J. Gurung
A. Maataoui
M. Khan
A. Wetter
M. Harth
V. Jacobi
T. J. Vogl

Automated Detection of Lung Nodules in Multidetector CT: Influence of Different Reconstruction Protocols on Performance of a Software Prototype

Automatisierte Detektion von Lungenrundherden mittels Mehrzeilen-Detektor-Spiral-CT: Einfluss unterschiedlicher Rekonstruktionsprotokolle auf die Leistung eines Softwareprototyps

Zusammenfassung

Ziel: Evaluation der Leistungsfähigkeit eines Softwareprototyps zur automatisierten Detektion (CAD) von Lungenrundherden mittels Mehrzeilen-Detektor-Spiral-CT in Abhängigkeit von der rekonstruierten Schichtdicke. **Material und Methoden:** Computertomogramme des Thorax von 15 Patienten mit bekannten Lungenrundherden wurden mit 5,0, 2,0 und 1,0 mm Schichtdicke sowie 1,5, 1,0 und 0,5 mm Rekonstruktionsinkrement nachberechnet. Die rekonstruierten, verblindeten Datensätze wurden sowohl mittels des Softwareprototyps „Nodule Enhanced Viewing" (NEV) als auch durch 2 erfahrene Radiologen (A und B) ausgewertet. Die gefundenen Rundherde wurden entsprechend ihrer Größe (Durchmesser > 10, 5 – 10, < 5 mm) zugeordnet und die Ergebnisse der Radiologen und der CAD mit einem unabhängigen Bezugsstandard verglichen. Die statistische Auswertung erfolgte mittels „Receiver Operating Characteristic" – ROC Curve Analysis, t-test und des 2-Rater-Cohen's-Kappa-Koeffizienten. **Ergebnisse:** Insgesamt 103 Rundherde wurden mittels Referenzstandard festgestellt. Die Detektionsrate von CAD war bei einer Schichtdicke von 5,0 mm niedriger als die der Radiologen (AUC = 0,522 für A und 0,517 für B bzw. 0,497 für CAD). Bei 2,0 mm Schichtdicke war die Detektionsrate von CAD besser als die der Radiologen (AUC = 0,524 für A und B bzw. 0,614 für CAD), ohne statistische Signifikanz zu erlangen. Statistisch signifikant überlegen zeigte sich die Software bei einer Schichtdicke von 1,0 mm (AUC = 0,537 für A und 0,531 für B bzw. 0,675 für CAD). Die Sensitivität bei 1,0 mm Schichtdicke wurde mit 66,99 % für A, 68,93 % für B und 80,58 % für CAD kalkuliert. Die durchschnittliche Zeit, die für die Evaluation der Datensätze benötigt wurde,

Abstract

Purpose: To evaluate the accuracy of software for computer-aided detection (CAD) of lung nodules using different reconstruction slice thickness protocols in multidetector CT. **Materials and Methods:** Raw image data sets for 15 patients who had undergone 16-row multidetector CT (MDCT) for known pulmonary nodules were reconstructed at a reconstruction thickness of 5.0, 2.0 and 1.0 mm with a reconstruction increment of 1.5, 1.0 and 0.5 mm, respectively. The "Nodule Enhanced Viewing" (NEV) tool of LungCare for computer-aided detection of lung nodules was applied to the reconstructed images. The reconstructed images were also blinded and then evaluated by 2 radiologists (A and B). Data from the evaluating radiologists and CAD was then compared to an independent reference standard established using the consensus of 2 independent experienced chest radiologists. The eligible nodules were grouped according to their size (diameter > 10, 5 – 10, < 5 mm) for assessment. Statistical analysis was performed using the receiver operating characteristic (ROC) curve analysis, t-test and two-rater Cohen's Kappa co-efficient. **Results:** A total of 103 nodules were included in the reference standard by the consensus panel. The performance of CAD was marginally lower than that of readers at a 5.0-mm reconstruction thickness (AUC = 0.522, 0.517 and 0.497 for A, B and CAD, respectively). In the case of 2.0-mm reconstruction slices, the performance of CAD was better than that of the readers (AUC = 0.524, 0.524 and 0.614 for A, B and CAD, respectively). CAD was found to be significantly superior to radiologists in the case of 1.0-mm reconstruction slices (AUC = 0.537, 0.531 and 0.675 for A, B and CAD, respectively). The sensitivity at a recon-

war für CAD bei 1,0 mm Schichtdicke am geringsten (t = 4 min). Die Leistung der Radiologen wurde durch den Einsatz von CAD signifikant erhöht. CAD war am effektivsten in der Detektion von Rundherden < 10 mm. **Schlussfolgerung:** CAD hat bei einer Schichtdicke von 1,0 mm eine höhere Detektionsrate als Radiologen. Die relativ kurze Evaluationszeit von CAD erlaubt den Einsatz als Second Reader.

struction thickness of 1.0 mm was determined to be 66.99%, 68.93% and 80.58% for A, B and CAD, respectively. The time required for detection was shortest for CAD at reconstruction slices of 1.0 mm (mean t = 4 min). The performance of radiologists was greatly enhanced when using CAD: sensitivity 91.26% and 94.17% for CAD+A and CAD+B, respectively (AUC = 0.889 and 0.917). CAD was most advantageous in the detection of nodules < 10 mm. **Conclusion:** At a 1.0-mm reconstruction thickness, CAD's ability to detect nodules < 10 mm is superior to that of radiologists and its relatively short evaluation time makes it a viable second reader.

Key words
MDCT · lung nodules · CAD

Introduction

Some of the most common indications for CT examination of the chest are signs and symptoms suspicious for cancer, follow-up for possible metastases or to evaluate possible nodules seen in conventional chest film. Despite recent advances in diagnosis and treatment strategies, lung cancer still remains the leading cause of cancer-deaths in the western world with a mean 5-year survival rate of 14% for all stages. Although stage-1 lung cancer has a good prognosis, only 15% of patients are diagnosed at this stage [1]. The development of pulmonary metastases is another common cause of death in a patient with malignancy. This makes detection of even small nodules in the lungs mandatory although most of these nodules are benign in nature [2]. Studies show that CT is the method of choice for detection of lung nodules [3]. Identification of lung nodules by CT is influenced by methods used for image data acquisition. Using multidetector CT (MDCT) with short gantry revolution time and simultaneous acquisition of up to 16 × 0.75 mm thin slices with unprecedented spatial resolution in a single breath-hold, the detection of increasing quantities of small lung nodules is now a real possibility. Although, detection of nodules as small as 1 mm is possible with thin-slice CT, detection rates of such small nodules are usually low [4]. Naidich et al. detected only 63% of all nodules in his study involving examination of nodules of 1 – 7 mm size, with detection rates as low as 48% and 1% for nodules of diameter below 3.0 and 1.5 mm respectively [5]. There are other studies that show limited sensitivity of radiologists (below 70%) in detection of small pulmonary nodules at baseline scans [6]. Hence there is an increasing demand for automated diagnosis tools in detection and follow-up of such nodules. With advance in medical technology, there are many CAD programs developed to help radiologists in their day to day practice with the aim to increase their sensitivity. Use of CAD as a second reader has been shown to result in significant increase in sensitivity in the interpretation of mammograms [7] and chest radiographs [8]. There are many other studies focused on CAD software for detection of lung nodules. Our aim in this study is to evaluate the influence of different reconstruction slice thickness protocols on performance of a knowledge-based CAD in comparison to 2 radiologists and to test the influence of CADs on nodule detection of radiologists.

Materials and methods

Study design
The data produced from this study was not intended for clinical use hence approval of institutional review board was not required. This is a retrospective study performed on a total of 15 patients with previously known lung nodules to compare the accuracy and usefulness of a CAD software – "Nodule Enhanced Viewing" (NEV) – (LungCare – Software VB10A, Siemens Medical Solutions, Erlangen, Germany) for detection of lung nodules in MDCT, with the performance of two radiologists (J. G. and M. F. K. with 4 and 5 years of experience as radiologists in training, respectively) using different reconstruction images. MDCT image data-sets of the patients were reconstructed in three different slice thicknesses. These reconstructed images were evaluated by NEV and the two radiologists independently. The data obtained from them were compared against an independent reference standard.

Image data acquisition
MDCT of chest was performed on 15 patients with previously known lung nodules (mean age 56.1 years, range: 13 – 64 years). All these scans were performed using Somatom Sensation (Siemens Medical Solutions, Forchheim, Germany) 16-row multidetector CT (MDCT) with 0.75 mm slice collimation. The tube current was set at 80 mA with a tube-voltage of 120 kV at 0.5 s rotation time and table-feed of 24 mm/rot. Image data were reconstructed using lung filter kernel (B60f) and slice thicknesses of 5.0, 2.0 and 1.0 mm with 1.5, 1.0 and 0.5 mm reconstruction increments. These reconstructed images were not available for regular clinical routine.

Evaluation by radiologists
All the image data-sets were evaluated independently by two radiologists (J. G. and M. F. K., hereafter mentioned as A and B) after being blinded to the reconstruction protocol. Evaluation was done on a graphic workstation (Leonardo®, Siemens Inc., Erlangen, Germany) with 512 × 512 matrix size using the interactive cine mode. Computer aided diagnosis (CAD) tool was not available to the readers and no slab techniques such as MIP were used. Confidence of the radiologists in detection of nodules was determined

using a 3-point scale (1 = negative; 2 = uncertain; 3 = positive) and a percentage (out of 10 %) was allotted for their level of confidence. The 3-point scale values were used for statistical analysis. Position, diameter, consistency of nodule were also documented. The image data-sets were re-evaluated by the radiologists with the help of CAD results after a gap of 10 days to avoid recall-bias.

Evaluation by NEV

"Nodule Enhanced Viewing" (NEV) algorithm of LungCare (Software VB10A, Siemens Medical Solutions, Erlangen, Germany) is a knowledge-based CAD prototype developed for detection of lung nodules. After loading of image data into the LungCare programme, the NEV is activated automatically which starts detection of lung nodules from the loaded image data. The potential nodules detected by NEV are encircled within a volume of interest (VOI). From this VOI, a 3D characterisation of the potential nodules is achieved which allows exclusion of vessels, pleura or thoracic wall (Fig. 1). Finally, the well segmented nodule can be assessed using "evaluate nodule" function of LungCare to determine its diameter, density, volume and area (Fig. 1).

Independent reference standard

As there are no gold-standard that could be used to compare our study, a reference standard was established using the consensus of two independent experienced radiologists (V. J. and T. J. V.) with more than 15 years of experience in chest radiology, who evaluated the image data from all reconstruction images in consensus with results of evaluating radiologists and NEV. A reference standard was established for each reconstruction data-set. Comparison of the performance of NEV with the performance of radiologists for a particular individual reconstruction protocol was done by using the reference standard containing all detected nodules from the same reconstruction data-set as ground truth and are described as relative sensitivity and relative performance determined by ROC curve analysis. Ultimately, the reference standard at 1.0 mm reconstruction data-set with all detected nodules was used for comparison of accuracy or performance of NEV at different reconstruction data-sets and are described as absolute sensitivity and absolute performance determined by ROC curve analysis.

Nodules were documented based on size, demarcation or location (central or peripheral) and density (density above that of surrounding lung parenchyma). The diameter of nodules detected by NEV was used as ground-truth. For nodules missed by NEV, diameter of reader-nodules was used as ground-truth. The nodules were sub-grouped according to diameter: > 10 mm, 5 – 10 mm and < 5 mm. Pleural or sub-pleural densities with attachment to the pleura were avoided from inclusion into reference standard to allow proper definition of intrapulmonary nodules.

Statistical analysis

Two-Rater Cohen's Kappa co-efficient (k) was used to determine inter-observer agreement. Receiver operating characteristic or ROC curve analysis was used to determine the performance of radiologists and NEV. T-test was used to determine difference between diameter of detected and non-detected nodules and detected central and peripheral nodules. A p-value of less than 0.05 indicates statistical significance.

Results

Our consensus panel detected and included a total of 103 nodules in the reference standard from the 1.0 mm reconstruction image data-set after reviewing both CAD and radiologist results (average of 6.86 nodules per patient). They also identified 89 nodules (average 5.93 per patient) and 50 nodules (average 3.33 per patient) from 2.0 mm and 5.0 mm reconstruction image data-sets, respectively. More than 95 % of nodules included by our consensus panel
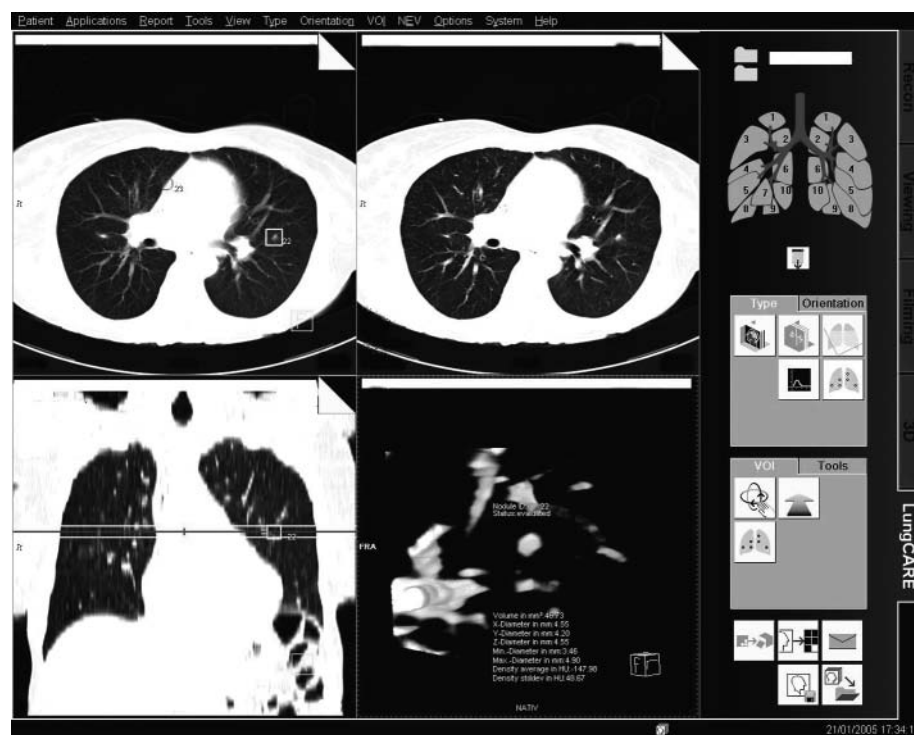
Fig. 1 Example of NEV-surface showing detected nodule (within circle) and evaluated nodule (within square) with its volume, diameter and density.

in the reference standard were solid or having a density of more than lung parenchyma. Of the 103 nodules, 19 nodules were > 10 mm in size, 46 nodules were 5 – 10 mm in size and 33 nodules were < 5 mm in size. Confidence of the evaluating radiologists was very good with ratings of 9.6 %, 9.5 % and 10 % for 1.0, 2.0 and 5.0 mm reconstruction data-sets, respectively.

Absolute sensitivity

The best performance of CAD was seen with images reconstructed at 1.0 mm thick slices with 0.5 mm reconstruction increment. Absolute sensitivity for CAD at 1.0 mm was 80.58 % and a positive predictive value (PPV) of 90.22 % (Fig. **2**). In comparison to CAD, the performance of the readers was significantly lower: absolute sensitivity for A = 66.99 % and PPV of 93.24 % absolute sensitivity for B = 68.93 % and PPV of 92.21 %. In other words, CAD outperformed the evaluating radiologists at 1.0 mm reconstruction. The performance of CAD was also better than readers for images at 2.0 mm reconstruction thickness: absolute sensitivity for CAD = 78.64 % absolute sensitivity for A and B = 59.22 %. Radiologists outperformed CAD at 5.0 mm reconstruction images with an absolute sensitivity of 41.75 % reader A and an absolute sensitivity of 42.72 % for reader B. Absolute sensitivity for CAD was calculated at 28.16 % (Fig. **3**). A detailed representation is shown in Table **1**.

Time required by CAD for detection was determined at each reconstruction image data-set: CAD required 12 – 15 minutes, 6 – 8 minutes and 3 – 5 minutes with 5.0, 2.0 and 1.0 mm reconstruction images, respectively.

Absolute performance determined by ROC analysis

Compared to 103 reference standard nodules, the absolute performance of CAD was best at 1.0 mm reconstruction images with an area under curve (AUC) of 0.675 ROC analysis (AUC for A = 0.537; AUC for B = 0.531) and a false positive rate of 0.6 per scan. (FP = 9.78, 6.75 and 7.79 % for CAD, A and B, respectively). Using 2.0 mm reconstruction thickness, the absolute performance of CAD reduced to AUC of 0.614 and a false positive rate of 1.0 per scan leading to significant deterioration of CAD performance. The performance deteriorated further with the use of 5.0 mm reconstruction images (AUC = 0.497 and a high false positive rate of 3.7 per scan). On the contrary, the absolute performance of the readers improved only slightly by reducing reconstruction slice thickness from 5.0 mm through 2.0 to 1.0 mm with no significant influence in the ROC curve analysis. Instead, there was an increase in

reader false positive rate with decreasing reconstruction slice thickness: FP increased from 1 to 5 for A and from 1 to 6 for B (Table **1**).

Relative performance determined by ROC analysis

Performance of readers and CAD at specific reconstruction slice thicknesses was determined. The readers outperformed CAD at 5.0 mm RT which was attributed largely to the high false positive and false negative rates of CAD. The relative performance was established in the form of area under curve (relative AUC) at 0.786 for A and AUC of 0.667 for B by ROC curve analysis and a false positive rate of 0.06 per scan. Whereas the relative AUC of CAD at 5.0 mm reconstruction images was calculated at 0.531 with a false positive rate of 3.7 per scan (Fig. **4**). Thus the relative sensitivity of the readers was significantly higher than CAD. The relative performance of readers was found to be lower than CAD at both 2.0 and 1.0 mm reconstruction images: relative AUC at 2.0 mm was 0.564 for A, 0.536 for B and 0.625 for CAD; relative AUC at 1.0 mm was 0.537 for A, 0.531 for B and 0.667 for CAD. At 2 mm CAD false positive rate was higher than the readers. The detailed analysis of relative performance rates of CAD in comparison to the readers for each reconstruction protocol is shown in Table **1**.

Inter-observer agreement

Inter-observer agreement between the readers was very good for all reconstruction protocols, especially at 5.0 mm reconstruction thickness (k = 0.683, 0.611 and 0.987 at 1.0, 2.0 and 5.0 mm, respectively). On the contrary, inter-observer agreement between CAD and readers was significantly poor for all reconstruction protocols, especially at 5.0 mm: k = 0.078, 0.031 and 0.012 between CAD & A and k = 0.156, 0.013 and 0.009 between CAD & B at 1.0, 2.0 and 5.0 mm, respectively (Table **2**).

Influence of different reconstruction protocol on size of detected nodules

Influence of various reconstruction protocols on CAD and the evaluating radiologists with regard to detected nodule size was determined: For 1.0 mm reconstruction images mean diameter of CAD detected nodules was 7.95 mm ± 5.39 mm standard deviation (SD) and mean diameter of nodules missed by CAD was 7.94 mm ± 3.47 mm SD (comparative p-value = 0.477). Hence there was found to be no significant difference between detected and non-detected nodule size. To the contrary, the mean diameter of reader-detected nodules at 1.0 mm was 7.86 mm ± 4.62 mm SD

Fig. **2** Right apical nodule (arrow) which was missed by NEV at 5.0 mm (**a**) but detected at 2.0 mm (**b**) and 1.0 mm (**c**) reconstruction slices (within circles).
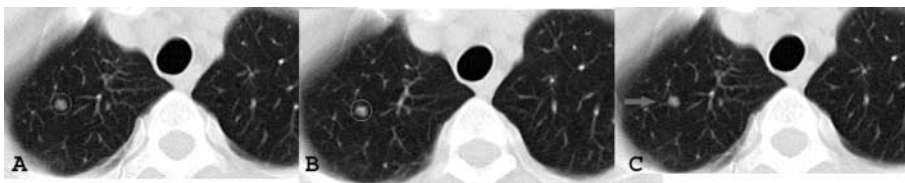

Fig. **3** Apical nodule detected NEV at 5.0 mm (**a**, within circle), 2.0 mm (**b**, within circle) but missed at 1.0 (**c**, arrow) reconstruction slices.

Table **1** Tabular representation of nodule detection, sensitivity and ROC analysis; PPV – predictive positive value; AUC – Area under curve of ROC curve analysis

| reader	recon. thickness (mm)	reference standard (n)	ground truth (n) at specific thickness	detected nodules (n)	false positive (n)	absolute sensitivity (confidence interval)	PPV (%)	absolute AUC (confidence interval)	relative sensitivity (confidence interval)	relative AUC (confidence interval)
A	1.0	103	103	69	5	66.99 (57.03–75.94)	93.24	0.537 (0.436–0.636)	66.99 (57.03–75.94)	0.537 (0.436–0.636)
B	1.0	103	103	71	6	68.93 (59.06–77.69)	92.21	0.531 (0.430–0.630)	68.93 (59.06–77.69)	0.531 (0.430–0.630)
NEV	1.0	103	103	83	9	80.58 (71.62–87.72)	90.22	0.675 (0.576–0.764)	80.58 (71.62–87.72)	0.675 (0.576–0.764)
A	2.0	103	89	61	3	59.22 (49.10–68.80)	95.31	0.524 (0.423–0.623)	68.54 (57.83–77.97)	0.564 (0.454–0.668)
B	2.0	103	89	61	3	59.22 (49.10–68.80)	95.31	0.524 (0.423–0.623)	68.54 (57.83–77.97)	0.536 (0.427–0.642)
NEV	2.0	103	89	81	15	78.64 (69.47–86.10)	84.38	0.614 (0.513–0.708)	91.01 (83.05–96.04)	0.625 (0.516–0.725)
A	5.0	103	50	43	1	41.75 (32.10–51.88)	97.73	0.522 (0.421–0.621)	86.00 (73.26–94.18)	0.786 (0.647–0.889)
B	5.0	103	50	44	1	42.72 (33.02–52.85)	97.78	0.517 (0.416–0.617)	88.00 (75.69–95.47)	0.667 (0.519–0.794)
NEV	5.0	103	50	29	55	28.16 (19.73–37.87)	34.52	0.497 (0.396–0.597)	58.00 (53.21–71.81)	0.531 (0.385–0.674)
NEV + A	1.0	103	103	94	0	91.26 (84.06–95.93)	100.0	0.889 (0.812–0.942)	91.26 (84.06–95.93)	
NEV + B	1.0	103	103	97	0	94.17 (87.75–97.83)	100.0	0.917 (0.846–0.962)	94.17 (87.75–97.83)	
NEV + A	2.0	103	89	85	2	82.52 (73.79–89.30)	97.70	0.750 (0.655–0.830)	95.51 (88.89–98.76)	
NEV + B	2.0	103	89	87	2	84.47 (76.00–90.85)	97.75	0.812 (0.724–0.883)	95.75 (92.12–99.73)	
NEV + A	5.0	103	50	46	2	44.66 (34.86–54.78)	95.83	0.588 (0.486–0.684)	92.00 (80.77–97.78)	
NEV + B	5.0	103	50	45	1	43.69 (33.94–53.82)	97.83	0.451 (0.451–0.650)	90.00 (78.19–96.67)	
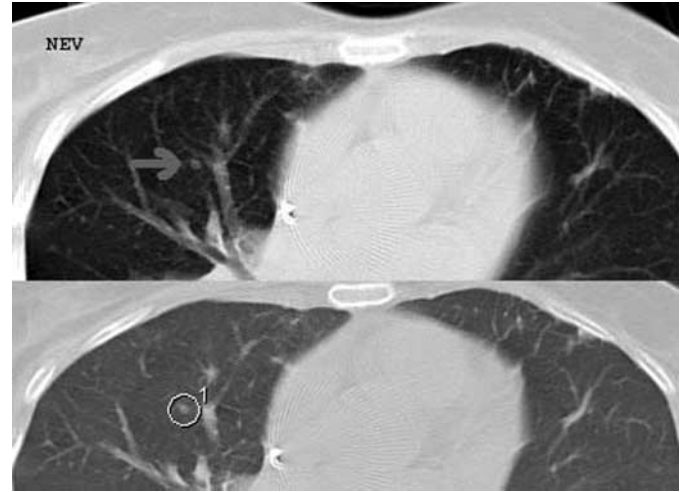
Fig. **4** Nodule detected by readers (bottom within circle) but missed by NEV (top, arrow) at 5 mm reconstruction slices.

Table **2** Two-Rater Cohen's kappa co-efficient (k) values for inter-observer agreement

reconstruction thickness	A vs B (k)	A vs NEV (k)	B vs NEV (k)
1.0 mm	0.683	0.0787	0.156
2.0 mm	0.611	0.0319	0.0129
5.0 mm	0.987	0.0125	0.009

and non-detected nodules was 3.8 mm ± 1.14 mm SD (comparative p-value = 0.159) indicating that nodules detected by readers were larger than the ones missed by them. The nodules missed by readers were also significantly smaller when compared to ones missed by CAD. Using 5.0 mm reconstruction slices, nodules detected by CAD and readers were larger than non-detected nodules. Individual analysis of nodules with regard to their size is mentioned in Table **3**.

Influence of different reconstruction protocol on detection of nodule according to their locations

Influence on location of nodules detected by CAD and evaluating radiologists by different reconstruction protocols were also determined. At 1.0 mm reconstruction images A and B detected on average 27 nodules at the centre and 43 at the periphery of the lung tissue (p = 0.002) whereas CAD detected 44 nodules centrally and 39 at the periphery. Hence there was no significant difference in detection of central and peripheral nodules (p = 0.294) by CAD but CAD was significantly better in detection of central nodules than A and B. Similarly at 2.0 mm reconstruction images CAD detected 40 central and 41 peripheral nodules (p = 0.722) whereas A and B detected on average 23 central and 38 peripheral nodules (p = 0.012). At 5.0 mm reconstruction images CAD detected 13 central and 16 peripheral nodules (p = 0.429) while A and B detected on average 20 central and 23 peripheral nodules (p = 0.191).

CAD and the effect of combined reading by CAD and radiologists

Individual evaluation of CAD and its effect on sensitivity of radiologists were as follows: CAD detected 12.62 % more nodules than A

reader	recon. thickness (mm)	n nodules (d or m)	minimum (mm)	maximum (mm)	mean (mm)	standard deviation (mm)	SEM	p-value
A and B	1.0	140 – d	1.0	21.0	7.86	4.62	0.390	0.159
A and B	1.0	66 – m	1.0	5.2	3.800	1.138	0.140	
A and B	2.0	122 – d	1.0	21.0	8.17	4.61	0.416	0.158
A and B	2.0	84 – m	1.8	6.0	4.05	1.25	0.136	
A and B	5.0	87 – d	2.1	21.0	9.33	5.04	0.540	0.153
A and B	5.0	119 – m	1.8	6.8	4.45	1.34	0.123	
NEV	1.0	83 – d	1.0	21.0	7.95	5.39	0.591	0.477
NEV	1.0	20 – m	1.0	13.0	7.94	3.47	0.776	
NEV	2.0	81 – d	1.5	21.0	7.68	5.11	0.568	0.691
NEV	2.0	22 – m	1.0	16.4	8.62	4.44	0.948	
NEV	5.0	29 – d	3.3	21.0	11.21	4.97	0.924	0.415
NEV	5.0	74 – m	1.0	21.0	7.84	5.22	0.607	

Table **3** Tabular representation of size of detected and missed nodules and their comparison using t-test (d – detected nodules; m – missed nodules; SEM – standard error of mean; p-value of t-test)

and B at 1.0 mm reconstruction images. The readers with the assistance of CAD detected more nodules than CAD alone: A + CAD detected 13.25 % and B + CAD detected 16.86 % more nodules than CAD alone. Using 2 mm reconstruction images, CAD detected 19.42 % more nodules than both A and B. Again with the help of CAD, A detected 4.93 % and B detected 7.40 % nodules more than CAD when used alone. CAD detected 13.7 % nodules less than A and B using 5.0 mm reconstruction images and detection rate of A improved by 6.97 % with the help of CAD and the detection rate of B increased only negligibly by 2.27 % with the support of CAD (Table **1**).

Finally, influence of CAD on the absolute performance rate of readers for every reconstruction protocol was determined. As seen in Table **1**, the absolute performance rate of readers improved to a significant extent with the support of CAD for 1.0 mm reconstruction data-sets with absolute sensitivity calculated at 91.26 % and PPV of 100 % and absolute sensitivity of 94.17 % and PPV of 100 % for A + CAD and B + CAD, respectively, which was significantly higher than independent sensitivities of CAD and readers (AUC = 0.889 for A + CAD; AUC = 0.917 for B + CAD). Combined reading by radiologists and CAD also led to reduction in false positive nodule detection.

Discussion

Big nodules are obvious to the radiologists but the smaller ones may be missed. The aim of CADs is to help radiologists increase their sensitivity by detecting the missed small nodules taking over the functions of a second reader. CAD has already been used for functions such as quantification of interstitial lung disease, automated measurement of lung nodule volumes, areas and densities and for detection of pulmonary embolism [9 – 11]. Lately, trials are on for the use of CAD in detection of polyps and tumours in the large intestine with virtual colonoscopy. Studies show that CAD can serve as guide to radiologists to detect suspicious nodules [12]. To our knowledge, there has been only one study besides ours where the performance of a CAD using different reconstruction protocols has been evaluated [13]. The study shows that CAD can be used to replace second reader when used in 0.75 mm reconstruction slice thickness with 0.6 mm reconstruction increment resulting in a high detection rate of 76.2 %,

thereby, increasing the overall detection rate of the radiologists to 91.85 %. Other studies are based on other modalities of CADs. Armato et al. in their studies applied multiple thresholding-based CAD tool with a sensitivity of 70 % [14]. Zhao et al. on the other hand applied multiple thresholding along with feature extraction and classification with a sensitivity of 84.2 % [4]. Wormanns et al. based their study on region-growing designed for detection of nodules > 5 mm diameter but with a low sensitivity of only 38 % [15].

In our study, we used all nodules detected by our consensus panel of experienced radiologists (V. J. and T. J. V.) at 1.0 mm reconstruction image data-sets as reference standard. The best performance of CAD was at 1.0 mm reconstruction images with sensitivity of 80.58 % and false positive detection rate of 0.6 per scan which is well within the range of results described by other studies. The performance results increased significantly when CAD was used in combination with evaluating radiologists leading to a high sensitivity rate and minimal false positive detection. This leads to a high confidence level among radiologists making CAD a good candidate to replace the second reader. The performance of CAD independently at 2.0 mm reconstruction thickness was inferior to one at 1.0 mm but superior to 5.0 mm reconstruction images indicating that the best results for CAD can only be achieved at thin slice reconstruction image data-sets. Use of CAD at slice thickness of 5.0 mm is not recommended. Also the detection time for CAD was minimal at 1.0 mm images and increases to maximum at 5.0 mm thick images. This strongly supports our recommendation to use only thin slices for CADs. A possible hypothesis for the increasing detection time could be the increased amount of time consumed by segmentation of false positive nodules when using thick reconstruction slices.

There was poor inter-observer agreement between the readers and CAD but very high absolute performance rate when used in combination. The relative performance rate of the readers was best at 5.0 mm whereas it was poor for CAD (relative sensitivity of 86 %, 88 % and 58 % for A, B and CAD respectively). However, relative performance of CAD was best at 1.0 mm reconstruction slices and low for the readers (relative sensitivity of 66.99 %, 68.93 % and 80.58 % for A, B and CAD respectively). CAD detected a significantly higher number of smaller nodules at 1.0 mm and 2.0 mm reconstruction slices but missed smaller nodules at

76

5.0 mm reconstruction slices. On the contrary, the detection ability of the readers deteriorated with decreasing section thickness due to the reason that the readers were not able to detect smaller nodules. This suggests a significant role of combined application of NEV and readers in detection of smaller nodules. This also explains the very low false positive nodules at 1.0 mm reconstruction slices (FP = 0) in combined evaluation by A or B and CAD. In other words, the failure to detect a nodule by a reader was compensated by CAD and vice-versa. The major advantage of CAD was that it could detect nodules shadowed by partial volume effect from blood vessels and soft tissues which were missed by radiologists (Fig. **5**). Moreover, there was no significant difference in detection of central and peripheral nodules by CAD when compared to A and B who detected more peripheral nodules than central nodules. The weakness of CAD was the inconsistency in detection of nodules with regard to their size. In other words, CAD detected a nodule of a particular size but missed another of the same size in the same patient sometimes.

Our study was not performed using a low-dose protocol for lung CT examinations. Hence influence of low-dose protocol on sensitivity of CADs still remains an area for potential study.

Conclusion

Marten et al. reported in their studies that the performance of radiologists increased when using CAD prototype software in conjunction with normal CT [13]. Wormanns et al. also reported significantly increased sensitivity when using CAD as second reader [16]. Our study supports the hypothesis that CAD programs for lung nodule detection can be used as second reader in day to day clinical practice. Using thin reconstruction slices it is possible to obtain high accuracy detection with CAD in a reasonable short time. CAD completely outperformed the readers at 1.0 mm reconstruction slices in detection of nodules especially the smaller nodules. However, we do not recommend use of CAD at 5.0 mm reconstruction thickness owing to the high number of false positive nodules and increased detection time. CAD has a positive role to play at present and in future in detection and follow-up of early stage cancers and lung metastases. However, more work needs to be done with regards to reducing false positive detection by CADs. Future developments of MDCT technology with increased number of slices and increased rotation speed promise further improvement in sensitivity of CADs.

References

[1] Ries L et al. SEER Cancer statistics review, 1993 – 1996. Bethesda: National Cancer Institution, 1999
[2] Fischbach F, Knollmann F, Griesshaber V et al. Detection of pulmonary nodules by multislice computed tomography: improved detection rate with reduced slice thickness. Eur Radiol 2003; 13: 2378 – 2383
[3] Bolte H, Muller-Hulsbeck S, Riedel C et al. Ex-vivo injection technique for implanting solid pulmonary nodules into porcine lungs for multislice CT. Fortschr Röntgenstr 2004; 176: 1380 – 1384
[4] Zhao B, Gamsu G, Ginsberg MS et al. Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm. J Appl Clin Med Phys 2003; 4: 248 – 260
[5] Naidich DP, Rusinek H, McGuinness G et al. Variables affecting pulmonary nodule detection with computed tomography: evaluation with three-dimensional computer simulation. J Thorac Imaging 1993; 8: 291 – 299
[6] Diederich S, Semik M, Lentschig MG et al. Helical CT of pulmonary nodules in patients with extrathoracic malignancy: CT-surgical correlation. AJR Am J Roentgenol 1999; 172: 353 – 360
[7] Zheng B, Ganott MA, Britton CA et al. Soft-copy mammographic readings with different computer-assisted detection cuing environments: preliminary findings. Radiology 2001; 221: 633 – 640
[8] MacMahon H. Improvement in detection of pulmonary nodules: digital image processing and computer-aided diagnosis. Radiographics 2000; 20: 1169 – 1177
[9] Wormanns D, Kohl G, Klotz E et al. Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. Eur Radiol 2004; 14: 86 – 92
[10] Uppaluri R, Hoffman EA, Sonka M et al. Interstitial lung disease: A quantitative study using the adaptive multiple feature method. Am J Respir Crit Care Med 1999; 159: 519 – 525
[11] Masutani Y, MacMahon H, Doi K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. IEEE Trans Med Imaging 2002; 21: 1517 – 1523
[12] Armato SG 3rd, Giger ML, MacMahon H. Automated detection of lung nodules in CT scans: preliminary results. Med Phys 2001; 28: 1552 – 1561
[13] Marten K, Grillhosl A, Seyfarth T et al. Computer-assisted detection of pulmonary nodules: evaluation of diagnostic performance using an expert knowledge-based detection system with variable reconstruction slice thickness settings. Eur Radiol 2005; 15: 203 – 212
[14] Armato SG 3rd, Giger ML, Moran CJ et al. Computerized detection of pulmonary nodules on CT scans. Radiographics 1999; 19: 1303 – 1311
[15] Wormanns D, Fiebich M, Saidi M et al. Automatic detection of pulmonary nodules at spiral CT: clinical application of a computer-aided diagnosis system. Eur Radiol 2002; 12: 1052 – 1057
[16] Wormanns D, Beyer F, Diederich S et al. Diagnostic performance of a commercially available computer-aided diagnosis system for automatic detection of pulmonary nodules: comparison with single and double reading. Fortschr Röntgenstr 2004; 176: 953 – 958

Fig. **5** Nodule detected only by NEV (within circle) but missed by readers at all reconstruction slices.