Lung Cancers Missed on Chest Radiographs: Results Obtained with a Commercial Computer-aided Detection Program

Feng Li, MD, PhD
Roger Engelmann, MS
Charles E. Metz, PhD
Kunio Doi, PhD
Heber MacMahon, MB, BCh

Purpose:
To retrospectively determine the sensitivity of and number of false-positive marks made by a commercially available computer-aided detection (CAD) system for identifying lung cancers previously missed on chest radiographs by radiologists, with histopathologic results as the reference standard.

Materials and Methods:
Institutional review board approval was obtained for this HIPAA-compliant study; the requirement for informed patient consent was waived. A CAD nodule detection program was applied to 34 posteroanterior digital chest radiographs obtained in 34 patients (21 men, 13 women; mean age, 69 years). All 34 radiographs showed a nodular lung cancer that was apparent in retrospect but had not been mentioned in the report. Two radiologists identified these radiologist-missed cancers on the chest radiographs and graded them for visibility, location, subtlety (extremely subtle to extremely obvious on a 10-point scale), and actionability (actionable or not actionable according to whether the radiologists probably would have recommended follow-up if the nodule had been detected). The CAD results were analyzed to determine the numbers of cancers and false-positive nodules marked and to correlate the CAD results with the nodule grades for subtlety and actionability. The χ² test or Fisher exact test for independence was used to compare CAD sensitivity between the very subtle (grade 1–3) and relatively obvious (grade > 3) cancers and between the actionable and not actionable cancers.

Results:
The CAD program had an overall sensitivity of 35% (12 of 34 cancers), identifying seven (30%) of 23 very subtle and five (45%) of 11 relatively obvious radiologist-missed cancers (P = .21) and detecting two (25%) of eight missed not actionable and ten (38%) of 26 missed actionable cancers (P = .33). The CAD program made an average of 5.9 false-positive marks per radiograph.

Conclusion:
The described CAD system can mark a substantial proportion of visually subtle lung cancers that are likely to be missed by radiologists.
Lung cancer nodules are frequently missed on chest radiographs by radiologists in clinical practice, with reported error rates of 20%–90% (1–6). Even in observer performance studies in which the radiologists are aware that many lung cancers are included, they still may fail to detect up to 40% of previously missed cancers (7). Therefore, it is important to develop more effective methods of detecting lung cancers on radiographs. Computer-aided detection (CAD) systems can alert radiologists to the location of possible lung nodules, and we previously reported that CAD schemes can assist radiologists in the detection of nodules in observer studies (8,9). However, the CAD schemes used in those previously published observer studies were developed in our laboratories or at other academic institutions and are not available for clinical use. Also, the nodules evaluated in those studies usually were selected by using subjective criteria, so the results may not be directly applicable to clinical practice. To our knowledge, to date there have been no published results on the performance of a Food and Drug Administration–approved, commercially available CAD system when it is applied to chest radiographs that depict lung cancers previously missed by radiologists. Thus, the purpose of our study was to retrospectively determine the sensitivity of and number of false-positive marks made by a commercially available CAD system for detecting lung cancers previously missed on chest radiographs by radiologists, with histopathologic results as the reference standard.

Materials and Methods

Institutional review board approval was obtained, and the requirement for informed patient consent was waived. Our study was compliant with the Health Insurance Portability and Accountability Act. Three authors (H.M., K.D., C.E.M.) are shareholders in R2 Technology (Sunnyvale, Calif). Two authors (H.M., K.D.) are consultants for Riverain Medical (Miamisburg, Ohio). One author (F.L.), who is not an employee of or consultant for Riverain Medical, had control of the data and information submitted for this study.

Patient Database, CAD System, and Reference Standard

The records stored in the cancer registry of the University of Chicago Hospitals were reviewed to identify all patients (n = 821) who received a diagnosis of lung cancer from January 2001 to November 2004. The records of all patients (n = 314) with chest radiographs on file prior to treatment were reviewed, and the relevant radiographs and reports were analyzed. From these 314 patients, we identified 34 patients—30 with posteroanterior and 4 with posterolateral digital radiographs (Fig 1)—who had a nodular cancer that was apparent on the posteroanterior or lateral image in retrospect but had not been mentioned in the report. If a nodular cancer had been mentioned in the report but was misdiagnosed, the associated chest radiograph was not included. If a patient had multiple posteroanterior radiographs with the same nodular cancer that was not mentioned by radiologists in any report, only the first radiograph in the series was used in our study. If a nodular cancer was visible on the lateral images only, the case was excluded from the study because the CAD system was designed to be used with frontal radiographs. We excluded portable anteroposterior radiographs from our study, because these images are often acquired in complex clinical situations or to answer specific questions. The location of each radiologist–missed cancer on the 34 posteroanterior chest radiographs was identified in consensus by two radiologists (F.L., H.M., 15 and 30 years experience, respectively, in chest radiograph interpretation) and confirmed at computed tomography. All cancers were confirmed by using biopsy or surgery as the reference standard.

The 34 patients with radiologist–missed lung cancers on chest radiographs were 21 men and 13 women with a mean age of 69 years (age range, 47–87 years). Fifteen of the 34 missed cancers were located in the right lung (nine in right upper lobe, one in right middle lobe, five in right lower lobe), whereas 19 were located in the left lung (12 in left upper lobe, seven in left lower lobe). The histopathologic diagnoses were non–small cell carcinoma for 32 cancers (14 adenocarcinomas, eight squamous cell carcinomas, 10 others) and small cell carcinoma for two cancers. The radiographs were obtained with a computed radiography system (Fuji Medical Systems, Stamford, Conn) by using 110 kVp and 2.5–16.0 mAs. These images were analyzed by a Food and Drug Administration–approved CAD nodule detection program (Riverain Medical, Miamisburg, Ohio).

**Implication for Patient Care**

- A currently available CAD system can detect visually subtle lung cancers that may be missed by radiologists.
Radiologists' Subjective Judgments
Before the CAD program was applied, two radiologists (F.L., H.M.) independently graded the radiologist-missed cancers for subtlety (extremely subtle to extremely obvious) on a 10-point scale, and their numeric ratings were averaged. The two radiologists also classified the missed cancers as actionable or not actionable, first independently and then in consensus. A cancer was classified as not actionable if the finding was judged to be so subtle or nonspecific that it probably would not have been acted on even if it had been identified and scrutinized.

Data Analysis
The 34 radiologist-missed lung cancers were further categorized by location into two regions (upper and lower) and two zones (medial and lateral) on the chest radiographs. The mean diameter (average of length and width) of the cancers on the 34 radiographs was determined by one radiologist (F.L.). The radiographs were also classified according to the presence or absence of underlying abnormalities such as other lung diseases or medical devices. For determination of the true-positive and false-positive CAD results for each chest radiograph, all CAD marks were analyzed by comparing the x and y coordinates of the center of each mark with the cancer locations determined by the two radiologists. In our study, the CAD localization was considered correct if the center of the CAD mark was located within the boundary of the cancer identified by the radiologists. The false-positive findings marked by the CAD program were classified as noncancerous opacities or structure-related findings by one radiologist (F.L.).

Statistical Analyses
The CAD results were analyzed to determine the number of true cancers marked (sensitivity), as well as the number of false-positive detections. The relationships between CAD performance and radiologist-assigned nodule groups and ratings and between CAD performance and presence of underlying abnormalities were analyzed. The distribution of false-positive CAD marks was compared between the right and left lungs, between the upper and lower regions, and between the medial and lateral zones on the chest radiographs.

A conventional statistical test for the inference about a proportion (10) was used to compare the difference in distribution of radiologist-missed cancers between the two lungs, between the two lung regions, and between the two lung zones. The $\chi^2$ test for independence or Fisher exact test for independence was used to compare differences in CAD sensitivity and number of false-positive CAD marks between the radiographs with and those without underlying abnormalities for detection of the cancers. These two tests for independence were also used to compare the difference in CAD sensitivity between cancer sizes (small vs large), between the two subtlety groups (very subtle vs relatively obvious), and between the two actionability groups (actionable vs not actionable) for these cancers. The single proportion test (10) also was used to compare differences in the number of false-positive CAD marks between the two lungs, between the two regions, and between the two zones.

We report two $P$ values for the comparison of cancer distributions and 10 $P$ values for the comparison of CAD performance among the various types and locations of nodules. Therefore, in accor-

![Figure 1](image1.png)

**Figure 1:** Flow diagram of patient inclusion and exclusion. $PA =$ posteroanterior.

![Figure 2](image2.png)

**Figure 2:** Posteroanterior chest radiograph shows sites of 34 radiologist-missed cancers (15 cancers in right lung, 19 in left lung). $\bullet =$ CAD-marked cancers, $\bigcirc =$ CAD-missed cancers. White horizontal and vertical lines divide each lung into upper and lower regions and medial and lateral zones. The number of radiologist-missed cancers in the upper lobes ($n = 23, 68\%$) was substantially greater than that in the lower lobes ($n = 11, 32\%$) ($P = .04$).
dance with the Bonferroni correction for multiple comparisons (11), which is at least somewhat conservative, an estimated difference in cancer distribution reported herein should be considered significant at the 95% confidence level when \( P < .025 \), whereas an estimated difference in CAD performance should be considered significant at the 95% confidence level when \( P < .005 \).

**Results**

**Features of Radiologist-missed Cancers**

We observed no significant difference in number of radiologist-missed cancers (\( n = 34 \)) between the left (\( n = 19, 56\% \)) and right (\( n = 15, 44\% \)) lungs (\( P = .50 \)) (Fig 2). However, more missed cancers were located in upper lung regions (\( n = 23, 68\% \)) than in lower regions (\( n = 11, 32\% \)) (\( P = .04 \)). Equal numbers of cancers were found in the medial (\( n = 17, 50\% \)) and lateral (\( n = 17, 50\% \)) lung zones (\( P > .99 \)). Many of the missed cancers coincided with anatomic structures (mainly bones and vessels) (Fig 2).

Among the 34 radiographs, 17 (50%) showed underlying abnormalities. Five radiographs showed postsurgical scarring, five showed catheter or external devices, three showed diffuse lung disease such as edema, five showed focal opacities such as rib fractures or scars, six showed pleural changes such as effusions, and three showed other anomalies such as cardiomegaly. Nine radiographs had at least one finding, and eight had two or more findings.

The missed lung cancers had a mean diameter of 14.3 mm (Table 1). In terms of the two radiologists’ ratings for cancer subtlety, 23 cancers were assigned a grade of 1–3 and 11 were assigned a grade of 3.5–6.5 (relatively obvious). No cancers were assigned a subtlety rating of 7 or higher, so all cancers were rated as moderately to very subtle. The two radiologists judged 26 cancers to be actionable and eight to be not actionable. Of the 26 cancers rated as actionable, 10 were judged to be relatively obvious nodules (grade > 3).

**CAD Performance**

For all radiographs, the CAD program had a sensitivity of 35% for cancer detection (Table 2), identifying the missed cancer on 12 of the 34 images, and the...
sensitivity for both radiographs with and those without underlying abnormalities was 35% (six of 17 radiographs) as well. We observed no significant difference in the number of false-positive CAD marks \((n = 199)\) between the radiographs with \((91 \{46\%\} \text{ marks}, 5.4 \text{ marks per image})\) and those without \((108 \{54\%\} \text{ marks}, 6.4 \text{ marks per image})\) underlying abnormalities \((P = .23)\).

Of the total of 199 false-positive marks on the 34 chest radiographs, 114 (57\%) marks were made in the right lung and 85 (43\%) were made in the left lung \((P = .04)\). In addition, the number of false-positive CAD marks was greater in the upper lung region \((115 \{58\%\} \text{ marks})\) than in the lower lung region \((84 \{42\%\} \text{ marks})\) \((P = .03)\) and greater in the medial lung zone \((134 \{67\%\} \text{ marks})\) than in the lateral lung zone \((65 \{33\%\} \text{ marks})\) \((P < .001)\). One hundred ninety \((96\%)\) false-positive CAD marks were related to anatomic structures \((Table 2)\), and nine \((4\%)\) were related to other objects such as noncancerous opacities or medical devices. Seventy percent \((n = 139)\) of the false-positive CAD marks were related to bones \((\text{rib, clavicle, and scapula; 39 marks})\), small and medium-size vessels \((nine \text{ marks})\), or both \((91 \text{ marks})\).

Figure 3

**Figure 3:** Posteroanterior chest radiographs obtained in 82-year-old man. \(a\) A radiologist-missed non–small cell carcinoma \((\text{arrow})\) with a mean diameter of 16 mm is located in right upper lobe. This cancer was assigned an average subtlety grade of 3.0 and was judged to be actionable by the two radiologists. \(b\) The CAD system marked the cancer \((\text{arrow pointing to star})\), as well as seven false-positive nodules \((\times)\) —five related to bones and/or vessels and two related to hila.

The sensitivity of the CAD program was 22\% \((\text{four of 18 cancers})\) for the detection of 7–15-mm cancers and 50\% \((\text{eight of 16 cancers})\) for the detection of 16–26-mm cancers \((P = .18)\) \((Table 3)\). The sensitivity of CAD detection of relatively obvious cancers \((\text{grade} \geq 3)\) was 45\% \((\text{five of 11 cancers})\), which was slightly higher than the sensitivity for the detection of very subtle cancers \((\text{grade} 1–3)\): 30\% \((\text{seven of 23 cancers})\) \((P = .21)\). CAD sensitivity was 38\% \((\text{10 of 26 cancers})\) for the detection of actionable cancers \((\text{Fig 3})\) and 25\% \((\text{two of eight cancers})\) for the detection of not actionable cancers \((P = .33)\) \((Table 3)\).

**Criteria for Validating Detections Made by CAD Software**

Figure 4b shows a cancer contained within a circle produced by the CAD program. However, this was not considered a true detection according to our criterion that the center of a CAD circle had to be within the boundary of a cancer for the mark to be considered a true detection. Figure 5b shows an example of one true-positive detection and one false-positive detection.

**Discussion**

The term missed cancer can refer to a lesion that was detected but was misin-
terpreted by radiologists (5). In our investigation, we included only those cancers that had been missed by radiologists owing to detection errors (rather than interpretation errors), because our goal was to evaluate the potential effect of CAD in such cancers. Radiologist-missed lung cancers on chest radiographs have been reported to share the following characteristics: (a) Most missed nodular cancers are visually subtle, but they are not always very small (median diameter, 16–20 mm); (b) missed cancers are located predominantly in the upper lobes; (c) superposing structures and distracting lesions are frequently present; and (d) image quality is commonly poor (3–6). The findings in this radiologist-missed cancer series were similar to those reported in previous studies (5,6), although image quality was generally high.

Kundel (12) reported that peripheral lung cancers are rarely detected until they are 8–10 mm in diameter. Woodring (13) summarized the findings of many studies with high error rates in the detection of lung cancers, determining that many radiologist-missed cancers are extremely small and/or poorly defined. For such nonspecific opacities, radiologists probably would not recommend any action, even if the lesions were detected. Among the eight cancers that were categorized as not actionable in our series, seven were categorized as extremely subtle nodules (grade 1–2) by the two radiologists. However, the mean diameter of these not actionable cancers was not particularly small (13 mm). In addition, 50% of patients had underlying abnormalities compared with 63% of patients with underlying abnormalities in a previous study (5). This suggests that a complex background, including strong structured noise, and additional findings can have a role as important as that of small lesion size in the failure of radiologists to detect a lung cancer on chest radiographs. It is important to note that detection aids such as CAD systems probably should be focused on the identification of lesions in the subgroup of radiologist-missed lung cancers that are recognizable and actionable. If the CAD program would mark these cancers accurately, it is likely that radiologists would identify them and recommend further work-up, thereby decreasing the number of lung cancers that are missed.

CAD systems have yielded a wide...
range (26%–75%) of rates for the detection of breast cancers missed by radiologists on screening mammograms (14–16). An early CAD program (17) marked approximately 40% of the lung cancers that were missed by radiologists on photofluorographic images in a lung cancer screening program, but it made an average of 15 false-positive marks per image. In our study, we used a commercially available CAD nodule detection system, and the sensitivity for all radiologist-missed lung cancers was 35%, with 5.9 false-positive CAD marks per chest radiograph. The sensitivity of CAD improved with increased tumor diameter (50% for 16–26-mm cancers) and with decreased subtlety (45% for actionable cancers). These results suggest that the system has reasonable sensitivity for the detection of more obvious and actionable cancers, although these particular nodules were missed by radiologists.

Although complex background with underlying disease can have an important role in the failure of radiologists to detect lung cancer on chest radiographs, these factors did not seem to greatly affect the performance of the CAD system in our study. In fact, the sensitivity for detection of radiologist-missed cancers was the same (35%) for images with and those without complex background abnormalities. Because of these apparent differences in sensitivity between computers and radiologists, CAD may be of greater benefit in such patients with underlying disease.

The use of newer techniques such as digital chest radiography, image processing, energy subtraction, temporal subtraction, and CAD has been shown to enhance the detection of lung nodules, and in observer performance studies (18–26), the use of CAD schemes has been shown to improve radiologists’ detection of early lung cancers. Freedman (18) suggested that these techniques, used alone or in combination, are likely to benefit patients who undergo chest radiography in situations in which cancer is not suspected. However, the commercial CAD system used in our study was not designed for use with other types of images such as energy subtraction and lateral radiographs. The findings of a preliminary (nonpublished) analysis of the results obtained by using this CAD program with 21 patients who had been imaged with a dual-energy device (not the subject of our study) suggest that the sensitivity and specificity of CAD can be substantially improved when dual-energy images are used. This will be the subject of a separate investigation.

**Figure 5**

Posteroanterior chest radiographs obtained in 66-year-old woman. (a) A radiologist-missed squamous cell carcinoma (black arrow) is located in the left upper lobe, and a rib fracture (white arrow) is seen below it. (b) The cancer (black arrow) is at the center (+) of the circle; thus, this CAD mark was counted as true-positive. The fracture (white arrow) was counted as a false-positive finding.
A major limitation of the described CAD system is the high average number of false-positive marks (5.9 per chest image). However, it seems unlikely that these marks would lead to a false-positive diagnosis since 96% of the false-positive CAD marks made in our study were clearly related to anatomic structures. Another problem is that the marks applied by the CAD system are 5-cm-diameter circles. Therefore, with an average of six to seven circles per image, the covered lung area can be as high as 25% of the total area. Because this could cause a substantial number of lesions to be marked by chance, we used a criterion by which the center of the circle had to be within the lesion boundary for the mark to be counted as a true detection. This greatly reduced the chance of "accidental" markings of cancer. Our study was designed to determine the sensitivity of and number of false-positive CAD marks made by a commercial CAD program in patients with cancers previously missed on chest radiographs by radiologists. We did not attempt to determine how the radiologists' performance might be influenced by the CAD results in these patients.

We conclude that a currently available CAD system can mark many visually subtle lung cancers that may be missed by radiologists. Although false-positive detections are numerous and potentially distracting, the majority of them are clearly due to anatomic structures, and the reduction of false-positive marks should be a priority in the development of CAD programs. The actual effects of CAD on radiologist accuracy and productivity in routine clinical practice remain to be determined, and further studies are required.

References