FULL PAPER

The impact of greyscale inversion for nodule detection in an anthropomorphic chest phantom: a free-response observer study

[1,2]JOHN D THOMPSON, PhD, MSc, [1]NIGEL B THOMAS, MB BS, FRCR, [1,3]DAVID J MANNING, PhD, DSc and [1,4]PETER HOGG, MPhil

[1]Directorate of Radiology, University of Salford, Manchester, UK
[2]Radiology, University Hospitals of Morecambe Bay NHS Foundation Trust, Barrow-in-Furness, UK
[3]Lancaster Medical School, Lancaster University, Lancaster, UK
[4]Karolinska Institute, Stockholm, Sweden

Address correspondence to: Dr John David Thompson
E-mail: j.d.thompson@salford.ac.uk

Objective: The aim of this work was to assess the impact of greyscale inversion on nodule detection on posteroanterior chest X-ray images. Previous work has attempted this, with no consensus opinion formed. We assessed the value of "fast-flicking" between standard and inverted display modes for nodule detection.

Methods: Six consultant radiologists (with 5–32 years' reporting experience) completed an observer task under the free-response paradigm. An anthropomorphic chest phantom was loaded with 50 different configurations of simulated nodules (1–4 nodules per case) measuring 5, 8, 10 and 12 mm in spherical diameter; each configuration represented a single case. In addition, 25 cases contained no nodules. Images were displayed in three modes: (i) standard, (ii) inverted and (iii) fast-flicking between standard and inverted display modes. Each observer completed the study in a different order of display (i, ii, iii) using a calibrated 5-megapixel monitor. Nodules were localized with mouse clicks and ratings assigned using a 1–10 discrete slider-bar confidence scale. Rjafroc (Pittsburgh, PA) was used for data analysis; differences in nodule detection performance were considered significant at 0.05.

Results: The observer-averaged weighted jackknife alternative free-response receiver-operating characteristic figures of merit were 0.715 (standard), 0.684 (inverted) and 0.717 (fast-flicking). Random-reader fixed-case analysis revealed no statistically significant difference between any treatment pair [$F(2,8) = 1.22$; $p = 0.345$].

Conclusion: No statistically significant difference in nodule detection was found for the three display conditions.

Advances in knowledge: We have investigated the impact of fast-flicking between standard and inverted display modes for the detection of nodules. We found no benefit.

INTRODUCTION

Chest X-ray (CXR) remains the most commonly performed radiological investigation[1,2] and it is frequently used when there is a clinical suspicion of malignancy. Despite the low cost, low radiation dose and good availability relative to other imaging modalities, the CXR is limited by the inherent relative contrast attributes of lung nodules; this is reflected by the error rate for detection—20–50%.[3] The interpretation errors that occur are typically classified as search (detection), recognition or decision.[4] Errors in CXR reporting have been attributed to errors in decision rather than detection, where an eye-tracking study has shown that while most nodules are fixated (detected), it is a decision error that prevents them from being called clinically significant.[5] This is potentially detrimental to patients and a cause for concern for radiologists, since the majority of malpractice cases brought against radiologists result from a failure to correctly identify lung cancer on CXRs.[6,7]

Recent improvements in technology now give us the chance to improve the detection of clinically significant lung nodules. Digital imaging is now well established, providing improved detective quantum efficiency and better low-contrast detectability in comparison with film–screen combinations.[1,8] Edge enhancement is available with digital imaging with advantages recognized in the neonatal chest,[9] but this has been associated with an increase in quantum mottle and a "textured" appearance to the lung.[10] Dual-energy imaging,[11] image subtraction techniques[12,13] and computer-aided detection[14] have all been advocated to improve nodule detection, but none of these techniques are

available to all radiologists. Greyscale inversion is a universally available image manipulation technique that is expected to hold some advantages. It has been suggested that viewing the inverted image (black on white, photographic positive) may provide improved contrast perception, on the basis of the physiological attributes of the human visual system (HVS).[3]

The value of the inverted image for nodule detection has been investigated by several authors,[3,15–17] but no consensus opinion has been formed. We have anecdotal evidence from radiologists that suggests that "fast-flicking" between standard and inverted display modes may provide some advantages for nodule detection. We suspect that this process of fast-flicking between the two display modes may simulate movement in the image as a whole and, in doing so, help draw attention to suspicious areas, especially in the lung periphery. This is to some degree supported by the global-focal search model proposed by Kundel et al[18] and the function of the HVS. The peripheral visual field is more sensitive to motion than the central, foveal zone and also plays a role in attracting visual attention to candidate perturbations during the global evaluation of an image. Kundel et al[18] have proposed that radiologists adopt a visual strategy that captures the gist of the image in minimal time (global phase) and identifies suspicious areas to evaluate in greater detail in the focal phase. However, they comment that this search strategy may have limited the potential for nodule detection. If the attention of a radiologist can be enhanced by a simulation of

movement (*i.e.* by fast-flicking) in the global phase, then it is possible that search and nodule detection could be improved.

In this nodule detection study, we aimed to determine the impact of greyscale inversion on performance by assessing the precise localization of nodules in three distinct display modes: standard display, inverted display and fast-flicking between standard and inverted display.

METHODS AND MATERIALS
Ethical approval for this study was granted by the University of Salford (HSCR13/27). At the time of the study, Governance Arrangements for Research Ethics Committees (GAfREC) approval was not required. Permission to conduct the study was granted by the Director of Research and Innovation at Central Manchester University Hospitals Foundation Trust. A full and comprehensive study of contrast sensitivity[19] has guided the underlying principles adopted in this work. In the present study, we aimed to detect differences in the nodule detection performance; if differences are found, a follow-up evaluation should consider the impact on search.

Image acquisition
A digital flat-panel detector with a resolution of 3000×3000 pixels (pixel size 0.143 mm) and an X-ray tube (1.5-mm Al filtration, 1.2-mm focal spot size) (Varian Medical Systems, Palo Alto, CA) was used to acquire posteroanterior (PA) images of an anthropomorphic chest phantom (Lungman N1 Multipurpose

Figure 1. The positions of all nodules within the anthropomorphic chest phantom: nodules were distributed randomly throughout the lung fields of the phantom.
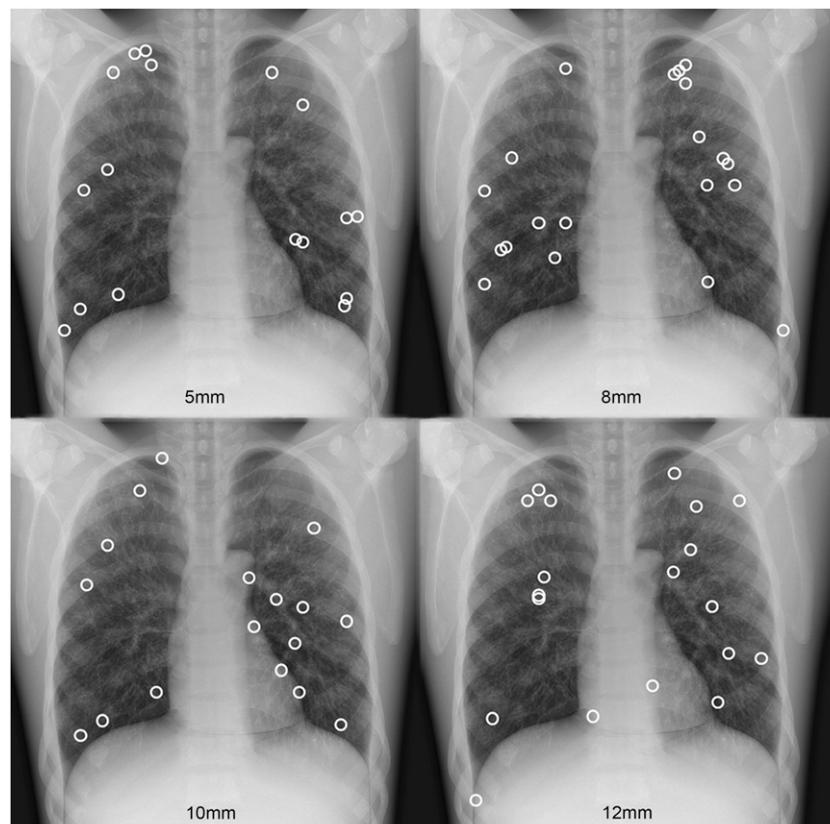
Table 1. The evaluation order of the six observers

Observer (years of experience)	Evaluation 1	Evaluation 2	Evaluation 3
1 (12)	Standard	Inverted	Fast-flicking
2 (30)	Inverted	Fast-flicking	Standard
3 (7)	Fast-flicking	Standard	Inverted
4 (32)	Standard	Fast-flicking	Inverted
5 (14)	Fast-flicking	Inverted	Standard
6 (5)	Inverted	Standard	Fast-flicking

Images were displayed in a different randomized order for each of the 18 image evaluations.

Chest Phantom; Kyoto Kagaku Company Ltd, Kyoto, Japan). All images were acquired at a source/detector distance of 180 cm using 120-kV potential and automatic exposure control with the right chamber of the automatic exposure control (AEC) selected. No grid was used. The phantom was loaded with spherical nodules of 5-, 8-, 10- and 12-mm diameter and of 100-Hounsfield units contrast. 50 different configurations of nodule positions were simulated and each nodule size could appear only once in each case. Each configuration of nodules represented a single case. Abnormal cases contained 1–4 nodules; 26 cases contained 1 nodule, 15 cases contained 2 nodules, 6 cases contained 3 nodules and 3 cases contained 4 nodules. In total, there were 22×5-mm nodules, 21×8-mm nodules, 22×10-mm nodules and 21×12-mm nodules. The average size was 8.7 ± 2.6 mm in diameter. 25 normal cases contained no nodules. Figure 1 explains the distribution of nodules on the basis of size. We distributed the nodules evenly within the lung fields of the chest phantom. Previous studies have used clinical images or clinically relevant positioning. Since we had only four nodule types, there would have been an excessive repetition of nodule positions/appearances if we had positioned them predominantly in the outer third of the lower lobes, where lesions are typically reported.[20] A phantom simulation allows us to compare the three display modes without the influence of case variation while the conspicuity and true nodule positions are tightly controlled. However, this is balanced against a consequent restriction of the search challenge because of a case test set that is more predictable than images of patients.

Image display

Images were displayed on a 23-inch 10-bit digital imaging and communications in medicine-compliant monitor (Eizo Radi-Force GS521 23-inch 10-bit digital imaging and communications in medicine-compliant monitor; Eizo Corporation, Japan) (2048×2560 pixels, 5-megapixel resolution) using ROCView (Bury St Edmunds, UK)[21] for the collection of free-response data (mark-rating pairs). The monitor was calibrated to the greyscale display function standard using the UX1 Sensor (Eizo Corporation, Hakui, Japan). A calibration check was performed prior to all image evaluations. An image inversion function was developed for ROCView to allow fast-flicking between normal and inverted states. Observers were restricted from making any further processing of the image.

Removing case variation (in a phantom study, the case is same for every CXR, only the nodule positions are changing) required us to ensure that advancing from one case to the next in the observer study did not allow the observers to see a nodule "appearing" in the phantom where it had not previously been—*i.e.* when advancing from a normal case containing no nodules to an abnormal case. To overcome this, we introduced a blanking image between cases. We used a white snow (noise) image (replicating the random dot pixel pattern of an analogue television) for 2.5 s between each case. Testing with one of the co-authors of this article revealed that this was successful.

Figure 2. Case 33 from the observer study in standard and inverted display modes containing three simulated nodules (arrows): a 10-mm nodule is identified in the right lung field, a 5-mm nodule is identified behind the left clavicle and an 8-mm nodule is identified in the left 6/7th anterior intercostal space.

Table 2. The measured diameter of all nodules

Nodule size (mm)	5	8	10	12
Mean diameter (pixels)	37.3 ± 2.9	58.4 ± 2.2	76.9 ± 4.5	93.9 ± 3.9
Mean diameter (mm)	5.3	8.4	11.0	13.4

The mean of the largest nodule (12 mm) was used to set the acceptance radius.

Observer performance study

Six consultant radiologists (with $16 ± 11$ years' reporting experience, range 7–32 years), who routinely use greyscale inversion when reporting, completed the observer performance study. All observers had significant and current experience in a wide range of radiological procedures, including CXR reporting. In addition, Observers 2 and 4 have greater expertise in head and neck imaging and nuclear medicine, respectively. Each observer completed three image evaluations: (i) standard display, (ii) inverted display and (iii) fast-flicking between standard and inverted display. The study was completed in a different order by each observer (Table 1). Image evaluations lasted approximately 75 min, breaks were permitted and no time restriction was enforced. All observers received training to introduce them to the normal appearance of the phantom (containing no nodules) and several different nodule appearances in normal and inverted states. The training images were not used in the final study. We did not expect case memory to have an impact, since the case remained the same and only the nodule positions changed. However, we did enforce a minimum 2-week gap between evaluations.

Observers were informed that there were 25 normal cases and 50 abnormal cases, where the abnormal cases could contain 1–4 nodules (Figure 2). Each abnormal case contained a different configuration of nodules. The free-response receiver-operating characteristic method was used to collect mark-rating pairs. Localizations were made using mouse clicks; this would then prompt a slider-bar confidence scale (1–10) to appear. An acceptance radius was used to classify localizations as lesion localization or non-lesion localization. The acceptance radius was determined by the largest nodules, as per current recommendations in using this methodology.[22] To establish the correct acceptance radius, nodule diameter was measured in ImageJ (Bethesda, MD).[23] Owing to magnification, the measured nodule sizes were slightly larger than the manufacturer specification. Table 2 describes the mean nodule diameter for all nodule sizes, where the size in millimetres was calculated from the size in pixels when the pixel size was known to be 0.143 mm. Since the average diameter of the largest nodule was close to 100 pixels, we set the acceptance radius at 50 pixels.

Statistical analysis

Observer data were analyzed using the weighted jackknife alternative free-response receiver-operating characteristic (wJAFROC) figure of merit (FOM). The wJAFROC FOM defines the weighted empirical probability that a lesion is rated higher than any mark made on a normal case (no lesion).[24] An R[25] (statistical programming language) implementation of jackknife alternative free-response receiver-operating characteristic analysis was used for the random-reader fixed-case analysis. This is available at (https://cran.r-project.org/web/packages/RJafroc/index.html). We analyzed three different display modes: (i) standard display, (ii) inverted display and (iii) fast-flicking between standard and inverted display. For a statistically significant difference in nodule detection to be declared significant, the p-value of the overall F-test should be <0.05 and the 95% confidence interval (CI) of individual treatment (display mode) pairs should not include 0. Test alpha was set at 0.05 to control the probability of Type I error.

RESULTS

The image display mode had no significant impact on nodule detection in this study $(F(2,10) = 2.34; \ p = 0.146)$. The wJAFROC FOMs for all display methods are summarized in Table 3 and Figure 3. Table 4 lists the p-values and FOM differences including the 95% CI for each display mode pair; since the result of the overall F-test is not significant and the 95% CI of each pair includes 0, the difference between each pair is not significant. The observer-averaged empirical weighted alternative free-response receiver-operating characteristic curves are presented in Figure 4. We also looked at the impact of evaluation order on the FOM. Each observer completed the study in a different display mode order. 5/6 observers achieved his/her highest FOM in his/her third (final) image evaluation; no such trend was observed in the first and second evaluations. The observer-averaged FOM for the first, second and third evaluation was 0.690, 0.675 and 0.725, respectively. The three most experienced observers achieved the highest FOM using fast-flicking; for Observers 1 and 6, this was his/her final image evaluation and for Observer 4, this was his/her second image evaluation.

Table 3. The observer-averaged weighted jackknife alternative free-response receiver-operating characteristic (wJAFROC) figure of merit (FOM) and 95% confidence interval (CI) for all image display modes

Display mode	wJAFROC FOM (95% CI)	Sensitivity	Specificity
Standard	0.710 (0.694, 0.726)	71%	84%
Inverted	0.672 (0.621, 0.723)	75%	74%
Fast-flicking	0.709 (0.657, 0.761)	69%	90%

Sensitivity and specificity were calculated using the sensitivity (HrSe) and specificity (HrSp) FOMs in Rjafroc (Pittsburgh, PA).

Figure 3. The weighted jackknife alternative free-response receiver-operating characteristic (wJAFROC) figure of merit (FOM) and 95% confidence interval (CI) for each display mode. FF, fast-flicking; Inv, inverted display; Std, standard display.
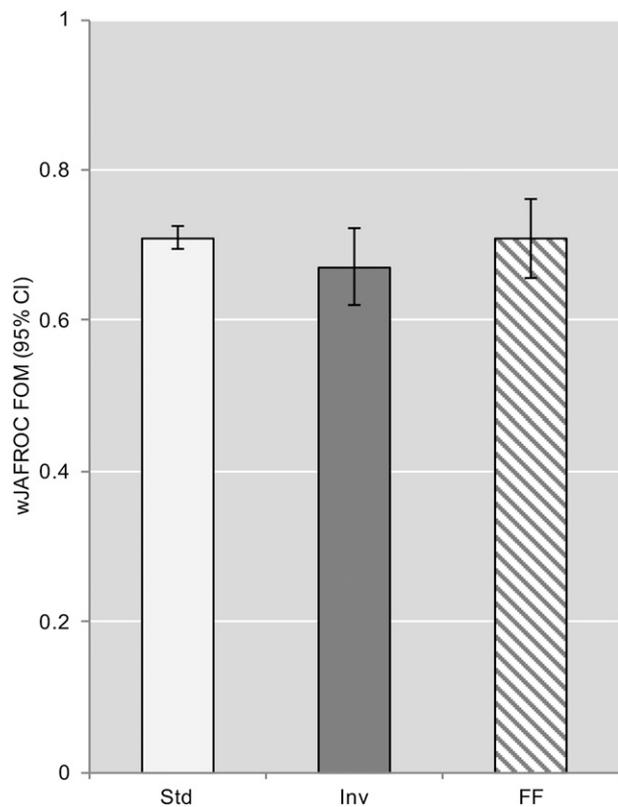
DISCUSSION

This phantom study has assessed nodule detection on PA CXR images in three display modes: standard display, inverted display and fast-flicking between standard and inverted display. We were unable to detect a significant difference in nodule detection performance on the basis of display mode. This finding is of interest, given that we removed the influence of case variability by using a phantom and we employed a robust method to assess the precise localization of nodules.

Previous work in this area has not formed a consensus opinion. Our work supports the findings of Lungren et al[3] and De Boo et al,[15] who find no advantage in nodule detection with inverted display. As in the present study, both of these studies used a location-sensitive analysis method. This makes an interesting comparison with the work of Kirchner et al[17] and Robinson et al,[16] who

both suggested an advantage of using greyscale inversion for nodule detection. They both used the receiver-operating characteristic paradigm, thus not accounting for the precise location of the nodules. Robinson et al[16] also used a small number of clinical cases ($n = 30$) and it is possible that case memory could have influenced the outcome, although they do not provide information on the order in which the images were evaluated.

We consider our work to have some methodological advantages over previous work. The result of our study is at odds with those who do not use a location-sensitive method of analysis and while Lungren et al[3] used the region-of-interest receiver-operating characteristic paradigm, this is a regional analysis and still does not account for precise localization and is associated with a limited number of ratings per case. De Boo et al[15] used the free-response receiver-operating characteristic method and jackknife alternative free-response receiver-operating characteristic analysis, but a manual (written) data collection method is suboptimal. All previous work used patient images with either real nodules[15,17] or nodules that were digitally superimposed.[3,16] The method to assess the impact of greyscale inversion has also been variable; standard display has always been compared with inversion, as isolated modalities,[16,17] isolated modalities with a third mode giving a choice between standard and inverted display[3] and with inversion as an adjunct to standard display.[15]

We used a phantom in our study and this presents challenges and provides advantages in equal measure. While we maintain a tightly controlled study, an unchanging background can cause a phenomenon where a nodule appears to "pop out". The global-focal model[18] implies that nodule "pop out" is a routine feature of radiological expertise, occurring as a perturbation from the cognitive template of a normal image, thus accounting for the fast identification of a nodule before the search phase has been completed. Clinical studies have confirmed this, where radiologists have been found to perform well in lesion detection when viewing a CXR for only 200 ms[26] and experienced readers were able to recognize most true lesions in mammography before the search phase had been completed.[27] As such, "pop out" is not exclusive to phantom studies, but the effect can be amplified when the background is always the same and this was the motivation for using a white snow (noise) image between images, to limit the influence of this effect in the observer study. However, human observers are not efficient in signal known statistically tasks, since they are influenced by inherent location uncertainty.[28]

The present study used fast-flicking between standard and inverted display modes as an additional and novel method in the

Table 4. The weighted jackknife alternative free-response receiver-operating characteristic (wJAFROC) figure of merit (FOM) differences and 95% confidence interval (CI) and *p*-value for all display mode pairs

Display mode pair	wJAFROC FOM difference (95% CI)	*p*-value
Standard *vs* inverted	0.038 (−0.007, 0.083)	0.086
Standard *vs* fast-flicking	0.001 (−0.044, 0.046)	0.955
Inverted *vs* fast-flicking	−0.037 (−0.082, 0.008)	0.095

FOM differences can be declared significant only if the 95% CI does not include 0 and the overall *F*-test is significant.

Figure 4. The observer-averaged weighted alternative free-response receiver-operating characteristic curves for the three display modes. FF, fast-flicking (yellow); Inv, inverted display (green); Std, standard display (blue). For colour image see online.
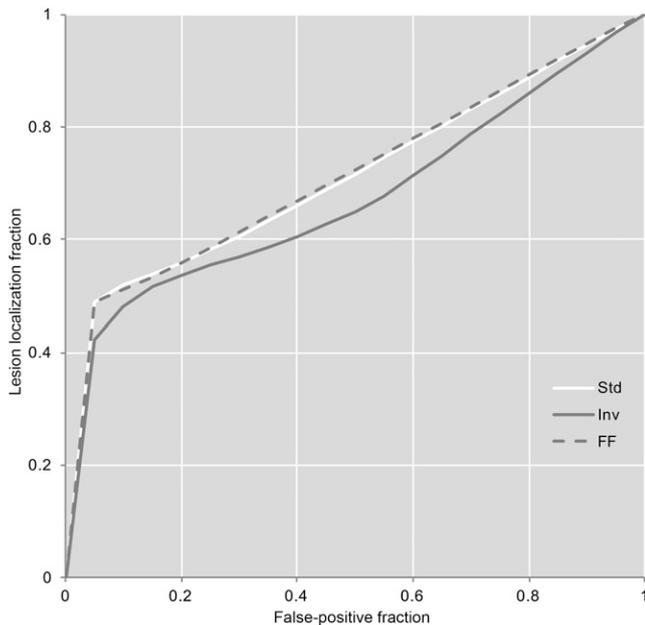
assessment of simulated lung nodule detection. Despite not finding a statistical advantage, it is valuable to explain the motivation for testing this strategy.

Fast-flicking between display modes was hypothesized to simulate motion and thus draw attention to a nodule. The peripheral visual field is more sensitive to the detection of motion than the central, foveal vision. It is a contributing factor in attracting visual attention to the peripheral field during the initial global view of a visual scene for finding, localizing and orientating candidate objects of interest. When such a stimulus occurs, the gaze is more likely to be directed towards it if movement enhances the scale of the stimulus.[29]

The retinal rod cells are more sensitive to motion than cones and are distributed towards the periphery; so, simulation of motion by fast-flicking may capture peripheral attention during the global interpretation of the image. Saccadic eye movements are a continuous phenomenon of the HVS that quickly redirect our eyes to targets such that visual acuity is maximized by the fovea;[30] these movements can be reflexive and we proposed that continuous microsaccades may be responsive to the stimulus of the changing display mode to give the illusion of motion. Eye movements tend not to be random and observers tend to move his/her eyes towards regions that are of interest to the current task;[31] for the interpretation of CXR images, the observer should identify regions in which a tumour may be hiding.[32] The present study failed to demonstrate any increased performance in nodule detection but was intended as a test of effect on performance of flicking in an observer study because of its *de facto* use in practice. It should be noted that all of our observers were experienced in CXR reporting, and all routinely used greyscale inversion. It is possible that expertise is an influential factor, but more work is required to determine whether radiologists are immune to this type of manipulation. Our follow-up study using eye-tracking and other methods of induced motion will test the effects of change detection on search in a patient-based radiology task with a variety of different lesion types.

There appears to be a growing body of evidence to suggest that greyscale inversion does not improve nodule detection in the thorax. Familiarity with the standard display mode may be an overriding factor.

CONCLUSION
We were unable to detect a statistically significant advantage of greyscale inversion or fast-flicking over standard display for the detection of simulated nodules in a chest phantom.

ACKNOWLEDGMENTS
We are grateful to the six radiologists who took part in this research.

REFERENCES

1. Bacher K, Smeets P, Bonnarens K, De Hauwere A, Verstraete K, Thierens H. Dose reduction in patients undergoing chest imaging: digital amorphous silicon flat-panel detector radiography *versus* conventional film-screen radiography and phosphor-based computed radiography. *AJR Am J Roentgenol* 2003; **181**: 923–9.
2. Raoof S, Feigin D, Sung A, Raoof S, Irugulpati L, Rosenow EC 3rd. Interpretation of plain chest roentgenogram. *Chest* 2012; **141**: 545–58. doi: http://dx.doi.org/10.1378/chest.10-1302
3. Lungren MP, Samei E, Barnhart H, McAdams HP, Leder RA, Christensen JD, et al. Gray-

scale inversion radiographic display for the detection of pulmonary nodules on chest radiographs. *Clin Imaging* 2012; **36**: 515–21. doi: http://dx.doi.org/10.1016/j.clinimag.2012.01.009
4. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol* 1978; **13**: 175–81. doi: http://dx.doi.org/10.1097/00004424-197805000-00001
5. Manning DJ, Ethell SC, Donovan T. Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. *Br J Radiol* 2004; **77**: 231–5.

6. Baker SR, Patel RH, Yang L, Lelkes VM, Castro A 3rd. Malpractice suits in chest radiology: an evaluation of the histories of 8265 radiologists. *J Thorac Imaging* 2013; **28**: 388–91. doi: http://dx.doi.org/10.1097/RTI.0b013e3182a21be2
7. Fardanesh M, White C. Missed lung cancer on chest radiography and computed tomography. *Semin Ultrasound CT MR* 2012; **33**: 280–7. doi: http://dx.doi.org/10.1053/j.sult.2012.01.006
8. Bacher K, Smeets P, Vereecken L, De Hauwere A, Duyck P, De Man R, et al. Image quality and radiation dose on digital chest imaging: comparison of amorphous silicon

and amorphous selenium flat-panel systems. *AJR Am J Roentgenol* 2006; **187**: 630–7.

9. Goo HW, Kim HJ, Song KS, Kim EA, Kim KS, Yoon CH, et al. Using edge enhancement to identify subtle findings on soft-copy neonatal chest radiographs. *AJR Am J Roentgenol* 2001; **177**: 437–40.

10. McAdams HP, Samei E, Dobbins J 3rd, Tourassi GD, Ravin CE. Recent advances in chest radiography. *Radiology* 2006; **241**: 663–83. doi: http://dx.doi.org/10.1148/radiol.2413051535

11. Kuhlman JE, Collins J, Brooks GN, Yandow DR, Broderick LS. Dual-energy subtraction chest radiography: what to look for beyond calcified nodules. *Radiographics* 2006; **26**: 79–92.

12. Tsukuda S, Heshiki A, Katsuragawa S, Li Q, MacMahon H, Doi K. Detection of lung nodules on digital chest radiographs: potential usefulness of a new contralateral subtraction technique. *Radiology* 2002; **223**: 199–203.

13. Kano A, Doi K, MacMahon H, Hassell DD, Giger ML. Digital image subtraction of temporally sequential chest images for detection of interval change. *Med Phys* 1994; **21**: 453–61. doi: http://dx.doi.org/10.1118/1.597308

14. Shiraishi J, Abe H, Engelmann R, Doi K. Effect of high sensitivity in a computerized scheme for detecting extremely subtle solitary pulmonary nodules in chest radiographs: observer performance study. *Acad Radiol* 2003; **10**: 1302–11.

15. De Boo DW, Uffmann M, Bipat S, Boorsma EF, Scheerder MJ, Weber M, et al. Gray-scale reversal for the detection of pulmonary nodules on a PACS workstation. *AJR Am J Roentgenol* 2011; **197**: 1096–100. doi: http://dx.doi.org/10.2214/AJR.11.6625

16. Robinson JW, Ryan JT, McEntee MF, Lewis SJ, Evanoff MG, Rainford LA, et al. Grey-scale inversion improves detection of lung nodules. *Br J Radiol* 2013; **86**: 20110812. doi: http://dx.doi.org/10.1259/bjr/27961545

17. Kirchner J, Gadek D, Goltz JP, Doroch-Gadek A, Stückradt S, Liermann D, et al. Standard *versus* inverted digital luminescence radiography in detecting pulmonary nodules: a ROC analysis. *Eur J Radiol* 2013; **82**: 1799–803. doi: http://dx.doi.org/10.1016/j.ejrad.2013.05.001

18. Kundel HL, Nodine CF, Thickman D, Toto L. Searching for lung nodules. A comparison of human performance with random and systematic scanning models. *Invest Radiol* 1987; **22**: 417–22.

19. Barten PG. *Contrast sensitivity of the human eye and its effects on image quality*. Vol. **72**: SPIE Press; 1999. doi: http://dx.doi.org/10.1117/3.353254

20. Zerhouni E. Computed tomography of parenchymal lung disease. In: Higgins CB, Pettersson H, eds. *Chest and Cardiac Radiology Volume 1*. London, UK: Merit Communications; 1991. p. 49–54.

21. Thompson J, Hogg P, Thompson S, Manning D, Szczepura K. ROCView: prototype software for data collection in jackknife alternative free-response receiver operating characteristic analysis. *Br J Radiol* 2012; **85**: 1320–6. doi: http://dx.doi.org/10.1259/bjr/99497945

22. Haygood TM, Ryan J, Brennan PC, Li S, Marom EM, McEntee MF, et al. On the choice of acceptance radius in free-response observer performance studies. *Br J Radiol* 2013; **86**: 42313554. doi: http://dx.doi.org/10.1259/bjr/42313554

23. Rasband W. *ImageJ*. Bethesda, Maryland, USA: U. S. National Institutes of Health; 2012. Available from: //imagej.nih.gov/ij/

24. Chakraborty DP, Berbaum KS. Observer studies involving detection and localization: modeling, analysis, and validation. *Med Phys* 2004; **31**: 2313–30. doi: http://dx.doi.org/10.1118/1.1769352

25. R Core Team. *R: a language and environment for statistical computing*, R Found Stat Comput; 2016. http://www.r-project.org/

26. Kundel HL, Nodine CF. Interpreting chest radiographs without visual search. *Radiology* 1975; **116**: 527–32. doi: http://dx.doi.org/10.1148/116.3.527

27. Nodine CF, Mello-Thoms C, Kundel HL, Weinstein SP. Time course of perception and decision making during mammographic interpretation. *AJR Am J Roentgenol* 2002; **179**: 917–23. doi: http://dx.doi.org/10.2214/ajr.179.4.1790917

28. Park S, Clarkson E, Kupinski MA, Barrett HH. Efficiency of the human observer detecting random signals in random backgrounds. *J Opt Soc Am A Opt Image Sci Vis* 2005; **22**: 3–16. doi: http://dx.doi.org/10.1364/JOSAA.22.000003

29. Monaco W, Kalb J, Johnson CA. Motion detection in the far peripheral visual field. Report ARL-MR-0684; 2007.

30. Barrett B, Elliott DB. Assessment of binocular vision. *Clinical procedures in primary eye care*: Philadelphia, PA: Elsevier; 2007. pp. 151–219.

31. Foulsham T, Kingstone A. Asymmetries in the direction of saccades during perception of scenes and fractals: effects of image type and image features. *Vision Res* 2010; **50**: 779–95. doi: http://dx.doi.org/10.1016/j.visres.2010.01.019

32. Nodine C, Kundel H. Using eye movements to study visual search and to improve tumor detection. *Radiographics* 1987; **7**: 1241–50. doi: http://dx.doi.org/10.1148/radiographics.7.6.3423330