

The 23rd Lung Cancer Mass Screening Seminar

CT Screening for Lung Cancer: Update 2008

Claudia I. Henschke¹; Rowena Yip¹; David F. Yankelevitz¹

ABSTRACT — Screening for a cancer should be considered when the cancer is significant in terms of incidence and mortality, treatment of early stage disease is better than treatment of late stage disease, and there is a screening regimen that provides for earlier diagnosis rather than later, symptom-prompted diagnosis. Lung cancer qualifies as it kills more people than any other cancer worldwide. In the United States it kills more people than colon, breast, and prostate cancer combined and more women than breast cancer. The fundamental concepts of screening are presented. Screening for a cancer is a repetitive process, starting with the *baseline* round followed by *repeat* rounds of screening at set intervals. The *regimen of screening* defines the initial diagnostic test and the sequence of tests to be performed leading to a rule-in diagnosis of the cancer. The regimen should provide *lead time* of the diagnosis of the cancer. The regimen for the first, baseline round may be different from the regimen for the repeat rounds as the former is inherently different from the subsequent repeat rounds. Baseline screening identifies a greater proportion of cancers with a longer latent (asymptomatic) phase than repeat screening, called *length bias*. Length bias exists for any screening program, regardless of the design of the study or the cancer. Repeat rounds of screening identify the same proportion of cancer diagnoses found in absence of screening for people having the same risk of the cancer and these repeat rounds of screening can be pooled. It is also a consequence of length bias that cancers found in repeat rounds are ear-

lier in their latent phase than those of the baseline round, a less frequently mentioned consequence. *Overdiagnosis bias*, another bias of screening, can occur in two ways: 1) a 'cancer' detected by the screening, pathologically proven, that is not life-threatening even when not resected and 2) a genuine life-threatening cancer that is diagnosed and treated but the person dies of another disease or accident. This bias can be addressed in various ways, including by the regimen of screening and by following those who refuse treatment. As screening is pursuit of early diagnosis followed by early treatment, both the diagnostic and treatment performance can be addressed separately. Key diagnostic performance measures of the regimen are: 1) the proportion of *screen-diagnoses* among all diagnoses and 2) the *stage distribution* of the diagnosed cancers. These performance measures are unaffected by the frequency of lung-cancer diagnoses and thus also unaffected by the enrollment criteria. The key prognostic performance measure is the curability rate which is provided by the long-term follow-up of all diagnosed cases of lung cancer, regardless of stage and treatment. It can also be estimated by the proportion in Stage I multiplied by the curability rate in Stage I. Finally this report provides a summary of the diagnostic and prognostic performance measures available from the screening trials to date.

(*JJLC*. 2009;49:477-490)

KEY WORDS — CT Screening for Lung Cancer, Early Lung Cancer Action Project (ELCAP), New York-ELCAP (NY-ELCAP), International-ELCAP (I-ELCAP)

INITIAL CT SCREENING STUDIES

In Japan, CT was added to an already long existing practice of screening for lung cancer using CXR,^{1,2} pre-

sumably for similar reasons that led to the initiation of the research study comparing CT with CXR called the Early Lung Cancer Action Project (ELCAP) in New York City.³⁻⁵ Both started in 1993, although neither one

¹Department of Radiology, New York Presbyterian Hospital-Weill Medical College, New York.
Reprints: Claudia I. Henschke, Department of Radiology, New

York Presbyterian Hospital-Weill Medical College, 525 East 68th Street, New York, NY 10065 (e-mail: chensch@med.cornell.edu).

knew of the other study. In a third study in Nagano Prefecture,⁶⁻⁸ starting in 1996, screening was provided by a mobile van which performed population wide screening using CT and chest radiography.

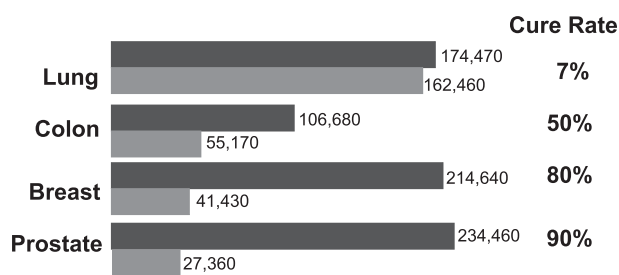
These three studies show that the lung cancer rate depends on risk characteristics (e.g., age and smoking history) and it ranged from 0.2%-0.6% per 1,000 having annual repeat screening. Sputum cytology only identified a few additional lung cancers not identified by CT. All of the studies showed a high proportion of Stage I diagnoses ranging from 79%-100%. Interim diagnoses of lung cancer were few. All three studies provided both CXR and CT to all participants, and CT was markedly superior to CXR in all three, CXR missing about 80% of the Stage I lung cancers found on CT.

KEY CONCEPTS AND DEFINITIONS

Screening has been defined as the pursuit of early diagnosis of the cancer before symptoms occur.³⁻⁵ The purpose of early diagnosis is to provide for early treatment which potentially prevents death from the cancer. The usefulness of screening depends on how early the cancer can be diagnosed and how many deaths can be prevented by the early treatment as compared with later symptom-prompted diagnosis and treatment. To fully understand the current discussions on the evidence for lung cancer screening, key concepts of screening and definitions are needed.

Screening for a cancer should be considered when the cancer is significant in terms of incidence and mortality, treatment of early stage disease is better than treatment of late stage disease, and there is a test that provides for earlier diagnosis — *lead time* — than later, symptom-prompted diagnosis.^{9,10} Lung cancer qualifies as it kills more people than any other cancer worldwide. In the United States it kills more people than colon, breast, and prostate cancer combined and more women than breast cancer (Figure 1).¹¹

For lung cancer, the staging system is based on differences in lung cancer survival.^{12,13} The curability rate as estimated by the 10-year survival rate for Stage I is high, particularly when the cancer diameter is 10 mm or less¹⁴ (Figure 2) and this rate decreases as the tumor size increases.¹⁵⁻¹⁷ Although the curability rate is high



2006 Cancer Facts and Figures, American Cancer Society

Figure 1. Annual incidence, mortality and cure rate for the 4 leading cancers (blue (■) represents new cases and red (■) deaths).

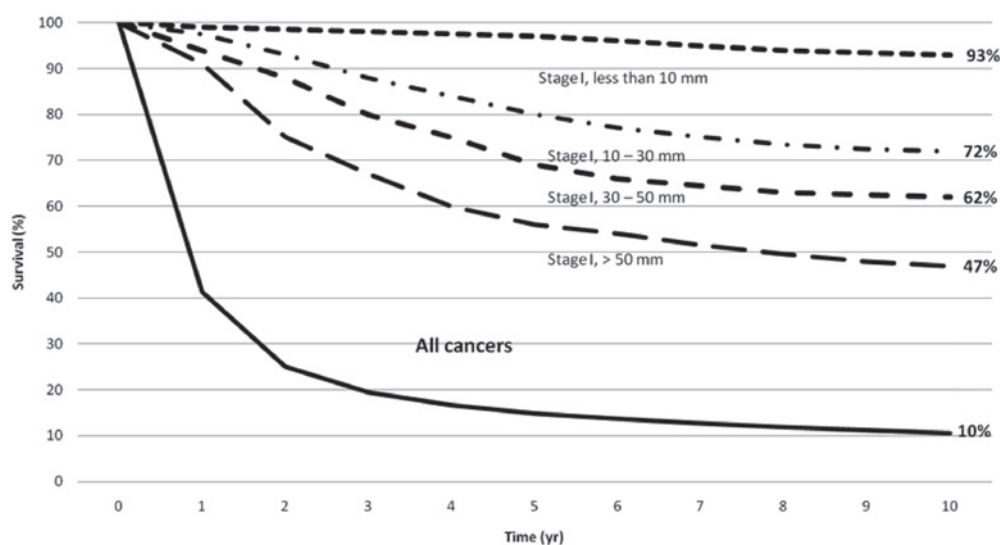


Figure 2. Curability of lung cancer within Stage I by tumor diameter and for all stages combined as estimated by 10-year survival rates.

for Stage I lung cancer, less than 15% are diagnosed in that stage and so the overall curability rate for lung cancer for all stages combined is below 10%.¹⁶

Screening for a cancer is a repetitive process, starting with the *baseline* round followed by *repeat* rounds of screening (Figure 3) at intervals defined by a *regimen of screening*. The regimen also defines the initial diagnostic test (e.g., PAP smear, mammogram, fecal occult blood, chest radiograph, CT) and the sequence of tests to be performed leading to a rule-in diagnosis of the cancer. The regimen for the first, baseline round may be different from the regimen for the repeat rounds as no prior results are available for the former, and for other reasons discussed below.

A diagnosed case of the cancer is classified as a *screen-diagnosis* if the diagnosis resulted from the work-up of an abnormality identified on the initial test of the regimen (Figure 3). A diagnosis of cancer resulting from findings identified in the baseline regimen is classified as

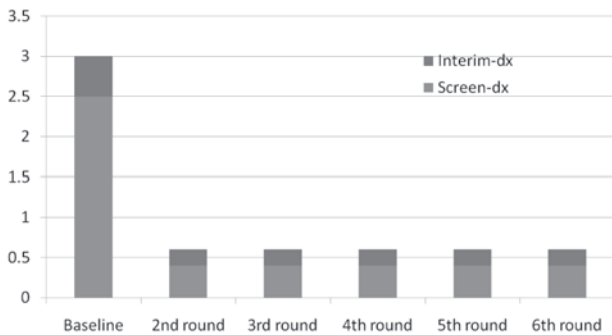


Figure 3. Depiction of first, baseline round and the subsequent repeat rounds of screening. Each round has both *screen-diagnosed* and *interim-diagnosed* cases of cancer.

a baseline cancer. Similarly when the diagnosis is made as a result of findings identified in the repeat regimen, it is classified as a repeat cancer. If the diagnosis of cancer resulted because of symptom-prompted work-up before the next screening, it is classified as an *interim-diagnosis*. These are standard definitions and it is critical to ensure that the same definitions are being used when comparing results of different studies.

The baseline round is inherently different from the repeat rounds of screening because it is the first round and no prior results are available for comparison. One consequence of baseline screening is that cancers with a longer latent (asymptomatic) phase are more frequently identified. This has been called *length bias*⁹ and exists for any screening program, regardless of the design of the study or the cancer.^{9,18} While this difference exists between baseline and repeated screening, it does not exist for repeat rounds and thus repeat rounds can be pooled. The other consequence is that cancers found in repeat rounds are found earlier in their latent phase than in the baseline round,¹⁸ a fact not usually stated. To address the consequences of this 'length bias', the baseline round should be reported separately from the repeat rounds.

In repeat screenings, the frequency of all cancer diagnoses should reflect that found in usual care, or in the absence of screening, for people having the same risk of lung cancer. The proportion by cell-type should also reflect the proportion found in usual care. For lung cancer, the proportion by cell-type in repeat rounds of screening¹⁹ is similar to that found in clinical practice,²⁰ and differs from the proportion found in the baseline round (Figure 4). The size distribution, however, of cancers

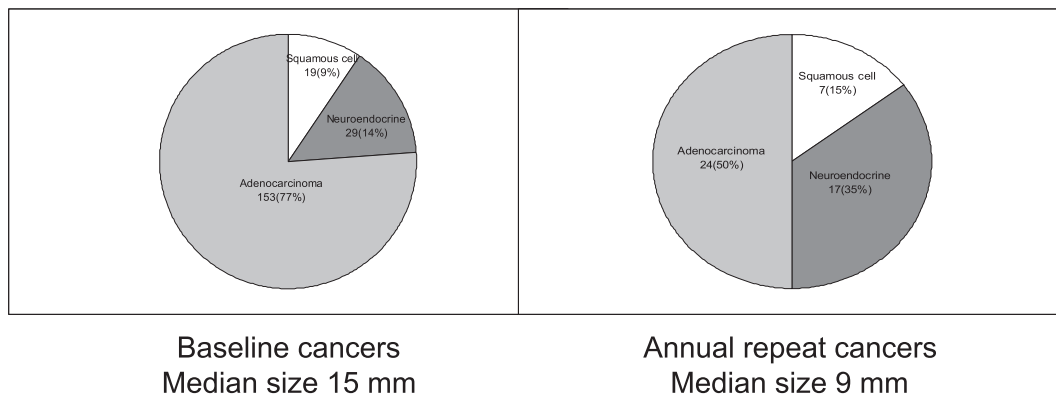


Figure 4. Slower-growing cancers are found more frequently in the baseline round, but the cancers in the repeat rounds are found when they are smaller. Cell-type of cancers found in repeat rounds reflect the proportion found in the absence of screening.

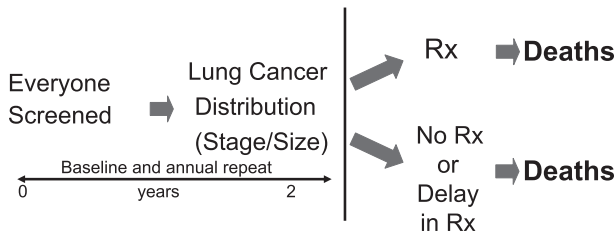


Figure 5. Sequential approach to evaluation of the effectiveness of screening.

found by screening is smaller than in clinical practice.

As screening is pursuit of early diagnosis followed by early treatment, both the diagnostic and treatment performance need to be determined. Key diagnostic performance measures of the regimen are: 1) the proportion of *screen-diagnoses* among all diagnoses, 2) the *stage distribution* of the diagnosed cancers and 3) the estimated *lead time* given by the ratio of the number of diagnoses in the baseline round to those in a single repeat round. If the ratio is 1, meaning that the screening regimen does not provide lead time, it will not provide for earlier diagnosis. Different from diagnosis, treatment potentially changes the natural course of the disease and thus to determine its effectiveness, a comparison is needed.²¹⁻²³

THE ELCAP APPROACH

The ELCAP investigators wanted to develop an efficient methodology to provide an ever-accumulating, continually updated body of evidence for evaluation of emerging new technologies for screening for cancer.^{3,24,25} To provide optimal screening, a regimen for the diagnostic work-up must be specified. It starts with the definition of the initial test, its positive result, the work-up for a positive result leading to a diagnosis of cancer. Once cancer is diagnosed, treatment is typically performed according to usual care standards and is documented. Screening, by definition, involves asymptomatic participants, however, the minimum age, smoking exposure, or other such admissibility criteria are flexible and set by each participating institution as these criteria determine the frequency of cancer diagnoses but do not affect the performance measures.

Initially a head-to-head comparison of the new screening regimen with the previous regimen (e.g., the initial test being CT instead of chest radiography (CXR)) is performed. To maximize the efficiency of the study, high-risk participants can be enrolled, but all should be

free of recognized symptoms and signs of cancer. Such a study would require a baseline and at least one repeat round of screening. If this limited study shows that the diagnostic performance—stage distribution, proportion of screen-to-interim diagnoses, lead time—are promising, then the regimen is updated and this updated regimen is then provided to participants at different institutions. Following further confirmation of the diagnostic performance, screening can then be provided to an expanded group of participants at a lower risk of the cancer. In the course of these successive studies, new technologies (e.g., PET, PET/CT, computer-aided diagnostics) can be integrated into the regimen. The data from these studies can be pooled to determine the curability of those diagnosed early and treated, provided a common protocol and the proper quality assurance procedures needed for such a collaboration are in place.^{25,26}

For curability determination, a comparison group is needed. The comparison group may be formed by randomly assigning people with screen-diagnosed lung cancer to immediate or delayed treatment (Figure 5) as was done for prostate cancer.²² The randomization could be further stratified by clinical stage and CT appearance of the cancer or perhaps even based on the results of percutaneous needle biopsy. The latter approach provides a direct assessment of the extent of over-diagnosis of lung cancer resulting from screening. Alternatively a quasi-experimental control group can be used consisting of participants diagnosed with lung cancer who have refused or delayed their treatment even though they are candidates for it. This is a valid approach as long as the choice of the treatment or lack of it is independent of the cancer prognosis or other factors that might influence the ultimate outcome and these factors are documented at the time of enrollment into the screening program and not at the time of diagnosis or treatment.²² A third alternative is to compare the mortality rate in the screening program once sufficient deaths have occurred to that in a non-screened comparison group which has a similar risk profile for lung cancer. Finally, a fourth approach is to analyze the temporal pattern of the deaths in the screened cohort after initiation of screening and compare the deaths in the early years to the deaths in the later years when the benefit of screening should become apparent if the screening is effective.^{18,25,26}

The ELCAP approach provides for further efficiencies. Follow-up of participants is only required for those

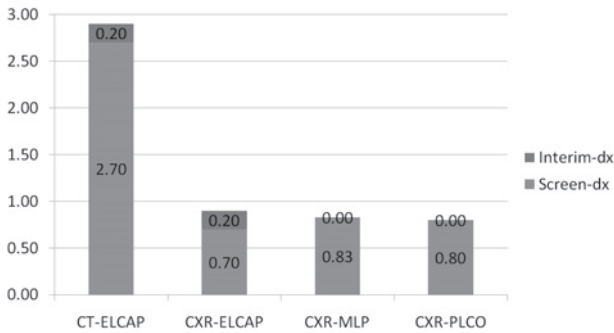


Figure 6. Percentage of baseline screen- and interim-diagnoses in ELCAP using CT and CXR as compared to the baseline diagnoses in MLP²⁹ and PLCO²⁸ (Baseline interim-diagnoses were not reported for MLP and PLCO).

diagnosed with lung cancer, usually some 1%-6%, depending on the risk of the participants. This markedly reduces the follow-up requirements. To fully document all interim-diagnoses occurring between the annual rounds of screening, the protocol requires that each person who has not returned for the repeat screening to be followed up to 18 months after their prior screening.^{18,26} If no cancer has been identified either by symptom-prompted work-up or for any other reason, then the person is considered to have stopped participation in the screening program and no further follow-up is required.

ELCAP TO NY-ELCAP TO I-ELCAP

Prior to starting ELCAP, we estimated that enrollment of 1,000 very-high risk participants would yield some 200-300 people with nodules and some 15-30 cancers to address the diagnostic performance of CT.³ We also asked Dr. Flehinger to use the model she and her co-workers had developed based on prior randomized screening trials²⁷ to estimate the potential benefit of CT screening which suggested that CT screening might decrease the deaths from lung cancer by as much as 80%.

ELCAP enrolled 1,000 participants at two institutions in New York at high-risk of lung cancer because of their age and smoking history (60 years of age and older with a history of at least 10 pack-years of cigarette smoking).⁴ In the baseline round, each participant received the CT and chest radiograph (CXR) which were read independently. If the result was positive, work-up proceeded according to a stated regimen. In the baseline round, 27 (2.7%) lung cancer cases were screen-diagnosed and 2 (0.2%) interim-diagnosed among the 1,000 participants. CXR identified 7 (0.7%) cancers, 3 of Stage I and also

missed the same 2 (0.2%) interim-diagnosed cases. Thus, CXR missed 20 (74%) of the 27 screen-diagnosed cancers found by CT (Figure 6). More importantly, CXR missed 20 (80%) of the 25 clinical Stage I cancers, so that CXR screening was stopped.

The baseline ELCAP-CXR results are similar to those found in the Mayo Lung Project (MLP)²⁷ and to the recently reported baseline results of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Trial²⁸ (Figure 6) which implies that the three cohorts had similar risk characteristics for lung cancer. Interim diagnoses must be presented in the MLP and PLCO baseline round, that is prior to the first repeat round, but they were not reported.

Annual repeat rounds of screening in ELCAP resulted in 7 screen-diagnoses of lung cancer and no interim diagnoses (7/1,184 = 0.59%) (Figure 7).⁵ Of the 7, 6 (86%) were of Stage I (Table 1). The frequency of repeat diagnoses should be the same as that found in the absence of screening among people with the same risk characteristics. Given that the baseline frequency of lung cancer diagnoses for ELCAP using CXR was quite similar to that found in the MLP,²⁹ it is to be expected that the frequency of diagnoses in the repeat rounds of ELCAP-CT of 0.59% would be similar to that in the MLP of 0.55%.³⁰

Given the high proportion of Stage I diagnoses in the baseline round of ELCAP, the prognostic prediction of a curability rate of 60%-80% because of CT screening was made²⁴ as had originally been predicted by the model based on prior randomized trials.⁴ ELCAP also showed that the regimen of screening minimized additional procedures to rates similar to those found in mammography screening for breast cancer.³¹ As a result, screening was rapidly expanded to other institutions in New York (NY-ELCAP)³² which confirmed the ELCAP results. In view of the demand for screening, other institutions joined I-ELCAP³³ which enrolled younger individuals and also people who had never smoked but were exposed to carcinogens by their occupation and/or second-hand smoke. Figure 7 shows that the frequency of baseline and annual repeat diagnoses of lung cancer decreased as the risk characteristics (i.e., age and smoking history) of the screenees decreased (i.e., lower age, lower smoking history or more ex-smokers) in the subsequent studies.

The diagnostic performance measures are unaffected

Table 1. Summary of Diagnostic and Performance of CT Screening Trials

Summary

1. Lung cancer prevalence rate depends on risk characteristics. It ranged from 0.1%-0.8% per 1,000 screened depending on age and smoking history. *Stage I diagnoses include non-small and small-cell cancers without lymph node metastases and multiple adenocarcinoma without lymph node metastases (based on pathology if resected).
2. 7 studies showed consistency in finding a high proportion of Stage I diagnoses ranging from 71%-100% and the frequency depends on the regimen of screening and adherence to it.
3. Two studies reported overall long-term survival rates 71% or higher. Survival was 92% or better if in Stage I and resected. Survival rate reflects proportion in Stage I.
4. Few interim diagnoses of lung cancer (symptom-prompted between screening) and few detected only by sputum cytology.
5. For comparison, a consistent definition of baseline, repeat cancers, and interim cancers is needed: a. baseline cancer: nodule is identified on initial CT at baseline, b. annual cancer: nodule is first identified on initial CT at annual repeat, and c. interim cancer: symptom-prompted diagnoses between screenings.

Project Name, Years Criteria for Enrollment	Enrolled Baseline Annual	Patients with Lung Cancer Baseline + Interim + Sputum only Annual + Interim + Sputum only	LC Prevalence		Ratio Baseline Annual	% Stage I *		Survival Rate Lung Cancer
			Baseline Annual	Annual		Baseline Annual	Annual	
ELCAP, New York, NY 1993-94 ⁵ Age > 60 yrs; Pk-yrs > 10	1,000 Median age: 67 Median pack-years: 45 %current smokers: 43% %male: 54%	27 + 2 + not done	2.9%	0.6%	4.9	86%	86%	
Nagano, Japan 1996-8 ^{6,8} Age > 40 yrs; Pk-yrs not req'd	5,480 Median age: 64 Median pack-years: NR %current smokers: 46% %male: 55%	37 + 0 + 1	0.7%	0.2%	3.2	100%	86%	10-yr overall: 86%; path Stage I: 90%
ALCA-NCC, Japan 1993-2001 ² Age > 40 yrs; Pk-yrs not req'd	7,891 Median age: NR Median pack-years: NR %current smokers: 86% %male: 88%	13 + 0 + 1	0.9%	0.3%	3.1	79%	82%	5-yr overall: 71%; baseline: 76%; annual 65%
Hitachi, Japan 2001-2 ⁷ Age > 40 yrs; Pk-yrs not req'd	7,956 Median age: NR Median pack-years: NR %current smokers: 62% %male: 79%	36 + 0 + 0	0.5%	0.1%	6.3	86%	100%	
Mayo Clinic 1999-2004 ⁷⁷ Age > 50 yrs; Pk-yrs > 20; quit < 10 yrs ago	5,568 Median age: 59 Median pack-years: 51 %current smokers: 61% %male: 52%	4 + 0 + 0	0.1%	0.8%	2.6	77% ^a	71% ^a	path Stage I: 94%
Istituto Tumori, Italy 2000-1 ⁷⁸ Age > 50 yrs; Pk-yrs > 20	4,472 Median age: 58 Median pack-years: 40 %current smokers: 86% %male: 71%	30 + 0 + 1	2.0%	0.5%	3.5	72%	100%	path Stage I: 94%
International-ELCAP 1993-2006 ³² Age > 40 yrs; Pk-yrs not req'd	996 Median age: 61 Median pack-years: 30 %current smokers: 37% %male: 58%	12 ^b + 6 ^c + 0 + not done	1.7%	1.3%	4.8	85% ^d	86% ^d	10-yr overall: 80%; timely resected Stage I: 92%

^aIncluded 6 limited small cell in Stage I for consistency with other studies. ^b1 typical carcinoma included. ^c6 had been identified at baseline low-dose CT, either < 5 mm or > 5 mm and considered benign and were identified as cancers on first annual. ^dclinical staging prior to surgery. ^eRevised in 2007 (see reference 7).

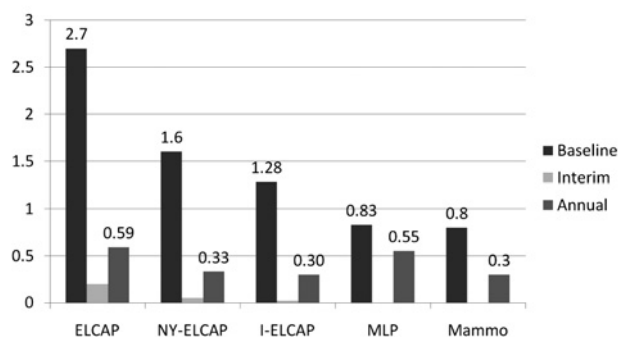


Figure 7. Frequency of baseline and annual repeat diagnoses of lung cancer in ELCAP, NY-ELCAP, I-ELCAP, Mayo Lung Project (MLP)^{29,30} and mammography screening.³¹

by the frequency of lung-cancer diagnoses. The proportion in Stage I in all three studies remained around 85%, for both baseline and annual repeat rounds of screening in NY-ELCAP³² and I-ELCAP.³³ Interim-diagnoses were rare and found prior to the first annual repeat screening, but not between the repeat rounds of screening. The estimated lead time for CT screening, based on the ratio of baseline to repeat screening also remained around 4.5 in all 3 studies, much higher than the 1.5 years estimated using the same approach for CXR screening using the MLP^{29,30} or the 2.5 years estimated for mammography screening for breast cancer.³¹

The diagnostic performance is dependent on the imaging used in the regimen and the regimen itself. The introduction of multidetector-row CT scanners and the reading of images on computer monitors instead of film after ELCAP should increase the proportion of Stage I diagnoses and decrease the interim diagnoses. It did slightly increase the proportion in Stage I, but it substantially increase the proportion of screen-diagnoses in the baseline round or otherwise said it decreased the proportion of baseline interim-diagnoses (symptom-prompted) from 7% (2/29) in ELCAP to 3% (3/104) in NY-ELCAP to 1% (5/405) in I-ELCAP.

These 3 studies provided important information for updating the regimen. The implications of finding non-solid and part-solid nodules was recognized,³⁴ and their significance,³⁵ work-up of nodules improved,^{36,37} usefulness of growth as an indication for biopsy was tested³⁸⁻⁴² and the work-up of other findings such as mediastinal masses,⁴³ cardiac calcifications,⁴⁴ and emphysema⁴⁵ were formulated.

Long-term follow-up of ELCAP, NY-ELCAP, and I-

ELCAP provided the estimated curability rate of 80% for all diagnoses (screen- and interim-combined), regardless of stage and treatment.³³ When diagnosed in early stage and having prompt treatment, the estimated curability rate was 92% while all those diagnosed with Stage I lung cancer who refused treatment died of it. The number of deaths prevented by early treatment of lung cancers diagnosed by CT screening was estimated by the proportion in Stage I multiplied by the curability rate in Stage I ($85\% \times 92\% = 78\%$) or alternatively by the overall curability rate of 80%. This compares to some 7% of deaths that are currently prevented in the absence of screening given by ratio of the number of deaths to new cases of lung cancer each year (164,000/174,000, Figure 1).

The initial rapid decline in the survival rate for all lung cancer patients³³ is due to deaths from lung cancer which mostly occurred in the first 4 years after diagnosis. They primarily occurred in those asymptomatic people who already had late-stage lung cancer when it was diagnosed and thus the screening did not prevent their deaths. The people whose death from lung cancer is prevented are those found with early stage cancer by screening who would have otherwise been diagnosed 5 to 6 years later, given the lead time provided by CT and then died within 2-4 years of diagnosis, overall some 7-10 years later. As these deaths are prevented by early treatment following screen-diagnosis provided by the CT screening, both the survival rate and the cumulative mortality rate will ultimately reach a plateau.

CT screening in ELCAP and NY-ELCAP had fewer late-stage lung cancers as compared with CXR screening in the MLP and Memorial-Sloan Kettering Lung Project (MSKLP).⁴⁶ In this comparison, the focus must be on the repeat rounds of screening. The overall frequency of ELCAP lung cancer diagnoses is close to that of MLP and NY-ELCAP is close to that of MSKLP, but the absolute and proportional number of late-stage cancers is significantly less for ELCAP and NY-ELCAP (Figure 8). Figure 8 shows that the reduction from 0.29% in MLP as compared with 0.08% in ELCAP and 0.22% in MSKLP as compared with 0.05% in NY-ELCAP. This significant stage shift to early stage cancers is highly suggestive of a decrease in the mortality rate of CT screening when compared to CXR screening. Such a shift to earlier stage cancers with a subsequent decrease in deaths had been demonstrated for cervical

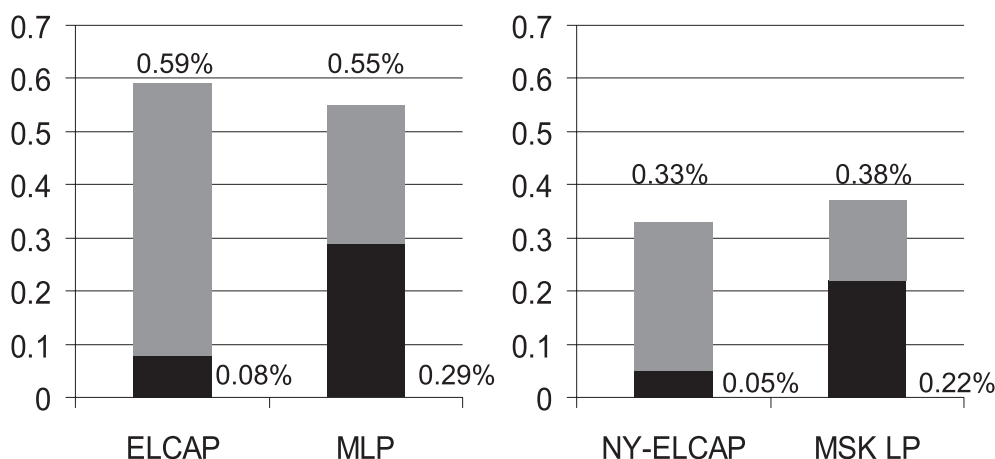


Figure 8. Comparison of late-stage lung cancers (black) in repeat rounds of screening in ELCAP vs. MLP and in NY-ELCAP vs. MSKLP.

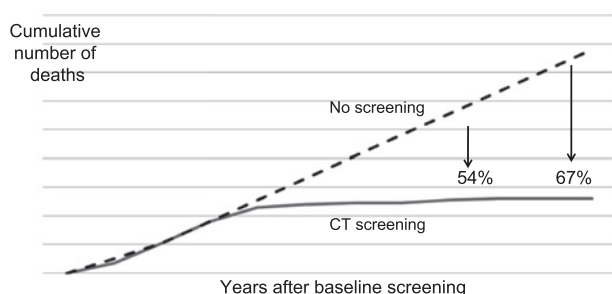


Figure 9. Schematic of cumulative deaths after baseline enrollment (solid line) as compared to the deaths in the absence of screening (dashed line).

cancer screening.^{47,48}

Consider the schematic graph shown in Figure 9. The cumulative number of deaths is shown on the Y axis and the years of screening relative to baseline screening is shown on the X axis. As previously shown,³³ the deaths in years 2 through 4 occur in those asymptomatic individuals whose cancer is found in late stage. This cumulative number of deaths in the first 4 years reflects the deaths that would be found in an unscreened cohort with same risk of lung cancer. Projecting this same rate of deaths over time (dashed line), the number of deaths in the absence of screening can be compared to that actually observed in a screening program (solid line) and the difference between the dashed and solid lines shows the reduction in deaths from lung cancer. Only after the fourth year of screening does the rate start to decrease. With increasing follow-up (as long as screening continues) there is a further reduction in deaths.

Figure 9 shows why screening needs to continue for

long enough to demonstrate the actual reduction in deaths that is provided by screening. If follow-up only extends to year 4, as in the National Lung Screening Trial (NLST)^{49,50} whose median follow-up will be 4 years by 2009, less than a 20% reduction can be anticipated. To see such a small reduction would require protocol compliance (of the two arms) and no delay in diagnosis and treatment.⁵¹ As the NLST requires a 20% mortality reduction to be able to reject the null hypothesis of no benefit from the screening,⁴³ this is highly unlikely to be reached until there is longer follow-up.⁵²

Figure 9 also illustrates one of the key problems in recent paper by Bach et al.⁵² which focused mainly on the early years (the first 3-4 years after baseline). However, even in that analysis, the decrease in the cumulative mortality rate in years 5 and 6 can already be seen. The focus on the appropriate time when the reduction in mortality can reasonably be expected has already been highlighted in breast cancer screening^{53,54} and in colorectal screening,^{55,56} both of which illustrated the need for longer screening and follow-up.

ADVANTAGES OF THE ELCAP APPROACH

The ELCAP approach was designed to utilize screenings performed either as part of a research project or as part of practice oriented research, both for efficiency and cost considerations as well as for rapid translation into clinical guidelines. Its efficiency is illustrated by the Early Lung Cancer Action Project (ELCAP) which provided information on the diagnostic performance of CT screening²⁴ and showed the prognostic potential so that

it could be expanded to New York State³² and other sites throughout the world.³³ This accumulating body of evidence has been accomplished by a very modest initial funding for 1,000 baseline and repeat screenings in ELCAP and the added enrollment of 30,000 participants, much less than the initial funding of the NLST comparing CT with chest radiography.^{49,50}

The ELCAP approach does not have the problems that have been recognized as occurring in randomized trials.^{23,57-66} Biases of randomized trial results include having: 1) an insensitive outcome measure—the cumulative mortality rate which does not focus on the relevant time when the decrease in deaths prevented by screening are seen^{53,57-61}; 2) an inadequate number of rounds of screening so that the decrease in the mortality rate cannot be seen^{27,52}; 3) protocol non-adherence^{52,58,65,67}; 4) delays in diagnosis and/or treatment or participants choosing not to be treated,⁶⁴ 5) without any analysis of whether this proportion was the same in both arms of the trial, and 6) reliance on death certificates.⁶⁶ By the time the randomized screening trials are completed, there has also been considerable technology drift so that the results are no longer relevant. This is demonstrated by the PLCO trial⁶⁸ which started in 1993 and will report the results of CXR towards the end of the next decade, when the CXR is no longer relevant. Having an inadequate number of rounds of screening was clearly illustrated by the Minnesota Colorectal Study which required extension to 10 years of screening from the originally planned 5 years.⁵⁶ Items 3-6 are particularly troublesome if the frequency of occurrence is different in the two arms of the randomized trial.^{51,67} Although randomization is used to provide comparability of the two arms at enrollment, it does not ensure comparability in those diagnosed with lung cancer (only 1%-3% of all participants), nor as to whether all had timely diagnosis or treatment. To help overcome these problems, a large number of participants are required, markedly increasing the cost and time required for such trials. An alternative that was suggested by the designers of some of these large randomized screening studies⁶⁷ was to perform a limited mortality analyses of these trials focusing only on the relevant cases in each arm of the randomized trials instead of all of the cases.

Probably, because of the considerable cost and time and inherent difficulties of these trials, only one randomized screening study for colorectal cancer (Minnesota)

study⁵⁶ and one for breast cancer,⁵³ and one for lung cancer study (which evolved into three separate smaller studies⁶⁹ have been completed in the United States. Consequently, emerging, promising modalities are not evaluated scientifically. Often the tests are simply integrated into the medical care system. For colon cancer screening, new tests (e.g., colonoscopy and/or virtual colonoscopy with CT has essentially replaced sigmoidoscopy and the latter is used in the ongoing PLCO, but the efficacy of these new tests has not been tested using a randomized screening trial design. For lung cancer, CT was already available when the CXR and sputum trials were started in the mid 1980's, but the CXR is still being tested in the PLCO^{28,68} and these results will not be reported in next decade. For coronary artery calcification screening with CT, no randomized trial has ever been performed, although large national cohort studies have been started. Thus, frequently scientific evidence for formulation of national policies are not available on these emerging technologies.

CONCERNS ABOUT THE ELCAP APPROACH

Concerns about biases in ELCAP have been raised and addressed previously.¹⁸ The key bias of the concern in the ELCAP design is 'overdiagnosis'^{18,69} as the other two biases—length and lead time—affect both ELCAP and the randomized screening trials.

Length bias

Length bias affects all screening programs, ELCAP as well as randomized screening trials. It is introduced by the very process of screening. The baseline round is inherently different from the repeat rounds of screening because cancers with a longer latent (asymptomatic) phase are more frequently identified in the baseline round,⁹ but the cancers found in repeat rounds are found earlier in their latent phase than in the baseline round.¹⁸ While this difference exists between baseline and repeated screening, it does not exist for repeat rounds and thus repeat rounds can be pooled. The solution is to report the results of the baseline round separately from the results of the repeat rounds as illustrated in Figure 4.

Lead-time bias

Screening for cancer is done to provide for earlier diagnosis and earlier treatment when it is more effective. In the ELCAP design, a bias is introduced when there is insufficiently long follow-up are presented. The ELCAP

curability rate is subject to lead-time bias if the Kaplan-Meier survival curve has not yet reached its asymptote (that point when the curve reaches a plateau and no longer decreases). If there is lead-time bias the estimated curability rate is higher than it will be when it reaches its asymptote. Kaplan-Meier survival analysis is the standard approach used in oncology trials. It adjusts for incomplete follow-up of those diagnosed with the cancer (i.e., that not all patients have been followed for the same length of time). I-ELCAP waited to report 10-year survival rates so that the plateau was clearly reached which was around 5-6 years after diagnosis.³³ Its curability estimate has no lead-time bias.

Randomized screening trials also have lead-time bias when they do not provide for sufficient follow-up in their design, but in this case the lead-time bias exists because the mortality reduction is underestimated. The period of screening and follow-up needs to be long enough (dependent on lead-time) so as to focus on the time period when the decrease in the deaths due to screening can be anticipated and when there is no longer a lead-time bias.^{51,53,55}

Lead-time bias is also introduced when comparing a treatment which has lead time relative to treatment without lead time. I-ELCAP did not do this. I-ELCAP compared effectiveness of treatment contrasted with no treatment in screened patients, all of whom have the same lead time.

Overdiagnosis

The other concern is that the curability rate might be inflated as “over-diagnosed” lung cancers might be included.^{9,18,69} As randomized screening trials focus only on deaths, this bias is not of concern, but it is a concern for the ELCAP approach.¹⁸

Overdiagnosis occurs in two ways: 1) a ‘cancer’ is detected by the screening which would have never been life-threatening even when not resected, but because of screening it was detected and the person is thus subjected to diagnostic tests and treatment when in fact the person was not at risk of dying of the cancer; 2) a genuine life-threatening cancer is diagnosed but the person dies of another disease or accident so that the screening has not saved the life. In other words the person dies of a competing cause^{70,71}; this concern needs to be addressed in the eligibility criteria so that the life-expectancy of participants is sufficient to justify the screening. Of these two concerns, the main concern is

that those with slow-growing cancer may inflate the estimated curability rate and that these patients would have also undergone surgery for a non-life threatening lesion.

For lung cancer, multiple reports have shown that identifying slow-growing cancers is not a significant concern and does not account for the survival differences reported in the 3 randomized screening trials for lung cancer. Almost all of those who were diagnosed with Stage I lung cancer as a result of CXR screening and refused treatment died of their disease as demonstrated by Flehinger et al.,⁶⁴ Yankelevitz et al.,⁷¹ and by Sobue et al.⁷² in Japan. In the absence of screening, this was also found in the analyses by Henschke et al.⁷³ using the data in the Surveillance and End Results (SEER) registry and by Raz et al. based on California registry data.⁷⁴

The I-ELCAP protocol²⁶ directly addresses the issue in several ways: 1) In both baseline and repeat rounds of screening, the regimen requires demonstration of *in-vivo* growth at a malignant rate prior to recommending biopsy. 2) All resected specimen are reviewed by an international panel of pathology experts who confirmed they are all genuine lung cancer, and that 95% of them are already invasive.¹⁹ 3) Those who delayed their diagnosis or treatment showed progression of their disease in NY-ELCAP³² and I-ELCAP.³³

In ELCAP, NY-ELCAP, and I-ELCAP, the lung-cancer diagnosis rate in annual repeat rounds of screening was essentially the same as in prior screening trials performed in the 1970's (Figure 6, 7). Thus there is no evidence of overdiagnosis in the repeat rounds of screening as repeat cancers in ELCAP, by definition, were not seen on the prior screening one year earlier. Thus, the volume doubling time of the cancer must be 200 days or less, if the lower limit of nodule detectability is one with a diameter of 2 mm.⁷⁰ A cancer with a volume doubling time of 200 days is an aggressive cancer, and does not fit the profile of an “overdiagnosed” one.

For baseline screening, we determined that 87% of clinical Stage I patients had genuine, life-threatening cancers defined as the cancer having doubling time faster than 400 days.³⁵ When considered separately by nodule consistency, it was 96% for cancers manifesting as solid nodules (e.g., typical carcinoids), 90% as part-solid nodules and 67% as nonsolid nodules.³⁵ Only adenocarcinoma is diagnosed in nonsolid nodules, adenocarcinoma diagnosed in nonsolid nodules are the most sus-

pect cases for being 'overdiagnosed' cancer.¹⁹ Thus, we identified that some of the adenocarcinoma manifesting as nonsolid nodules were slow-growing in addition to those already well-known cell-types, such as typical carcinoids which manifest as solid nodules, which are known to be slow-growing.

If these steps are not sufficient to address the concern about overdiagnosis, then ethically a randomized treatment trial (RCT) could be performed in which patients with potentially "overdiagnosed" screen-diagnosed lung cancer are randomly assigned to either immediate treatment or delayed treatment as was done for prostate cancer.²²

The second issue of competing causes of death is addressed by setting reasonable admissibility requirements for screening based on actual data obtained from I-ELCAP. It is suggested that as the lead time provided by CT is about 4.5 years, the life-expectancy of a person undergoing CT screening should be at least 10 years. ELCAP further addressed competing causes of death directly by performing a survival analysis focusing on non-lung cancer deaths which showed that older smokers and former smokers have a high life expectancy if they do not die of lung cancer. Their 10-year survival rate for death other than lung cancer was 93%.⁷⁵

OTHER CT SCREENING TRIALS

Subsequent to the three initial CT screening trials already discussed,¹⁻⁸ others have reported similar results⁷⁶⁻⁷⁸ (Table 1). To enhance the comparison, consistent definitions of baseline, repeat cancers, and interim cancers defined at the beginning of this paper were used whenever possible.

All of the studies showed a high proportion of Stage I diagnoses ranging from 71% -100%. The frequency, of course, depends on the regimen of screening and adherence to it in addition to the definition of Stage I. This definition of Stage I should be clearly defined in each study and should consider the results in the context of screening. Stage I diagnoses should include both non-small and small-cell cancers without lymph node metastases and but include multiple adenocarcinomas without lymph node metastases, as the latter should have the benefit of resection.^{70,79}

Stimulated by the demand for CT screening resulting from the ELCAP publication, the NLST was developed.^{49,50} It was designed by the PLCO investigators to-

gether with the American College of Radiology Imaging Network (ACRIN) and used the same design as prior randomized trials for lung cancer in the US.^{29,46,56,80} The control arm of the NLST is identical to the control arms of the MSKLP and JHHLP where all received annual CXR screening.⁸⁰ The intervention arm of the NLST was provided annual CT. Three rounds of screening were provided, baseline and 2 annual repeat rounds of screening. Enrollment in the NLST started in 2002 and ended by mid 2004, so that when follow-up ends in 2008 the median follow-up time after diagnosis is only some 4 years in 2009 when the results are to be reported. The regimen of screening was not well defined and not enforced or checked at the participating institutions. The traditional outcome measure of cumulative mortality rate is to be used and there is no mention of performing a limited mortality analysis which would focus on the timeliness of diagnosis or treatment or on the deaths in the relevant time interval during which the screening benefit is expected, an important limitation of the traditional analysis of such trials that had been recognized by the designers of the PLCO and NLST as early as 1983.⁵⁷ Reanalysis of the MLP, focusing only on the lack of protocol non-adherence showed that the CXR might have provided as much as a 43% reduction in deaths from lung cancer.⁶⁵ A pilot study of 3,000 participants was performed prior to the start of the NLST⁸¹ to demonstrate that randomization was feasible, but it also suggested lack of adherence, particularly for those randomized to the CXR arm. A regimen for the work-up of screen-detected nodules was specifically not included in that study and it resulted in a low proportion of Stage I diagnoses in the CT arm. Further the proportion of cancer diagnoses among those having invasive procedures was low in comparison to ELCAP, NY-ELCAP and I-ELCAP, suggesting that there was poor understanding of the importance of a work-up protocol and also poor compliance with any of the known recommended work-up algorithms.²⁶

Different cost-effectiveness analyses have shown that CT screening for lung cancer is very cost-effective⁸¹⁻⁸⁶ with the exception of one theoretical study.⁸⁷ These analyses look at the overall cost per life-year saved, but ideally a cost-effectiveness assessment would be done on an individualized basis.⁸⁸ Such an individualized assessment would determine the life expectancy of the person (in light of the personal risk indicators) and the

risk of competing causes of death in order to determine how many years of life might be saved by the screening round that is being contemplated.

REFERENCES

- Kaneko M, Eguchi K, Ohmatsu H, Kakinuma R, Naruke T, Suemasu K, et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. *Radiology*. 1996;201:798-802.
- Sobue T, Moriyama N, Kaneko M, Kusumoto M, Kobayashi T, Tsuchiya R, et al. Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. *J Clin Oncol*. 2002;20:911-920.
- Henschke CI, Miettinen OS, Yankelevitz DF, Libby DM, Smith JP. Radiographic screening for cancer. Proposed paradigm for requisite research. *Clin Imaging*. 1994;18:16-20.
- Henschke CI, McCauley DI, Yankelevitz DF, Naidich DP, McGuinness G, Miettinen OS, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet*. 1999;354:99-105.
- Henschke CI, Naidich DP, Yankelevitz DF, McGuinness G, McCauley DI, Smith JP, et al. Early lung cancer action project: initial findings on repeat screenings. *Cancer*. 2001;92:153-159.
- Sone S, Takashima S, Li F, Yang Z, Honda T, Maruyama Y, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet*. 1998;351:1242-1245.
- Sone S, Li F, Yang ZG, Honda T, Maruyama Y, Takashima S, et al. Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer*. 2001;84:25-32.
- Sone S, Nakayama T, Honda T, Tsushima K, Li F, Haniuda M, et al. Long-term follow-up study of a population-based 1996-1998 mass screening programme for lung cancer using mobile low-dose spiral computed tomography. *Lung Cancer*. 2007;58:329-341.
- Hutchison GB, Shapiro S. Lead time gained by diagnostic screening for breast cancer. *J Natl Cancer Inst*. 1968;41:665-681.
- Morrison AS. The effects of early treatment, lead time and length bias on the mortality experienced by cases detected by screening. *Int J Epidemiol*. 1982;11:261-267.
- American Cancer Society. Statistics for 2006. Cancer Facts and Figures. http://www.cancer.org/docroot/stt/stt_0.asp. Accessed 1/16/07
- Mountain CF. Revisions in the International System for Staging Lung Cancer. *Chest*. 1997;111:1710-1717.
- Inoue K, Sato M, Fujimura S, Sakurada A, Takahashi S, Usuda K, et al. Prognostic assessment of 1310 patients with non-small-cell lung cancer who underwent complete resection from 1980 to 1993. *J Thorac Cardiovasc Surg*. 1998;116:407-411.
- Martini N, Bains MS, Burt ME, Zakowski MF, McCormack P, Rusch VW, et al. Incidence of local recurrence and second primary tumors in resected stage I lung cancer. *J Thorac Cardiovasc Surg*. 1995;109:120-129.
- Buell PE. The importance of tumor size in prognosis for resected bronchogenic carcinoma. *J Surg Oncol*. 1971;3:539-551.
- SEER Relative Survival Rates by Stage at Diagnosis. SEER 9 Registries for 1988-2003. Fast Stats: Lung and Bronchus Cancer. <http://seer.cancer.gov/faststats/sites.php?site=Lung%20and%20Bronchus%20Cancer&stat=Survival>. Assessed 12/19/07
- Wisnivesky JP, Yankelevitz D, Henschke CI. The effect of tumor size on curability of stage I non-small cell lung cancers. *Chest*. 2004;126:761-765.
- Henschke CI, Smith JP, Miettinen OS. Response to letters to the editor. *N Engl J Med*. 2007;356:743-747.
- Carter D, Vazquez M, Flieder DB, Brambilla E, Gazdar A, Noguchi M, et al. Comparison of pathologic findings of baseline and annual repeat cancers diagnosed on CT screening. *Lung Cancer*. 2007;56:193-199.
- Ginsberg MS, Grewal RK, Heelan RT. Lung cancer. *Radiol Clin North Am*. 2007;45:21-43.
- Hill AB. Suspended judgment. Memories of the British Streptomycin Trial in Tuberculosis. The first randomized clinical trial. *Control Clin Trials*. 1990;11:77-79.
- Bill-Axelson A, Holmberg L, Ruutu M, Häggman M, Andersson SO, Bratell S, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med*. 2005;352:1977-1984.
- Campbell DT, Stanley JC. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Company; 1963.
- Henschke CI, Yankelevitz DF, Wisnivesky JP, Smith JP, Libby D, Pasmantier M. CT Screening for Lung Cancer: The diagnostic-prognostic approach. In: Pass HI, Carbone DP, Johnson DH, Minna JD, Turrisi AT, eds. *Lung Cancer: Principles and Practice*. 3rd ed. Lippincott, Williams and Wilkins; 2005.
- Henschke CI, Yankelevitz DF, Smith JP, Miettinen OS, ELCAP Group. Screening for lung cancer: the early lung cancer action approach. *Lung Cancer*. 2002;35:143-148.
- International Early Lung Cancer Action Program: Enrollment and Screening Protocol. <http://www.ielcap.org/professionals/docs/ielcap.pdf>.
- Flehinger BJ, Kimmel M, Polyak T, Melamed MR. Screening for lung cancer. The Mayo Lung Project revisited. *Cancer*. 1993;72:1573-1580.
- Oken MM, Marcus PM, Hu P, Beck TM, Hocking W, Kvale PA, et al. Baseline chest radiograph for lung cancer detection in the randomized Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. *J Natl Cancer Inst*. 2005;97:1832-1839.
- Fontana RS, Sanderson DR, Taylor WF, Woolner LB, Miller WE, Muhm JR, et al. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Mayo Clinic study. *Am Rev Respir Dis*. 1984;130:561-565.
- Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, Muhm JR. Lung cancer screening: the Mayo program. *J Occup Med*. 1986;28:746-750.

31. Quality Determinants of Mammography Guideline Panel. Quality Determinants of Mammography. Clinical Practice Guideline, No 13. US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research (Publication No 95-0632). 1994:82-86.
32. 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.
33. International Early Lung Cancer Action Program Investigators, Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, et al. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med*. 2006;355:1763-1771.
34. Henschke CI, Yankelevitz DF, Mirtcheva R, McGuinness G, McCauley D, Miettinen OS, et al. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *AJR Am J Roentgenol*. 2002;178:1053-1057.
35. Henschke CI, Shaham D, Yankelevitz DF, Kramer A, Kostis WJ, Reeves AP, et al. CT screening for lung cancer: significance of diagnoses in its baseline cycle. *Clin Imaging*. 2006;30:11-15.
36. Henschke CI, Yankelevitz DF, Naidich DP, McCauley DI, McGuinness G, Libby DM, et al. CT screening for lung cancer: suspiciousness of nodules according to size on baseline scans. *Radiology*. 2004;231:164-168.
37. Libby DM, Wu N, Lee IJ, Farooqi A, Smith JP, Pasmantier MW, et al. CT screening for lung cancer: the value of short-term CT follow-up. *Chest*. 2006;129:1039-1042.
38. Yankelevitz DF, Gupta R, Zhao B, Henschke CI. Small pulmonary nodules: evaluation with repeat CT—preliminary experience. *Radiology*. 1999;212:561-566.
39. Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B, Henschke CI. Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. *Radiology*. 2000;217:251-256.
40. Kostis WJ, Reeves AP, Yankelevitz DF, Henschke CI. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. *IEEE Trans Med Imaging*. 2003;22:1259-1274.
41. Reeves AD, Chan AB, Yankelevitz DF, Henschke CI, Kressler B, Kostis WJ. On measuring the change in size of pulmonary nodules. *IEEE Trans Med Imaging*. 2006;25:435-450.
42. Kostis WJ, Yankelevitz DF, Reeves AP, Fluture SC, Henschke CI. Small pulmonary nodules: reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up CT. *Radiology*. 2004;231:446-452.
43. Henschke CI, Lee IJ, Wu N, Farooqi A, Khan A, Yankelevitz D, et al. CT screening for lung cancer: prevalence and incidence of mediastinal masses. *Radiology*. 2006;239:586-590.
44. Shemesh J, Henschke CI, Farooqi A, Yip R, Yankelevitz DF, Shaham D, et al. Frequency of coronary artery calcification on low-dose computed tomography screening for lung cancer. *Clin Imaging*. 2006;30:181-185.
45. Kostis WJ, Fluture S, Yankelevitz DF, Henschke CI. Method for analysis and display of distribution of emphysema in CT scans. *SPIE Medical Imaging*. 2003;5032:199-206.
46. Melamed MR, Flehinger BJ, Zaman MB, Heelan RT, Perchick WA, Martini N. Screening for early lung cancer. Results of the Memorial Sloan-Kettering study in New York. *Chest*. 1984;86:44-53.
47. Kjellgren O. Mass screening in Sweden for cancer of the uterine cervix: effect on incidence and mortality. An overview. *Gynecol Obstet Invest*. 1986;22:57-63.
48. Johannesson G, Geirsson G, Day N. The effect of mass screening in Iceland, 1965-74, on the incidence and mortality of cervical carcinoma. *Int J Cancer*. 1978;21:418-425.
49. National Lung Screening Trial. <http://www.nci.nih.gov/NLST>. Accessed 1/25/2003
50. Aberle DR, Black WC, Goldin JG. (2003) Contemporary screening for the detection of lung cancer protocol [NLST], 10 May 2002. American College of Radiology Imaging Network (ACRIN sharp6654). http://www.acrin.org/6654_protocol.html. Accessed 12/19/2007
51. Henschke CI, Yankelevitz DF, Kostis WJ. CT screening for lung cancer. *Semin Ultrasound CT MR*. 2003;24:23-32.
52. Bach PB, Jett JR, Pastorino U, Tockman MS, Swensen SJ, Begg CB. Computed tomography screening and lung cancer outcomes. *JAMA*. 2007;297:953-961.
53. Miettinen OS, Henschke CI, Pasmantier MW, Smith JP, Libby DM, Yankelevitz DF. Mammographic screening: no reliable supporting evidence? *Lancet*. 2002;359:404-405.
54. Chu KC, Smart CR, Tarone RE. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial. *J Natl Cancer Inst*. 1988;80:1125-1132.
55. Hanley JA. Analysis of mortality data from cancer screening studies: looking in the right window. *Epidemiology*. 2005;16:786-790.
56. Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *N Engl J Med*. 1993;328:1365-1371.
57. Baker SG, Kramer BS, Prorok PC. Statistical issues in randomized trials of cancer screening. *BMC Med Res Methodol*. 2002;2:11.
58. Connor RJ, Prorok PC. Issues in the mortality analysis of randomized controlled trials for cancer screening. *Control Clin Trials*. 1994;15:81-99.
59. Aron JL, Prorok PC. An analysis of the mortality effect in a breast cancer screening study. *Int J Epidemiol*. 1986;15:36-43.
60. Prorok PC, Kramer BS, Gohagan JK. Screening theory and study design: the basics. Chapter 3. In: Kramer BS, Gohagan JK, Prorok PC, eds. *Cancer Screening*. New York: Marcel Dekker, Inc; 1999.
61. Tarone RE, Gart JJ. Significance tests for cancer screening trials. *Biometrics*. 1989;45:883-890.
62. Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, Muhm JR, et al. Screening for lung cancer. A critique of the Mayo Lung Project. *Cancer*. 1991;67

- (Suppl):1155-1164.
63. Fontana RS. The Mayo Lung Project. *International Conference on Prevention and Early Diagnosis of Lung Cancer, Varese, Italy, Dec 9-10*. 1998:15-20.
 64. Flehinger BJ, Kimmel M, Melamed MR. The effect of surgical treatment on survival from early lung cancer. Implications for screening. *Chest*. 1992;101:1013-1018.
 65. Miettinen OS. Screening for lung cancer. *Radiol Clin North Am*. 2000;38:479-486.
 66. Black WC, Haggstrom DA, Welch HG. All-cause mortality in randomized trials of cancer screening. *J Natl Cancer Inst*. 2002;94:167-173.
 67. Prorok PC, Chamberlain J, Day NE, Hakama M, Miller AB. UICC Workshop on the evaluation of screening programmes for cancer. *Int J Cancer*. 1984;34:1-4.
 68. Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000;21(Suppl):273S-309S.
 69. Black WC. Overdiagnosis: An underrecognized cause of confusion and harm in cancer screening. *J Natl Cancer Inst*. 2000;92:1280-1282.
 70. Consensus statements. International Conferences on Screening for Lung Cancer. <http://www.ielcap.org>.
 71. Yankelevitz DF, Kostis WJ, Henschke CI, Heelan RT, Libby DM, Pasmantier MW, et al. Overdiagnosis in chest radiographic screening for lung carcinoma: frequency. *Cancer*. 2003;97:1271-1275.
 72. Sobue T, Suzuki T, Matsuda M, Kuroishi T, Ikeda S, Naruke T. Survival for clinical stage I lung cancer not surgically treated. Comparison between screen-detected and symptom-detected cases. The Japanese Lung Cancer Screening Research Group. *Cancer*. 1992;69:685-692.
 73. Henschke CI, Wisnivesky JP, Yankelevitz DF, Miettinen OS. Small stage I cancers of the lung: genuineness and curability. *Lung Cancer*. 2003;39:327-330.
 74. Raz DJ, Zell JA, Ou SH, Gandara DR, Anton-Culver H, Jablons DM. Natural history of stage I non-small cell lung cancer: implications for early detection. *Chest*. 2007;132:193-199.
 75. Henschke CI, Yip R, Yankelevitz DF, Miettinen OS. Computed tomography screening for lung cancer: prospects of surviving competing causes of death. *Clin Lung Cancer*. 2006;7:323-325.
 76. 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.
 77. Swensen SJ, Jett JR, Hartman TE, Midthun DE, Mandrekar SJ, Hillman SL, et al. CT screening for lung cancer: five-year prospective experience. *Radiology*. 2005;235:259-265.
 78. Pastorino U, Bellomi M, Landoni C, De Fiori E, Arnaldi P, Picchio M, et al. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet*. 2003;362:593-597.
 79. Vazquez M, Carter D, Brambilla E, Gazdar A, Noguchi M, Travis WD, et al. Solitary and multiple resected adenocarcinomas after CT screening for lung cancer: histopathologic features and their prognostic implications. *Lung Cancer*. 2009;64:148-154.
 80. Berlin NI, Buncher CR, Fontana RS, Frost JK, Melamed MR. The National Cancer Institute Cooperative Early Lung Cancer Detection Program. Results of the initial screen (prevalence). Early lung cancer detection: Introduction. *Am Rev Respir Dis*. 1984;130:545-549.
 81. Miettinen OS. Screening for lung cancer: can it be cost-effective? *CMAJ*. 2000;162:1431-1436.
 82. Gohagan J, Marcus P, Fagerstrom R, Pinsky P, Kramer B, Prorok P, et al. 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.
 83. 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.
 84. Marshall D, Simpson KN, Earle CC, Chu C. Potential cost-effectiveness of one-time screening for lung cancer (LC) in a high risk cohort. *Lung Cancer*. 2001;32:227-236.
 85. Marshall D, Simpson KN, Earle CC, Chu CW. Economic decision analysis model of screening for lung cancer. *Eur J Cancer*. 2001;37:1759-1767.
 86. Chirikos TN, Hazelton T, Tockman M, Clark R. Screening for lung cancer with CT: a preliminary cost-effectiveness analysis. *Chest*. 2002;121:1507-1514.
 87. 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.
 88. I-ELCAP Investigators. CT screening for lung cancer: individualising the benefit of the screening. *Eur Respir J*. 2007;30.