CrossMark

CHEST

Effect of display type and room illuminance in chest radiographs

Esa Liukkonen[1,2] · Airi Jartti[1,2] · Marianne Haapea[1,2] · Heljä Oikarinen[1,2] ·
Lauri Ahvenjärvi[1,2] · Seija Mattila[1,2] · Terhi Nevala[1,2] · Kari Palosaari[1,2] ·
Marja Perhomaa[1,2] · Miika T. Nieminen[1,2,3]

Abstract

Objectives To compare diagnostic accuracy in the detection of subtle chest lesions on digital chest radiographs using medical-grade displays, consumer-grade displays, and tablet devices under bright and dim ambient light.

Methods Five experienced radiologists independently assessed 50 chest radiographs (32 with subtle pulmonary findings and 18 without apparent findings) under bright (510 lx) and dim (16 lx) ambient lighting. Computed tomography was used as the reference standard for interstitial and nodular lesions and follow-up chest radiograph for pneumothorax. Diagnostic accuracy and sensitivity were calculated for assessments carried out in all displays and compared using the McNemar test. The level of significance was set to $p < 0.05$.

Results Significant differences in sensitivity between the assessments under bright and dim lighting were found among consumer-grade displays in interstitial opacities with, and in pneumothorax without, Digital Imaging and Communication in Medicine-Grayscale Standard Display Function (DICOM-GSDF) calibration. Compared to 6 megapixel (MP) display under bright lighting, sensitivity in pneumothorax was lower in the tablet device and the consumer-grade display. Sensitivity in interstitial opacities was lower in the DICOM-GSDF calibrated consumer-grade display.

Conclusions A consumer-grade display with or without DICOM-GSDF calibration or a tablet device is not suitable for reading digital chest radiographs in bright lighting. No significant differences were observed between five displays in dim light.

Key Points
• *Ambient lighting affects performance of consumer-grade displays (with or without DICOM-GSDF calibration).*
• *Bright light decreases detection of pneumothorax on non-medical displays.*
• *Bright light decreases detection of interstitial opacities on DICOM-GSDF-calibrated, consumer-grade displays.*
• *Dim light is sufficient to detect subtle chest lesions from all displays.*

Keywords Digital radiography · Data display · Illumination · Calibration · Chest radiography

✉ Esa Liukkonen
  esa.liukkonen@ppshp.fi

  Airi Jartti
  airi.jartti@ppshp.fi

  Marianne Haapea
  marianne.haapea@oulu.fi

  Heljä Oikarinen
  helja.oikarinen@ppshp.fi

  Lauri Ahvenjärvi
  lauri.ahvenjarvi@ppshp.fi

  Seija Mattila
  seija.mattila@ppshp.fi

  Terhi Nevala
  terhi.nevala@ppshp.fi

  Kari Palosaari
  kari.palosaari@ppshp.fi

  Marja Perhomaa
  marja.perhomaa@ppshp.fi

  Miika T. Nieminen
  miika.nieminen@oulu.fi

[1] Department of Diagnostic Radiology, Oulu University Hospital, P.O. BOX 50, Fin-90029 OYS Oulu, Finland

[2] Medical Research Center Oulu, Oulu University Hospital, Oulu, Finland

[3] Medical Imaging, Physics and Technology consortium, University of Oulu, Oulu, Finland

Abbreviations

LCD	Liquid Crystal Display
Lx	Lux
MP	Megapixels
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
GSDF	Grayscale Standard Display Function
PACS	Picture Archiving and Communication Systems
AAPM	American Association of Physicists in Medicine
CR	Computed Radiography
DR	Direct Radiography

Introduction

The quality of the display has an important role in interpretation in digital radiology. Medical-grade displays offer significant advantages for diagnostic imaging compared with consumer-grade displays or tablet devices. Consumer-grade displays offer limited resolution and typically lower maximum luminance. For many applications, contrast is even more important than luminance. The number of available shades of gray on most consumer-grade displays is limited to 256 (8 bit). Medical-grade displays have a grayscale range of up to 4096 shades of gray (12 bit). The medical-grade displays also contain look-up tables calibrated for viewing images, and these systems are supported by the proper configuration and quality control tools. These tools are lacking in consumer-grade displays and tablet devices. In addition, the size of the tablet devices is much smaller (e.g. 9.7 inches) than medical-grade displays (e.g. 23 inches), which means that the image cannot be assessed as full size (pixel to pixel).

To optimize the usage of grayscale value, gamma correction of images is used by taking advantage of the non-linear manner in which humans perceive light and colour. The American Association of Physicists in Medicine (AAPM) has provided national guidelines regarding acceptable grayscale calibration to be used for radiographic interpretation [1, 2]. This guideline provides requirements that determine whether a display is suited for medical use. The basic principle is that the displays must show grayscale images according to the DICOM (Digital Imaging and Communication in Medicine) standard, using the Grayscale Standard Display Function (GSDF) [3]. This relationship ensures that the differences in grayscale are optimized for the human eye. The GSDF curve was derived from Barten's experiments with human observers determining their contrast thresholds over the complete grayscale range [4]. In contrast to the medical-grade displays, consumer-grade displays and tablet devices are not adjusted according to the DICOM-GSDF standard, which can compromise image interpretation. Most of these displays are designed based on a gamma value of 2.2.

The purpose of a radiologist's work is to detect and identify findings in an imaging examination, leading to a correct diagnosis and subsequent treatment. If diseases are not discovered, the consequences for patients can be dramatic. Chest radiography is a common examination and is often used as the initial diagnostic tool. Also, subtle lesions such as small pneumothorax or subtle interstitial disease may have important diagnostic value. Digital chest radiographs are viewed from various types of displays in various conditions, especially by clinicians. However, there is a limited knowledge of how the type of consumer-grade displays, with or without DICOM-GSDF calibration, or tablet devices, or levels of bright ambient light affect observer performance in chest radiography.

To our knowledge, four studies have determined or compared the diagnostic accuracy of displays to detect chest lesions in optimal ambient light level (20–50 lx) [5–8]: MacEntee et al concluded that for the task of identifying pulmonary nodules, the use of tablet device does not significantly change performance in a DICOM-GSDF calibrated off-the-shelf LCD [6]. Salazar et al compared a medical-grade grayscale display and two consumer grade colour displays with respect to accuracy performance with and without DICOM-GSDF calibration. For the chest conditions (interstitial opacities, pneumothorax, and nodules) and selected observers included in their study, no significant differences were observed [7]. Yin et al concluded that the observers' performances in detecting pulmonary nodules by radiologists were comparable between 2 MP, 3 MP, and 5 MP medical-grade displays [8]. Abboud et al reported that there is no difference in optimal lighting conditions between a consumer-grade display and a tablet device (second-generation iPad) in the reader's decision when diagnosing tuberculosis from digital chest radiographs; however, reading on the tablet device (a second-generation iPad) was slower [5]. There has been no study of how a more advanced tablet device (a third-generation iPad) affects the diagnostic accuracy of observing pulmonary nodules in chest radiographs. Moreover, the effect of bright ambient light, display type, and display calibration methods in chest radiographs have not been investigated within a single study. Therefore, the aim of this study was to compare observer performance in the detection of subtle chest lesions in digital chest radiographs using five different displays including a more advanced tablet device, in two different ambient light conditions: bright (510 lx) and dim (16 lx). There are a number of other studies which are focused on tablet devices and consumer-grade displays (with or without DICOM-GSDF calibration), but these studies are associated with modalities other than digital chest radiography, e.g. [9–12].

The research hypothesis was that the sensitivity and accuracy of the medical-grade displays are better than those of consumer-grade displays and tablet devices in dim and especially in bright ambient light conditions. The aim of this study

was to investigate the link between set technical display specifications and ambient light and the detectability of subtle chest lesions.

Material and methods

Image acquisition

Fifty digital chest radiographs were acquired with various computed radiography (CR) and direct radiography (DR) systems, and all the images were archived at a minimum 10 bits to the PACS (Picture Archiving and Communication System) of Department of Diagnostic Radiology, *BLINDED* (Table 1). The minimum matrix size was 3.1 megapixels (MP) and maximum 15.1 MP. Images were archived with lossless packing. The digital archives, PACS systems (neaPACS, Neagen Ltd, Finland), a custom-made case selection system (Neagen Ltd, Oulu, Finland), and a digital patient information system (ESKO, Oulu University Hospital, Oulu, Finland) were used in conjunction with an HTML4/5 viewer software (neaLink, Neagen Ltd, Oulu, Finland).

Displays and calibration

For the evaluation of chest radiographs, two identical sets of five displays were used for convenience and to save interpretation time. A standard PC (Lifebook S-761 VPro, Fujitsu, Japan, integrated graphic card: Esprimo C5731E) was connected to the consumer-grade displays. The consumer-grade displays (Fujitsu P23T6IPS) were adjusted according to the DICOM-GSDF standard and γ 2.2 (gamma 2.2) in preparation for this study using the manufacturer's internal adjustments [3, 13]. The tablet device used was the third-generation model (MD368KS/A, Apple Inc., Cupertino, CA, USA), iPad3. A 6 MP display was connected to the computer (Fujitsu Celsius R570, Fujitsu, Japan) with graphic card Barco 5200, and a 3 MP display was connected to the computer (Fujitsu Celsius R570, Fujitsu, Japan) with graphics card Nvidia Quadro FX 1800 (Table 2). In all the displays (including tablet devices) IPS (In-Plane Switching) technology was used. These displays were chosen because they are the typical products on the market.

The digital radiographs were evaluated on a 3 MP medical-grade grayscale display (Eizo Radiforce GX320-CL) and 6 MP medical-grade colour display (Barco Coronis Fusion 6MP DL), both adjusted according to the DICOM-GSDF standard.

Prior to the study, comparable maximum luminance was set by adjustments of the displays between identical displays, i.e. maximum luminance of display pairs was adjusted according to the lower maximum luminance of a display-pair. Displays luminance ranges were adjusted, which is typical of the

Table 1 Computed radiography systems and stored from chest radiographs

Manufacturer	Radiography system	Frequency	Bits stored 10	12	14
Agfa	CR	4	0	4	0
Agfa-Gevaert AG	CR	3	0	3	0
Canon Inc.	DR	2	0	2	0
Fujifilm Co.	CR	16	16	0	0
Fujifilm Corporation	CR	13	0	10	3
Philips Medical Systems	DR	10	0	10	0
Siemens	DR	2	0	2	0
Total		50	16	31	3

CR Computed Radiography, *DR* Direct Radiography

respective display type. Constant luminance was adjusted on the consumer-grade displays and tablet device using a luminance meter (RaySafe Xi; Unfors; Billdal; Sweden). Luminance for the diagnostic displays was used with factory settings. In accordance with the objectives of the study, the same luminance without auto-adjusting, as they were adjusted at the beginning of the study, was used in order to determine displays differences in characteristics at dim and bright ambient light conditions. Characteristics of the displays are summarized in Table 2. Displays were acceptance-tested for characterization purposes by a medical physicist using the AAPM TG18 test patterns [1].

Case selection and image reading

The images were taken as part of patient treatment at the Department of Diagnostic Radiology, Oulu University Hospital. The inclusion criteria for the routine digital chest radiographs for the study were that the lesions were subtle, but still distinctly visible, had clinical importance and could be validated either by computed tomography (CT) or follow-up chest radiographs. An experienced chest radiologist evaluated the suitability of the images according to the inclusion criteria and retrospectively selected 42 postero-anterior digital chest radiographs and eight antero-posterior bedside digital chest radiographs (N=50), including 32 radiographs (64 %) with lung disease findings and 18 (36 %) without apparent findings (Table 3). Cases were randomly selected without repetition and were included in the sample if chest CT scans were available to establish the reference standard. There were 18 normal cases to achieve a sample distribution similar to the patient distribution in our hospital. The total number of images selected in our study was based on previous studies containing 30–100 chest X-ray images [6, 14, 15]. All patients with interstitial lung disease and nodular opacities and one with

Table 2 Technical specifications for two identical sets of five displays and ambient light conditions

Display Group	Manufacturer	Model	Gamma correction[a]	Active screen size, diagonal (cm)	Screen type	Colour	Resolution (pixels)	Ambient light, bright (lx)	Error (%)	Ambient light, dim (lx)	Error (%)	Max. luminance of displays (cd/m²)	Error (%)
Consumer grade 1	Fujitsu	P23T6IPS	γ.2.2	58.4	LED	Yes	1920×1080	510.2	0.00	16.2	1.22	206.2	0.34
Consumer grade 2	Fujitsu	P23T6IPS	γ.2.2	58.4	LED	Yes	1920×1080	510.2		16.4		205.5	
Tablet device 1	Apple	MD368KS	γ.2.2	24.6	LED	Yes	2048×1536	513.7	0.57	16.2	0.61	344.8	3.55
Tablet device 2	Apple	MD368KS	γ.2.2	24.6	LED	Yes	2048×1536	510.8		16.3		357.5	
Consumer grade 1	Fujitsu	P23T6IPS	GSDF	58.4	LED	Yes	1920×1080	510.0	0.22	16.2	0.61	180.0	0.45
Consumer grade 2	Fujitsu	P23T6IPS	GSDF	58.4	LED	Yes	1920×1080	508.9		16.3		179.2	
6 MP 1	Barco	MDCC6130DL	GSDF	77.2	CCFL	Yes	3280×2048	509.5	0.02	16.3	1.88	391.7	4.54
6 MP 2	Barco	MDCC6130DL	GSDF	77.2	CCFL	Yes	3280×2048	509.4		16.0		374.7	
3 MP 1	Eizo	Radioforce GX320-CL	GSDF	54.1	CCFL	Monochrome	1536×2048	508.4	0.06	16.1	0.00	427.0	1.27
3 MP 2	Eizo	Radioforce GX320-CL	GSDF	54.1	CCFL	Monochrome	1536×2048	508.7		16.1		432.5	

γ.2.2 gamma2.2, GSDF Gray Scale Standard Display Function, LED Light-Emitting Diode, CCFL Cold Cathode Fluorescent Lamp, MP megapixels, Error difference between two identical displays

a Gamma correction: γ.2.2 and GSDF refer to the respective calibration (association) between pixel grayscale value and pixel intensity on display surface

Table 3 Findings of chest radiographs used in this study

Findings	n	(%)	Age of patients (years)	
			Mean (SD)	Range
Nodular opacities	9	(18)	68.9	(48-83)
Reticular opacities and pleural fluid	16	(32)	63.8	(19-81)
Pneumothorax	7	(14)	62.3	(36-85)
No findings	18	(36)	44.1	(24-65)

SD Standard Deviation

pneumothorax had chest CT scans in order to establish the reference standard. Four of the pneumothorax patients were treated with pleural drainage tubes and three were treated conservatively. All the patients with pneumothorax had follow-up chest radiography to ensure recovery. The size of the nodular opacities, as determined from the CT, varied from 6 mm to 25 mm (three cases under 7 mm, three cases 7-15 mm and three cases larger than 15 mm). The size of the pneumothorax varied from 8 mm to 28 mm. The group of interstitial diseases consisted of three cases with interstitial oedema, four cases with interstitial pneumonia, seven patients with interstitial pneumonitis (i.e. usual interstitial pneumonia, non-specific interstitial pneumonia, or desquamative interstitial pneumonia), one case with sarcoidosis, and one case with vasculitis.

Five radiologists with more than eight years of experience in general radiology were recruited. In previous studies five to eight radiologists have been used [6, 14, 15]. The radiologists were blinded to the patients' identities, conditions and findings. Each observer evaluated 50 radiographs from five displays in two ambient lighting conditions. The observers wrote their statements of findings which were later dichotomized as a finding or not a finding.

In the first session, the observers assessed images from the consumer-grade display without DICOM-GSDF calibration and tablet device, in the second session from the DICOM-GSDF-calibrated consumer-grade display, and in the third session from the DICOM-GSDF-calibrated 6 MP colour display and a 3 MP monochrome display (Table 4). The observations were made under standardized conditions, in bright (510 lx) and dim (16 lx) ambient lighting conditions. Ambient light was measured from the surface of the display in the direction of the viewer using a luminance meter (RaySafe Xi; Unfors; Billdal; Sweden). The radiographs were displayed in random order in each evaluation so as to minimize the memory effect. An evaluation time of 1 min per image was allowed. Prior to the study, the observers were familiarized with the software interface and the score sheets. To prevent potential learning bias on the part of the observers, an interval of at least 2 weeks was respected between successive evaluation sessions. Each observer evaluated each of the 50 radiographs altogether 10 times (Table 4).

Table 4 Reading sessions and viewing conditions

| Session | Reading | Images[a] | Viewing conditions (Gamma correction of displays[b]) | |
			Bright (510 lx)	Dim (16 lx)
1	1	51	CG (γ2.2)	
	2	51	Tablet device (γ2.2)	
	3	51		CG (γ2.2)
	4	51		Tablet device (γ2.2)
2	5	51	CG (GSDF)	
	6	51		CG (GSDF[b])
3	7	51	6MP (GSDF)	
	8	51	3MP (GSDF)	
	9	51		6MP (GSDF)
	10	51		3MP (GSDF)
	Total	510		

CG Consumer-grade display, *MP* megapixels, *γ2.2* gamma2.2, *GSDF* Grayscale Standard Display Function

[a] Images including QC-test and digital chest radiographs

[b] γ.2.2 and GSDF refer to the respective calibration (association) between pixel grayscale value and pixel intensity on display surface

Statistical methods

Frequency distributions of the diagnostic findings were calculated. Sensitivity, specificity, and accuracy (i.e. any finding vs. no finding) were calculated, as well as sensitivity within each diagnostic finding separately, for all the displays in both ambient lighting conditions. Only the images that were successfully read, i.e. were given a statement of a finding by an observer, in each display in both ambient lighting conditions were included in order to assure comparable results in separate analyses. The differences in sensitivities, specificities and accuracies, and sensitivity within different diagnostic findings, between dim and bright lighting were analysed using McNemar's test, as well as between the 6 MP display and other displays and between the consumer-grade display with and without DICOM-GSDF calibration. Kappa for multiple raters was calculated in order to evaluate the reliability between the observers [16]. Kappa statistics were interpreted as follows: 0.00–0.20, slight; 0.21–0.40, fair; 0.0.41–0.60, moderate; 0.61–0.80, substantial; 0.81–0.99, almost perfect agreement [17].

Results

Fifty images were read by five radiologists (i.e. 250 in total) in bright and dim ambient lighting. Because of technical problems, especially at the first reading session in which only 88 % of the image readings resulted in successful ratings, altogether 129/160 (81 %) images with findings, [34/45 (76 %) with nodular opacities, 66/80 (83 %) with interstitial opacities, and 29/35 (83 %) with pneumothorax], and 73/90 (81 %) images without findings, 202/250 (81 %) images in total, were used in the analyses.

Overall diagnostic accuracy

Sensitivities, specificities, and accuracies are presented in Table 5. Overall, sensitivity was significantly higher in dim compared to bright lighting with a consumer-grade display (70 % vs. 57 %, *p*<0.001) and DICOM-GSDF-calibrated consumer-grade display (69 % vs. 58 %, *p*=0.004), and non-significantly higher in dim compared to bright lighting with a tablet device (67 % vs. 62 %, *p*=0.263). On 6 MP or 3 MP displays there were no differences between the ambient lighting conditions (71 % vs. 70 % and 72 % vs. 71 %, respectively). The accuracies were concordant with the sensitivities, with higher accuracy in dim compared to bright lighting with a consumer-grade display (75 % vs. 67 %, *p*=0.005) and a DICOM-GSDF-calibrated consumer-grade display (75 % vs. 68 %, *p*=0.016). There were no statistically significant differences between the ambient lighting conditions in specificity in any of the displays.

With 6 MP in bright lighting, the sensitivity was 70 %, which was significantly higher as compared to the sensitivity with consumer-grade display (*p*=0.004) and a DICOM-GSDF-calibrated consumer-grade display (*p*=0.004). There were no statistically significant differences when other displays were compared with 6 MP, or between consumer-grade and DICOM-GSDF-calibrated consumer-grade displays.

Inter-reader agreement between the observers was mainly moderate (Table 5).

Visibility of different diagnostic findings

Results of the visibility of nodular opacities, interstitial opacities and pneumothorax are shown in Table 6. There were no significant differences in sensitivities between dim and bright lighting in detecting nodular opacities in any of the displays. Sensitivity was significantly higher in dim compared to bright lighting in detecting interstitial opacities with a consumer-grade display (76 % vs. 61 %, *p*=0.002) and pneumothorax in DICOM-GSDF-calibrated consumer-grade display (76 % vs. 55 %, *p*=0.031). There were no other significant differences between dim and bright lighting in detecting interstitial opacities or pneumothorax in other displays.

In bright lighting, the sensitivity of detecting interstitial opacities was significantly higher with 6 MP compared to DICOM-GSDF-calibrated consumer-grade display (68 % vs. 56 %, *p*=0.039), and for detecting pneumothorax it was significantly higher with 6 MP compared to

Table 5 Overall diagnostic accuracy in different displays under bright and dim lighting conditions

129 finding/202 images	Positives ("finding"), n=129			Negatives ("no finding"), n=73			Total, n=202			Kappa
	N	Sensitivity (95 % CI)	p	N	Specificity (95 % CI)	p	N	Accuracy (95 % CI)	p	
Consumer grade			<0.001			>0.999			0.005	
Bright	74	0.57 (0.49-0.66)		61	0.84 (0.73-0.90)		135	0.67 (0.60-0.73)		0.40
Dim	90	0.70 (0.61-0.77)		61	0.84 (0.73-0.90)		151	0.75 (0.68-0.80)		0.48
Tablet PC			0.263			0.125			>0.999	
Bright	80	0.62 (0.53-0.70)		67	0.92 (0.83-0.96)		147	0.73 (0.66-0.78)		0.47
Dim	86	0.67 (0.58-0.74)		62	0.85 (0.75-0.91)		148	0.73 (0.67-0.79)		0.37
Dicom calibrated consumer grade			0.004			>0.999			0.016	
Bright	75	0.58 (0.50-0.66)		62	0.85 (0.75-0.91)		137	0.68 (0.61-0.74)		0.42
Dim	89	0.69 (0.61-0.76)		62	0.85 (0.75-0.91)		151	0.75 (0.68-0.80)		0.40
6 MP			>0.999			>0.999			>0.999	
Bright	90	0.70 (0.61-0.77)		62	0.85 (0.75-0.91)		152	0.75 (0.69-0.81)		0.46
Dim	91	0.71 (0.62-0.78)		61	0.84 (0.73-0.90)		152	0.75 (0.69-0.81)		0.46
3 MP			0.791			0.453			>0.999	
Bright	91	0.71 (0.62-0.78)		61	0.84 (0.73-0.90)		152	0.75 (0.69-0.81)		0.45
Dim	93	0.72 (0.64-0.79)		58	0.79 (0.69-0.87)		151	0.75 (0.68-0.80)		0.40

N Number of readings concordant with golden standard, *p* Significance of the difference between bright and dim lighting conditions from McNemar's test, *Kappa* Fleiss' multirater kappa, *95 % CI* 95 % confidence interval, *MP* megapixels

consumer-grade display (72 % vs. 41 %, $p=0.012$) and tablet device (72 % vs. 48 %, $p=0.039$). There were no other significant differences in detecting specific findings when a 6 MP display was compared with the other displays. In dim lighting, sensitivity of detecting interstitial opacities was significantly higher with a consumer-grade display compared to a DICOM-GSDF-calibrated consumer-grade display (76 % vs. 65 %, $p=0.039$).

Table 6 Sensitivity under different light conditions for three different findings

	Nodular opacities, n=34			Interstitial opacities, n=66			Pneumothorax, n=29		
	N	Sensitivity (95 % CI)	p	N	Sensitivity (95 % CI)	p	N	Sensitivity (95 % CI)	p
Consumer-grade			>0.999			0.002			0.063
Bright	22	0.65 (0.48-0.79)		40	0.61 (0.49-0.71)		12	0.41 (0.26-0.59)	
Dim	23	0.68 (0.51-0.81)		50	0.76 (0.64-0.84)		17	0.59 (0.41-0.74)	
Tablet			0.250			>0.999			0.687
Bright	22	0.65 (0.48-0.79)		44	0.67 (0.55-0.77)		14	0.48 (0.31-0.66)	
Dim	25	0.74 (0.57-0.85)		45	0.68 (0.56-0.78)		16	0.55 (0.38-0.72)	
Dicom-calibrated consumer-grade			0.625			0.146			0.031
Bright	22	0.65 (0.48-0.79)		37	0.56 (0.44-0.67)		16	0.55 (0.38-0.72)	
Dim	24	0.71 (0.54-0.83)		43	0.65 (0.53-0.76)		22	0.76 (0.58-0.88)	
6MP			>0.999			>0.999			>0.999
Bright	24	0.71 (0.54-0.83)		45	0.68 (0.56-0.78)		21	0.72 (0.54-0.85)	
Dim	25	0.74 (0.57-0.85)		45	0.68 (0.56-0.78)		21	0.72 (0.54-0.85)	
3MP			0.500			0.289			0.125
Bright	24	0.71 (0.54-0.83)		47	0.71 (0.59-0.81)		21	0.69 (0.51-0.83)	
Dim	26	0.76 (0.60-0.88)		43	0.65 (0.53-0.76)		20	0.83 (0.65-0.92)	

N Number of true positives, *p* Significance of the difference between bright and dim lighting conditions from McNemar's test, *95 % CI* 95 % confidence interval, *MP* megapixels

Discussion

Recently, many studies have attempted to determine the applicability of new display devices in radiology [5, 7, 8, 10, 12, 14, 15, 18–20]. However, to date, there has been no study of how a more advanced tablet device (a third-generation iPad) affects the diagnostic accuracy of observing pulmonary nodules in chest radiographs. Moreover, the effects of bright ambient light, display type, and display calibration methods in chest radiographs have not been studied within a single study.

The results provide evidence that in the case of consumer-grade display with or without DICOM-GSDF calibration, the ambient light conditions have a significant impact on the observer's performance. With these displays, sensitivity and accuracy were significantly better in dim compared to bright ambient lighting conditions. This is similar to previous studies which indicated that ambient light that is too bright degrades the quality of the image on the display by lowering contrast and causing reflections [21–23].

In bright light conditions, sensitivity was significantly better with 6 MP and 3 MP displays compared to the consumer-grade displays with or without DICOM-GSDF calibration. On the other hand, the ambient light conditions had no influence on the tablet device, 3 MP and 6 MP displays, or else this was negligible. In addition, ambient light conditions had no influence on specificity with any displays.

The backlight for the 6 MP displays was a cold cathode fluorescence lamp (CCFL) while the other displays used light-emitting diode (LED) backlights. In a previous study on the effect of backlight, the researchers found no differences between CCFL and LED backlights in terms of diagnostic performance in chest radiology [24]. It is noteworthy that maximum luminance varied between displays, being highest for the medical-grade displays.

In the present study, nodular opacities were detected equally well in all displays in bright and dim ambient lighting. This is coherent with Pollard et al [14]. They suggest that a controlled increase of ambient lighting within 1 to 50 lx does not appear to have a statistically significant effect on nodule detection performance with a 5 MP medical-grade monochrome display. In the present study, the quality of some displays was poorer than in the study by Pollard et al, and the disparity between the lighting conditions to be compared was larger (16 lx vs. 510 lx). Moreover, McEntee et al [6] concluded that there are no significant differences between second-generation tablet devices (iPad2) compared with consumer-grade displays with DICOM-GSDF calibration in terms of identifying lung nodules on digital chest radiographs. It is noteworthy that a second-generation tablet device and in this study the third-generation tablet device used differ technically primarily in respect of the resolutions (1024×768 and 2048×1536, respectively) and pixels per inch (PPI) (132 and 264, respectively) (Table 2).

We found that contrary to the 6 MP and 3 MP displays, the bright ambient lighting (510 lx) had a significant impact on detecting interstitial opacities with a consumer-grade display and pneumothorax with a DICOM-GSDF-calibrated consumer-grade display and tablet device. In these cases, it is more advisable to use a 6 MP or 3 MP medical-grade display rather than a tablet device or a consumer-grade display. In conclusion, nine more images (out of 21) detect pneumothorax on a medical-grade display than with a consumer-grade display, and seven more images (out of 21) on a tablet device (Table 6).

Our results indicate that there are no significant differences between displays in dim light conditions. This is identical with the two studies by Salazar et al [15, 25]. They compared a medical-grade grayscale display and two consumer-grade colour displays with and without DICOM-GSDF calibration with respect to accuracy performance. For the chest lesions and a method to quantify pneumothorax size, with selected observers included in their study, no significant differences were observed. In their study, ambient light was set to 20 lx, while for our study it was set to 16 lx in dim light conditions. In their studies no third-generation tablet device was compared.

Some previous studies suggest the potential application of a tablet device for clinical purposes, such as radiological image evaluation. In these studies, the versatility of a tablet device regarding image quality and diagnostic performance was assessed for the review of tuberculosis diagnosis from digital chest radiographs [5], CT/MRI images [9–11, 26], and dental radiography [19, 20]. Hammon et al [12] concluded that the third-generation tablet device could be more useful for patient consultation, clinical demonstration or educational, and teaching purposes rather than diagnostic practice.

In this study, evaluation of the images was divided into three sessions between which there was a period of more than 2 weeks. In addition, in every session and every reading (from different displays) the images were randomized. Despite this, some kind of learning bias may have occurred because the observers evaluated the same images altogether 10 times. An attempt was made to reduce learning bias by determining the evaluation order such that the first evaluated display was technically of the lowest quality and the last evaluated display technically was of the highest quality, i.e. medical-grade displays. Learning bias may affect the results improving accuracy of medical-grade displays. However, another kind of viewing order would have resulted in more learning bias. Moreover, the loss of cases might have had an effect on the significance of outcomes.

In this study, experienced radiologists were selected because they have learned to use displays with different quality in different ambient light conditions. The aim of this study was not to compare the impact of radiologists' experience on the results. For this reason, in this study, we have not stated

how inexperienced radiologists and other physicians detect lesions in the different quality of displays.

We use in this study both CR and DR systems. It has been concluded that DR performs better than CR in terms of dose and image quality [27] and has a higher detection rate than film-screen mammography in dense breasts and for tumours of high grade [28]. However, the purpose of this study was to compare the diagnostic accuracy of displays, regardless of the system used to produce the images. After all, it would have been interesting to study how a different system affects the results in this kind of a study design. All CR and DR systems used to produce images for this study were subject to regular quality assurance tests.

Today, the major issue with using a tablet device for viewing radiological images is the lack of calibration and quality assurance and screen size, despite the fact that some applications have been developed to calibrate tablet devices to conform to the DICOM standard [29]. In the present study, we conclude that the third-generation tablet device or consumer-grade display (with or without DICOM-GSDF calibration) are not suitable replacements for medical-grade displays in chest radiology practice.

Conclusions

Subtle chest lesions may have as much clinical importance as more apparent findings and, consequently, consumer-grade display with or without DICOM-GSDF calibration or the third generation tablet device are not suitable for reading digital chest radiographs in bright ambient light conditions. For the chest conditions and selected observers included in this study, no significant differences were observed between five different displays in dim light. The effect of ambient light on observer performance with diagnostic displays was negligible as compared to consumer-grade displays with or without DICOM-GSDF calibration.

References

1. Samei E, Badano A, Chakraborty D et al (2005) Assessment of display performance for medical imaging systems: executive summary of AAPM TG18 report. Med Phys 32:1205–1225
2. Norweck JT, Seibert JA, Andriole KP et al (2013) ACR-AAPM-SIIM technical standard for electronic practice of medical imaging. J Digit Imaging 26:38–52
3. Hemminger B, Muller K (1997) A performance metric for evaluating conformance of medical image displays with the proposed ACR/NEMA display function standard. SPIE, Bellingham
4. Blume H, Steven P, Ho A et al (2003) Characterization of liquid-crystal displays for medical images - part 2. In: Galloway R (ed) Proceedings of SPIE: medical imaging 2003, vol 5029. International Society for Optical Engineering Edn, Bellingham, pp 449–473
5. Abboud S, Weiss F, Siegel E, Jeudy J (2013) TB or not TB: interreader and intrareader variability in screening diagnosis on an iPad versus a traditional display. J Am Coll Radiol 10:42–44
6. McEntee MF, Lowe J, Butler ML et al (2012) iPads and LCDs show similar performance in the detection of pulmonary nodules. In: Abbey CK, Mello- Thomas CR (ed) Proceedings of SPIE
7. Salazar AJ, Romero J, Bernal O, Moreno A, Velasco S, Diaz X (2013) Effects of the DICOM grayscale standard display function on the accuracy of medical-grade grayscale and consumer-grade color displays for telemammography screening. In: Brieva Je B (ed) Proceedings of SPIE, p 89220r
8. Yin J, Guo Q, Zhang W et al (2012) Effect of greyscale liquid crystal displays of different resolutions on observer performance during detection of small solitary pulmonary nodules. Br J Radiol 85:E549–E555
9. Yoshimura K, Nihashi T, Ikeda M et al (2013) Comparison of liquid crystal display monitors calibrated with gray-scale standard display function and with gamma 2.2 and iPad: observer performance in detection of cerebral infarction on brain CT. Am J Roentgenol 200: 1304–1309
10. Park JB, Choi HJ, Lee JH, Kang BS (2013) An assessment of the iPad 2 as a CT teleradiology tool using brain CT with subtle intra-cranial hemorrhage under conventional illumination. J Digit Imaging 26:683–690
11. Tewes S, Rodt T, Marquardt S, Evangelidou E, Wacker FK, Von Falck C (2013) Evaluation of the use of a tablet computer with a high-resolution display for interpreting emergency CT scans. Rofo 185:1066–1072
12. Hammon M, Schlechtweg PM, Schulz-Wendtland R, Uder M, Schwab SA (2014) iPads in breast imaging - a phantom study. Geburtshilfe Frauenheilkd 74:152–156
13. Barten PGJ (1999) Contrast sensitivity of the human eye and its effects on image quality. SPIE-The International Society For Optical Engineering, Washington, USA
14. Pollard BJ, Samei E, Chawla AS et al (2012) The effects of ambient lighting in chest radiology reading rooms. J Digit Imaging 25:520–526
15. Salazar AJ, Aguirre DA, Ocampo J, Camacho JC, Diaz XA (2014) DICOM gray-scale standard display function: clinical diagnostic accuracy of chest radiography in medical-grade gray-scale and consumer-grade color displays. Am J Roentgenol 202:1272–1280
16. Fleiss J, Levin B, Paik M (2003) Statistical methods for rates and proportions. John Wiley & Sons, Inc, Hoboken
17. Landis J, Koch G (1977) Measurement of observer agreement for categorical data. Biometrics 33:159–174
18. Yoshimura K, Shimamoto K, Ikeda M, Ichikawa K, Naganawa S (2011) A comparative contrast perception phantom image of brain CT study between high-grade and low-grade liquid crystal displays

(LCDs) in electronic medical charts rid B-5163-2010. Phys Med 27:109–116

19. Kallio-Pulkkinen S, Haapea M, Liukkonen E, Huumonen S, Tervonen O, Nieminen M (2014) Comparison of consumer grade, tablet and 6MP-displays: observer performance in detection of anatomical and pathological structures in panoramic radiographs. Oral Surg Oral Med Oral Pathol Oral Radiol 118:135–141

20. Pakkala T, Kuusela L, Ekholm M, Wenzel A, Haiter-Neto F, Kortesniemi M (2012) Effect of varying displays and room illuminance on caries diagnostic accuracy in digital dental radiographs. Caries Res 46:568–574

21. Pollard BJ, Chawla AS, Delong DM, Hashimoto N, Samei E (2008) Object detectability at increased ambient lighting conditions. Med Phys 35:2204–2213

22. Fetterly KA, Blume HR, Flynn MJ, Samei E (2008) Introduction to grayscale calibration and related aspects of medical imaging grade liquid crystal displays. J Digit Imaging 21:193–207

23. Samei E, Ranger N, Delong DM (2008) A comparative contrast-detail study of five medical displays. Med Phys 35:1358–1364

24. Lim H, Chung MJ, Lee G et al (2013) Interpretation of digital chest radiographs: comparison of light emitting diode versus cold cathode fluorescent lamp backlit monitors. Korean J Radiol 14:968–976

25. Salazar AJ, Aguirre DA, Ocampo J, Camacho JC, Diaz XA (2014) Evaluation of three pneumothorax size quantification methods on digitized chest x-ray films using medical-grade grayscale and consumer-grade color displays. J Digit Imaging 27:280–286

26. Mcnulty JP, Ryan JT, Evanoff MG, Rainford LA (2012) Flexible image evaluation: iPad versus secondary-class monitors for review of MR spinal emergency cases, a comparative study. Acad Radiol 19:1023–1028

27. Bosmans H, De Hauwere A, Lemmens K et al (2013) Technical and clinical breast cancer screening performance indicators for computed radiography versus direct digital radiography. Eur Radiol 23:2891–2898

28. Seradour B, Heid P, Esteve J (2014) Comparison of direct digital mammography, computed radiography, and film-screen in the French national breast cancer screening program. Am J Roentgenol 202:229–236

29. De Paepe L, De Bock P, Vanovermeire O, Kimpe T (2012) Performance evaluation of a visual display calibration algorithm for ipad. Medical Imaging 2012: Advanced Pacs-Based Imaging Informatics and Therapeutic Applications 8319: 831909