Note: This copy is for your personal non-commercial use only. To order presentation-ready copies for distribution to your colleagues or clients, contact us at www.rsna.org/rsnarights.

Pulmonary Nodules: Estimation of Malignancy at Thin-Section Helical CT—Effect of Computer-aided Diagnosis on Performance of Radiologists[1]

Kazuo Awai, MD
Kohei Murao, PhD
Akio Ozawa, BS
Yoshiharu Nakayama, MD
Takeshi Nakaura, MD
Duo Liu, MD
Koichi Kawanaka, MD
Yoshinori Funama, PhD
Shoji Morishita, MD
Yasuyuki Yamashita, MD

Purpose: To evaluate the effect of a computer-aided diagnosis (CAD) system on the diagnostic performance of radiologists for the estimation of the malignancy of pulmonary nodules on thin-section helical computed tomographic (CT) scans.

Materials and Methods: The institutional review board approved use of the CT database; informed specific study-related consent was waived. The institutional review board approved participation of radiologists; informed consent was obtained from all observers. Thirty-three (18 malignant, 15 benign) pulmonary nodules of less than 3.0 cm in maximal diameter were evaluated. Receiver operating characteristic (ROC) analysis with a continuous rating scale was used to compare observer performance for the estimation of the likelihood of malignancy first without and then with the CAD system. The participants were 10 board-certified radiologists and nine radiology residents.

Results: For all 19 participants, the mean area under the best-fit ROC curve (A_z) values achieved without and with the CAD system were 0.843 ± 0.097 (standard deviation) and 0.924 ± 0.043, respectively. The difference was significant ($P = .021$). The mean A_z values achieved without and with the CAD system were 0.910 ± 0.052 and 0.944 ± 0.040, respectively, for the 10 board-certified radiologists ($P = .190$) and 0.768 ± 0.078 and 0.901 ± 0.036, respectively, for the nine radiology residents ($P = .009$).

Conclusion: Use of the CAD system significantly ($P = .009$) improved the diagnostic performance of radiology residents for assessment of the malignancy of pulmonary nodules; however, it did not improve that of board-certified radiologists.

© RSNA, 2006

[1] From the Department of Diagnostic Radiology, Graduate School of Medical Sciences (K.A., Y.N., T.N., D.L., K.K., S.M., Y.Y.), and Department of Radiological Technology, School of Health Sciences (Y.F.), Kumamoto University, 1-1-1 Honjyo, Kumamoto 860-8556, Japan; and Bio-IT Business Development Group, Fujitsu, Chiba, Japan (K.M., A.O.). Received February 6, 2005; revision requested April 6; revision received May 10; final version accepted June 13. **Address correspondence to** K.A.

© RSNA, 2006

Screening for lung cancer by using low-radiation-dose helical computed tomography (CT) (low-dose CT) has gained attention in the past years. With low-dose CT, the detection rate has been observed to be 0.4%–2.7%, a value that is 2.6–10-fold higher than that for detection with chest radiography (1–10). Moreover, noncalcified nodules have been detected at 5%–66% of the examinations, a rate that is 1.7–3.0-fold higher than that for detection with chest radiography (1,2,6).

Sobue et al (11) reported that small lung cancers are associated with a better patient survival rate and that the 5-year survival rate was almost 100% in patients with nodules that were 9 mm or smaller. According to Henschke et al (6), in whose study 233 participants with one to six noncalcified nodules underwent low-dose CT, the largest nodule was 2–5 mm in 58%, 6–10 mm in 30%, 11–20 mm in 9%, and greater than 20 mm in 2% of the patients. Diederich et al (7) also used low-dose CT for the detection of pulmonary nodules; in their study, the largest nodule was 2–5 mm in 73%, 6–10 mm in 23%, and greater than 10 mm in 3% of the 1001 participants. These findings suggest that low-dose CT is a promising method for the detection of small lung cancers.

Only 2.5%–11.6% of detected noncalcified nodules, however, prove to be lung cancer, and screening with low-dose CT results in many false-positive findings (3–7,9). Also, in a 1-year follow-up study with low-dose CT, noncalcified nodules were detected at 2.5%–3.9% of the examinations; only 3%–23% of these nodules were identified as lung cancer (1,5,10). Therefore, rational algorithms that facilitate the accurate diagnosis of noncalcified nodules detected at lung cancer screening with low-dose CT must be developed.

Thin-section CT has been recommended as the next step when a noncalcified nodule is detected at low-dose CT screening. At present, however, there are no clear diagnostic criteria for identification of malignant nodules detected by using thin-section CT, and in most instances, the interpretation of thin-section CT findings relies on the knowledge and experience of the radiologist who is performing the interpretation. The independent interpretation of noncalcified pulmonary nodules by two or more experienced radiologists and the use of a computer-aided diagnosis (CAD) system (12–36) for estimation of the malignancy of the nodules may assist radiologists in determination of a correct diagnosis. Thus, the purpose of our study was to evaluate the effect of a CAD system on the diagnostic performance of radiologists for the estimation of the malignancy of pulmonary nodules on thin-section helical CT scans.

Advance in Knowledge

- Use of the CAD system significantly ($P = .009$) improved the diagnostic performance of radiology residents for assessment of the malignancy of pulmonary nodules; however, it did not improve that of board-certified radiologists.

Materials and Methods

All patients who underwent CT examinations at Kumamoto University Hospital, Kumamoto, Japan, gave prior informed consent for the use of their CT images in future retrospective studies. Our institutional review board approved the use of the CT database, and informed specific study-related consent was waived. Our institutional review board also approved the participation of radiologists in this observer performance study. Informed consent for the observer performance study was obtained from all observers.

Two authors (K.M. and A.O.) were employees of Fujitsu, and two other authors (K.A. and Y.Y.) had control of inclusion of all data and information for this study.

Nodule Selection

One chest radiologist (K.A.) with 18 years of chest CT experience reviewed the records of 171 consecutive patients who were suspected of having pulmonary nodules and underwent thin-section helical CT of the chest at our institute during a 24-month period from January 2002 to December 2003. He did not participate in the observer performance study. He selected all patients who satisfied the following criteria: They had only one pulmonary nodule that did not exceed 3 cm, and there was neither consolidation caused by the presence of organizing tissues after pneumonia nor consolidation associated with idiopathic pulmonary fibrosis around the nodules because it was difficult to define the boundary of the nodule. On the basis of these criteria, 21 malignant and 26 benign or possibly benign nodules were identified. Because three malignant nodules for which a histologic diagnosis was not determined were excluded, 18 malignant nodules were available for the observer performance test. Histologically, all 18 nodules were identified as primary lung cancer; 14 were well-differentiated adenocarcinomas and four were moderately differentiated adenocarcinomas. The diagnosis of these nodules was based on findings at video-assisted thoracic surgery in 14 nodules, findings at CT-guided transcutaneous biopsy in two nodules, and findings at bronchoscopic transbronchial biopsy in two nodules. We did not encounter any patients with solitary pulmonary metastasis during the period between January 2002 and December 2003.

Of the 26 possibly benign nodules, 15 were chosen for the observer performance test. In seven of these nodules, a

Published online before print
10.1148/radiol.2383050167

Radiology 2006; 239:276–284

Abbreviations:
A_z = area under the best-fit ROC curve
CAD = computer-aided diagnosis
ROC = receiver operating characteristic
ROI = region of interest

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

See Materials and Methods for pertinent disclosures.

histologic diagnosis was established at video-assisted thoracic surgery. Among the histologically proved benign nodules, two were inflammatory lesions, two were focal fibrosis, one each was a granuloma and a pulmonary hamartoma, and one nodule was caused by anthracosilicosis. The other eight nodules either were histologically undetermined but manifested no change in size and internal and marginal characteristics at chest CT performed in the course of more than 2 years (six nodules) or disappeared at chest CT within 3 months (two nodules). The other 11 of 26 possibly benign nodules also manifested no change in size and internal or marginal characteristics at chest CT; however, observation periods for these nodules were shorter than 2 years (average, 10.7 months; range, 4–20 months). Therefore, we did not include these 11 nodules in the observer performance test.

The mean size in the xy (or transverse) plane of the 18 malignant nodules was 20.1 mm ± 8.9 (standard deviation), and that of the 15 benign nodules was 17.4 mm ± 9.1. According to the results of the two-tailed Student *t* test, there was no significant difference in size between malignant nodules and benign nodules ($P = .394$). The 33 patients were 15 men and 18 women who were 25–79 years old (mean age, 61.8 years); the mean age of patients with malignant nodules was 62.9 years (range, 25–79 years), and that of patients with benign nodules was 60.6 years (range, 39–72 years). According to the results of the two-tailed Student *t* test, there was no significant difference between the age of patients with malignant nodules and that of patients with benign nodules ($P = .581$).

CT Imaging

CT imaging was performed with one of two scanners (LightSpeed QX/i, GE Medical Systems, Milwaukee, Wis; Somatom Volume Zoom, Siemens Medical Systems, Erlangen, Germany). Scanning parameters for thin-section helical CT with the LightSpeed scanner were as follows: detector row width, 1.25 mm; helical pitch, 0.75; section thickness, 1.25 mm; section interval, 1.25 mm; rotation time, 0.8 second; tube voltage, 120 kVp; and tube current, 160–200 mA. For scanning with the Somatom Volume Zoom machine, the parameters were: detector row width, 1.0 mm; helical pitch (beam pitch), 0.75; section thickness, 1.0 mm; section interval, 1.0 mm; rotation time, 0.75 second; tube voltage, 120 kVp; and tube current, 160–200 mA.

Computerized Scheme for Differentiation of Pulmonary Nodules

First, one radiologist (K.A.) specified the region of interest (ROI) on a section that included the target nodule (Fig 1). The radiologist did not need to specify the ROI on the other sections because extension of the nodule along the z-axis was automatically estimated from the size of the ROI on the section. We determined the size of the ROI to encompass the nodule in all x-, y-, and z-axis directions and changed the size of the ROI in each nodule. For example, we used an ROI of 4–6 cm for the nodule that was 3 cm in maximal diameter. In the observer performance test described later, we set up the ROI for each case in advance and presented it on the screen of the cathode ray–tube monitor. Once the ROI was indicated, segmentation of the nodule was performed automatically with our CAD system.

The segmentation process was divided into three steps as follows:

1. Interpolation of image data was performed so that the pixel sizes in the x-, y-, and z-axis became isotropic. This preprocessing was necessary to segment the nodule with a smooth boundary and to quantify accurately the morphologic features of the pulmonary nodule.

2. Classification of the nodule type was then determined. Nodules were considered as solid, as a ground-glass opacity, or as that with a cavity. This classification was achieved by calculating the averaged CT number (C_{av}) of the small-volume area, which was a 1.5-mm cube around the center of the ROI. We calculated the C_{av} of the cube and defined the types as solid for C_{av} of −150 HU or greater, as ground-glass opacity for C_{av} of greater than −150 HU to −800 HU or greater, and as a nodule with a cavity for C_{av} of greater than −800 HU.

3. Elimination of structures tangent to the nodule, such as vessels and tho-

Figure 1

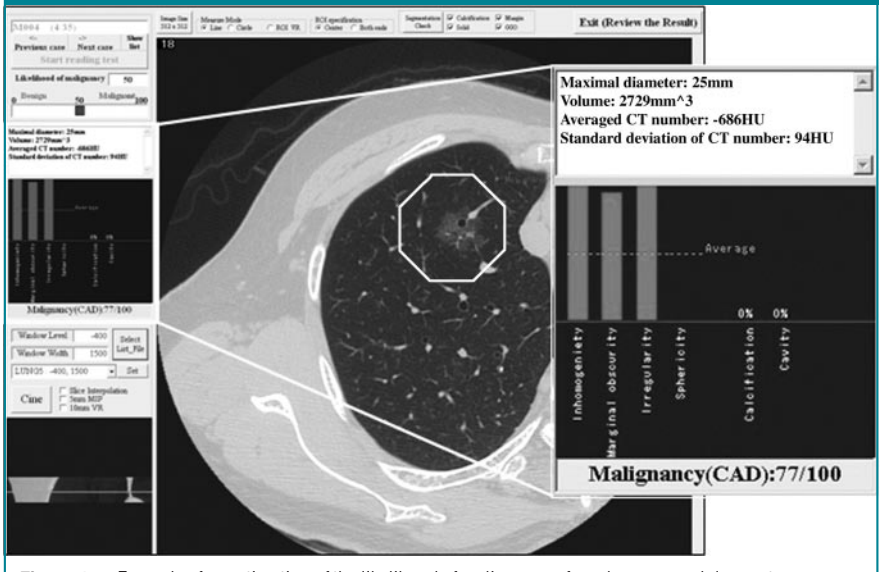


Figure 1: Example of an estimation of the likelihood of malignancy of a pulmonary nodule on a transverse thin-section CT image. In the main window, the target nodule is indicated by the octagon. The pathologic diagnosis was well-differentiated adenocarcinoma. In the enlarged window at right are the results of the estimation. The last row shows the estimation of likelihood of malignancy.

racic walls, was performed. This elimination was achieved by using several image processing techniques, such as "snakes" (37) and mathematical morphology (38). For example, vessels around the nodule were eliminated by using the "geodesic erosion/dilation" technique of mathematical morphology (38). At first, a lesion area within the ROI was extracted according to the threshold of the CT number. Then, the margin of the lesion area was eroded by a structure element that had almost the same diameter as the vessels adjacent to the nodule. Finally, the lesion area was dilated so as to reproduce the accurate margin of the lesion area without vessels.

In the next step, morphologic features of the nodule were quantified on the basis of the segmentation results with our CAD system (Table 1). We selected morphologic features that were effective for estimation of the malignancy of the nodule by investigating the strength of the neural network connection (39). In the training stage, a neural network (40,41) was established by using an iterated calculation to minimize the difference between pathologic results and the outputs from the morphologic features. Thirty-four nodules were used for the training of our CAD system, and these cases were not included among the cases used for the observer performance test.

Table 1

Morphologic Features for Quantification of Pulmonary Nodules

Morphologic Feature	Notes
Density	
Averaged CT number	None
Standard deviation of CT number	None
Internal inhomogeneity of CT number	Nonuniformity of gradient vectors inside solid part
Marginal obscurity	Volume rate of lower CT number part tangent to solid part
Shape	
Volume	None
Sphericity	None
Irregularity of surface	Nonuniformity of normal vectors on surface of nodule
Maximal diameter	None
Calcification volume*	Volume rate of calcification to volume of whole nodule
Cavitary volume*	Volume rate of cavity to volume of whole nodule

* The value was calculated as a percentage.

Figure 2

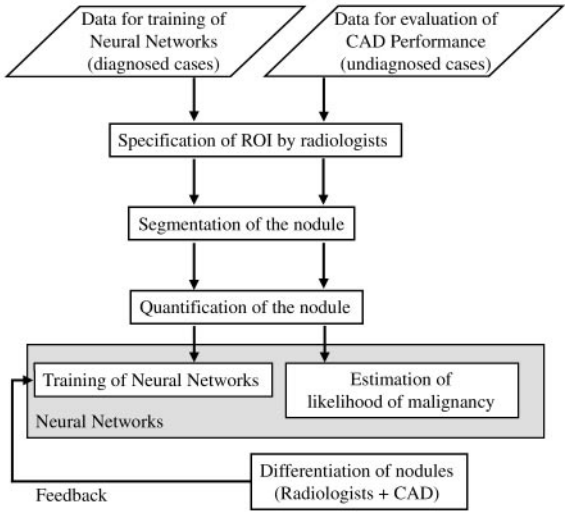


Figure 2: Diagram of computerized scheme for quantification of pulmonary nodules on thin-section helical CT images.

In the differentiation stage, the morphologic features of undiagnosed nodules were fed to the neural networks. The output was the likelihood of malignancy that was represented by the continuous values ranging from 0% to 100%. At the observer performance test, we also showed the results of analysis of the morphologic features. The values of maximal diameter, volume, mean, and standard deviation of the CT number of the nodule were shown in a text style so that radiologists could check whether the segmentation results were acceptable in comparison with their visual observations. The values of features such as inner inhomogeneity of CT number, marginal obscurity, irregularity of the surface, and sphericity were shown in a bar chart. The height of the bar was normalized according to the statistics of the training data set for the CAD system. The value of the feature F was normalized to $\tilde{F}$ according to the following equation:

$$\tilde{F} = \frac{F - F_{av}}{2F_{\sigma}} + 50.$$

Here, F_{av} and F_{σ} were the average and the standard deviation, respectively, of the morphologic features of the training set for use of the CAD system (Fig 2).

A computer workstation (Primergy; Fujitsu, Tokyo, Japan) with dual 1.4-GHz processors (Pentium III; Intel, Santa Clara, Calif) was used in this study, and the average calculation time for the CAD system to analyze each nodule was about 2.5 minutes.

Observer Performance Study

With a sequential-test method, we employed receiver operating characteristic

(ROC) analysis to evaluate the performance of our CAD system and that of the 19 participating radiologists in distinguishing between benign and malignant nodules without and with the CAD output (42). The observers were 10 board-certified radiologists with 8–26 years of experience (mean, 17.0 years) and nine radiology residents with 1–4 years of experience (mean, 2.1 years). All board-certified radiologists specialized in body imaging and read thoracic CT images on a regular basis.

Two clinical parameters (patient age and sex) were disclosed to the observers on the cathode-ray tube monitor. They were allowed to change the level and width of the window on the monitor; reading time was not limited. The observers were not given specific criteria for judging a nodule as benign or malignant; rather, they were asked to use their existing knowledge. In response to the question of whether they judged the nodule to be benign or malignant, the radiologists marked their confidence level in regard to the likelihood of malignancy on a continuous rating scale (43,44). By using a mouse, they indicated their judgment on a horizontal bar displayed on the screen; "definitely malignant" and "definitely benign" were marked at the right and left ends of the bar, respectively.

CT images were first presented without the CAD output. After each radiologist marked the initial level of confidence, the computer output for the results of quantitative measurements and of the analysis for likelihood of malignancy were displayed on the monitor. Each observer then had a chance to change the previously indicated confidence level.

Before training and test taking, the participating radiologists were informed that the purpose of the experiment was to evaluate the potential benefit of using the CAD system to help in the differentiation between malignant and benign pulmonary nodules on thin-section helical CT scans. They were also told that 33 nodules would be shown randomly, that approximately 50% of the nodules were malignant, and that the accuracy of the computer output was about 73% when a threshold level of 50% for the estimated degree of malignancy with the computer was used as the measure of the likelihood of malignancy. They were instructed to use the rating scale consistently and uniformly. All participants were trained with three training cases of pulmonary nodules before they participated in the observer performance test to make sure that they could operate the observer interface and they knew how to take into account the computer output in their decision making. Cases used for training were not included among those used for the observer performance test.

To assess the diagnostic difficulty level of each observer performance test case, we recorded the number of radiologists who, without the CAD system, incorrectly answered whether the nodule was malignant or benign in cases in which a threshold level of 50% was used as the measure of the likelihood of malignancy.

Figure 3

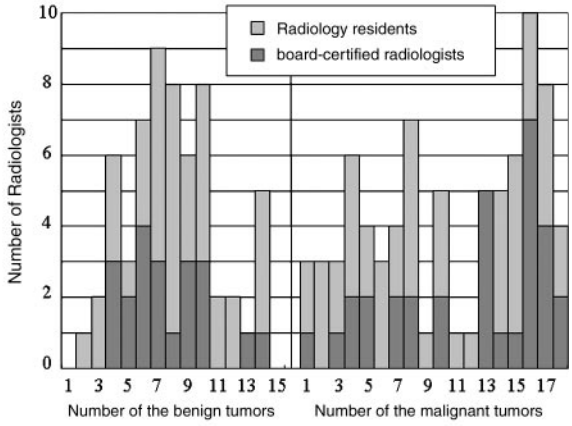


Figure 3: Graph shows number of radiologists who incorrectly identified 33 nodules as malignant or benign without CAD system.

Table 2

Sensitivity, Specificity, and Positive Predictive Value of Board-certified Radiologists, Radiology Residents, and CAD System Alone for Diagnosis of Malignant Pulmonary Tumors

Statistic	Board-certified Radiologists	Radiology Residents	All Radiologists	CAD System
Sensitivity	76.1	67.3	71.9	72.2
Specificity	89.3	74.8	82.5	75.0
Positive predictive value	90.0	77.2	84.0	81.3

Note.—All numbers are percentages.

Statistical Analysis

ROC analysis was used to compare the radiologists' performance without and with CAD output in distinguishing between benign and malignant pulmonary nodules. A binormal ROC curve was fitted to each radiologist's confidence rating data from two reading conditions with quasi–maximum likelihood estimation (43). A computer program (LABMRMC; Charles E. Metz, University of Chicago, Chicago, Ill) was used to obtain binormal ROC curves from the ordinal-scale rating data (43). The area under the best-fit ROC curve (A_z) plotted in the unit square was calculated for each fitted curve. The statistical significance of the difference between the ROC curves obtained without CAD output and those obtained

Figure 4

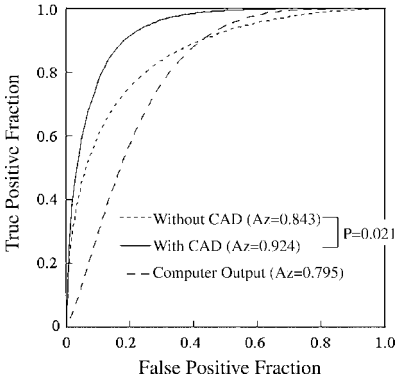


Figure 4: Mean ROC curves for all observers who distinguished between benign and malignant nodules without and with CAD output and ROC curve for the CAD output alone. The mean A_z values for all radiologists increased from 0.843 ± 0.097 without CAD output to 0.924 ± 0.043 with CAD output. The difference was significant (P = .021). The A_z value for the CAD output alone was 0.795.

with CAD output was tested by using the same computer program; differences were estimated by using analysis of variance for pseudovalues of A_z calculated from the rating scores of all 19 participating radiologists.

The significance of the difference between A_z values was determined. These differences were between the board-certified radiologists who performed assessments without or with the CAD system and the radiology residents who performed assessments with the CAD system. They were also between the board-certified radiologists who performed assessments without the CAD system and the radiology residents who performed assessments without the CAD system. Evaluation of these differences was performed with multiple comparisons by using the Tukey-Kramer method. Statistical analyses were performed with a statistical software package (SPSS, version 11.0; SPSS, Chicago, Ill). P values of less than .05 were considered to indicate a significant difference.

We also calculated the sensitivity, specificity, and positive predictive values for the diagnosis of malignant nodules determined by the board-certified radiologists and by the radiology residents.

Table 3

A_z Values for Board-certified Radiologists and Radiology Residents for Diagnosis without and with CAD Output

Observer	A_z Value without CAD Output	A_z Value with CAD Output
Board-certified radiologists		
1	0.871	0.890
2	0.808	0.888
3	0.871	0.939
4	0.883	0.896
5	0.927	0.953
6	0.939	0.981
7	0.953	0.974
8	0.929	0.980
9	0.940	0.947
10	0.984	0.992
Average*	0.910 ± 0.052	0.943 ± 0.040
Radiology residents		
11	0.836	0.931
12	0.811	0.874
13	0.776	0.885
14	0.761	0.874
15	0.668	0.902
16	0.754	0.883
17	0.653	0.935
18	0.749	0.860
19	0.905	0.969
Average*	0.768 ± 0.078	0.901 ± 0.036
Average for all*	0.843 ± 0.097	0.924 ± 0.043

* Data are the mean ± standard deviation.

Results

Of the 18 malignant nodules, three were incorrectly diagnosed by more than three board-certified radiologists, and nine were incorrectly diagnosed by more than three radiology residents. Of the 15 benign nodules, five were incorrectly diagnosed by more than three board-certified radiologists, and seven were incorrectly diagnosed by more than three residents (Fig 3). When a threshold level of 50% was used for the measure of the likelihood of malignancy in regard to the confidence level, the sensitivity, specificity, and positive predictive values of our CAD system alone were 72.2%, 75.0%, and 81.3%, respectively (Table 2). All sensitivity, specificity, and positive predictive values for the CAD system were between the values for board-certified radiologists and those for the radiology residents (Table 2). The A_z value for the CAD system alone was 0.795 (Fig 4). The A_z values for all 19 observers were significantly higher with CAD output than they were without CAD output (Table 3). Results of the analysis of overall performance of the 19 observers for distinguishing between benign and malignant pulmonary nodules (Fig 4) indicated that the averaged A_z values for all radiologists increased from 0.843 ± 0.097 without CAD output to 0.924 ± 0.043 with CAD output; the difference was significant (P = .021) (Table 3).

For the board-certified radiologists, the mean A_z values (Figs 5, 6) obtained without and those obtained with CAD output were 0.910 ± 0.052 and 0.944 ± 0.040, respectively; the difference was not significant (P = .190) (Table 3). On the other hand, for the group of residents, the mean A_z values obtained without CAD output and those obtained

with CAD output were 0.768 ± 0.078 and 0.901 ± 0.036, respectively; these values were significantly different (P = .009) (Table 3).

There was a significant difference between the mean A_z values recorded for board-certified radiologists and those recorded for the residents when assessment was performed without the CAD system ($P < .001$). There was no significant difference, however, between the mean A_z values recorded for board-certified radiologists who did use the CAD system or those who did not use the CAD system, on one hand, and those for the group of residents who used the CAD system in their assessment, on the other (P = .892 and P = .101, respectively).

Figure 5

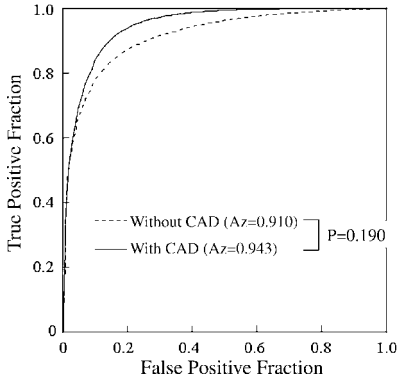


Figure 5: Mean ROC curves for the 10 board-certified radiologists who distinguished between benign and malignant nodules without and with CAD output. The mean A_z values obtained without and with the CAD output were 0.910 ± 0.052 and 0.943 ± 0.040, respectively. The difference was not significant (P = .190).

Figure 6

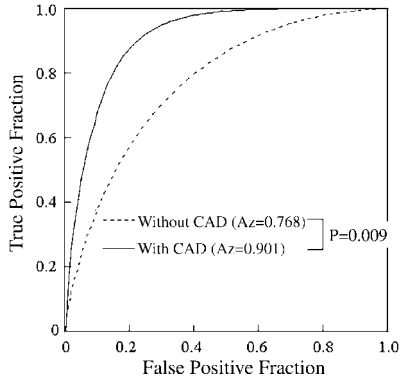


Figure 6: Mean ROC curves for the nine radiology residents who distinguished between benign and malignant nodules without and with CAD output. The mean A_z values without and with the CAD output were 0.768 ± 0.078 and 0.901 ± 0.036, respectively. The difference was significant (P = .009).

Discussion

The spread of lung cancer screening with low-dose CT has drawn attention to the importance of the use of a CAD system for the assessment of pulmonary nodules (12–21,23–31,33–36,39,45,46). We found that, when radiology residents used the CAD system, their diagnostic abilities advanced to the same level as that of board-certified radiologists who performed the assessment with or without the CAD system. On the other hand, use of the CAD system did not significantly enhance the diagnostic abilities of board-certified radiologists. Of 18 malignant nodules, three (17%) were incorrectly diagnosed by more than three board-certified radiologists, as were five (33%) of the 15 benign nodules. On the other hand, nine (50%) malignant nodules and seven (47%) benign nodules were incorrectly diagnosed by more than three residents. This finding suggests that the data sets presented in our observer performance test may have been relatively easy to diagnose for board-certified radiologists and may have been relatively difficult for the radiology residents to diagnose. The A_z value for our CAD system was smaller than that for the board-certified radiologists as a group and larger than that for the residents as a group. To render our CAD system useful for not only less experienced residents but also experienced board-certified radiologists, its performance must be raised to the diagnostic ability level of board-certified radiologists.

Our CAD system displays not only estimations of malignancy in percentages but also typical parameters in numeric values or in histogram formats. This may make it possible for the observer to use the displays for assessment of the appropriateness of the malignancy estimations calculated with the CAD system. We did not, however, confirm that results of morphologic feature analysis actually contributed to final decisions of radiologists.

With our CAD system, estimates of the likelihood of malignancy that are based on morphologic features of the pulmonary nodules on thin-section helical CT scans were calculated. The addition of clinical information, such as patient age, sex, and smoking history, may improve the diagnostic performance of our system. Furthermore, although the neural network was used with our CAD system to estimate the likelihood of malignancy on the basis of various morphologic features, Bayesian analysis (47–49) is a viable alternative.

With our CAD system, estimates of the likelihood of malignancy are determined by using learning input data. When the CAD system contains data on benign nodules, such as a pulmonary hamartoma, but not on metastatic pulmonary tumors, the latter may be estimated as benign. Therefore, it is important to recognize that the estimation of the likelihood of malignancy that is based on the assessment of only CT images is not totally reliable. Even when the CAD system returns a diagnosis of a benign pulmonary nodule, if this estimation is based on a single CT image, it is advisable to monitor the nodule for temporal changes. Our CAD system can be used to monitor temporal changes from various perspectives because the results of analysis that can be performed in regard to internal density and marginal characteristics, as well as volumetry, are considered.

We used a threshold level of 50% for the measure of the likelihood of malignancy with our CAD system. This threshold level was used to assess the sensitivity, specificity, and positive predictive values for our CAD system, for board certified radiologists, and for radiology residents for the estimation of malignancy. Much lower probabilities of malignancy can and should result in a change in patient treatment. For example, some decision-analysis studies suggest that a probability of 5%–10% or

greater should result in a change in treatment (50,51). The likelihood of malignancy that was estimated by using the CAD system represents only a reference for probability of malignancy, and radiologists or pulmonologists must determine the appropriate treatment for pulmonary nodules.

We used our CAD system to perform quantitative analysis of pulmonary nodules displayed on thin-section helical CT scans. If thin-section helical CT is employed for lung cancer screening, the detection and characterization of tumors can be achieved simultaneously, small faint nodules can be identified (52), and the misidentification of structures, such as blood vessels, as nodules (45) may cease. Thin-section helical CT, however, generates huge volumes of image data and increases the amount of radiation exposure to patients.

There were some limitations in our study. The nodules that were investigated were almost 3 cm in maximal diameter. We chose tumors that were less than 3 cm in diameter because this size corresponds with stage T1 tumors according to the Union Internationale Contre le Cancer classification. Patients whose nodules are 2 cm or larger and are without calcification usually undergo biopsy or surgery immediately (4), and most nodules that require CT for estimation of their benign or malignant nature are relatively small (2 cm in maximal diameter). We are preparing to conduct an observer performance test in which we include nodules with a diameter of only 2 cm or smaller. Also, only 33 lesions were presented in our observer performance test to allow individual observers to perform their interpretation in a relatively short time. On the other hand, 19 radiologists with different levels of experience participated in our study; this number of observers was sufficiently large for validation of the reliability of our statistical results.

In conclusion, use of our CAD system significantly improved the diagnostic performance of radiology residents but not of board-certified radiologists. Our future challenge is to increase the performance level of the CAD system to a level identical to, or higher than, the performance level of the board-certified radiologists.

Acknowledgments: The authors are grateful to Lorenzo Pesce, PhD (University of Chicago, Department of Radiology, Kurt Rossmann Laboratories, Chicago, Ill), for providing the computer program (LABMRMC, version 1.4B3) for statistical analysis and for helpful advice about the statistics.

References

1. Sone S, Li F, Yang ZG, et al. Results of 3-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. Br J Cancer 2001;84:25–32.
2. Sone S, Li F, Yang ZG, et al. Characteristics of small lung cancers invisible on conventional chest radiography and detected by population based screening using spiral CT. Br J Radiol 2000;73:137–145.
3. Sone S, Takashima S, Li F, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 1998;351:1242–1245.
4. Swensen SJ. CT screening for lung cancer. AJR Am J Roentgenol 2002;179:833–836.
5. Nawa T, Nakagawa T, Kusano S, Kawasaki Y, Sugawara Y, Nakata H. Lung cancer screening using low-dose spiral CT: results of baseline and 1-year follow-up studies. Chest 2002;122:15–20.
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. 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.
8. Diederich S, Wormanns D, Heindel W. Lung cancer screening with low-dose CT. Eur J Radiol 2003;45:2–7.
9. MacRedmond R, Logan PM, Lee M, Kenny D, Foley C, Costello RW. Screening for lung cancer using low dose CT scanning. Thorax 2004;59:237–241.
10. Henschke CI, Naidich DP, Yankelevitz DF, et al. Early lung cancer action project: initial findings on repeat screenings. Cancer 2001; 92:153–159.
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. Aoyama M, Li Q, Katsuragawa S, Li F, Sone S, Doi K. Computerized scheme for determination of the likelihood measure of malignancy for pulmonary nodules on low-dose CT images. Med Phys 2003;30:387–394.
13. Armato SG 3rd, Altman MB, La Riviere PJ. Automated detection of lung nodules in CT scans: effect of image reconstruction algorithm. Med Phys 2003;30:461–472.
14. Armato SG 3rd, 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.
15. Armato SG 3rd, Giger ML, MacMahon H. Automated detection of lung nodules in CT scans: preliminary results. Med Phys 2001; 28:1552–1561.
16. Armato SG 3rd, Giger ML, Moran CJ, Blackburn JT, Doi K, MacMahon H. Computerized detection of pulmonary nodules on CT scans. RadioGraphics 1999;19:1303–1311.
17. Armato SG 3rd, Li F, Giger ML, MacMahon H, Sone S, Doi K. Lung cancer: performance of automated lung nodule detection applied to cancers missed in a CT screening program. Radiology 2002;225:685–692.
18. Croisille P, Souto M, Cova M, et al. Pulmonary nodules: improved detection with vascular segmentation and extraction with spiral CT—work in progress. Radiology 1995; 197:397–401.
19. Giger ML, Bae KT, MacMahon H. Computerized detection of pulmonary nodules in computed tomography images. Invest Radiol 1994;29:459–465.
20. Goo JM, Lee JW, Lee HJ, Kim S, Kim JH, Im JG. Automated lung nodule detection at low-dose CT: preliminary experience. Korean J Radiol 2003;4:211–216.
21. Ko JP, Betke M. Chest CT: automated nodule detection and assessment of change over time—preliminary experience. Radiology 2001;218:267–273.
22. Ko JP, Rusinek H, Jacobs EL, et al. Small pulmonary nodules: volume measurement at chest CT—phantom study. Radiology 2003; 228:864–870.
23. Kostis WJ, Reeves AP, Yankelevitz DF, Henschke CI. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. IEEE Trans Med Imaging 2003;22:1259–1274.
24. Kostis WJ, Yankelevitz DF, Reeves AP, Fluture SC, Henschke CI. Small pulmonary nodules: reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up CT. Radiology 2004; 231:446–452.
25. Lawler LP, Wood SA, Pannu HK, Fishman

EK. Computer-assisted detection of pulmonary nodules: preliminary observations using a prototype system with multidetector-row CT data sets. J Digit Imaging 2003;16:251–261.

26. McNitt-Gray MF, Hart EM, Wyckoff N, Sayre JW, Goldin JG, Aberle DR. A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution CT: preliminary results. Med Phys 1999;26:880–888.

27. McNitt-Gray MF, Wyckoff N, Goldin JG, Suh R, Sayre JW, Aberle DR. Computer-aided diagnosis of the solitary pulmonary nodule imaged on CT: 2D, 3D and contrast enhancement features. In: Sonka M, Hanson KM, eds. Proceedings of SPIE: medical imaging 2001—image processing. Vol 4322. Bellingham, Wash: International Society for Optical Engineering, 2001; 1845–1852.

28. Reeves AP, Kostis WJ. Computer-aided diagnosis of small pulmonary nodules. Semin Ultrasound CT MR 2000;21:116–128.

29. 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.

30. Wyckoff N, McNitt-Gray MF, Goldin JG, Suh R, Sayre JW. Classification of solitary pulmonary nodule (SPNs) imaged on high resolution CT using contrast enhancement and three dimensional quantitative image features. In: Hanson KM, ed. Proceedings of SPIE: processing medical imaging 2000—image processing. Vol 3979. Bellingham, Wash: International Society for Optical Engineering, 2000; 1107–1115.

31. Zhao B, Gamsu G, Ginsberg MS, Jiang L, Schwartz LH. Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm. J Appl Clin Med Phys 2003;4:248–260.

32. Yankelevitz DF, Gupta R, Zhao B, Henschke CI. Small pulmonary nodules: evaluation with repeat CT—preliminary experience. Radiology 1999;212:561–566.

33. Kawata Y, Niki N, Ohmatsu H, et al. Classification of pulmonary nodules in thin-section CT images based on shape characterization. In: Proceedings of IEEE International Conference on Image Processing. Vol 3. Santa Barbara, Calif: Institute of Electrical and Electronics Engineers, 1997; 528–531.

34. Revel MP, Lefort C, Bissery A, et al. Pulmonary nodules: preliminary experience with three-dimensional evaluation. Radiology 2004;231:459–466.

35. Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B, Henschke CI. Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. Radiology 2000;217:251–256.

36. Awai K, Komi M, Hori S, Murao K, Ozawa A, Yamanaka T. Integrated CAD working on PACS for CT lung cancer screening: concepts and preliminary results (abstr). Radiology 2001; 221(P):692.

37. Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. Int J Comput Vis 1988;1:321–331.

38. Dougherty E. Digital image processing method. In: Thompson BJ, ed. Optical engineering. New York, NY: Dekker, 1994; 77–85.

39. Murao K, Awai K, Ozawa A, Yamanaka T, Komi M, Hori S. Three-dimensional analysis of pulmonary nodules using thin slice helical CT data: automated nodule classification and estimation of malignancy. In: Sixteenth International Congress and Exhibition: computed assisted radiology and surgery. Paris, France: Springer, 2002; 1099.

40. Gurney JW, Swensen SJ. Solitary pulmonary nodules: determining the likelihood of malignancy with neural network analysis. Radiology 1995;196:823–829.

41. Asada N, Doi K, MacMahon H, et al. Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung diseases: pilot study. Radiology 1990;177:857–860.

42. Kobayashi T, Xu XW, MacMahon H, Metz CE, Doi K. Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs. Radiology 1996;199:843–848.

43. Metz CE, Herman BA, Shen JH. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. Stat Med 1998;17:1033–1053.

44. Rockette HE, Gur D, Metz CE. The use of continuous and discrete confidence judgments in receiver operating characteristic studies of diagnostic imaging techniques. Invest Radiol 1992;27:169–172.

45. 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.

46. Cheung YC, Ng SH, Chang JW, Tan CF, Huang SF, Yu CT. Histopathological and CT features of pulmonary sclerosing haemangiomas. Clin Radiol 2003;58:630–635.

47. Cummings SR, Lillington GA, Richard RJ. Estimating the probability of malignancy in solitary pulmonary nodules: a Bayesian approach. Am Rev Respir Dis 1986;134:449–452.

48. Gurney JW. Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis. I. Theory. Radiology 1993;186:405–413.

49. Gurney JW, Lyddon DM, McKay JA. Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis. II. Application. Radiology 1993;186:415–422.

50. Gambhir SS, Shepherd JE, Shah BD, et al. Analytical decision model for the cost-effective management of solitary pulmonary nodules. J Clin Oncol 1998;16:2113–2125.

51. Cummings SR, Lillington GA, Richard RJ. Managing solitary pulmonary nodules: the choice of strategy is a "close call." Am Rev Respir Dis 1986;134:453–460.

52. Li F, Sone S, Abe H, MacMahon H, Armato SG 3rd, Doi K. Lung cancers missed at low-dose helical CT screening in a general population: comparison of clinical, histopathologic, and imaging findings. Radiology 2002;225:673–683.