Computer-Aided Nodule Detection on Digital Chest Radiography: Validation Test on Consecutive T1 Cases of Resectable Lung Cancer

Shuji Sakai, M.D., Ph.D.,¹ Hiroyasu Soeda, M.D.,² Naoki Takahashi, M.D.,² Takashi Okafuji, M.D.,² Tadamasu Yoshitake, M.D.,² Hidetake Yabuuchi, M.D., Ph.D.,² Ichiro Yoshino, M.D., Ph.D.,³ Keiji Yamamoto, B.S.,⁴ Hiroshi Honda, M.D., Ph.D.,² and Kunio Doi, Ph.D.⁵

Purpose. To evaluate the usefulness of a commercially available computer-assisted diagnosis (CAD) system on operable T1 cases of lung cancer by use of digital chest radiography equipment. **Materials and Methods.** Fifty consecutive patients underwent surgery for primary lung cancer, and 50 normal cases were selected. All cancer cases were histopathologically confirmed T1 cases. All normal individuals were selected on the basis of chest computed tomography (CT) confirmation and were matched with cancer cases in terms of age and gender distributions. All chest radiographs were obtained with one computed radiography or two flat-panel detector systems. Eight radiologists (four chest radiologists and four residents) participated in observer tests and interpreted soft copy images by using an exclusive display system without and with CAD output. When radiologists diagnosed cases as positives, the locations of lesions were recorded on hard copies. The observers' performance was evaluated by receiver operating characteristic analysis. **Results.** The overall detectability of lung cancer cases with CAD system was 74% (37/50), and the false-positive rate was 2.28 (114/50) false positives per case for normal cases. The mean A_z value increased significantly from 0.896 without CAD output to 0.923 with CAD output ($P = 0.018$). The main cause of the improvement in performance is attributable to changes from false negatives without CAD to true positives with CAD (19/31, 61%). Moreover, improvement in the location of the tumor was observed in 1.5 cases, on average, for radiology residents. **Conclusion.** This CAD system for digital chest radiographs is useful in assisting radiologists in the detection of early resectable lung cancer.

KEY WORDS: Chest radiography, lung cancer, computer-aided nodule detection, screening, computer-assisted diagnosis, computer-assisted image interpretation, PACS

INTRODUCTION

Patients with surgically resected lung cancers show a better prognosis than nonresected patients.¹⁻⁴ Detection of T1 cases, which have a better prognosis, is important in reducing mortality due to lung cancer.²⁻⁴ Chest radiography is the most frequently used modality for screening of lung lesions because it is popular and convenient. However, diagnosis of early lung cancer via chest

¹From the Department of Health Sciences, School of Medicine, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan.

²From the Department of Clinical Radiology, Graduate School of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan.

³From the Department of Surgery and Science, Graduate School of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan.

⁴From the Mitsubishi Space Software, 5-4-36, Tsukaguchi Honmachi, Amagasaki 661-0001, Japan.

⁵From the Kurt Rossmann Laboratory, Department of Radiology, University of Chicago, 5841 South Maryland Avenue, MC2026, Chicago, IL 60637, USA.

Correspondence to: Shuji Sakai, M.D., Ph.D., Department of Health Sciences, School of Medicine, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan; tel: +81-92-6426727; fax: +81-92-6426727; e-mail: sakai@shs.kyushu-u.ac.jp

Copyright © 2006 by SCAR (Society for Computer Applications in Radiology)

Online publication 14 June 2006

doi: 10.1007/s10278-006-0626-4

radiographs may be a difficult task for radiologists, because a large part of the lung may be concealed by the overlying thoracic cage, mediastinum, large vessels, and diaphragm. Moreover, contrast resolution in chest radiography is too low to allow the detection of subtle opacities caused by small lung tumors.⁵⁻⁸ Detection failure is primarily attributed to indications being overlooked by interpreters.⁸ To improve lung cancer detection on chest radiograph, it is important to reduce the number of overlooked lesions by interpreters. The purpose of using computer-aided detection (CAD) systems is to achieve efficiency in diagnostic work by reducing reading time and by improving the accuracy. Several previous studies showed that CAD systems for lung nodules could significantly improve the diagnostic accuracy on chest radiographs.⁹⁻¹² However, most previous studies were based on a consistency test in which the same cases were used for training and testing the CAD system, and no observer study was performed on consecutive cases of resectable early lung cancer. Verifications for the usefulness of practical CAD systems must be performed by using such cases, because one important task of a screening system for lung cancer is to detect patients with better prognosis.⁵⁻⁸

In May 2002, we developed an interpretation system for chest radiographs in which we refer to CAD output images. A commercially available system for nodule detection on chest radiographs was integrated into the hospital's picture archiving and communications system (PACS). CAD has a potential to make great progress in the future by linking imaging equipment such as computed radiography (CR) or flat-panel detectors (FPD) to PACS.^{13,14} The purpose of our study was to evaluate the usefulness of a commercially available CAD system on operable T1 cases by using digital chest radiography equipment.

MATERIALS AND METHODS

Patients

From August 2002 to December 2003, 97 patients underwent curative operations for primary lung cancer. In all cases, the diagnoses of primary lung cancer and their stages were histologically confirmed. Our institutional review board approved this study, and informed consent was obtained from all study subjects. One chest radiologist and one thoracic surgeon

closely examined each case by referring to chest radiograph, computed tomography (CT), surgical records, and histopathology records. Among the 97 cases, 47 were omitted for the following reasons:¹ 36 cases due to T factor of 2 or more,² 6 cases due to multiple lesions, and³ 5 cases due to non-visualization of lung cancer on a chest radiograph. When there were two or more chest radiographs for one case, the one nearest to an operation day was used. Finally, 50 consecutive cases with primary lung cancers and 50 normal cases were selected. All cancer cases were histopathologically confirmed T1 cases (3 cm or less in greatest dimension of the tumor) and have only solitary pulmonary lesion. All normal cases were selected based on chest CT confirmation and were matched with cancer cases in terms of age and gender. Gender and age of subjects were as follows: 25 male and 25 female and 44–84 years (mean 66.5) for lung cancer cases, and 24 male and 26 female and 52–82 years (mean 65.4) for normal cases.

Chest Radiography

Chest radiographs were obtained via CR (FCR 5501D, Fuji Photo Film, Tokyo, Japan) or digital radiography with a FPD (CXDI-11, Canon, Tokyo, Japan, and Thorax FD, Siemens, Erlangen, Germany). Image data from CR and FPD were stored in the hospital's PACS with a 2K × 2K matrix, 10-bit grayscale and a 2K × 2K matrix, 12-bit grayscale, respectively. Image data used were obtained for 30 cases with CXDI-11, 17 cases with FCR 5501D, and 3 cases with Thorax FD for cancer cases, and 30, 18, and 2 for normal cases, respectively.

CAD System and Observer Performance Study

The CAD system (EpiSight/XR, Mitsubishi Space Software, Amagasaki, Japan) was designed to detect tumors of 3 cm or less on posteroanterior chest radiographs. Eight radiologists (four chest radiologists and four residents) participated in observer studies and interpreted soft copy images by using an exclusive display system with and without CAD output. In the observer performance study, only posteroanterior chest radiographs were displayed on an exclusive system with two 3-megapixel monochromatic liquid crystal display monitors (Truedia/XR, Mitsubishi Space Software, Amagasaki, Japan, and DOME C3i, Planar Systems, Beaverton, OR, USA). The observers were permitted to manipulate the monitor brightness and contrast with a mouse. This system is able to automatically calculate the area under the ROC (receiver operating characteristic) curve by using the computer program LABROC5 provided by Metz et al.¹⁵ A continuous rating scale with a line-marking method was used to record each observer's confidence level regarding the presence or absence of lung cancer. Chest radiographs with lung cancer and normal cases were presented in random order and the sequence was the same for all observers. Each observer interpreted all images in one session. The radiographs were first shown for conventional interpretation, and the observer marked his/her confidence level by marking on a bar on the lower part of the monitor screen. When observers diagnosed cases as positives, the locations of lesions were recorded on paper hard copies as a numeral 1 with

a red pencil. The CAD output image was then displayed on the other monitor. The observer then viewed the CAD output image and the radiographs and marked his or her confidence level on the same bar. When observers diagnosed cases as positives and changed the location of a nodule, the new location of the lesion was recorded on hard copies as 2 without erasing numeral 1.

Analysis

The overall detectability of lung cancer and the false positive rate were calculated as output of the CAD system. ROC analysis was used for comparison of observer performance in detecting lung cancer with and without CAD output images. Estimates of the area under the ROC curve (A_z value) were computed by using the computer program LABROC5. The observers' performance was evaluated by ROC analysis in the detection of lung cancer with and without CAD output images. The difference in A_z values between with and without CAD output images was analyzed by using a paired t test. The influence of CAD, including changes in the diagnosis between analysis with and without CAD, and the location of lesions detected by the observers were analyzed.

RESULTS

For lung cancer cases, the overall detection rate of the CAD system was 74% (37/50). There were 139 detections of lung cancer on 50 radiographs. Among them, 37 detections were true positives. For normal cases, there were 114 detections on 50

Table 1. A_z Value for Each Observer

Observer	A_z Value	
	Without CAD	With CAD
Chest radiologists		
1	0.930	0.950
2	0.900	0.926
3	0.958	0.984
4	0.942	0.940
Mean	0.933	0.950
Resident		
5	0.836	0.873
6	0.893	0.916
7	0.797	0.880
8	0.910	0.918
Mean	0.859	0.897
All	0.896	0.923

radiographs. Thus, the false-positive rate for lung cancer and normal cases was 2.04 and 2.28 false positives per case, respectively. The mean A_z value for all observers increased significantly from 0.896 without CAD to 0.923 with CAD output ($P = 0.018$) (Fig. 1). Increases in mean A_z values ranged from 0.859 to 0.897 for residents and from 0.933 to 0.950 for chest radiologists (Table 1).

The relevant change in diagnosis for all observers, including residents and chest radiologists, were analyzed (Fig. 2). The main cause for the improvement in performance was traced to

Fig 1. ROC curves for detection of lung cancer with and without CAD output images.

Fig 2. Total relevant change between images with and without CAD output in diagnosis by all observers, including residents and chest radiologists.

changes from false negatives without CAD to true positives with CAD (19/31, 61%) (Fig. 3). For residents, considerable changes from false positive to true negative were noted (10/22, 45%). Ratio of beneficially changed cases to detrimentally changed cases was higher for chest radiologists (9 to 2) than for residents (22 to 13). Most of deteriorated cases (13/15, 87%) in performance were observed among residents. For residents,

changes from true negative to false positive ($n = 7$) and from true positive to false negative ($n = 6$) were almost equal. On the other hand, there were only 2 deteriorated cases for chest radiologists. Moreover, improvement in the location of the tumor was observed in 1.5 cases on average (a total of 6 cases) for residents. For chest radiologists, there was no case for which the location of the tumor was changed by CAD.

Fig 3. An 82-year-old female patient with adenocarcinoma, stage IA (T1N0M0). Original chest radiograph obtained with CXDI-11 shows left apical nodule. (b) Output image with CAD indicates two candidates of lung nodules (arrowheads). In one candidate (left apex), early lung cancer is located precisely. In this case, the diagnosis by three observers was improved from false negative to true positive by reference to the CAD output.

DISCUSSION

Prerequisites for a suitable modality in cancer screening are simplicity, low cost, high detectability, and high specificity. Chest radiography is predominant over other modalities with regard to simplicity and low cost. There are arguments both for and against the usefulness of chest radiography as a screening modality for lung cancer.^{7,16-18} A negative aspect of chest radiography as screening modality for lung cancer is mainly low detectability of lung nodules.^{19,20} The most frequent cause of failure in lesion detection on chest radiograph is missed lesions by interpreters.^{8,21} To improve the detectability of lung cancer on chest radiographs, it is important to reduce the number of overlooked lesions by interpreters. One method to reduce overlooking of subtle lung cancers by interpreters is CAD of lung nodules. Recently, several previous studies reported on the usefulness of CAD systems for chest radiography.^{9,10} However, these studies used consistency test or validation test by using selected cases to evaluate the usefulness of CAD systems. In these studies, case selection bias cannot be denied. Furthermore, in consistency tests, observer performance studies were performed by use of positive data that were used to train the CAD system. Under such conditions, we cannot evaluate the true performance of CAD systems. Thus, we designed the observer performance test of this study by using consecutive lung cancer cases.

The detectability of lung cancer and the false positive rate of this CAD system were 74% and 2.28 per case, respectively. Kakeda et al.¹⁰ reported on an observer performance study in which they used the same CAD system. In their study, the detectability and the false positive rate of the CAD system were 73% and 3.15 per case. In their study and ours, the detectability rate was almost the same, but we obtained a better false-positive rate. We surmise that one of the reasons for our better rate was that we used a newer version of the CAD system. Furthermore, Kakeda et al used chest radiographs obtained by CR only, whereas in our study, most images were obtained via FPDs. Although we cannot directly compare the results of both studies, it was considered that the improvement could be attributed to the use of an upgraded CAD version. Usually, a commercially available CAD system is needed that shows a stable

performance for chest radiographs obtained with any equipment. In the past, observer performance studies used only CR or converted data from conventional radiographs by a film digitizer.^{9,10} We verified the usefulness of the CAD system on chest radiographs obtained by CR and FPDs. We believe that this CAD system makes it possible to improve the detectability of lung cancer and to decrease the false-positive rate by further technological developments. However, further studies are needed to examine the performance of this system for each modality of chest radiography.

Results of our study showed a significant improvement (from 0.896 to 0.923) in the A_z value of ROC with the use of CAD output images. Improvements in the A_z value were similar for residents (from 0.859 to 0.897) and chest radiologists (from 0.933 to 0.950). In designing a CAD system, it is very important to select a decreased false-positive rate while maintaining detectability, or an improved detectability while maintaining the false-positive rate. There were several cases in which residents changed their diagnosis due to CAD output images; they detrimentally changed their diagnosis with CAD in some cases at a percentage that cannot be ignored (13/35, 37%). On the other hand, although there were nine cases for which the chest radiologists changed their diagnosis with CAD, there were only two cases for which they changed their diagnosis detrimentally. These differences between residents and chest radiologists might result from the extent of their experience. Chest radiologists are not perplexed by the CAD output and can use that output efficiently. Furthermore, there was a total of six cases in which the locations of tumors were improved by residents with the CAD output. These changes were considerable, because only ROC analysis could not manifest these improvements.

Strategies directed at improving the outcome of lung cancer are screening, early detection, and picking up more operable cases. It has been suggested that if lesions are discovered while they are still small, the patients have an early stage of the disease, and this should result in lower lung cancer mortality.^{1,2} Recently, new screening trials have been performed with low-dose spiral CT. Prevalence data from these trials have shown that CT can detect smaller nodules compared to conventional chest radiographs.²²⁻²⁴ Nevertheless, it remains unclear whether detection of small can-

cers by CT will indicate earlier-stage disease and translate into a statistically significant reduction of lung cancer mortality.²⁵

Our study has several limitations. First, the observer performance study was carried out in a particular situation. In clinical practice, the distribution of positive and negative cases of lung cancer is extremely uneven. Usually, most cases are negative on cancer screening. The cases in our study had only a single lung cancer or no other lesion. On occasion, an inflammatory lesion, benign tumor, rib fracture, or metastatic tumor cannot be differentiated from primary lung cancer. Furthermore, although cases with multiple lesions are frequently encountered in clinical practice, we omitted such cases from our study. Second, our study consisted of observer tests with few observers. Observer performance test with many observers are more reliable. Our study was limited to chest radiologists in our hospital. Third, our observer performance study was done with interpretation based on soft copy chest radiographs, whereas past observer studies with CAD systems used hard copies of chest radiographs. However, several previous studies had shown no clear differences in the accuracy of diagnosis between soft copy and hard copy image quality.^{26,27} Hereafter, interpretation based on soft copy chest radiographs should be considered sufficiently practicable. Fourth, our study was retrospective. An ideal study design is a prospective study of cases that have never been diagnosed. Such prospective studies allow true evaluation of CAD systems.

In conclusion, we evaluated the usefulness of a commercially available CAD system on consecutive T1 cases of operable lung cancer by using several items of digital chest radiography equipment. The overall detectability rate of this CAD system for lung cancer cases was 74%, and the false-positive rate was 2.28 false positives per case for normal cases. In this study, the mean A_z value increased significantly from 0.896 without CAD to 0.923 with CAD output. The main cause for the improvement in performance was attributable to changes of 61% false negatives without CAD to true positives with CAD. Improvement in determining the location of tumor was observed in 1.5 cases on average for the residents, but there was no case for which the location of the tumor was changed by the chest radiologists. This CAD system for digital chest radiographs is useful in

assisting radiologists for detection of early resectable lung cancer.

ACKNOWLEDGMENTS

The authors are grateful to express to Miho Ochi, M.D., Hirofumi Ihara, M.D., Eiki Nagao, M.D., and Daisuke Okamoto, M.D. for participating as observers; Mrs. Elisabeth Lanzl for improving the manuscript. K. Doi is a shareholder of R2 Technology, Inc. (Los Altos, CA, USA). CAD technologies developed in the Kurt Rossmann Laboratories have been licensed to companies including R2 Technology, Deus Technologies, Riverain Medical Group, Mitsubishi Space Software Co., Median Technologies, General Electric Corporation, and Toshiba Corporation. It is the policy of the University of Chicago that investigators disclose publicly actual or potential significant financial interests that may appear to be affected by research activities.

REFERENCES

1. Jazieh AR, Kyasa MJ, Sethuraman G, et al.: Disparities in surgical resection of early-stage non-small cell lung cancer. *J Thorac Cardiovasc Surg* 123:1173–1176, 2002
2. Flehinger BJ, Kimmel M, Melamed MR: The effect of surgical-treatment on survival from early lung cancer implications for screening. *Chest* 101:1013–1018, 1992
3. Patz FE, Rossi S, Harpole HD, Herndon EJ, Goodman CP: Correlation of tumor size and survival in patients with stage IA non-small cell lung cancer. *Chest* 117:1568–1571, 2000
4. Pasic A, Postmus EP, Sutedja GT: What is early lung cancer? A review of the literature. *Lung Cancer* 45:267–277, 2004
5. Chotas HG, Ravin CE: Chest radiography: estimated lung volume and projected area obscured by the heart, mediastinum, and diaphragm. *Radiology* 193:403–404, 1994
6. Webb WR: Radiologic evaluation of the solitary pulmonary nodule. *AJR* 154: 701–708, 1990
7. Fontana RS, Sanderson DR, Woolner LB, et al.: Screening for lung cancer. A critique of the Mayo Lung Project. *Cancer* 67:1155–1164, 1991
8. Austin JH, Romney BM, Goldsmith LS: Missed bronchogenic carcinoma: Radiographic findings in 27 patients with a potentially resectable lesion evident in retrospect. *Radiology* 182:115–122, 1992
9. Kobayashi T, Xu XW, MacMahon H, et al.: Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs. *Radiology* 199: 843–848, 1996
10. Kakeda S, Moriya J, Sato H, et al.: Improved detection of lung nodules on chest radiographs using a commercial computer-aided diagnosis system. *AJR* 182:505–510, 2004
11. Shiraishi J, Katsuragawa S, Ikezoe J, et al.: Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *AJR* 174:71–74, 2000

12. Freedman MT, Lo SCB, Osicka T, et al.: Computer-aided detection of lung cancer on chest radiographs: effect of machine CAD false-positive locations on radiologists' behavior. *Proc SPIE* 4684:1311–1319, 2002
13. MacMahon H: Digital chest radiography: Practical issues. *J Thorac Imag* 18:138–147, 2003
14. Morishita J, Katsuragawa S, Kondo K, et al.: An automated patient recognition method based on an image-matching technique using previous chest radiographs in the picture archiving and communication system environment. *Med Phys* 28:1093–1097, 2001
15. Metz CE, Herman BA, Shen JH: Maximum-likelihood estimation of receiver operating characteristic (ROC) curves from continuously distributed data. *Stat Med* 17:1033–1053, 1998
16. Strauss GM, Gleason RE, Sugarbaker DJ: Chest X-ray screening improves outcome in lung cancer. A reappraisal of randomized trials on lung cancer screening. *Chest* 107: 270–279, 1995
17. Flehinger BJ, Kimmel M, Polyak T, Melamed MR: Screening for lung cancer. The Mayo Lung Project revisited. *Cancer* 72:1573–1580, 1993
18. Sobue T, Suzuki T, Naruke T: A case-control study for evaluating lung-cancer screening in Japan. Japanese Lung-Cancer-Screening Research Group. *Int J Cancer* 50:230–237, 1992
19. Gohagan JK, Marcus PM, Fagerstrom RM, et al.: Final results of the Lung Screening Study, a randomized feasibility study of spiral CT versus chest X-ray screening for lung cancer. *Lung Cancer* 47:9–15, 2005
20. Hirsch FR, Franklin WA, Gazdar AF, Bunn PA, Jr.: Early detection of lung cancer: clinical perspectives of recent advances in biology and radiology. *Clin Cancer Res* 7:5–22, 2001
21. Shah PK, Austin JH, White CS, et al.: Missed non-small cell lung cancer: radiographic findings of potentially resectable lesions evident only in retrospect. *Radiology* 226: 235–241, 2003
22. Kaneko M, Eguchi K, Ohmatsu H, et al.: Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. *Radiology* 201:798–802, 1996
23. Sone S, Takashima S, Li F, et al.: Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet* 351:1242–1245, 1998
24. Henschke CI, McCauley DI, Yankelevitz DF, et al.: Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet* 354:99–105, 1999
25. Swensen JS, Jett RJ, Hartman ET, et al.: CT screening for lung cancer: Five-year prospective experience. *Radiology* 235:259–265, 2005
26. Weatherburn GC, Ridout D, Strickland NH, et al.: A comparison of conventional film, CR hard copy and PACS soft copy images of the chest: analyses of ROC curves and inter-observer agreement. *Eur J Radiol* 47:206–214, 2003
27. Balassy C, Prokop M, Weber M, Sailer J, Herold CJ, Schaefer-Prokop C: Flat-panel display (LCD) versus high-resolution gray-scale display (CRT) for chest radiography: an observer performance study. *AJR* 184:752–756, 2005