Heart and Vessels

An International Journal

Springer

Heart Vessels (2004) 19:161–163
DOI 10.1007/s00380-003-0759-z

ORIGINAL ARTICLE

Yasutaka Itani · Shusuke Sone · Tomio Nakayama
Takaichiro Suzuki · Shigeru Watanabe · Ken-ichi Ito
Shodayu Takashima · Hajime Fushimi · Hideki Sanada

Coronary artery calcification detected by a mobile helical computed tomography unit and future cardiovascular death: 4-year follow-up of 6120 asymptomatic Japanese

Received: July 22, 2003 / Accepted: December 12, 2003

Abstract In the present study, we performed a prospective follow-up study in a population which underwent chest computed tomography (CT) screening. A total of 6120 participants underwent a chest CT medical examination for lung cancer and tuberculosis in Nagano Prefecture, Japan, between 1996 and 1997. Computed tomography scanning was performed from the apex of the lung to the diaphragm at a tube voltage of 120 kV and a tube current of 50 mA. We measured the CT density of the coronary arteries in 5–7 slices where coronary arteries were detected. The CT density threshold for determining coronary artery calcification (CAC) was above +110 HU. In 2000, we investigated the number of deaths due to cardiac and noncardiac disease among the participants. Of the 6120 participants, 14 died of cardiac disease (9, myocardial infarction; 4, heart failure; and 1, angina pectoris) and 64 died of other diseases. Coronary artery calcification was detected in 10 of the patients who died of cardiac disease, and in 31 of those who died of other diseases. The prevalence of CAC was higher in the former than in the latter (71.4% vs 48.4%, $P = 0.084$). The relative risk of CAC for cardiac death was 2.66 (95% confidence interval: 0.76, 9.37). The findings of this study suggested that CAC detected in a mass chest CT screening by a mobile helical CT unit was predictive of future cardiovascular death.

Key words Coronary artery calcification · Mobile helical computed tomography · Prospective study

Y. Itani (✉)
Itani Medical Clinic, 5-11-1 Shishibone, Edogawa-ku, Tokyo 133-0073, Japan
e-mail: itani-y@mtg.biglobe.ne.jp

S. Sone · K. Ito
Azumi General Hospital, Nagano, Japan

T. Nakayama · T. Suzuki
Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka, Japan

S. Watanabe
Utase Medical Clinic, Chiba, Japan

S. Takashima
Department of Radiology, Shinshu University School of Medicine, Nagano, Japan

H. Fushimi
Nagano Health Promotion Corporation, Nagano, Japan

H. Sanada
Nagano Prefecture Matsumoto Public Health Center, Nagano, Japan

Introduction

It is well known that a close relationship exists between coronary artery calcification (CAC) detected by computed tomography (CT) and coronary artery disease (CAD).[1–9] Recently, a mobile helical CT unit was developed in Japan and has been used in a mass chest CT screening for lung cancer and tuberculosis. It is effective in diagnosing not only respiratory diseases using the lung display, but also morphological abnormalities of cardiovascular organs using the mediastinum display. We have been conducting a mass chest CT screening for an asymptomatic population since 1995, and have reported that a strong relationship exists between CAC and a prior CAD event history (myocardial infarction, coronary angioplasty, and coronary artery bypass grafting) in males under the age of 60 years.[10] On the other hand, a previous study reported that the population with CAC had significantly higher morbidity of myocardial infarction than those without CAC in the asymptomatic population.[11] Therefore, in the present study we investigated the relationship between CAC and death due to cardiac disease in a 4-year follow-up in a population which underwent this CT examination.

Subjects and methods

A total of 6120 subjects (3377 men, 2743 women, mean age 61.4 ± 11.3 years) were invited to participate in a chest CT examination at a mobile helical CT unit to screen for lung cancer and tuberculosis. None of the participants com-

plained of chest symptoms, or had a prior CAD event history (myocardial infarction, coronary angioplasty, and coronary artery bypass grafting). All medical examinations were performed by the Nagano Anti-Tuberculosis Association between 1996 and 1997. Informed consent was obtained from all participants. The mobile helical CT unit used consists of a remodeled bus with a built-in helical CT scanner, CT-W950SR (Hitachi Medical, Tokyo, Japan). Computed tomography scanning was performed at a tube voltage of 120kV and a tube current of 50mA, with 10-mm consecutive sections, 10mm/s table speed, and slight breathing. The scanning range was from the apex of the lung to the diaphragm, and it took about 30s to scan the entire area. Radiation exposure was 3.6mSv. All CT data were stored on compact discs, and monitored using a personal computer. TM viewer software (Kissei Comtec, Nagano, Japan) was used to analyze the CT data. We measured the CT density of coronary arteries in 5–7 slices where coronary arteries were detected. Coronary artery calcification was defined as above +110 HU of the CT density, given that +110 HU was the lowest density at which CAC could be detected visually on the monitor of the personal computer. Between May and October 2000, we examined the death certificates of participants at the public offices of Nagano Prefecture with permission from the Ministry of Public Management of Japan. Accidental death and suicides were excluded from the present study. All statistical analyses were performed using a personal computer with SPSS for Windows version 10.0 (SPSS Japan). The χ^2 test was used to investigate the relationships between CAC and cardiac, and CAC and noncardiac death. A P value of less than 0.05 was considered statistically significant.

Results

The prevalence of CAC among the 6120 participants was 19.7% (men: 24.6%; women; 13.7%). Fourteen patients (mean age 72.1 ± 6.2 years) died of cardiac disease, and 9 of myocardial infarction, 4 of cardiac failure, and 1 of angina pectoris. Coronary artery calcification was detected in 10 of these 14 patients (Table 1). On the other hand, 64 patients (mean age 71.0 ± 8.4 years) died of other diseases, and CAC was detected in 31 of these patients. The prevalence of CAC was higher in those who died of cardiac death than in those with noncardiac death (71.4% vs 48.4%, $P = 0.084$) (Fig. 1). The mean time interval from the CT examination to death was 13.9 months in the former and 21.0 months in the latter. The relative risk of CAC for cardiac death was 2.66 [95% confidence interval (CI): 0.76, 9.37].

Discussion

Several previous studies have reported on the relationship between CAC and the long-term prognosis. Naito et al. investigated a 4-year follow-up in a population without chest pain who underwent a plain chest conventional CT scan. In this study, the population with CAC had significantly higher morbidity of myocardial infarction than those without CAC (men: 5.5% vs 0%; women: 3.7% vs 0%; $P < 0.05$). In terms of mortality, in men there was no significant difference between the two groups (13% vs 12%). However, in women, the population with CAC had significantly higher mortality than those without CAC (26% vs 8.9%, $P < 0.05$).[11]

In the study of Arad et al. using electron-beam CT (EBCT), a 19-month follow-up of 1173 asymptomatic subjects was performed. During the follow-up periods, there were 7 cases of nonfatal myocardial infarction, 1 cardiac death, 1 thromboembolic stroke, 8 coronary artery bypass procedures, and 9 cases of coronary angioplasty in 18 patients. The total calcium score (TCS) was significantly higher for the patients with cardiac events than for those without events ($P < 0.0001$).[12] It was also reported that the incidence of nonfatal myocardial infarction plus coronary death and the incidence of all coronary events increased as TCS increased,[13] and that patients with a higher TCS

Table 1. Patients who died of cardiac disease in this study

Patient no.	Age (years)	Sex	Cause of death	CAC
1	61	M	Acute myocardial infarction	LAD
2	71	F	Acute myocardial infarction	LAD
3	72	M	Cardiac failure	LAD, RCA
4	80	M	Cardiac failure	LAD, RCA
5	69	M	Acute myocardial infarction	–
6	82	M	Acute myocardial infarction	LAD, LCx
7	69	M	Acute myocardial infarction	LAD
8	71	F	Acute myocardial infarction	RCA
9	71	M	Acute myocardial infarction	LAD
10	76	F	Acute myocardial infarction	–
11	78	M	Cardiac failure	LAD, LCx, RCA
12	71	M	Cardiac failure	–
13	61	M	Acute myocardial infarction	–
14	77	M	Angina pectoris	LAD, RCA

CAC, coronary artery calcification; LAD, left anterior descending branch; LCx; left circumflex branch; RCA, right coronary artery; LMT, left main trunk

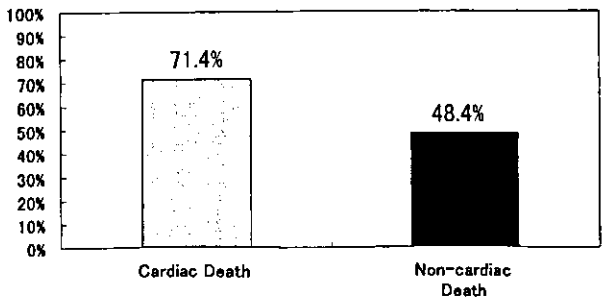

Fig. 1. Percentage of cardiac and noncardiac deaths in this study

(≧100) had a significantly poorer long-term prognosis than those with a lower TCS (>100).[14] The relative risk for predicting CAD by detecting CAC was 2.83 (95% CI: 0.38, 21.05).[15] These findings suggest that CAC is highly predictive of future cardiac events and adds valuable prognostic information. In the present study, we performed CT scanning at a lower tube voltage (50 mA), longer scan times, and with a thicker slice than those of EBCT. We could not measure the area of calcification to calculate TCS, given that cardiac motion artifact was not prevented under these CT scanning conditions. Therefore, we must conclude that our method is less sensitive in detecting CAC than EBCT. It is thought that this is the most important study limitation. However, the relative risk of CAC for cardiac death in the present study is almost the same as that of a previous study.[15] Accordingly, we suggest that CAC detected by the mobile helical CT unit may be also predictive of a future cardiac event. If CAC is detected in a mass chest CT screening, we recommend persons with CAC to undergo secondary examinations (e.g. treadmill test, ergometer, and exercise stress scintigram test) for the diagnosis of CAD. We further recommend coronary angiography (CAG) for all persons who show positive signs on secondary examinations.

In 6971 persons who underwent the mass chest CT examination, we detected 11 cases (0.16%) with asymptomatic thoracic aortic aneurysm.[16] We previously showed that secondary examinations would be required to diagnose CAD in the population with CAC. However, aortic aneurysm can be diagnosed directly by CT. Therefore, we suggest that it is more effective to diagnose aortic aneurysms using CT than predicting CAD by detecting CAC. According to the results of our studies, we conclude that the mobile helical CT unit is also useful to screen for cardiovascular diseases. If the CT screening for lung and cardiovascular organs is performed simultaneously, great advancement in the prevention of disease can be expected.

References

1. Masuda Y, Naito S, Takasu J, Watanabe S, Morooka N, Inagaki Y (1990) Coronary artery calcification detected by CT. Clinical significance and angiographic correlates. Angiology 41:1037–1047
2. Masuda Y, Naito S, Aoyagi Y, Watanabe S, Morooka N, Inagaki Y (1991) Clinical importance of coronary calcification detected by CT. Acta Cardiol 26:51–56
3. Naito S, Watanabe S, Masuda Y, Morooka N, Aoyagi Y, Inagaki Y (1992) Evaluation of coronary calcification by computed tomography. J Jpn Coll Angiol 32:715–722
4. Agaston A, Janowitz W, Viamonte M, Gasso J, Kaplan G (1990) Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol 15:827–832
5. Tannenbaum S, Kondos G (1989) Detection of calcific deposits in coronary arteries by ultrafast computed tomography and correlation with angiography. Am J Cardiol 63:837–871
6. Breen J, Sheedy P, Schwarz R, Stanton A, Kaufman R, Moll P (1992) Coronary artery calcification detected with ultrafast CT as an indication of coronary artery disebÉe. Radiology 185:435–439
7. Rumberger J, Sheedy P, Breen J, Schwarz R, Fitzpatrick L (1995) Coronary calcium as determined by electron beam computed tomography, and coronary disease on arteriogram effect of patients sex on diagnosis. Circulation 91:1363–1367
8. Shemesh J, Apter S, Rozenman J, Lusky A, Rath S, Motro S (1995) Calcification of coronary arteries: detection and quantification with double-helix CT. Radiology 197:779–783
9. Brodelick L, Shemesh J, Wilensky R, Eckert G, Zhou X, Torres W (1996) Measurement of coronary artery calcium with dual-slice helical CT compared with coronary angiography: evaluation of CT: scoring methods, interobserver variations, and reproducibity. Am J Radiol 167:439–444
10. Itani Y, Watanabe S, Masuda Y, Hanamura K, Asakura K, Sone S (2001) Coronary artery calcification by a mobile helical CT unit in a mass screening: the frequency and relationship to coronary risk factors and coronary artery disease. Chiba Med J 77:123–131
11. Naito S, Watanabe S, Masuda Y, Morooka N, Aoyagi Y, Inagaki Y (1990) Progression to ischemic heart disease in subjects with coronary artery calcification as evaluated by computed tomography. J Cardiol 20:249–258
12. Arad Y, Spadaro L, Goodman K, Liedo-Perez A, Sherman S, Lerner G (1996) Predictive value of electron beam computed tomography of the coronary arteries 19-month follow-up of 1173 asymptomatic subjects. Circulation 93:1951–1953
13. Arad Y, Spadaro L, Goodman K, Newstein D, Guerci A (2000) Prediction of coronary events with electron beam tomography. J Am Coll Cardiol 36:1253–1260
14. Detrano R, Hsiai T, Wang S, Puentes G, Fallavollita J, Shields P (1996) Prognostic value of coronary calcification and angiographic stenoses in patients undergoing coronary angiography. J Am Coll Cardiol 27:285–290
15. Keelan P, Bielak L, Ashai K, Jamjoum L, Denktas A, Rumberger J (2001) Long-term prognostic value of coronary calcification detected by electron-beam computed tomography in patients undergoing coronary angiography. Circulation 104:412–417
16. Itani Y, Watanabe S, Masuda Y, Hanamura K, Asakura K, Sone S (2002) Measurement of aortic diameters and detection of asymptomatic aortic aneurysm in a mass screening program using a mobile helical computed tomography unit. Heart Vessels 16:42–45

Springer-Verlag Tokyo, Incorporated

3-13, Hongo 3-chome, Bunkyo-ku, Tokyo 113-0033, Japan • Telephone (03)3812-0617 Fax (03)3812-4699

Feng Li, MD, PhD
Shusuke Sone, MD
Hiroyuki Abe, MD, PhD
Heber MacMahon, MD
Kunio Doi, PhD

Malignant versus Benign Nodules at CT Screening for Lung Cancer: Comparison of Thin-Section CT Findings[1]

Index terms:
Cancer screening
Computed tomography (CT), thin-section
Lung neoplasms, CT, 60.1211
Lung neoplasms, diagnosis, 60.31, 60.32

Published online before print
10.1148/radiol.2333031018
Radiology 2004; 233:793–798

Abbreviations:
GGO = ground-glass opacity
PPV = positive predictive value

[1] From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637 (F.L., H.A., H.M., K.D.), and J. A. Azumi General Hospital, Ikeda, Nagano, Japan (S.S.). From the 2002 RSNA scientific assembly. Received June 30, 2003; revision requested September 9; final revision received February 27, 2004; accepted April 12. Supported in part by USPHS grant CA62625. Address correspondence to F.L. (e-mail: *fli@kurt.bsd.uchicago.edu*).

H.M. and K.D. are shareholders of R2 Technology, Los Altos, Calif. K.D is a shareholder of Deus Technology, Rockville, Md.

Author contributions:
Guarantors of integrity of entire study, F.L., S.S., K.D.; study concepts and design, F.L., K.D.; literature research, H.M., K.D.; clinical studies, F.L., S.S., H.A., H.M.; data acquisition, F.L., S.S.; data analysis/interpretation, F.L., H.A., H.M., K.D.; statistical analysis, F.L.; manuscript preparation, F.L.; manuscript definition of intellectual content, F.L., H.M., K.D.; manuscript editing, K.D., H.M.; manuscript revision/review and final version approval, all authors.

PURPOSE: To evaluate thin-section computed tomographic (CT) characteristics of malignant nodules on the basis of overall appearance (pure ground-glass opacity [GGO], mixed GGO, or solid opacity) in comparison with the appearance of benign nodules.

MATERIALS AND METHODS: Institutional review board approval and patient consent were obtained. Follow-up diagnostic CT was performed in 747 suspicious pulmonary nodules detected at low-dose CT screening (17 892 examinations). Of 747 nodules, 222 were evaluated at thin-section CT (1-mm collimation), which included 59 cancers and 163 benign nodules (3–20 mm). Thin-section CT findings of malignant versus benign nodules with pure GGO (17 vs 12 lesions), mixed GGO (27 vs 29 lesions), or solid opacity (15 vs 122 lesions) were analyzed. Fisher exact test for independence was used to compare differences in shape, margin, and internal features between benign and malignant nodules. Positive predictive value (PPV) was analyzed when a category was significantly different from the others.

RESULTS: Among nodules with pure GGO, a round shape was found more frequently in malignant lesions (11 of 17, 65%) than in benign lesions (two of 12, 17%; $P = .02$; PPV, 85%); mixed GGO, a subtype with GGO in the periphery and a high-attenuation zone in the center, was seen much more often in malignant lesions (11 of 27, 41%) than in benign lesions (two of 29, 7%; $P = .004$; PPV, 85%). Among solid nodules, a polygonal shape or a smooth or somewhat smooth margin was present less frequently in malignant than in benign lesions (polygonal shape: 7% vs 38%, $P = .02$; smooth or somewhat smooth margin: 0% vs 63%, $P < .001$), and 98% (46 of 47) of polygonal nodules and 100% (77 of 77) of nodules with a smooth or somewhat smooth margin were benign.

CONCLUSION: Recognition of certain characteristics at thin-section CT can be helpful in differentiating small malignant nodules from benign nodules.

Computed tomographic (CT) screening has increased the detection rate of early peripheral lung cancer in the United States and Japan (1,2). The results of the Early Lung Cancer Action Project, or ELCAP (1), suggested that nodules with pure (nonsolid) or mixed (partially solid) ground-glass opacity (GGO) at thin-section CT are more likely to be malignant than are those with solid opacity; among 44 nodules with GGO (19% of 233 nodules identified at baseline screening), 15 (34%) were confirmed to be malignant. On the other hand, most of the benign lesions were solid at CT, although some (approximately 15%) contained elements of GGO. According to the ELCAP data, 18% of nodules (five of 28) with pure GGO were malignant and 63% of nodules (10 of 16) with mixed GGO were malignant (1). To our knowledge, there are no previous studies that specifically compare thin-section CT characteristics between malignant lesions and benign lesions with pure GGO, mixed GGO, and solid opacity.

A 3-year lung cancer screening program has recently been completed in Japan by using low-dose CT and follow-up thin-section CT. We have previously reported that among 59

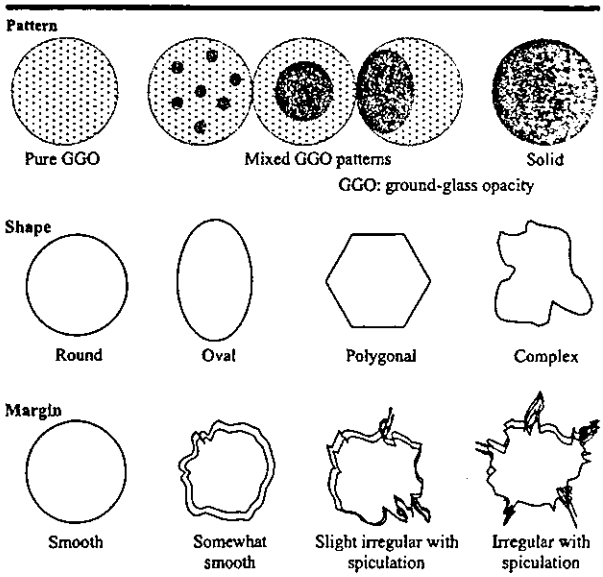

Figure 1. Typical appearance of the three patterns, four shapes, and four margins used to classify lesions in this study.

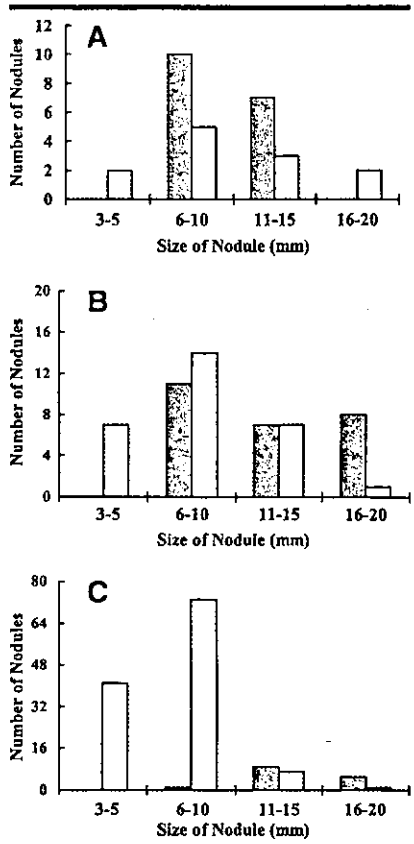

Figure 2. Graphs show distribution of sizes among, *A*, 29 nodules with pure GGO (17 malignant and 12 benign); *B*, 56 nodules with mixed GGO (27 malignant and 29 benign); and *C*, 137 nodules with solid opacity (15 malignant and 122 benign). Gray bars = malignant nodules, white bars = benign nodules. For pure and mixed GGO lesions, the size of benign nodules extensively overlaps that of malignant nodules in the 6–15-mm range.

small (6–20 mm) lung adenocarcinomas, only 16 nodules (27%) showed solid opacity and the rest (73%) showed pure or mixed GGO at thin-section CT in this screening program (2). In another study (3), thin-section CT characteristics were compared between 25 very small (≤10 mm) cancers, 24 of which were adenocarcinomas, and 40 benign lesions, most of which were solid nodules. We found that by using a single CT feature, namely polygonal shape, and a three-dimensional ratio (maximum transverse diameter to maximum z-axis dimension of a lesion, which was measured as the difference between the cephalic extent and the caudal extent of the lesion in coronal reformation) greater than 1.78, 100% specificity was shown for benign nodules (3). However, these features were not necessarily applicable to benign lesions with GGO, especially not to those larger than 10 mm. Thus, the purpose of our study was to evaluate the thin-section CT characteristics of malignant nodules on the basis of the overall appearance (pure GGO, mixed GGO, or solid opacity) compared with the appearance of benign nodules.

MATERIALS AND METHODS

Study Nodules

From May 1996 to March 1999, 17 892 examinations were performed in 7847 individuals (4288 men and 3559 women; mean age, 61 years) as part of an annual low-dose CT screening program for lung cancer in Nagano, Japan. A mobile unit equipped with a CT scanner (W950SR; Hitachi, Tokyo, Japan) was used to scan the chest with 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 of 2), 10-mm collimation, and a 10-mm reconstruction interval. The program was sponsored and supported by the Telecommunications Advancement Organization of Japan and was completed after 3 years. All subjects gave informed consent. Approval for review and research of the cases used in this study was obtained from our institutional review board at the University of Chicago.

Among those undergoing the examinations, 605 patients with 747 suspicious pulmonary nodules detected at low-dose CT underwent follow-up diagnostic CT. Diagnostic work-up CT, which included thin-section CT, was performed within 3 months of low-dose CT screening; follow-up CT examinations were performed at 3, 6, 12, 18, and 24 months, as needed. Most of the follow-up CT examinations were performed at Shinshu University Hospital, and some were performed at local hospitals. The results for follow-up work were accrued until December 1999.

The follow-up results for the 747 nodules include six categories, as follows: 76 primary lung cancers confirmed at biopsy; 11 atypical adenomatous hyperplasias confirmed at biopsy; 444 lesions, which included 167 resolved nodules, 230 nodules that were stable for 2 years or more, 38 nodules with benign-pattern calcifications (diffuse, central, popcorn, and laminar or concentric calcification), and nine nodules that were resected and confirmed as benign; 27 nodules with findings suspicious for malignancy at thin-section CT but not confirmed at biopsy; 176 nodules suspected of being benign but with insufficient follow-up; and 13 indeterminate nodules.

For this study, we used a database of thin-section CT images obtained from Shinshu University Hospital as part of the Nagano CT screening program for lung cancer. A helical scanner (HiSpeed Advantage; GE Medical Systems, Milwaukee, Wis) was used for scanning the nodules with a 200-mA tube current, 1 second per tube rotation, table speed of 1

TABLE 1
Thin-Section CT Findings in Malignant versus Benign Lesions with Pure GGO

Feature	Malignant (n = 17)	Benign (n = 12)
Shape		
Round	11	2
Oval	3	1
Polygonal	0	3
Complex	3	6
Margin		
Smooth	1	0
Somewhat smooth	9	5
Slightly irregular with spiculation	7	7
Irregular with spiculation	0	0

TABLE 2
Thin-Section CT Findings in Malignant versus Benign Lesions with Mixed GGO

Feature	Malignant (n = 27)	Benign (n = 29)
Central opacity		
Present	11	2
Absent	16	27
Air component		
Present	16	9
Absent	11	20
Shape		
Round	10	2
Oval	1	2
Polygonal	3	9
Complex	13	16
Margin		
Smooth	0	0
Somewhat smooth	4	7
Slightly irregular with spiculation	9	14
Irregular with spiculation	14	8

TABLE 3
Thin-Section CT Findings in Malignant versus Benign Solid Nodules

Feature	Malignant (n = 15)	Benign (n = 122)
Air component		
Present	7	5
Absent	8	117
Shape		
Round	7	39
Oval	0	24
Polygonal	1	46
Complex	7	13
Margin		
Smooth	0	27
Somewhat smooth	0	50
Slightly irregular with spiculation	7	37
Irregular with spiculation	8	8

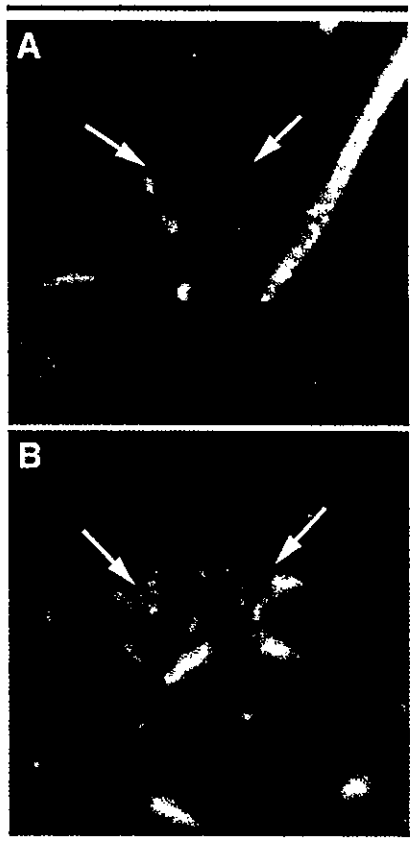

Figure 3. Transverse thin-section CT images. *A*, Image shows a malignant pure GGO lesion (adenocarcinoma) with a round shape (arrows). *B*, Image shows a benign pure GGO lesion (resolved within 3 months) with a polygonal-complex shape (arrows) that is confined to a secondary lobule.

mm/sec, 1-mm collimation, and 0.5-mm interval with a bone reconstruction algorithm. The database consisted of studies performed in 222 patients with 222 confirmed malignant or confirmed benign nodules, which were small in size (3–20 mm) on the first thin-section CT image obtained within 3 months of low-dose CT screening. Among the 222 patients, there were 14 patients with two nodules in different lung lobes, in which case the larger of the two nodules was selected for this study. Patients with two nodules in the same lung lobe and patients with more than two nodules were not included. On thin-section CT images, nonnodular lesions such as linear or scattered opacities, which had been regarded as suspicious on the original 10-mm collimation screening CT images, were excluded from the analysis. Nodules with benign-pattern calcifications were also excluded. This database contained cases of 96 pulmonary nodules that were used in two previous studies (2,3).

Among the 222 patients (mean age, 62.4 years; age range, 30–84 years), there were 119 men (mean age, 62.8 years; age range, 30–84 years) and 103 women (mean age, 61.9 years; age range, 34–75 years).

Data Analysis

Thin-section CT images for the 222 nodules were displayed and interpreted with use of "stacked" mode on a monochrome cathode ray tube monitor at a width and level of 1500 HU and −550 HU, respectively. The images of 222 nodules were randomly arranged for a reading sequence, and the final diagnosis for the nodules, which included the his-

topathologic results, was blinded to the radiologists. Three radiologists with 20, 18, and 17 years of experience in general radiology (F.L. and H.A. included) independently viewed these images and subjectively classified the nodules as one of three patterns: pure GGO, mixed GGO, or solid opacity. They also independently determined the overall shape (round, oval, polygonal, or complex) and margin (smooth, somewhat smooth, somewhat irregular with slight spiculation, or irregular with spiculation) of the nodules, as well as the internal features. Internal features included a specific mixed GGO pattern characterized by GGO in the periphery, with a high-attenuation zone in the center and the presence or absence of air (air bronchogram, cavitation, or focal emphysema) within the nodule on thin-section CT images. The typical appearance of the three patterns, four shapes, and four margins used to classify the lesions is illustrated in Figure 1.

For pattern, shape, and margins of the nodules, the same judgment was made by all three radiologists for 75%, 40%, and 13% of cases, respectively, and the same judgment was made by any two of the radiologists for 99%, 91%, and 73% of cases, respectively. Two radiologists (F.L., H.A.) worked together to reach a consensus for the remaining 83 features in 76 nodules; these nodules were initially classified differently by each of the three radiologists. For internal features, the same judgment was made by all three radiologists in 67% of cases and by any two radiologists in 100% of cases. The final decision regarding the CT findings was based on the consensus of at least two radiologists. The mean size (average length and width) and clinical outcome for 222 nodules were recorded by one radiologist (F.L.).

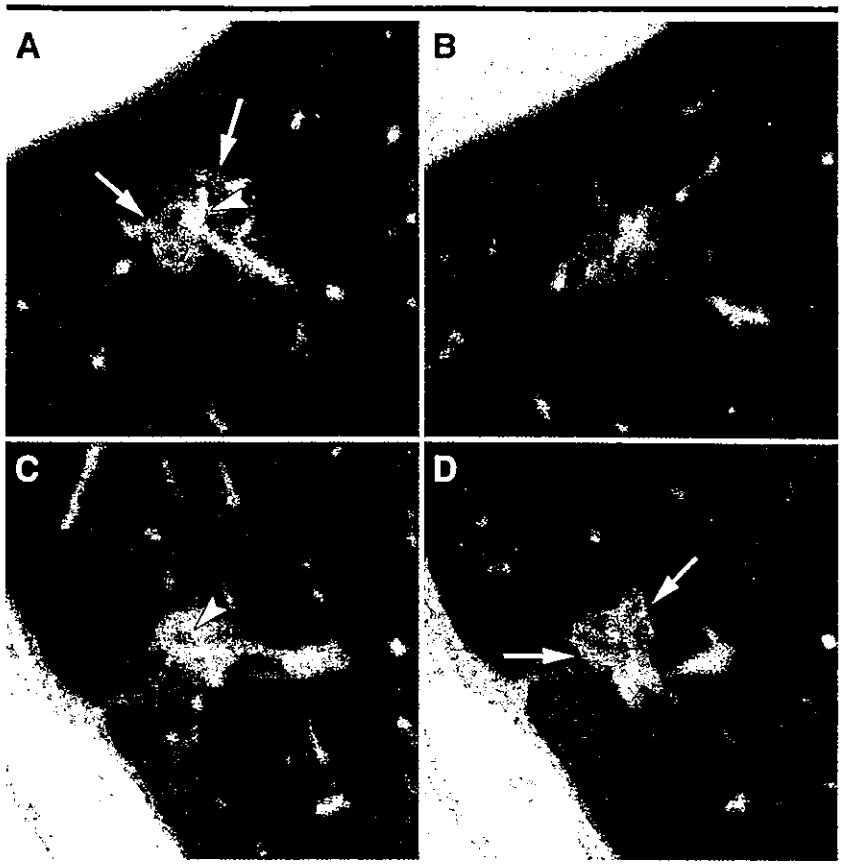

Figure 4. Transverse thin-section CT images. *A, B,* Images show a malignant mixed GGO lesion (adenocarcinoma) with irregular margins. The nodule shows both GGO in the periphery (arrows) and a high-attenuation zone (arrowhead) in the center. *C, D,* Images show a benign mixed GGO lesion (nodular fibrosis) with irregular margins. In *C,* a small air collection (arrowhead) is seen in the nodule. In *D,* the nodule (arrows) is seen on another section.

Statistical Analysis

Statistical analysis was performed by using the Student *t* test for comparison of differences in size between benign and malignant nodules. The χ^2 test for independence was used independently for comparison of the differences in patterns (nodules with and those without GGO) between the benign nodules and the malignant nodules. The data presented in Tables 1–3 were analyzed first by using the Fisher exact test for independence to determine whether there were any significant differences in the proportion of malignant lesions and benign lesions in the categories of shape, margin, and internal features. If such differences were established (the difference was significant at $P \leq .05$), additional Fisher exact tests were performed to determine which categories were significantly different from the others. Fisher exact test was used instead of χ^2 test because of the small sample size. Positive predictive value (PPV) was further analyzed when a category was significantly different from the others.

RESULTS

Of the 222 patients evaluated, 59 (27 men and 32 women; mean age, 64.6 years) had malignant nodules and 163 (92 men and 71 women; mean age, 61.6 years) had benign nodules. The mean size of the 59 malignant nodules (12.3 mm) was larger than that of the 163 benign nodules (7.2 mm, $P < .001$). Among 59 malignant nodules, there were 17 with pure GGO, 27 with mixed GGO, and 15 with solid opacity. Among 163 benign nodules, 12 showed pure GGO, 29 showed mixed GGO, and 122 showed solid opacity. The number of lesions with GGO was greater in the group of malignant nodules than in the group of benign nodules ($P < .001$).

All 17 malignant nodules with pure GGO were well-differentiated adenocarcinomas. Among 27 malignant nodules with mixed GGO, 26 were well-differentiated adenocarcinomas and one was a moderately differentiated adenocarcinoma. Of the 15 malignant nodules with solid opacity, four were well-differentiated adenocarcinomas, seven were other adenocarcinomas, two were squamous cell carcinomas, and two were small cell carcinomas. All 12 benign nodules with pure GGO had resolved at the 3-month follow-up examination. Among 29 benign nodules with mixed GGO, nodular fibrosis was confirmed at surgery in three cases, was re-

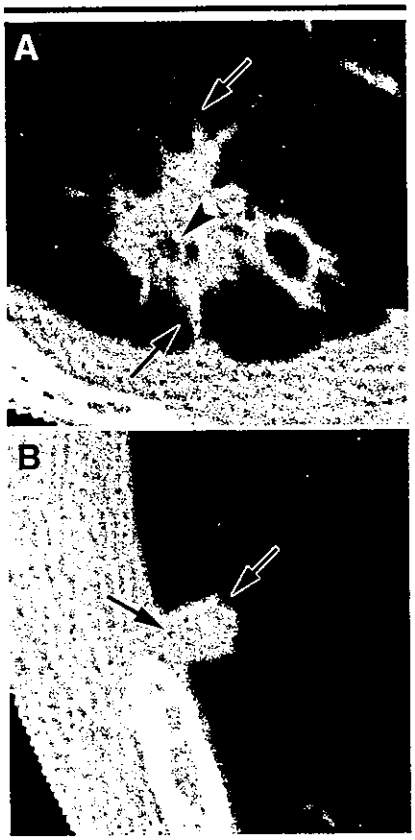

Figure 5. Transverse thin-section CT images. *A*, Image shows a malignant nodule (squamous cell carcinoma) with air components (arrowhead) and an irregular margin and gross spiculation (arrows). *B*, Image shows a small benign solid nodule (stable for more than 2 years) with a polygonal shape (arrows) and somewhat smooth margin.

solved at 3 months or more of follow-up in 17 cases, and showed no change for 2 years or more in nine cases. Among the 122 benign solid nodules, five cases (one case each of inflammatory granuloma, cryptococcoma, focal organizing pneumonia, inflammatory pseudotumor, and sclerosing hemangioma) were confirmed at surgery, 19 cases were resolved at 3 months or more of follow-up, and 98 cases showed no change for 2 years or more. All malignant nodules were confirmed at surgery.

The distribution of sizes among 29 nodules with pure GGO, 56 with mixed GGO, and 137 with solid opacity is shown in Figure 2. For GGO lesions, there was extensive overlap between the size of benign nodules and that of malignant nodules. On the other hand, for solid lesions, there was relatively limited overlap between the size of benign nodules and that of malignant nodules.

Table 1 shows the thin-section CT findings for malignant versus benign lesions with pure GGO; Figure 3 shows a malignant nodule and a benign nodule with pure GGO obtained at thin-section CT. The overall Fisher exact test indicated a significant association between lesion shape and malignancy ($P = .008$) but indicated no significant association between margins and malignancy ($P = .826$). At further examination, we found a significant association between malignancy and round nodules ($P = .022$); the number of round nodules was greater in the malignant group (65%, 11 of 17) than in the benign group (17%, two of 12) of pure GGO lesions. If round shape was used to discriminate between malignant lesions and benign lesions with pure GGO, the PPV (probability that a nodule is malignant, given that it is round) of such a test would be 85% (95% confidence interval: 54.55%, 98.08%) in this data set.

Table 2 lists thin-section CT findings for malignant lesions versus benign lesions with mixed GGO; Figure 4 shows a malignant nodule and a benign nodule with mixed GGO obtained at thin-section CT. The overall Fisher exact test again showed a significant association between nodule shape and malignancy ($P = .020$) but showed no significant association between margins and malignancy ($P = .174$). The association between round nodules and malignancy was found to be significant ($P = .009$), and the proportion of round nodules was higher among malignant lesions (37%, 10 of 27) than among benign lesions (7%, two of 29). The PPV was 83% (95% confidence interval: 51.59%, 97.91%). Furthermore, the presence of central opacity with mixed GGO was significantly associated with malignancy ($P = .004$), with a higher proportion of nodules with this feature in the malignant group (41%, 11 of 27) than in the benign group (7%, two of 29). The PPV of this test was 85% (95% confidence interval: 54.55%, 98.08%). However, the presence of air components within lesions was not significantly associated with malignancy ($P = .059$).

Table 3 lists thin-section CT findings for malignant versus benign solid nodules; Figure 5 shows malignant nodules and benign nodules obtained at thin-section CT. Fisher exact test showed a significant association between shape and malignancy ($P < .001$), as well as between margins and malignancy ($P < .001$). However, a round shape was not found to be associated with malignancy in solid nodules ($P = .262$), which is in contrast to the results found with pure and mixed GGO lesions. An oval shape was not significantly associated with malignancy ($P = .073$). The association between a complex shape and malignancy was found to be significant ($P = .002$)—the proportion of nodules with complex shape was higher among malignant lesions (47%, seven of 15) than among benign lesions (11%, 13 of 122). However, the PPV of this test was only 35% (95% confidence interval: 15.39%, 59.22%). The proportion of nodules with a polygonal shape was greater among benign lesions (38%, 46 of 122) than among malignant lesions (7%, one of 15; $P = .019$). There were 47 polygonal nodules, 46 (98%) of which were benign. When the margin classifications were dichotomized into "smooth or somewhat smooth" and "slightly irregular or irregular" categories, there was a significant difference between benign nodules and malignant nodules ($P < .001$). The proportion of smooth or somewhat smooth margins among malignant lesions was lower (0%, none of 15) than it was among benign lesions (63%, 77 of 122). There were 77 smooth or somewhat smooth nodules, and all 77 were benign. Furthermore, the presence of air components within these solid lesions was significantly associated with the malignant group (47%, seven of 15; $P < .001$) in comparison with the benign group (4%, five of 122). The PPV of this test was 58% (95% confidence interval: 27.67%, 84.83%).

DISCUSSION

Comparison of various CT features such as contour, margins, and internal characteristics of pulmonary nodules with pathologic specimens can be helpful for developing criteria to distinguish between cancers and benign lesions (1,2,4–7). In CT screening programs, however, most benign nodules are not confirmed at pathologic diagnosis. Because of this limitation, we were not able to make a detailed radiologic-pathologic comparison. Therefore, we chose to investigate two internal patterns, namely *(a)* nodules with both GGO in the periphery and a high-attenuation zone in the center and *(b)* nodules with an area of air, such as an air bronchogram, that is frequently found in small well-differentiated adenocarcinomas (2,5). Also, we classified all nodules into one of four subcategories of shape and margins on the basis of the predominant CT appearance. In our

study, we found that differences in the CT features between benign lesions and malignant lesions were observed for each of the three patterns on thin-section CT images.

Results of previous clinical CT studies (8–10) have shown that malignant nodules commonly contain solid opacity and that benign nodules have higher attenuation, often with visible calcifications, than do malignant nodules. Siegelman et al (10) reported that 61% of 279 benign nodules (including 153 nodules with diffuse calcifications) had smooth or moderately smooth margins and 65% of 283 primary malignant tumors had irregular shapes with spiculation. Kuriyama et al (5), in a study of 20 peripheral lung cancers and 20 benign nodules less than 20 mm in diameter, reported that an air bronchogram was not observed as frequently in small benign lesions, such as hamartoma and tuberculoma, as it was in adenocarcinomas.

The number of solid benign nodules was much greater than the total number of malignant nodules in our database, which was obtained from a lung cancer CT screening program, and the frequency of some features, such as internal air bronchograms, a complex shape, and an irregular margin, was much less in common in benign lesions than in malignant lesions. However, these observations do not necessarily mean that these features are reliable for differentiating benign nodules from malignant nodules, because the absolute numbers of benign nodules with such features may be comparable to the numbers of malignant nodules with similar features. For example, the frequency of an irregular margin in solid nodules was 7% (eight of 122) for benign nodules and 53% (eight of 15) for malignant nodules. However, if a radiologist encountered such a case in a screening examination, there would be an approximately 50% (eight of 16) likelihood that the lesion was malignant, if all other factors were equal. We found that a polygonal shape or a smooth or somewhat smooth margin (98%–100% likelihood of benignity) could be more helpful for differentiating solid benign nodules from malignant nodules than would internal air bronchograms, a complex shape, or an irregular margin.

There were some limitations to this study. For instance, no malignant lesions 5 mm or smaller were found; this is probably because the database used here was compiled from images obtained with low-dose single-detector row CT at a 10-mm section thickness. Second, many of the benign GGO lesions detected at the initial screening CT had resolved before thin-section diagnostic CT was performed. In a previous study, we reported that among 108 benign nodules (54, 27, and 27 of which showed pure GGO, mixed GGO, and solid opacity, respectively, at low-dose CT), 92 (85%) resolved within 3 months (11). Also, a large variance was noted in the judgment for CT features by three radiologists, especially for margins of the nodules; this is probably because most nodules used in current study were smaller than 10 mm.

The margins and size of nodules were not useful for differentiating benign from malignant GGO lesions in this series, and benign lesions with GGO were more difficult to distinguish from malignant nodules than were those with solid opacity. However, certain features, such as a round shape or a combination of GGO in the periphery with a high-attenuation zone in the center, were observed much more frequently in malignant GGO nodules. Therefore, we believe that familiarity with the different features of benign nodules and malignant nodules can be useful to radiologists in the management of indeterminate nodules. Also, short-term follow-up imaging can be helpful for differentiating benign from malignant nodules with GGO patterns, because all 12 of the benign pure GGO lesions in this series, as well as the majority of benign mixed GGO lesions, had partially or completely resolved within 3 months.

Acknowledgments: The authors are grateful to Chaotong Zhang, MD, for participating as an image reviewer, Qiang Li, PhD, for helpful suggestions, Masha Kocherginsky, PhD, for assistance with statistical analysis, Roger Engelmann, MS, for his useful work on the display software, and Elisabeth Lanzl for the English editing of the manuscript.

References

1. Henschke CI, Yankelevitz DF, Mirtcheva R, et al. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. AJR Am J Roentgenol 2002; 178:1053–1057.
2. Yang ZG, Sone S, Takashima T, et al. High-resolution CT analysis of small peripheral lung adenocarcinomas revealed on screening helical CT. AJR Am J Roentgenol 2001; 176:1399–1407.
3. Takashima S, Sone S, Li F, et al. Small solitary pulmonary nodules (< or = 1 cm) redetected at population-based CT screening for lung cancer: reliable high-resolution CT features of benign lesions. AJR Am J Roentgenol 2003; 180:955–964.
4. Zwirewich CV, Vedal S, Miller RR, Müller NL. Solitary pulmonary nodule: high-resolution CT and radiologic-pathologic correlation. Radiology 1991; 179:469–476.
5. Kuriyama K, Tateishi R, Doi O, et al. Prevalence of air bronchograms in small peripheral carcinomas of the lung on thin-section CT. AJR Am J Roentgenol 1991; 156:921–924.
6. Li F, Sone S, Takashima S, et al. Correlations between high-resolution computed tomographic, magnetic resonance and pathological findings in cases with non-cancers but suspicious lung nodules. Eur Radiol 2000; 10:1782–1789.
7. Kohno N, Ikezoe J, Johkoh T, et al. Focal organizing pneumonia: CT appearance. Radiology 1993; 189:119–123.
8. Proto AV, Thomas SR. Pulmonary nodules studied by computed tomography. Radiology 1985; 156:149–153.
9. Zerhouni EA, Stitik FP, Siegelman SS, et al. CT of the pulmonary nodule: a cooperative study. Radiology 1986; 160:319–327.
10. Siegelman SS, Khouri NF, Leo FP, Fishman EK, Braverman RM, Zerhouni EA. Solitary pulmonary nodules: CT assessment. Radiology 1986; 160:307–312.
11. Li F, Sone S, Takashima S, Maruyama Y, Hasegawa M, Yang ZG. Roentgenologic analysis of 108 non-cancerous focal lung lesions detected in screening CT for lung cancer. Jpn J Lung Cancer 1999; 39:369–380.

Radiologists' Performance for Differentiating Benign from Malignant Lung Nodules on High-Resolution CT Using Computer-Estimated Likelihood of Malignancy

Feng Li[1]
Masahito Aoyama[2]
Junji Shiraishi[1]
Hiroyuki Abe[1]
Qiang Li[1]
Kenji Suzuki[1]
Roger Engelmann[1]
Shusuke Sone[3]
Heber MacMahon[1]
Kunio Doi[1]

OBJECTIVE. The purpose of our study was to evaluate whether a computer-aided diagnosis (CAD) scheme can assist radiologists in distinguishing small benign from malignant lung nodules on high-resolution CT (HRCT).

MATERIALS AND METHODS. We developed an automated computerized scheme for determining the likelihood of malignancy of lung nodules on multiple HRCT slices; the likelihood estimate was obtained from various objective features of the nodules using linear discriminant analysis. The data set used in this observer study consisted of 28 primary lung cancers (6–20 mm) and 28 benign nodules. Cancer cases included nodules with pure ground-glass opacity, mixed ground-glass opacity, and solid opacity. Benign nodules were selected by matching their size and pattern to the malignant nodules. Consecutive region-of-interest images for each nodule on HRCT were displayed for interpretation in stacked mode on a cathode ray tube monitor. The images were presented to 16 radiologists—first without and then with the computer output—who were asked to indicate their confidence level regarding the malignancy of a nodule. Performance was evaluated by receiver operating characteristic (ROC) analysis.

RESULTS. The area under the ROC curve (A_z value) of the CAD scheme alone was 0.831 for distinguishing benign from malignant nodules. The average A_z value for radiologists was improved with the aid of the CAD scheme from 0.785 to 0.853 by a statistically significant level ($p = 0.016$). The radiologists' diagnostic performance with the CAD scheme was more accurate than that of the CAD scheme alone ($p < 0.05$) and also that of radiologists alone.

CONCLUSION. CAD has the potential to improve radiologists' diagnostic accuracy in distinguishing small benign nodules from malignant ones on HRCT.

CT screening has led to early detection of peripheral lung cancer and also detection of a large number of false-positives (i.e., noncalcified benign nodules) [1–5]. The false-positive rate at screening has been reported as 87–93% with low-dose single-detector CT at 10-mm slice thickness [1–3] and 98–99% with single-detector CT or MDCT at 5-mm slice thickness [4, 5]. Also, simultaneous or additional diagnostic high-resolution CT (HRCT) is needed for the distinction between malignant and benign lung nodules detected as suspicious or indeterminate lesions on screening CT [1–5]. This high false-positive rate because of benign nodules is likely to reduce the benefit of CT screening for early detection of lung cancer [6]. Therefore, it is important to differentiate benign from malignant nodules to reduce the number of false-positive findings on screening CT and to reduce follow-up examinations for diagnostic HRCT.

We developed an automated computerized scheme [7] for determination of the likelihood measure of malignancy by using various objective features of the nodules in our a database of thick-section low-dose CT; one or two slices were used for image analysis of each nodule. The low-dose CT database consisted of 489 nodules obtained from a mass screening for lung cancer in Nagano, Japan [2]. All of these nodules were considered as suspicious or indeterminate lesions when detected by radiologists on low-dose CT screening. With the use of receiver operating characteristic (ROC) analysis, our computerized scheme achieved an area under the ROC curve (A_z value) of 0.846 for distinction between 76 malignant and 413 benign lung nodules.

Recently, we further developed another computerized scheme for distinction be-

Received November 10, 2003; accepted after revision January 30, 2004.

Supported in part by USPHS grant CA62625.

H. MacMahon and K. Doi are shareholders of R2 Technology, Inc., Los Altos, CA. K. Doi is a shareholder of Deus Technology, Inc., Rockville, MD.

[1]Department of Radiology, Kurt Rossmann Laboratories for Radiologic Image Research, MC-2026, The University of Chicago, 5841 S Maryland Ave., Chicago, IL 60637. Address correspondence to F. Li (fli@kurt.bsd.uchicago.edu).

[2]Department of Intelligent Systems, Faculty of Information Sciences, Hiroshima City University, Hiroshima 731-3194, Japan.

[3]Azumi General Hospital, Ikeda, Nagano 399-8695, Japan.

AJR 2004;183:1209–1215

0361–803X/04/1835–1209

tween malignant and benign lesions derived from multiple slices of HRCT (1-mm collimation) based on 2D and 3D volume data. The HRCT database consisted of 244 small noncalcified (3–20 mm) nodules obtained as part of follow-up diagnostic work for suspicious or indeterminate lesions detected on low-dose CT in the same screening program.

In the present study, we assessed observer performance using ROC analysis to evaluate the effectiveness of our computer-aided diagnosis (CAD) scheme to assist radiologists in distinguishing small benign from malignant lung nodules in various patterns at HRCT. The malignant lung cancers included nodules with pure ground-glass opacity, mixed ground-glass opacity, and solid opacity; the benign nodules were selected by matching their size and pattern to the cancers on HRCT in this observer study.

Materials and Methods

Our institutional review board approved the use of this database and the participation of radiologists in this observer performance study. Informed consent for use of cases was waived. Informed consent for the observer performance study was obtained from all observers.

Database

The diagnostic HRCT database used in this study consisted of 59 patients (27 men, 32 women; mean age, 64.6 years) with 61 malignant nodules and 169 patients (99 men, 70 women, mean age 61.6 years) with 183 benign nodules. The database was obtained as part of an annual 3-year low-dose CT screening for lung cancer from 17,892 examinations on 7,847 individuals in Nagano, Japan [2]. HRCT scans were obtained on a helical scanner (HiSpeed Advantage, GE Healthcare) with a standard tube current (200 mA) to cover the entire nodule lesion, 1-mm collimation, and a bone reconstruction algorithm with a 0.5-mm interval.

Two features concerning the size and pattern type of the pulmonary nodules on HRCT were subjectively determined by radiologists for the purpose of grouping nodules in our database. The mean size (average length and width) was recorded by one radiologist. The three types of patterns of these nodules—pure ground-glass opacity, mixed ground-glass opacity, and solid opacity—were viewed independently and grouped by three radiologists without knowledge of the final diagnosis, and a consensus was reached through discussion. Nodules with benign-pattern calcifications (diffuse, central, popcorn, and laminar, or concentric calcification) were excluded. The range of nodule sizes for the 61 malignant and 183 benign nodules was 6–19 mm (mean, 12 mm) and 3–20 mm (mean, 7 mm), respectively. Among the 61 malignant nodules, there were 18 nodules with pure ground-glass opacity, 28 with mixed ground-glass opacity, and 15 with solid opacity, whereas 183 benign nodules included 12 with pure ground-glass opacity, 30 with mixed ground-glass opacity, and 141 with solid opacity.

All malignant nodules were primary lung cancers confirmed by surgery, including 49 well-differentiated adenocarcinomas, eight other adenocarcinomas, two squamous cell carcinomas, and two localized small cell carcinomas. Among the 183 benign nodules, nine (four cases of nodular fibrosis; and one case each of inflammatory granuloma, cryptococcoma, focal organizing pneumonia, inflammatory pseudotumor, and sclerosing hemangioma) were confirmed by surgery, 51 had resolved at follow-up examination, and 123 showed no change for 2 or more years.

CAD

With our CAD scheme, the nodules were segmented automatically using a dynamic programming technique [7]. Forty-one and 15 image features based on 2D and 3D volume data, respectively, were determined from quantitative analysis of the nodule outline and pixel values. Linear discriminant analysis was used to distinguish benign from malignant nodules. The performance of this CAD scheme was evaluated on the basis of a leave-one-out testing method by use of 61 malignant and 183 benign lung nodules. For the input of the linear discriminant analysis, we selected many combinations from 56 features and two clinical parameters (patient age and sex). The following features were used in this study: effective diameter, contrast of the segmented nodule on the HRCT image, overlap measures of two gray-level histograms for the inside and outside regions of the segmented nodule on the HRCT image, overlap measures of two gray-level histograms for the inside and outside regions of the segmented nodule on the edge-gradient image, radial gradient index for the inside region of the segmented nodule on the HRCT image, peak value of the histogram for the inside region of the segmented nodule on the edge-gradient image, pixel value at the peak of the histogram for the inside region of the segmented nodule on the edge-gradient image, and pixel value at the peak of the histogram for the inside region of the segmented nodule on the HRCT image.

Our computerized classification method outputs a percentage, from 1% to 99%, that indicates the likelihood of malignancy. The performance of the classification scheme yielded an A_z value of 0.937 (0.919 for nodules with pure ground-glass opacity, 0.852 for nodules with mixed ground-glass opacity, and 0.957 for solid nodules) for distinction between 61 malignant and 183 benign lung nodules.

Observer Study

The data used in this observer study consisted of 28 malignant nodules that were randomly selected from the 61 primary lung cancers and 28 benign nodules that were selected from the 183 benign nodules by matching in size and pattern to the cancers. For both malignant and benign lesions, nine nodules ranged from 6 to 10 mm and 19 nodules ranged from 11 to 20 mm. The patterns involved were eight nodules with pure ground-glass opacity, 12 with mixed ground-glass opacity, and eight with solid opacity. Examples of cases used for this observer study are shown in Figure 1.

Sixteen radiologists participated in this observer study. The 16 radiologists, seven chest radiologists and nine other radiologists, have a mean of 14 years of experience (range, 7–26 years). Consecutive region-of-interest HRCT images for each nodule were displayed for interpretation using the cine mode on a cathode ray tube monitor (1,280 × 1,024 resolution). The window settings were initially at a width of 1,500 H and a level of –550 H, but the settings could be adjusted by the observer. In addition, zooming capability was provided. Two clinical parameters (patient age and sex) were displayed to the observer on the monitor.

The observers were told that the purpose of this observer study was to assist radiologists in distinguishing benign from malignant lesions on HRCT by using a CAD scheme. The instructions for the observers were an explanation of the role of CAD output as a second opinion. The observers were told that 28 malignant lesions (6–10 mm, nine cases; 11–20 mm, 19 cases; pure ground-glass opacity, eight cases; mixed ground-glass opacity, 12 cases; and solid opacity, eight cases) and 28 benign lesions (matched in size and pattern to the malignant lesions) were included in this study and that the sensitivity and specificity of our CAD scheme, using a threshold of 0.50 (50%) likelihood of malignancy, are 80% and 75%, respectively.

The observers were instructed to click on a bar (left, benignancy; right, malignancy) on the screen using a mouse to indicate confidence level regarding the malignancy (or benignancy) of a lesion first without and then with computer output, and after indicating your confidence (without and with CAD), click on one of the four following clinical actions: return to annual screening; follow-up in 6 months; follow-up in 3 months; or biopsy or surgery.

For a training session before the test, we provided five cases so that the observers could learn how to operate the cine mode interface and how to take into account the computer output in their decision. The review time was not limited. The average review time was 46 min (range, 28–100 min).

Data Analysis

The confidence level ratings from each observer were analyzed using receiver operating characteristic (ROC) methodology, and a quasimaximum likelihood estimation of the binormal distribution was fitted to the radiologists' confidence ratings [8]. The statistical significance of the difference in A_z values between observer interpretations without and with the CAD scheme was tested using the Dorfman-Berbaum-Metz method [9]; this method included both observer variation and case sample variation by means of an analysis-of-variance approach. The sta-

tistical significance of the difference in A_z values between the computer outputs and observer interpretations (without and with the CAD scheme) was tested by means of confidence interval method by taking into account observer variation alone [10]. The effect of the computer output on the rating scores and also the change in scores that were due to the use of the CAD scheme were analyzed. The distributions of the radiologists' ratings and of the computer outputs were compared for the malignant and benign nodules.

The statistical significance of the difference in clinical actions between the beneficial and detrimental effect of the CAD scheme for each of the malignant and benign nodules was estimated using the Student's paired *t* test for 16 radiologists.

Results

For the cases selected for this observer study, the A_z value of the CAD scheme alone was 0.831 for distinguishing 28 malignant and 28 benign nodules (0.910 for nodules with pure ground-glass opacity, 0.814 for nodules with mixed ground-glass opacity, and 0.783

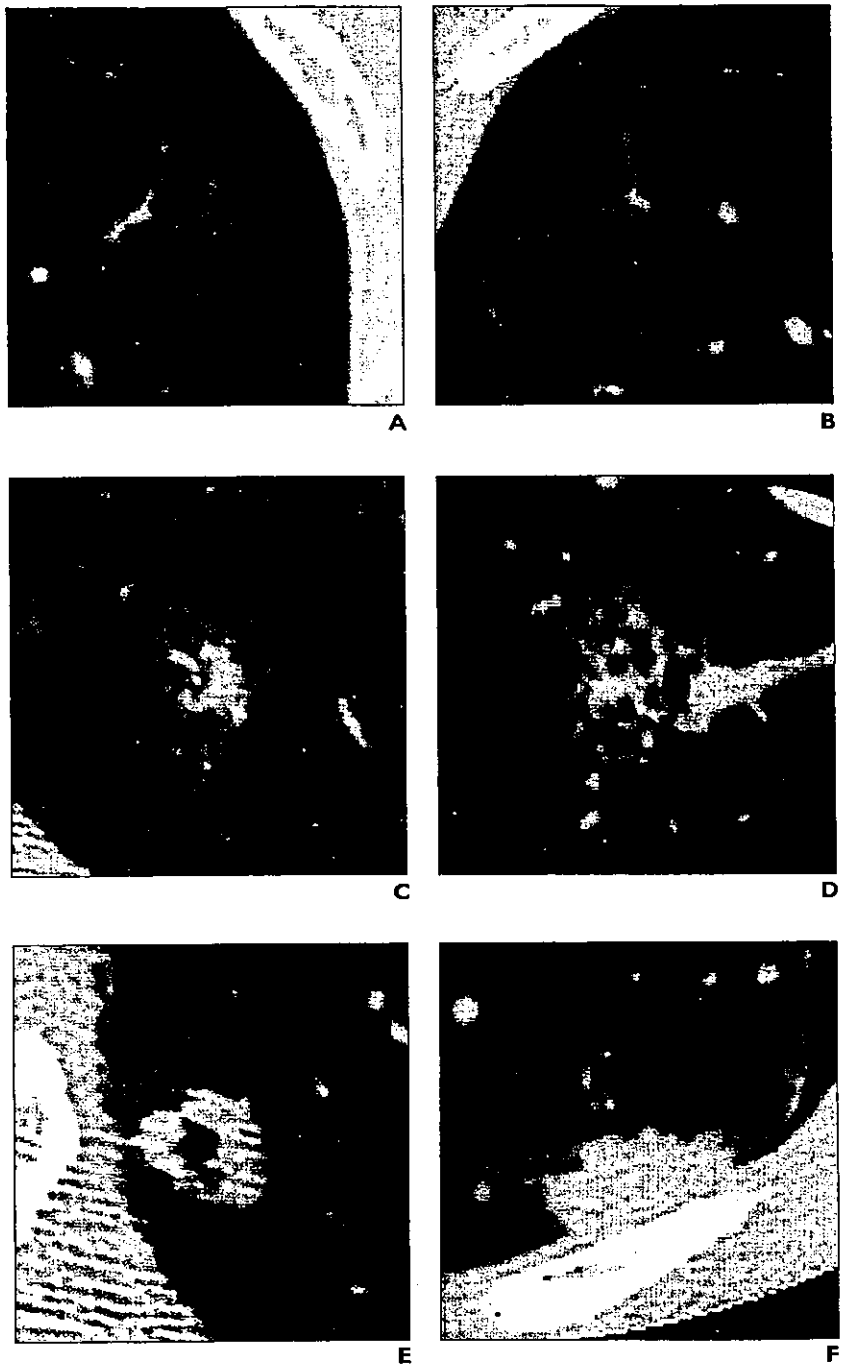

Fig. 1.—Radiologists' average ratings without and with computer output for six cases used in observer study. Note that difference in likelihood of malignancy between computer output and initial radiologists' ratings was not large in cases shown here. Radiologists' interpretation with computer-aided diagnosis (CAD) scheme was, in general, more accurate than radiologists without CAD scheme in most malignant and benign nodules.

A, High-resolution CT (HRCT) scan of 55-year-old woman with lung cancer shows pure ground-glass opacity. Computer output was 0.66; radiologists' ratings without CAD, 0.64; and radiologists' ratings with CAD, 0.71.

B, HRCT scan of 57-year-old woman with benign nodule shows pure ground-glass opacity. Computer output was 0.24; radiologists' ratings without CAD, 0.32; and radiologists' ratings with CAD, 0.27.

C, HRCT scan of 73-year-old man with lung cancer shows mixed ground-glass opacity. Computer output was 0.90; radiologists' ratings without CAD, 0.75; and radiologists' ratings with CAD, 0.85.

D, HRCT scan of 79-year-old man with benign nodule shows mixed ground-glass opacity. Computer output was 0.57; radiologists' ratings without CAD, 0.48; and radiologists' ratings with CAD, 0.56.

E, HRCT scan of 57-year-old man with lung cancer shows solid opacity. Computer output was 0.78; radiologists' ratings without CAD, 0.66; and radiologists' ratings with CAD, 0.76.

F, HRCT scan of 68-year-old man with benign nodule shows solid opacity. Computer output was 0.36; radiologists' ratings without CAD, 0.37; and radiologists' ratings with CAD, 0.36.

TABLE 1 Values for Area Under Receiver Operating Characteristic Curve (A_z) of 16 Radiologists for Interpretation Without and With CAD Scheme

Radiologist	A_z	
	Without CAD	With CAD
A	0.798	0.871
B	0.736	0.898
C	0.793	0.837
D	0.763	0.871
E	0.790	0.861
F	0.833	0.844
G	0.706	0.874
H	0.695	0.812
I	0.826	0.881
J	0.823	0.840
K	0.768	0.819
L	0.840	0.883
M	0.849	0.857
N	0.781	0.826
O	0.807	0.835
P	0.757	0.833
Mean	0.785	0.853

Note.—The difference for values without and with CAD scheme was statistically significant with a p value of 0.016. CAD = computer-aided diagnosis.

for solid nodules). Table 1 shows the A_z values without and with the CAD scheme for each radiologist. The average A_z value for the 16 radiologists was improved from 0.785 to 0.853 (from 0.812 to 0.892 for nodules with pure ground-glass opacity, from 0.819 to 0.863 for nodules with mixed ground-glass opacity, and from 0.784 to 0.844 for solid nodules) by a statistically significant level ($p = 0.016$) with the aid of the CAD scheme. The average ROC curves for the performance of the computer alone and the overall performance of the 16 radiologists without and with the CAD scheme for distinction between malignant and benign nodules are shown in Figure 2. The radiologists' diagnostic performance with the CAD scheme was more accurate than that of the CAD scheme alone ($p = 0.0005$). The A_z value for the CAD scheme was also greater than that of the radiologists alone ($p = 0.00006$).

Figure 3 shows the correlation between the computer outputs and the average radiologists' ratings without (Fig. 3A) and with (Fig. 3B) the CAD scheme for indicating the malignancy and benignancy of lung nodules. The radiologists' interpretations with the computer aid were, in general, more accurate than those of the radiologists alone for most of the malignant and benign nodules (Fig. 1). Note, however, that there were some cases for which the radiologists' ratings without CAD scheme were correct and the likelihood of malignancy in the computer output was incorrect. In those cases, the radiologists gave the correct ratings with the CAD scheme, as illustrated by three cancer cases (black circles) in the upper left quadrant and three benign cases (white circles) in the lower right quadrant in Figure 3B. Sample cases are shown in Figure 4.

The effect of the computer output on the average change in rating score due to the CAD is illustrated in Figure 5. The relationship between the likelihood of malignancy and the average change in confidence level (average change in ratings from without to with CAD) for each nodule by the 16 radiologists has a large correlation coefficient ($r = 0.927$). The radiologists increased their confidence level toward malignancy when the likelihood of malignancy was greater than 0.50 and decreased the confidence level toward benignancy when the likelihood measure was less than 0.50 for most of the malignant and benign nodules.

For the four clinical actions—return to annual screening, follow-up in 6 months, follow-up in 3 months, or biopsy or surgery, we attempted to quantify the changes in clinical action that were due to the CAD scheme. For malignant nodules, the average number of nodules for which clinical actions were changed by the 16 radiologists toward a beneficial effect (step up) (mean, 4.1 nodules) was greater than that toward a detrimental effect (step down) (mean, 1.2 nodules) ($p = 0.003$). For benign nodules, the number of nodules affected by the CAD scheme toward a beneficial effect (step down) and detrimental effect (step up) was 3.1 and 2.1, respectively ($p = 0.15$). Table 2 shows only the cases for which the clinical action was changed to or from the two extreme situations—that is, from biopsy or surgery to screening and from screening to biopsy or surgery. For malignant nodules, the difference was statistically significant between the change to (1.9 cases) and the change

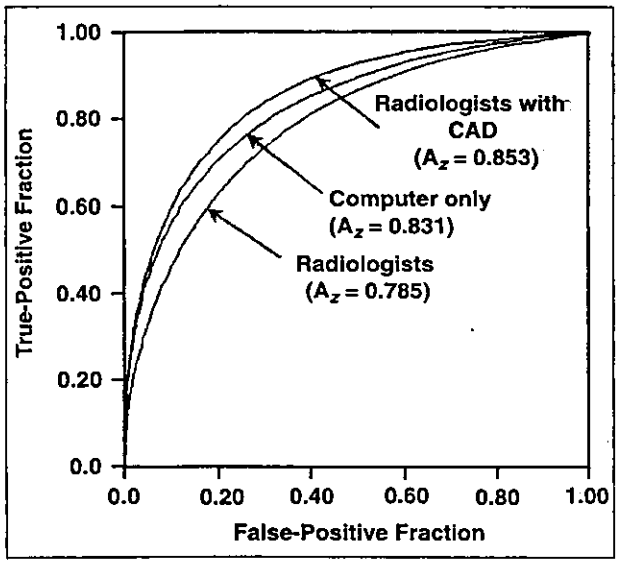

Fig. 2.—Graph shows receiver operating characteristic (ROC) curves for performance of computer alone and average performance of 16 radiologists without and with computer-aided diagnosis (CAD) scheme. Note that difference was statistically significant between radiologists without and with CAD scheme ($p = 0.016$), between computer alone and radiologists' performance without ($p = 0.00006$), and between computer alone and radiologists' performance with CAD scheme ($p = 0.0005$).

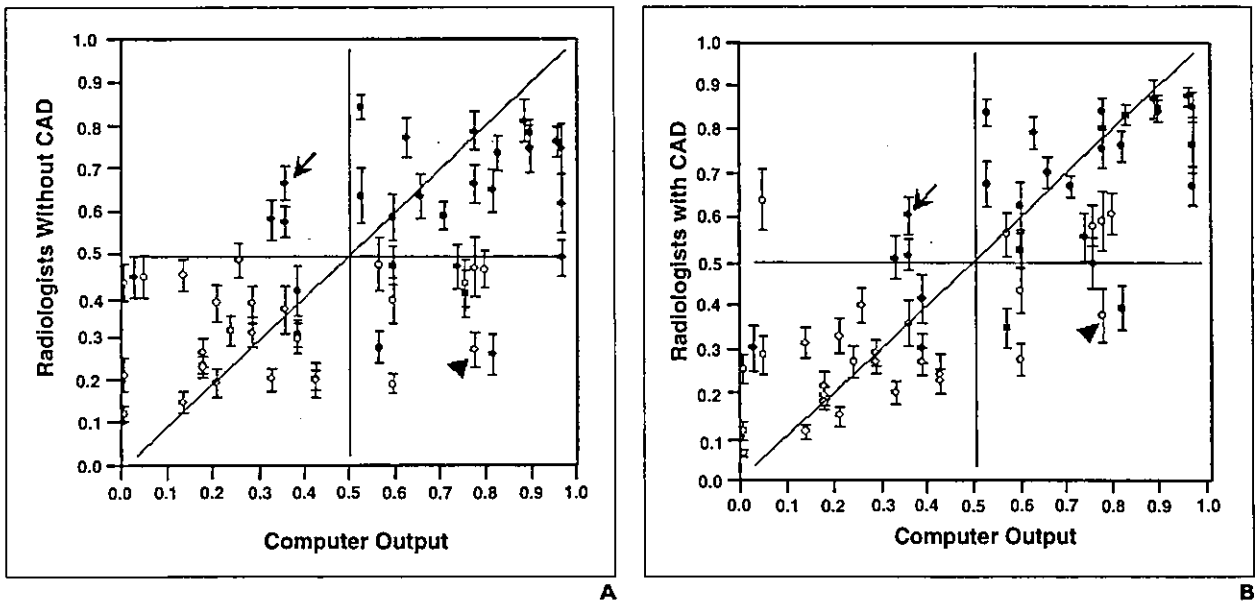

Fig. 3.—Graphs show correlation between computer output and average radiologists' ratings without and with computer-aided diagnosis (CAD) scheme for indicating likelihood of malignancy for lung nodules. ● = average radiologists' ratings for malignant nodules, ○ = average radiologists' ratings for benign nodules, horizontal lines = range of radiologists' ratings for each nodule.
A and B, Graphs show correlation between computer outputs and average radiologists' ratings without CAD (A) ($r = 0.514$) and with CAD (B) ($r = 0.784$). Note that radiologists' ratings without CAD scheme in some malignant (*upper left quadrant*) and benign (*lower right quadrant*) nodules were obviously correct, whereas likelihood of malignancy based on computer outputs alone was incorrect; even with incorrect CAD outputs, radiologists retained correct ratings. One malignant case (*arrow*) and one benign case (*arrowhead*) shown here are illustrated in Figure 4.

from (0.8 cases) biopsy or surgery ($p = 0.007$) and between the change from (0.7 cases) and the change to (0.1 cases) screening ($p = 0.02$). For benign nodules, there was no statistically significant difference between them.

Discussion

Evaluation of specific morphologic features of solitary pulmonary nodules on CT, particularly on HRCT, can help radiologists in differentiating benign from malignant lesions [11–16]. Zwirewich et al. [12] reported that increased nodule size and the presence of coarse spiculation, lobulation, and inhomogeneous central attenuation were observed with significantly greater frequency among malignant lesions, which generally appeared as solid nodules on HRCT. However, CT screening frequently detected a number of early peripheral lung adenocarcinomas, and these cancers generally appeared as nodules with pure and mixed ground-glass opacity on diagnostic HRCT [14, 15]. Some benign lesions such as nodular fibrosis also showed an HRCT pattern similar to that of adenocarcinomas and appeared as mixed ground-glass opacity nodules with a spiculated margin [16]. In this observer study, the benign lung nodules were matched in size and pattern to the malignant lung nodules, including those with pure ground-glass opacity, mixed ground-glass opacity, and solid opacity. We believe that the differential diagnosis of both benign and malignant pulmonary nodules similar in size and pattern can be difficult, and it is important to verify that a CAD scheme can assist radiolo-

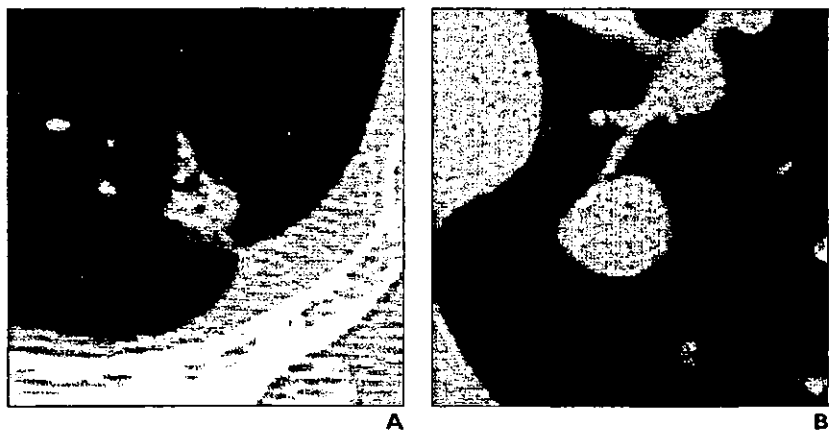

Fig. 4.—High-resolution CT (HRCT) scans show one malignant case and one benign case. Note that radiologists' interpretations without computer-aided diagnosis (CAD) scheme were correct in these cases, whereas likelihoods of malignancy based on computer outputs only were obviously incorrect; even with incorrect CAD outputs, radiologists retained correct ratings.
A, HRCT scan shows malignant lung nodule in 68-year-old man. Computer output was 0.36; radiologists' ratings without CAD, 0.67; and radiologists' ratings with CAD, 0.61.
B, HRCT scan shows benign lung nodule in 35-year-old woman. Computer output was 0.78; radiologists' ratings without CAD, 0.27; and radiologists' ratings with CAD, 0.38.

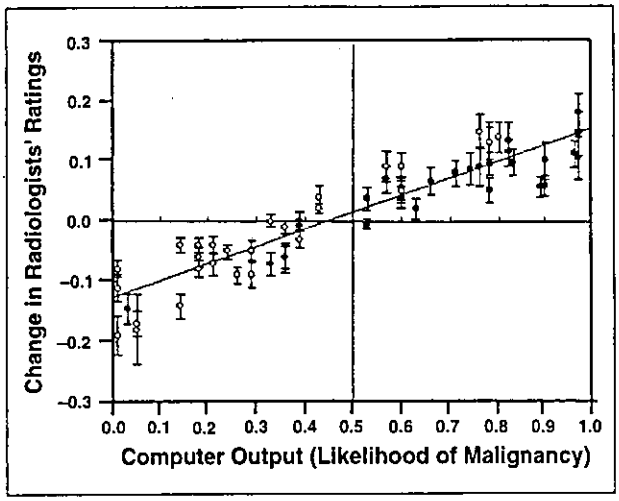

Fig. 5.—Graph shows correlation (r = 0.925) between likelihood of malignancy and average change in confidence level (rating scores) for each nodule by 16 radiologists. Malignant and benign nodules are marked by black circles and white circles, respectively. ● = average change in confidence level for malignant nodules, ○ = average change in confidence level for benign nodules, horizontal lines = range of radiologists' ratings for each nodule. Note that radiologists increased their confidence level when likelihood of malignancy was greater than 0.50 and decreased their confidence level when likelihood was less than 0.50 for most malignant and benign nodules.

TABLE 2 Cases in Which Important Clinical Actions Related to Biopsy or Screening Were Changed by 16 Radiologists as a Result of Computer-Aided Diagnosis (CAD) Scheme

Radiologist	Malignant Nodules		Benign Nodules	
	Beneficial Effect (to biopsy/from screening)	Detrimental Effect (from biopsy/to screening)	Beneficial Effect (from biopsy/to screening)	Detrimental Effect (to biopsy/from screening)
A	0/0	0/0	1/0	0/0
B	5/3	0/0	0/2	0/1
C	2/0	2/0	3/1	1/5
D	0/2	0/0	0/1	0/1
E	0/1	0/0	2/0	0/2
F	1/1	0/1	1/0	3/0
G	2/2	1/0	2/1	2/2
H	5/1	2/0	3/0	1/1
I	3/0	0/0	2/1	1/2
J	0/0	0/0	0/0	1/0
K	3/1	0/0	0/1	0/1
L	4/0	4/0	2/1	0/0
M	0/0	0/0	0/0	0/0
N	1/0	0/0	0/0	0/0
O	1/0	0/0	0/0	0/0
P	4/0	3/0	1/1	1/0
Mean (± SD)	1.9 ± 1.8[a] / 0.7 ± 0.9[b]	0.8 ± 1.3[a] / 0.1 ± 0.3[b]	1.1 ± 1.1 / 0.6 ± 0.6	0.6 ± 0.9 / 0.9 ± 1.3

Note.—Clinical actions included return for annual screening, follow-up in 6 months, follow-up in 3 months, and biopsy or surgery. The first entry for radiologist B "5/3" means the following: five cases that had been classified as annual screening, 6-month follow-up, 3-month follow-up, or biopsy or surgery were reclassified as biopsy on the basis of the CAD output, and three cases that had been classified as screening were reclassified to 6-month follow-up, 3-month follow-up, or biopsy or surgery on the basis of CAD output.

[a]The difference for all radiologists was statistically significant with a p value of 0.007 between beneficial effect and detrimental effect among malignant lesions.

[b]The difference for all radiologists was statistically significant with a p value of 0.02 between beneficial effect and detrimental effect among malignant lesions.

gists in distinguishing these benign from malignant nodules on HRCT.

Previous studies indicated several methods for determining the probability of malignancy in masses on mammography [17, 18] and solitary pulmonary nodules on chest radiography [19–22] and chest CT [7, 23–25]. Automated feature-extraction techniques have been applied in CAD schemes for classification of malignant and benign masses on breast and lung images [7, 17, 18, 22]. Several observer studies indicated that the likelihood-of-malignancy measures can improve radiologists' diagnostic accuracy in distinguishing benign from malignant lesions on radiographs [17, 18, 23, 26] and low-dose CT scans [27]. A recent study indicated that the use of an artificial neural network (ANN) as a computer aid based on attending radiologists' subjective rating scores improved radiologists' performance in terms of A_z value from 0.831 to 0.959 in differentiating benign from malignant pulmonary nodules on HRCT [25]. The performance of our automated feature-extraction scheme for all nodules in our database (A_z = 0.937) was comparable to that of the ANN by use of subjective ratings (A_z = 0.951) [25]. Our observer study indicates the usefulness of our automated computerized scheme in the classification of pulmonary nodules on HRCT images. In the future, therefore, an automated computerized scheme as second opinion may be acceptable to radiologists in clinical situations.

Our automated computerized scheme is based on various objective features (size, contrast, shape, margin, internal opacity, and internal features) of the nodules. The performance of the CAD scheme was evaluated on the basis of a leave-one-out testing method using 61 malignant and 183 benign nodules. In the computer output, a misclassification by the CAD system was observed to occur in large benign solid nodules (Fig. 4B) and in nodules with mixed ground-glass opacity, including benign (Fig. 1D) and malignant lesions (Fig. 4A). These misclassifications probably occurred because our database was obtained from a CT screening program in which all (15 lesions) solid malignant lesions were more than 10 mm, 94% (133/141) of solid benign nodules were 10 mm or less, and in a nodule with mixed ground-glass opacity, it was more difficult to differentiate benign from malignant by the CAD scheme. Also, there was a limitation in this observer study because the 56 nodules were included for developing the CAD scheme. The number of nodules, especially malignant nodules in our database, was not enough to divide training and test groups in this study,

and we plan to use an independent database from other CT screening programs to test the usefulness of our CAD scheme in the future.

Our results in this study showed that the radiologists' performance with CAD scheme (0.853) was greater than that of either radiologists alone (0.785) or computer output alone (0.831), with statistically significant differences in A_z values. The radiologists generally increased or decreased their confidence level when the likelihood of malignancy was above or below 0.50, respectively, and the changes based on CAD output for most nodules were toward a beneficial effect. Important findings are that the radiologists' initial ratings without CAD were clearly correct for some nodules and that even when the computer output indicated incorrect results, no serious detrimental effect to the radiologists' ratings as a result of the CAD output occurred. Thus, radiologists were able to maintain their correct judgments when nodules appeared obviously benign or malignant despite an incorrect CAD output. In addition, the correct computer output was able to assist radiologists in improving their decisions on many subtle cases. Therefore, this study indicated that a synergistic improvement in observers' interpretation by use of a CAD scheme as a second opinion was possible, because the radiologists were able to maintain their own correct opinions on some obvious cases, whereas the computer output assisted in improving their decisions on the majority of subtle cases.

In this study, we quantified the changes due to the CAD scheme in two extreme situations—that is, changes to or from biopsy or screening, which are important decisions in cancer screening. The results indicate the benefit of the computer aid to radiologists in making correct recommendations for malignant lesions. However, no significant benefit of the computer aid to radiologists was observed for benign nodules. Possible reasons might be that because this study was based on lung cancer CT screening, radiologists were highly alerted to avoid making underinterpretations for subtle pulmonary nodules regardless of the result of the CAD scheme.

Acknowledgments

We thank Takeshi Kobayashi, Kazuto Ashizawa, Naohiro Matsuyama, Hajime Abiru, Tetsuji Yamaguchi, Chaotong Zhang, Peter MacEneaney, Ulrich Bick, Christopher Straus, Edward Michale, Gregory Scott Stacy, Akiko Egawa, Tomoaki Okimoto, Kazunori Minami, and Shuji Sakai, for participating as observers; Shigehiko Katsuragawa, for helpful suggestions; and Elisabeth Lanzl 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. 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
5. 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
6. Mahadevia PJ, Fleisher LA, Frick KD, Eng J, Goodman SN, Powe NR. Lung cancer screening with helical computed tomography in older adult smokers. *JAMA* 2003;289:313–322
7. 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
8. Metz CE, Herman BA, Shen JH. Maximum-likelihood estimation of receiver operating (ROC) curves from continuously distributed data. *Stat Med* 1998;17:1033–1053
9. 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
10. Metz CE. Quantification of failure to demonstrate statistical significance: the usefulness of confidence intervals. *Invest Radiol* 1993;28:59–63
11. Erasmus JJ, Connolly JE, McAdams HP, Roggli VL. Solitary pulmonary nodules. I. Morphologic evaluation for differentiation of benign and malignant lesions. *RadioGraphics* 2000;20:43–58
12. Zwirewich CV, Vedal S, Miller RR, Müller NL. Solitary pulmonary nodule: high-resolution CT and radiologic-pathologic correlation. *Radiology* 1991;179:469–476
13. Takashima S, Sone S, Li F, et al. Small solitary pulmonary nodules (≤ 1 cm) detected at population-based CT screening for lung cancer: reliable high-resolution CT features of benign lesions. *AJR* 2003;180:955–964
14. Yang Z-G, Sone S, Takashima S, et al. High-resolution CT analysis of small peripheral lung adenocarcinomas revealed on screening helical CT. *AJR* 2001;176:1399–1407
15. Henschke CI, Yankelevitz DF, Mirtcheva R, et al. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *AJR* 2002;178:1053–1057
16. Li F, Sone S, Maruyama Y, et al. Correlation between high-resolution computed tomographic, magnetic resonance and pathological findings in cases with non-cancerous but suspicious lung nodules. *Eur Radiol* 2000;10:1782–1791
17. Chan HP, Sahiner B, Helvie MA, et al. Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: an ROC study. *Radiology* 1999;212:817–827
18. Huo Z, Giger ML, Vyborny CJ, Metz CE. Breast cancer: effectiveness of computer-aided diagnosis observer study with independent database of mammograms. *Radiology* 2002;224:560–568
19. Gurney JW. Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis. I. Theory. *Radiology* 1993;186:405–413
20. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules: application to small radiologically indeterminate nodules. *Arch Inter Med* 1997;157:849–855
21. Nakamura K, Yoshida H, Engelmann R, et al. Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules with use of artificial neural networks. *Radiology* 2000;214:823–830
22. Aoyama M, Li Q, Katsuragawa S, MacMahon H, Doi K. Automated computerized scheme for distinction between benign and malignant solitary pulmonary nodules on chest images. *Med Phys* 2002;29:701–708
23. Henschke CI, Yankelevitz DF, Mateescu I, Brettle DW, Rainey TG, Weingard FS. Neural networks for the analysis of small pulmonary nodules. *Clin Imaging* 1997;21:390–399
24. McNitt-Gray MF, Hart EM, Wyckoff N, Sayre JW, Goldin JG, Aberle D. A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution CT: preliminary results. *Med Phys* 1999;26:880–888
25. Matsuki Y, Nakamura K, Watanabe H, et al. Usefulness of an artificial neural network for differentiating benign from malignant pulmonary nodules on high-resolution CT: evaluation with receiver operating characteristic analysis. *AJR* 2002;178:657–663
26. Shiraishi J, Abe H, Engelmann R, Aoyama M, MacMahon H, Doi K. Computer-aided diagnosis to distinguish benign from malignant solitary pulmonary nodules on radiographs: ROC analysis of radiologists' performance—initial experience. *Radiology* 2003;227:469–474
27. Li Q, Li F, Shiraishi J, Katsuragawa S, Sone S, Doi K. Investigation of new psychophysical measures for evaluation of similar images on thoracic CT for distinction between benign and malignant nodules. *Med Phys* 2003;30:2584–2593

International Congress Series 1268 (2004) 946–951

ELSEVIER

www.ics-elsevier.com

Usefulness of computerized scheme for differentiating benign from malignant lung nodules on high-resolution CT

Feng Li[a,*], Qiang Li[a], Masahito Aoyama[b], Junji Shiraishi[a], Hiroyuki Abe[a], Kenji Suzuki[a], Roger Engelmann[a], Shusuke Sone[c], Heber MacMahon[a], Kunio Doi[a]

[a] *Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, MC-2026, 5841 South Maryland Avenue, Chicago, IL 60637, USA*
[b] *Department of Intelligent Systems, Faculty of Information Sciences, Hiroshima City University, Hiroshima 731-3194, Japan*
[c] *Azumi General Hospital, Ikeda, Nagano 399-8695, Japan*

Abstract. A computer-aided diagnosis (CAD) scheme for determination of the likelihood of malignancy of 244 nodules on high-resolution CT (HRCT) was developed. The performance (A_z) for 16 radiologists was improved from 0.785 to 0.853 ($P=0.02$) with the aid of the CAD scheme by use of 56 nodules, including 28 cancerous and 28 benign nodules which were matched in size and pattern to the cancers. Our purpose in this study was to investigate further whether a CAD scheme can assist radiologists in distinguishing benign from malignant nodules in different groups. The results indicated that A_z values for radiologists without and with the CAD scheme were improved from 0.770 to 0.855 for general radiologists ($P=0.01$) and from 0.805 to 0.850 for chest radiologists ($P=0.12$); from 0.717 to 0.821 for nodules at 6–10 mm ($P=0.04$) and from 0.837 to 0.901 for nodules at 11–20 mm ($P=0.04$); and from 0.812 to 0.892 for nodules with pure ground-glass opacity (GGO) ($P=0.149$), from 0.819 to 0.863 for nodules with mixed GGO ($P=0.196$), and from 0.784 to 0.844 for solid nodules ($P=0.334$). CAD has the potential to improve the diagnostic accuracy in distinguishing benign nodules from malignant ones in different groups on HRCT.

Keywords: Lung nodule; Computer-aided diagnosis (CAD); Receiver operating characteristic (ROC) curve

1. Introduction

Low-dose spiral CT has been applied for cancer screening and led to early detection of lung cancer in some countries [1–3]. However, simultaneous or additional diagnostic high-resolution CT (HRCT) was needed for distinction between early cancers, most of

* Corresponding author. Tel.: +1-773-834-5093; fax: +1-773-702-0371.
E-mail address: fli@kurt.bsd.uchicago.edu (F. Li).

0531-5131/
doi:10.1016/j.ics.2004.03.083

which were nodules with ground-glass opacity (GGO), and a large number of benign nodules, which were detected as suspicious lesions by screening with low-dose CT [4,5]. It is important to differentiate benign from malignant nodules on the first diagnostic HRCT in order to reduce the number of follow-up HRCT examinations. The database used in this study was obtained as part of follow-up diagnostic work in a lung cancer screening program in Japan [1]. The database consisted of 244 small, non-calcified (3–20 mm) nodules, including nodules with pure GGO, mixed GGO, and solid opacity on the HRCT (1 mm collimation). We found that certain HRCT characteristics, which were reviewed subjectively by radiologists, could be helpful in differentiating small malignant from benign nodules [5]. Recently, we developed an automated computerized scheme for distinction between malignant and benign lesions. Our observer study indicated that the average A_z value for 16 radiologists was improved from 0.785 to 0.853 by a statistically significant amount ($P=0.02$) with the aid of a computer-aided diagnosis (CAD) scheme for 28 malignant nodules and 28 benign nodules, which were matched in size and pattern to the cancers [6]. In this study, we further investigated whether a CAD scheme could assist radiologists in distinguishing benign from malignant nodules in different groups on HRCT.

2. Methods

2.1. Database

The HRCT database used in this study consisted of 61 malignant (mean 12 mm; range, 6–19 mm) and 183 benign nodules (mean 7 mm; range, 3–20 mm). The patterns of malignant and benign nodules included pure GGO (malignant vs. benign: 18 vs. 12), mixed GGO (28 vs. 30) and solid opacity (15 vs. 141), which were subjectively reviewed by three radiologists. The mean size of all solid malignant lesions was more than 10 mm, and the mean size of 133 (94%) solid benign nodules was 10 mm or less.

2.2. Computer-aided diagnosis

We developed an automated computerized scheme for determination of the likelihood of malignancy of lung nodules, which was based on various objective features (size, contrast, shape, margin, internal opacity, and internal features) of the nodules. The CAD scheme was based on multiple slices of HRCT derived from 2D and 3D volume data by use of linear discriminant analysis. The performance of our CAD scheme yielded an A_z value of 0.937 in the distinction between malignant and benign nodules for all three patterns. The A_z values of the classification scheme were 0.919 for nodules with pure GGO, 0.852 for nodules with mixed GGO, and 0.957 for solid nodules.

2.3. Observer study

The cases used in this observer study consisted of 28 malignant nodules, which were randomly selected from 61 primary lung cancers, and 28 benign nodules, which were