Cumulative Incidence of False-Positive Test Results in Lung Cancer Screening
A Randomized Trial
Jennifer M. Croswell, MD, MPH; Stuart G. Baker, ScD; Pamela M. Marcus, PhD; Jonathan D. Clapp, BS; and Barnett S. Kramer, MD, MPH

Background: Direct-to-consumer promotion of lung cancer screening has increased, especially low-dose computed tomography (CT). However, screening exposes healthy persons to potential harms, and cumulative false-positive rates for low-dose CT have never been formally reported.

Objective: To quantify the cumulative risk that a person who participated in a 1- or 2-year lung cancer screening examination would receive at least 1 false-positive result, as well as rates of unnecessary diagnostic procedures.

Design: Randomized, controlled trial of low-dose CT versus chest radiography. (ClinicalTrials.gov registration number: NCT00006382)

Setting: Feasibility study for the ongoing National Lung Screening Trial.

Patients: Current or former smokers, aged 55 to 74 years, with a smoking history of 30 pack-years or more and no history of lung cancer (n = 3190).

Intervention: Random assignment to low-dose CT or chest radiography with baseline and 1 repeated annual screening; 1-year follow-up after the final screening. Randomization was centralized and stratified by age, sex, and study center.

Measurements: False-positive screenings, defined as a positive screening with a completed negative work-up or 12 months or more of follow-up with no lung cancer diagnosis.

Results: By using a Kaplan–Meier analysis, a person’s cumulative probability of 1 or more false-positive low-dose CT examinations was 21% (95% CI, 19% to 23%) after 1 screening and 33% (CI, 31% to 35%) after 2. The rates for chest radiography were 9% (CI, 8% to 11%) and 15% (CI, 13% to 16%), respectively. A total of 7% of participants with a false-positive low-dose CT examination and 4% with a false-positive chest radiography had a resulting invasive procedure.

Limitations: Screening was limited to 2 rounds. Follow-up after the second screening was limited to 12 months. The false-negative rate is probably an underestimate.

Conclusion: Risks for false-positive results on lung cancer screening tests are substantial after only 2 annual examinations, particularly for low-dose CT. Further study of resulting economic, psychosocial, and physical burdens of these methods is warranted.

Primary Funding Source: National Cancer Institute.

For author affiliations, see end of text.
who would die of a competing cause first (6, 7)—and the risk for false-positive results. False-positive results are important because they may have negative psychological effects (8, 9), affect future adherence to other preventive health measures (10, 11), and generate physical harms and economic costs from surveillance visits and confirmatory procedures.

Although the Lung Screening Study has reported the total positivity rate, we have not previously examined cumulative false-positivity rates by using formal statistical methods, and the false-positive component most accurately represents a clinically important burden of screening. We focus on the probability of false-positive test results and resulting diagnostic procedures when chest radiography and CT are used as early detection strategies for lung cancer.

**Methods**

**Design**

The Lung Screening Study was a 2-year study conducted by 6 centers participating in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. It was a feasibility study for the ongoing National Lung Screening Trial (12, 13).

**Setting and Participants**

The Lung Screening Study had a goal of randomly assigning 3000 participants at elevated risk for lung cancer. Enrollment was achieved through mass mailings of recruitment materials, along with public service announcements, posters, and physician recruitment efforts. A total of 3318 persons were randomly assigned from September 2000 to January 2001 to have chest radiography or low-dose CT. All participants signed a consent form approved by the institutional review board before randomization.

Eligible participants were aged 55 to 74 years, had a cigarette smoking history of 30 pack-years or more, and were current smokers or had quit in the past 10 years. Exclusion criteria included chest CT within 24 months of enrollment, previous lung cancer, removal of part or all of a lung, current treatment of any cancer except nonmelanoma skin cancer, and ongoing participation in a cancer prevention or screening trial other than a smoking cessation study.

**Randomization and Interventions**

Once eligibility was established and consent was obtained by a study center, participants were randomly assigned to a treatment group through a single centralized, secure, Web-based system (which generated random code) operated by the trial coordinating center. This process ensured allocation concealment for study site investigators. Randomization was stratified by age group (in 5-year categories), sex, and study center by using variable block sizes. Once randomization occurred, participants and study investigators were not blinded to the screening method received.

Two screening examinations were possible: baseline (T0) and repeated examination (T1) 1 year later. Participants were eligible for the second screening if they did not receive a diagnosis of lung cancer after the first examination. For inclusion in this analysis of false-positive results, participants had to adhere to at least 1 screening.

Low-dose CT scans were obtained with the following technical parameters: 120 to 140 kV peak, 60 ma, scan time of 1 s, 5-mm collimation, pitch of 2 or equivalent, and contiguous reconstructions. Chest radiography consisted of single posteroanterior views and was obtained by using high-kilovolt equipment at a tube-to-receiver distance of 6 to 10 feet.

Each study center had 1 or more (range, 1 to 14) board-certified radiologists interpreting the examinations. A second radiologist blinded to the initial interpretation as a quality-control measure reread a small sample of films (n = 20) at each center.

**Outcomes and Follow-up**

For CT, the definition of a positive screening result changed slightly between the T0 and T1 scans to better match accumulating prognostic evidence: Noncalcified nodules larger than 3 mm at T0 scan or 4 mm or larger at T1 scan were considered suspicious for cancer. Other abnormalities (including spiculated noncalcified nodules of any size; focal parenchymal opacification; endobronchial lesions; hilar, mediastinal, bony, or pleural masses; and major atelectasis) could also be deemed positive according to the radiologist’s judgment. For chest radiography, nodules with circular opacity of 3.0 cm or less in diameter, masses greater than 3.0 cm, hilar or mediastinal lymph node enlargement (excepting calcified nodes), major atelectasis, in-
filtrates or consolidation, and pleural masses were considered suggestive of cancer.

We defined a false-positive screening result as a positive screening with a completed negative work-up or follow-up of at least 12 months with no diagnosis of lung cancer. Because performing biopsy of all screening-detected lung abnormalities was impractical and undesirable (because of the potential for harm), we had to choose a definition of a false-positive test result that relied on diagnostic work-ups for suspicious examinations. For persons who did not receive definitive testing, given that lung cancer as a general rule is one of the more aggressive tumors, we felt that a 12-month monitoring period was a reasonable cutoff for a false-positive result.

We defined a false-negative examination result as a negative screening associated with a diagnosis of lung cancer within 1 year of the examination. This definition is limited in its ability to discern between types of cancer truly missed by screening and aggressive interval tumors that may develop between tests. Furthermore, because the Lung Screening Study was a feasibility trial, follow-up of negative examination findings was not done in the same systematic manner as for positive test results. Persons in whom screening was negative at T1 did not continue to have follow-up in the trial; reported false-negative rates are limited to the period between the T0 and T1 examinations.

All positive results were communicated by telephone and mailed to participants and their designated physician within 3 weeks. The Lung Screening Study did not specify a diagnostic algorithm for follow-up of positive results; centers would provide recommendations for diagnostic action if requested. Center personnel abstracted medical records relating to follow-up of positive screening results. This process began after a positive screening result and continued until a conclusive diagnosis was made or 12 months had passed. In addition, study participants completed a study update form at the T1 screening to identify any interval cases of lung cancer.

Classifications for diagnostic follow-up were divided into categories by author consensus: imaging examinations (noninvasive), minimally invasive procedures (bronchoscopy), moderately invasive procedures (for example, biopsy, thoracentesis, video-assisted thoracoscopy), and major surgical procedures (thoracotomy or lung resections).

Statistical Analysis

Our study attempts to answer the question, “What is the probability that a person entering a lung cancer screening program involving 1 or 2 screening tests will have at least 1 false-positive CT or chest radiography?” A person could contribute to the cumulative risk curve only once, after a first false-positive test result; this avoided double-counting of suspicious nodules and artificial inflation of the curve. Kaplan–Meier analysis generated cumulative incidence curves based on estimated probability of a first false-positive result at baseline or second screening received (14). For the base-case analysis, we considered persons with incomplete follow-up (<12 months) after a positive screening to have received a false-positive result. As a sensitivity analysis, we assumed that a proportion of persons with insufficient follow-up after a positive examination result did have cancer, in which the proportion was estimated as published positive predictive values (from other trials) of CT (7%) and chest radiography (2%) for screening detection of lung cancer (15, 16).

Logistic regression was done through 2 models to identify potential participant characteristics associated with increased odds of a false-positive examination result after the first screening or the second screening (if the first screening was negative). Variables included age, current versus former smoking status, and smoking history of 60 pack-years or more versus 30 to 59 pack-years. We adjusted logistic regression analyses for screening center. We calculated odds ratios (ORs) separately for CT and chest radiography. We also examined variance in the false-positive rate by study center.

Role of the Funding Source

The Lung Screening Study was performed under contracts awarded by the National Cancer Institute, and all data were collected under those contracts. Most of the authors who designed, analyzed, and wrote this study are federal employees.

RESULTS

Demographic Characteristics

Table 1 shows demographic characteristics of the study population. Participants were more likely to be men (59%) and current smokers (57%). As expected, randomization achieved balance in baseline patient characteristics.

### Table 1. Demographic Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Low-Dose CT (n = 1610), n (%)</th>
<th>Chest Radiography (n = 1580), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65–74 y</td>
<td>516 (32)</td>
<td>502 (32)</td>
</tr>
<tr>
<td>55–64 y</td>
<td>1094 (68)</td>
<td>1078 (68)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>675 (42)</td>
<td>637 (40)</td>
</tr>
<tr>
<td>Male</td>
<td>935 (58)</td>
<td>943 (60)</td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>929 (58)</td>
<td>897 (57)</td>
</tr>
<tr>
<td>Former</td>
<td>681 (42)</td>
<td>683 (43)</td>
</tr>
<tr>
<td><strong>Smoking history</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥60 pack-years</td>
<td>673 (42)</td>
<td>687 (43)</td>
</tr>
<tr>
<td>30–59 pack-years</td>
<td>937 (58)</td>
<td>893 (57)</td>
</tr>
</tbody>
</table>

CT = computed tomography.
Adherence Rates

A total of 1610 participants underwent at least 1 CT, and 1580 participants underwent at least 1 chest radiography (Figure 1). A total of 1374 participants in the CT group (97%) and 1287 participants in the chest radiography group (95.2%) received both examinations. Adherence was lower in both groups for the second screening than for the baseline test.

Figure 1. Study flow diagram.

False-Positive and False-Negative Results

A total of 31% (n = 506) of participants in the CT group and 14% (n = 216) of participants in the chest radiography group received at least 1 false-positive result. In comparison, screening CT was true-positive in 38 instances (2% of participants) and chest radiography was true-positive in 16 instances (1% of participants).

Figure 2 illustrates the cumulative risk that a person would receive at least 1 false-positive result in a lung cancer screening program over several years. At baseline screening, the risk for a false-positive result is 21% (95% CI, 19% to 23%) for CT and 9% (CI, 8% to 11%) for chest radiography. These risks increase to 33% (CI, 31% to 35%) and 15% (CI, 13% to 16%), respectively, at second examination. Sensitivity analysis, which assumed that 7% of participants in the CT group and 2% of participants in the chest radiography group with insufficient follow-up after a positive screening had true-positive results, yielded identical rates. Four examinations with false-negative results were reported between the baseline and T1 examinations. All were in the chest radiography group (0.2% of participants in this group).

Diagnostic Follow-up and Invasive Procedures

Table 2 shows absolute rates of diagnostic follow-up by category. Of persons with at least 1 false-positive CT or chest radiography, 61% (n = 308) and 51% (n = 110), respectively, received 1 or more secondary imaging tests. The overall percentage of participants who had at least 1 invasive procedure as a result of a false-positive result was 7% for CT and 4% for chest radiography. Bronchoscopies were the most common invasive procedure resulting from false-positive CT (5% of participants with a false-positive result). The proportion of participants who had a moderately invasive procedure was 4% for false-positive CT and 3% for false-positive chest radiography. Rates of participants who had major surgical procedures for benign disease were similar (2%), although absolute numbers differed by a factor of 2 (8 CT recipients and 4 chest radiography recipients).

Participant Characteristics and False-Positive Risk

We did multivariable analyses to investigate potential participant characteristics associated with increased odds of a false-positive result after first or second screening received (in which the first screening was negative). Our models included age (65 to 74 years vs. 55 to 64 years), smoking status (current vs. former), smoking history (≥60 pack-years vs. 30 to 59 pack-years), and study center.

First Screening Received

For CT, older versus younger age demonstrated a non–statistically significant trend toward increased odds of
a false-positive screening result (OR, 1.28 [CI, 0.98 to 1.67]). This effect was not seen in the chest radiography group (OR, 1.11 [CI, 0.77 to 1.60]). Current versus former smoking status did not show an association with false-positive results (OR, 1.15 [CI, 0.88 to 1.49] for CT and 1.31 [CI, 0.92 to 1.88] for chest radiography). Greater number of pack-years smoked was not associated with increased odds of a false-positive result (OR, 0.92 [CI, 0.71 to 1.19] for CT and 1.02 [CI, 0.72 to 1.45] for chest radiography).

The estimated probability of a false-positive result on the first screening received varied by center. For the combined categories of age (65 to 74 years), current smoker, and smoking history of 60 pack-years or more, the estimate for CT varied from 10% to 42% across centers. For chest radiography, the estimated false-positive rate for the combined categories of age (65 to 74 years), current smoker, and smoking history of 60 pack-years or more varied from 3% to 19% across centers. The Appendix (available at www.annals.org) provides further details.

Second Screening Received

Older versus younger age was not associated with odds of false-positive CT (OR, 1.20 [CI, 0.83 to 1.74]) but was associated with twice the odds of false-positive chest radiography (OR, 2.03 [CI, 1.23 to 3.36]). Current versus former smoking status was not associated with odds of false-positive CT (OR, 1.07 [CI, 0.75 to 1.52]) or chest radiography (OR, 0.85 [CI, 0.51 to 1.40]). More pack-years of smoking was associated with 1.5 times increased odds of a false-positive result for CT (OR, 1.53 [CI, 1.08 to 2.18]), but this was not observed in the chest radiography group (OR, 1.12 [CI, 0.68 to 1.85]).

**DISCUSSION**

To our knowledge, our study is the first to formally evaluate a person’s cumulative risk for receiving at least 1 false-positive test result in a program of low-dose CT screening for lung cancer over several years. The probability of a false-positive result is substantial after 2 annual examinations (33%). Of participants with a false-positive CT scan, 7% had an unnecessary invasive procedure and 2% had major surgery for benign disease.

Cumulative false-positive rates associated with screening tests have been infrequently reported; most studies of this nature have focused on mammography (17–20). An

### Table 2. Rates of False-Positive Results That Prompted Diagnostic Follow-up*

<table>
<thead>
<tr>
<th>Procedure Type</th>
<th>Low-Dose CT (n = 506)†</th>
<th>Chest Radiography (n = 216)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Participants, n</td>
<td>False-Positive Results, %</td>
</tr>
<tr>
<td>Imaging examinations</td>
<td>308</td>
<td>61</td>
</tr>
<tr>
<td>Minimally invasive procedure‡</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>Moderately invasive procedure§</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>Major surgical procedure†</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Any invasive procedure¶</td>
<td>33</td>
<td>7</td>
</tr>
</tbody>
</table>

CT = computed tomography.
* Approximately 3% of all participants with a positive screening result received no diagnostic follow-up.
† Total number of persons with at least 1 false-positive result in each group. A person could contribute only once to each category, even if he or she had more than 1 false-positive result.
‡ Bronchoscopy.
§ Procedures included lung biopsy (any method except open surgery), lymph node biopsy, mediastinoscopy, mediastinotomy, thoracectomy, and video-assisted thoracoscopy.
¶ Procedures included lung resection and thoracotomy.

A person could contribute to this group only once, regardless of the number of procedures he or she had. Imaging examinations were not considered invasive procedures.
estimate of the cumulative false-positive rate for chest radiography in lung cancer screening as part of a larger evaluation of false-positive rates in screening programs that used multiple testing methods was about 14% after 2 screenings and 22% after 4 screenings, which aligns with our findings (21). Previous studies have generally reported discovery rates of all noncalcified nodules in persons who had CT screening. These rates have varied widely, depending on screening frequency and other factors; the average range is 25% to 50% of participants (16, 22).

Clinical experience suggests that, in general, the larger the nodule, the greater the suspicion that the lesion is or will become cancerous (23). Despite this rule of thumb, no uniform consensus on how best to categorize and manage lesions detected on CT exists. Several studies of CT for lung cancer screening used protocols that called for diagnostic follow-up of varying intensity for all detected nodules (24–29), although most lesions detected by low-dose CT are smaller than 4 mm (30). Fleischner Society guidelines for management of incidental nodules detected on nonscreening CT suggest that lesions smaller than 4 mm pose minimal risk, and as such, generally recommend no further follow-up for low-risk patients and a single repeated screening at 12 months for high-risk patients with no intervention if the lesion is unchanged (31). The definition of a positive test result used in the Lung Screening Study (≥3 mm or 4 mm, depending on year) attempted to rule out lesions that were least likely to be indicative of cancer; the study design sought to minimize the number of false-positive results. The relatively high cumulative risk for a false-positive result after 2 examinations may represent a conservative estimate of rates in community practice, in which all nodules, regardless of size, may be more likely to receive follow-up.

Multislice scanners were emerging at the time of the Lung Screening Study, and some, but perhaps not all, of the study sites had 4-row scanners in use. It is not known what the effect of single versus multirow detectors would be on false-positive rates, nor the effect of additional (for example, 16 or 64) slices.

Previous studies have reported on surgery rates for benign disease; the proportion of persons with a screening-detected nonmalignant lesion who had surgery has ranged from about 0.6% to 2.7% for CT (26–28, 32–35). This is consistent with our findings.

More than half of participants with false-positive chest radiography or CT had at least 1 additional imaging examination—some at higher radiation doses than that of the original test—which exposed these persons to a theoretical risk for radiation-induced carcinogenesis. This is potentially concerning in the target population because current evidence indicates that radiation and smoking damage interact synergistically, and the interaction is near multiplicative (36, 37). Consensus has not been reached on a single diagnostic algorithm for positive screening examinations; however, repeated scans (of varying dose) at 1, 3, 6, 12, and/or 24 months have been advocated (23–29, 38). Although no long-term longitudinal data on the effects of cumulative radiation exposure are available, estimates of cancer risks have been done by using relevant data from cohort studies of survivors of long-term atomic bomb exposure (39, 40). One estimate found that among female smokers aged 60 years, a series of 3 low-dose chest CTs would generate an excess lung cancer risk for about 1.5 women per 1000 exposed (39). The National Lung Screening Trial is collecting information on cumulative radiation exposure among participants and should help clarify this important issue.

Negative psychological consequences of false-positive screening are also of concern. Although data for lung cancer screening are limited, 1 study examining the effect of CT on quality of life found that about half of participants had “discomfort and dread” while waiting for confirmation of screening results (41). Because positive lung cancer screening results may encompass diagnostic uncertainty for up to 24 months, further investigation of the long-term psychological effects of false-positive test results would be important.

A high rate of false-positive test results may place economic burdens on persons and the health care system. One study that investigated medical care expenditures triggered by false-positive test results found that the adjusted mean difference in spending after such a test was an additional $1024 for a woman and $1171 for a man in 1 year (42). Because 31% of participants in the CT group received at least 1 false-positive result in 2 years, the effect these rates could have on the cost-effectiveness of CT (if proven effective) is apparent.

Our study has several limitations. As with all cancer screening trials, a “healthy volunteer effect” is likely to be present. This effect has been documented in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial, which includes study centers that participated in the Lung Screening Study (43).

The Lung Screening Study was a pilot study and, as such, was limited to 2 rounds of screening and 1 year of follow-up after the last examination. Although it is possible that this study might overestimate the cumulative risk for a false-positive result, it is unlikely that a large proportion of false-positive results would convert to diagnoses of cancer. A small proportion (0.04% for each method) of Lung Screening Study participants was lost to follow-up after positive examination findings; this study’s baseline assumption was that those findings were false positive. Because this could potentially overestimate the cumulative false-positive risk, a sensitivity analysis was done; however, the cumulative risk estimates remained unchanged.

Because the Lung Screening Study was a feasibility trial, negative screening results were not systematically followed in the same manner as positive screening results. Recording of false-negative results was limited to the period between the T0 and T1 examinations, and the estimated false-negative rate is crude and probably an underestimate.
The analysis of associations between patient characteristics and false-positive risk should be considered exploratory. The Lung Screening Study had a relatively modest sample size; subgroup populations were small, and, as in any multivariable analysis, multiplicity potentially inflated the probability of a type I error. The variation in estimated false-negative rates by center is probably due to the effects of interobserver variability in the assessment of a positive examination among centers. Studies of interobserver agreement on interpretation of chest CT scans have demonstrated notable variations among readers (44–46). Semi-automated volumetric determinations of nodule size or computer-aided detection programs have been proposed as methods that could theoretically reduce observer variability, although the ultimate degree of effect this might have on false-negative rates is unknown.

Given the relatively high probability of a false-positive low-dose CT lung cancer screening examination, it is important that providers have careful discussions with patients who request this technology to help them weigh known harms against currently theoretical benefits. Further investigation into the physical, psychological, and economic ramifications of false-positive low-dose CT screening test results is warranted.

From National Institutes of Health, National Cancer Institute, Bethesda, and Information Management Services, Rockville, Maryland.

Portions of this work were presented at the annual meeting of the American Society of Clinical Oncology, Orlando, Florida, 29 May–2 June 2009 (Abstract 1502).

Disclaimer: The views expressed in this article are those of the authors and do not necessarily represent the views of the U.S. federal government or the National Institutes of Health.

Acknowledgment: The authors thank all participants in the Lung Screening Study and the centers and health professionals involved in this trial. They also thank Dr. Phil Prorok, Dr. Christine Berg, the National Lung Screening Trial Executive Committee, and Dr. Thomas Louis for their helpful input during study development.

Grant Support: The Lung Screening Study was funded by the National Cancer Institute; no monies or grants were awarded to any of the authors to perform this study.

Potential Conflicts of Interest: Dr. Crosswell: Financial relationships involving spouse: Husband owns shares in Johnson & Johnson. Dr. Kramer: Money received: Journal of the National Cancer Institute. Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M09-1802.

Reproducible Research Statement: Study protocol: Available at http://prevention.cancer.gov/programs-resources/groups/ed/programs/lis/manual. Statistical code: Written in Mathematica 7.0 (Wolfram Research, Champaign, Illinois) and available by written request to Dr. Baker (e-mail, bakers@mail.nih.gov). Data set: Available by written request (to include the reader’s objectives and the specific data required) to Dr. Baker (e-mail, bakers@mail.nih.gov).

www.annals.org

False-Positive Test Results in Lung Cancer Screening

Requests for Single Reprints: Jennifer M. Crosswell, MD, MPH, Office of Medical Applications of Research, National Institutes of Health, 6100 Executive Boulevard, Suite 2B-03, Bethesda, MD 20892; e-mail, crosswellj@od.nih.gov.

Current author addresses and author contributions are available at www.annals.org.

References


False-Positive Test Results in Lung Cancer Screening


**Current Author Addresses:** Dr. Croswell: Office of Medical Applications of Research, National Institutes of Health, 6100 Executive Boulevard, Suite 2B-03, Bethesda, MD 20892.
Dr. Baker: Biometry Research Group, Division of Cancer Prevention, National Cancer Institute, EPN 3131, 6130 Executive Boulevard, Bethesda, MD 20892-7354.
Dr. Marcus: Division of Cancer Prevention, National Cancer Institute, 6130 Executive Boulevard, Suite 3131, Bethesda, MD 20892-7354.
Mr. Clapp: Information Management Services, 6110 Executive Boulevard, Suite 310, Rockville, MD 20852.
Dr. Kramer: Office of Disease Prevention, National Institutes of Health, 6100 Executive Boulevard, Suite 2B-03, Bethesda, MD 20892.

**Author Contributions:** Conception and design: J.M. Croswell, B.S. Kramer.
Drafting of the article: J.M. Croswell, B.S. Kramer.
Critical revision of the article for important intellectual content: J.M. Croswell, S.G. Baker, P.M. Marcus, J.D. Clapp.
Statistical expertise: S.G. Baker.
Collection and assembly of data: J.M. Croswell, P.M. Marcus, J.D. Clapp.

**Appendix**

This appendix describes the statistical calculations in more detail.

**Summary of the Data**

The possible outcomes were coded as 0, indicating no false-positive result (no FP); 1, indicating a false-positive result (FP); 2, indicating that the test was missing (missing); or 3, indicating a positive test result with insufficient follow-up (PWIF) (Appendix Table 1).

The data are summarized in Appendix Tables 1 to 9.

**Formulas for Computing Estimate False-Positive Rates**

Let \( x_i \) denote a number with outcome \( i \) at round \( T_0 \) and outcome \( j \) at round \( T_1 \), for a particular screening method, either chest radiography or low-dose CT. These are the numbers in the cells of Appendix Table 1. Let \( x_{1p} = x_{00} + x_{01} + x_{12} + x_{13} \) denote the row sums.

We let \( f \) denote the probability that a person who received a positive result on screening with insufficient follow-up had a true-positive result (that is, a disease). Under the base case, we set \( f = 0 \). Under the sensitivity analysis, we set \( f = 0.02 \) for chest radiography and \( f = 0.07 \) for low-dose CT. The estimate of \( f \), which comes from other studies, is the probability of disease given a positive screening test result, which is called the positive predictive value. One minus the positive predictive value is the probability of no disease given a positive screening test result.

Our formulas are as follows.

For a first screening at \( T_0 \):

\[
\begin{align*}
y_{1a} &= \text{number with a first FP at } T_0 = x_{1p} + x_{32} (1 - f) \\
n_1 &= \text{number at risk for FP at } T_0 = x_{0p} + x_{1p} + x_{32}
\end{align*}
\]

For a first screening at \( T_1 \) because of missing at \( T_0 \):

\[
\begin{align*}
y_{1b} &= \text{number with a first FP at } T_1 = x_{21} + x_{23} (1 - f) \\
n_{1b} &= \text{number at risk for first FP at } T_1 = x_{20} + x_{21} + x_{23}
\end{align*}
\]

We can then write:

\[
\begin{align*}
y_1 &= y_{1a} + y_{1b} &= \text{number with a first FP} \\
n_1 &= n_{1a} + n_{1b} &= \text{number at risk for a first FP}
\end{align*}
\]

For a second screening at \( T_1 \) after no FP at \( T_0 \):

\[
\begin{align*}
y_2 &= \text{number with FP at } T_1 = x_{01} + x_{03} (1 - f) \\
n_2 &= \text{number at risk for FP at } T_1 = x_{00} + x_{01} + x_{03}
\end{align*}
\]

From the above quantities, we computed:

\[
\begin{align*}
h_1 &= \text{hazard for first FP} = h_{1} = y_1/n_1 \\
h_2 &= \text{hazard for FP after FP} = 0 = h_2 = y_2/n_2
\end{align*}
\]

The estimated probability of no FPs on 2 screenings is

\[
S = (1 - h_1)(1 - h_2)
\]

Therefore the estimated probability of at least one FP on 2 screenings is

\[
P = 1 - S
\]

From the binomial distribution, \( \text{var}(h_i) = h_i (1 - h_i)/n_i \). By using the delta method,

\[
\text{var}(P) \approx \text{var}(S) = S^2 \text{var}(\log(S))
\]

\[
= S^2 \frac{\text{var}(h_1) + \text{var}(h_2)}{(1 - h_1)^2 + (1 - h_2)^2} = x_i/n_i (1 - x_i) + x_j/n_j (1 - x_j) \]

which is the Greenwood formula.

**Computed False-Positive Rates**

For low-dose CT, using the formulas and data cells described, \( h_1 = 335/1600 = 0.21 \). This is the estimated probability of a false-positive result on the first screening received. \( h_2 = 171/114 = 0.15 \); this is the estimated probability of a false-positive result on the second screening received, given that the first result was negative. The cumulative probability of a false-positive result is therefore \( h_t (0.21) \) for the first test received and \( 1 - (1 - h_1)(1 - h_2) = 1 - (1 - 0.21)(1 - 0.15) = 0.33 \) for the second test received.

For chest radiography, \( h_1 = 147/1580 = 0.09 \). This is the estimated probability of a false-positive result on the first screening received. \( h_2 = 69/1183 = 0.06 \); this is the estimated probability of a false-positive result on the second screening received, given that the first result was negative. The cumulative probability of a false-positive result is therefore \( h_t (0.09) \) for the first test received and \( 1 - (1 - h_1)(1 - h_2) = 1 - (1 - 0.09)(1 - 0.06) = 0.15 \) for the second test received.

**Logistic Regression Model**

For each screening method, we fit the following logistic regression models:

\[
\text{logit}[\text{pr(FP on first screening received})] = \beta_{age} I(age) + \beta_{smoker} I(smoker) + \Sigma_i \beta_{CI} I(center i) + \beta_py I(\text{pack-years})
\]

\[
\text{logit}[\text{pr(FP on second screening received after no FP on first screening received})] = \beta_{age} I(age) + \beta_{smoker} I(smoker) + \Sigma_i \beta_{CI} I(center i) + \beta_py I(\text{pack-years})
\]

in which

\[
I(smoker) = 1 \text{ if current smoker, and 0 if former smoker,}
\]

\[
I(\text{pack-years}) = 1 \text{ if pack-years, and 0 otherwise}
\]
If the outcome at T0 in the original pair was 1(FP), then the outcome of the new pair was coded as {1, 2} because the outcome at the second screening is not in the risk set. If original pair = {1, 0}, then new pair = {1, 2}
If original pair = {1, 1}, then new pair = {1, 2}
If original pair = {1, 2}, then new pair = {1, 2}
If original pair = {1, 3}, then new pair = {1, 2}
If the outcome at T0 in the original pair was 2 (missing), then the outcome of the new pair was coded as [outcome at T1, 2] because the screening at T1 is now the first screening, and the second screening is not in the risk set. Also PWIF was coded as FP under the base case.
If original pair = {2, 0}, then new pair = {0, 2}
If original pair = {2, 1}, then new pair = {1, 2}
If original pair = {2, 2}, then new pair = {2, 2}
If original pair = {2, 3}, then new pair = {1, 2}
If the outcome at T0 in the original pair was 3 (PWIF), then the outcome of the new pair was coded as {1, 2} because PWIF is coded as 1(FP) under the base case, and the second screening is not in the risk set.
If original pair = {3, 0}, then new pair = {1, 2}
If original pair = {3, 1}, then new pair = {1, 2}
If original pair = {3, 2}, then new pair = {1, 2}
If original pair = {3, 3}, then new pair = {1, 2}

Recoding Data for Fitting Logistic Regression Models

Before fitting the logistic regressions we recoded the original data as follows. We assumed the base case in which a PWIF is an FP. We denote the original pair of data by {outcome at T0, outcome at T1}, in which 0 indicated no false-positive result (no FP), 1 indicated a false-positive result (FP), 2 indicated a missing test, and 3 indicated a positive result with insufficient follow-up (PWIF).

We recoded the original pair as a new pair, {outcome for first screening received and outcome for second screening received after no FP on first screening}, in which 0 indicated no false-positive result (no FP), 1 indicated a false-positive result (FP), 2 indicated a missing result (Missing), and 3 indicated information outside of the risk set.

The rules for recoding were as follows.

If the outcome at T0 in the original pair was 0 (no FP), then the outcome of the new pair was identical to that for the original pair, except that PWIF was coded as FP under the base case.
If original pair = {0, 0}, then new pair = {0, 0}
If original pair = {0, 1}, then new pair = {0, 1}
If original pair = {0, 2}, then new pair = {0, 2}
If original pair = {0, 3}, then new pair = {0, 1}

If the outcome at T0 in the original pair was 1(FP), then the outcome of the new pair was coded as {1, 2} because the outcome at the second screening is not in the risk set.
If original pair = {1, 0}, then new pair = {1, 2}
If original pair = {1, 1}, then new pair = {1, 2}
If original pair = {1, 2}, then new pair = {1, 2}
If original pair = {1, 3}, then new pair = {1, 2}
If the outcome at T0 in the original pair was 2 (missing), then the outcome of the new pair was coded as [outcome at T1, 2] because the screening at T1 is now the first screening, and the second screening is not in the risk set. Also PWIF was coded as FP under the base case.
If original pair = {2, 0}, then new pair = {0, 2}
If original pair = {2, 1}, then new pair = {1, 2}
If original pair = {2, 2}, then new pair = {2, 2}
If original pair = {2, 3}, then new pair = {1, 2}
If the outcome at T0 in the original pair was 3 (PWIF), then the outcome of the new pair was coded as {1, 2} because PWIF is coded as 1(FP) under the base case, and the second screening is not in the risk set.
If original pair = {3, 0}, then new pair = {1, 2}
If original pair = {3, 1}, then new pair = {1, 2}
If original pair = {3, 2}, then new pair = {1, 2}
If original pair = {3, 3}, then new pair = {1, 2}
### Appendix Table 3. Odds Ratio (95% CI) for the Effect of Age, Smoking Status, and Smoking History on the False-Positive Rate on First Screening Received, Based on Logistic Regression Model*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low-Dose CT (n = 1160)</th>
<th>Chest Radiography (n = 1580)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65–74 y</td>
<td>1.28 (0.98–1.67)</td>
<td>1.11 (0.77–1.60)</td>
</tr>
<tr>
<td>55–64 y</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>1.15 (0.88–1.49)</td>
<td>1.31 (0.92–1.88)</td>
</tr>
<tr>
<td>Former</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥60 pack-years</td>
<td>0.92 (0.71–1.19)</td>
<td>1.02 (0.72–1.45)</td>
</tr>
<tr>
<td>30–59 pack-years</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

CT = computed tomography.

* Estimates are adjusted for center.

### Appendix Table 4. Estimated False-Positive Rate on First Screening Received for Low-Dose Computed Tomography, Based on Logistic Regression Model*

<table>
<thead>
<tr>
<th>Center</th>
<th>0, O, 0</th>
<th>0, O, P</th>
<th>0, S, O</th>
<th>0, S, P</th>
<th>A, O, 0</th>
<th>A, O, P</th>
<th>A, S, O</th>
<th>A, S, P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.32</td>
<td>0.30</td>
<td>0.35</td>
<td>0.33</td>
<td>0.37</td>
<td>0.35</td>
<td>0.40</td>
<td>0.39</td>
</tr>
<tr>
<td>2</td>
<td>0.08</td>
<td>0.07</td>
<td>0.09</td>
<td>0.08</td>
<td>0.09</td>
<td>0.09</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td>3</td>
<td>0.10</td>
<td>0.09</td>
<td>0.11</td>
<td>0.10</td>
<td>0.12</td>
<td>0.11</td>
<td>0.13</td>
<td>0.12</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>0.09</td>
<td>0.11</td>
<td>0.10</td>
<td>0.12</td>
<td>0.11</td>
<td>0.13</td>
<td>0.12</td>
</tr>
<tr>
<td>5</td>
<td>0.22</td>
<td>0.21</td>
<td>0.25</td>
<td>0.23</td>
<td>0.27</td>
<td>0.25</td>
<td>0.30</td>
<td>0.28</td>
</tr>
<tr>
<td>6</td>
<td>0.13</td>
<td>0.12</td>
<td>0.14</td>
<td>0.13</td>
<td>0.16</td>
<td>0.15</td>
<td>0.17</td>
<td>0.16</td>
</tr>
</tbody>
</table>

* A means I(age) = 1; S means I(smoker) = 1; P means I(pack-years of smoking) = 1. Analysis based on 1160 participants who received at least 1 computed tomography examination.
### Appendix Table 5. Estimated False-Positive Rate on First Screening Received for Chest Radiography, Based on Logistic Regression Model*

<table>
<thead>
<tr>
<th>Center</th>
<th>0, 0, 0</th>
<th>0, 0, P</th>
<th>0, S, 0</th>
<th>0, S, P</th>
<th>A, 0, 0</th>
<th>A, 0, P</th>
<th>A, S, 0</th>
<th>A, S, P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.14</td>
<td>0.14</td>
<td>0.17</td>
<td>0.17</td>
<td>0.15</td>
<td>0.15</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>0.08</td>
<td>0.08</td>
<td>0.10</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>3</td>
<td>0.05</td>
<td>0.06</td>
<td>0.07</td>
<td>0.07</td>
<td>0.06</td>
<td>0.06</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>0.10</td>
<td>0.13</td>
<td>0.13</td>
<td>0.11</td>
<td>0.11</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>5</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>6</td>
<td>0.08</td>
<td>0.08</td>
<td>0.10</td>
<td>0.11</td>
<td>0.09</td>
<td>0.09</td>
<td>0.11</td>
<td>0.12</td>
</tr>
</tbody>
</table>

* A means I(age) = 1; S means I(smoker) = 1; P means I(pack-years of smoking) = 1. Analysis based on 1580 participants who received at least 1 chest radiography examination.

### Appendix Table 6. Parameter Estimates (±SE) for the Logistic Regression Model for the False-Positive Rate on Second Screening Received*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter Estimate (±SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-Dose CT (n = 1114)</td>
</tr>
<tr>
<td></td>
<td>Chest Radiography (n = 1183)</td>
</tr>
<tr>
<td>( \beta_{\text{age}} )</td>
<td>0.18 ± 0.019</td>
</tr>
<tr>
<td>( \beta_{\text{smoker}} ) (current smoker)</td>
<td>-0.07 ± 0.18</td>
</tr>
<tr>
<td>( \beta_{\text{center 2}} )</td>
<td>-2.75 ± 0.28</td>
</tr>
<tr>
<td>( \beta_{\text{center 3}} )</td>
<td>-3.02 ± 0.33</td>
</tr>
<tr>
<td>( \beta_{\text{center 4}} )</td>
<td>-1.38 ± 0.25</td>
</tr>
<tr>
<td>( \beta_{\text{center 5}} )</td>
<td>-1.59 ± 0.24</td>
</tr>
<tr>
<td>( \beta_{\text{center 6}} )</td>
<td>-2.76 ± 0.29</td>
</tr>
<tr>
<td>( \beta_{\text{py}} ) (≥60 pack-years)</td>
<td>0.43 ± 0.18</td>
</tr>
</tbody>
</table>

### Appendix Table 7. Odds Ratio for the Effect of Age and Smoking Status on False-Positive Rate on Second Screening Received*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low-Dose CT (n = 1114)</th>
<th>Chest Radiography (n = 1183)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>65–74 y: 1.20 (0.83–1.74)</td>
<td>2.03 (1.23–3.36)</td>
</tr>
<tr>
<td></td>
<td>55–64 y: 1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Smoking status</td>
<td>Current: 1.07 (0.75–1.52)</td>
<td>0.85 (0.51–1.40)</td>
</tr>
<tr>
<td></td>
<td>Former: 1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Smoking history</td>
<td>≥60 pack-years: 1.53 (1.08–2.18)</td>
<td>1.12 (0.68–1.85)</td>
</tr>
<tr>
<td></td>
<td>30–59 pack-years: 1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

CT = computed tomography.
* When baseline examination was negative.
† Estimates adjusted for center.

---

CT = computed tomography.
* When baseline examination was negative.
† Estimates adjusted for center.
### Appendix Table 8. Marginal Counts, Crude Rates (FP Probabilities), and Crude ORs for Each Variable Included in Appendix Table 3

<table>
<thead>
<tr>
<th>Variable, by FP Result</th>
<th>Low-Dose CT (n = 1610)</th>
<th>Chest Radiography (n = 1580)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FP = 0</td>
<td>FP = 1</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65–74 y</td>
<td>393</td>
<td>123*</td>
</tr>
<tr>
<td>55–64 y</td>
<td>882</td>
<td>212‡</td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>724</td>
<td>205‡‡</td>
</tr>
<tr>
<td>Former</td>
<td>551</td>
<td>130**</td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥60 pack-years</td>
<td>535</td>
<td>138‡‡</td>
</tr>
<tr>
<td>30–59 pack-years</td>
<td>740</td>
<td>197</td>
</tr>
</tbody>
</table>

CT = computed tomography; FP = false-positive; OR = odds ratio.  
* Probability of FP if aged 65–74 y = 0.24; crude OR = 1.30.  
† Probability of FP if aged 65–74 y = 0.10; crude OR = 1.12.  
‡ Probability of FP if aged 55–64 y = 0.19.  
§ Probability of FP if aged 55–64 y = 0.09.  
‖ Probability of FP if current smoker = 0.22; crude OR = 1.20.  
¶ Probability of FP if current smoker = 0.10; crude OR = 1.35.  
** Probability of FP if former smoker = 0.19.  
†† Probability of FP if former smoker = 0.08.  
‡‡ Probability of FP if ≥60 pack-years = 0.21; crude OR = 0.97.  
§§ Probability of FP if ≥60 pack-years = 0.09; crude OR = 1.00.  
|| Probability of FP if 30–59 pack-years = 0.21.  
††† Probability of FP if 30–59 pack-years = 0.09.  

### Appendix Table 9. Marginal Counts, Crude Rates (FP Probabilities), and Crude ORs for Each Variable Included in Appendix Table 7

<table>
<thead>
<tr>
<th>Variable, by FP Result</th>
<th>Low-Dose CT (n = 1114)</th>
<th>Chest Radiography (n = 1580)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FP = 0</td>
<td>FP = 1</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65–74 y</td>
<td>271</td>
<td>61*</td>
</tr>
<tr>
<td>55–64 y</td>
<td>672</td>
<td>110‡</td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>522</td>
<td>98</td>
</tr>
<tr>
<td>Former</td>
<td>421</td>
<td>73**</td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥60 pack-years</td>
<td>372</td>
<td>83‡‡</td>
</tr>
<tr>
<td>30–59 pack-years</td>
<td>571</td>
<td>88</td>
</tr>
</tbody>
</table>

CT = computed tomography; FP = false-positive; OR = odds ratio.  
* Probability of FP if aged 65–74 y = 0.18; crude OR = 1.38.  
† Probability of FP if aged 65–74 y = 0.09; crude OR = 1.95.  
‡ Probability of FP if aged 55–64 y = 0.14.  
§ Probability of FP if aged 55–64 y = 0.05.  
‖ Probability of FP if current smoker = 0.16; crude OR = 1.08.  
¶ Probability of FP if former smoker = 0.05; crude OR = 0.78.  
** Probability of FP if former smoker = 0.15.  
†† Probability of FP if 30–59 pack-years = 0.07.  
‡‡ Probability of FP if 30–59 pack-years = 0.13.  
§§ Probability of FP if 30–59 pack-years = 0.06.