Sensitivity and Specificity of a CAD Solution for Lung Nodule Detection on Chest Radiograph with CTA Correlation

William Moore,¹ Jennifer Ripton-Snyder,¹ George Wu,¹ and Craig Hendler¹

The objective of this research was to determine the sensitivity and specificity of a commercially available computer-aided detection (CAD) system for detection of lung nodule on posterior–anterior (PA) chest radiograph in a varied patient population who are referred to computed tomographic angiogram (CTA) of the chest as a reference standard. Patients who had a PA chest radiograph with concomitant CTA of the chest were included in this retrospective study. The PA chest radiograph was analyzed by a CAD device, and results were recorded. A qualitative assessment of the CAD results was performed using a 5-point Likert scale. The CTA was then reviewed to determine if there were correlative nodules. The presence of a correlative nodule between 0.5 cm and 1.5 cm was considered a positive result. The baseline sensitivity of the system was determined to be 0.707 (95% CI = 0.52–0.86), with a specificity of 0.50 (95% CI = 0.38–0.76). Positive predictive value was 0.30 (95% CI = 0.24–0.49), with a negative predictive value of 0.858 (95% CI = 0.82–0.95), and accuracy of 0.555 (95% CI = 0.40–0.66). When excluding nodules that were qualitatively determined by a thoracic radiologist to be false positives, the specificity was 0.781 (95% CI = 0.764–0.839), the positive predictive value was 0.564 (95% CI = 0.491–0.654), the negative predictive value was 0.829 (95% CI = 0.819–0.878), and the accuracy was 0.737 (95% CI = 0.721–0.801). The use of CAD for lung nodule detection on chest radiograph, when used in conjunction with an experienced radiologist, has a very good sensitivity, specificity, and accuracy.

KEY WORDS: Chest CT, chest radiographs, computer assisted detection, lung neoplasms

INTRODUCTION

Lung cancer is the leading cause of cancer-related death among both men and women in the United States. In 2007, there were an estimated 171,900 cases of lung cancer diagnosed in the USA and 1.2 million cases worldwide.^{1,2} A total of 163,000 deaths from lung cancer were estimated in

the USA in 2007,¹ which is more deaths than from the second, third, and fourth leading causes of cancer-related deaths combined. The early detection of lung cancer may decrease mortality, with greater than 90% 10-year survival after surgical resection of early stage lung cancers.³ This data underscores the need to detect small lung nodules, which could be lung cancers.

Chest radiographs are commonly performed studies⁴ used to diagnose a variety of diseases.⁵ Lung nodules are frequently identified on these studies. Several studies have looked back at missed lung cancers⁶ and found that the cancers were often visible on the initial chest radiograph. The reason for the lack of diagnosis in these cases is usually multi-factorial. Some studies have shown that the addition of a computer-aided detection (CAD) system might increase the sensitivity of the general radiologist to the detection of lung cancers.⁷ Although there are several commercially available CAD platforms, the exact sensitivity and specificity of these devices have not been fully tested in a general population. The focus of this study was to determine the performance of a chest radiograph CAD system in the varied population referred to computed tomographic angiogram (CTA) of the chest as a reference standard. In addition to the baseline CAD data,

¹From the Stony Brook University Hospital, 100 Nicolls Road, HSC Level IV, Room 120, Stony Brook, NY, 11733, USA.

Correspondence to: William Moore, Stony Brook University Hospital, 100 Nicolls Road, HSC Level IV, Room 120, Stony Brook, NY, 11733, USA; tel: +1-631-4447955; fax: +1-631-4447538; e-mail: William.Moore@StonyBrook.edu

Copyright © 2010 by Society for Imaging Informatics in Medicine

Online publication 31 March 2010

doi: 10.1007/s10278-010-9284-7

the performance of the CAD system was also assessed when an experienced user interacted with the CAD system.

MATERIAL AND METHODS

All CTA of the chest performed from November 2007 to November 2008 were collected. These were then cross referenced for patients who had a posterior–anterior (PA) radiograph within a 24-h time frame. A total of 240 patients fit this selection criterion; all were included in the study. All 240 of the chest radiographs were manually sent to an FDA-approved chest radiograph CAD system (IQQA Chest v2.0 EDDA Technology, Princeton, NJ) as DICOM image. All 240 patients were included in this study regardless of age or image quality. All cases were outpatients, and inpatients were not included in this study.

All 240 cases yielded a CAD result. A subspecialty trained thoracic radiologist with 5 years of experience who had approximately 10 months of experience with the CAD device then reviewed all chest radiograph CAD images. All CAD findings were given a score from 1 to 5; where 1 was definitely not a nodule; 2, probably not a nodule; 3, unsure; 4, probably a nodule; and 5, definitely a nodule. The CTA and chest radiograph CAD results were interrogated at a separate time to avoid bias. All CTA images were performed on a GE Lightspeed 4 detector CT, (GE Medical Systems, Milwaukee, WI). Each CTA was reviewed in lung windows (center 500, window 1500) at 1.25-mm thickness without overlap. Images were displayed on a Dome Monitor (Beaverton, OR, USA) with 1,536 by 2,048 pixel resolution.

Only nodules between 0.5 and 1.5 cm were analyzed in this study; this CAD system is designed to ignore any nodule outside of these parameters.

A CTA was considered positive when a nodule between 0.5 and 1.5 cm was identified. A negative CTA was defined as a CTA where there were no nodules seen between 0.5 and 1.5 cm. A nodule was defined by CTA as a localized area of increased density; airspace disease and scars were not considered nodules. All nodules found by CTA were determined by the same experienced thoracic radiologist who evaluated the chest radiographs.

The CTA and the chest radiographs were evaluated at different times by the same thoracic radiologist to avoid bias. Any nodule greater than 1.5 cm or less than 0.5 cm was excluded from analysis, as the CAD system is designed to generate a region of interest (ROI) for nodules between 0.5 and 1.5 cm in size. The ROIs generated by the CAD system vary slightly in size on each case. A positive CAD result was considered to be any nodule in which the CAD identified a ROI. A negative CAD result was defined as any chest radiograph no ROI was identified by the CAD system. True positives were defined as both a positive CTA finding as defined above and a positive CAD result in the same region of the lung. True negatives were defined as a negative CTA with a negative CAD result. False positives were defined as CAD-generated ROIs without a corresponding nodule on CTA. False negatives were defined as nodules seen on the CTA without a corresponding CAD ROI. If there were nodules outside of the CAD reference range on CTA or chest radiography, these nodules were ignored.

All chest radiographs were exposed at 100 kV with a 10:1 grid and were obtained using a computed radiography system (FCR 9501, Fuji Photo Film, Tokyo, Japan). The imaging plate (ST-V, Fuji Photo Film) was 35×43 cm (matrix size, 1,760×2,140; gray level, 10 bit; pixel size, 200 μ m). The CAD result was sent to the PACS system, (Dynamic Imaging, Princeton, NJ) as an additional series. The system allows the radiologist to interact with the CAD results in a DICOM full resolution image, which is overlaid on the monitor. The image can be fully manipulated including enlargement of the image; window and level values can be adjusted, and the image can be inverted. Additional tools are available including measurement tools and windows designed to enhance nodule detection. These tools are fully integrated into the PACS system.

All CTA were performed using smart mA, and with intravenous contrast. The exact parameters were slightly variable, but the average mA was in the 150–300 range, dependent on patient size. All kVp were set to 120. The detector configuration was 4×1.25 mm with a pitch of 1:1.375, with a table speed of 10 mm/s, 0.5-s rotation times used in all cases. All imaging studies were performed from the lung apex to the lung bases; there were no cases where portions of the lungs were not

included. This retrospective study was approved by our institutional IRB.

Statistical analysis including a Spearman rank correlation and sensitivity, specificity, negative predictive value, positive predictive value, and accuracy were calculated using WinSTAT® for Microsoft® Excel version 2006.1.

RESULTS

Participants

A total of 240 patients were included and ranged in age from 16 to 88 years, with a mean age of 43.7 years. There were 86 male and 154 female patients included in this cohort.

CAD Results Without Radiologist

There were a total of 69 CTA confirmed nodules in 49 patients. Of these nodules, two were found to be malignant, one at 1.2 cm and the other at 1.4 cm; both were adenocarcinoma. Four nodules were densely calcified, (likely benign granulomas); these were included as true nodules, and all four were identified by CTA and chest radiograph. All other nodules are currently in a follow-up protocol. Figure 1a shows a nodule in the right lower lobe, which was identified by the CAD system, and Figure 1b shows the correlative CTA findings. The mean size of the nodules identified on CTA was 0.98 cm (± 0.3 cm SD).

Table 1. CAD Results Without Radiologist

	Positive CTA	Negative CTA
Positive CAD	49	116
Negative CAD	20	121

There were a total of 165 CAD findings or ROIs on chest radiograph, of which 49 represented actual nodules seen by CTA. These nodules were seen in 36 patients. Two patients had four nodules each, and seven patients each had two nodules. There were five nodules in four patients that were qualitatively marked as a 1 or 2 and were found to be true nodules by CTA.

No definitive correlation with size of the nodule and CAD detection of the nodule was noted. There were 121 cases in which there was no CAD ROI, and the CTA showed no nodules between 5 and 15 mm. A total of 116 false positive ROI were identified, see Table 1 and Figure 2. Therefore, the system achieved an overall sensitivity of 0.707 (95% CI=0.52–0.86), with a specificity of 0.50 (95% CI=0.38–0.76), positive predictive value of 0.30% (95% CI=0.24–0.49), with a negative predictive value of 0.86 (95% CI=0.82–0.95), and accuracy of 0.555 (95% CI=0.4–0.66). If the four granulomas are excluded, the sensitivity decreases to 0.692 (95% CI=0.51–0.87).

CAD Results with Radiologist

Of the 165 CAD findings, 87 received a score of 1 or 2 by the radiologist, see Table 2 and Figure 3.

Fig 1. 60-year-old woman with a history of smoking who had a chest x-ray for chest pain in the ED. Figure 1a shows a chest radiograph CAD result with a region of interest in the right lower lobe overlapping with the breast shadow; this area was scored as a 4. An additional region of interest was highlighted in the left lower chest, which was scored as a 1, felt to represent overlapping ribs. Figure 1b shows the correlative CTA preformed on the same day identifying the nodule in the right lower lobe in this patient.

Fig 2. Bar graph showing the total number of true positives, false positive, true negative, and false negative ROIs without interaction from a radiologist.

In our experience, nodules scored at this level can be easily dismissed; these mostly represented either crossing ribs or prominent blood vessels. This represented a total of 71.5% of false positive ROIs as designated by the CAD software. Figure 4 shows a CAD result in the right lower chest that was assigned a score of 1; it was felt to represent to be artifact from overlapping ribs; the CTA (not shown) demonstrated no nodule or other abnormality in this region. These ROI would not have led to a follow-up study. If excluded from the data set, the resultant specificity was 0.781 (95% CI=0.764–0.839); positive predictive value, 0.564 (95% CI=0.491–0.654); negative predictive value, 0.829 (95% CI=0.819–0.878); and accuracy, 0.737 (95% CI=0.721–0.801). There was a good Spearman rank correlation between radiologist nodule score and the presence of nodule on CTA with $r=0.73$, $p<0.01$. Five ROIs marked as either a 1 or 2 by the radiologist at CTA were determined to be true nodules.

An average of 0.48 false positive ROIs were seen per chest radiograph. One hundred twenty-one of the 240 cases had no ROI suggested by CAD. All ROIs were concentrated in the remaining 119 patients; of these patients, 85 patients had false positive ROIs. One case had four false positive ROIs, six cases had three false positive

Fig 3. Bar graph showing the total number of true positives, false positive, true negative, and false negative ROIs after interaction from a radiologist and dismissing ROIs scored as whether a 1 or 2 on a Likert scale. Note that five true nodules were dismissed using this scheme.

ROIs, and 15 cases had two false positive ROIs. The remaining 64 patients had only one ROI.

DISCUSSION

We found that this chest radiograph CAD system had a moderate sensitivity of 71% and a very low specificity at baseline. However, if the CAD was used as an interactive tool, we found a

Fig 4. 50-year-old woman who presented to the ED with shortness of breath. There are two regions of interest on this chest radiograph CAD result both in the right lower hemithorax. The more superior was considered to be a 1 on the Likert scale and thought to represent overlapping ribs, while in the more inferior ROI was considered less specific and given a 3 on the scale. CTA (not shown) confirmed that both of these regions of interest were false positives.

Table 2. CAD Results with Radiologist

	Positive CTA	Negative CTA
Positive CAD	44	34
Negative CAD	25	121

significant increase in the specificity to 78.1%. Accuracy of 73.7% was achieved when the CAD device was used with an experienced radiologist.

Measurements from this variable patient population who are referred to CTA studies as summarized above help to evaluate how CAD systems could fit into clinical workflow and potentially impact clinical outcome. The specificity of the CAD system increases to 80% when easy-to-dismiss nodules (nodules which receive a score of 1 or 2 on the quantitative assessment scale) are excluded from the dataset. However, five CAD ROIs that were determined to be true nodules by CTA were dismissed; this represents 10.2% of all true nodules originally found by the CAD system. van Beek et al.¹⁶ found an increase in the sensitivity in detection of lung nodules among chest radiologists when using this CAD device.

In applying CAD to real daily practice, a low number of false positive regions of interest are crucial. The importance of a low false positive number is related to reader fatigue. Large numbers of false positives could cause the radiologist to dismiss all regions of interest, including true nodules, thus making the system useless. Computer-aided detection of lung nodules has been shown to be an effective method for detecting lung nodules and lung cancers.^{6-13,15} In one study by Kakeda et al.,¹⁷ a radiologist using a chest radiographs CAD device was able to find more lung nodules than without use of a CAD device. Similar findings have also been reported in patients with metastatic disease of the chest.¹⁸ To date however, there have been no peer-reviewed articles examining the performance of a chest radiograph CAD system in this patient population. These results are extremely promising for integration of a chest radiograph CAD system into a radiology practice.

There are several limitations to this study. We did not attempt to determine if there was correlation between the chest radiograph reports and CAD results. If all of the nodules found by the CAD system were also seen by the radiologist, there would be little benefit to the system. Additionally, there could be a potential down side to the use of CAD, which was not evaluated here. If there were no additional findings and there was an average of 0.48 false positives per case, there could be a deleterious effect on the specificity of the radiologist. Additionally, although CTA with

1.25-mm images is an excellent method to evaluate the lungs, it is by no means a gold standard. This is the best available current reference standard, as thin axial, contrast-enhanced CT has the highest contrast resolution currently available for imaging of the lung. Finally, a potential population bias could have been introduced via the emergency department physician ordering trend. If the ED physicians noted a lung nodule on chest radiography, they may have been more likely to order a CTA. Additionally, most of the previously performed studies evaluating the use of CAD have focused on patients who were either high risk of a primary lung malignancy or had a previously known extra-thoracic malignancy with suspected metastases.

CONCLUSION

When this chest radiograph CAD system is used as an interactive tool, a reasonably good specificity and accuracy can be obtained. The blind use of any CAD system in any modality would not be appropriate and potentially could be deleterious to patients. The performance of this system is very good in this patient population, with a low number of false positives per case at 0.48. When used appropriately it also has a very good specificity (78.1%) for detecting nodules measuring between 5 and 15 mm. Detecting nodules of this size is critical as they are considered actionable and would thus increase the level of care that the patient receives. This study did not assess possible increase in the sensitivity of the radiologist; however, other studies have addressed this issue.^{14,18}

REFERENCES

1. American Cancer Society: Cancer Facts and Figures 2007. Atlanta, GA, American Cancer Society, 2007
2. Levi F, Lucchini F, Negri E, La Vecchia C: Worldwide patterns of cancer mortality: 1990-1994. *Eur J Cancer Prev* 8:381-400, 1999
3. The international Early Lung cancer Action program Investigators: Survival of patients with stage I lung cancer detected on CT screening. *NEJM* 355:1763-1771, 2006
4. National Institutes of Standards and Technology, an agency of the U.S. Commerce Department's Technology Administration

5. McCloud TC, Davis SD, Aquino SD, Batra PV, Goodman PC, Haramati LB, Khan A, Leung AN, Rosado de Chritenson ML, Rozenshtein A, White CS, Kaiser LR, Raoof S: Expert Panel on Thoracic Imaging. Routine admission and preoperative chest radiography. [online publication]. Reston (VA): American College of Radiology (ACR), 2006
6. 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
7. 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
8. Giger ML, Doi K, MacMahon H, Metz CE, Yin FF: Pulmonary nodules: computer-aided detection in digital chest images. *Radio Graphics* 10:41–51, 1990
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. Reeves AP, Kostis WJ: Computer-aided diagnosis for lung cancer. *Radiol Clin North Am* 38:497–509, 2000
11. Freedman MT, Lo SCB, Lure F, et al: Computer aided detection of lung cancer on chest radiographs: algorithm performance vs. radiologists' performance by size of cancer. *Proc SPIE* 4319:150–159, 2001
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. Giger ML, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography. III. Automated detection of nodules in peripheral lung field. *Med Phys* 15:158–166, 1988
14. 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
15. Jin Z, Ma D, Song W, Fan L, et al: Improving radiological interpretation of chest digital radiograph images using a real-time interactive pulmonary nodule analysis system: a cross-center study, RSNA, Chicago, November, 2005
16. van Beek E, Mullan B, Thompson B: Evaluation of a real-time interactive pulmonary nodule analysis system on chest digital radiographic images: a prospective study, RSNA, Chicago, November, 2006
17. 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
18. Van Beek E: Evaluation of a real-time interactive pulmonary nodule analysis system on chest digital radiographic images: a prospective study. *Acad Rad* 15(5):571–575, 2008