

Overstating the Evidence for Lung Cancer Screening

The International Early Lung Cancer Action Program (I-ELCAP) Study

H. Gilbert Welch, MD, MPH; Steven Woloshin, MD, MS; Lisa M. Schwartz, MD, MS; Leon Gordis, MD, MPH, DrPH; Peter C. Gøtzsche, MD, MPH; Russell Harris, MD, MPH; Barnett S. Kramer, MD, MPH; David F. Ransohoff, MD

Last year, the *New England Journal of Medicine* ran a lead article reporting that patients with lung cancer had a 10-year survival approaching 90% if detected by screening spiral computed tomography. The publication garnered considerable media attention, and some felt that its findings provided a persuasive case for the immediate initiation of lung cancer screening. We strongly disagree. In this article, we highlight 4 reasons why the publication does not make a persuasive case for screening: the study had no control group, it lacked an unbiased outcome measure, it did not consider what is already known about this topic from previous studies, and it did not address the harms of screening. We conclude with 2 fundamental principles that physicians should remember when thinking about screening: (1) survival is always prolonged by early detection, even when deaths are not delayed nor any lives saved, and (2) randomized trials are the only way to reliably determine whether screening does more good than harm.

Arch Intern Med. 2007;167(21):2289-2295

There is no cancer for which screening ought to be more seriously considered than lung cancer. Each year, 160 000 Americans die from the disease—a number that surpasses that of the next 4 most deadly solid tumors combined (colon, breast, prostate, and melanoma).¹ Most people diagnosed as having lung cancer will ultimately succumb to it because most present late in their disease course—when the effectiveness of treatment is limited. And most of this disease burden is attributable to a single exposure: cigarette smoking.² Substantial disease burden, ineffective treatment for clinically diagnosed disease, and a readily identifiable high-risk group combine to make the conditions extremely favorable for screening.

Against this backdrop, the October 26, 2006, issue of the *New England Journal of Medicine's* lead article on lung cancer screening³ reporting a 10-year survival rate approaching 90% was bound to garner considerable attention. And it did. Among the 10 highest-circulation newspapers in the country, 7 carried the story. So did both of the major wire services and 4 of the 5 major television networks. It was tempting to conclude that the International Early Lung Cancer Action Program (I-ELCAP) investigation on spiral computed tomographic (CT) scanning had made a strong case for initiating lung cancer screening now.

This conclusion would lend support to groups already promoting widespread lung cancer screening. For example, plaintiffs in separate class action suits against Phillip Morris USA (Richmond, Virginia) (the manufacturer of Marlboro cigarettes) are seeking annual low-dose spiral CT scanning as the remedy for people with 20 or more pack-years of smoking who reside in the states of New York and Massachusetts. And because of their exposure to secondhand smoke, there is also an initiative to start screening flight attendants.⁴

Author Affiliations: VA Outcomes Group, White River Junction VA Medical Center, White River Junction, Vermont, and the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth Medical School, Hanover, New Hampshire (Drs Welch, Woloshin, and Schwartz); Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, and Department of Pediatrics, Johns Hopkins School of Medicine, Baltimore, Maryland (Dr Gordis); Nordic Cochrane Center, Rigshospitalet, Copenhagen, Denmark (Dr Gøtzsche); University of North Carolina School of Medicine, Chapel Hill (Drs Harris and Ransohoff); and Office of Disease Prevention, National Institutes of Health, Bethesda, Maryland (Dr Kramer).

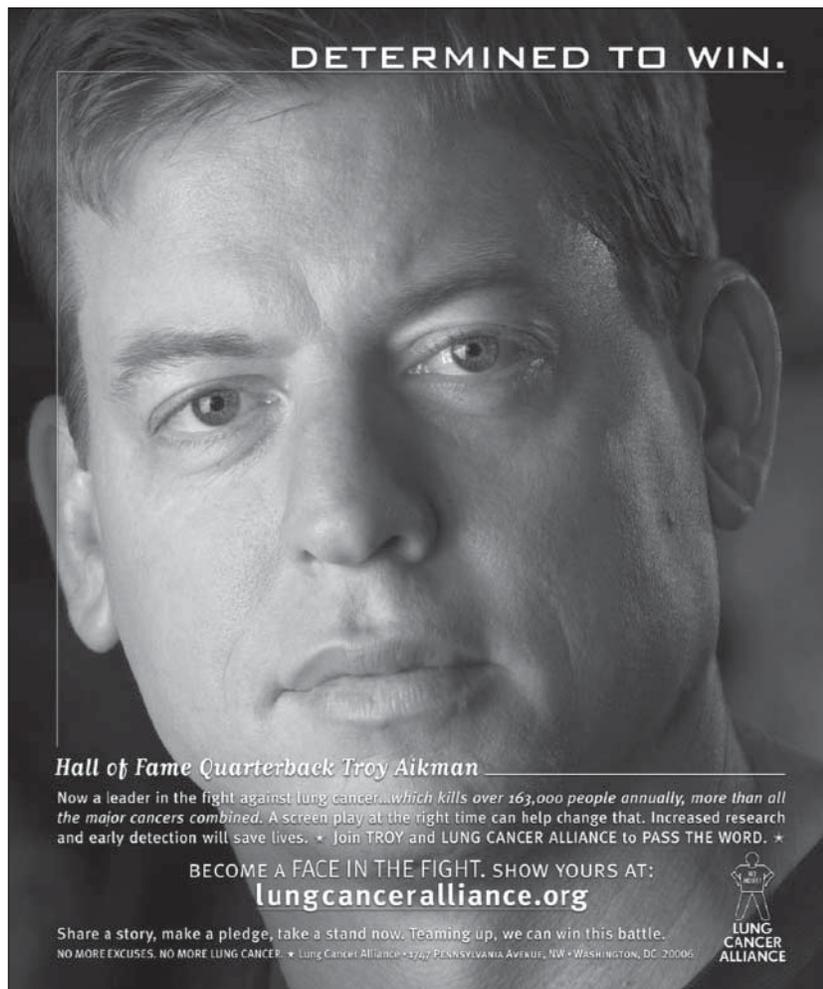


Figure 1. Lung Cancer Alliance sports celebrity advertisement.

The I-ELCAP data were even used to affirm the value of early detection in a recent Senate resolution passed by unanimous consent.⁵

Finally, the Lung Cancer Alliance, a major patient advocacy group, is lobbying Congress and President George W. Bush to “make early detection, treatment and chemoprevention of lung cancer a national priority.”⁶ The organization encourages CT screening for lung cancer as a way to save lives and uses the I-ELCAP findings as compelling evidence.⁷ Despite the recent publication of a controlled observational study⁸ suggesting that screening could do more harm than good (which the Lung Cancer Alliance president characterized as “another delaying tactic to deny people at high risk for lung cancer the chance to have it detected at an early, treatable stage”^{9(p11)}), the Lung Cancer Alliance is now launching an adver-

tising campaign featuring sports celebrities encouraging the public to make the right call and get screened (**Figure 1**).¹⁰

In this article, we discuss the reasons why the I-ELCAP data cannot be used to recommend lung cancer screening. We focus on 4 essential elements missing from the publication³: a control group, an unbiased outcome measure, information that is already known about this topic from previous studies, and the harms of screening.

WHAT IS MISSING

A Control Group

A basic step in assessing the validity of an inference is to determine the design of the study on which it is based. This is so fundamental that many journals now require that the study design be explicitly identi-

fied in the abstract. And one of the first things medical students are taught to look for is a design that provides a control group (ie, randomized trials and cohort and case-controls studies). This point warrants such emphasis for a simple reason: only a control group can provide insight about what would happen in the absence of intervention.

The I-ELCAP had no control group. Although it provided lengthy follow-up for a large number of people who were screened for lung cancer, there were no people who were *not* screened for lung cancer. The I-ELCAP study design is a case series. Because there was no control group, readers cannot know what would have been observed in the absence of screening.

In making a scientific inference, large numbers of participants and lengthy follow-up cannot compensate for the absence of a control group. Of course, although never explicitly stated, there was an implied control group—the experience of current practice. In other words, what was observed in the I-ELCAP case series could be compared with what is observed in current practice. Which brings us to the question, “In terms of what outcome?” The primary outcome of I-ELCAP was survival, and that is a fundamental problem.

An Unbiased Outcome Measure

In I-ELCAP, the 10-year survival rate of patients with stage I lung cancer approached 90%. For most physicians, this is a stunning percentage. Their experience is almost exactly the reverse: the vast majority of lung cancer patients die within a few years of diagnosis. In fact, the most representative data source of current cancer outcomes in the United States (the Surveillance, Epidemiology, and End Results [SEER] Program data from the National Cancer Institute) reports that the 10-year survival rate of patients with lung cancer diagnosed in current practice is roughly 10%.¹¹

There is, of course, a big difference between survival rates of 90% and 10%. The difference is so big, in fact, that it is tempting to conclude that any concerns raised about the

study design represent a sort of quibbling—trivial concerns that cannot threaten the overall conclusion.

But survival is a remarkably misleading metric. Most of us interpret the words “higher survival” as “extended life” or “delayed death.” In other words, we assume that a higher survival rate means a lower death rate. This is understandable given the conventional definitions of the words. However, the actual measurement of survival is powerfully affected by the mechanism of diagnosis, whereas the measurement of a death rate is not. If the mechanism of diagnosis is changing (precisely what is being studied in I-ELCAP), survival can skyrocket, even if no one had their death delayed.

Two phenomena explain this paradox. The first is *lead-time bias*, illustrated in **Figure 2**. Imagine a group of patients with lung cancer currently diagnosed at age 67 years, all of whom die at age 70 years. Each survived only 3 years, so their 10-year survival rate is 0%. Now imagine that the same group is all diagnosed earlier by spiral CT—at age 59 years—but they all still die at age 70 years. All have now survived 11 years, and thus their 10-year survival is 100%. Even though the survival rates have changed dramatically, nothing has changed about the time of death—whether diagnosed at age 67 years or at 59 years—all patients die at age 70 years. This simple example demonstrates how survival can be increased by advancing the time of diagnosis, even if no deaths were delayed.

Lead-time bias is a mathematical certainty associated with any successful effort to detect disease early. If a test does not introduce lead time, then it fails the most basic criterion of a screening test—the ability to detect disease early. Because survival is measured from the time of diagnosis, any screening test that advances the time of diagnosis will bias the measure of survival—the lead-time bias.

The second phenomenon that leads to a spuriously high survival rate is the so-called *overdiagnosis bias*, illustrated in **Figure 3**. Overdiagnosis is the detection of pseudodisease—screen-detected abnormalities that meet the pathologic definition of

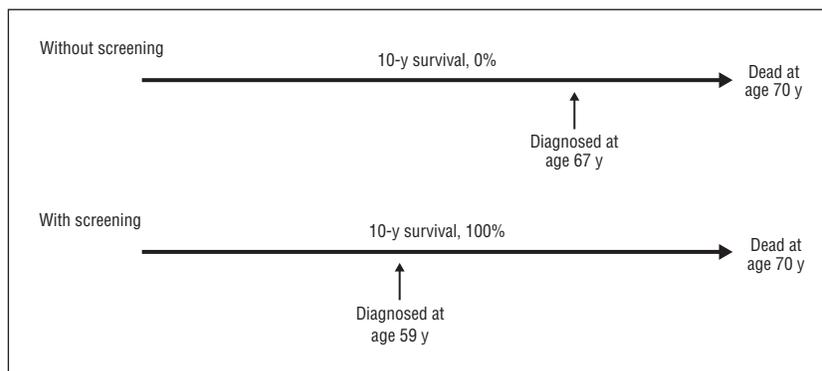


Figure 2. Lead-time bias. The diagram shows how earlier diagnosis will increase the survival statistic, even if death is not delayed.

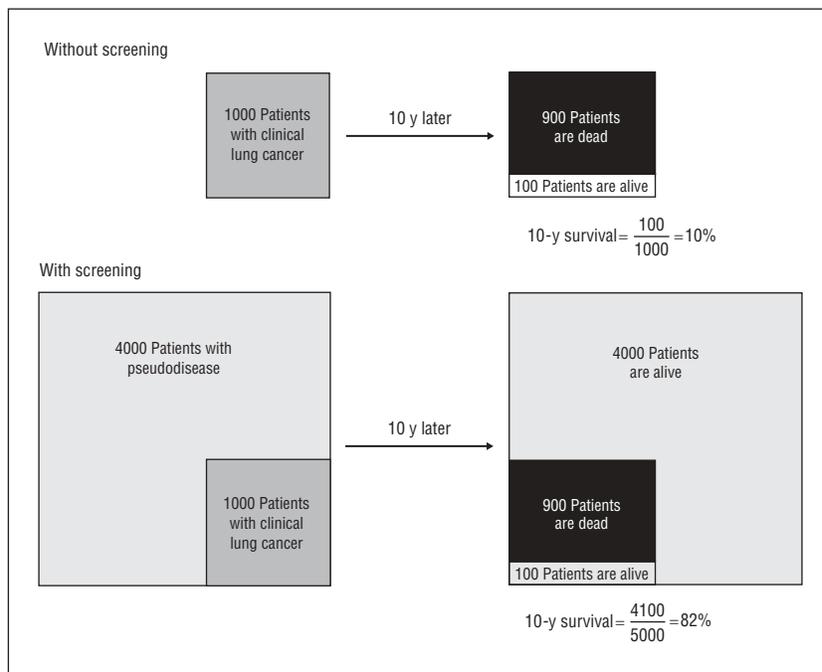


Figure 3. Overdiagnosis bias. The diagram shows how the detection of pseudodisease inflates the survival statistic even when the number of deaths is stable.

lung cancer but will never progress to cause symptoms. Although this concept may seem implausible to physicians, basic scientists have begun to uncover biological mechanisms that halt the progression of cancer.¹²⁻¹⁴ Imagine that there is no pseudodisease detected in current practice and that among 1000 patients diagnosed, only 100 are alive 10 years later (ie, the 10-year survival rate is 100 divided by 1000, or 10%). Now imagine that in addition to identifying these cancers, spiral CT also identifies 4000 patients with pseudodisease, all of whom survive 10 years. The new 10-year survival rate will include these patients in both the numerator and denominator, leading to a 10-year survival rate of

4100 divided by 5000, or 82%. Note that even though the survival rate has changed dramatically, the number of people who die has not changed: under either condition, 900 patients have died. This example demonstrates how survival can be increased by more lung cancer diagnosis, even if no one avoids death.

Although overdiagnosis is not a necessary criterion for a successful screening test, it is a common adverse effect. Overdiagnosis has now been associated with early diagnosis in a variety of cancers, including neuroblastoma,^{15,16} melanoma,¹⁷ and thyroid,¹⁸ breast,¹⁹⁻²² and prostate cancer.²³ In fact, some degree of overdiagnosis in cancer screening is probably the rule, not the exception.

Table 1. Lung Cancer Survival and Mortality in the Mayo Lung Project Randomized Trial^a

| Result | Patients Who Had Chest Radiograph Screening | Control Group |
|---|---|---------------|
| Total participants randomized, No. | 4618 | 4593 |
| Lung cancers detected during the trial, No. | 206 | 160 |
| Early stage at diagnosis, % | 46 | 31 |
| 10-y survival rate, % | 29 | 14 |
| Mortality rate per 1000 persons, y ^b | 4.4 | 3.9 |

^aSee Black²⁵ and Marcus et al.^{28,29}

^bTwo-sided *P* value for difference, .09.

Table 2. Risk of Lung Cancer in a 3-Year Spiral CT Screening Program Among Smokers and Nonsmokers^a

| Group | Lung Cancers Detected, No. | Patients Screened, No. | Risk of Lung Cancer (per 1000 Patients) ^b |
|------------|----------------------------|------------------------|--|
| Smokers | 29 | 2529 | 11.5 |
| Nonsmokers | 31 | 2954 | 10.5 |

Abbreviation: CT, computed tomography.

^aSee Sone et al.³⁰ ("Nonsmokers" indicates those who have never smoked.)

^bRelative risk for smokers vs nonsmokers = $(11.5/10.5) = 1.1$. Adjusting for the small difference in the mean number of scans for smokers (2.49) and nonsmokers (2.54) had no effect on the relative risk (1.1).

With screening, these 2 biases combine to inflate the survival statistic (eg, 5- or 10-year survival rates) even if the mortality rate is unchanged. This apparent paradox relates to the difference in the 2 metrics: mortality is the number of patients who died divided by the number of participants in the study population; the survival rate is the number alive at a specified time following diagnosis divided by the number diagnosed. The problem with survival is that the numerator is inflated by advancing the time of diagnosis (lead time) and both the numerator and denominator are inflated by the increased likelihood of being diagnosed (overdiagnosis). The mortality rate, on the other hand, is unaffected by either. This distinction makes a survival rate an unreliable statistic to evaluate progress against cancer over time: in patients with stomach cancer, for example, survival has changed minimally, whereas mortality has dropped precipitously; in patients with melanoma, survival has increased dramatically, whereas mortality has actually risen. In fact, changes in 5-year survival rates have no relationship to changes in mortality for the 20 most common solid

tumors in the United States over the last 50 years.²⁴

In the context of screening, survival is a biased metric. The forgoing examples of lead time and overdiagnosis, however, may strike some as rather extreme. Does spiral CT screening really advance the time of diagnosis this much? And is substantial overdiagnosis a real issue in lung cancer? This brings us to the third deficit in the I-ELCAP report: the context of information that is already known about lung cancer screening.

Information That Is Already Known About This Topic

It is important for readers to have the context of prior studies of lung cancer screening. Since 1960, there have been 3 randomized trials of lung cancer screening comparing chest radiographs with usual care.²⁵⁻²⁸ None found a statistically significant difference in lung cancer mortality. In fact, 2 of them reported nominally higher lung cancer mortality in the screened group. Why? Because there was more surgery in the screened group, and lung cancer resection confers a substantial mortality risk.

One of these trials, performed at the Mayo Clinic (Rochester, Minnesota),^{25,28,29} demonstrated the problems with survival outlined in the "An Unbiased Outcome Measure" subsection. As shown in **Table 1**, the 10-year survival rate in the screened group was twice as high as that in controls (29% vs 14%). Because mortality was essentially the same, the Mayo trial^{25,28,29} demonstrated that the increased survival was entirely the result of lead time and overdiagnosis. The Mayo trial also provided direct evidence of overdiagnosis associated with chest radiography screening. The excess of 46 cases in the group exposed to screening (206 vs 160) persists after more than 20 years of follow-up,²⁹ suggesting that overdiagnosis associated with chest radiography inflates the number of cases by about 30% (46/160).

There are reasons to believe that the problem of overdiagnosis in lung cancer can get much larger with spiral CT screening. Because chest radiographs superimpose multiple structures on a 2-dimensional image, their ability to identify small tumors (<3 cm) is relatively poor. But the problem of superimposition is avoided by spiral CT, which allows radiologists to "see" much smaller tumors, even those as small as 2 to 3 mm in diameter. It is these small tumors that are likely to present the greatest problem of overdiagnosis.

Investigators in Japan have highlighted this issue. They studied mass spiral CT screening in 5483 volunteers.³⁰ After the first round (ie, prevalence screen), they reported finding almost 10 times as much lung cancer as they had previously found in the same population using chest radiographs.³¹ Amazingly, the rate of lung cancer detection after completion of this 3-year screening program was virtually the same in smokers as in nonsmokers (those who have never smoked) (**Table 2**). Because the wealth of epidemiologic investigation has demonstrated that the risk of smokers dying from lung cancer is at least 10 times higher than that of nonsmokers, the Japanese data³⁰ provide powerful evidence that overdiagnosis can be a substantial problem with spiral CT screening.

Overdiagnosis in lung cancer screening is now well established in the literature.³² In fact, the problem is so substantial that the I-ELCAP investigators³ found it necessary to change the conventional screening paradigm. Instead of biopsying every abnormality, they developed a strategy of “watchful waiting” for small lesions. These lesions would be biopsied only if there was some evidence of growth. The I-ELCAP investigators deserve credit for being willing to change the paradigm of immediate biopsy. At the same time, they did not fully describe the harms from the diagnostic process—that is, the number of people without cancer who needed follow-up scans and biopsies—and the harms from overdiagnosis.

The Harms

Overdiagnosis means that some patients are being harmed. To the extent that it occurs, patients are being treated who do not have a disease to be cured. These patients cannot benefit from treatment because their “disease” posed no threat. Thus, they can only be harmed by treatment.

Although unneeded radiation or chemotherapy pose unmistakable harms, the most obvious harm from lung cancer treatment is the operative mortality associated with resection. The 30-day mortality following lung cancer resection in I-ELCAP was reported to be 0.5%. But there are reasons to worry that operative mortality would be considerably higher in the real world of community practice. Over three-quarters of the lung cancer diagnoses in I-ELCAP³ occurred in patients aged 60 to 79 years, and most underwent lobectomy. In the Medicare population aged 65 to 79 years, the 30-day mortality rate for lobectomy is almost an order of magnitude higher than reported in I-ELCAP,³ ranging from 3.6% (for those aged 65-69 years) to 6.1% (for those aged 75-79 years).³³ Furthermore, because the elevated death risk associated with cancer-directed surgery extends well beyond 30 days,³⁴ the 30-day mortality rate itself is probably an underestimate of the excess risk of death from surgery.

Moreover, it is unclear whether I-ELCAP’s measurement³ of treat-

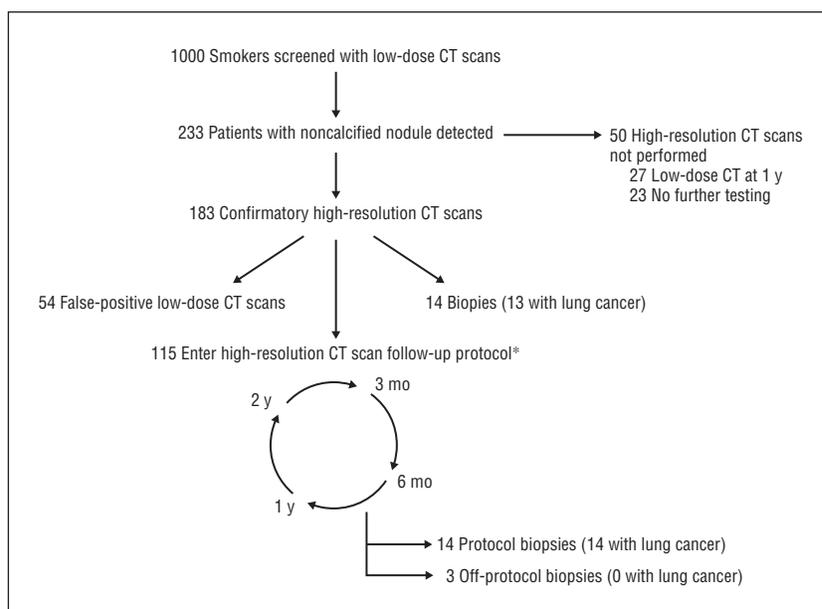


Figure 4. Cycle of scans required in initial case series of spiral computed tomographic (CT) screening.^{42,43} The asterisk indicates that 9 patients were recommended to have only 1 year of follow-up.

ment-related deaths was complete. In fact, there are reasons to wonder about the completeness and accuracy of the study’s measurement of lung cancer deaths (ie, both disease- and treatment-related deaths)—the most basic input to the primary outcome. Two questions are critical in this regard: (1) How is it known whether a study subject is alive or dead (the ascertainment of the fact of death)? (2) How is it determined whether someone died from lung cancer—or its treatment (the ascertainment of the cause of death)? No systematic method to do either was described in the I-ELCAP report.³ The fact of death was determined “when a participant was known to have died”^{3(p1767)}—thus, deaths may have been missed. The cause of death was determined “by the participant’s physicians, family members or both”^{3(p1767)}; thus, lung cancer deaths may have been misclassified. Valid measurement requires a systematic follow-up protocol to capture all deaths (eg, checking the national death index) and an independent death review committee to adjudicate its cause. There is no way to know whether treatment-related mortality has been adequately captured.

The diagnostic efforts to find cancer also can produce harm. The resolution of spiral CT is so great that many without cancer will nonethe-

less have abnormalities detected by scans. Many will have to wait to see whether growth occurs. Others undergo biopsy and face a substantial risk of requiring a chest tube for a pneumothorax (the reported risk ranges from 4% to 32% among major US medical centers).³⁵⁻⁴¹ And all of these individuals have to deal with the anxiety of wondering whether they have a deadly cancer.

Although the I-ELCAP report in the *New England Journal of Medicine*³ provides limited information about the diagnostic process, the group’s earlier work provides some of the relevant details.^{42,43} The cycle of testing is highlighted in **Figure 4**. Nearly one-quarter of all who were screened were told they had an abnormal finding on a scan (generally a nodule). More than 10% were entered in the high-resolution CT follow-up protocol, which requires multiple follow-up scans (at 3 months, 6 months, 1 year, and 2 years) to determine if growth was occurring. Given current concerns about missing cancers, the rates of abnormal findings on scans, follow-up scans, and biopsies would likely be higher in clinical practice.

Although the I-ELCAP investigators³ modified this protocol (adding trials of antibiotic therapy and positron-emission tomography [PET] scans), they did not report how this affected the cycle of test-

ing: how many CT and PET scans were performed, how often antibiotics were prescribed, or how many off-protocol biopsies were performed after screening. They also did not report the number of incidental findings in the heart, thyroid, liver, and kidneys. In a separate study by the I-ELCAP investigators,⁴⁴ nearly two-thirds of asymptomatic people were found to have coronary artery calcifications on their spiral CT. Many of these abnormalities do not translate into clinically important cardiovascular disease, and it is not known whether detecting such abnormalities (and telling people about them) does more good than harm.

HOW TO DO BETTER

The contribution of the I-ELCAP report³ was challenging for readers to understand. Its real contribution was the prudent protocol that limited the number of biopsies obtained. The study clearly suggests that this protocol misses very few cancers.

The finding of a high survival rate in patients who have cancers detected by spiral CT, by itself, provides remarkably little information. But it is totally understandable why readers might think it did, because their primary source is the journal article.

In retrospect, we believe many of the problems described herein could have been addressed with 3 editorial alterations to the abstract. The first is an additional heading for the type of study design (many journals already do this). The addition of 4 words—"Study Design: case series"—could have avoided much confusion. The second is an additional heading for limitations (at least 2 journals require this). In this study, an entry such as "Limitations: increased survival may not translate into reduced mortality" could have helped temper the inferences made. Third, the conclusion needed to be expanded beyond the single sentence published: "Conclusion: Annual spiral CT screening can detect lung cancer that is curable."^{3(p1763)} While technically true, the conclusion failed to highlight 2 fundamental unresolved questions: Would screening "cure" the 160 000 can-

cers that people die from now? Or would it "cure" cancers that never needed to be cured?

To be fair, assessing the value of screening is challenging.⁴⁵ These challenges are not limited to lung cancer screening but apply to breast, colon, prostate, and ovarian cancer screening as well. But the burden of proof must be on proponents of screening^{46,47}: although symptomatic patients generally ask for our help, by advocating screening, physicians are asserting that asymptomatic people need our help. To avoid confusion in the future, it is important to remember 2 basic screening principles.

First, survival is always prolonged by early detection. Whenever the time of diagnosis is advanced, survival will be prolonged. But that does not mean that death has been delayed or lives have been saved. Physicians need to be inoculated to recognize that the juxtaposition of the words "early detection" and "survival" equals bias. Special effort is needed to educate physicians and the public about how to interpret (and not overinterpret) survival statistics.

Second, a randomized trial is the only way to reliably determine whether screening does more good than harm. The reality of screening is that even though it offers the potential of benefit for a few (ie, those destined to develop advanced lung cancer), many more will experience no benefit at all, and some will be harmed. Consequently, the potential benefit must be demonstrated to exist and then weighed against the harms. The harms of screening are always there: some people will experience false-positive results that will cause anxiety, more testing, and, for some, physical harm (eg, in this case, pneumothoraces). Others—with pseudo-disease—will be diagnosed with cancer unnecessarily and can only be harmed (and, occasionally, killed) by therapy. The only way to know the balance of benefit and harm is a true experiment: a randomized trial. Luckily, 2 trials examining spiral CT are currently under way: the National Lung Screening Study funded by the National Cancer Institute⁴⁸ and the Netherlands-Leuven Longkanker Screenings Onderzoek trial in the Netherlands and Belgium⁴⁹ (the Pros-

tate, Lung, Colon, and Ovary Trial is examining chest radiography screening⁵⁰). We eagerly await the results, although they are at least 3 years away. In the meantime, because neither the I-ELCAP³ nor any other published data constitute sufficient evidence of its benefit, it would be a mistake to screen for lung cancer using spiral CT.

Accepted for Publication: June 29, 2007.

Correspondence: Lisa M. Schwartz, MD, MS, VA Outcomes Group, White River Junction VA Medical Center, 215 N Main St, VA Outcomes Group, 11B, White River Junction, VT 05009 (lisa.schwartz@dartmouth.edu).

Author Contributions: *Study concept and design:* Welch, Woloshin, Schwartz, Gordis, and Ransohoff. *Acquisition of data:* Woloshin. *Analysis and interpretation of data:* Welch, Woloshin, Schwartz, Gordis, Gøtzsche, Harris, Kramer, and Ransohoff. *Drafting of the manuscript:* Welch, Woloshin, and Schwarz. *Critical revision of the manuscript for important intellectual content:* Welch, Woloshin, Schwartz, Gordis, Gøtzsche, Harris, Kramer, and Ransohoff. *Statistical analysis:* Gordis. *Administrative, technical, and material support:* Welch and Schwartz. *Study supervision:* Welch.

Financial Disclosure: None reported.

Disclaimer: The views expressed herein do not necessarily represent the views of the Department of Veterans Affairs, National Institutes of Health, or the US government.

Additional Information: Drs Harris and Kramer are members of the National Lung Screening Trial. Dr Kramer is also on the Steering Committee of the Prostate, Lung, Colorectal, and Ovarian trial.

Additional Contributions: David L. Sackett, MD, MSc, provided encouragement and comments on an early draft.

REFERENCES

1. Cancer Statistics Review, section I: overview: estimated new cancer cases and deaths for 2006. SEER Web site. http://seer.cancer.gov/csr/1975_2003/index.html. Accessed December 22, 2006.
2. Peto R, Lopez A, Borehan J, Thun M, Heath C.

- Mortality From Smoking in Developed Countries 1950-2000: Indirect Estimates From National Vital Statistics.* Oxford, England: Oxford University Press; 1994.
3. Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, Miettinen OS; International Early Lung Cancer Early Action Program Investigators. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med.* 2006; 355(17):1763-1771.
 4. Health effects of secondhand smoke studied in first-of-its-kind series of clinical studies \$8.7 million Flight Attendant Medical Research Institute (FAMRI) grant to fund effort to enhance early detection and treatment: flight attendants, food and entertainment industry professionals, and other nonsmokers exposed to smoking environments recruited [news release]. Weill Cornell Medical College Web site. http://news.med.cornell.edu/wcmc/wcmc_2007/02_01b_07.shtml. Accessed February 1, 2007.
 5. Designating November 2006 as "National Lung Cancer Awareness Month" Resolution of 2006, S 620, 109th Cong, 2nd Sess (2006).
 6. No more excuses: no more lung cancer: sign the petition. Lung Cancer Alliance Web site. http://www.lungcanceralliance.org/involved/sign_the_petition.html. Accessed February 1, 2007.
 7. Screen for lung cancer: challenges. Lung Cancer Alliance Web site. <http://www.screenforlungcancer.org/challenges.html>. Accessed February 1, 2007.
 8. Bach PB, Jett J, Pastorino U, Tockman M, Swensen S, Begg C. Computed tomography screening and lung cancer outcomes. *JAMA.* 2007;297(9):953-961.
 9. Lung cancer alliance maintains strong support for early detection for lung cancer [news release]. Lung Cancer Alliance Web site. http://www.lungcanceralliance.org/news/documents/BachJAMAResponse3_6_07.pdf. Accessed August 29, 2007.
 10. Three sports legends team up to battle #1 cancer killer [news release]. Lung Cancer Alliance Web site. <http://www.lungcanceralliance.org/press/faceinthefight/index.html>. Accessed August 29, 2007.
 11. Cancer Statistics Review, section 15: lung and bronchus. SEER Web site. http://seer.cancer.gov/csr/1975_2003/index.htm. Accessed December 22, 2006.
 12. Folkman J, Kalluri R. Cancer without disease. *Nature.* 2004;427(6977):787.
 13. Mooi WJ, Peeper DS. Oncogene-induced cell senescence—halting on the road to cancer. *N Engl J Med.* 2006;355(10):1037-1046.
 14. Serrano M. Cancer regression by senescence. *N Engl J Med.* 2007;356(19):1996-1997.
 15. Schilling F, Spix C, Berthold F, et al. Neuroblastoma screening at one year of age. *N Engl J Med.* 2002;346(14):1047-1053.
 16. Yamamoto K, Hanada R, Kikuchi A, et al. Spontaneous regression of localized neuroblastoma detected by mass screening. *J Clin Oncol.* 1998; 16(4):1265-1269.
 17. Welch HG, Woloshin S, Schwartz L. Skin biopsy rates and incidence of melanoma: population based ecological study. *BMJ.* 2005;331(7515):481-484.
 18. Davies L, Welch H. The increasing incidence of thyroid cancer in the United States, 1973-2002. *JAMA.* 2006;295(18):2164-2167.
 19. Gøtzsche P, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev.* 2006;(4):CD001877.
 20. Welch HG, Schwartz LM, Woloshin S. Ramifications of screening for breast cancer: 1 in 4 cancers detected by mammography are pseudo-cancers. *BMJ.* 2006;332(7543):727.
 21. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne J. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ.* 2006; 332(7543):689-692.
 22. Zahl PH, Strand B, Maehlen J. Breast cancer incidence in Norway and Sweden during introduction of nation-wide screening: prospective cohort study. *BMJ.* 2004;328(7445):921-924.
 23. Etzioni R, Penson D, Legler J, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. *J Natl Cancer Inst.* 2002;94(13):981-990.
 24. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA.* 2000;283(22):2975-2978.
 25. Black WC. Lung cancer. In: Kramer BS, Gohagan JK, Prorok PC, eds. *Cancer Screening: Theory and Practice.* New York, NY: Marcel Dekker; 1999.
 26. Brett GZ. Earlier diagnosis and survival in lung cancer. *Br Med J.* 1969;4(5678):260-262.
 27. Kubik A, Parkin D, Khat M, Erban J, Polak J, Adamec M. Lack of benefit from semi-annual screening for cancer of the lung: follow-up report of a randomized controlled trial on a population of high-risk males in Czechoslovakia. *Int J Cancer.* 1990; 45(1):26-33.
 28. Marcus PM, Bergstralh E, Fagerstrom R, et al. Lung cancer mortality in the Mayo Lung Project: impact of extended follow-up. *J Natl Cancer Inst.* 2000;92(16):1308-1316.
 29. Marcus PM, Bergstralh E, Zweig M, Harris A, Oford K, Fontana R. Extended lung cancer incidence follow-up in the Mayo Lung Project and overdiagnosis. *J Natl Cancer Inst.* 2006;98(11):748-756.
 30. Sone S, Li F, Yang Z, Honda T, Maruyama Y, Takashima S. Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer.* 2001;84(1):25-32.
 31. Sone S, Takashima S, Li F, Yang Z, Honda T, Maruyama Y. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet.* 1998;351(9111):1242-1245.
 32. Black WC. Overdiagnosis: an underrecognized cause of confusion and harm in cancer screening. *J Natl Cancer Inst.* 2000;92(6):1280-1282.
 33. Finlayson EV, Birkmeyer J. Operative mortality with elective surgery in older adults. *Eff Clin Pract.* 2001; 4(4):172-177.
 34. Welch HG, Black W. Are deaths within one month of cancer-directed surgery attributed to cancer? *J Natl Cancer Inst.* 2002;94(14):1066-1070.
 35. Brown KT, Brody L, Getrajdman G, Napp T. Outpatient treatment of iatrogenic pneumothorax after needle biopsy. *Radiology.* 1997;205(1):249-252.
 36. Collings CL, Westcott J, Banson N, Lange R. Pneumothorax and dependent versus nondependent patient position after needle biopsy of the lung. *Radiology.* 1999;210(1):59-64.
 37. Cox JE, Chiles C, McManus C, Aquino S, Choplin R. Transthoracic needle aspiration biopsy: variables that affect risk of pneumothorax. *Radiology.* 1999;212(1):165-168.
 38. Geraghty PR, Kee S, McFarlane G, Razavi MK, Sze DY, Dake MD. CT-guided transthoracic needle aspiration biopsy of pulmonary nodules: needle size and pneumothorax rate. *Radiology.* 2003;229(2):475-481.
 39. Gupta S, Krishnamurthy S, Broemeling L, et al. Small (≤ 2 -cm) subpleural pulmonary lesions: short- versus long-needle-path CT-guided biopsy: comparison of diagnostic yields and complications. *Radiology.* 2005;234(2):631-637.
 40. Kazerooni EA, Lim F, Mikhail A, Martinez F. Risk of pneumothorax in CT-guided transthoracic needle aspiration biopsy of the lung. *Radiology.* 1996;198(2):371-375.
 41. Ko JP, Shepard J, Drucker E, et al. Factors influencing pneumothorax rate at lung biopsy: are dwell time and angle of pleural puncture contributing factors? *Radiology.* 2001;218(2):491-496.
 42. Henschke CI, McCauley D, Yankelevitz D, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet.* 1999;354(9173):99-105.
 43. Woloshin S, Schwartz L, Welch H. Tobacco money: up in smoke? *Lancet.* 2002;359(9323):2108-2111.
 44. Shemesh J, Henschke C, Feroqi A, et al. Frequency of coronary artery calcification on low-dose computed tomography screening for lung cancer. *Clin Imaging.* 2006;30(3):181-185.
 45. Welch H. *Should I Be Tested for Cancer? Maybe Not and Here's Why.* Berkeley: University of California Press; 2004.
 46. Cochrane AL, Holland W. Validation of screening procedures. *Br Med Bull.* 1971;27(1):3-8.
 47. Lee JM. Screening and informed consent. *N Engl J Med.* 1993;328(6):438-440.
 48. Clark K, Gierada D, Moore S, et al. Creation of a CT image library for the Lung Screening Study of the National Lung Screening Trial. *J Digit Imaging.* 2007;20(1):23-31.
 49. van Lersel CA, de Koning H, Draisma G, et al. Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian randomized lung cancer multi-slice CT screening trial (NELSON). *Int J Cancer.* 2007;120(4):868-874.
 50. Oken MM, Marcus P, Hu P. Baseline chest radiograph for lung cancer detection in the randomized Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. *J Natl Cancer Inst.* 2005; 97(24):1832-1839.