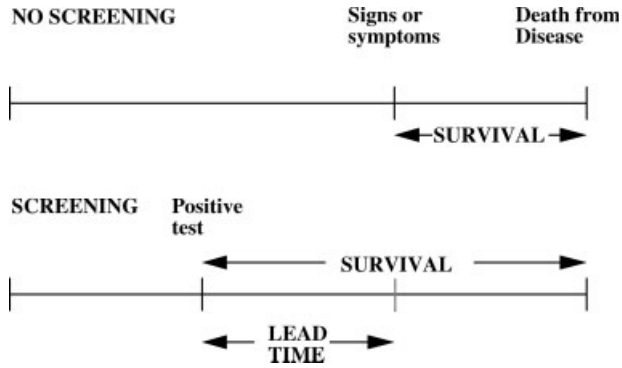

FIGURE 2. Lead time bias. Without screening, diagnosis occurs when clinical signs or symptoms develop. With screening, the time of diagnosis is advanced by lead time provided by a positive test result. If earlier diagnosis has no effect on the time of death from disease, then survival with testing is equal to survival without testing plus lead time. Reprinted with permission from the *American Journal of Roentgenology* (see Black and Welch, 1997[23]).

Lead time bias. Lead time bias pertains to comparisons that are not adjusted for the timing of diagnosis (Fig. 2). If survival is measured from the time of diagnosis, which is the usual approach, then the comparison between screen-detected cases and cases that are diagnosed clinically is biased, regardless of the real effect of earlier diagnosis. In the simple case in which earlier diagnosis has no real effect on the length of survival, screening will appear to prolong survival by the lead time. (In the more general case in which earlier diagnosis has a real effect on survival, either positive or negative, screening will appear to prolong survival by the sum of the lead time and the real effect.) Unfortunately, adjusting for lead time is problematic, because it is usually unknown and variable, which is related to a second bias.

Length bias. Length bias pertains to comparisons that are not adjusted for the rate of disease progression. The probability that a case will be detected by screening is directly proportional to the length of its

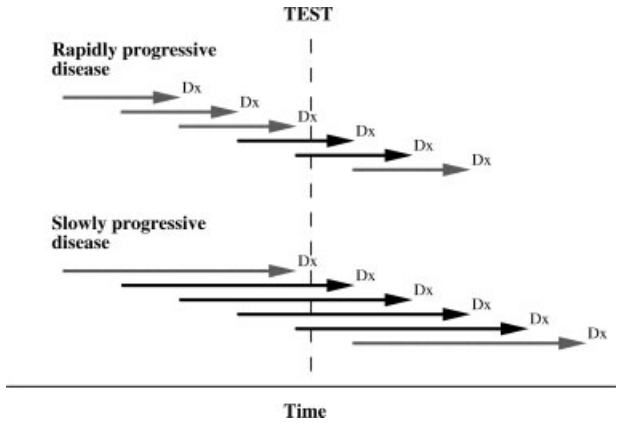

FIGURE 3. Length bias. The probability of detection is related to the rate of disease progression. The length of each arrow represents the length of detectable preclinical phase, from initial detectability to clinical diagnosis (Dx). Testing at a single moment in time detects 4 patients with slowly progressive disease but only 2 patients with rapidly progressive disease (black arrows). Patients who do not have disease detected by the test (gray arrows) are diagnosed clinically, either before or after the time of testing. Reprinted with permission from the *American Journal of Roentgenology* (see Black and Welch, 1997[23]).

DPCP, which is related inversely to its rate of progression (Fig. 3). Therefore, cases detected by screening are more likely to be slowly progressive than those that are not detected by screening and ultimately present clinically. This principle is evident in reports of CT screening for lung cancer in which there is a strong over-representation of slowly growing adenocarcinomas, especially BAC, and an under-representation of rapidly growing small cell carcinomas.[32]

Overdiagnosis bias. Overdiagnosis bias pertains to comparisons that are not adjusted for overdiagnosis, which can be a major cause of confusion and harm in cancer screening.[18] In addition to overestimations of survival and cure rates, overdiagnosis causes overestimations of the accuracy of the screening test, the prevalence and incidence of the target disease, and the stage shift because of screening. Ironically, overdiagnosis is often the single greatest source of harm in cancer screening, but it is the harm least familiar to the general population.[44]

Secondary Outcomes of Screening

Survival

Although survival statistics should not be used as a primary outcome measure of screening effectiveness, they may be useful as secondary outcome measures if placed in the proper context, such as in simulation models that adjust for lead time, length, and overdiagnosis biases.

Morbidity

The morbidity caused or prevented by screening may be substantial. Although most imaging tests that are used for cancer screening are noninvasive and painless for most participants, some may experience significant discomfort. The most common cause of morbidity is a false-positive screening test, which may lead to invasive testing, including percutaneous needle biopsies or even surgery in addition to increased anxiety about the cancer. Perhaps the most common cause of serious morbidity is overdiagnosis, which results in unnecessary treatment in addition to the diagnostic workup. However, screening also may prevent some morbidity from the cancer and its treatment if the cancer is detected when it is in an early stage. In addition, a true-negative screening result may relieve cancer anxiety. Screening also may have a negative or positive effect or morbidity through the detection of nontarget cancers or other diseases.

Cost

Costs related to screening are important, because healthcare resources are finite.[45] Direct medical costs include those related to the screening tests and subsequent diagnostic evaluations, treatment for the target cancer (and other conditions first detected through the screening process), and complications from testing and treatment. Nonmedical and opportunity costs include lost wages, traveling and lodging costs for the original screening tests and all subsequent evaluations, and treatment for the screenee (and caregiver).

Study Designs

Observational studies

Three major types of observational studies have been used in the evaluation of screening interventions. Correlation studies describe the relation between screening frequency and disease-specific mortality rates in populations across time and location.[46] In cohort studies, a group of individuals usually is followed prospectively over time with or without a nonrandomized control group.[47] In case-control studies, cases are identified after they have experienced an adverse outcome, such as the death from the target cancer, and are matched to controls (usually on the basis of age and other relevant variables) who have not experienced the adverse outcome; then, the screening histories of the cases and controls are compared.[48] All observational study designs are vulnerable to selection bias, in which the risk for the target disease may be higher or lower in the screened population than in the control group (assuming there is a defined control

group). If the risk is higher in the screened population, then the selection bias will be against screening. Conversely, if the risk is lower in the screened population, then the selection bias will favor screening. Observational studies also are vulnerable to confounding by other known and unknown differences in the populations, such as access to medical care or an unknown genetic predisposition to the target disease. Consequently, the observed differences in screened and unscreened populations may be cause by factors other than screening.

Although observational study designs do not provide reliable estimates of screening effectiveness, they may provide useful estimates of test performance and secondary outcomes if follow-up is sufficient. Close clinical follow-up of participants who screen negative is necessary to determine the number of interval cancers and test sensitivity. Close clinical follow-up of participants who screen positive is necessary to determine the number of false-positive results, test specificity, complications of screening and treatment, survival in patients with screen-detected cancer, quality of life, and costs.

Randomized controlled trials

RCTs are considered the best method for determining the effectiveness of any intervention, because they distribute the known and unknown confounding variables equally among the different groups, thereby ensuring that differences in outcome can be attributed to differences in intervention.[49] RCTs are particularly appropriate for screening, because they eliminate the early detection biases and the potential for confounding by variables associated with access to screening.[50] To date, 17 RCTs of cancer screening that used the disease-specific mortality endpoint have been completed, including 14 studies that pertained to imaging.[42,51] In addition, 2 large RCTs of cancer screening with imaging currently are underway in the United States: the PLCO Cancer Screening Trial[11] and the NLST.[52]

Although RCTs are considered the best method for determining the effectiveness of cancer screening, there are many particular details to consider in the planning, execution, and interpretation of screening RCTs.[50] The sample size of participants required for a statistically significant effect is a function several variables, including the disease-specific mortality in the screen-eligible population and compliance.[53] Because participants in screening RCTs must have no clinically evident disease at the time of entry, their disease-specific mortality will be much less than that of patients with clinical disease. Consequently, screening trials generally require far more participants and longer periods of observation than treatment trials. Informed consent concerning the potential benefits and harms and confidentiality issues is critical for ethical reasons and for the integrity of the RCT.[54] If they are informed properly before randomization, then participants will be more likely to comply subsequently with the assigned regimen and follow-up. Because screening RCTs are analyzed according to randomization irrespective of compliance, lack of compliance in either group will result in an underestimate of the screening effect (positive or negative).

The interpretation strategy should be defined well with respect to the workup of positive findings. In addition, the treatments for different stages of the cancer should be made explicit. The screening intervention is more than simply a performance of the screening test. The screening intervention is a full regimen that begins with the screening examination and also includes the subsequent workup of positive findings and treatment of the cancers diagnosed. The number of screening rounds and their frequency also must be chosen.[39,55] The intensity and duration of screening should be sufficient to produce an effect on disease-specific mortality (assuming there is an effect). To the extent that costs or other constraints limit the number of screening rounds and the duration of observation, the RCT will underestimate the potential impact of screening. Computer simulations may be helpful in exploring the magnitude of this underestimation.

After randomization, both groups should be followed closely to obtain accurate records of their health and medical care related to cancer and screening regimen. Follow-up should continue for at least 5 years from randomization and perhaps several years longer for screening regimens with longer lead times. Because the major endpoint is disease-specific mortality, cancer deaths and the circumstances surrounding them should be well documented with death certificates and medical records; because, otherwise, the cause of death may be difficult to verify. In addition, all morbid events related to the disease and complications of screening should be well documented.

Results of Cohort Studies

General

Since the early 1990s, there have been 12 reported observational studies of CT screening, all of which have been single-arm cohort studies (Table 1).[12,13,17,19,56–70] All of those studies involved volunteers who were at high risk for lung cancer on the basis of smoking, occupational exposures to carcinogens, or age.

TABLE 1
Studies of Computed Tomography (CT) Screening: Study Population, CT Technique, and Compliance

Name or institution	Reference(s)	Year started	Population			CT technique			Compliance incidence screen: no. (%)[‡]
			No.*	Age, y[†]	Smoker, %	kVp	mA	Slice thickness, mm	
ELCAP, US	Henschke et al., 1999[12]	1992	1000	67	100	140	40	10	841 (84)
I-ELCAP, International	Henschke et al., 2006[17]	1993	31,567	61	87	NR	NR	NR	NR
ALCA, Japan	Sobue et al., 2002[66]	1993	1611	60[§]	86	120	50	10	1180 (73)
University of Munster, Germany	Diederich et al., 2002, 2004[57,58]	1995	817	53	100	120	50	5	668 (82)
Shinshu University, Japan	Sone et al., 1998, 2001[67,68]	1996	5483	64	46	120	50	10	4425 (81)
Finnish Institute of Occupational Health, Finland	Tiitola et al., 2002[70]	1998	602	63	100	140	125	10	NA
Mayo Clinic, US	Swensen et al., 2002, 2003, 2005[13,19,69]	1999	1520	59	100	120	40	5	1478 (97)
Hitachi Health Care Center, Japan	Nawa et al., 2002[64]	1998	7956	70	62	120	50	10	5568 (70)
PALCAD, Ireland	Macredmond et al., 2004, 2006[61,62]	2000?	449	55	100	130	50	10	413 (92)[‖]
University of Milano, Italy	Pastorino et al., 2003[65]	2000	1035	58	100	140	40	10	996 (96)
Nuclear Fuel Workers, US	Miller et al., 2004[63]	2000	3598	>40	66	NR	NR	NR	NR
NY-ELCAP, US	NY-ELCAP, 2007[56]	2000	6295	66	100	120	40	1.25–10	5134 (82)
LSS, US	Bach et al., 2001, 2007[59,60]	2000	1660[¶]	60[#]	100	120–140	60	5	1398 (84)

ELCAP indicates Early Lung Cancer Action Program; NR, not reported; ALCA, Anti-Lung Cancer Association; NA, not applicable; PALCAD, ProActive Lung Cancer Detection; LSS, Lung Screening Study (CT arm only).

* The number of study participants is the same as the number completing prevalence screen unless indicated otherwise.

† Median unless otherwise indicated.

‡ Number of participants at first incidence screen.

§ Average of upper and lower age range (median not reported).

‖ Number participants completing first 2 incidence screens (number completing first screen was not reported).

¶ Only 1586 participants (96% of those randomized to CT) completed prevalence screen.

Approximated from reported distribution (median not reported).

In all of the studies, screening was performed without intravenous contrast and, in all but 1 study, with a low-dose (≤50 mA) technique. Although none of the studies provided a reliable estimate of the effect of CT screening on lung cancer mortality, several of the studies provided useful information on test performance and secondary outcomes. Among the 9 studies that reported compliance (the percentage of participants that returned for the first repeat screen), the median compliance was 82% (range, 70%–97%).

Test performance

Sensitivity. Among the 8 studies that reported the number of interval cases during all screening rounds, the sensitivity ranged from 81% to 100%, and the median value was 96% (Tables 2, 3). In the study that had the fewest participants lost to follow-up (only 1 of 1520 participants),[19] the sensitivity was 92% (61 screened-detected cases, 3 interval cases, and 2 cases detected by sputum cytology only). In the same study, the sensitivity ranged from 50% for lung cancers that measured 4 to 8 mm in greatest dimension to 100% for lung cancers that measured >8 mm.[13] In the I-ELCAP study,[17] the sensitivity was 99% (479 screen-detected vs 5 interval cancers), but the investigators did not report the number of participants with negative screening examinations who were lost to follow-up.

When sensitivity was estimated from a retrospective review of screen-detected cases, it was much lower. In the largest study that used this approach,[71] 83 lung cancers were detected at the first incidence screen. In retrospect, 32 of those cancers had been missed on the prevalence screen, 20 because of detection errors and 12 because of interpretation errors. Thus, the sensitivity in that study was only 61% (51 of 83 cancers). However, most of these missed lung cancers still were in the early stage at the incidence screening, and there were no reported interval cases.

Prevalence, incidence, and prevalence ratio. The prevalence of screen-detected lung cancer was reported in all 12 studies (Table 2). The median prevalence was 0.9% (range, 0.4%–2.7%). This wide range probably was caused largely by the variation in age and smoking history among the participants. The study with the lowest prevalence had the lowest proportion of smokers, 46%, whereas the study with the second lowest prevalence had the lowest median age, 55

TABLE 2
Studies of Computed Tomography Screening: Results of Prevalence Screening

Name or institution	Reference(s)	No. of positive results (%)	No. of lung cancers (%)	NSCLC, %	Stage I, %	Adeno-Ca, %	No. of interval cases	Sensitivity, %*	Specificity, %
ELCAP, US	Henschke et al., 1999[12]	233 (23)	27 (2.7)	96	88	78	2	93	79
I-ELCAP, International	Henschke et al., 2006[17]	NR	405 (1.3)	NR	86†	76‡	5	99	NR
ALCA, Japan	Sobue et al., 2002[66]	186 (12)	13 (0.8)	100	77	77	NR	NR	89
University of Munster, Germany	Diederich et al., 2002, 2004[57,58]	350 (43)	11 (1.3)	91	70	45	5	69	58
Shinshu University, Japan	Sone et al., 1998, 2001[67,68]	279 (5)	23 (0.4)	100	100	83	NR	NR	95
Finnish Institute of Occupational Health, Finland	Tiitola et al., 2002[70]	111 (18)	5 (0.8)	100	0	40	0	100	82
Mayo Clinic, US	Swensen et al., 2002, 2003, 2005[13,19,69]	782 (51)	30 (2)	93	75	77	1§	97	50
Hitachi Health Care Center, Japan	Nawa et al., 2002[64]	541 (7)	37 (0.45)	95	89	95	NR	NR	94
PALCAD, Ireland	Macredmond et al., 2004, 2006[61,62]	109 (24)	2 (0.4)	50	100	NR	1	67	76
University of Milano, Italy	Pastorino et al., 2003[65]	61 (6)	11 (1.1)	100	55	91	0	100	95
Nuclear Fuel Workers, US	Miller et al., 2004[63]	1139 (32)	22 (0.6)	100	NR	NR	NR	NR	69
NY-ELCAP, US	NY-ELCAP, 2007[56]	906 (14)	101 (1.6)	94	97	67	3	97	87
LSS, US	Bach et al., 2001, 2007[59,60]	325 (20)	30 (1.8)	97	55	63	2	94	78

NSCLC indicates nonsmall cell lung cancer; Adeno-Ca, adenocarcinoma; ELCAP, Early Lung Cancer Action Program; NR, not reported; ALCA, Anti-Lung Cancer Association; NA, not applicable; PALCAD, ProActive Lung Cancer Detection; LSS, Lung Screening Study (computed tomography arm only).

* Based on interval method (see text).

† Based on all detected lung cancers, because NSCLCs were not reported.

‡ Based only on stage I lung cancers.

§ Detected by sputum cytology only.

TABLE 3
Studies of Computed Tomography Screening: Results of Incidence Screenings

Name or institution	Reference(s)	Incidence screens	No. of positive results (%)	No. of lung cancers (%)	NSCLC, %	Stage, %	Adeno-Ca, %	No. of interval cases	Sensitivity, %*	Specificity, %
ELCAP, US	Henschke et al., 2006[12]	1184	40 (3.4)	7 (0.6)	86	83	NR	0	100	97
I-ELCAP, International	Henschke et al., 2006[17]	27,456	NR	74 (0.3)	NR	86†	48‡	0	100	NR
ALCA, Japan	Sobue et al., 2002[66]	7891	721 (9.1)	19 (0.2)	95	83	74	NR	NR	91
University of Munster, Germany	Diederich et al., 2002, 2004[57,58]	1735	89 (5.1)	10 (0.6)	100	70	NR	0	100	95
Shinshu University, Japan	Sone et al., 1998, 2001[67,68]	8303	309 (3.7)	37 (0.4)	92	94	86	NR	NR	97
Finnish Institute of Occupational Health, Finland	Tiitola et al., 2002[70]	NA	NA	NA	NR	NR	NR	NA	NA	NA
Mayo Clinic, US	Swensen et al., 2002, 2003, 2005[13,19,69]	5365	NR	31 (0.6)	90	61	42	4§	89	NR
Hitachi Health Care Center, Japan	Nawa et al., 2002[64]	5568	148 (2.7)	4 (0.07)	100	100	100	0	100	97
PALCAD, Ireland	Macredmond et al., 2004, 2006[61,62]	826‖	NR	3 (0.4)	33	100	NR	0	100	NR
University of Milano, Italy	Pastorino et al., 2003[65]	996	34 (3.4)	11 (1.1)	100	100	64	0	100	98
Nuclear Fuel Workers, US	Miller et al., 2004[63]	NR	NR	NR	NR	NR	NR	NR	NR	NR
NY-ELCAP, US	NY-ELCAP, 2007[56]	6014	361 (6.0)	20 (0.3)	100	85	35	0	100	94
LSS, US	Bach et al., 2001, 2007[59,60]	1398	360 (25.8)	8 (0.6)	63	40	63	0	100	75

NSCLC indicates nonsmall cell lung cancer; Adeno-Ca, adenocarcinoma; ELCAP, Early Lung Cancer Action Program; NR, not reported; ALCA, Anti-Lung Cancer Association; NA, not applicable; PALCAD, ProActive Lung Cancer Detection; LSS, Lung Screening Study (computed tomography arm only).

* Based on interval method (see text).

† Based on all detected lung cancers, because NSCLCs were not reported.

‡ Based only on stage I lung cancers.

§ Three interval cases and 1 detected by sputum cytology alone.

‖ Number $\geq$826 but not reported.

years. ELCAP,[12] the study with the highest prevalence, had the second highest median age, 67 years, and the highest reported median pack-years of smoking, 45 pack-years (not shown in Table 1).

The incidence of screen-detected lung cancer was lower than the prevalence in 9 of the 10 studies that reported both, and the 2 measures were correlated only weakly (Table 3). The median incidence was 0.4% (range, 0.1%–1.1%), and the median prevalence ratio was 3.4 (range, 0.9–6.5).

Stage distribution. The percentage of screen-detected nonsmall cell lung cancers (NSCLCs) that were stage I at diagnosis was reported in 10 of the 12 studies (Tables 2, 3). Excluding the 2 studies with <10 NSCLCs, the median proportion of stage I NSCLCs was 85% (range, 68%–96%). There was no consistent difference in these percentages at the prevalence versus incidence screens. Although the median percentage was much higher than that reported in previous screening studies of chest radiography—39% in the Mayo Lung Project[72]—this high median percentage was not necessarily an indication of effectiveness. Instead, the high median percentage may reflect a high level of surveillance of the participants before screening and overdiagnosis.

Specificity. Specificity could be estimated from the prevalence screens in 11 of the 12 studies (Table 2) and from the incidence screens in 7 studies (Table 3). At the prevalence screen, the median specificity was 82% (range, 50%–95%). This wide range was largely caused by variations in the definition of a positive test result, which has evolved over time. At the Mayo Clinic, which had the lowest specificity, all noncalcified nodules were considered positive, regardless of their size. However, this specificity increased to about 75% when it was recalculated using the 5-mm greatest dimension threshold (39% of the noncalcified nodules measured <4 mm, and 50% measured 4–8 mm).[13] The specificity was 87% in the most recently reported study, the New York-ELCAP, which used a 5-mm threshold. The remaining difference between the specificities in the Mayo Clinic and New York-ELCAP studies may have been because of the higher prevalence of histoplasmosis at the former site and differences in interpretation that have yet to be elucidated.

The specificities, as expected, were higher in the incidence screens (Table 3) than in the prevalence screen for all 7 studies in which both percentages could be estimated. The median specificity was 97% per screening (range, 91%–98% per screening). These percentages do not include results from the Mayo Clinic, because their study did not report false-positive results for the incidence screens on a per individual basis separate from the prevalence screen. However, the Mayo Clinic study did report that 69% of all participants had at least 1 false-positive screening result after 1 prevalence screen and 4 incidence screens. This false-positive rate probably would have been close to 35% had the study used the 5-mm positivity threshold for pulmonary nodules.

Positive predictive value. The positive predictive value could be estimated from the prevalence screens in 11 of the 12 studies and from the incidence screens in 7 studies. At the prevalence screens, the median positive predictive value was 6% (range, 2%–15%). At the incidence screens, the median positive predictive value was 10% (range, 3%–24%).

Primary outcomes: Mortality

The lung cancer-specific mortality rate was estimated in only 1 of the 12 cohort studies: the Mayo Clinic study.[19] Among the 1453 participants who had at least 1 incidence screen and a mean observation time of 3.8 years, the incidence lung cancer-specific mortality rate was 1.6 per 1000 person-years, and the total mortality rate was 6.0 per 1000 person-years. The lung cancer mortality also was calculated for a subset of this population—men aged >50 years—and was compared with the lung cancer mortality derived from a similar subset from the Mayo Lung Project.[72] With 4 years of follow-up, there was no difference in the incidence lung cancer mortality rates in these 2 populations (2.8 vs 2.0 per 1000 person-years; $P = .43$). However, the investigators acknowledged that this comparison may have been subject to a substantial degree of confounding, because it was not based on a randomized control group, and the length of follow-up may have been insufficient to observe an effect from CT screening.

Deaths resulting from surgery for lung cancer were reported in 4 of the studies. In the Anti-lung Cancer Association (ALCA) study,[66] there were 2 deaths among 29 surgeries (6%). Both deaths resulted from infection 6 months and 9 months after surgery and occurred in patients with stage I lung cancer. In the German study,[57] there were 2 postoperative deaths among 20 surgeries (10%), and both deaths occurred in patients with stage I lung cancer. In the Mayo Clinic study,[19] there was 1 postoperative death among 53 surgeries (2%) for lung cancer. In the Italian study,[65] there were no postoperative deaths among 21 surgeries. In the I-ELCAP study,[17] there were 2 deaths among 411 lung cancer surgeries (0.5%). For comparison, the operative mortality for

lobectomy, the most commonly performed procedure for resection of lung cancer, is approximately 4% in the United States overall.[73]

Secondary outcomes: Survival

Three of the 12 cohort studies reported survival statistics. In the ALCA study,[66] the 5-year overall survival rate (counting deaths from all causes) for the 36 participants with screen-detected lung cancers (32 detected by CT, 4 detected by sputum cytology only) was 71%. The 5-year survival rate was higher for the prevalence cases than the incidences cases (76% vs 65%), but this difference was not statistically significant. Survival was significantly higher for adenocarcinomas, which comprised 75% of the CT-detected cancers, than for other cell types.

In the Finnish study,[70] which consisted of a single CT screening for smokers with asbestos exposure, the 5-year survival rate for screen-detected lung cancer was 0%. All 5 participants with screen-detected lung cancers died from their disease within 21 months of detection. However, none of those cancers were stage I at the time of diagnosis. Perhaps the asbestos exposure caused pleural and parenchymal changes that decreased the sensitivity of CT screening for early-stage lung cancer in that study.

In the I-ELCAP study,[17] the 10-year lung cancer-specific survival (treating deaths from other causes as censored) was estimated for various subgroups of the 484 participants who were diagnosed with lung cancer (479 screen-detected cases and 5 interval cases). The reported 10-year survival rate was 80% for the entire group, 88% for the subgroup with clinical stage I lung cancer, and 92% for the subgroup with clinical stage I lung cancer who underwent surgical resection within 1 month of diagnosis. In addition, as mentioned above, the reported sensitivity of CT screening was 99%. Although these reportedly high survival rates and sensitivity suggest than CT screening may be very effective in reducing lung cancer mortality, these statistics are subject to several forms of bias. The survival rates and sensitivity almost certainly were overestimated because of overdiagnosis, although the extent to which overdiagnosis occurred is unknown. In addition, loss to follow-up was reported neither for participants with known lung cancer nor for the vast majority of participants without the diagnosis. Loss to follow-up of these 2 groups may have caused an overestimation of survival and sensitivity, respectively. Furthermore, the death review process was not explicit for those who were known to have died, and ascertainment bias may have caused a further overestimation of survival.

No other complications were reported as commonly as postoperative mortality. None of the studies reported on quality of life or costs.

Results of Analytic Studies

Longitudinal analysis

One major limitation of the single-arm cohort studies described above is that they had no internal control group (or explicit external control group, with the exception of the Mayo Clinic Study) to allow for any comparison of lung cancer outcomes. To address this limitation, Bach et al.[74] used a validated lung cancer prediction model based on age, sex, and smoking history to estimate the expected numbers of various lung cancer outcomes among a combined cohort of 3 single-arm studies (including the Mayo Clinic and Italian studies described above and a study from the Moffitt Cancer Center in Tampa, Florida, which has not been reported elsewhere). To assess the effectiveness of CT screening, the investigators compared the observed numbers of lung cancer outcomes with those predicted. They observed a >3-fold increase in the number of new lung cancer cases (144 observed vs 44.5 predicted) and a 10-fold increase in lung cancer resections (109 vs 10.9). However, they observed no decrease in advanced lung cancer cases (42 vs 33.4) or in lung cancer deaths (38 vs 38.8). Bach et al acknowledged that a longer duration of screening or follow-up (median, 3.9 years) may have revealed a benefit to screening and that there is uncertainty about the accuracy of the prediction model. Nevertheless, they cautioned that, to date, there is no strong evidence that CT screening is effective.

Cost-effectiveness analyses

Several cost-effectiveness analyses of CT screening for lung cancer have been published.[45] The results, which were reported in terms of baseline incremental cost-effectiveness ratios, have ranged from very favorable (<$2500 per life-year saved[75,76]) to marginal (>$100,000 per quality-adjusted life-year saved[77,78]). The 2 analyses that produced the marginal results made adjustments for overdiagnosis, the effects of smoking on competing mortality, and quality of life; whereas the other analyses did not. However, all of these analyses were based on "shallow" stage-shift models, which have not been validated.[45] More sophisticated "deep" models, which simulate the natural history of lung cancer, are under development. Some of these models eventually will be used in conjunction with the results of the ongoing NLST to help inform screening policy.

Randomized controlled trials

Lung screening study. Two small RCTs of CT screening have been conducted to demonstrate the feasibility of conducting a large RCT with a lung cancer mortality endpoint.[59,79] In the larger study, the Lung Screening Study (LSS),[59] 3318 smokers ages 55 to 74 years were randomized to receive screening with either low-dose CT or posterioanterior (PA) chest radiograph. In the CT arm, compliance with screening was 96% at baseline and 93% at 1-year; whereas, in the chest radiograph arm, compliance was 93% and 80%, respectively.[60] The sensitivity (based on prevalent and interval cases) was higher for CT (94%; 30 of 32 cancers) than chest radiography (64%; 7 of 11 cancers). In addition, CT screening detected more stage I lung cancers than chest radiographic screening (18 vs 8 cancers) and more stage II and IV lung cancers (20 vs 8 cancers). The specificity (based on prevalence and incidence screens) was higher in the chest radiograph arm than in the CT arm (91% vs 78%; however, about 50% of all nodules that were detected by CT were measured <5 mm in greatest dimension). More participants in the CT arm underwent an invasive biopsy than in the chest radiograph arm (64 vs 22 patients[59,60]), and approximately 33% of those procedures were resections or open surgical biopsies.[80] Although the LSS confirmed the feasibility of conducting a larger RCT, it was not designed to determine the effect of CT screening on mortality.

National lung screening trial. The NLST[52] is an ongoing multicenter RCT funded by the National Cancer Institute (NCI). The primary objective of the study is to determine whether screening with chest CT reduces lung cancer mortality relative to screening with chest radiography in a high-risk cohort. The study is designed to have 90% power to detect a lung cancer mortality reduction of 20% about 6 years after randomization.

The eligibility criteria included ages 55 to 74 years, ≥30 pack-years of smoking, no symptoms or signs of lung cancer, no prior history of lung cancer or any other cancer within 5 years, and no chest CT within 18 months. Between August 2002 and April 2004, 53,464 participants were recruited into the study from 33 trial sites across the United States. Participants were randomized to receive 3 rounds (1 prevalence and 2 annual incidence) of screening with either low-dose chest CT or PA chest radiography. The CT technique included low-dose (<80 mA) and multidetector arrays (4-16 rows) with a maximal slice thickness of 2.5 mm. All CT and radiographic findings were recorded on standardized forms, and the management of participants with positive findings (noncalcified nodules ≥4 mm) was guided by a detailed workup algorithm. All participants have been and will continue to be followed by questionnaires at least annually. Death certificates are obtained for all deaths; and, when there is concern that death may have been related to lung cancer, treatment of lung cancer, or the screening process (based on an explicit algorithm), the cause of death is determined by an independent committee that is blinded to the study arm. Vital status and cause of death will be obtained on participants who are lost to follow-up through the National Death Index. It is expected that the primary outcome, cumulative lung cancer mortality through August 2008, will be reported in 2010.

Secondary objectives of the NLST are to determine the effect of CT screening on all-cause mortality, lung cancer stage at diagnosis, medical utilization, quality of life, and smoking behavior.[81] Detailed collection of information on medical use and quality of life from a subset of trial sites will permit rigorous cost-effectiveness analyses of CT screening for lung cancer. The NLST will be working in collaboration with the NCI's Cancer Intervention and Surveillance Modeling Network to model the impact of CT screening on population trends in lung cancer.

NELSON trial. The NELSON trial,[82] which was launched in 2003, 1 year later than the NLST, is an ongoing multicenter RCT in the Netherlands and Belgium. The primary objective of the study is to determine whether screening with chest CT reduces lung cancer mortality in a high-risk cohort. However, unlike the NLST, the control group in the NELSON trial receives no screening. Another difference between the NELSON trial and NLST is that the 2 annual incidence screens occur 1 year and 3 years after the prevalence screen in the former trial as opposed to 1 year and 2 years after the prevalence screen in the latter trial. With regard to the screening technique, the NELSON trial uses 16-row multidetector CT exclusively and semiautomated 3-dimensional volumetric assessment of nodular growth.[83] For volumetric assessment of small, solid, parenchymal nodules, it has been demonstrated that this semiautomated measurement process is highly reproducible[84] Through October 2005, the study recruited 15,428 smokers and former smokers ages 50 to 75 years. When the NELSON trial is combined with a similar trial of 4000 participants in Denmark, the combined study will have 80% power to detect a lung cancer mortality reduction of 25% 10 years after randomization.

Current Recommendations

Although several medical organizations make recommendations about screening for cancer, the U.S. Preventive Services Task Force (USPSTF) is the most authoritative in the United States.[85] The USPSTF systematically reviews evidence pertaining to the benefits, harms, and costs of screening, placing greatest weight on the results of RCTs, and it makes recommendations based on explicit criteria.[86] Currently, the USPSTF states that the evidence is insufficient to recommend for or against lung cancer screening[87] but acknowledges that the results of the ongoing PLCO and NLST will provide critical information in the future[88] The American College of Chest Physicians also uses an evidence-based approach to develop clinical practice guidelines. This organization currently recommends that individuals should be screened for lung cancer only in the context of well-designed clinical trials[89] The Physician Data Query Screening Editorial Board[90] also uses an explicit, evidence-based approach to produce summaries of evidence for the benefits and harms of cancer screening but does not make recommendations. Because noncalcified nodules that could represent lung cancer are detected commonly on chest CT performed for nonscreening purposes, such as the evaluation of shortness of breath, the Fleischner Society of chest radiologists recently published guidelines for the management of these incidentally detected nodules[38] Those guidelines take into account smoking history and the strong relation between nodule size and the probability of malignancy.

Summary

CT screening for lung cancer is a "hot topic" and is under intense investigation. Although it has been demonstrated convincingly that low-dose, multidetector row CT is highly sensitive for the detection of small lung cancers, uncertainty remains about whether screening with this modality will decrease lung cancer mortality and do so sufficiently to offset the harms and costs of screening. The results of 2 large, ongoing RCTs combined with simulation modeling should resolve this uncertainty in a few years and help guide the implementation of CT screening if it is proven to be effective.

REFERENCES

1. American Cancer Society. *Cancer Facts and Figures 2007*. Atlanta, Ga: American Cancer Society; 2007.
2. Shopland DR, Eyre HJ, Pechacek TF. Smoking-attributable cancer mortality in 1991: is lung cancer now the leading cause of death among smokers in the United States? *J Natl Cancer Inst*. 1991;83:1142–1148.
3. Morabia A, Wynder EL. Cigarette smoking and lung cancer cell types. *Cancer*. 1991;68:2074–2078.
4. Morabia A, Wynder EL. Relation of bronchioloalveolar carcinoma to tobacco. *BMJ*. 1992;304:541–543.
5. Neuberger JS, Field RW. Occupation and lung cancer in nonsmokers. *Rev Environ Health*. 2003;18:251–267.
6. Miller DP, Liu G, De Vivo I, et al. Combinations of the variant genotypes of GSTP1, GSTM1, and p53 are associated with an increased lung cancer risk. *Cancer Res*. 2002; 62:2819–2823.
7. Ziegler RG, Mayne ST, Swanson CA. Nutrition and lung cancer. *Cancer Causes Control*. 1996;7:157–177.
8. The Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. *N Engl J Med*. 1994;330:1029–1035.
9. Omenn GS, Goodman GE, Thornquist MD, et al. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N Engl J Med*. 1996;334: 1150–1155.
10. Prorok PC, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial. *Control Clin Trials*. 2000;21:273S–309S.
11. Gohagan JK, Prorok PC, Hayes RB, Kramer BS. The Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial of the National Cancer Institute: history, organization, and status. *Control Clin Trials*. 2000;21:251S–272S.
12. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening [see comments]. *Lancet*. 1999;354:99–105.
13. Swensen SJ, Jett JR, Sloan JA, et al. Screening for lung cancer with low-dose spiral computed tomography. *Am J Respir Crit Care Med*. 2002;165:508–513.
14. Mascalchi M, Belli G, Zappa M, et al. Risk-benefit analysis of X-ray exposure associated with lung cancer screening in the Italung-CT trial. *AJR Am J Roentgenol*. 2006;187:421–429.
15. National Council on Radiation Protection and Measurements (NRCP). *Ionizing Radiation Exposures of the Population of the United States*. Report no. 93. Washington, DC: National Council on Radiation Protection and Measurements; 1987.
16. Maher MM, Kalra MK, Toth TL, Wittram C, Saini S, Shepard J. Application of rational practice and technical advances for optimizing radiation dose for chest CT. *J Thorac Imaging*. 2004;19:16–23.
17. Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, Miettinen OS. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med*. 2006; 355:1763–1771.
18. Black WC. Overdiagnosis: an underrecognized cause of confusion and harm in cancer screening. *J Natl Cancer Inst*. 2000;92:1280–1282.
19. Swensen SJ, Jett JR, Hartman TE, et al. CT screening for lung cancer: five-year prospective experience. *Radiology*. 2005;235:259–265.
20. Brenner DJ. Radiation risks potentially associated with low-dose CT screening of adult smokers for lung cancer. *Radiology*. 2004;231:440–445.
21. Gazelle GS, McMahon PM, Siebert U, Beinfeld MT. Cost-effectiveness analysis in the assessment of diagnostic imaging technologies. *Radiology*. 2005;235:361–370.
22. Hillman BJ, Black WC, D'Orsi C, Hauser B, Smith R. The appropriateness of employing imaging screening technologies: report of the Methods Committee of the ACR Task Force on Screening Technologies. *JACR*. 2004;1:861–864.

23. Black WC, Welch HG. Screening for disease. *AJR Am J Roentgenol*. 1997;168:3–11.
24. Morrison AS. The natural history of disease in relation to measures of disease frequency. In: *Screening in Chronic Disease*. 2nd ed. New York, NY: Oxford University Press; 1992:21–42.
25. Collins VP, Loeffler RK, Tivey H. Observations on growth rates of human tumors. *Am J Roentgenol Radium Ther Nucl Med*. 1956;76:988–1000.
26. Geddes DM. The natural history of lung cancer: a review based on rates of tumour growth. *Br J Dis Chest*. 1979;73:1–17.
27. Winer-Muram HT, Jennings SG, Tarver RD, et al. Volumetric growth rate of stage I lung cancer prior to treatment: serial CT scanning. *Radiology*. 2002;223:798–805.
28. Morrison AS. Early detection: sensitivity and lead time. In: *Screening in Chronic Disease*. 2nd ed. New York, NY: Oxford University Press; 1992:43–73.
29. Folkman J. The role of angiogenesis in tumor growth. *Semin Cancer Biol*. 1992;3:65–71.
30. Spratt JS, Meyer JS, Spratt JA. Rates of growth of human solid neoplasms: part I. *J Surg Oncol*. 1995;60:137–146.
31. Thunnissen FB, Kerr KM, Brambilla E, et al. EU-USA pathology panel for uniform diagnosis in randomised controlled trials for HRCT screening in lung cancer. *Eur Respir J*. 2006;28:1186–1189.
32. Lindell RM, Hartman TE, Swensen SJ, et al. Five-year lung cancer screening experience: CT appearance, growth rate, location, and histologic features of 61 lung cancers. *Radiology*. 2007;242:555–562.
33. Doll R, Peto R, Wheatley K, Gray R, Sutherland I. Mortality in relation to smoking: 40 years' observations on male British doctors [see comments]. *BMJ*. 1994;309:901–911.
34. Phillips AN, Wannamethee SG, Walker M, Thomson A, Smith GD. Life expectancy in men who have never smoked and those who have smoked continuously: 15 year follow up of large cohort of middle aged British men. *BMJ*. 1996; 313:907–908.
35. Day NE, Williams DRR, Khaw KT. Breast cancer screening programmes: the development of a monitoring and evaluation system. *Br J Cancer*. 1989;59:954–958.
36. Black WC. Anatomic extent of disease: a critical variable in reports of diagnostic accuracy. *Radiology*. 2000;217:319–320.
37. Cole P, Morrison AS. Basic issues in population screening for patients. *J Natl Cancer Inst*. 1980;64:1263–1272.
38. MacMahon H, Austin JH, Gamsu G, et al. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. *Radiology*. 2005;237:395–400.
39. Prorok PC, Kramer BS, Gohagan JK. Screening theory and study design: the basics. In: Kramer BS, Gohagan JK, Prorok PC, eds. *Cancer Screening: Theory and Practice*. New York, NY: Marcel Dekker; 1999:29–53.
40. Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. *BMJ*. 2002; 324:827–830.
41. Sarfati D, Howden-Chapman P, Woodward A, Salmond C. Does the frame affect the picture? A study into how attitudes to screening for cancer are affected by the way benefits are expressed. *J Med Screen*. 1998;5:137–140.
42. Black WC, Haggstrom DA, Welch HG. All-cause mortality in randomized trials of cancer screening. *J Natl Cancer Inst*. 2002;94:167–173.
43. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA*. 2000;283:2975–2978.
44. Schwartz LM, Woloshin S, Sox HC, Fischhoff B, Welch HG. US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: cross sectional survey. *BMJ*. 2000;320:1635–1640.
45. Knudsen AB, McMahon PM, Gazelle GS. Use of modeling to evaluate the cost-effectiveness of cancer screening programs. *J Clin Oncol*. 2007;25:203–208.
46. Hennekens CH, Buring JE. Descriptive studies. In: *Epidemiology in Medicine* Boston, Mass: Little, Brown; 1987:101–131.
47. Hennekens CH, Buring JE. Cohort studies. In: *Epidemiology in Medicine*. Boston, Mass: Little, Brown; 1987:168–173.
48. Hennekens CH, Buring JE. Case–control studies. In: *Epidemiology in Medicine*. Boston, Mass: Little, Brown; 1987: 132–152.
49. Hennekens CH, Buring JE. Intervention studies. In: *Epidemiology in Medicine*. Boston, Mass: Little, Brown; 1987: 178–212.
50. Black WC. Randomized clinical trials for cancer screening: rationale and design considerations for imaging tests. *J Clin Oncol*. 2006;24:3252–3260.
51. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet*. 2006;368:2053–2060.
52. National Lung Screening Trial. Available at URL: http://www.cancer.gov/nlst Accessed January 3, 2006.
53. Gohagan JK, Prorok P, Kramer B, Cornett J. Prostate cancer screening in the prostate, lung, colorectal, and ovarian cancer screening trial of the National Cancer Institute. *J Urol*. 1994;152:1905–1909.
54. Weed DL. Ethics and consent. In: Kramer BS, Gohagan JK, Prorok PC, eds. *Cancer Screening: Theory and Practice*. New York, NY: Marcel Dekker; 1999:89–118.
55. Hu P, Zelen M. Planning clinical trials to evaluate early detection programmes. *Biometrika*. 1997;84:817–829.
56. New York Early Lung Cancer Action Project Investigators. CT screening for lung cancer: diagnoses resulting from the New York Early Lung Cancer Action Project. *Radiology*. 2007;243:239–249.
57. Diederich S, Thomas M, Semik M, et al. Screening for early lung cancer with low-dose spiral computed tomography: results of annual follow-up examinations in asymptomatic smokers. *Eur Radiol*. 2004;14:691–702.
58. Diederich S, Wormanns D, Semik M, et al. Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers. *Radiology*. 2002;222:773–781.
59. Gohagan J, Marcus P, Fagerstrom R, Pinsky P, Kramer B, Prorok P. Baseline findings of a randomized feasibility trial of lung cancer screening with spiral CT scan vs chest radiograph: the Lung Screening Study of the National Cancer Institute. *Chest*. 2004;126:114–121.
60. Gohagan JK, Marcus PM, Fagerstrom RM, et al. Final results of the Lung Screening Study, a randomized feasibility study of spiral CT versus chest X-ray screening for lung cancer. *Lung Cancer*. 2005;47:9–15.
61. MacRedmond R, Logan PM, Lee M, Kenny D, Foley C, Costello RW. Screening for lung cancer using low dose CT scanning. *Thorax*. 2004;59:237–241.

62. MacRedmond R, McVey G, Lee M, et al. Screening for lung cancer using low dose CT scanning: results of 2 year follow up. *Thorax*. 2006;61:54–56.
63. Miller A, Markowitz S, Manowitz A, Miller JA. Lung cancer screening using low-dose high-resolution CT scanning in a high-risk workforce: 3500 nuclear fuel workers in 3 US states. *Chest*. 2004;125:152S–153S.
64. Nawa T, Nakagawa T, Kusano S, Kawasaki Y, Sugawara Y, Nakata H. Lung cancer screening using low-dose spiral CT: results of baseline and 1-year follow-up studies. *Chest*. 2002;122:15–20.
65. Pastorino U, Bellomi M, Landoni C, et al. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet*. 2003;362: 593–597.
66. Sobue T, Moriyama N, Kaneko M, et al. Screening for lung cancer with low-dose helical computed tomography: Anti-Lung Cancer Association project. *J Clin Oncol*. 2002;20: 911–920.
67. Sone S, Li F, Yang ZG, et al. Results of 3-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer*. 2001;84:25–32.
68. Sone S, Takashima S, Li F, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet*. 1998;351:1242–1245.
69. Swensen SJ, Jett JR, Hartman TE, et al. Lung cancer screening with CT: Mayo Clinic experience. *Radiology*. 2003;226: 756–761.
70. Tiitola M, Kivisaari L, Huuskonen MS, et al. Computed tomography screening for lung cancer in asbestos-exposed workers. *Lung Cancer*. 2002;35:17–22.
71. Li F, Sone S, Abe H, MacMahon H, Armato SG 3rd, Doi K. Lung cancers missed at low-dose helical CT screening in a general population: comparison of clinical, histopathologic, and imaging findings. *Radiology*. 2002;225:673–683.
72. Fontana RS, Sanderson DR, Woolner LB, et al. Screening for lung cancer: a critique of the Mayo Lung Project. *Cancer*. 1991;67(suppl):1155–1164.
73. Bach PB, Cramer LD, Schrag D, Downey RJ, Gelfand SE, Begg CB. The influence of hospital volume on survival after resection for lung cancer. *N Engl J Med*. 2001;345:181–188.
74. Bach PB, Jett JR, Pastorino U, Tockman MS, Swensen SJ, Begg CB. Computed tomography screening and lung cancer outcomes. *JAMA*. 2007;297:953–961.
75. Marshall D, Simpson KN, Earle CC, Chu C. Potential cost-effectiveness of 1-time screening for lung cancer (LC) in a high risk cohort. *Lung Cancer*. 2001;32:227–236.
76. Wisnivesky JP, Mushlin AI, Sicherman N, Henschke C. The cost-effectiveness of low-dose CT screening for lung cancer: preliminary results of baseline screening. *Chest*. 2003; 124:614–621.
77. Mahadevia PJ, Fleisher LA, Frick KD, Eng J, Goodman SN, Powe NR. Lung cancer screening with helical computed tomography in older adult smokers: a decision and cost-effectiveness analysis. *JAMA*. 2003;289:313–322.
78. Manser R, Dalton A, Carter R, Byrnes G, Elwood M, Campbell DA. Cost-effectiveness analysis of screening for lung cancer with low dose spiral CT (computed tomography) in the Australian setting. *Lung Cancer*. 2005;48:171–185.
79. Garg K, Keith RL, Byers T, et al. Randomized controlled trial with low-dose spiral CT for lung cancer screening: feasibility study and preliminary results. *Radiology*. 2002;225: 506–510.
80. Pinsky PF, Marcus PM, Kramer BS, et al. Diagnostic procedures after a positive spiral computed tomography lung carcinoma screen. *Cancer*. 2005;103:157–163.
81. ACRIN 6654.Contemporary screening for the detection of lung cancer. Available at URL: http://www.acrin.org/pdf_file2.html?file=protocol_docs/A6654partial_summary.pdf Accessed January 4, 2006.
82. van Iersel CA,de Koning HJ, Draisma G, et al. Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian Randomised Lung Cancer Multi-slice CT Screening Trial (NELSON). *Int J Cancer*. 2006;120:868–874.
83. Xu DM, Gietema H, de Koning H, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. *Lung Cancer*. 2006;54:177–184.
84. Gietema HA, Wang Y, Xu D, et al. Pulmonary nodules detected at lung cancer screening: interobserver variability of semiautomated volume measurements. *Radiology*. 2006; 241:251–257.
85. US Preventive Services Task Force. About USPSTF: the new US Preventive Services Task Force. Available at URL: http://www.ahrq.gov/clinic/uspstfab.htm Accessed April 12, 2007.
86. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med*. 2001;20:21–35.
87. US Preventive Services Task Force. Screening for Lung Cancer. Available at URL: http://www.ahrq.gov/clinic/uspstf/uspslung.htm Accessed April 12, 2007.
88. Humphrey LL, Teutsch S, Johnson M. Lung cancer screening with sputum cytologic examination, chest radiography, and computed tomography: an update for the US Preventive Services Task Force. *Ann Intern Med*. 2004;140:740–753.
89. Bach PB, Silvestri GA, Hanger M, Jett JR. Screening for lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132:69S–77S.
90. Physician Data Query, National Cancer Institute Common Scientific Outline partners. Available at: http://www.cancer.gov/cancertopics/pdq/screening/overview/healthprofessional Accessed January 5, 2006.