Eur. Radiol. 9, 1826–1834 (1999)

European
Radiology

Original article

Image quality for five modern chest radiography techniques: a modified FROC study with an anthropomorphic chest phantom

L. G. Månsson[1], **S. Kheddache**[2], **B. Lanhede**[1], **U. Tylén**[2]

[1] Department of Radiation Physics, Göteborg University, Sahlgrenska University Hospital, S-413 45 Göteborg, Sweden
[2] Department of Radiology, Sahlgrenska University Hospital, S-413 45 Göteborg, Sweden

Received: 11 September 1997; Revised: 25 June 1998; Accepted: 18 February 1999

Abstract. The purpose of the study was to compare the image quality for one conventional and four digital chest radiography techniques. Three storage phosphor systems, one selenium drum system, and one film-screen system were compared using a modified receiver-operating-characteristics method. Simulated pathology was randomly positioned over the parenchymal regions and the mediastinum of an anthropomorphic phantom. Eight observers (four chest radiologists, one specialist in general radiology, one hospital physicist, and two radiographers) evaluated 60 images for each technique. The selenium drum system (Philips, Eindhoven, The Netherlands) rated best for the detection of parenchymal nodules. Together with the storage phosphor system of generation IIIN (Philips/Fuji), the selenium drum system also rated best for detection of thin linear structures. The storage phosphor system of generation V (Fuji) rated best for the detection of mediastinal nodules. The first generation of the storage phosphor system from Agfa (Mortsel, Belgium) rated worst for the detection of parenchymal nodules and thin linear structures. These differences were significant ($p < 0.0001$). Averaging the results for all test objects, the selenium drum system and the storage phosphor system of generation V were significantly better than the other systems tested. The film/screen system performed significantly better than the first-generation storage phosphor system from Agfa, equal to the generation IIIN storage phosphor system (Philips/Fuji) and significantly worse than the selenium drum system (Philips) and the generation-V storage phosphor system (Fuji). The conclusion is therefore that the image quality of selenium-based digital technique and of the more recent generations of storage phosphor systems is superior to both conventional technique and storage phosphor systems using image plates of older types.

Correspondence to: L. G. Månsson

Key words: Phantoms – Digital radiography – Technology – Receiver-operating-characteristics curve – Thorax

Introduction

During the past 15 years many efforts have been made to develop new imaging techniques in radiology. Much of the work has been focused on the development of techniques for chest imaging. Chest radiography is the most common single X-ray examination and it represents a challenge in imaging due to the large variations in density and X-ray scatter as well as the presence of structure noise (e. g., the ribs). Moreover, the anatomical and pathological structures to be visualized differ considerably in size, shape, and contrast.

The development of analog techniques include the advanced multiple-beam equalization radiography (AMBER) system [1] and asymmetric film-screen systems [2]. Digital techniques, in clinical use presently, started with digital luminescence radiography (DLR) using storage phosphor imaging plates [3]. A DLR system uses a europium-activated barium fluorohalide coating on an image plate which is exposed to radiation. Part of the energy from the exposure is stored in metastable states in the phosphor and is emitted in the form of light photons when stimulated by a laser beam. More recently, a digital radiography system for chest radiography based on a selenium drum detector has been presented [4]. The selenium detector system utilizes a drum detector that rotates during charging and readout. Radiation modulates the initially homogeneous charge distribution of the selenium drum. Readout of the charge distribution is not performed directly but by sweeping electrometer probes closely over the surface of the drum. Other digital systems under development are the IMIX [5], the Add-On Bucky from Swiss-Ray, and the direct radiography (DR) system presented by Du Pont (now Sterling Diagnostic Imaging) [6]. The

IMIX system utilizes one intensifying screen, optics, and a CCD camera. The Add-On Bucky consists of a system of scintillators, lenses, light guides, and intensified CCD cameras to form the final image from smaller parts of the image detector. The DR system from Sterling consists of a selenium plate with direct readout of the charge distribution modulated by the X-ray photons passing through the patient. Alternatives to the selenium flat panel image detectors have been presented by, for example, dpiX [7] and Trixell [8] based on light-sensitive silicon detectors in connection with Gd_2O_3- or CsI-intensifying screens.

For a long time there were doubts concerning the ability of digital methods in chest imaging to replace analog imaging techniques [9, 10, 11, 12]. The limiting factor for previous generations of DLR systems has been unfavorable noise properties [13]. One consequence is that no dose savings has been achieved with these systems. On the contrary, higher doses than for conventional techniques have been necessary for comparable image quality [14, 15].

Later generations of storage phosphor plates and the selenium drum technique are reported to have better noise properties than older generations of image plates [4, 16]. The aim of this study, therefore, was to evaluate the image quality – with respect to the visualization of low-contrast nodules and thin linear structures – of both older and newer generations of storage phosphor plates and the selenium detector system, and to compare them with a conventional film/screen system.

In this study the detection of simulated pathology using an anthropomorphic chest phantom was evaluated using a modified version of free response receiver operating characteristic (FROC) analysis. Commercially available digital systems for radiography were compared with a conventional film/screen system. Three types of storage phosphor systems, one selenium drum detector system, and one conventional film/screen system were evaluated.

Materials and methods

Radiographic equipment

The present study compares five radiographic chest systems: three storage phosphor systems [17, 18] [Agfa ADC 70 (using image plates of the first Agfa generation), Agfa Gaevert, Mortsel, Belgium; Philips PCR Ace, Philips Medical Systems, Eindhoven, The Netherlands (using image plates of generation IIIN); Fuji 9501, Fuji, Tokyo, Japan (using image plates of generation V)]; one selenium drum system (Philips Thoravision; Philips Medical Systems, Hamburg, Germany); and one film-screen system (Agfa Curix/Orto Medium HT-L, Agfa Gaevert, Mortsel, Belgium).

The Philips Thoravision and the Fuji 9501 were dedicated chest stands. A stationary 10:1 anti-scatter grid was used with Thoravision in addition to its standard 15-cm air gap, and a moving 10:1 grid was used with Fuji 9501. Images from the Agfa ADC 70 system, the Philips PCR Ace system, and the film-screen system were all produced on the same radiographic equipment – the Philips AOrigo chest stand with a moving 12:1 grid. All digital equipment produced images with a $2\,K \times 2\,K$ matrix. The screen-film system had a sensitivity corresponding to a nominal speed class of 190.

Exposure parameters

The entrance surface doses to the phantom for the Agfa ADC 70, the Fuji 9501, and the Philips PCR were 160 μGy; 180 μGy was used for the Philips Thoravision. The entrance surface dose values include backscattered radiation. When taking into account the different distances (and air gaps) used with the digital systems, and also correcting for the differences in depth doses in the phantom due to different distances, the difference in radiation dose to the image detectors between Philips Thoravision and the three other digital systems was less than 4%. To achieve optimal density, 120 μGy was used for the film-screen system. The difference in exposure between the screen-film system and the digital systems corresponds to approximately 1-mAs step.

The tube potential for all systems except the Philips Thoravision was 141 kV. These systems all had an additional filtration of 0.1 mm Cu. For the Philips Thoravision, 150 kV and an additional filtration of 0.2 mm Cu were used.

The focus-image detector distance was 150 cm for all systems except for Philips Thoravision, for which the distance was 200 cm.

Image processing

For all digital images the basic image processing parameters proposed by the manufacturers were used (see Appendix).

Hard-copy units

It would have been desirable to observe the images on video monitors of the same type and quality in order to avoid the influence of laser cameras; however, at the time of the study it was not possible to present digital images from different manufacturers on one workstation. For this reason the images had to be observed as hard copies produced by the specific laser camera coupled to each imaging system; thus, all digital images were produced on laser cameras. The Fuji 9501 images were produced on a Fujix FL-IM 2636, the Agfa ADC 70 and the Philips PCR Ace on an Agfa Matrix Compact L, and the Philips Thoravision on an Agfa Matrix LR 3300. The image formats were 43×35 cm for the ADC, PCR, and the Thoravision, 35×35 cm for the conventional radiograms, and 25×30 cm for the Fuji images. The maximum optical density for all cameras was set to be equal ($OD_{max} = 3.0$), and all cameras were

set to linearly reproduce the contrast-enhanced image information given by the respective imaging systems.

Phantom and simulated pathology

Simulated pathology was randomly positioned in three independent test regions on the back of an anthropomorphic chest phantom (RSD Torso Imaging Phantom, Radiology Support Devices Inc., Long Beach, Calif.) [19]. Three positioning matrices – not seen in the images – were defined by dividing each region into small numbered squares, the sizes of which were suitable for the different test objects. The positioning matrices covered the whole area of the right and left lung and the mediastinum. The random positioning and number of structures in these squares was calculated using a computer program. Three types of pathology were used: a large nodule (18-mm lucite hemisphere) over the mediastinum, a small nodule (6-mm lucite hemisphere) over the left lung, and a thin linear structure (0.5-mm Al thread, length 4 cm) over the right lung. Nodules were simulated by hemispheres because of the diffuse radiographic reproduction of the edges of a hemisphere, thereby more closely simulating a tumor than with a sphere.

Observer performance study

Observer performance was evaluated using a modified version of FROC analysis, the free-response forced error experiment (FFE) [20]. In an attempt to overcome the statistical limitations attributed to FROC analysis, Chakraborty [21] suggested the use of the false-positive image (FPI) in the evaluation of FROC experiments. The FPI is an image that produces one or more false-positive responses. Chakraborty proposed the area A_1 under a plot of the true-positive fraction of detected and localized objects (TPF_F) vs (p)FPI – the probability of generating an FPI – as an accuracy index. Such a plot is termed alternative FROC (AFROC). The underlying idea of the FFE experiment [20] is an analogy between the interpretation of A_z in ROC and A_1 in AFROC. Like A_z, with its two-alternative forced-choice (2AFC) counterpart, A_1 is believed to have a counterpart in an observer performance experiment "with an element of forcing to it." Furthermore, the image is characterized by a "highest-noise value," i.e., a location of "noise" somewhere in the image that looks most like a "signal." The FPI occurs if the observer sets a decision threshold below that of the highest noise-generating stimulus in the image. Thus, in an FFE experiment, the observer is asked to rank the test structures in decreasing order of confidence until he or she makes an error, thereby causing a false-positive image to be produced. A_1 can be shown to be equal to the average fraction of test structures detected in each image before the observer makes one false-positive error:

$$A_1 = \sum_{i=1}^{I} (TPF_{F,i})/I, \quad (1)$$

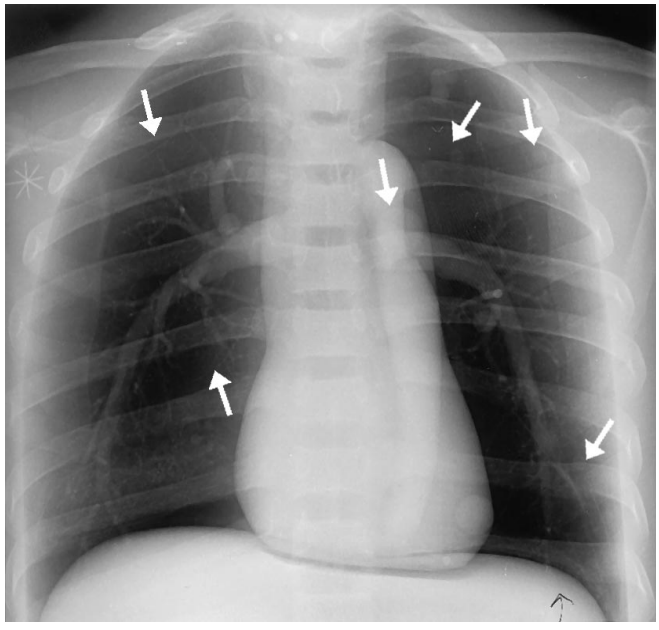

Fig. 1. Example of phantom image with test structures indicated (arrows). Left lung: small nodules; mediastinum: large nodules; right lung: thin linear structures

where $(TPF_{F,i}) = m_i/n_i$, the quotient between the number of nodules detected before the observer makes the false-positive error (m) and the total number of nodules (n) in image i. I is the total number of images in the experiment. Each image then provides a sample of A_1, and the value of A_1 for the ensemble of images is achieved by simply averaging A_1 for all images. Random performance in an FFE experiment has an A_1 value of 0, whereas perfect performance has an A_1 value of 1.

In this study 60 images of each modality were produced. An example of a phantom image with indicated test structures is shown in Fig. 1. Between one and ten test structures were positioned in the numbered squares for each test region using the computer program for random positioning. The number of structures in one region was determined independently of the number of structures in the other regions. Each linear structure was oriented in one out of four possible orientations, as given by the computer program. The four different orientations of the linear structures were known to the readers, but the possible locations of the test objects were not. In total, 320 small nodules, 319 large nodules, and 322 linear structures were used in the 60 images, giving a mean value of 5.3 structures of each type per image. The observers did not know the maximum or minimum number of structures for each region or image. The observers did not know the possible locations of the structures, only the locations in rough outline, i.e., "left lung," "right lung," and "mediastinum," respectively; thus, a search process was included in the reading.

Images from all modalities were coded, randomly mixed, and batched ten by ten. The images were decoded of their identity. For every image, the true configura-

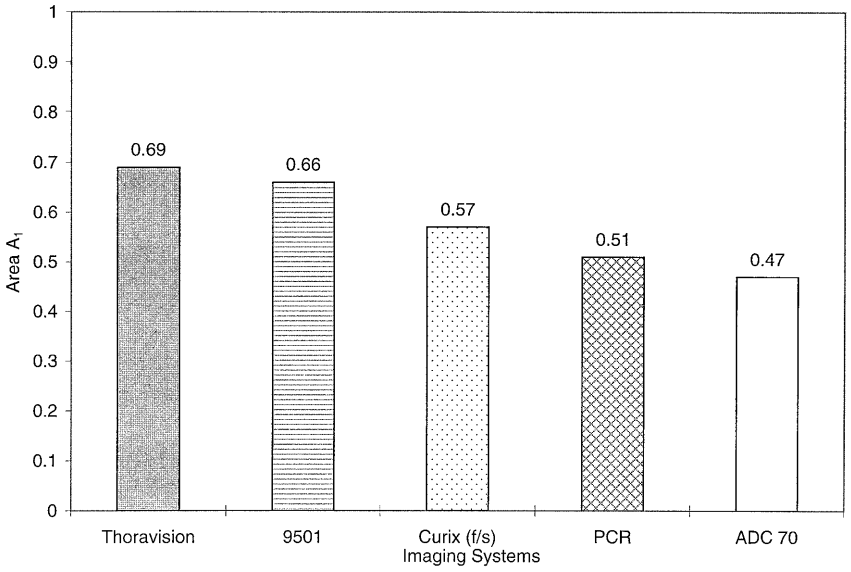

Fig. 2. Area A_1 for parenchymal nodules (6 mm) for five imaging modalities. Significant differences ($p < 0.05$) are indicated by different shadings of the bars and ranked from best (highest bar) to worst (lowest bar). Systems represented by bars with the same shading cannot be statistically separated from each other

tion of test objects was copied onto a transparent sheet to be used in the scoring procedure. During two subsequent sessions (2 days each), all the images were evaluated by eight observers: four chest radiologists, one specialist in general radiology, one hospital physicist, and two radiographers. Each observer viewed the images one by one on a personal masked light box used throughout the sessions. The readings were carried out in a room with dimmed light. The observers recorded their findings on a transparent overlay and ranked them in decreasing order of confidence (1, 2, 3, etc.). After a reading of a batch of ten images, the overlays were scored into true- and false positives in order to see if the forcing principle of the FFE experiment was met. If no false positives were found, the observer was asked to reread the actual image again using a less strict decision criterion.

A total of 2400 images were observed. The results were then calculated for each observer and modality according to Eq. (1), giving the area (A_1) under the AFROC curve for the three regions. The significance of differences between modalities was calculated using the analysis of variance (ANOVA) in conjunction with a method for multiple comparisons in order to reduce the risk of random significances, the Newman-Keuls test [22].

Results

The results of the FFE experiment are shown in Figs. 2, 3, 4, 5 and 6. The bars in the figures indicate the area, A_1, under the AFROC curve. According to the ANOVA, statistically significant differences exist between modalities within each group of structures ($p < 0.0001$). The Newman-Keuls test shows where these differences are found. Significant differences between modalities are indicated by different shading of the bars in the figures. Modalities with equal shading cannot be statistically separated from each other.

Small nodules in the parenchyma

For the detection of small nodules, statistically significant differences were seen between all five imaging systems (see Fig. 2). Philips Thoravision rated best with an A_1 value of 0.69. Fuji 9501 rated second best ($A_1 = 0.66$), followed by the film-screen system and Philips PCR, $A_1 = 0.57$ and 0.51, respectively. Agfa ADC 70 rated worst with an A_1 value of 0.47.

Large nodules in the mediastinum

For the detection of mediastinal nodules, a statistically significant difference was found only for Fuji 9501 ($A_1 = 0.83$), in comparison with all other systems, having A_1 values ranging between 0.73 and 0.76 (see Fig. 3).

Linear structures in the parenchyma

For the detection of thin linear structures, Philips Thoravision and Philips PCR rated best with A_1 values of 0.89 and 0.87, respectively (see Fig. 4). Fuji 9501 and the film-screen system rated second best with A_1 values of 0.83 for both systems. Agfa ADC 70 rated significantly worst with an A_1 value of 0.77.

All structures

Averaging the results for all simulated pathology, Philips Thoravision and Fuji 9501 rated significantly best with A_1 values of 0.78 and 0.77, respectively (see Fig. 5). The film-screen system and Philips PCR rated second best with A_1 values of 0.72 and 0.71, respectively. Agfa ADC 70 rated significantly worse with an A_1 value of 0.67.

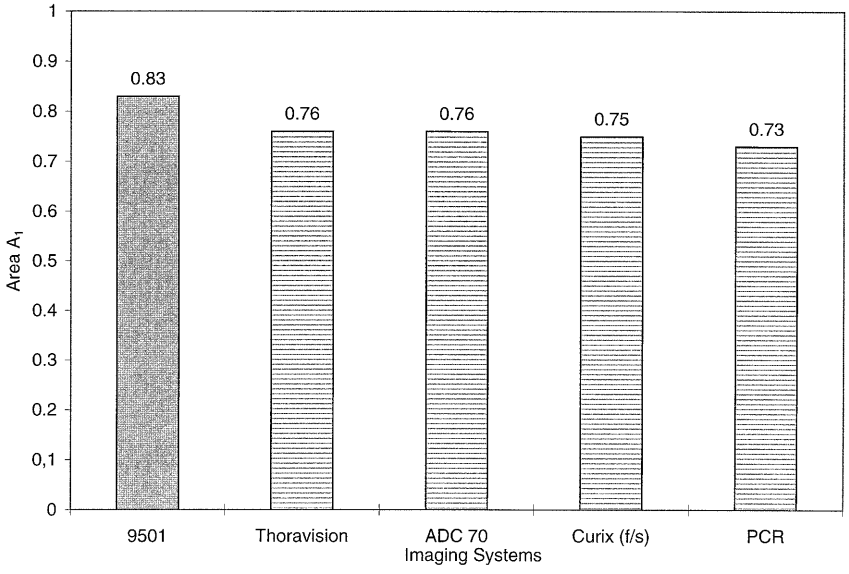

Fig. 3. Area A_1 for mediastinal nodules (18 mm) for five imaging modalities. Significant differences ($p < 0.05$) are indicated by different shadings of the bars and ranked from best (highest bar) to worst (lowest bar). Systems represented by bars with the same shading cannot be statistically separated from each other

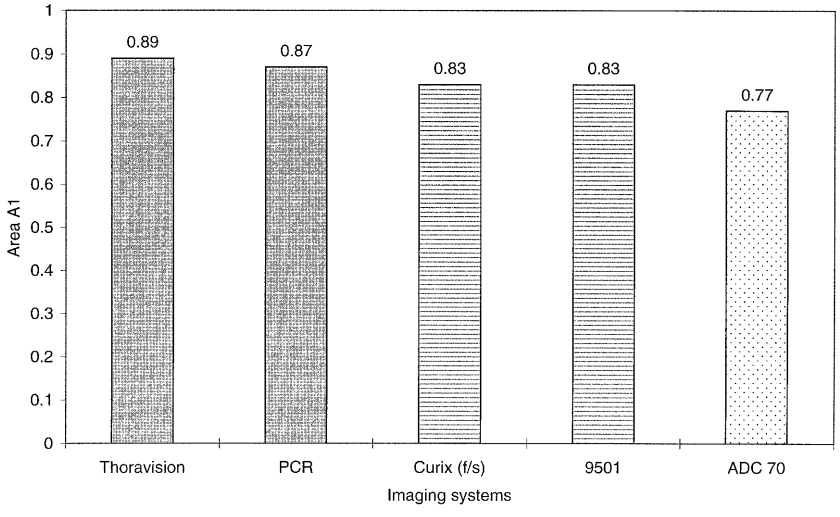

Fig. 4. Area A_1 for thin linear structures (0.5 mm) for five imaging modalities. Significant differences ($p < 0.05$) are indicated by different shadings of the bars and ranked from best (highest bar) to worst (lowest bar). Systems represented by bars with the same shading cannot be statistically separated from each other

Individual observers

In Fig. 6 the results for individual observers are seen. Averaging the results for all simulated pathology, the results are given for each observer and each imaging system. As is seen, there is no difference in detection rate between the hospital physicist and the radiologists. The detection rate of the radiographers is somewhat lower but the same trend as for the other observers is seen, with a better detection rate for Philips Thoravision and Fuji 9501 compared with the other systems. The rate of re-reading, due to the absence of a false-positive result in the first reading, was at most 10 %.

In summary, Philips Thoravision rated best for the parenchymal nodules and for the thin linear structures. Fuji 9501 rated best for the mediastinal nodules. Agfa ADC 70 rated worst for the parenchymal nodules and the thin linear structures. Averaging the results for all simulated pathology, the selenium drum system (Philips Thoravision) and generation five of the storage phosphor systems (Fuji 9501) were statistically superior to the other systems studied. The film/screen system performed significantly better than the Agfa ADC 70, equal to the Philips PCR, and significantly worse than the Philips Thoravision and Fuji 9501.

Discussion

In subjective evaluations of radiographic systems, radiologists often maintain a "conservative" attitude. They generally prefer images that they are used to, and work with in daily practice, as shown in a several studies [23, 24]. The basis for radiographic work is the detection of relevant structures. The quality of new imaging modalities should therefore always be evaluated by objective detection studies. The acceptance of a new technique producing images with resolution, contrast, and noise properties different from what radiologists are used to will therefore to a large extent depend on objective proof that detection of relevant structures with the new technique is at least as good as the old one.

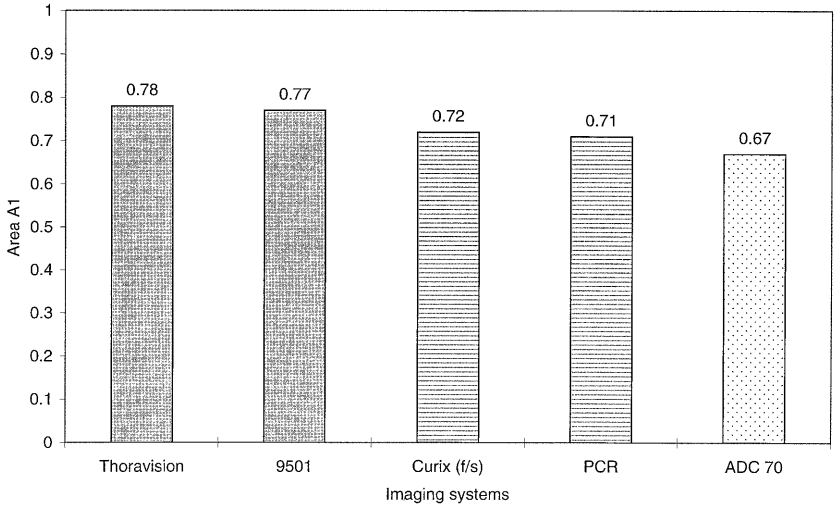

Fig. 5. Area A_1 for all test objects (parenchymal nodules 6 mm), mediastinal nodules (18 mm), and thin linear structures (0.5 mm) for five imaging modalities. Significant differences ($p < 0.05$) are indicated by different shadings of the bars and ranked from best (highest bar) to worst (lowest bar). Systems represented by bars with the same shading cannot be statistically separated from each other

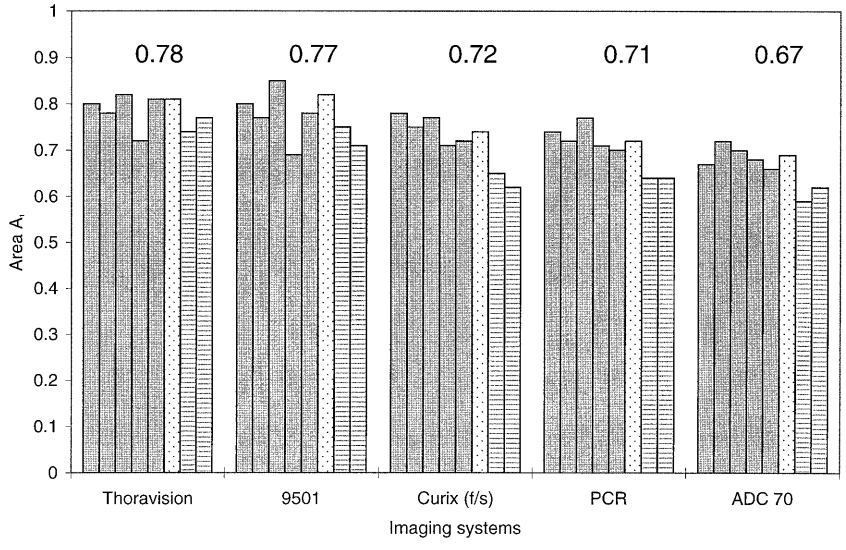

Fig. 6. A_1 values for each observer and imaging system averaged over all test structures. Observers 1–5 *(shaded gray)* are specialists in radiology, observer 6 *(dotted)* is a hospital physicist, and observers 7 and 8 (with *horizontal lines*) are radiographers. The *figures above the bars* are the average areas A_1 for all observers and all test structures for that particular imaging system

The quality of an imaging system is determined by its ability to reproduce anatomy and pathology correctly and to avoid the formation of noise-generated false-positive events. Thus, it is equally important to map the true-positive as well as the false-positive responses. As a consequence, in the evaluation of an imaging system, images have to be read to the limit of lesion detection. Although false-positive detection is a prerequisite in all ROC-based methods, it is especially crucial in the FFE experiment [20]. Here, the observer is forced to "interpret to the limit," i.e., read every image until he or she makes at least one false-positive error. In other ROC-based experiments, the observer may be reluctant to indicate other than obvious lesions. In an ROC experiment, this would result in data degeneracy problems, whereas the design of an FFE experiment ensures that both true and false-positive responses are delivered. In fact, the FFE measure could be regarded as an empirical detection accuracy measure independent of any observer performance models (ROC or FROC). If a physician detects a greater fraction of lesions before making his or her false-positive error when using images from system A than from system B, then system A is clearly superior.

In this study none of the eight observers had any difficulty in accepting the principle of forced reading. As an estimate, depending on the observer, approximately 10 % of all images had to be reread.

Calculating the results of an FFE experiment involves only an averaging of a set of numbers. In all its simplicity, this contributes to an attractive experimental design. By contrast, the calculations of conventional ROC and FROC are very complex. On the other hand, the FFE experiment is logistically demanding. Observers together with those who scored the answers had to be assembled for at least two sessions of 2 days each. The observers read the randomly batched images ten by ten. For every image a preliminary scoring was done to see that at least one false-positive answer per region had been given; if not, the observer had to read the image again. For the scoring work, sheets with the correct answers had been prepared in advance for every constellation of simulated pathology.

In FROC analysis a prerequisite for adequate use of the methodology is that the total area of test structures is small compared with the sample area. This condition is easily met for the parenchymal structures where the total area of structures is negligible in comparison with the parenchymal area. The total area of the mediastinal structures constitute at most 20 % of the mediastinal area. We have no reason to believe that this would influence the results in a significant way, especially since the FFE measure – the value of A_1 – is being used as a comparative tool only.

The observers in this study were trained radiologists as well as non-radiologists with no experience in judging chest radiograms. Several authors [25, 26] have found that if the detection tasks are well defined and involve only the detection of limited, specific pathological conditions, non-radiologists perform as well as professional radiologists. This was also the case in this study.

The parenchymal regions and the mediastinum of the anthropomorphic phantom used in this study contained little interstitial tissue, the peripheral vascular tree was sparse, and the mediastinum lacked some of the normal structures. Computed tomography measurements of the phantom give CT numbers of approximately –900 HU indicating that the parenchyma of the phantom was somewhat too "air-like" in comparison with real parenchyma (average CT numbers –700 to –800 HU) [27]. It is unlikely, however, that this would alter the outcome in studies of relative performance as in this study. The test objects used in the FFE experiment are intended to simulate pathology such as tumors and pneumothorax. They have been used in previous studies and have proven to be sufficiently difficult to detect [28].

Ideally, in a study of image quality using specific test structures, individual image processing for optimal visualization of each structure would be preferable in order to ascertain the "limit" quality of each type of equipment. However, since several items of equipment with different types of image processing parameters participated in this study, it would have been a formidable task to optimize all of them for each test object. It was therefore decided to use the image processing recommended for chest imaging by the manufacturers of the items of equipment tested. In this way, the quality for normal clinical use, rather than for a very specific setting of image processing parameters, was studied. A general description of the image processing used in the systems studied here has recently been given by Prokop and Schaefer-Prokop [29]. The Fuji-based systems in this study (9501, PCR) were not equipped with dynamic range compression (DRC). On the contrary, the non-linear unsharp masking employed in these systems almost eliminates the effect of dynamic range compression. The image processing of the Agfa system has an advantage in its ability to enhance objects according to their local contrast, independent of their size (MUSI enhancement). As a result, the dynamic range is decreased without leading to edge artifacts. However, the noise level of the first generation of the Agfa image plates was substantial, probably explaining the results for the Agfa ADC 70 in this study. The non-linear unsharp masking used for Thoravision, called dynamic range reduction (DRR), has the ability to exhibit images with wide density latitude while retaining high local contrast. However, the dynamic range compression achieved with this technique is not as extensive as with the Agfa's MUSI enhancement, which is an advantage in the visualization of the mediastinum.

The entrance surface dose used for the film-screen system is normal for a nominal 200 speed system (190 in this study). In this case the entrance surface dose corresponds to a maximum density value of 1.6 in the lung parenchyma of the phantom image. The digital image detectors studied here are essentially noise limited, i. e., their image quality is determined mainly by the radiation dose. For the detection experiment, an essential aim was therefore to expose the phantom to as equal a radiation dose as possible for all systems. With the exception of the screen-film system this was accomplished. In all cases, the normal clinical dose settings for the X-ray systems were used, as set or recommended by the manufacturers. A reduction in dose corresponding to 1 mAs step for the digital systems – thereby equalizing all doses to the image detectors – would increase the root mean square noise by approximately 12 %. In an ROC study with 16 viewers on nodule detection in digitized chest images where the noise level was artificially increased by 50 % – corresponding to halving the radiation dose – no significant difference in terms of A_z value could be found ($A_z = 0.767$ for normal and $A_z = 0.775$ for added noise) [30]. Although it is unlikely that the 12 % difference in root mean square noise present here would alter the relative results of the systems in a significant way, it cannot be excluded that the detection rate of subtle structures of the type used in this study might be slightly increased with the higher dose used for the digital systems. In any event, there would be no difference between the digital systems since they were obtained with the same radiation dose.

As described previously, all digital images had to be observed as hard copies produced by the specific laser camera coupled to each imaging system. In the preparation of the study, efforts were made to reduce the influence of the different laser cameras by setting an equal maximum optical density for all cameras and reproducing the image information given by the imaging systems linearly for all cameras. Still, despite our efforts to optimize the parameters of all laser cameras, the latter may have contributed somewhat to the differences seen between the systems.

The Fuji images were presented in a smaller format than the other images: 25×30 cm as compared with 35×35 cm for the others. All test structures were therefore presented in smaller dimensions on the Fuji images. The 6-mm nodule became 4 mm, the 18 mm nodule 14 mm. The thin linear structure of 0.5 mm became even thinner. Despite this, the Fuji system rated best for the mediastinal nodules and rated second best for the parenchymal nodules. Probably the results for the thin linear structure for the Fuji system can be explained by the fact that 0.5 mm is difficult to see when reduced to an even smaller format.

Until recently, the image quality of digital techniques in chest radiography has most often proven to be at best equal to analog techniques [9, 31, 32]. Schaefer-Prokop et al., however, have shown an improved detection of fine linear and low-contrast micronodular details for the selenium detector technique compared with both conventional and asymmetric film-screen technique as well as storage phosphor technique (Fuji, generation IIIN) [33]. This is in accordance with the results of our study which show that the selenium drum technique and also the newer generations of image plates for computed radiography, such as generation V from Fuji, are able to produce better visualization of low-contrast nodules and thin linear structures than conventional technique, even with standard settings of image processing parameters. Superior image quality with the selenium drum technique has also been reported by others [4, 34], whereas there still seems to be only a few studies of generation-V image plates. The long-awaited breakthrough for digital techniques in medical imaging with regard to image quality seems to have become a reality.

Conclusion

In conclusion, the image quality of selenium drum digital technique and of newer generations of image plates is superior both to conventional film-screen technique and to storage phosphor systems with older image plates.

Appendix: Image processing parameters

The following image processing parameters were used for the digital imaging systems in the study:

Agfa ADC: multiple scale image contrast amplification (MUSICA) parameters

Image processing parameter	Parameter value
MUSI contrast	4
Edge contrast	0
Latitude reduction	0
Noise reduction	0

Philips PCR and Fuji 9501

Image processing parameter	Parameter value
Gradient amount GA	0.8
Gradient type GT	E
Gradient center GC	1.6
Gradient shift GS	–0.20
Frequency rank RN	4
Frequency type RT	R
Frequency enhancement RE	0.5
Dynamic Range Compression	Not available

Thoravision: algorithm type „standard_grid_pa“

Image processing parameter	Parameter value
Lung density	1.7
Abdomen density	0.3
Gamma lower limit	1.7
Gamma upper limit	4.0
Detail contrast enhancement	0.8
Noise compensation	0.5

Acknowledgements. This article was supported in part by a grant from the Radiation Protection Institute, Stockholm, Sweden. The authors thank the following radiologists, physicists, and radiographers for their participation in the study: E. Bjurklint, L. Björneld, L. Denbratt, S. Gustavsson, J. Hansson, A.-S. Hoff, E. Kivilo-Carlsson, A. Lundström, C. Moudi, and A. Haglund-Olmarker.

References

1. Vlasbloem H, Schultze Kool LJ (1988) AMBER: a scanning multiple-beam equalisation system for chest radiography. Radiology 169: 29–34
2. Van Metter R (1991) Describing the signal-transfer characteristics of asymmetrical radiographic screen-film systems. Med Phys 19: 53–58
3. Sonoda M, Takano M, Miyahara J, Kato H (1983) Computed radiography utilizing scanning laser stimulated luminescence. Radiology 148: 833–838
4. Neitzel U, Maack I, Günther-Kohfahl S (1994) Image quality of a digital chest radiography system based on a selenium detector. Med Phys 21: 509–516
5. Månsson LG, Varjonen V (1995) Evaluation of a new digital method for direct chest imaging. ECR'95: Eur Radiol 5 (Suppl): S223
6. Lee DL, Cheung LK, Jeromin LS (1995) A new digital detector for projection radiography. Proc SPIE vol 2432 (Med Imaging 1995): 237–249
7. Weisfeld RL, Street RA, Apte R, Moore A (1997) An improved page-size 127 μm pixel amorphous-silicon image sensor for X-ray diagnostic medical imaging applications. Proc SPIE vol 3032 (Med Imaging 1997): 14–21
8. Chaussat C, Chabbal J, Ducourant T et al. (1998) New CsI/a-Si 17" x 17" X-ray flat panel detector provides superior detectivity and immediate direct digital output for general radiography systems. Proc SPIE vol 3336 (Med Imaging 1998): 45–56
9. Blume H, Jost RG (1992) Chest imaging within the radiology department by means of photostimulable phosphor computed radiography: a review. J Digit Imaging 5: 67–78
10. Schaefer CM, Greene R, Llewellyn HJ et al. (1991) Interstitial lung disease: impact of postprocessing in digital storage phosphor imaging. Radiology 178: 733–738
11. Schaefer CM, Greene R, Hall DA et al. (1991) Mediastinal abnormalities: detection with storage phosphor digital radiography. Radiology 178: 169–173
12. Kido S, Ikezoe J, Takeuchi N et al. (1993) Interpretation of subtle interstitial lung abnormalities: conventional versus storage phosphor radiography. Radiology 187: 527–533
13. Hillen W, Schiebel U, Zaengel T (1987) Imaging performance of a digital storage phosphor system. Med Phys 14: 744–751
14. Busch H P, Georgi M (eds) (1992) Digital radiography. Clinical experiences with digital image intensifier and storage phosphor radiography. Blackwell, Berlin, pp 29–34
15. Dobbins JR, Rice JJ, Beam CA et al. (1992) Threshold perception performance with computed and screen-film radiography: implications for chest radiography. Radiology 183: 179–187
16. Workman A, Cowen AR (1993) Signal, noise, and SNR transfer properties of computed radiography. Phys Med Biol 38: 1789–1808
17. Matsuda T, Arakawa S, Kohda K et al. (1993) New technological developments in the FCR 9000. Fuji Computed Radiography Technical Review no. 2, Fuji Photo Film Co., Ltd., Tokyo
18. Vuylsteke P, Schoeters EP (1994) Multiscale image contrast amplification (MUSICA). Proc SPIE vol 2167: 551–560

19. International Commission on Radiation Units and Measurement (1992) ICRU report 48: Phantoms and computational models in therapy, diagnosis, and protection. ICRU, Bethesda, Maryland, p 110
20. Chakraborty DP, Winter LHL (1990) Free response methodology: alternate analysis and a new observer-performance experiment. Radiology 174: 873–881
21. Chakraborty DP (1989) Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data. Med Phys 16: 561–568
22. Miller RG Jr (1980) Simultaneous statistical inference, 2nd edn. Springer, Berlin Heidelberg New York
23. Vucich J, Goodenough DJ, Lewicki A, Briefel E, Weaver KE (1980) Use of anatomical criteria in screen-film selection for portable X-ray procedures. In: Cameron (ed) Optimization of chest radiography. HHS Publication (FDA) 80–8124: 237–248
24. Burgess AE, Hicken P (1982) Comparative performance of X-ray intensifying screens. Radiology 143: 551–556
25. Gray JE, Taylor KW, Hobbs BB (1978) Detection accuracy in chest radiography. Am J Roentgenol 131: 247–253
26. Rackow PW, Spitzer VM, Hendee WR (1987) Detection of low-contrast signals: a comparison of observers with and without radiology training. Invest Radiol 22: 311–314
27. Rosenblum LJ, Maucer RA, Wellenstein DE et al. (1978) Computed tomography of the lung. Radiology 129: 521–524
28. Kheddache S, Månsson LG, Angelhed JE et al. (1991) Effects of optimization and image processing in digital chest radiography: an ROC study with an anthropomorphic phantom. Eur J Radiol 13: 143–150
29. Prokop M, Schaefer-Prokop CM (1997) Digital image processing. Eur Radiol 7 (Suppl 3):73–82
30. Herrmann C (1994) Quantenrauschen und visuelle Detailerkennbarkeit in medizinischen Röntgenaufnahmen. PTB-Bericht Opt-44. Physikalische Technische Bundesanstalt, Braunschweig
31. Cook LT, Insana MF, McFadden MA et al. (1994) Comparison of the low-contrast detectability of a screen-film system and third generation computed radiography. Med Phys 21: 691–695
32. Busch HP (1995) Digital radiography: comparison of different methods for imaging of the thorax and the gastrointestinal tracts. J Digit Imaging 8: 8–10
33. Schaefer-Prokop CM, Prokop M, Schmidt A, Neitzel U, Galanski M (1996) Selenium radiography versus storage phosphor and conventional radiography in the detection of simulated chest lesions. Radiology 201: 45–50
34. van Heeswijk HP, Neitzel U, van der Graaf Y et al. (1995) Digital chest imaging with a selenium detector: comparison with conventional radiography for visualization of specific anatomic regions of the chest. Am J Roentgenol 165: 535–540