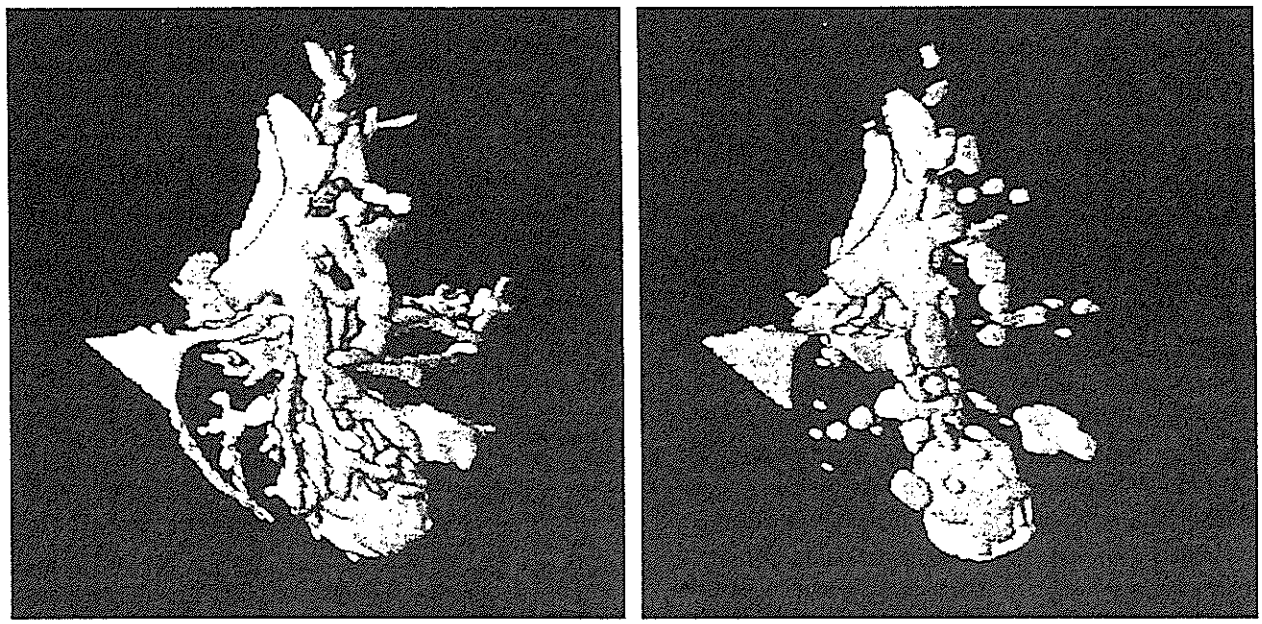


A. B.

Figure 4. **(a)** A single three-dimensional structure identified at a low gray-level threshold. **(b)** The three-dimensional structures that disassociate from the single structure in **a** as identified at a higher gray-level threshold. Many of these smaller structures will satisfy the volume criterion to become nodule candidates. (Reprinted from Armato et al.[44] with permission.)

the automated nodule detection method was evaluated through free-response receiver operating characteristic (FROC) curves (42), which are generated by plotting overall nodule detection sensitivity as a function of the number of false-positive structures identified per section. ROC and FROC curves were obtained by incrementally altering the operating point represented by the selected LDA output threshold. Both types of analysis are required because not all actual lung nodules in the database become members of the set of nodule candidates, a fact not captured by ROC analysis but reflected in FROC analysis.

Analysis of Complete Database

The automated lung nodule detection method was applied to the complete 393-case database of low-dose, lung cancer screening CT scans. For this analysis, all 470 lung nodules were the detection targets, regardless of nodule malignancy status, subtlety, or radiographic opacity. The initial set of nodule candidates generated by the method through the multiple gray-level thresholding stage included 415 of the 470 nodules (88.3%) contained within the database along with 211,813 non-nodule structures. Rule-based and linear discriminant classifiers were then applied to this initial set of nodule candidates to distinguish candidates that corresponded to actual nodules from those that corresponded to non-nodules (i.e., false positives). A jackknife training/testing approach was used in which the nodule candidates from 197 randomly chosen cases were used (1) to develop the thresholds for the rule-based classifier and (2) to establish the linear discriminant function. The rule-based and linear discriminant classifiers trained in this manner were applied to the candidates of the remaining 196 cases, and overall nodule detection results were obtained for these cases. The "training" and "testing" of the automated nodule detection method were thus performed separately on subsets of nodule candidates from independent patients.

Lung nodules demonstrate a spectrum of radiologic appearances (43). Accordingly, the performance of the automated detection method, as trained and tested on the complete 393-case database, was decomposed into performance on the basis of nodule malignancy status, size, subtlety, and radiographic opacity.

Analysis of Malignant Nodules

The methods reported in the previous section are based on the task of optimizing detection performance for all 470 lung nodules, regardless of nodule malignancy status,

size, subtlety, or radiographic opacity. A method that is trained and tested specifically on malignant nodules would be expected to yield improved performance on the more important clinical task of automated detection of *malignant* nodules. Among the 393 cases in the complete database, 66 cases contained at least one confirmed malignant nodule. These 66 cases contained all 69 malignant nodules in the database, along with 7 of the 401 benign nodules. When applied to this subset of cases, the initial set of nodule candidates generated by the method through the multiple gray-level thresholding stage included 66 of the 69 malignant nodules (95.7%) along with 35,447 non-nodule structures. Rule-based and linear discriminant classifiers were then applied to this initial set of nodule candidates to distinguish candidates that corresponded to actual malignant nodules from those that corresponded to non-nodules (i.e., false positives). A jackknife training/testing approach was used in which the nodule candidates from 33 randomly chosen malignant cases were used (1) to develop the thresholds for the rule-based classifier and (2) to establish the linear discriminant function. The rule-based and linear discriminant classifiers trained in this manner were applied to the candidates of the remaining 33 malignant cases, and detection results for malignant nodules were obtained for these cases. The "training" and "testing" of the automated nodule detection method were thus performed independently on independent subsets of nodule candidates from malignant cases. Detection of all malignant nodules was evaluated, and then this performance was decomposed into performance by radiographic opacity.

RESULTS

Complete Database

The overall performance of the method was obtained from the average performance of 10 random partitions of the nodule candidate set into training and testing subsets and is presented in the first row of Table 2. At an overall nodule detection sensitivity of 70% regardless of nodule malignancy status, size, subtlety, or radiographic opacity, a mean false-positive rate of 1.6 ± 0.2 per section was obtained across all 10 partitions. This sensitivity level was selected for consistency with our other published automated nodule detection studies, which did not report performance on the basis of malignancy status, size, subtlety, or radiographic opacity.

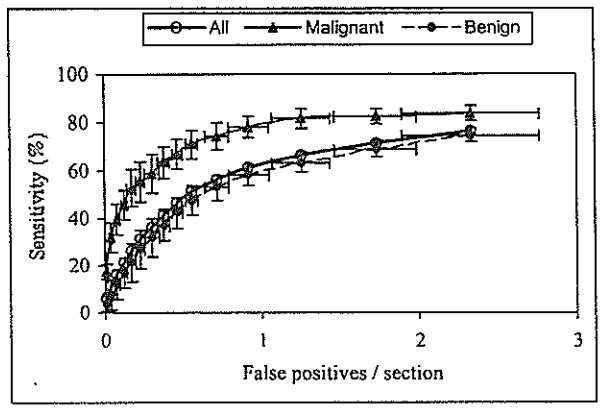
Table 2
Overall Performance of the Automated Nodule Detection Method Obtained From the Average of 10 Jackknife Analyses of the Entire 393-Case Database With a Mean False-Positive Rate of 1.6 per Section

Nodule Subset	Nodule Detection Sensitivity (%)
All nodules	70
Malignant	83 ± 5.0
Benign	68 ± 0.9
Not subtle	91 ± 2.5
Subtle	73 ± 2.1
Very subtle	52 ± 2.9
Solid	76 ± 2.0
Part solid	78 ± 2.5
Nonsolid	51 ± 3.7
Effective diameter >5 mm	74 ± 2.7

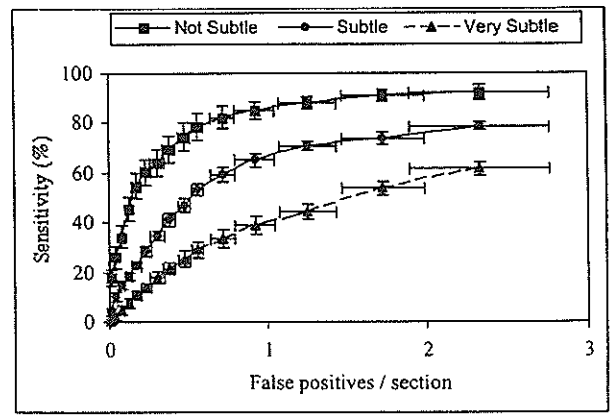
Reported sensitivities are at a mean false-positive rate of 1.6 per section. For this analysis, the method was trained to detect all nodules regardless of malignancy status, size, subtley, or radiographic opacity.

Table 2 decomposes this *overall* performance into performance by nodule malignancy status, size, subtlety, and radiographic opacity. Note that the automated detection method does not *classify* nodules on the basis of malignancy status, size, subtlety, or radiographic opacity, but rather, in this analysis, the ability of the method to *detect* nodules within these various categories was evaluated. For example, given that 70% of all nodules in the test sets were detected, 83% ± 5.0% of the malignant nodules among the nodules of the test sets were detected (at the same false-positive rate of 1.6 per section), whereas only 68% ± 0.9% of the benign nodules among the nodules of the test sets were detected. As expected, nodule detection sensitivity increased as nodule subtlety decreased, with a detection sensitivity of 91% ± 2.5% for nodules rated as "not subtle" and a detection sensitivity of only 52% ± 2.9% for "very subtle" nodules. The detection sensitivity for solid and part-solid nodules were comparable (76% ± 2.0% and 78% ± 2.5%, respectively), whereas the detection sensitivity for nonsolid nodules was substantially lower at 51% ± 3.7%. The detection sensitivity for nodules with an effective diameter greater than 5 mm was 74% ± 2.7% at the same false-positive rate.

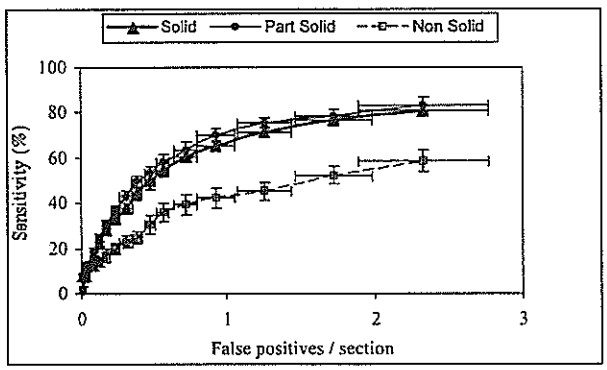
Figure 5 presents FROC curves that represent automated nodule detection sensitivity as a function of false-positive detections per section obtained from the average of 10 jackknife analyses of the entire 393-case database as described above. FROC curves are shown for (a) all



a.



b.



c.

Figure 5. FROC curves representing nodule detection sensitivity as a function of false-positive detections per section obtained from the average of 10 jackknife analyses of the entire 393-case database. For this analysis, the method was trained to detect all nodules regardless of malignancy status, size, subtlety, or radiographic opacity. FROC curves are shown for (a) all nodules and nodules divided into malignancy status, (b) nodules divided into subtlety, and (c) nodules divided into radiographic opacity. The data in Table 2 correspond to a single operating point from each of these nine curves.

nodules (which this analysis was intended to detect) and nodules divided into malignancy status, (b) nodules divided into subtlety, and (c) nodules divided into radiographic opacity. The data in Table 2 correspond to a single operating point from each of these nine curves. These curves illustrate the disparity in the ability of the method (when trained to detect all nodules regardless of malignancy status, size, subtlety, or radiographic opacity) to detect malignant and benign nodules; not subtle, subtle, and very subtle nodules; and solid or part-solid nodules and nonsolid nodules. The ability of the method to detect nodules that have been diagnosed as malignant is better than its ability to detect the benign nodules in the database (Fig. 5a), the ability of the method to detect nodules in the database decreases as nodule subtlety increases (Fig. 5b), and whereas the ability of the method to detect solid and part-solid nodules is comparable, its ability to detect nonsolid nodules is lower (Fig. 5c).

Scans with Malignant Nodules

The results reported in the previous section were based on the task of optimizing detection performance for all 470 lung nodules, regardless of nodule malignancy status, size, subtlety, or radiographic opacity. These results exhibited an interesting differentiation: 83% of the malignant nodules were detected, whereas only 68% of the benign nodules were detected. The preponderance of benign nodules in the database (401 of 470, or 85%, of the nodules were benign), combined with this malignancy status–based differential in detection performance, then yielded an overall nodule detection performance of 70% with an average of 1.6 false-positives per section. With

Table 3
Performance of the Automated Nodule Detection Method Specifically for Malignant Nodules Obtained From the Average of 20 Jackknife Analyses of the 66 Cases That Contained at Least One Malignant Nodule

Malignant Nodule Subset	Nodule Detection Sensitivity (%)
All malignant nodules	80
Solid	90 ± 7
Part solid	80 ± 3
Nonsolid	56 ± 9
Not subtle	94 ± 5
Subtle	78 ± 6
Very subtle	46 ± 10

Reported sensitivities are at a mean false-positive rate of 0.85 per section. For this analysis, the method was trained to detect malignant nodules regardless of size, subtlety, or radiographic opacity.

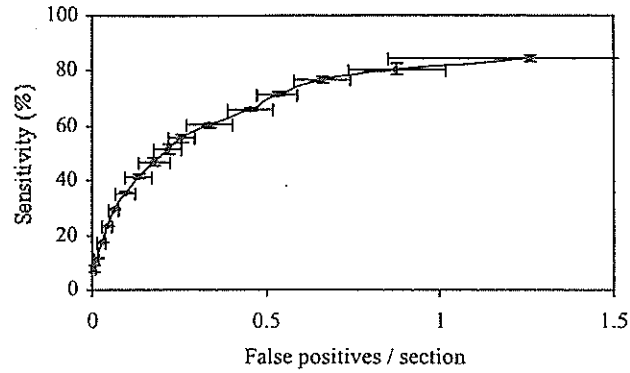


Figure 6. FROC curve representing *malignant* nodule detection sensitivity as a function of false-positive detections per section obtained from the average of 20 jackknife analyses of the 66 cases with at least one malignant nodule. For this analysis, the method was trained to detect all malignant nodules regardless of size, subtlety, or radiographic opacity.

the method trained and tested specifically on malignant nodules, improved performance was observed for the automated detection of malignant nodules.

The overall performance of the method for the detection of malignant nodules was obtained from the average performance of 20 random partitions of the nodule candidate set (obtained from the 66 cases with at least one malignant nodule) into training and testing subsets and is presented in the first row of Table 3. At a malignant nodule detection sensitivity of 80% ± 2.0%, a mean false-positive rate of 0.85 ± 0.14 per section was obtained across all 20 partitions. Figure 6 shows the corresponding FROC curve. This performance may be decomposed into performance by nodule radiographic opacity so that, on average, 90% ± 7% of the solid malignant nodules were detected, 80% ± 3% of the part-solid malignant nodules were detected, and 56% ± 9% of the nonsolid malignant nodules were detected at the same mean false-positive rate of 0.85 per section (Table 3).

DISCUSSION

The automated method presented in this study was developed to identify any focal abnormality that generally may be regarded as a lung nodule. While lung cancers represent the most clinically important subset of such focal abnormalities, the automated identification of any focal abnormality for radiologists' consideration may be useful. In the clinical practice of the future, automated nodule *detection* would likely be integrated with automated *classification* to distinguish between malignant and benign detected nodules as part of a comprehensive CAD system (36).

A limitation of the results obtained from this study is the 10-mm collimation of the CT scans that comprised the database. Such thick sections hinder the ability of the method to more accurately capture the three-dimensional contiguity of anatomic and pathologic structures. Moreover, the structures themselves are more sharply represented with state-of-the-art scanners that generate images with submillimeter axial resolution. We believe that thinner-section CT scans inherently will provide for improved performance of our method. A database of lung nodule cases from such scanners that is as well characterized as the database used in this study is certainly required to establish the clinical utility of these methods.

Despite the suboptimal nature of thick-section CT scans, a key observation from this study is that, as expected, the ability of a single automated method to detect the complete spectrum of lung nodules may be limited. Consequently, the importance of this study is the *relative* performance of the method in the detection of lung nodules in the various malignancy status, size, subtlety, and radiographic opacity categories. Based on this analysis, the automated method requires modification to improve the detection sensitivity of those lesions most likely to be missed by radiologists, namely, very subtle nodules and nonsolid nodules. The category-based performance analysis we present serves to underscore the importance of a full characterization of the lung nodules used by investigators when reporting results of CAD methods.

The disparity between the ability of the method to detect malignant and benign nodules is substantial. Despite the argument that it could be advantageous for an automated nodule detection method to detect as many nodules (both malignant and benign) as possible while leaving the diagnostic task to the radiologist (with the eventual assistance of another automated method for nodule classification), an automated nodule detection method that preferentially detects malignant nodules is expected to provide additional clinical benefits. As may be observed from Figure 1, the effective diameters of malignant nodules tended to be larger than the effective diameters of benign nodules (a t test for difference in means between the two distributions of effective diameters yielded a significant difference with $P < 0.01$); accordingly, the effect on method performance of nodule size versus the radiologic characteristics specific to malignant nodules remains for further investigation.

It is worth noting that the overall performance of the automated detection method applied to low-dose CT scans is consistent with the overall performance of the detection method applied to a database of diagnostic CT scans (71% sensitivity with 1.5 false-positives per section) reported earlier (24). Although the rule-based and linear discriminant classifiers were established separately for the diagnostic and low-dose databases, the consistent levels of performance demonstrate robustness of the general methodology.

As expected, the overall performance of the automated lung nodule detection method decreased for nodules of increased subtlety and for nodules with an increased nonsolid component. The clinical utility of automated nodule detection methods, of course, will depend on the ability of such methods to detect those nodules most likely to be overlooked or misinterpreted by radiologists. The goal of ongoing research is to reduce the number of false-positive detections, which at present limit the clinical utility of the method, and to reduce the variability of the method across different nodule categories.

CONCLUSION

We have developed an automated lung nodule detection method that we evaluated with a large number of low-dose CT scans from a lung cancer screening program. This evaluation used a jackknife paradigm for the training and testing of the automated classifiers that distinguished nodule candidates that corresponded to actual nodules from those that corresponded to non-nodules. Because nodules demonstrate a spectrum of radiologic appearances, the performance of the automated method was evaluated on the basis of nodule malignancy status, size, subtlety, and radiographic opacity. The category-based performance analysis we present serves to underscore the importance of a full characterization of the lung nodules used by investigators when reporting results of CAD methods. Such computerized lung nodule detection methods are expected to become important parts of CT-based lung cancer screening programs (44).

ACKNOWLEDGMENTS

The authors would like to thank Maryellen L. Giger, PhD, for insightful discussions and Roger Engelmann, MS, for development of the interface used to identify the location of actual nodules and for development of the relational database used to organize the CT images. S. G. Armato, H. MacMahon, and K. Doi are shareholders in R2 Technology, Inc. (Sunnyvale, CA). K. Doi is a shareholder in Deus Technologies, Inc. (Rockville, MD).

REFERENCES

1. Miettinen OS, Henschke CI. CT screening for lung cancer: coping with nihilistic recommendations. Radiology 2001; 221:592–596.
2. Patz EF Jr, Black WC, Goodman PC. CT screening for lung cancer: not ready for routine practice. Radiology 2001; 221:587–591.
3. Melamed MR, Flehinger BJ, Zaman MB, et al. Screening for early lung cancer: results of the Memorial Sloan-Kettering study in New York. Chest 1984; 86:44–53.
4. Fontana RS, Sanderson DR, Woolner LB, et al. Lung cancer screening: the Mayo program. J Occup Med 1986; 28:746–750.
5. Tockman MS. Survival and mortality from lung cancer in a screened population: the Johns Hopkins Study. Chest 1986; 89:3245–3255.
6. Fontana RS, Sanderson DR, Woolner LB, et al. Screening for lung cancer: a critique of the Mayo Lung Project. Cancer 1991; 67:1155–1164.
7. Eddy DM. Screening for lung cancer. Ann Intern Med 1989; 111:232–237.
8. Potchen EJ, Austin JHM. Problems and pitfalls in the diagnosis of early lung cancer. In: Potchen EJ, Grainger RG, Greene R, eds. Pulmonary radiology. Philadelphia, PA: WB Saunders, 1993:315–328.
9. Strauss GM, Gleason RE, Sugarbaker DJ. Screening for lung cancer: another look; a different view. Chest 1997; 111:754–768.
10. Flehinger BJ, Kimmel M, Polyak T, et al. Screening for lung cancer: the Mayo Lung Project revisited. Cancer 1993; 72:1573–1580.
11. Sone S, Takashima S, Li F, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 1998; 351:1242–1245.
12. Kaneko M, Eguchi K, Ohmatsu H, et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. Radiology 1996; 201:798–802.
13. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999; 354:99–105.

14. Rusinek H, Naidich DP, McGuinness G, et al. Pulmonary nodule detection: low-dose versus conventional CT. Radiology 1998; 209:243–249.
15. Diederich S, Lenzen H, Windmann R, et al. Pulmonary nodules: experimental and clinical studies at low-dose CT. Radiology 1999; 213:289–298.
16. Swensen SJ, Jett JR, Hartman TE, et al. Lung cancer screening with CT: Mayo Clinic experience. Radiology 2003; 226:756–761.
17. Giger ML, Bae KT, MacMahon H. Computerized detection of pulmonary nodules in computed tomography images. Invest Radiol 1994; 29: 459–465.
18. Kanazawa K, Kubo M, Niki N, et al. Computer assisted lung cancer diagnosis based on helical images. In: Chin RT, Ip HHS, Naiman AC, Pong T-C, eds. Image analysis applications and computer graphics. Proceedings of the Third International Computer Science Conference. Berlin: Springer-Verlag, 1995:323–330.
19. Ryan WJ, Reed JE, Swensen SJ, et al. Automatic detection of pulmonary nodules in CT. In: Lemke HU, Vannier MW, Inamura K, Farman AG, eds. Computer assisted radiology (CARS '96). Amsterdam: Elsevier Science, 1996:385–389.
20. Okumura T, Miwa T, Kako J, et al. Image processing for computer-aided diagnosis of lung cancer screening system by CT (LSCT). Proc SPIE 1998; 3338:1314–1322.
21. Lou S-L, Chang C-L, Lin K-P, et al. Object-based deformation technique for 3-D CT lung nodule detection. Proc SPIE 1999; 3661:1544–1552.
22. Satoh H, Ukai Y, Niki N, et al. Computer aided diagnosis system for lung cancer based on retrospective helical CT image. Proc SPIE 1999; 3661:1324–1335.
23. Taguchi H, Kawata Y, Niki N, et al. Lung cancer detection based on helical CT images using curved surface morphology analysis. Proc SPIE 1999; 3661:1307–1314.
24. Armato SG III, Giger ML, MacMahon H. Automated detection of lung nodules in CT scans: preliminary results. Med Phys 2001; 28:1552–1561.
25. Brown MS, McNitt-Gray MF, Goldin JG, et al. Patient-specific models for lung nodule detection and surveillance in CT images. IEEE Trans Med Imaging 2001; 20:1242–1250.
26. Fan L, Novak CL, Qian J, et al. Automatic detection of lung nodules from multi-slice low-dose CT images. Proc SPIE 2001; 4322:1828–1835.
27. Fiebich M, Wietholt C, Renger BC, et al. Automatic detection of pulmonary nodules in low-dose screening thoracic CT examinations. Proc SPIE 1999; 3661:1434–1439.
28. Gurcan MN, Sahiner B, Petrick N, et al. Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system. Med Phys 2002; 29:2552–2558.
29. Ko JP, Betke M, Chest CT. Automated nodule detection and assessment of change over time—preliminary experience. Radiology 2001; 218:267–273.
30. Lee Y, Hara T, Fujita H, et al. Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. IEEE Trans Med Imaging 2001; 20:595–604.
31. Suzuki K, Armato SG III, Li F, et al. Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography. Med Phys 2003; 30:1602–1617.
32. Li Q, Sone S, Doi K. Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans. Med Phys 2003; 30:2040–2051.
33. Armato SG III, Altman MB, LaRivière PJ. Automated detection of lung nodules in CT scans: effect of image reconstruction algorithm. Med Phys 2003; 30:461–472.
34. Armato SG III, Li F, Giger ML, et al. Lung cancer: performance of automated lung nodule detection applied to cancers missed in a CT screening program. Radiology 2002; 225:685–692.
35. Sone S, Li F, Yang Z-G, et al. Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. Br J Cancer 2001; 84:25–32.
36. Armato SG III, Altman MB, Wilkie J, et al. Automated lung nodule classification following automated nodule detection on CT: A serial approach. Med Phys 2003; 30:1188–1197.
37. Austin JHM, Mueller NL, Friedman PJ, et al. Glossary of terms for CT of the lungs: Recommendations of the Nomenclature Committee of the Fleischner Society. Radiology 1996; 200:327–331.
38. Johnson RA, Wichern DW. Applied multivariate statistical analysis. Englewood Cliffs, NJ: Prentice Hall, 1992.
39. Armato SG III, Giger ML, MacMahon H. Analysis of a three-dimensional lung nodule detection method for thoracic CT scans. Proc SPIE 2000; 3979:103–109.
40. Metz CE. ROC methodology in radiologic imaging. Invest Radiol 1986; 21:720–733.
41. Metz CE, Herman BA, Shen J-H. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. Stat Med 1998; 17:1033–1053.
42. Bunch PC, Hamilton JF, Sanderson GK, et al. A free response approach to the measurement and characterization of radiographic observer performance. Proc SPIE 1977; 127:124–135.
43. Armato SG III, McLennan G, McNitt-Gray MF, et al. Lung Image Database Consortium: developing a resource for the medical imaging research community. Radiology 2004; 232:739–748.
44. Armato SG III, Giger ML, Blackburn JT, et al. Three-dimensional approach to lung nodule detection in helical CT. Proc SPIE 1999; 3661: 553–559.

Feng Li, MD, PhD
Hidetaka Arimura, PhD
Kenji Suzuki, PhD
Junji Shiraishi, PhD
Qiang Li, PhD
Hiroyuki Abe, MD, PhD
Roger Engelmann, MS
Shusuke Sone, MD, PhD
Heber MacMahon, MD
Kunio Doi, PhD

Published online
10.1148/radiol.2372041555
Radiology 2005; 237:684–690

Abbreviations:
A_z = area under ROC curve
CAD = computer-aided detection
GGO = ground-glass opacity
LROC = localization ROC
ROC = receiver operating characteristic

[1] From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, MC-2026, University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637 (F.L., K.S., J.S., Q.L., H.A., R.E., H.M., K.D.); Department of Health Sciences, Faculty of Medicine, Kyushu University, Fukuoka, Japan (H.A.); and J. A. Azumi General Hospital, Nagano, Japan (S.S.). From the 2003 RSNA Annual Meeting. Received September 8, 2004; revision requested October 29; revision received December 27; accepted January 21, 2005. Supported in part by U.S. Public Health Service grants CA62625 and CA98119. Address correspondence to F.L. (e-mail: *feng@uchicago.edu*).

See Materials and Methods for pertinent disclosures.

Author contributions:
Guarantors of integrity of entire study, F.L., K.D.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, F.L., K.D.; clinical studies, F.L., H.A., S.S.; statistical analysis, F.L., K.S., J.S., Q.L., K.D.; and manuscript editing, F.L., H.A., K.S., Q.L., R.E., H.M., K.D.

Computer-aided Detection of Peripheral Lung Cancers Missed at CT: ROC Analyses without and with Localization[1]

PURPOSE: To retrospectively evaluate whether a difference-image computer-aided detection (CAD) scheme can help radiologists detect peripheral lung cancers missed at low-dose computed tomography (CT).

MATERIALS AND METHODS: Institutional review board approval and informed patient and observer consent were obtained. Seventeen patients (eight men and nine women; mean age, 60 years) with a missed peripheral lung cancer and 10 control subjects (five men and five women; mean age, 63 years) without cancer at low-dose CT were included in an observer study. Fourteen radiologists were divided into two groups on the basis of different image display formats. Six radiologists (group 1) reviewed CT scans with a multiformat display, and eight radiologists (group 2) reviewed images with a "stacked" cine-mode display. The radiologists, first without and then with the CAD scheme, indicated their confidence level regarding the presence (or absence) of cancer and the most likely position of a lesion on each CT scan. Receiver operating characteristic (ROC) curves were calculated without and with localization to evaluate the observers' performance.

RESULTS: With the CAD scheme, the average area under the ROC curve improved from 0.763 to 0.854 for all radiologists ($P = .002$), from 0.757 to 0.862 for group 1 ($P = .04$), and from 0.768 to 0.848 for group 2 ($P = .01$). The average sensitivity in the detection of 17 cancers improved from 52% (124 of 238 observations) to 68% (163 of 238 observations) for all radiologists ($P < .001$), from 49% (50 of 102 observations) to 71% (72 of 102 observations) for group 1 ($P = .02$), and from 54% (74 of 136 observations) to 67% (91 of 136 observations) for group 2 ($P = .006$). The localization ROC curve also improved.

CONCLUSION: Lung cancers missed at low-dose CT were very difficult to detect, even in an observer study. The use of CAD, however, can improve radiologists' performance in the detection of these subtle cancers.

In the past decade, low-dose single–detector row computed tomography (CT) with 10-mm-thick sections has been used to screen asymptomatic smoking and nonsmoking populations for lung cancer (1–3). The results of these CT screening studies showed that early peripheral lung cancers usually appeared as solitary noncalcified lesions with or without areas of ground-glass opacity (GGO), and the detection of these lesions at an early stage was greatly improved with use of low-dose CT rather than chest radiography. Most CT scans obtained in screening programs, however, showed only minor benign abnormalities, including noncancerous abnormalities such as diffuse lung disease (emphysema and interstitial changes) and focal lung disease (active infections, scars, and calcified nodules); in addition, 5%–27% of patients had noncalcified nodules that were detected at baseline screening with use of low-dose CT and 10-mm-thick sections (1–3). When reading images obtained in a CT screening program, radiologists must search for suspicious

noncalcified lung nodules, differentiate these lesions from benign nodules and lung cancer, and, finally, recommend follow-up actions for the detected lesions. In the studies mentioned above, no additional clinical information was provided to radiologists reviewing the CT scans except for age, sex, and smoking status.

At baseline CT screening performed in a general population that included smokers and nonsmokers in Nagano, Japan(2), the fraction of lung cancers among the detected noncalcified lesions was 9% and the prevalence of cancers was only 0.48%. The corresponding data were 12% and 2.7%, respectively, for smokers in the U.S. Early Lung Cancer Action Project (3). In CT screening programs, however, 32%–39% of lung cancers (4,5) were missed in previous years, and the numbers of these missed cancers were not included in the determination of the prevalence of lung cancers in these studies. We previously reported (5) that 32 missed lung cancers were very difficult to detect in the Nagano series; in general, they were very subtle and appeared as small, faint nodules with GGO that overlapped normal structures or as opacities in a complex background of other diseases.

When an automated lung nodule–detection method (6) was used, 84% of these missed lung cancers in the Nagano series were marked by the computer; however, the false-positive rate was high (1.0 false-positive marks per section, 28 false-positive marks per study), and this is not acceptable to radiologists. Recently, we developed a computer-aided detection (CAD) scheme (7) that is based on a difference-image technique for enhancing lung cancers and suppressing most normal background structures, and the false-positive rate has improved to about 3.0 marks per study (sensitivity, 87%) with use of a multiple massive training artificial neural network (8). Thus, the purpose of our study was to retrospectively evaluate whether a difference-image CAD scheme can help radiologists detect peripheral lung cancers missed at low-dose CT.

MATERIALS AND METHODS

H.M. and K.D. are shareholders in R2 Technology, Sunnyvale, Calif. K.D. is a shareholder in Deus Technology, Rockville, Md. CAD technologies developed in the Kurt Rossmann Laboratories have been licensed to companies including R2 Technology, Deus Technologies, Riverain Medical Group, Mitsubishi Space Software, Median Technologies, GE, and Toshiba.

Database

An annual low-dose CT screening program for lung cancer in Nagano, Japan, began in May 1996 and ended in March 1999. In the program, 17 892 examinations were performed in 7847 individuals (4288 men, 3559 women; mean age, 61 years; age range, 19–92 years). All individuals gave informed consent to undergo CT screening and for use of the data for research purposes. The database used in this study consisted of data from 38 low-dose CT examinations performed in 31 patients with missed peripheral lung cancers. All of the CT studies had been performed as part of the 3-year lung cancer screening program (5,6). Twenty-three cancers were missed because of detection errors, and 15 cancers were missed because of interpretation errors.

As described previously (5), the locations of missed lung cancers on sections obtained at 39 CT examinations (one examination was excluded from this study because of technical error) were determined in consensus by two radiologists (F.L. and S.S., with 20 and 42 years of experience, respectively). One radiologist (F.L.) measured the length and width of cancers on at least one section. Three radiologists (F.L., H.A., and H.M., with 20, 18, and 29 years of experience, respectively) first independently classified the low-dose CT scans with the 38 cancers into three patterns, and the final judgment was based on agreement by at least two radiologists. The mean diameter of the 38 lesions missed at low-dose CT was 12 mm (range, 6–26 mm). The following patterns were noted: 10 nodules had pure GGO (nonsolid), 16 had mixed GGO (part solid), and 12 had solid opacity.

The 31 missed cancers, which included 28 adenocarcinomas, two small cell carcinomas, and one squamous cell carcinoma, were confirmed with surgery. The CT examinations were performed with a mobile scanner (CT-W950SR; Hitachi Medical, Tokyo, Japan) with use of a low-dose protocol and a tube current of 25 or 50 mA, a scanning time of 2 seconds per rotation of the x-ray tube (tube rotation time, 2 seconds), a table speed of 10 mm/sec (pitch, 2), 10-mm collimation, and a 10-mm reconstruction interval. The mean number of sections per study was 30, and the pixel size was 0.586 or 0.684 mm for scans with a 512 × 512 image matrix size. The use of this database and the participation of radiologists in this observer performance study were approved by the University of Chicago Institutional Review Board. Informed consent for the observer performance study was obtained from all observers.

CAD Scheme

Our scheme was based on a difference-image technique (7,9,10) that enhances the lung nodules and suppresses most of the background normal structures. The difference image for each CT study was obtained by subtracting the nodule-suppressed image processed with a ring average filter from the nodule-enhanced image processed with a matched filter. By applying a multiple-gray-level threshold technique to the difference image, on which most nodules showed strong enhancement, the initial nodule candidates were identified. A number of false-positive findings were removed by using the two rule-based schemes on the localized image features related to morphologic characteristics and gray levels, and a false-positive rate of 15.8 per study was achieved (7). Most (81%) of the remaining false-positive findings were eliminated without removing any true-positive findings by using a multiple massive training artificial neural network trained to reduce various types of false-positive findings (8). The CAD scheme had a sensitivity of 87% (33 of 38 cancers) for 38 missed cancers, with an average of 3.0 false-positive findings per study (7,8).

Observer Study

Among 23 studies in which cancer was missed due to detection errors, 17 studies in 17 patients (eight men and nine women; mean age, 60 years; age range, 48–69 years) were performed the year before the cancers were found; the other six studies, including three that were performed in the same 17 patients 2 years before the cancers were found and three that revealed a coexisting benign nodule (diameter, 4–5 mm), were not used in this investigation. All 17 cancers were adenocarcinomas. At low-dose CT, six nodules had pure GGO, 10 nodules had mixed GGO, and one nodule had solid opacity. The mean diameter of the 17 missed cancers was 10 mm (range, 6–17 mm). Fifteen studies in which cancer was missed due to interpretation errors were also excluded from the observer study. In addition, we included studies obtained in 10 control subjects (five men and five women;

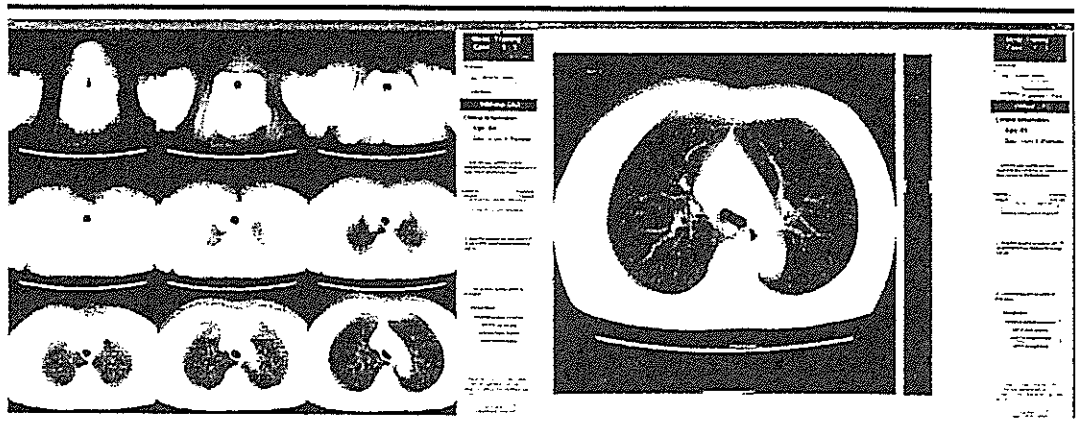


Figure 1. (a) Example of the multiformat display used by the six radiologists in group 1. From the top to the bottom of the entire lung for each patient, 27 consecutive transverse CT sections were displayed in a multiformat (3 × 3) mode on three high-spatial-resolution monitors. (b) Stacked cine-mode display used by the eight radiologists in group 2. Magnified and stacked transverse CT sections were displayed on one monitor.

TABLE 1
A_z Values for 14 Radiologists in Detection of Missed Cancers without and with CAD Scheme

Group and Observer	A_z Value	
	Without CAD	With CAD
Group 1*		
1	0.856	0.962
2	0.792	0.825
3	0.613	0.824
4	0.723	0.865
5	0.851	0.876
6	0.706	0.818
Group 2†		
7	0.589	0.828
8	0.723	0.834
9	0.818	0.824
10	0.865	0.936
11	0.811	0.899
12	0.826	0.856
13	0.786	0.822
14	0.728	0.784
All observers	0.763	0.854

Note.—The difference in A_z values without and with the CAD scheme was statistically significant, with a P value of .002 for all radiologists, .04 for group 1, and .01 for group 2. No statistically significant difference in A_z values between the two viewing modes was found for observers without and with the CAD scheme.

* This group used a multiformat display. The mean A_z value for this group without CAD was 0.757; the mean value with CAD was 0.862.

† This group used a cine-mode display. The mean A_z value for this group without CAD was 0.768; the mean value with CAD was 0.848.

mean age, 63 years; age range, 49–69 years) without cancer who had participated in the same screening program and whose ages and sexes closely matched those of the patient group; findings in these subjects were confirmed with 2-year follow-up. Some of the 27 studies revealed other abnormal findings such as scars, focal interstitial lung lesions, and small (<3 mm) benign nodules. The CAD scheme had a sensitivity of 82% (14 of 17 cancers), with 3.0 false-positive findings per study (range, zero to eight) for patients with missed cancers and 2.4 false-positive findings per study (range, zero to five) for the 10 control subjects (7,8).

Two image display formats were used in this investigation: a multiformat display and a "stacked" cine-mode display (Fig 1). For the multiformat display, from the top to the bottom of the entire lung for each patient, 27 consecutive sections with the original matrix size at low-dose CT were displayed in a multiformat display (3 × 3) on three high-spatial-resolution (1600 × 1200 pixels) liquid crystal display color monitors (CCL202; Totoku Electric, Tokyo, Japan). For cine-mode display, the same 27 CT sections for each study were magnified and stacked on one monitor. The speed or sequence of the image display for cine-mode display was controlled manually by the observer. The windowing in the two image display formats was initially set at lung settings but could be adjusted by the observer to bronchial or mediastinal settings. Two clinical parameters (age and sex) were provided to the observer on the monitor.

The 14 radiologists who participated in this observer study were classified into two groups according to type of display. Observers who used multiformat display (group 1) consisted of five general radiologists with 7–18 years of experience (mean, 12 years) and one 3rd-year radiology resident. Observers who used cine-mode display (group 2) consisted of three chest radiologists with 16, 17, and 45 years of experience (mean, 26 years), four general radiologists with 5–16 years of experience (mean, 13 years), and one 4th-year radiology resident. The observers in group 2 had more experience than did the observers in group 1.

Radiologists were given the following instructions: "*(a)* We wish to evaluate radiologists' performance in detecting lung cancer without and with a CAD scheme on low-dose CT scans obtained from a screening program. *(b)* The role of the CAD output is that of a 'second opinion.' *(c)* Twenty-seven low-dose CT studies (with 10-mm-thick sections) that did not or did contain lung cancer and/or noncancerous abnormalities such as benign nodules and scars are included in this observer study. *(d)* The observer in this study will be blinded to the number of patients with lung cancer and the performance level of the CAD scheme. *(e)* Click on the screen by using a mouse (i) to indicate on a bar your confidence level regarding the presence (or absence) of a lung cancer and (ii) to locate the most likely position on each CT scan. You may indicate the cancer location first and *(f)* click on one of the following four clinical actions: (i) Return to annual screening, (ii) diagnostic thin-section CT in 6 months, (iii) diagnostic thin-section CT in 3 months, or (iv) diagnostic thin-section CT immediately." The radiologists made their judgments first without and then with the CAD scheme.

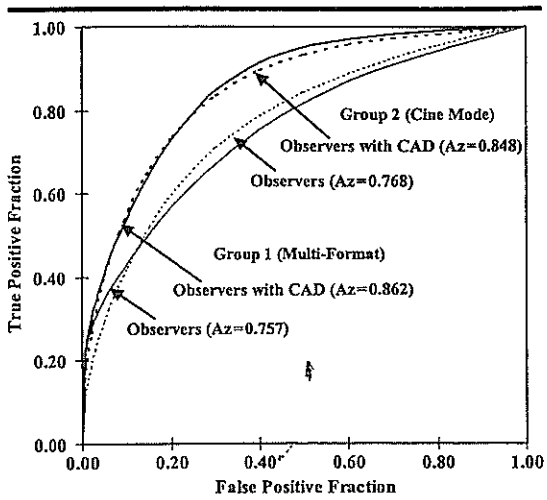


Figure 2. Graph shows ROC curves for detecting cancers missed at CT without and with use of the CAD scheme and for the two display modes. With the CAD scheme, the average A_z value improved significantly from 0.757 to 0.862 for group 1 (P = .04) and from 0.768 to 0.848 for group 2 (P = .01).

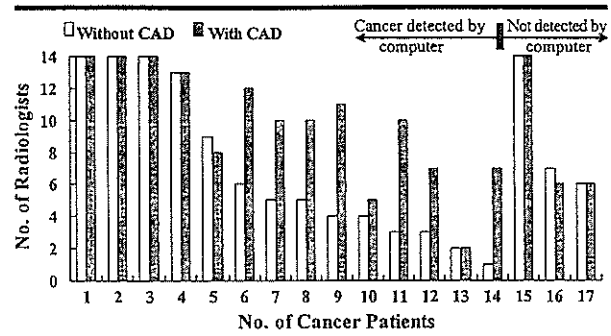


Figure 3. Bar graph shows the number of radiologists who correctly detected cancer in each of 17 patients with lung cancer with and without the use of the CAD scheme. In eight patients, the CAD scheme had a beneficial effect for one to seven radiologists. In two patients, the use of CAD had a detrimental effect for two radiologists. (See details in Discussion.)

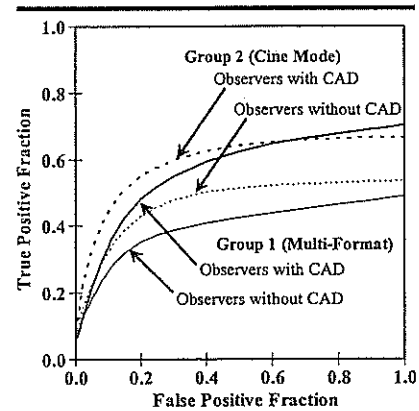


Figure 4. Graph shows LROC curves in the detection of cancers missed at CT for radiologists with and without use of the CAD scheme and with two display modes. The LROC curve was improved for groups 1 and 2 with use of the CAD scheme.

For a training session before the test, we provided five different cases (that were not part of the study set of 27) so that radiologists could learn how to operate the cine-mode interface and how to take into account the computer output in their decision. The reading time was not limited in this study. The average reading time was 48 minutes (range, 27–61 minutes; 1.8 minutes per case).

Statistical Analysis

The confidence level ratings from each observer were analyzed with use of the receiver operating characteristic (ROC) method, and a quasi-maximum-likelihood estimation of the binormal distribution was fitted to the radiologists' confidence ratings (11). The statistical significance of the difference in the area under the ROC curve (A_z) between observer readings without and with the CAD scheme was tested with use of the Dorfman-Berbaum-Metz method (12), which included both reader variation and case sample variation by means of an analysis of variance approach. Localization ROC (LROC) curves (13) for observers without and with the CAD scheme were also determined for each reading condition.

The "proper" binormal model (14) was used to fit the ROC and LROC curves (Metz CE, written communication, 2004). In this study, localization was considered correct if the center of the cancer lesion was located within 15 mm from the point marked by the observer. The distance criterion of 15 mm was based on the fact that our database contained lesions with diameters as large as 26 mm. The distance was computed automatically by the user interface program. The sensitivity in this study was defined on the basis of the number of cancer lesions that were correctly located by an observer regardless of the confidence level ratings. The statistical significance of the difference in sensitivities between the computer outputs and the observer readings without and with the CAD scheme was tested by means of a confidence interval method by taking into account reader variation alone (15). The statistical significance of the difference in sensitivities between radiologists without and with the CAD scheme and in clinical actions between a beneficial and a detrimental effect of the CAD scheme for each of the studies that did or did not contain a lung cancer was estimated with use of the Student paired t test for the 14 radiologists. In general, $P < .05$ was considered to indicate a statistically significant difference.

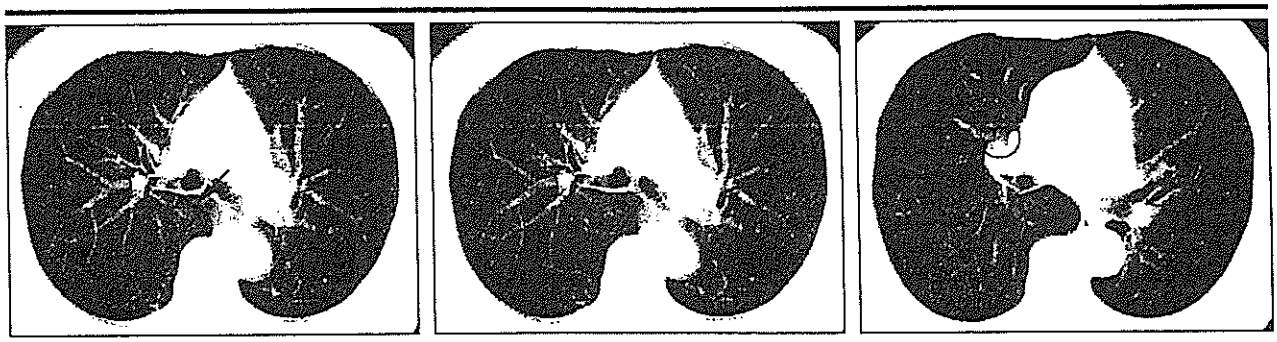
RESULTS

Radiologist Performance

With use of the CAD scheme, the average A_z value improved significantly from 0.763 to 0.854 for the 14 radiologists (P = .002), from 0.757 to 0.862 for group 1 (P = .04), and from 0.768 to 0.848 for group 2 (P = .01) (Table 1, Fig 2). No significant difference in the average A_z values between the two groups was found for radiologists without (P = .82) and with (P = .63) CAD.

In eight of the 17 patients with lung cancer, the CAD scheme helped from one to seven radiologists find the cancers (Fig 3). In two patients, CAD had a detrimental effect for two radiologists. The average LROC curves for the 14 radiologists without and with the CAD scheme in the two groups are shown in Figure 4. Figure 5 shows images from a patient in whom the use of CAD helped seven radiologists detect a cancer lesion.

With use of the CAD scheme (sensitivity, 82% [14 of 17 cancers]), the average sensitivity in the detection of 17 cancers improved significantly—from 52% (124 of 238 observations) to 68% (163 of 238 observations) for the 14 radiologists ($P <$



a. b. c.

Figure 5. Images obtained in a 69-year-old woman in whom CAD was helpful. Transverse CT scans show (a) a missed lung cancer (arrow) with pure GGO in the right upper lobe, (b) the cancer, and (c) a false-positive finding. Circles in b and c indicate the computer detections. In this patient, 10 radiologists did not detect the cancer without CAD, whereas CAD helped seven radiologists find the cancer.

TABLE 2
Number of Patients in Whom Important Clinical Action Related to Follow-up Was Changed by 14 Radiologists Owing to CAD

Observer No.	Patients with CAD-revealed Lung Cancer*		Patients without Lung Cancer†	
	Beneficial Effect	Detrimental Effect	Beneficial Effect	Detrimental Effect
1	2	0	0	1
2	2	1	0	0
3	3	1	0	0
4	5	0	1	0
5	0	0	0	0
6	6	0	1	0
7	2	0	0	0
8	2	2	0	0
9	0	0	0	0
10	0	0	0	0
11	5	0	0	0
12	1	0	0	1
13	0	0	0	1
14	2	0	1	0

Note.—"Beneficial effect" indicates a change to follow-up for a patient with lung cancer and a change to screening for a patient without lung cancer; "detrimental effect" indicates a change to screening for a patient with lung cancer or a change to follow-up for a patient without lung cancer.

* The mean number of such patients in whom the use of CAD had a beneficial effect was 2.1 ± 2.0 (standard deviation); the mean number in whom it had a detrimental effect was 0.3 ± 0.6. The difference between a beneficial effect and a detrimental effect among these patients was statistically significant ($P = .005$).

† Patients in whom CAD revealed a condition other than lung cancer. The mean number of such patients in whom the use of CAD had a beneficial effect was 0.2 ± 0.4; the mean number in whom it had a detrimental effect was also 0.2 ± 0.4.

.001), from 49% (50 of 102 observations) to 71% (72 of 102 observations) for group 1 ($P = .02$), and from 54% (74 of 136 observations) to 67% (91 of 136 observations) for group 2 ($P = .006$). The sensitivity of the CAD scheme alone was greater than that of the radiologists alone ($P < .001$) and that of the radiologists with CAD ($P < .001$), although the specificity was lower. No significant difference in sensitivities was found between the two viewing modes for radiologists without ($P = .44$) and with ($P = .71$) the use of CAD.

Clinical Actions

For the four clinical actions described earlier (ie, return to annual screening or perform diagnostic thin-section CT in 6 months, 3 months, or immediately), we attempted to quantify the changes in clinical action attributable to use of the CAD scheme. For patients with a lung cancer, the average number for whom clinical actions were changed for a beneficial effect (ie, a "step up") (3.3) was greater than the number for whom clinical actions were changed for a detrimental effect (ie, a "step down") (0.4) ($P < .001$). For patients without a lung cancer, the average numbers affected by the CAD scheme for a beneficial effect (step down) and a detrimental effect (step up) were 0.5 and 0.3, respectively ($P = .27$).

Table 2 shows the number of patients for whom the important clinical action related to follow-up was influenced positively or negatively by the 14 radiologists. For these patients, the difference between the mean number of patients in whom the action was changed from screening to follow-up (2.1 patients) and the mean number of patients in whom the action was changed from follow-up to screening (0.3 patients) was significant ($P = .005$). For patients without a lung cancer, no statistically significant difference between a beneficial effect (a change from follow-up to screening [in 0.2 patients]) and a detrimental effect (a change from screening to follow-up [in 0.2 patients]) owing to use of the CAD scheme was found for the radiologists ($P > .99$).

DISCUSSION

It has been reported (16–21) that the use of CAD has the potential to improve diagnostic accuracy in the detection of lung nodules on chest radiographs and CT scans. In previous studies with chest radiography, however, some abnormal cases—each with one lung nodule—and some normal cases were used for the observer test, and the radiologists' performance in terms of their confidence level regarding the presence or absence of a nodule was evaluated by means of ROC analysis without localization (16,17). For developing CAD schemes for use with relatively thick (18,19) or thin CT sections (20,21), the number of lung nod-

ules was generally not limited to a single nodule in each examination, and the sensitivity with which radiologists correctly detected the nodule, regardless of the confidence level, was commonly used as a measure of the radiologists' performance. In previous CT-based studies (18–21), the truth for the nodules was established by radiologist consensus—not according to pathologic results—because most small nodules are benign and do not undergo biopsy or resection.

There were some differences between the present study and the previous studies, as follows: In the present study, *(a)* the CAD scheme was developed by using missed lung cancers, which were confirmed at surgery; *(b)* the mean diameter of the cancers was 12 mm (all were at least 6 mm), and the CT findings for the cancers included lesions with pure GGO, mixed GGO, and solid opacity; and *(c)* ROC, LROC, and sensitivity analyses were used to evaluate radiologists' performance in the detection of subtle cancers without and with CAD. The importance of these differences is discussed in the next paragraphs.

Missed lung cancers include the most difficult cases for detection in clinical work and mass screening programs, and several investigators have reported the possible reasons for missing lung cancers on CT scans (4,5,22,23). In our series (5), lung cancers were missed mainly because they had low attenuation (eg, they were of small size and/or were faint lesions with GGO) or because of the presence of large structured noise elements (normal structures and/or complex backgrounds caused by other disease) or both. In addition, the cancers had poor conspicuity as defined by Kundel and Revesz (24). In general, the missed cancers corresponded to earlier visible findings in the same locations at previous examinations—findings that had been identified as abnormal according to radiologists' consensus. However, in a previous study by Austin et al (25) of radiologists' performance alone, each of six radiologists, who were biased by knowledge that the patients had lung cancers that were missed on chest radiographs, missed cancer in a mean of 26% of 22 patients. The main purpose of our study was to identify whether unassisted radiologists could identify these previously missed cancers in the context of an observer study and to evaluate whether a CAD scheme could help them detect the cancers missed on CT scans.

Diederich et al (26) reported that more than 70% of noncalcified nodules are 5 mm or smaller, and no lung cancers were found among those small lesions in CT screening programs for lung cancer at baseline. Similar findings have also been reported by Swensen et al (27). Henschke et al (28) reported that the frequency with which malignancy was or could have been diagnosed when the largest noncalcified nodule was smaller than 5 mm in diameter was very low (0 of 378). The nodules with pure or mixed GGO on CT scans in lung cancer screening programs were more likely to be malignant than were solid nodules (29,30). Although there was a limitation in the low-dose CT protocol with 10-mm-thick sections used in our study, the cancers were at least 6 mm in diameter, and the CT findings for the cancers included lesions with pure GGO, lesions with mixed GGO, and lesions with solid opacity. We believe, therefore, that it may be more important for a CAD scheme used as a "second opinion" to detect relatively large nodules with or without GGO; such nodules include primary lung cancers more frequently than small nodules, most of which are benign lesions, do.

Basically, ROC analysis without localization (11,12) can help correctly evaluate observer performance in the detection of the presence (or absence) of a lesion on medical images when each image does not include obvious false-positive findings, provided that the number of patients is sufficiently large. However, because chest CT scans may contain pulmonary vessels or focal lung diseases that have an appearance that is similar to that of nodular lesions, high positive confidence level ratings by radiologists for a given CT study do not always correspond to true-positive findings (lung cancers) but instead sometimes correspond to false-positive findings. With use of LROC analysis (13), only the responses with correct localization are evaluated for each reading condition, although a proper statistical test for practical use in evaluating the difference between the curves is still unavailable. The shortcoming of LROC analysis for estimating sensitivity is that the radiologist's performance is evaluated only for patients with true-positive findings and not for patients with true-negative findings. Therefore, in this study, we decided to evaluate the performance with three methods—that is, ROC, LROC, and sensitivity analysis—and the results obtained with all three methods showed that the diagnostic accuracy of the radiologists improved with use of the CAD scheme.

Although the radiologists in our study were able to recognize the presence of some subtle lung cancers, they could not be sure whether the CT features of the lesion were indicative of malignancy even when the computer marked the lesion. The possible reasons why the sensitivity for radiologists who used CAD did not reach at least 82% include the fact that the radiologists were not familiar with the appearance of early lung cancers at CT, especially at thick-section CT. In addition, the sensitivity of the radiologists for detecting cancer lesions was affected by some findings such as scars and vertically oriented pulmonary vessels, which had an appearance similar to that of nodular lesions on CT scans in this observer study. In addition, false-positive computer findings would have an effect on radiologists' performance in the detection of lung cancer. We noted that radiologists tended to ignore the CAD output more frequently for studies with a large number of false-positive findings (eight per study, the largest in our scheme) than for those with a small number of false-positive findings. In a previous observer study of the use of LROC analysis in the detection of clustered microcalcifications on mammograms, Chan et al (31) reported that radiologists' diagnostic accuracy with CAD was further improved by reducing the computer's false-positive rate (from four to one false-positive finding per image).

In this observer study, the use of CAD had a detrimental effect in two patients for two radiologists. In one patient, a radiologist detected a cancer lesion without CAD with a confidence level of 0.46 and made a recommendation to follow up the cancer with diagnostic thin-section CT in 3 months. The computer indicated the cancer lesion and eight false-positive findings. With use of CAD in the same patient, a different radiologist changed the location from cancer to a false-positive finding (vertical pulmonary vessel) with a confidence level of 0.59 and did not change the clinical action. In another patient, another radiologist detected a cancer lesion without CAD with a confidence level of 0.31 and recommended follow-up with CT in 6 months. The computer did not mark the cancer lesion but indicated three false-positive findings. The radiologist who used CAD also changed the location from cancer to a false-positive finding (vertical pulmonary vessel) with a confidence level of 0.46 and did not change the clinical action. Therefore, when the CAD scheme yields false-positive findings that are very

similar to true-positive findings, it may have a detrimental effect on the observers' performance when the task involves the detection of only one lesion at CT in an observer study. If radiologists were allowed to identify more than one lesion in an observer study, however, it is possible that they might elect to keep the cancer as detected initially and add the false-positive finding as a further suspicious area.

In recent CT screening programs, most images were reviewed in a multiformat display (film- or monitor-based viewing) and/or a cine-mode display (1–3,26,27). The cancers in this observer study were missed in the Nagano lung cancer screening project, in which a multiformat display (3 × 4 or 4 × 4) on two high-spatial-resolution (1728 × 2304 matrix) monitors was used (5). A similar multiformat viewing mode was used in our study by the radiologists in group 1. In general, cine viewing of CT scans of the chest is believed to improve radiologists' ability to detect lung nodules compared with film-based viewing (32,33). Tillich et al (33), however, found no significant difference between cine and film-based viewing in the detection rate of pulmonary nodules (metastases) larger than 5 mm in diameter. We also did not find a significant difference between the two viewing modes in the detection of primary lung cancers (≥6 mm) missed in a CT screening program. The limitations of this study include the facts that the low-dose CT sections were thick (10 mm), rather than thin, and the radiologists differed in the two groups. It was not the purpose of our study to compare diagnostic accuracy with the cine or multiformat mode but rather to determine that the benefits of CAD were substantial, independent of the display mode used.

In summary, lung cancers missed at low-dose CT screening were very difficult to detect, even in an observer study; the use of CAD, however, improved the radiologists' performance in the detection of these subtle cancers. In addition, CAD can help radiologists make recommendations for follow-up.

Acknowledgments: The authors are grateful to Ulrich Bick, MD, Alexandra Funaki, DO, John Fennessy, MD, Gen Iinuma, MD, Edward Michals, MD, Masaki Matsusako, MD, Peter MacEneaney, MD, Sidney Regalado, MD, Christopher Straus, MD, Gregory Scott Stacy, MD, Shuji Sakai, MD, Taylor Stroud, MD, Ira Wolke, MD, and Chaotong Zhang, MD, for participating as observers; to Charles E. Metz, PhD, and Lorenzo Pesce, PhD, for the use of proper binormal model in ROC and LROC curves; and to Elisabeth Lanzl, AM, for improving the manuscript.

References

1. Kaneko M, Eguchi K, Ohmatsu H, et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. Radiology 1996;201:798–802.
2. Sone S, Takashima S, Li F, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 1998;351:1242–1245.
3. Henschke CI, MacCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999;354:99–105.
4. Kakinuma R, Ohmatsu H, Kaneko M, et al. Detection failures in spiral CT screening for lung cancer: analysis of CT findings. Radiology 1999;212:61–66.
5. Li F, Sone S, Abe H, MacMahon H, Armato SG III, Doi K. Missed lung cancers in low-dose helical CT screening obtained from a general population. Radiology 2002;225:673–683.
6. Armato SG III, Li F, Giger ML, MacMahon H, Sone S, Doi K. Performance of automated CT nodule detection on missed cancers from a lung cancer screening program. Radiology 2002;225:685–692.
7. Arimura H, Katsuragawa S, Suzuki K, et al. Computerized scheme for automated detection of lung nodules in low-dose CT images for lung cancer screening. Acad Radiol 2004;11:617–629.
8. Suzuki K, Armato SG III, Li F, Sone S, Doi K. Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose CT. Med Phys 2003;30:1602–1616.
9. Doi K, MacMahon H, Katsuragawa S, Nishikawa RM, Jiang Y. Computer-aided diagnosis in radiology: potential and pitfalls. Eur J Radiol 1999;31:97–109.
10. Doi K. Computer-aided diagnosis in digital chest radiography. In: Samei E, Flynn M, eds. 2003 Syllabus: categorical course in diagnostic radiology—physics. Oak Brook, Ill: Radiological Society of North America, 2003; 227–236.
11. Metz CE, Herman BA, Shen JH. Maximum-likelihood estimation of receiver operating (ROC) curves from continuously distributed data. Stat Med 1998;17:1033–1053.
12. Dorfman DD, Berbaum KS, Metz CE. ROC rating analysis: generalization to the population of readers and cases with the jackknife method. Invest Radiol 1992;27:723–731.
13. Starr SJ, Metz CE, Lusted LB, Goodenough DJ. Visual detection and localization of radiographic images. Radiology 1975;116:533–538.
14. Metz CE, Pan X. "Proper" binormal ROC curves: theory and maximum-likelihood estimation. J Math Psychol 1999;43:1–33.
15. Metz CE. Quantification of failure to demonstrate statistical significance: the usefulness of confidence intervals. Invest Radiol 1993;28:59–63.
16. MacMahon H, Engelmann R, Behlen FM, et al. Computer-aided diagnosis of pulmonary nodules: results of a large-scale observer test. Radiology 1999;213:723–726.
17. Kakeda S, Moriya J, Sato H, et al. Improved detection of lung nodules on chest radiographs using a commercial computer-aided diagnosis system. AJR Am J Roentgenol 2004;182:505–510.
18. Awai K, Murao K, Ozawa A, et al. Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance. Radiology 2004;230:347–352.
19. Wormanns D, Fiebich M, Saidi M, Diederich S, Heindel W. Automatic detection of pulmonary nodules at spiral CT: clinical application of a computer-aided diagnosis system. Eur Radiol 2002;12:1052–1057.
20. McCulloch CC, Kaucic RA, Mendonca PR, Walter DJ, Avila RS. Model-based detection of lung nodules in computer tomography exams. Acad Radiol 2004;11:258–266.
21. Brown MS, Goldin JG, Suh RD, McNitt-Gray MF, Sayre JW, Aberle DR. Lung micronodules: automated method for detection at thin-section CT—initial experience. Radiology 2003;226:256–262.
22. Gurney JW. Missed lung cancer at CT: imaging findings in nine patients. Radiology 1996;199:117–122.
23. White CS, Romney BM, Mason AC, Austin JH, Miller BH, Protopapas Z. Primary carcinoma of the lung overlooked at CT: analysis of findings in 14 patients. Radiology 1996;199:109–115.
24. Kundel HL, Revesz G. Lesion conspicuity, structured noise, and film reader error. AJR Am J Roentgenol 1976;126:1233–1238.
25. Austin JH, Romney BM, Goldsmith LS. Missed bronchogenic carcinoma: radiographic findings in 27 patients with potentially resectable lesion evident in retrospect. Radiology 1992;182:115–122.
26. Diederich S, Wormanns D, Semik M, et al. Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers. Radiology 2002;222:773–781.
27. Swensen SJ, Jett JR, Sloan JA, et al. Screening for lung cancer with low-dose spiral computed tomography. Am J Respir Crit Care Med 2002;165:508–513.
28. Henschke CI, Yankelevitz D, Naidich DP, et al. CT screening for lung cancer: suspiciousness of nodules according to size on baseline scans. Radiology 2004;231:164–168.
29. Henschke CI, Yankelevitz DF, Mirtcheva R, McGuinness G, McCauley D, Miettinen OS. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. AJR Am J Roentgenol 2002;178:1053–1057.
30. Li F, Sone S, Abe H, Macmahon H, Doi K. Malignant versus benign nodules at CT screening for lung cancer: comparison of thin-section CT findings. Radiology 2004;233:793–798.
31. Chan HP, Doi K, Vyborny CJ, et al. Improvement in radiologists' detection of clustered microcalcifications on mammography: the potential of computer-aided diagnosis. Invest Radiol 1990;25:1102–1110.
32. Seltzer SE, Judy PF, Adams DF, et al. Spiral CT of the chest: comparison of cine and film-based viewing. Radiology 1995;197:73–78.
33. Tillich M, Kammerhuber F, Reittner P, Riepl T, Stoeffler G, Szolar DH. Detection of pulmonary nodules with helical CT: comparison of cine and film-based viewing. AJR Am J Roentgenol 1997;169:1611–1614.

Subcentimeter Large Cell Neuroendocrine Carcinoma of the Lung

Takaomi Hanaoka, MD, Shusuke Sone, MD,† Hitoshi Ino, MD,‡ Fumiyoshi Takayama, MD,† Toshiyuki Sato, MD,* Hiroshi Kanaya, MD,* and Hiroyuki Ogata, MD§*

Abstract: To our knowledge, no report exists of a subcentimeter size large cell neuroendocrine carcinoma (LCNEC) of the lung. A 75-year-old man participating in a low-dose CT screening program for lung cancer was found incidentally to have a partly-solid nodule in the right upper lung. After treatment with antibiotics, a repeat CT showed resolution of the nodule, but a new solid nodule measuring 9 × 9 mm was detected in the left lower lobe. The lesion showed marked enhancement on dynamic contrast-enhanced MRI. Video-assisted thoracic surgery and frozen section biopsy was suggestive of malignant lesion, resulting in extension of surgery to lobectomy with nodal dissection. The final diagnosis was stage IA-LCNEC. The estimated volume doubling time of the tumor was 30.1 days. These aggressive tumors may rarely have doubling times that overlap with benign processes.

Key Words: large cell neuroendocrine carcinoma, CT screening, peripheral small lung cancer, rapid growth, volume doubling time

(*J Thorac Imaging* 2005;20:288–290)

It is not uncommon to find small nodules in the peripheral lung zone in asymptomatic individuals on mass screening for lung cancer using low-dose chest computed tomography (CT). We previously reported that a CT screening program increases the detection rate of primary lung cancer by about 10-fold compared with screening using chest radiographs.[1] However, there is controversy regarding the interval between repeat CT screening, although there seems to be some agreement on annual repeat CT screening. Before a decision is made on the interval, there is a need to accumulate information regarding the benefits associated with such interval in detecting lung cancers at a surgically curable stage with a wide range of growth rates. We report here a patient who was incidentally found to have a subcentimeter size rapidly growing nodule that proved to be large cell neuroendocrine carcinoma (LCNEC) in the left lung. The tumor was incidentally detected at the time of repeat work up CT on the conventional CT image that was taken prior to high-resolution CT (HRCT) to examine a low-dose CT screened-nodule in the right lung.

From the *Department of Surgery, JA. Azumi General Hospital, Nagano, Japan; †Department of Radiology, JA. Azumi General Hospital, Nagano, Japan; ‡Department of Respiratory Medicine, JA. Azumi General Hospital, Nagano, Japan; and §Department of Laboratory Medicine, JA. Azumi General Hospital, Nagano, Japan.

Reprints: Dr. T. Hanaoka, Department of Chest Surgery, JA. Azumi General Hospital, Nagano Prefectural Federation of Japanese Agricultural Co-operatives, 3207-1, Ikeda, Nagano 399-8695, Japan (e-mail: hanaoka-takaomi@umin.ac.jp).

CASE REPORT

A 75-year-old current smoker (68 pack-year) asymptomatic man underwent annual low-dose CT screening in September 2003, followed by high-resolution CT (HRCT) 2 weeks later for a new 6 × 8 mm partly-solid nodule in the periphery of the right upper lobe. After 1-month of antibiotic treatment (ciprofloxacin 600 mg/d), a repeat HRCT showed partial resolution of the nodule. A third HRCT performed 3 months later following the second HRCT showed further resolution. Complete resolution of the nodule was confirmed on the fourth HRCT, performed 7 months after the initial HRCT.

On that occasion, a solid nodule measuring 9 × 9 mm in diameter with homogeneous soft tissue density and a well-defined margin was newly identified in the periphery of the left lower lobe (Fig. 1B). A careful retrospective re-examination of the conventional CT taken 3 months earlier, at the occasion of third work up CT examination, showed a tiny (4 × 4 mm diameter) lesion (Fig. 1A). Based on the measurements of the nodule on CT images, the tumor volume doubling time (VDT) was estimated at 30.1 days using the method described by Hasegawa et al.[2] Lung tumor markers, including carcinoembryonic antigen (CEA), neuron-specific enolase (NSE), and pro-gastrin-releasing peptide (Pro-GRP) were within the normal range. After a 1-month course of antifungal treatment of a possible fungal infection (itraconazole, 200 mg/d), follow-up HRCT showed tumor growth with additional pleural tag formation (Fig. 1C). Gadolinium-enhanced magnetic resonance imaging (MRI) showed definite enhancement effect of the nodule, suggesting an active lesion. The nodule was biopsied by video-assisted thoracic surgery (VATS), and intraoperatively diagnosed as a malignant tumor (frozen-section method). Based on a provisional diagnosis of a rapidly growing cancer with highly proliferative activity on CT images, a complete lobectomy with nodal dissection (ND) 2a was performed. The postoperative course was uneventful.

The final pathologic diagnosis was stage IA-large cell neuroendocrine carcinoma, measuring 8 × 10 mm, p0, pm0, n0, and p-T1N0M0. Histopathological examination of hematoxylin-eosin stained sections showed an organoid pattern, nuclear palisading pattern, central necrosis, and abundant mitosis in the cancerous tissue with no invasion of the surrounding vasculature or lymphatic vessels (Fig. 2A). Immunohistochemistry showed positive staining for chromogranin A and S-100 (Fig. 2B), but negative for NSE.

DISCUSSION

LCNEC was added to the World Health Organization (WHO) classification of lung tumors in 1999. The aggressive clinical behavior and poor prognosis of LCNEC are well documented,[3,4] and novel therapeutic approaches are needed.

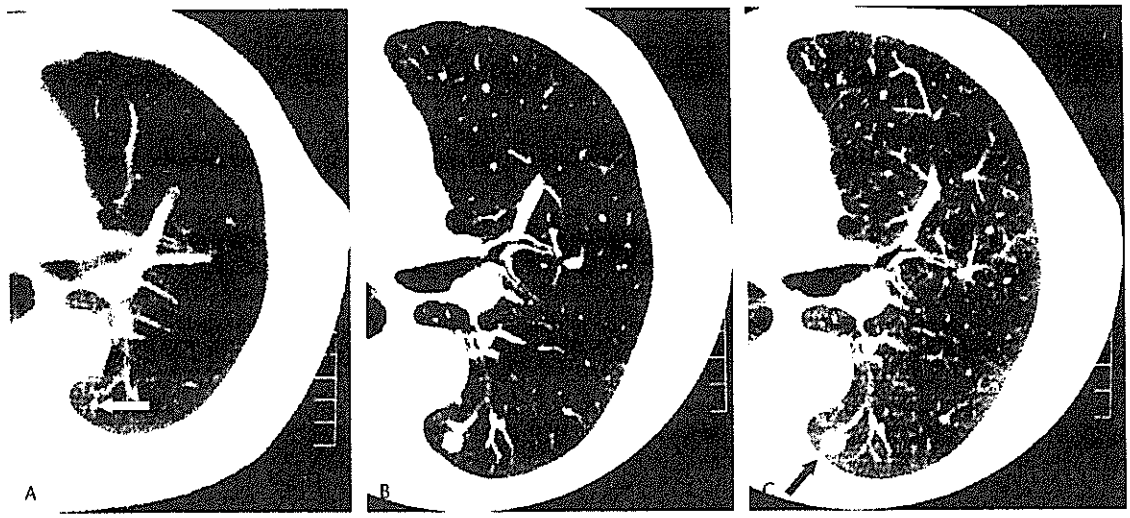


FIGURE 1. Computed Tomographic scan. A, The tumor appears as a tiny (4 × 4 mm) lesion in the periphery of the superior segment of the left lower lobe, which was identified retrospectively on a conventional CT image taken 3 months prior to that shown in (B). White arrow points to the tiny lesion. B, In this high-resolution CT (HRCT), the solid nodule was first recognized as a 9 × 9 mm sized, homogeneous one of soft tissue density with well-defined margin. C, Further growth of the lesion with additional pleural tag formation on HRCT taken 1 month after the image displayed in (B) following a 1-month course of anti-fungal treatment. Arrow shows the pleural tag.

LCNEC constitutes a minority of lung cancers, for example, only 2 cases out of a total 106 resected non-small cell lung cancers (1.9%) in our hospital during the last 4 years. Furthermore, resected subcentimeter LCNEC is extremely rare,[5] even in the CT-screening era.[6] Although the prognosis of patients with resected LCNEC is reported to be poorer than patients with the same stage of poorly differentiated non-small cell lung cancer and other large cell carcinoma even in stage I-disease,[3] because of the highly aggressive biologic behaviors,[7] complete resection of a subcentimeter LCNEC that is still in stage I is expected to be better than that in other stages.[4]

CT allows us to identify rapidly growing lung cancers like LCNEC in the localized stage so that surgery can be performed while the tumor is still in a curable stage. Follow-up and HRCTs within the context of a screening program may allow the incidental identification of rapidly growing lung tumors. The VDT of the LCNEC in our patient was 30.1 days, representing the most rapidly growing tumor of CT-screened lung cancers reported so far (for example, mean ± SD; adenocarcinoma 533 ± 381 days; squamous cell carcinoma, 129 ± 97 days; small cell carcinoma, 97 ± 46 days).[2] Such a short VDT is compatible with the aggressive

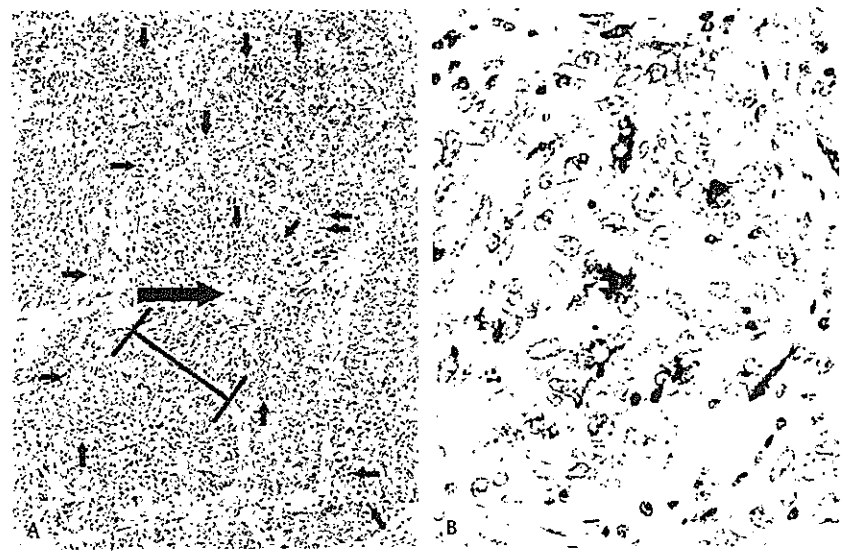


FIGURE 2. Histopathological findings. A, Note the presence of large mitotic cancer cells pointed to by small arrows, with organoid structures, central necrosis pointed to by a large arrow, and nuclear palisading patterns indicated by |—|. Magnification, ×10. B, Positive immunostaining for S-100 protein. Magnification, ×40.

clinicopathologic[3] and molecular biologic features of LCNEC reported previously.[7] Detection of subcentimeter tumors by CT scan may be the most critical point in saving the life of patients with LCNEC.

As to the interval between repeat CT screening to detect lung cancers at a surgically curable stage, although there seems to be some agreement on annual repeat CT screening, which has a length-time bias for detecting slower growing tumors and may exhibit a lower detection rate of rapidly growing tumors at a curable stage, which are more likely to be symptom-detected, repeat CT screening may become feasible at an appropriate interval to find cancers more effectively based on the detection rate, cost, radiation exposure, and curability. There is a need to accumulate more information regarding the benefits associated with such an interval to learn about clinical management approaches, presently mainly based on the surgical treatments, for such small solid lung cancer.

ACKNOWLEDGMENTS

The authors thank Ms. Kimberly K. Agnello and Prof. Claudia I. Henschke, Lung Cancer Screening Program, Department of Radiology, Weill Medical College, Cornell University, New York, for their help in the preparation of the manuscript.

REFERENCES

1. Sone S, Li F, Yang ZG, et al. Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer.* 2001;84:25–32.
2. Hasegawa M, Sone S, Takashima S, et al. Growth rate of small lung cancers detected on mass CT screening. *Br J Radiol.* 2000;73:1252–1259.
3. Takei H, Asamura H, Maeshima A, et al. Large cell neuroendocrine carcinoma of the lung: a clinicopathologic study of eighty-seven cases. *J Thorac Cardiovasc Surg.* 2002;124:285–292.
4. Zacharias J, Nicholson AG, Ladas GP, et al. Large cell neuroendocrine carcinoma and large cell carcinomas with neuroendocrine morphology of the lung: prognosis after complete resection and systematic nodal dissection. *Ann Thorac Surg.* 2003;75:348–352.
5. Asamura H, Suzuki K, Watanabe S, et al. Clinicopathological study of resected subcentimeter lung cancers: a favorable prognosis for ground glass opacity lesions. *Ann Thorac Surg.* 2003;76:1016–1022.
6. Wang JC, Sone S, Feng L, et al. Rapidly growing small peripheral lung cancers detected by screening CT: correlation between radiological appearance and pathological features. *Br J Radiol.* 2000;73:930–937.
7. Iyoda A, Hiroshima K, Moriya Y, et al. Pulmonary large cell neuroendocrine carcinoma demonstrates high proliferative activity. *Ann Thorac Surg.* 2004;77:1891–1895.

Minority Opinion

CT Screening for Lung Cancer

Claudia I. Henschke, MD, John H. M. Austin, MD,† Nathaniel Berlin, MD,‡ Thomas Bauer, MD,§ Salvatore Giunta, MD,‖ Fred Gannis, MD,¶ Michael Kalafer, MD,# Samuel Kopel, MD,** Albert Miller, MD,†† Harvey Pass, MD,‡‡ Heidi Roberts, MD,§§ Rakesh Shah, MD,‖‖ Dorith Shaham, MD,¶¶ Michael V. Smith, MD,## Shusuke Sone, MD,*** Richard Turner, MD,‡ David F. Yankelevitz, MD,* and Javier Zulueta, MD†††*

(*J Thorac Imaging* 2005;20:324–325)

The Society of Thoracic Radiology (STR) charged Dr. Stephen Swensen to form a committee to develop a consensus statement on screening for lung cancer. After numerous discussions, a consensus could not be reached. Instead, the STR decided to publish the majority opinion (this issue, page 321) and to allow me, as the dissenting opinion, to write this editorial.

While we agreed with point 1 of the majority consensus statement, we disagree with points 2–5 in the STR consensus statement as detailed below, and we also provide an alternative statement.

Majority Point 2: Screening for lung cancer with chest radiography has not been shown to lower disease-specific mortality. CT screening offers hope for earlier detection that could lower disease-specific mortality; it is unproven.

In the United States, the evidence against screening chest radiography is essentially based on a single 30-year-old study, the Mayo Lung Project. Moreover, it has been recognized generally that this study was flawed[1–3] and should not be used as a basis for public policy. Recently the American Cancer Society[4] and then the United States Preventive Services Task Force[5] have changed their previous recommendation against screening for lung cancer to one advising people to discuss the potential risks and benefits with their physician. The United States Preventive Services Task Force based its change on 6 case-control studies on screening using chest radiography in Japan, which showed a small, but real, benefit when compared with no screening.[5] Because the evidence now suggests a benefit for chest radiography screening, at least one sufficient to change the recommendations, a better diagnostic test should be of even greater benefit. CT provides for diagnosis of smaller, earlier lung cancers, as shown in multiple studies,[6–11] and its promise was used as a justification for the National Lung Screening Trial (NLST).[12]

Majority Point 3: Concerns have been raised regarding false-positive diagnoses, over diagnosis, cost, and morbidity and mortality related to intervention.

While these are important considerations for any screening program, numerous publications have shown that a well-thought-out regimen of CT screening can keep these at an acceptable level.[6–11,13–15] Thus, the real issue for high-quality screening is that a regimen should be used and adhered to as a matter of quality assurance.

To highlight one example, consider the issue of false-positives. The NLST defines all non-calcified nodules ≤4 mm to be a negative result of screening,[12] and I-ELCAP protocol calls for 1-year repeat screening when all non-calcified nodules are <5 mm in baseline screening.[8,16] Given either of these definitions, the majority of false-positives are eliminated. Similarly, all the other issues mentioned can be addressed and quantified.[8,17]

Majority Point 4: Promotion of CT screening to the general population by medical professionals with a financial interest in an enterprise is inappropriate.

This issue, we feel, is beyond the scope and expertise of our professional society, which focuses on thoracic radiology and not on public health or private practice policies and related economic issues. But, it should be noted that all academic radiologists also have an interest in the financial well-being of their department and institution.

We do not advocate CT screening for lung cancer to the general population, but we, as well as others,[3–5] consider it reasonable for a person at high-risk for lung cancer to be screened at an institution with sufficient experience in screening using an appropriate regimen of screening with quality assurance measures in place.

From *Weill Medical College of Cornell University, Cornell, NY; †Columbia University Medical Center, New York, NY; ‡Jackson Memorial Hospital, University of Miami, Miami, FL; §Christiana Care, Helen F. Graham Cancer Center, Newark, DE; ‖National Cancer Institute Regina Elena, Rome, Italy; ¶City of Hope National Medical Center, Duarte, CA; #Sharp Memorial Hospital, San Diego, CA; **Maimonides Medical Center, Brooklyn, NY; ††City University of New York at Queens College, Queens, NY; ‡‡New York University Medical Center, New York, NY; §§University of Toronto, Princess Margaret Hospital, Tornonto, Canada; ‖‖North Shore-Long Island Jewish Health System, New Hyde Park, NY; ¶¶Hadassah Medical Organization, Jerusalem, Israel; ##Georgia Institute for Lung Research, Atlanta, GA; ***Azumi Hospital, Nagano, Japan; †††Clinica Universitaria de Navarra, Pamplona, Spain.

Reprints: Claudia I. Henschke, PhD, MD, Department of Radiology, New York Presbyterian Hospital, Weill Cornell Medical Center, 525 East 68th Street, New York, New York 10021 (e-mail: chensch@med.cornell.edu).

Majority Point 5: There is insufficient evidence to justify recommending CT screening for lung cancer to patients, including those at high risk for lung cancer.

As stated above, we believe there is sufficient evidence for it to be reasonable for a person at high-risk for lung cancer with a sufficient life expectancy to pursue screening.

There is a growing body of evidence collected over the past 12 years that CT screening for lung cancer leads to a dramatic increase in the proportion of early stage genuine (fatal in the absence of early treatment) lung cancer relative to symptom-prompted diagnosis.

The proposed alternative statement is:

1. Lung cancer is the most common cause of cancer death in both men and women in the industrialized world. It, thus, is a major public health problem.[18,19]
2. Prior studies on screening for lung cancer were interpreted as not demonstrating a benefit of screening, but it is generally agreed that there were shortcomings in the methodology of those studies.[1–5]
3. It is accepted that the curability of Stage I lung cancers is very high relative to the curability of late-stage cancers; and within Stage I, cancers less than 3 cm in diameter (Stage IA) are more curable than those that are larger (Stage IB).[20,21]
4. Studies on annual CT screening have established that lung cancers are much more commonly diagnosed at Stage I and at smaller sizes than by chest radiography.[6–11]
5. Based on the points above, it is knowable that annual CT screening for lung cancer provides for prevention of death from lung cancer by early intervention. Quantitative assessment of the actual magnitude of this benefit is being pursued by studies in the US and elsewhere.
6. A person at high-risk for lung cancer yet free of suspicion-raising symptoms of it, who is interested in potentially being screened, should be fully apprised of the implications of screening and of the treatment that may result. In light of this, it is reasonable for the individual to choose to be screened by a suitably defined CT regimen.[22]

Point 5 follows from Points 3 and 4. Point 6 draws from point 5 together with the principle of Patients' Autonomy, recently enunciated by a prestigious European-US joint commission.[22]

REFERENCES

1. Strauss G, Dominioni L. Meeting report: International Conference on Prevention and Early Diagnosis of Lung Cancer. *Lung Cancer.* 1999;23: 171–172.
2. International Conference on Prevention and Early Diagnosis of Lung Cancer. Consensus statement. Varese, Italy. *Cancer.* 2000;89:2329–2330. Also at www.IELCAP.org.
3. Strauss GM, Dominioni L, Jett JR, et al. Como International Conference Position Statement. *Chest.* 2005;127:1146–1151.
4. Smith RA, von Eschenback AD, Wender R, et al. American Cancer Society guidelines for the early detection of cancer: update of early detection guidelines for prostate, colorectal, and endometrial cancer; also update 2001—testing for early lung cancer detection. *CA Cancer J Clin.* 2001;51:38–75.
5. Humphrey LL, Johnson M, Teutsch S. Lung cancer screening with sputum cytologic examination, chest radiography, and computed tomography: an update of the U.S. Preventive Services Task Force. *Ann Intern Med.* 2004; 140:738–753. Also at http://ahrg.gov/clinic/cps3dix.htm.
6. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet.* 1999;354:99–105.
7. Henschke CI, Naidich DP, Yankelevitz DF, et al. Early Lung Cancer Action Project: initial results of annual repeat screening. *Cancer.* 2001;92: 153–159.
8. Henschke CI, Yankelevitz DF, Smith JP, et al. CT Screening for Lung Cancer: assessing a regimen's diagnostic performance. *Clin Imaging.* 2004;28:317–321.
9. Sone S, Takahima S, Li F, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet.* 1998;351:1242–1245.
10. Sone S, Li F, Yang Z-G, et al. Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. *Br J Cancer.* 2001;84:25–32.
11. Sobue T, Moriyama N, Kaneko M, et al. Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. *J Clin Oncol.* 2002;20:911–920.
12. National Lung Screening Trial. website: http://www.cancer.gov/nlst and http://acrin.org.
13. Nawa T, Nakagawa R, Kusano S, et al. Lung cancer screening using low-dose spiral CT. *Chest.* 2002;122:15–20.
14. Pasterino U, Bellomi M, Landoni C, et al. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet.* 2003;362:593–597.
15. Bastarrika G, Garcia Velloso MJ, Lozano MD, et al. Early lung cancer detection using spiral CT and positron emission tomography. *American Review of Respiratory and Critical Care Medicine.* 2005;171:1378–1383.
16. Henschke CI, Yankelevitz DF, Naidich DP, et al. CT screening for lung cancer: suspiciousness of nodules according to size on baseline scans. *Radiology.* 2004;231:164–168.
17. Consensus Statement of the International Conference on Screening for Lung Cancer. Online: http://cancer.org/downloads/STT/CAFF_finalpw secured.pdf.
18. American Cancer Society Facts and Figures. Also at www.IELCAP.org.
19. Parkin D. Global cancer statistics in the year 2000. *Lancet Oncol.* 2001;2: 533–543.
20. Mountain CF. Revisions in the International System for Staging Lung Cancer. *Chest.* 1997;111:1710–1717.
21. American Joint Committee on Cancer (AJCC). Lung. In: Fleming ID, Cooper JS, Hensen DE, et al., eds. *Cancer Staging Manual,* 5th ed. Philadelphia: Lippincott-Raven; 1997;127.
22. Medical professionalism in the new millennium. A physician charter. *Ann Intern Med.* 2002;136:243–246.

外科治療 Vol. 93 No. 4 (2005：10)

特集 増え続ける肺癌—治療成績を向上させるために

肺癌検診の現状

曾根 脩輔　花岡 孝臣　高山 文吉

永井書店

特集 増え続ける肺癌－治療成績を向上させるために

肺癌検診の現状

Current situation of screening for lung cancer

曾根脩輔 SONE Shusuke　花岡孝臣* HANAOKA Takaomi　高山文吉** TAKAYAMA Fumiyoshi

肺がんは，がん死亡の第1原因である．罹患数の多さのみならず，その救命率の極端な低さによる．治療成績の向上には早期発見が不可欠である．肺野型肺がんの早期発見には胸部X線写真は非力である．低X線曝射方式のCT検診によりこれが可能になり肺がんの治療成績は飛躍的に向上するであろう．その普及には，受診者数を増加するための検診コストへの配慮と検診実施態勢，精度管理体制の確立が望まれる．

はじめに

平成16年度のわが国における肺がん死亡数は59,910名(男性43,910名，女性16,000名)，人口10万人対死亡率は47.5(男性71.3，女性24.8)で，がん死亡の第1位であった．罹患数と死亡数の差，すなわち救命数(率)が低い点でも肺がんはすい臓がんとともに特異であり，難治がんと目される状況にある．肺がんの5年生存率は米国でもおよそ14%でありわが国と同様に低い[1]．早期肺がんの手術成績は，病理学的病期IAでは約80%と報告されているにもかかわらず[2]，肺がんの85%以上が進行がん状態で発見されるために，全体としてはこのような低治癒率を示している[3]．肺がん対策を唱えるからには，早期発見と治療の実現，そのために有効な検診法の導入に目を向けるべきであり，現実には，肺がんの早期発見に役立つ低X線CT検診の有用性の理解と具体化に向けた動きが必要である．

I. 肺がん検診

一般にがん検診が有効に働く要件として，対象がんの罹患率が高く社会的関心や影響が大きいこと，現状では救命率(罹患率と死亡率の差が少ない)が低いが，早期発見できるテスト法があって，これによる発見と診断確定後に効果的治療が実施可能なこと，そして当該のテスト法が一般に普及可能な内容，すなわち安全性や侵襲性，利便性に問題のないことなどがあげられる[4]．今日，肺がんはまさにこのような条件に合致しており，がん検診の対象疾患として最適である．

従来は，胸部疾患一般の検診に胸部X線単純写真が利用され，上述の要求を満たすと考えられていたであろう．1987年来，老人保健事業における肺がん検診の胸部検査法として，痰の細胞診断法とともに定着している．しかし肺がんではこれ

JA長野厚生連中南信地区がん検診センター長・信州大学名誉教授 *JA長野厚生連安曇総合病院呼吸器外科 部長 **同病院放射線科 部長

Key words：肺がん／がん検診／CT検診

0433-2644/05/¥50/頁/JCLS