

University of Groningen

Methods and validation of nodule measurement in a lung cancer screening

Wang, Ying

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2010

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Wang, Y. (2010). *Methods and validation of nodule measurement in a lung cancer screening*. [Thesis fully internal (DIV), University of Groningen]. [s.n.].

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 6

Benefit of consensus double reading during baseline lung cancer

CT screening

Submitted for publication in Radiology

Benefit of consensus double reading during baseline lung cancer CT screening

Ying Wang¹, MD, Rob J. van Klaveren², MD PHD, Geertruida H de Bock³, PHD, Yingru Zhao¹, MD, Rene Vernhout², MD, Anne Leusveld¹, MD, Ernst T. Scholten⁴, MD, Johnny Verschakelen⁵, MD PHD, Willem Mali⁶, MD, PhD, Harry de Koning⁷, MD PHD, Matthijs Oudkerk¹, MD PHD

1. Department of Radiology, University Medical Center Groningen, the Netherlands
2. Department of Pulmonology, Erasmus Medical Center, Rotterdam, the Netherlands
3. Department of Epidemiology, University Medical Center Groningen, Groningen, the Netherlands
4. Department of Radiology, Kennemer Gasthuis, Haarlem, the Netherlands
5. Department of Radiology, University Hospital Gasthuisberg, Leuven, Belgium
6. Department of Radiology, University of Utrecht Medical Center, the Netherlands
7. Department of Public Health, Erasmus Medical Center, Rotterdam, the Netherlands

Abstract

Purpose: To investigate the benefit of consensus double reading compared to single reading during the baseline period of a low-dose CT lung cancer screening trial.

Materials and Methods: The screening program was approved by the Minister of Health and informed consent was obtained from all participants. Consensus double reading was performed in 7,557 participants. All participants have been followed up for 2 years. The benefit of consensus double reading compared to single reading was expressed by the percentage change in the cancer detection rate, the number of additional nodules detected, follow up recommendations and the change in sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).

Results: Of the 196 participants with a positive baseline test result, 73 were true positive for lung cancer after 2-years of follow after the baseline CT scan. One false negative case was found by retrospectively review. Due to consensus double reading, the lung cancer detection rate increased by 2.8% (2/71, 95% CI: 0.7%, 11.5%), the stage I detection rate by 2.2% (1/45, 95% CI: 0.3%, 16.1%), the nodule detection rate by 23.4% (1,635/6,988, 95% CI: 22.2%, 24.7%), the chest physician referral rate decreased by 9.3% (20/216, 95% CI: 5.9%, 14.6%) and the 3-month recall rate increased by 0.5% (8/1,551, 95% CI: 0.3%, 1.0%). The sensitivity, specificity, PPV and NPV were 96.0%, 98.1%, 32.9% and 99.9% for single reading and 98.6%, 98.4%, 37.2% and 99.9% for consensus double reading, respectively.

Conclusion: There is no benefit of consensus double reading with the use of NELSON nodule management strategy at baseline screening.

Introduction

Achieving a maximum diagnostic yield in cancer screening programs is not only dependent on the image quality but also on the appropriate reading of the images. Studies in breast cancer screening have shown that radiologists can sometimes substantially differ in the interpretations of mammograms and their recommendations for further management [1-3]. Efforts to improve accuracy and to reduce variability in the interpretation may potentially increase the effectiveness of a screening program.

Several studies have investigated the benefit of double reading in breast cancer screening and shown that double reading increased the cancer detection rate with 6-15% compared with single reading [4-12]. Taking the costs of double reading into account, double reading also appeared to be more cost-effective than a single reading policy [13;14]. Although double reading is recommended for breast cancer screening today, there is inconsistency in the data reported so far as some investigators only found an increase in the cancer detection rate of 2-5% after double reading [15-17].

Inter-observer variability in reader performance for the detection and characterization of pulmonary nodules has been found to be relatively high in chest CT imaging [18-21]. So far, the value of double reading in lung cancer CT screening has, as to our knowledge, not been investigated. Question is whether double reading could improve the diagnostic performance. The purpose of this study was, therefore, to investigate the benefit of consensus double reading compared to single reading at the baseline screening round of a low-dose CT lung cancer screening trial.

Materials and Methods

Study group

This study is a side-study of a multi-centre randomised controlled low-dose CT lung cancer screening trial [22]. Participants had to be current or former smokers with a history of >15 cigarettes per day for >25 years or >10 cigarettes per day for >30 years and between 50-75 years of age. Between April 2004 and December 2006, 7,557 participants were included. Their mean age was 59 ± 6 years and 16% were women. The study has been approved by the Dutch Minister of Health and the Ethical Committees of all four participating hospitals. Written informed consent was obtained from all participants. The original approval and informed consent for the screening study included the ability to use data for future research, including the current prospective side study.

Imaging methods

All scans were performed using 16 detector-row helical CT scanners (Sensation-16, Siemens Medical systems, Forchheim, Germany; Mx8000 IDT or Brilliance 16P, Philips Medical SYSTEMS, Cleveland, OH, USA) at low dose radiation levels with the following parameters: 0.5 sec tube rotation, 0.75 mm single slice collimation and 15 mm or 18 mm table feed per rotation (pitch = 1.3 -1.5). A caudo-cranial scan direction without contrast was used. Scans were obtained from the level of the lung bases (posterior recesses) to the lung apex with the help of a scout view. Exposure settings were 20-30 mAs and 100-140 kVp depending on the weight of the participants. This corresponds to an effective radiation dose <1.6 mSv. We reconstructed axial images with 1.0 mm thickness at 0.7 mm increments. A moderate kernel was used for reconstruction (Siemens B30 filter, Philips B filter). All scans were reconstructed with a field-of-view large enough to cover the complete lung cross-section. For nodule detection digital workstations (Leonardo®, Siemens Medical Solutions, Erlangen, Germany) were used in all screening sites with U.S. Food and Drug Administration (FDA) approved, commercially available software for semi-automated volume measurements (LungCare®, Siemens Medical Solutions, version Somaris/5: VA70C-W) [23;24]. Pulmonary nodules were identified on axial thin slab maximum intensity projections (MIPs) with cine mode. The slice thickness of MIPs was 6 mm or 8 mm. Images were initially displayed with a window level of -500 HU and window width of 1500 HU, but the readers were free to alter these settings at their discretion.

Nodule management protocol

A test was positive if any non-calcified nodule (NCN) on a CT scan had a solid component > 500 mm³ (>9.8 mm in diameter) and indeterminate if the volume of the largest solid nodule or the solid component of a partial-solid nodule was between 50-500 mm³ (4.6-9.8 mm in diameter) or > 8 mm for non-solid nodules. Otherwise, the test was negative. In participants with an indeterminate test result a 3-month follow-up scan was performed to assess whether the nodule had grown. In case of growth and a volume doubling time < 400 days, the final test result for this group was positive, otherwise negative. Growth was defined as a percentage volume change (PVC) of ≥ 25% between the first and the second scan. All test positives were referred to the chest physician for work-up and diagnosis. If lung cancer was diagnosed, the participant was treated appropriately; otherwise they were scheduled for the 2nd round CT scan [22,25].

Procedure of consensus double reading

All CT scans have subsequently been read by a first and a second reader. The level of

experience of the 13 first readers ranged from no experience to more than 20 years experience in reading thoracic CT scans (median 6 years), while the two second readers both had 6 years of experience.

In the first part of the evaluation procedure, CT scans were independently read by the first and second readers. Subsequently, the second readers matched the nodules they had identified with those identified by the first readers according to location and size. After matching the nodules, the second readers compared their screening test results with those of the first readers. In case of a discrepancy, the second readers informed the first readers and they re-evaluated the CT scan to reach consensus. If no consensus was reached, arbitration from an expert radiologist with more than 20 years of experience was performed.

Reference diagnosis

Reference standard for a positive baseline test result was based on the retrospective analysis of the lung cancers were present during a 2-years period following the baseline CT scans. All screen detected and interval cancer cases identified during this follow-up period were retrospectively matched to one particular baseline pulmonary nodule. If these cancers originated from participants with large ($>500 \text{ mm}^3$) or growing nodules with a VDT < 400 days these participants were regarded as true positives. A baseline or 2nd round interval lung cancer was defined as a lung cancer detected during the one year period following a negative baseline or 2nd round test result, respectively. They were identified through linkage of the participants with the national pathology database PALGA and by active information collection from appointments, general practitioners, letters and phone calls.

Statistical analysis

The effect of consensus double reading was first of all expressed by the change in cancer detection rate, defined as the number of cancer cases detected by consensus double reading minus the number detected by single reading divided by the number detected by single reading (%; 95% confidence interval (95% CI))[26]. This was also done for stage I and II cancers to investigate if consensus double reading led to the detection of more early stage disease than single reading. Furthermore, the effect of consensus double reading was expressed by the change in the nodule detection rate and the change in the rate of 3-months follow-up recommendations and referrals to a chest physician. The diagnostic accuracy of single and consensus double reading were expressed by the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) with their 95% CI's at the participant level. Sensitivity of single and consensus double reading was defined as the ratio between the number of true positive results (participants who were diagnosed with lung

cancer during a 2-years period after a positive baseline screening test) and the number of true positive results plus the number of false negative results (interval cancers detected during the same time period). Diagnostic performance evaluations have been performed in VassarStat online calculator (<http://faculty.vassar.edu/lowry/clin1.html>).

Results

Of the 196 participants with a positive baseline test result, 73 were true positive for lung cancer after 2-years of follow after the baseline CT scan. Seventy-one cancers were detected by both readings, 2 additional cases by consensus double reading and none by the single reading only. One false negative case interval cancer was found by retrospectively review in a participant with an interval cancer. The baseline CT of this participant showed a pleural mass in the right upper lobe, which had been classified as a benign pleural plaque by both readers. Of the 71 cancer cases detected by both readings, 63.4% (45/71) was pathological stage I and of the 2 cases detected by consensus double reading only, 50% (1/2) was stage I disease. The percentage change in cancer detection was 2.8% (2/71, 95% CI: 0.1%, 11.5%) (Table 1) and the percentage change in the detection of stage I disease was 2.2% (1/45, 95% CI: 0.3%, 16.1%). In total 8,623 nodules have been detected at baseline screening, 1,635 have been additionally detected by consensus double reading, and the percentage change in nodule detection was 23.4% (1,635/6,988, 95% CI: 22.2%, 24.7%) (Table 2). During the first screening round 9,096 CT scans have been performed including the follow-up scans. Consensus double reading upgraded the follow-up recommendations in 1.8% (161/9,096) and downgraded in 2.0% (180/9,096), which eventually led to 20 less referrals to a chest physician and 8 more 3-month follow-up scans compared with single reading (Table 3). The percentage change was 9.3% (20/216, 95% CI: 5.9%, 14.6%) less referrals to a chest physician and 0.5% (8/1551, 95% CI: 0.3%, 1.0%) more 3-month follow up scans.

The sensitivity, specificity, PPV and NPV were, respectively, 95.9% (95% CI: 87.8%, 98.9%), 98.2% (95% CI: 97.9%, 98.5%), 32.9% (95% CI: 28.1%, 41.4%), 99.9% (95% CI: 99.9%, 100%) for single reading and 98.6% (95% CI: 91.7%, 99.9%), 98.4% (95% CI: 98.08%, 98.6%), 37.2% (95% CI: 30.5%, 44.5%), 99.9% (95% CI: 99.9%, 100%) for consensus double reading, respectively (Table 4).

Table 1

Change in the lung cancer detection rate by single and consensus double reading at baseline screening of the low-dose CT screening trial (NELSON) at the participant level.

		Consensus double reading		
		Positive	Negative	Total
Single	Positive	71	0	71
Reading	Negative	2	1	3
	Total	73	1	74

Percentage change in lung cancer detection rate was 2.8% (2/71, 95% CI: 0.7%, 11.5%)

Discussion

At baseline screening of the NELSON lung cancer screening trial consensus double reading led, compared to single reading, to the identification of 2.8% more subjects with lung cancer and 2.2% more stage I disease lung cancer, the detection of 23.4% more nodules, a 9.3% decline in referrals, a 0.5% increase in 3-month recall CT scans, a 2.7% increase in sensitivity and a 0.2% increase in specificity.

The most important parameter in the assessment of the effectiveness of double reading in a cancer screening program is the degree by which the cancer detection rate increases. The 2.8% increase in our study is much lower than the 6-15% observed for breast cancer screening [4-10]. The reason for the only slight increase is two-fold. First of all, low-dose CT provides a high spatial and density resolution, which makes the identification of abnormalities more easy than for mammography in which lesion identification is relatively difficult due to the small difference in density between the lesion and the surrounding normal breast tissue. Secondly, evaluation of our CT scans was based on the NELSON nodule management strategy which is an objective, software driven approach [22;25] and, as such, far more objective than the interpretation of a mammography which is rather subjective according to the Breast Imaging Reporting and Data System [27]. We also found that the detection of stage I disease increased by only 2.2% as a result of consensus double reading. It demonstrates that the clinical benefit of consensus double reading was low, although this conclusion is based on a relatively small number of cases.

Table 2

Number of additional pulmonary nodules detected by consensus double reading compared to single reading at baseline low-dose CT screening of the NELSON trial.

Nodule category	Single reading	Additional nodules detected by consensus double reading	Total
Benign nodules and nodules <50 mm ³ (<4.6 mm)	5,090	1,166	6,256
50-500 mm ³ (4.6-9.8 mm)	1,796	440	2,236
>500 mm ³ (>9.8 mm)	102	29	131
Total	6,988	1,635	8,623

Percentage change in nodule detection rate was 23.4% (1,635/6,988, 95% CI: 22.2%, 24.7%)

Another observation is that consensus double reading led to the detection of 23.4% more nodules, which is higher than the 16% increase reported by Wormanns et al. [28]. The increase in the detection of nodules could possibly be explained by the fact that first readers pay especially attention to the larger, suspicious nodules and neglect the smaller ones. Theoretically, the more nodules identified, the higher the probability to detect cancer. However, the 23.4% increase in the nodule detection rate observed in our study only translated into a 2.8% increase in the cancer detection rate due to the fact that the majority of the additional nodules detected was < 50 mm³ (< 4.6 mm in diameter) because lung cancer risk in these small nodules is very low [29].

Another way to express the added value of double reading is to evaluate its effect on the sensitivity and specificity. Double reading can be performed by two different methods: independent double reading in which the final result is in favour of a positive diagnosis by either reader, and consensus double reading in which the final result is reached either by consensus between the first and second reader or after arbitration of a third reader. In breast cancer screening independent double reading has been reported to increase the sensitivity at the expense of the specificity[5;10;15;16], while consensus double reading has been reported to increase the sensitivity without changing specificity [6;7;9]. In our study double reading

Table 3

Effect of consensus double reading on the number of participants who received a recommendation for referral to a chest physician or a 3-months recall scan in 9,096 chest CT scans of 7,557 participants during baseline period

		Consensus double reading			Total
		referral to chest physician	3-month follow up	Annual follow up	
Single reading	Referral to chest physician	172	20	24	216
	3-month follow up	13	1,402	136	1,551
	annual follow up	11	137	7,181	7,329
Total		196	1,559	7,341	9,096

The percentage change was 9.3% (20/216, 95% CI: 5.9%, 14.6%) less referrals to a chest physician and 0.5% (8/1551, 95% CI: 0.3%, 1.0%) more 3-month follow up scans compared with single reading

increased both the sensitivity and the specificity, probably due to the fact that consensus double reading not only increased the detection of more suspicious lesions but also excluded more false positive cases. It could be hypothesized that during consensus double reading readers might perform less good because they are not blinded to each other, leading to carelessness of the first reader, prejudice of the second reader and eventually to a lower accuracy. However, we did not observe this phenomenon in our study because the final sensitivity increased by consensus double reading. Computer-aided detection (CAD) devices could serve as an attractive alternative for the enormous costs and efforts of double reading. The already commercially available CAD software might in the near future assist the reader in the detection of pulmonary nodules and improves their performance.[30-33].

Consensus double reading led only to a slight increase in the lung cancer detection rate without significant change in the detection of early stage disease despite the detection of 23% more nodules. All these additional nodules require radiological evaluation which increases the reading time and screening costs. The only advantage of consensus double reading was a decline in the referral rate to the chest physicians by 9% which reduces the screening costs.

Table 4A

Sensitivity and specificity of single reading during the baseline period of low-dose CT lung cancer screening trial NELSON including the 3-months follow-up scan.

		Participants with Lung Cancer		
		Yes	No	Total
Test Results	Positive	71	135	216
	Negative	3	7,338	7,341
	Total	74	7,483	7,557

Sensitivity: 95.9% (95% CI: 87.8%, 98.9%); specificity 98.2% (95% CI: 97.9%, 98.5%); PPV 32.9% (95% CI: 28.1%, 41.4%); NPV 99.9% (95% CI: 99.9%, 100%)

Table 4B

Sensitivity and specificity of consensus double reading during the baseline period of low-dose CT lung cancer screening trial NELSON including the 3-months follow-up scan.

		Participants with Lung Cancer		
		Yes	No	Total
Test Results	Positive	73	123	196
	Negative	1	7,360	7,361
	Total	74	7,483	7,557

Sensitivity: 98.6% (95% CI: 91.7%, 99.9%); specificity 98.4% (95% CI: 98.08%, 98.6%); PPV 37.2% (95% CI: 30.5%, 44.5%); NPV 99.9% (95% CI: 99.9%, 100%).

After weighing the advantages and disadvantages of consensus doubling reading, we do not recommend consensus double reading in lung cancer screening when using the NELSON nodule management strategy. Limitation of our study is that we did not quantify the extra costs associated with double reading. Therefore, it is not possible to provide a cost-

effectiveness recommendation for consensus double reading. Secondly, participants were followed up for this study for a period of 2 years. Although some “indolent” malignancies might have been missed longer follow-up data will not change the outcome of our study.

In conclusion, no benefit of consensus double reading could be demonstrated with the use of the NELSON nodule management strategy at baseline screening of a lung cancer CT screening trial.

References

1. Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretations of mammograms. *N Engl J Med* 1994; 331:1493-1499
2. Taplin SH, Rutter CM, Elmore JG, Seger D, White D, Brenner RJ. Accuracy of screening mammography using single versus independent double interpretation. *AJR Am J Roentgenol* 2000; 174:1257-1262
3. Sickles EA, Wolverton DE, Dee KE. Performance parameters for screening and diagnostic mammography: specialist and general radiologists. *Radiology* 2002; 224:861-869
4. Anderson ED, Muir BB, Walsh JS, Kirkpatrick AE. The efficacy of double reading mammograms in breast screening. *Clin Radiol* 1994; 49:248-251
5. Ciatto S, Ambrogetti D, Bonardi R, et al. Second reading of screening mammograms increases cancer detection and recall rates. Results in the Florence screening programme. *J Med Screen* 2005; 12:103-106
6. Warren RM, Duffy SW. Comparison of single reading with double reading of mammograms, and change in effectiveness with experience. *Br J Radiol* 1995; 68:958-962
7. Thurfjell EL, Lernevall KA, Taube AA. Benefit of independent double reading in a population-based mammography screening program. *Radiology* 1994; 191:241-244
8. H.E.Deans, D.Everington, C C, A.E.Kirkpatrick., E.Lindsay. Scottish experience of double reading in the National Breast Screening Programme. *The Breast* 1998; 7:75-79
9. Anttinen I, Pamillo M, Soiva M, Roiha M. Double reading of mammography screening films--one radiologist or two? *Clin Radiol* 1993; 48:414-421
10. Harvey SC, Geller B, Oppenheimer RG, Pinet M, Riddell L, Garra B. Increase in cancer detection and recall rates with independent double interpretation of screening mammography. *AJR Am J Roentgenol* 2003; 180:1461-1467
11. Shaw CM, Flanagan FL, Fenlon HM, McNicholas MM. Consensus review of discordant findings maximizes cancer detection rate in double-reader screening mammography: Irish National Breast Screening Program experience. *Radiology*. 2009; 250:354-62
12. Hofvind S, Geller BM, Rosenberg RD, Skaane P. Screening-detected breast cancers: discordant independent double reading in a population-based screening program. *Radiology*. 2009; 253:652-60
13. Brown J, Bryan S, Warren R. Mammography screening: an incremental cost effectiveness

- analysis of double versus single reading of mammograms. *BMJ* 1996; 312:809-812
14. Leivo T, Salminen T, Sintonen H, et al. Incremental cost-effectiveness of double-reading mammograms. *Breast Cancer Res Treat* 1999; 54:261-267
 15. Ciatto S, Del Turco MR, Morrone D, et al. Independent double reading of screening mammograms. *J Med Screen* 1995; 2:99-101
 16. Denton ER, Field S. Just how valuable is double reporting in screening mammography? *Clin Radiol* 1997; 52:466-468
 17. Matson M, Hibbert J, Field S. Double reporting of screening mammograms. *Clin Radiol* 1997; 52:567
 18. Wormanns D, Diederich S, Lentschig MG, Winter F, Heindel W. Spiral CT of pulmonary nodules: interobserver variation in assessment of lesion size. *Eur Radiol* 2000; 10:710-713
 19. Burns J, Haramati LB, Whitney K, Zelefsky MN. Consistency of reporting basic characteristics of lung nodules and masses on computed tomography. *Acad Radiol* 2004; 11:233-237
 20. Leader JK, Warfel TE, Fuhrman CR, et al. Pulmonary nodule detection with low-dose CT of the lung: agreement among radiologists. *AJR Am J Roentgenol* 2005; 185:973-978
 21. Gruden JF, Ouanounou S, Tigges S, Norris SD, Klausner TS. Incremental benefit of maximum-intensity-projection images on observer detection of small pulmonary nodules revealed by multidetector CT. *AJR Am J Roentgenol* 2002; 179:149-157
 22. Xu DM, Gietema H, de KH, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. *Lung Cancer* 2006; 54:177-184
 23. Bolte H, Riedel C, Müller-Hülsbeck S, et al. Precision of computer-aided volumetry of artificial small solid pulmonary nodules in ex vivo porcine lungs. *Br J Radiol* 2007;80:414-421
 24. Wormanns D, Kohl G, Klotz E, et al. Volumetric measurement of pulmonary nodules at multi-row detector CT: in vivo reproducibility. *Eur Radiol* 2004;14:86-92
 25. van Klaveren RJ, Oudkerk M, Prokop M, et al. Management of lung nodules detected by volume CT scanning. *N Engl J Med*. 2009 Dec 3;361(23):2221-9
 26. Williams LJ, Hartwood M, Prescott RJ. Methodological issues in mammography double reading studies. *J Med Screen* 1998; 5:202-206
 27. American College of Radiology. Breast imaging reporting and data system (BI-RADS), 3rd ed.. Reston, Va: American College of Radiology 1998
 28. Wormanns D, Ludwig K, Beyer F, Heindel W, Diederich S. Detection of pulmonary nodules at multirow-detector CT: effectiveness of double reading to improve sensitivity at

- standard-dose and low-dose chest CT. *Eur Radiol* 2005; 15:14-22
29. Xu DM, van Klaveren RJ, de Bock GH, et al. Limited value of shape, margin and CT density in the discrimination between benign and malignant screen detected solid pulmonary nodules of the NELSON trial. *Eur J Radiol* 2008; 68:347-352
 30. Das M, Muhlenbruch G, Mahnken AH, et al. Small pulmonary nodules: effect of two computer-aided detection systems on radiologist performance. *Radiology* 2006; 241:564-571
 31. Yuan R, Vos PM, Cooperberg PL. Computer-aided detection in screening CT for pulmonary nodules. *AJR Am J Roentgenol* 2006; 186:1280-1287
 32. Rubin GD, Lyo JK, Paik DS, et al. Pulmonary nodules on multi-detector row CT scans: performance comparison of radiologists and computer-aided detection. *Radiology* 2005; 234:274-283
 33. Awai K, Murao K, Ozawa A, et al. Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance. *Radiology* 2004; 230:347-352

