ORIGINAL ARTICLE

Radiation Medicine: Vol. 21 No. 3, 112–119 p.p., 2003

Effect of Temporal Subtraction Technique on the Diagnosis of Primary Lung Cancer with Chest Radiography

Takeshi Matsuda,* Yoshifumi Yasuhara,* Akiko Kano,**
Teruhito Mochizuki,* and Junpei Ikezoe*

Purpose: The purpose of this study was to assess the diagnostic accuracy of the temporal subtraction technique in the detection of primary lung cancers by readers with different levels of experience.
Methods: Previous and current chest radiographs from 40 patients with histologically proven lung cancer and 40 controls were studied. Temporal subtraction images were produced using an automated digital subtraction technique. We evaluated the effect of temporal subtraction images in the diagnosis of lung cancer with chest radiographs via an observer performance study with the use of receiver operating characteristic analysis. Six experienced radiologists and six residents participated as observers.
Results: Observer performance for all observers was superior when temporal subtraction images were used (mean A_Z value increased from 0.764 to 0.836, p=0.0006). Although the average A_Z value for residents increased significantly, from 0.707 to 0.795 (p=0.0038), the average A_Z value for experienced radiologists increased only from 0.821 to 0.878 (n.s.).
Conclusion: In conclusion, the temporal subtraction technique clearly improves diagnostic accuracy for the detection of primary lung cancer. The results indicated that the use of temporal subtraction images was more beneficial for the residents than for the experienced radiologists. This method would compensate to some extent for experience-dependent diagnostic accuracy in the detection of lung cancer.

Key words: diagnostic radiology, chest radiography, lung cancer, temporal subtraction, computer-aided diagnosis

INTRODUCTION

IN CLINICAL PRACTICE, chest radiography is currently the most frequently used screening procedure for pulmonary lesions. However, even for experienced radiologists, it is sometimes difficult to identify subtle abnormalities on chest radiographs, specifically the lesions overlapping normal structures such as the ribs, heart, diaphragm, and pulmonary vessels. This could happen when the lesion is very serious, as in the case of lung cancer. In lung cancer screening with low-dose spiral computed tomography (LDSCT), lung cancers were found more frequently than with conventional chest radiography.[1,2] Because chest radiographs are usually taken in advance of computed tomography (CT), it is very important to detect subtle, suspicious abnormalities on chest radiographs. Investigators have recently reported that the use of a temporal subtraction technique improves the diagnostic accuracy of pulmonary lesions on chest radiographs.[3-7] To our knowledge, no study has reported the usefulness of the temporal subtraction method in the detection of lung cancers in a relatively large number of cases. The purpose of this study was to assess the diagnostic accuracy of the temporal subtraction technique in the detection of primary lung cancers by readers with different levels of experience.

Received January 15, 2003; revision accepted March 4, 2003.
*Department of Radiology, Ehime University School of Medicine
**Medical & Graphic Company, Konica Corporation
Reprint requests to Takeshi Matsuda, M.D., Department of Radiology, Ehime University School of Medicine, Shitsukawa, Shigenobu-cho, Onsen-gun, Ehime 791-0295, JAPAN.

SUBJECTS AND METHODS

Subjects

Forty patients with lung cancer who had current (at the time of diagnosis) and previous posteroanterior chest radiographs at a minimum interval of one year were gathered by reviewing clinical records in our and affiliated hospitals between 1989 and 1998. The lung cancer group consisted of 26 patients with adenocarcinoma, seven with squamous cell carcinoma, five with small cell lung cancer, and two with large cell carcinoma. Pathologic proof of the diagnosis was obtained by means of surgical resection or fine-needle biopsy in all cases. There were 25 women and 15 men, with an age range of 37 to 89 years (mean age, 67.7 years).

The presence or absence of each shadow of lung cancer was established on the basis of the consensus of two experienced chest radiologists (T.M. and J.I.) who did not participate in the observer performance study. These radiologists used the radiographs together with additional data such as findings on CT, on follow-up radiographs, and on surgical records for verification. The mean size of lung cancer on chest radiographs was 32.1×24.7 mm (range: 6×6 to 96×60 mm). In the cases with miniature chest fluorophotographs, tumor size was calculated from the measured size and the factor of size reduction.

Forty controls who had current and previous posteroanterior chest radiographs at a minimum interval of one year were gathered from among the participants of a regional lung cancer-screening program using chest radiography conducted by the Koseiren Medical Check-up Center. There were 19 women and 21 men, with an age range of 53 to 83 years (mean age, 70.9 years). They had normal chest radiographs or inactive lung disease. To determine whether or not they had active lung disease, we reviewed their interviews and follow-up chest radiographs that were taken on an annual basis. In all cases in the observer performance study, there was no substantial disagreement regarding the presence and location of tumor shadows.

The types of chest radiographs included 25 screen-film radiographs, six computed radiographs, and nine miniature chest fluorophotographs in the lung cancer group, and 40 miniature chest fluorophotographs in the control group. The mean interval between current and previous chest radiographs was 2.1 years (range, 1 to 12 years) in the lung cancer group and 1.7 years (range, 1 to 5 years) in the control group.

Generating temporal subtraction images

All the pairs of current and previous films were digitized with a film digitizer (Konica Laser Film Digitizer LD-5500, Konica Corporation, Tokyo, Japan) that had a matrix size of 512×512 and 12-bit gray scale. These digitized data were transferred to a dedicated workstation (CC-650TS, Konica Corporation). Eighty temporal subtraction images were generated from the image pair data by means of an automated digital subtraction technique.[8] Details of the methodology were described previously.[8,9] First, the density and contrast of digitized images were normalized. Second, global matching to correct the variations in patient positioning between the two images was performed by shifting and rotating the previous image with a rib cage edge-detection algorithm. Third, the previous image was nonlinearly warped according to the local shift vectors of a coordinate system for “best matching,” which were determined by local matching of the small regions of interest between the current and previous images by means of a cross-correlation. Finally, a temporal subtraction image was produced by subtracting the warped previous image from the current image. The subtraction image was presented with an inverse gray scale. Both the digitized original radiographs and the resulting subtraction images were printed with a laser film printer (LI-7, Konica Corporation) with a matrix size of 512×512 and a 12-bit gray scale, with a reduced format (10×10 cm image area) (Figs. 1, 2).

Evaluation of the quality of subtraction images

Before the reading test, the quality of temporal subtraction images was evaluated on the basis of the consensus of two experienced chest radiologists (T.M. and J.I.). Subtraction images were evaluated as fitting one of the following four grading descriptions: “excellent”: few mismatch artifacts of pulmonary structures; “good”: some mismatch artifacts of pulmonary structures, but the image is still suitable for diagnosis; “acceptable”: in spite of the presence of some severe mismatch artifacts of pulmonary structures, an excess of fifty percent of the lung fields is suitable for diagnosis; and “not acceptable”: there is poor matching of the whole lung, and the results are not acceptable for diagnosis.

Observer performance study

Six radiologists who had six or more years of experience and six residents who had two or fewer years of experience as radiologists participated as observers. They evaluated two image sets: set A, current and previous radiographs only; and set B, current and previous radiographs with temporal subtraction images. The images were displayed on a view box. To reduce learning effects, a minimum of seven days was required between reading sessions. Prior to image interpretation, 10 training cases not included in the observer performance study were distributed to the observers to familiarize them with temporal subtraction images and

Fig. 1. A 50-year-old man with large cell carcinoma.
(a) Previous chest radiograph. (b) Current chest radiograph (1 year later). (c) Temporal subtraction image shows a solitary mass in the right middle lung zone.

a
b
c

Fig. 2. A 59-year-old woman with adenocarcinoma.
(a) Previous chest radiograph. (b) Current chest radiograph (1 year later). (c) Temporal subtraction image clearly demonstrates that a pulmonary nodule in the right upper lung zone has become larger.

a
b
c

Table 1. Evaluation of the quality of subtraction images

Grade	Right lung		Left lung	
	N	%	N	%
Excellent	25	31.2	27	33.8
Good	27	33.8	29	36.3
Acceptable	22	27.5	18	22.5
Not acceptable	6	7.5	6	7.5

Table 2. A_Z values of ROC curves for detection accuracy

	A_Z value	
Observer No.	Without subtraction images	With subtraction images
Radiologists		
1	0.917	0.906
2	0.847	0.893
3	0.800	0.846
4	0.799	0.800
5	0.806	0.922
6	0.757	0.900
Mean A_Z value	0.821	0.878
Standard deviation	0.055	0.046
Residents		
1	0.702	0.825
2	0.785	0.818
3	0.671	0.773
4	0.712	0.771
5	0.626	0.770
6	0.748	0.815
Mean A_Z value	0.707	0.795
Standard deviation	0.056	0.026
All observers		
Mean A_Z value	0.764	0.836
Standard deviation	0.080	0.056

the scoring system. No time constraints were placed on the observers, although the reading times were recorded. For each case, the presence or absence of a new potentially important finding was rated on a continuous rating scale by placing a mark at a point on a printed 10-cm-long line, in which the left end indicated complete confidence that the lung did not have any change and the right end indicated complete confidence that the lung had such findings. The scores were converted by means of direct measurement to a scale from 0 to 100 for data analysis.

Statistical analysis

A binormal receiver operating characteristic (ROC) curve was fitted to each observer's confidence rating by means of a maximum likelihood estimation.[10,11] The A_Z value, which represents the area under a binormal ROC curve when it is plotted in the unit square, was then calculated for each fitted curve. The statistical significance of differences between the A_Z value for set A and that for set B in addition to the A_Z value for experienced radiologists and that for residents were evaluated with Student's two-tailed t-test. In addition, the statistical significance of the difference between the reading times for sets A and B was also evaluated with Student's two-tailed t-test for paired data.

RESULTS

Results of the visual evaluation of the quality of temporal subtraction images are summarized in Table 1. Out of 80 images, 74 (92.5%) were assessed as acceptable for the reading test in both the right and left lungs.

The A_Z values for all 12 observers are shown in Table 2. ROC curves are illustrated in Figs. 3-5. The average A_Z value increased significantly from 0.764 without

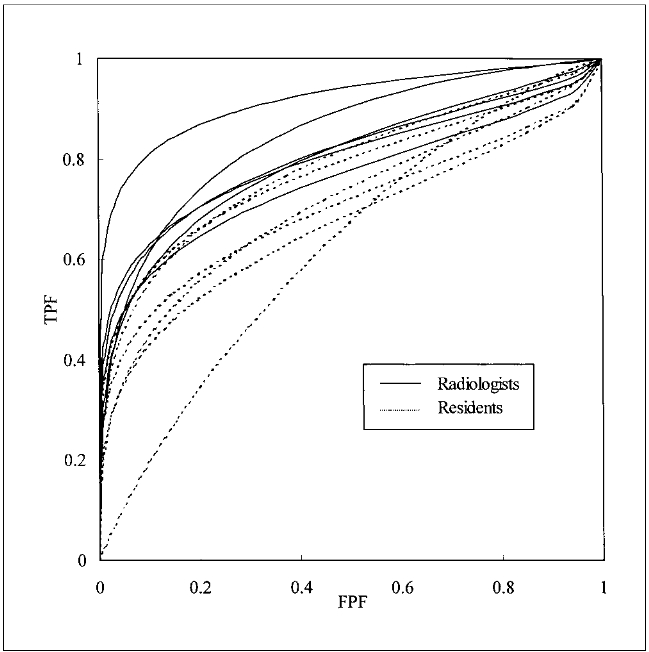

Fig. 3. ROC curves for individual observers for conventional comparison of current and previous radiographs without the temporal subtraction images. TPF, true-positive fraction; FPF, false-positive fraction.

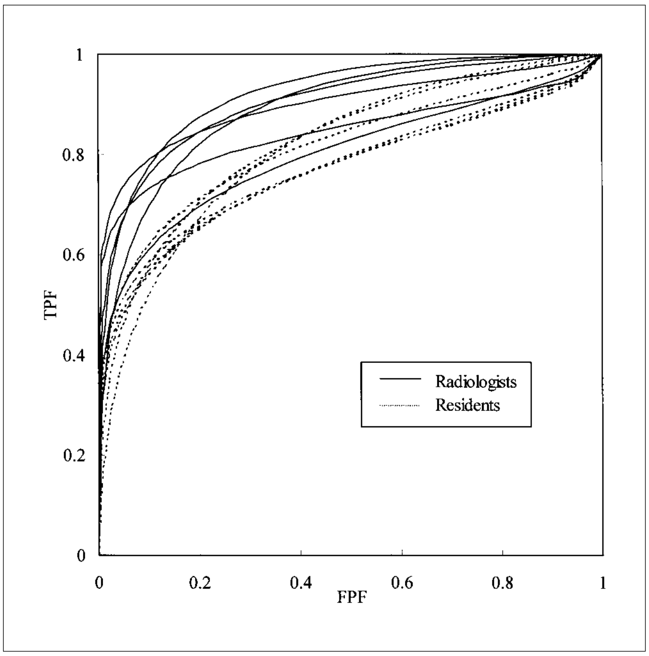

Fig. 4. ROC curves for individual observers for comparison of current and previous radiographs with the temporal subtraction images. TPF, true-positive fraction; FPF, false-positive fraction.

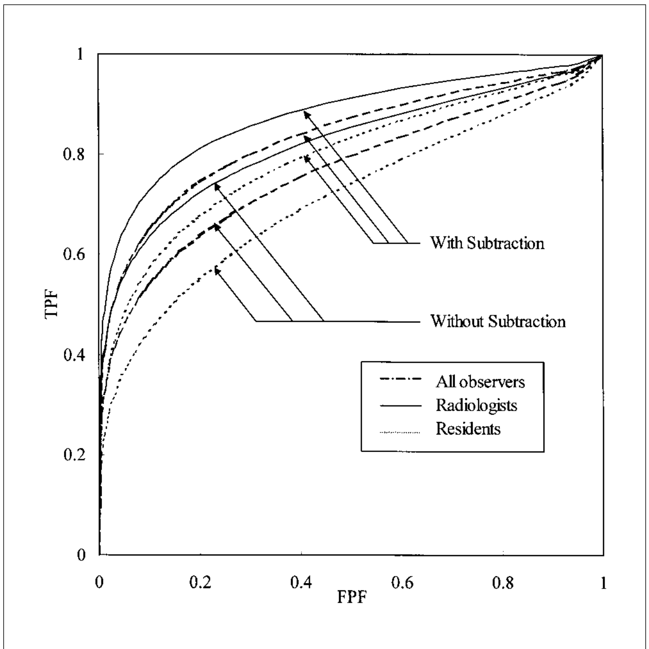

Fig. 5. Mean ROC curves show detection accuracy with and without the use of temporal subtraction images. TPF, true-positive fraction; FPF, false-positive fraction.

temporal subtraction images to 0.836 with temporal subtraction images (p=0.0006). Specifically, although the average A_Z value for residents increased significantly from 0.707 to 0.795 (p=0.0038), no statistically significant difference was seen between the average A_Z values without (0.821) and with (0.878) temporal subtraction images in the data from experienced radiologists (p=0.0721). In the comparison between the residents and experienced radiologists, the average A_Z value for the experienced radiologists was significantly higher than that of the residents in both reading tests without (p=0.0054) and with (p=0.0033) temporal subtraction images. There was no statistically significant difference between the A_Z value of the residents with the use of temporal subtraction images and that of the experienced radiologists without the use of temporal subtraction images (p=0.3273). For both groups, the mean reading time per case with temporal subtraction images (26.8 sec for experienced radiologists, 35.1 sec for residents) was significantly shorter than that without temporal subtraction images (38.8 sec for experienced radiologists, 46.4 sec for residents) ($p<0.05$) (Table 3).

Discussion

The appearance of primary lung cancer on chest radiographs is highly variable, reflecting the wide variety of the disease's histopathological nature as adenocarcinoma, squamous cell carcinoma, small cell carcinoma, large cell carcinoma, and so on. Therefore, the shadow of lung cancer on a chest radiograph could be very difficult to diagnose as lung cancer even for an experienced chest radiologist. The key to the correct diagnosis may be a subtle abnormality of a chest

Table 3. Mean reading time of each observation

Observer No.	Reading time (seconds) Without subtraction images	With subtraction images
Radiologists		
1	28.5	21.0
2	39.8	22.5
3	41.3	24.8
4	34.5	26.3
5	50.3	42.8
6	38.3	23.3
Mean	38.8	26.8
Standard deviation	7.3	8.0
Residents		
1	56.3	41.3
2	48.8	37.5
3	46.5	35.3
4	54.0	37.5
5	30.0	34.5
6	42.8	24.8
Mean	46.4	35.1
Standard deviation	9.4	5.6
All observers		
Mean	42.6	30.9
Standard deviation	8.9	7.9

radiograph, which is difficult to detect even in comparison with previous radiographs. The main reason is that three-dimensional structures such as the ribs, heart, diaphragm, and pulmonary vessels overlap the lesion on the projected image. For the evaluation of subtle changes on chest radiographs, subtraction of previous images from current images can be useful to enhance any changes in local opacity (Fig. 2). To take advantage of this potential aid in diagnosis, a computerized scheme for making temporal subtraction images based on a nonlinear geometrical warping technique has been developed.[8] This technique improves diagnostic accuracy in the detection of subtle abnormalities of the lung with various patterns, such as nodular lesions,[4-6] nonnodular lesions,[6] and lesions that create hazy pulmonary opacities.[7] To our knowledge, however, there has been no study in which the detection of primary lung cancers was evaluated with the temporal subtraction method for a relatively large number of cases. In this study, we demonstrated that the temporal subtraction technique clearly improved diagnostic accuracy in the detection of various histopathological types of primary lung cancers on chest radiographs. The results indicated that, when the primary lung cancer is the main target to be diagnosed, for example, in lung cancer screening with chest radiography, the temporal subtraction technique would be useful.

Primary lung cancer is the leading cause of cancer death in many countries. To improve the prognosis in cases of lung cancer, early diagnosis and treatment are necessary. Although mass screening for lung cancer with chest radiography is considered a method to gain early detection of the disease, its efficacy has been disputed by American and European investigators because the randomized controlled trials did not show any advances towards decreasing lung-cancer mortality.[12-15] On the other hand, other investigators have reported a high rate of lung cancer detection and a high ratio of early and small lung cancers among patients with lung cancer detected by LDSCT.[1,2,16] However, CT screening for lung cancer is not accepted as the standard because reduction of lung-cancer mortality has not been shown by any lung cancer-screening project using LDSCT. Problems such as cost-effectiveness and higher radiation exposure with LDSCT than chest radiography also need to be addressed.[17] If improved diagnostic accuracy for lung cancer contributes to a reduction of lung-cancer mortality, the temporal subtraction technique should be considered as a method for lung cancer screening. Its major advantage over LDSCT is that, since it is based on chest radiography, radiation exposure to the screener is minimal. Recently, a nested case-control study in Japan showed that a mass-screening program for lung cancer was capable of reducing by 46% the risk of death from carcinoma of the lung.[18] The program included miniature chest fluorophotography for all screeners and sputum cytology for those with a smoking index of more than 30 pack years. Since LDSCT has not been accepted in routine practice, lung cancer screening with chest radiography using the temporal subtraction technique may be a practical way to improve the diagnostic accuracy of lung cancer detection during lung cancer screening.

Our findings showed that the use of temporal subtraction images substantially improved the accuracy of diagnosing lung cancer, with the additional benefit of a reduction in mean reading time. This result is similar to those of previous studies.[3-7] In our study, six radiologists with six or more years of experience and six residents with two or fewer years of experience as radiologists participated as observers. The increase in A_Z value was significant only in the subgroup of residents. Johkoh *et al.* reported similar results in their experiments for newly detected pulmonary nodules.[5] The most remarkable findings in our study would be that the diagnostic accuracy with temporal subtraction images in the subgroup of residents increased nearly the same as did that without temporal subtraction images in the subgroup of experienced radiologists. There was no significant difference between the mean A_Z value of residents with the use of temporal subtraction images and that of radiologists without the use of temporal subtraction images. This means that the temporal subtraction technique would compensate to some extent for experience-dependent differences in diagnostic accuracy in the detection of lung cancer. Considering that chest radiography is usually the first screening method for chest diseases including lung cancer and that the reader of chest radiographs is not usually an experienced radiologist, we conclude that the temporal subtraction technique would be useful for reducing missed diagnoses of lesions in daily clinical practice.

In this study, we used various film-screen radiography images, computed radiography images, and miniature fluorophotography images. The time interval between the prior and current chest radiographs was at least one year. Under these conditions, misregistration and the difference in conditions of exposure between previous and current radiographs tended to be large and could easily cause false-positive findings that looked like pulmonary lesions on temporal subtraction images. However, the results of the visual evaluation regarding the quality of the temporal subtraction images showed that more than 90% of the images were acceptable for reading. The diagnostic accuracy of lung cancer was improved when the prior and current chest radiographs were read with temporal subtraction images. Another reason for this result seems to be that observers were not detrimentally influenced by the misregistration artifacts on the temporal subtraction images because observers could inspect the original chest radiographs to distinguish misregistration artifacts from pulmonary lesions. In other words, observers seemed to use temporal subtraction images effectively as a second opinion. These results suggest that reading chest radiographs with temporal subtraction images is an effective method for detecting changes on chest radiographs free from the difference of types of radiographs and time intervals between chest radiographs.

There were several limitations involved in our study.

First, the chest radiographs and temporal subtraction images for reading were small (10×10 cm). This was because the chest radiographs for the control group were miniature chest fluorophotographs. The size of other chest radiographs was reduced after digitization to adjust the size used in the reading test. This could have led to underestimation of the A_Z value in the detection of lung cancer. However, this would be unlikely to alter the overall conclusion because the condition would be the same with and without temporal subtraction images.

Second, we used various film-screen radiography images, computed radiography images, and miniature fluorophotography images in the lung-cancer group. However, as mentioned before, this would decrease the quality of temporal subtraction images and increase the uncertainty of any diagnosis of lung cancer. Therefore, the results of this study that the diagnostic accuracy was improved with temporal subtraction images would not be spoiled by the variety of types of images used.

Third, relatively large lesions were included in the lung cancer group. They could be easily diagnosed without temporal subtraction images. Kakeda *et al.* showed that the use of temporal subtraction images was more beneficial for the diagnosis of subtle lesions.[6] Although we consider our study to have been closer to the conditions of daily clinical practice than their setting, we may have underestimated the effect of using the temporal subtraction technique. Additional studies in a prospective clinical setting will be needed to further validate this technique.

In conclusion, the temporal subtraction technique clearly improves diagnostic accuracy for the detection of primary lung cancer. The results indicated that the use of temporal subtraction images was more beneficial for the residents than for the experienced radiologists and would compensate to some extent for experience-dependent differences in diagnostic accuracy in the detection of lung cancer.

ACKNOWLEDGMENTS

We thank Shinji Yoshioka, M.D., Shuichi Shinohara, M.D., Takashi Takeguchi, M.D., Tadashi Murakami, M.D., Masahiko Uemura, M.D., Hideki Mori, M.D., Nobuko Isshiki, M.D., Yoshiyasu Hiratsuka, M.D., Sadaaki Iwamura, M.D., Teruhito Kido, M.D., and Takanobu Okuno, M.D., for participating as observers.

References

1) 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.
2) Sone S, Takashima S, Li F, *et al.* Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet*, 351: 1242–1245, 1998.
3) Difazio MC, MacMahon H, Xu XW, *et al.* Digital chest radiography: effect of temporal subtraction images on detection accuracy. *Radiology*, 202: 447–452, 1997.
4) Uozumi T, Nakamura K, Watanabe H, *et al.* ROC analysis of detection of metastatic pulmonary nodules on digital chest radiographs with temporal subtraction. *Acad Radiol,* 8: 871–878, 2001.
5) Johkoh T, Kozuka T, Tomiyama N, *et al.* Temporal subtraction for detection of solitary pulmonary nodules on chest radiographs: evaluation of a commercially available computer-aided diagnosis system. *Radiology,* 223: 806–811, 2002.
6) Kakeda S, Nakamura K, Kamada K, *et al.* Improved detection of lung nodules by using a temporal subtraction technique. *Radiology*, 224: 145–151, 2002.
7) Tsubamoto M, Johkoh T, Kozuka T, *et al.* Temporal subtraction for the detection of hazy pulmonary opacities on chest radiography. *AJR Am J Roentgenol,* 179: 467–471, 2002.
8) Kano A, Doi K, MacMahon H, Hassell DD, Giger ML. Digital image subtraction of temporally sequential chest images for detection of interval change. *Med Phys*, 21: 453–461, 1994.
9) Ishida T, Katsuragawa S, Nakamura K, MacMahon H, Doi K. Iterative image warping technique for temporal subtraction of sequential chest radiographs to detect interval change. *Med Phys,* 26: 1320–1329, 1999.
10) Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method. *Invest Radiol,* 27: 723–731, 1992.
11) Rockette HE, Gur D, Metz CE. The use of continuous and discrete confidence judgments in receiver operating characteristic studies of diagnostic imaging techniques. *Invest Radiol,* 27: 169–172, 1992.
12) Taylor WF, Fontana RS, Uhlenhopp MA, Davis CS. Some results of screening for early lung cancer. *Cancer,* 47: 1114–1120, 1981.
13) Eddy DM. Screening for lung cancer. *Ann Intern Med,* 111: 232–237, 1989.
14) Buncher CR. *Final Report of the Central Statistical Group for Collaborative Study of Lung Cancer.* NCI Contract N01-CN-33937, Rockville, National Cancer Institute, 1985.
15) Smart CR. Critique of the early cancer detection guidelines of the US Preventive Services Task Force and of the National Cancer Institute. *Mayo Clin Proc,* 65: 892–898, 1990.
16) 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.
17) Diederich S, Wormanns D, Semik M, *et al.* Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers. *Radiology,* 222: 773–781, 2002.
18) Sagawa M, Tsubono Y, Saito Y, *et al.* A case-control study for evaluating the efficacy of mass screening program for lung cancer in Miyagi Prefecture, Japan. *Cancer,* 92: 588–594, 2001.