CT Screening for Lung Cancer: Alternative Definitions of Positive Test Result Based on the National Lung Screening Trial and International Early Lung Cancer Action Program Databases

Rowena Yip, MPH
Claudia I. Henschke, PhD, MD
David F. Yankelevitz, MD
James P. Smith, MD

Purpose:
To determine the usefulness of alternative nodule size thresholds in a population undergoing computed tomographic (CT) screening for lung cancer and to compare the reported International Early Lung Cancer Action Program (I-ELCAP) results with the National Lung Screening Trial (NLST) results.

Materials and Methods:
The institutional review board approved this retrospective analysis. Informed consent was obtained according to HIPAA compliance. Findings in the CT cohort in the NLST of 25,813 participants who underwent baseline CT in 2002–2004 were reviewed. The frequency of solid and part-solid pulmonary nodules and the lung cancer diagnoses using an alternative nodule threshold of 5.0, 6.0, 7.0, 8.0, and 9.0 mm were determined. Proportional reduction in the frequency of positive results and their 95% confidence intervals using each of the alternative thresholds were calculated.

Results:
The frequency of positive results in the baseline round in the CT arm of the NLST using the definition of a positive result of any parenchymal, solid or part-solid, noncalcified nodule of 5.0 mm or larger was 15.8% (4080 of 25,813). Using alternative thresholds of 6.0, 7.0, 8.0, and 9.0 mm, the frequencies of positive results were 10.5% (2700 of 25,813), 7.2% (1847 of 25,813), 5.3% (1362 of 25,813), and 4.1% (1007 of 25,813), respectively, and the corresponding proportional reduction in additional CT scans would have been 33.8% (1380 of 4080), 54.7% (2233 of 4080), 66.6% (2718 of 4080), and 73.8% (3013 of 4080), respectively. Concomitantly, the proportion of lung cancer diagnoses determined within the first 12 months would be delayed up to 9 months for 0.9% (two of 232), 2.6% (six of 232), 6.0% (14 of 232), and 9.9% (23 of 232) of the patients, respectively.

Conclusion:
The NLST results are similar to those previously reported for the I-ELCAP and suggest that, even for high-risk participants in the NLST, higher thresholds of nodule size should be considered and prospectively evaluated.

1 From the Icahn School of Medicine at Mount Sinai, 1 Gustave Levy Place, New York, NY 10029 (R.Y., C.I.H., D.F.Y.); and Weill Cornell Medical College, New York, NY (J.P.S.).
Received January 13, 2014; revision requested March 19; final revision received March 26; accepted April 22; final version accepted April 29. Address correspondence to C.I.H. (e-mail: Claudia.Henschke@mountsinai.org).
© RSNA, 2014
In computed tomographic (CT) screening for lung cancer, the positive result of any given round of screening is the indication for an additional low-dose CT scanning work-up before the next scheduled annual repeat screening. Thus, the definition is important, as the goal is to minimize unnecessary work-ups while still finding lung cancer early. This is particularly important in the first, baseline, round of screening, which has the highest frequency of positive results, as prior CT scans are not available for comparison. If the definition is too inclusive, there is an excessive frequency of unnecessary diagnostic testing. However, if it is too restrictive, the additional diagnostic work-up of patients who ultimately receive a diagnosis of lung cancer may be delayed.

We have addressed the concern to identify the most appropriate nodule size criterion to define a positive result and provided the results of applying alternative higher thresholds of nodule size using the International Early Lung Cancer Action Program (I-ELCAP) database, which includes a broad spectrum of age and history of smoking of its participants (1,2). Nodule size is defined as the average of the maximum length and width on a single transaxial image, as this two-dimensional measure better reflects the three-dimensional tumor volume than a one-dimensional measure of length, as used in the National Lung Screening Trial (NLST). We showed that, by increasing the nodule size criterion from the currently used threshold of 5.0 mm to 6.0, 7.0, 8.0, and 9.0 mm, the proportional reduction in obtaining an additional CT scan of the chest would have been 36%, 56%, 68%, and 75%, respectively. Concomitantly, lung cancer diagnoses would have been delayed by up to 9 months for 0%, 5.0%, 5.9%, and 6.7%, respectively, of the patients in whom lung cancer would have been diagnosed prior to the first annual repeat screening (2). The 9-month delay is based on the maximum time between a 3-month follow-up CT, the current recommended work up for a noncalcified nodule (NCN) of 5.0 mm but smaller than 15 mm in average diameter, and the first annual repeat screening in 12 months, which would be recommended if the nodule size does not meet the current threshold criterion for a positive result.

Recently, we obtained the necessary data from the NLST on their high-risk participants. Therefore, the purpose of this study was to determine the implication of alternative nodule size thresholds in a population undergoing CT screening for lung cancer and to compare the reported I-ELCAP results with the NLST results.

Materials and Methods

The institutional review board of Icahn School of Medicine at Mount Sinai (New York, NY) approved this retrospective analysis. All participants enrolled in the NLST followed a multicenter research protocol approved by the institutional review board of all participating centers. Informed consent was obtained according to Health Insurance Portability and Accountability Act compliance. All protected health information and site identifiers were removed electronically.

This report draws from the database of all the baseline screenings in the CT arm of the NLST, which were performed during its enrollment period between 2002 and 2004. Participants were 55–74 years of age with at least a 30 pack-year history of smoking and if not currently smoking had quit within 15 years of enrollment. Further details of the NLST enrollment and screening process have been detailed in a prior publication (3). Among the

Advances in Knowledge

- The analysis of the National Lung Screening Trial (NLST) results suggests that, for high-risk participants in the NLST, higher thresholds of nodule size of 6.0, 7.0, 8.0, and 9.0 mm should be considered and prospectively evaluated.
- The use of a nodule size threshold that is higher than 5.0 mm can markedly decrease the frequency of positive results in the baseline round by up to 74% and, thus, of additional work-up without causing undue delay in the diagnosis of lung cancers.
- The definition of a positive result needs to be continually evaluated and updated in light of emerging evidence from ongoing screening programs to reduce unnecessary diagnostic chest CT work-ups.

Implications for Patient Care

- Reduction in the frequency of a positive result in the baseline round of screening would result in dramatic decreases in the frequency of obtaining an additional low-dose chest CT scan and any subsequent studies and their attendant costs and potential harms.
- If a chest CT scan obtained at 3 months is eliminated, the post-scanning visit (and cost of the visit) to the internist or pulmonologist and potential additional diagnostic tests would also be eliminated.
- Continual evaluation and updating of the definition of a positive result can minimize unnecessary work-ups while still finding lung cancer early.
26722 participants in the CT arm of the NLST who signed the consent form, a baseline CT scan (time T0) was actually obtained in 26309. Among these 26309 participants, a scan with a negative result (according to the data set designation, code: scr_res0 = 1, 2, 3) was obtained in 19118, and a scan with a positive result (based on the NLST criterion of at least one noncalcified nodule 4.0 mm or larger in diameter measured in length only, code: scr_res0 = 4) was obtained in 7191, but the nodule width was documented for all nodules (both with positive and negative results). Thus, the frequency of a positive result in the baseline round of the NLST was 27.3%. Each NCN in the NLST data set was also classified as to its consistency: solid (soft tissue), part-solid (mixed), and nonsolid (ground-glass). In 496 participants, either the nodule size (n = 279) or consistency (n = 210), or both (n = 7), were missing, and these participants were excluded from this proportional reduction analysis. In 16 cases of lung cancer, the tumor size documented by using the resected pathologic specimen was so large compared with the largest CT-documented nodule that the calculated volume doubling times were less than 30 days, suggesting either a CT measurement or documentation error and, thus, for our report, these cases were assigned an original nodule length greater than 9 mm. Also, in the NLST data set, different from I-ELCAP, the nodule in which the cancer was manifest was not specified, although the lobe in which the cancer was found was identified. It was therefore presumed that the cancer manifested in the largest CT-documented nodule. Nonetheless, in this analysis, as in the I-ELCAP (2), all lung cancers diagnosed in any lobe were included. Included then were 25813 NLST participants, among whom were 6695 (calculated as 7191 – 496) (25.9%) participants who had a positive result according to the NLST definition and documentation of both the nodule size (length and width) and consistency and the remaining 19118 (74.1%) participants who had a negative result of baseline screening.

The current definition of a positive result in the baseline round in I-ELCAP includes participants with at least one parenchymal solid or part-solid NCN; it does not include those with only nonsolid nodules. Of the 6695 NLST participants who had at least one NCN of 4.0 mm or more in longest diameter of any consistency, 848 had only nonsolid NCNs, and these were not considered to be a positive result according to the I-ELCAP criteria. Thus, 5848 (calculated as 6695 – 847) NLST participants had at least one part-solid or solid NCN diameter as small as 2.5 mm. Table 1 also gives the breakdown of the data in 4080 participants who met the I-ELCAP criterion of a nodule diameter of 5.0 mm or more in average diameter. Among these 4080 participants, the largest part-solid or solid NCN was 5.0 mm or more but less than 9.0 mm in average diameter in 3013, 9.0 mm or more but less than 15.0 mm in 707, and 15.0 mm or more in 360. Within these latter three size categories, 23, 85, and 124 participants, respectively, received a diagnosis of lung cancer within 12 months of the first scheduled annual repeat screening. We determined the proportional reduction in the frequency of positive results by using each of the alternative thresholds and the 95% confidence interval (CI) and the proportional delay and its 95% CI in the diagnosis of lung cancer among those who received a diagnosis prior to the first scheduled annual repeat screening. We determined the proportional reduction in the frequency of positive results by using each of the alternative thresholds and the 95% confidence interval (CI) and the proportional delay and its 95% CI in the diagnosis of lung cancer among those who received a diagnosis prior to the first scheduled annual repeat screening.

Results

Table 1 shows all 5848 NLST participants who had at least one part-solid or solid NCN of 4.0 mm in length (longest diameter). Among them, the largest part-solid or solid NCN was less than 5.0 mm in average diameter (average of length and width) in 1768 participants (0.11%) the diagnosis of lung cancer would have been delayed by 5 months in one and 6 months in 26 of them. Among these 26, 118, and a scan with a positive result (based on the NLST criterion of at least one noncalcified nodule 4.0 mm or larger in diameter measured in length only, code: scr_res0 = 4) was obtained in 7191, but the nodule width was documented for all nodules (both with positive and negative results). Thus, the frequency of a positive result in the baseline round of the NLST was 27.3%. Each NCN in the NLST data set was also classified as to its consistency: solid (soft tissue), part-solid (mixed), and nonsolid (ground-glass). In 496 participants, either the nodule size (n = 279) or consistency (n = 210), or both (n = 7), were missing, and these participants were excluded from this proportional reduction analysis. In 16 cases of lung cancer, the tumor size documented by using the resected pathologic specimen was so large compared with the largest CT-documented nodule that the calculated volume doubling times were less than 30 days, suggesting either a CT measurement or documentation error and, thus, for our report, these cases were assigned an original nodule length greater than 9 mm. Also, in the NLST data set, different from I-ELCAP, the nodule in which the cancer was manifest was not specified, although the lobe in which the cancer was found was identified. It was therefore presumed that the cancer manifested in the largest CT-documented nodule. Nonetheless, in this analysis, as in the I-ELCAP (2), all lung cancers diagnosed in any lobe were included. Included then were 25813 NLST participants, among whom were 6695 (calculated as 7191 – 496) (25.9%) participants who had a positive result according to the NLST definition and documentation of both the nodule size (length and width) and consistency and the remaining 19118 (74.1%) participants who had a negative result of baseline screening.

Table 1

Table 1 shows all 5848 NLST participants who had at least one part-solid or solid NCN of 4.0 mm in length (longest diameter). Among them, the largest part-solid or solid NCN was less than 5.0 mm in average diameter (average of length and width) in 1768 participants, two of them received a diagnosis of lung cancer before the first scheduled annual repeat screening. Of these, 411 participants had an average nodule diameter that was smaller than 4.0 mm, with some nodules having a diameter as small as 2.5 mm. Table 1 also gives the breakdown of the data in 4080 participants who met the I-ELCAP criterion of a nodule diameter of 5.0 mm or more in average diameter. Among these 4080 participants, the largest part-solid or solid NCN was 5.0 mm or more but less than 9.0 mm in average diameter in 3013, 9.0 mm or more but less than 15.0 mm in 707, and 15.0 mm or more in 360. Within these latter three size categories, 23, 85, and 124 participants, respectively, received a diagnosis of lung cancer within the baseline round, that is, before the first scheduled annual repeat screening.
When the higher average nodule diameter thresholds of 6.0, 7.0, 8.0, and 9.0 mm were applied to the Nlst data set, as had previously been done for I-ELCAP (2), the results, given in Table 2, were as follows: For the threshold of 6.0 mm, the result would be positive in 10.5% (2700 of 25813) (95% CI: 10.1%, 10.9%) of the participants, as compared with 15.8% using the 5.0-mm threshold. The use of this threshold would have caused the number of 3-month follow-up CT scans to be reduced by 33.8% (calculated as \(\frac{4080 - 2700}{4080}\) (95% CI: 32.4%, 35.3%). Similarly, using thresholds of 7.0, 8.0, and 9.0 mm, the corresponding frequency of a positive result would be 7.2% (1847 of 25813), 5.3% (1302 of 25813), and 4.1% (1067 of 25813), respectively (95% CI: 6.8%, 7.5%; 5.0%, 5.6%; and 3.9%, 4.4%, respectively). This would have caused the work-up in the baseline round of screening to be reduced by 54.7% (95% CI: 53.2%, 56.3%), 66.6% (95% CI: 65.2%, 68.1%), and 73.8% (95% CI: 72.5%, 75.2%), respectively.

Of the 4080 Nlst participants with a positive result based on the current I-ELCAP nodule diameter threshold of 5.0 mm, the additional CT scan of the chest and subsequent diagnostic work-up would have led to the diagnosis of lung cancer in 232 (5.7%) participants within the baseline round. Had the higher threshold of 6.0 mm been used, 2700 participants would have had a test result that met the definition of a positive test result and would have received a recommendation for low-dose CT of the chest (Table 2). This would have led to the diagnosis of lung cancer in 230 patients within the baseline round, two fewer than received a diagnosis using the 5.0-mm threshold, a reduction of 0.9% (95% CI: 0.1%, 3.1%). The use of nodule diameter thresholds of 7.0 mm, 8.0 mm, or 9.0 mm instead of 5.0 mm would have decreased the number of participants who received a diagnosis of lung cancer within 12 months of obtaining the baseline CT scan from 232 to 226, 218, and 209, respectively (Table 2), a reduction of 2.6% (95% CI: 1.0%, 5.5%), 6.0% (95% CI: 3.3%, 9.9%), and 9.9% (95% CI: 6.4%, 14.5%), respectively.

In a high-risk population, such as that of the Nlst, an increase in the average nodule diameter threshold from 5.0 mm to 6.0, 7.0, 8.0, and 9.0 mm would decrease the obtaining of an additional CT scan of the chest by 33.8%, 54.7%, 66.6%, and 73.8%, respectively, and result in a delay in the diagnosis of lung cancer of 0.9%, 2.6%, 6.0%, and 9.9% of patients who actually received a diagnosis of lung cancer during the baseline round of the Nlst CT arm.

### Discussion

In light of ongoing concerns about excessive so-called false-positive results reported in all CT screening programs for lung cancer, it has been suggested that the definition of a positive result should be reevaluated on an ongoing basis. To that end, we reported the results of using alternative definitions of a positive result, first using the Eearly Lung Cancer Action Program, or ELCAP, database (1) and second, recently, using the larger I-ELCAP database (2). As expected, we showed that the frequency of a positive result in the baseline round of screening could be markedly decreased by using larger nodule-size thresholds than currently in use, and these reductions would result in dramatic decreases in the frequency of obtaining an additional CT scan of the chest and performance of possible subsequent diagnostic tests and their attendant costs and potential harms. The trade-off would be limited to a delay, by up to 9 months, in the diagnosis of lung cancer in some patients who were also reported (2). Our purpose here was to present a similar analysis of the high-risk participants in the CT arm of the Nlst to determine the proportional reduction in the frequency of a positive result and, thus, a reduction in obtaining an additional CT scan of the chest that would have resulted had higher nodule thresholds been used. We also wanted to know the delays in cancer diagnoses that would have resulted in the Nlst participants and to compare those with the I-ELCAP published results (2).
The current definition of a positive result in the baseline round of I-ELCAP includes only solid and part-solid nodules. The I-ELCAP protocol recommends 1-year follow-up CT (i.e., the first annual repeat round) for all participants with only nonsolid nodules, regardless of size. Our rationale derives from observations from our database over the past 20 years, which has shown that nonsolid nodules may resolve, decrease in size, or grow very slowly (4–8) with long doubling times.

The results from the NLST presented here support those from I-ELCAP and suggest that the use of a nodule-size threshold that is higher than 5.0 mm can markedly decrease the frequency of a positive result in the baseline round and thus of additional work-up without causing undue delay in the diagnosis of lung cancers. This statement, of course, requires that individual participants return for their first annual repeat CT screening which, in turn, requires that an effective recall system be embedded in the screening process of all CT screening programs.

We had previously attempted to address the key question of whether the use of a higher threshold would lead to possible stage progression of the small cancerous nodules in patients who would no longer undergo 3-month follow-up CT but instead be referred to the first annual repeat screening, with the diagnosis thus being potentially delayed by 9 months. There remains no direct evidence to answer this critical question. A review of nodules that at baseline were smaller than the threshold of a positive result, or which were missed or not worked up, in earlier CT screening trials would be useful in providing additional information to address the issue. Alternatively, prospective research in an ongoing screening program by using a larger nodule-size threshold in a round of baseline screening with careful follow-up at the first annual repeat screening of all nodules that did not meet the size threshold could provide the most direct evidence about potential stage progress from delay in diagnosis. The National Comprehensive Cancer Center Network has recently taken such a step in the modification of its latest guidelines for CT screening that recommend a nodule-size threshold of 6.0 mm (9,10). We also believe that the NLST data suggest that the use of a nodule size of at least 5.0 mm (in average diameter) is preferable to any lower threshold for the definition of a positive CT result in the baseline round of screening.

Limitations of this report include that the results of using higher nodule-size thresholds for the definition of a positive result derive, necessarily, from retrospective reviews of available data. Again, no direct evidence is available as to whether the delay in lung cancer diagnoses due to an actual increase in the threshold of a positive test result would have led to stage progression in any of the NLST or I-ELCAP participants and thus changed the patients’ outcome. In the NLST, participants with nodules for which the length, width, or consistency were not provided were excluded, but assuming these were randomly distributed, the proportional reduction given different thresholds for the definition of a positive result would be essentially the same.

The key point of this report is that the definition of a positive result needs to be continually evaluated and updated in light of emerging evidence from ongoing screening programs to reduce unnecessary CT scans of the chest and possible further diagnostic workup, including surgery, with their attendant potential harms and costs while maximizing the early diagnosis and treatment of curable lung cancers.

Acknowledgments: The authors thank the National Cancer Institute (NCI) for access to the data of the NCI collected by the NLST. Data used in this article were downloaded from the Cancer Data Access System on approval of the NCI to access the NLST data sets, a public database. All protected health information and site identifiers were removed electronically by the NCI prior to the download. All participants enrolled in the NLST followed a multicenter research protocol approved by the institutional review board of all participating centers. The statements contained herein are solely those of the authors and do not represent or imply concurrence or endorsement by the NCI.

Disclosures of Conflicts of Interest: R.Y. disclosed no relevant relationships. C.L.H. activities related to the present article: disclosed no relevant activities. Other activities related to the present article: disclosed no relevant relationships. R.Y. disclosed relationships related to the present article: disclosed no relevant relationships. Other relationships: named inventor on a number of patents and patent applications relating to the evaluation of pulmonary nodules on CT scans of the chest which are owned by Cornell Research Foundation (CRF), and since 2005, author does not accept any financial benefit from these patents including royalties and any other proceeds related to the patents or patent.
applications owned by CRF; is president and serves on the board of the Early Diagnosis and Treatment Research Foundation but does not receive compensation from the Foundation (the Foundation provides grants for projects, conferences, and public databases for research on early diagnosis and treatment of diseases; recipients include International Early Lung Cancer Action Program, among others, with funding from a variety of sources including philanthropic donations, grants, and contracts with agencies [federal and nonfederal], imaging, and pharmaceutical companies relating to image processing assessments; the various sources of funding exclude any funding from tobacco companies or tobacco-related sources). D.E.Y. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received grants from Flight Attendants Medical Research Institute and AstraZeneca; serves on the scientific advisory board (unpaid) for Give-A-Scan, Lung Cancer Alliance; is a named inventor on a number of patents and patent applications relating to the evaluation of diseases of the chest including measurement of nodules, and some of these, which are owned by CRF, are nonexclusively licensed to GE Healthcare; as an inventor of these patents, author is entitled to a share of any compensation which CRF may receive from its commercialization of these patents. Other relationships: disclosed no relevant relationships. J.P.S. disclosed no relevant relationships.

References


