DESIGN OF A 32” COLOR TFT LCD PACS MONITOR AND ITS CLINICAL EVALUATION THROUGH ROC ANALYSIS

by

Sadık Hakan Kayabaşı

B.S., in Industrial Engineering, Boğaziçi University, 1994

Submitted to the Institute of Biomedical Engineering

in partial fulfillment of the requirements

for the degree of

Master of Science

in

Biomedical Engineering

Boğaziçi University
April 2010

DESIGN OF A 32" COLOR TFT LCD PACS MONITOR AND ITS CLINICAL EVALUATION THROUGH ROC ANALYSIS

APPROVED BY:

Assoc. Prof. Albert Güveniş (Thesis Advisor)	……
Prof. Dr. Yekta Ülgen	……
Assoc. Prof. İpek Karaaslan	…….

DATE OF APPROVAL: 06 May 2010

ACKNOWLEDGEMENTS

Firstly, I would like to thank my thesis advisor Assoc. Prof. Dr. Albert Güveniş for his encouragement and support during my thesis study. His proposal for the selection of master thesis and vision motivated me very much to launch the project in a smooth path.

I would like to thank Acıbadem Hospital Group for their support and contribution through the evaluation period of the study. I specifically appreciate the efforts of Merve Temiz for her continuous support and kind approach as the coordinator of the project. I am fully grateful to Acıbadem Bakırköy Hospital-Radiology Department's doctors who examined the data set in both existing medical monitor and my monitor in their tight and stressful schedule. So special thanks are to Dr. Hüseyin Akçam, Dr. Ulaş Can and Dr. Ali Türk whose contributions and involvement helped me reach to the final phase of the thesis study.

My thanks and gratitude deep from my heart goes to my wife, Makbule Kayabaşı, without her support and motivation, I would not dare to start the thesis study after the long break.

ABSTRACT

DESIGN OF A 32” COLOR TFT LCD PACS MONITOR AND ITS CLINICAL EVALUATION THROUGH ROC ANALYSIS

One of the most important stages in digital radiology process is the transfer of the image to the observer, as the variations in light and color from a physical display. The common practice in observing these visual results is to use a medical grade LCD monitor. The main problem with medical grade monitors is their high cost. First objective of the study was to design a 32” LCD PACS monitor to be as compatible as possible to a medical grade LCD monitor with a remarkable cost advantage. The second was to test the hypothesis for a significant difference between a medical grade LCD monitor and the designed one in terms of diagnostic image quality.

After the design’s validation, 60 digital radiographs with definite findings were obtained in cooperation with the authors of a previous study. Three experienced radiologists from Acıbadem Hospital examined these radiographs both on a medical grade Reference branded LCD monitor and on the 32” design. To check observers’ performance, the receiver operating characteristic (ROC) curves for all monitor-reader cases were statistically compared by using the same content and observers. The area under curve (AUC) of each ROC curve was used as a metric for detecting lung modules in the radiographs. With 95% confidence interval, the hypothesis was tested for a significant statistical difference between the related monitors. AUC for Reference monitor for observer 1, 2 and 3 were calculated as 0.634, 0.703 and 0.755 respectively. AUC for 32” design were 0.811, 0.746 and 0.811 For observer 1, the 32”design showed superior performance. For observer 2 and 3, Though AUC was far better on behalf of the new design, no significant statistical difference could be proven.

As a result, it is possible to implement the new 32” design as a PACS monitor for medical diagnosis purposes without sacrificing any diagnostic value.

Keywords: PACS Monitor, ROC curve, AUC

ÖZET

32” RENKLİ TFT LCD PACS MONİTÖR TASARIMI VE ROC ANALİZİ KULLANILARAK KLİNİK DEĞERLENDİRİLMESİ

Dijital radyolojinin en önemli adımlarından biri, imajın fiziksel bir medya üzerinden, ışık ve renk değişimi şeklinde radyoloğa ulaştırılmasıdır. Görsel sonuçları incelemek için genellikle bir medikal LCD monitor kullanılmaktadır. Medikal Monitörlerin ana problemi ise yüksek maliyetli olmalarıdır. Bu çalışmanın ilk amacı, maliyet avantajı sağlayacak ve medical LCD monitörlere eşlenik bir 32” LCD PACS Monitör tasarlamaktır. İkinci amaç ise mevcut bir medical monitor ile yeni monitör arasında, görüntü kalitesi açısından kayda değer bir fark olmadığı hipotezini test etmektir.

Tasarımın onaylanmasından sonra, bir bilimsel çalışmada kullanılmış ve haklarındaki bulgular kesin olarak bilinen 60 dijital röntgen elde edilmiştir. Acıbadem Hastanesi Radyoloji bölümünden üç tecrübeli radyolog, bu röntgenleri bir Reference medikal monitorde ve yeni 32” tasarımda incelemişlerdir. Radyologların performansını kontrol etmek için, monitör-radyolog eşleşme sonuçları, ROC (Receiver Operating Characteristic) eğrisi kullanılarak istatistiksel olarak değerlendirilmiştir. Her ROC eğrisinin altında kalan alan (AUC), röntgenlerdeki akciğer nodüllerini tespit etmek için ana parametre olarak alınmıştır. %95 güven aralığıyla, monitörlerin aralarında kayda değer bir fark olup olmadığı istatistiksel olarak test edilmiştir. Reference monitörde AUC değerleri, radyolog 1, 2 ve 3 için sırasıyla, 0.634, 0.703 ve 0.755 olarak hesaplanmıştır. 32” tasarım için ise AUC değerleri, aynı sırayla 0.811, 0.746 ve 0.811’dir. İlk radyoloğun değerlendirmesinde, 32” tasarım istatistiksel olarak Reference monitöre göre daha üstün performans göstermiştir. Diğer radyologların değerlendirmelerinde, AUC değerleri yeni tasarım lehine olmakla beraber, istatistiksel olarak kayda değer bir fark bulunamamıştır.

Sonuç olarak, yeni 32” tasarımın, verilerinde kayba sebep olmaksızın, teşhis amaçlı bir PACS monitörü olarak kullanılabilmesi mümkündür..

Anahtar sözcükler: PACS Monitör, ROC eğrisi, AUC

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

H_0	Null hypothesis
H_1	Alternative hypothesis
SE	Standard Error
r	Correlation coefficient of two areas for the same sample pool
n_A	Number of abnormal cases
n_N	Number of normal cases
A	Area under curve
Q_1	Factor1 in SE calculation
Q_2	Factor2 in SE calculation

LIST OF ABBREVIATIONS

TFT	Thin Film Transistor
ROC	Receiver Operating Characteristic
AUC	Area Under Curve
CI	Confidence Interval
TPF / TPR	True Positive Fraction / Ratio
FPF / FPR	False Positive Fraction / Ratio
PACS	Picture Archiving and Communication System

1. INTRODUCTION

1.1 Background and Motivation

Digital radiology has many advantages in comparison with hard-copy based radiology. In digital radiology, there are four basic steps: data acquisition, image processing, data storage and display. One of the most important stages in the whole process is the transfer of the image to the observer. So it is vital to have high quality displays in order not to deteriorate any information obtained in data acquisition and image processing phases. With the rapid development of PACS in radiology, digital image interpretation is introduced at an ever-growing number of hospitals. The technical, financial, and practical advantages of digital radiology can be fully exploited only when image evaluation relies mainly on soft-copy display.

In recent years, most of the medical grade monitors have been converted to TFT LCD types, switching from CRT based medical monitors. Therefore, the common recent practice in observing the visual results of the data acquisition is to use a medical grade TFT LCD monitor. The main problem with medical grade TFT LCD monitors is their high cost. It is possible to have drastic cost advantage in case the consumer grade monitors with enhanced features could be used for medical diagnosis without sacrificing any diagnostic value.

In diagnostic radiology, medical monitors are generally recommended because of their higher luminance and better contrast ratio. Since TFT LCD Panel technology developed very rapidly in last 5 years from performance and technical specs improvement point of view, the quality level of images acquired by using medical grade TFT LCD monitors could be recently comparable to the quality level of images obtained by using consumer grade TFT LCD monitors. There are many previous studies on the comparison of medical LCD monitors with medical CRT monitors from the medical diagnosis

suitability point of view [4, 10]. There are also some recent articles which focus on the comparison of medical grade LCD monitors with consumer grade LCD monitors [4, 5]. The common finding in these studies is that the consumer grade monitors with some special technical specifications could be also used as a PACS monitor for medical diagnosis purpose.

1.2 Objectives

The main objective of this study was to design a 32" Color TFT LCD PACS Monitor to be compatible to existing TFT LCD medical monitors. Then the validation of the design was targeted to be realized through receiver operating characteristic (ROC) curve analysis by a clinical evaluation method.

1.3 Outline of the Thesis

Chapter 1 introduces the subject and draws the framework of the study. The design of the monitor and related technical data is specified in Chapter 2. The receiver operating characteristic curves and the applications are summarized in Chapter 3. The statistical methods for analyzing ROC curves are presented in Chapter 4. The design of the experiment and clinical test results are given in Chapter 5. The future work and conclusion can be found in Chapter 6.

2. DESIGN OF THE MONITOR

2.1 TFT LCD Panel Technology

TFT-LCD stands for thin-film transistor liquid-crystal display. The TFT LCD Panel Module uses liquid crystal to control the passage of light. The basic structure of a

TFT LCD panel may be thought of as two glass substrates sandwiching a layer of liquid crystal. The front glass substrate is fitted with a color filter, while the back glass substrate has transistors fabricated on it. When voltage is applied to a transistor, the liquid crystal is bent, allowing light to pass through to form a pixel. A light source is located at the back of the panel and belongs to a backlight unit. The front glass substrate is fitted with a color filter, which gives each pixel its own color. The combination of these pixels in different colors forms the image on the panel.

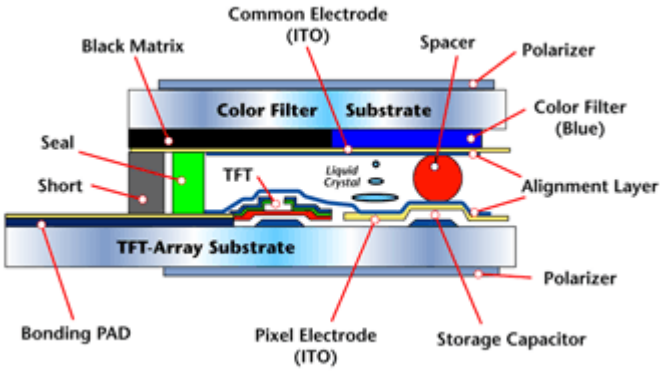

Figure 2.1 The basic structure of TFT LCD panel module [12]

A TFT array is composed of a matrix of pixels and ITO region (a transparent electric conducting film) each with a TFT device and is so called array. Thousands or millions of these pixels together create an image on the display. The diagram below shows the simple structure of a pixel.

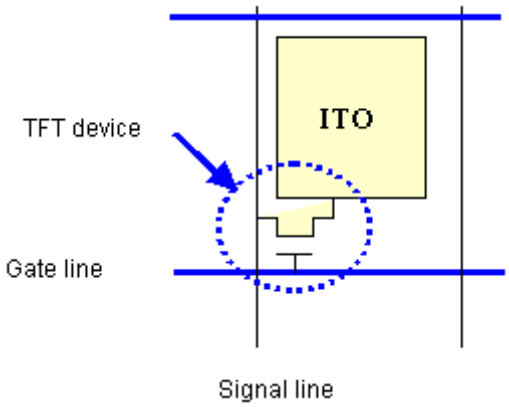

Figure 2.2 The basic structure of a pixel. [12]

TFT device is a switching device, which functions to turn each individual pixel on or off hence control the number of electrons flow into the ITO zone. As the number of electrons reaches the expected value, TFT turns off and these electrons can be kept within the ITO zone.

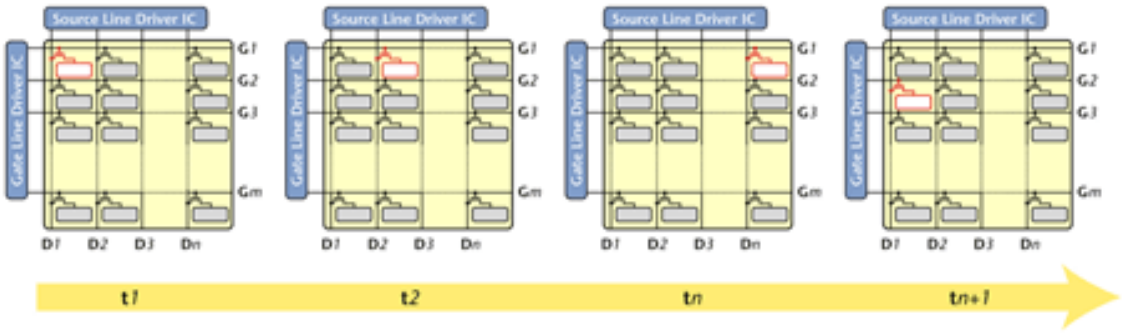

Figure 2.3 The flow of electricity for a TFT substrate [12]

The above diagram shows the flow of electricity for a TFT substrate, during time t1 to tn gate line driver IC continuously chooses to open G1 and allows each source line to charge the TFT pixel which lies along G1 starting from D1, D2 to Dn. Once completed, the gate line driver IC then chooses to open G2 and the process is repeated.

Liquid crystal, lying between TFT substrates and CF substrates, will rotate into different angles according to the charges applied to each pixel. When millions of pixels are charged differently, we could obtain millions of LC angles within the area. The reason why we need to control the LC's standing angle lying within millions of pixels is that we need to use LC's optic rotation nature to control the amount of light passing through the LCD panel. As shown at the following diagram, light from the back light module travels from the TFT panel through ITO electrode, turned by Liquid Crystal and reaches the LCD panel on the top.

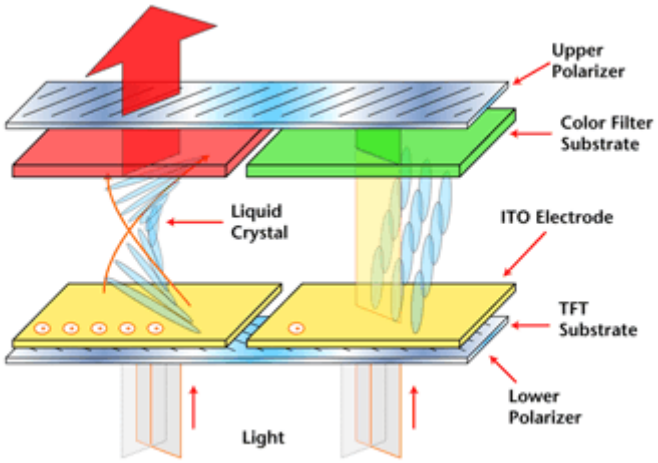

Figure 2.4 The light orientation through LC.

A few things can be discovered from the above diagram:

LC angles control the amount of light being rotated. From the above example: the bigger the standing angle the less the light passes through the top substrate. Light not rotated by LC will be absorbed by the top polarizer. In the Natural world, light travels in random directions; the function of polarizer is to filter most of them and allows only certain light with desired direction to pass through.

2.2 TFT LCD Panel and Monitor Performance Measures

The basic parameters in defining the characteristics of a TFT LCD Panel and Monitor are highly critical for application selection. The prominent parameters are specified below:

a. Luminance is a photometric measure of the luminous intensity per unit area of light travelling in a given direction. It is measured in candelas per square meter.

b. Viewable image size is measured diagonally in inches.

c. Aspect ratio is the ratio of the horizontal length to the vertical length. 4:3 is the standard aspect ratio, for example, so that a screen with a width of 1024 pixels will have a height of 768 pixels. If a widescreen display has an aspect ratio of 16:9, a display that is 1024 pixels wide will have a height of 576 pixels.

d. Display resolution is the number of distinct pixels in each dimension that can be displayed. (It does not mean currently displayed.) Maximum resolution is limited by dot pitch.

e. Dot pitch is the distance between pixels of the same color in millimeters. In general, the smaller the dot pitch, the sharper the picture will appear.

f. Refresh rate is the number of times in a second that a display is illuminated. Maximum refresh rate is limited by response time.

g. Response time is the time a pixel in a monitor takes to go from active (black) to inactive (white) and back to active (black) again, measured in milliseconds. Lower numbers mean faster transitions and therefore fewer visible image artifacts.

h. Contrast ratio is the ratio of the luminosity of the brightest color (white) to that of the darkest color (black) that the monitor is capable of producing. A high contrast ratio is a desired aspect of any display, but with the various methods of

measurement for a system or its part, remarkably different measured values can sometimes produce similar results

i. Power consumption is measured in watts.

j. Viewing angle is the maximum angle at which images on the monitor can be viewed, without excessive degradation to the images. It is measured in degrees horizontally and vertically.

2.3 The Basic Design Concept for the Monitor

A color TFT LCD monitor display simply consists of the plastic cabinet, TFT LCD panel, video board called as the chassis and internal LVDS cables as the interface between the panel and the signal board. Without tooling investment, the front and back plastic cabinets were obtained from some solution providers in Asia. Full HD TFT LCD Panels with resolution of 1920 x 1080 are recently available from major TFT LCD panel manufacturers. The most common monitor resolutions are shown in Table 2.3.1 The designed monitor will be a derivative of WUXGA resolution.

Table 2.3.1

The Common monitor resolutions

Name	Pixel array		Aspect ratio	Comment
VGA	640	480	4:3	
SVGA	800	600	4:3	
XGA	1024	768	4:3	
WXGA	1365	768	16:9	Wide-XGA; used for widescreen LCD TV displays (sometimes 1280×768 is called WXGA)
SXGA	1280	1024	5:4	This format is "squarer" than the others
WSXGA+	1680	1050	16:10	Wide-SXGA (plus a bit more)
UXGA	1600	1200	4:3	
WUXGA	1920	1200	16:10	Wide-UXGA
QXGA	2048	1536	4:3	Quad-XGA

The monitors' resolution is dependent on the resolution of LCD panel modules. It is a solid fact that the reason of these panels` existence is generally for TV applications. The worldwide TFT LCD TV sales volume is about 120 million pieces in a year by about 20% average annual growth in quantity. With this huge production volume, it is obvious that the cost of a commercial Full HD TFT LCD TV panel is highly competitive so as to maintain a feasible cost input for a TFT LCD Monitor display structure which is intended to be used for medical diagnosis purposes.

Based on the screen size to enable better visualization for the user and exploiting the cost advantage due to huge production volume of 32" TFT LCD Panels as covering one third of total worldwide TFT LCD panel volume, the design was targeted for a 32" TFT LCD Color Monitor. 32" Full HD TFT LCD Panel was supplied from AU Optronics in Taiwan. The plastic cabinet, TFT LCD Panel and the signal board were all integrated in Turkey so that the delivery would be quick in case of the commercialization of the design. The resolution of the designed monitor is about 2,074 Mega pixels. Furthermore, the proper chassis with some modification based on the requirements was also sourced from the design houses which have close cooperation with some concept IC manufacturers. The plastic cabinet, TFT LCD Panel and the signal board were all integrated in Turkey so that the delivery would be quick in case of the commercialization of the design.

2.3.1 The Chassis

The chosen chassis uses Novatek series chipset. The general specs of Novatek NT68667/NT68627 chipset is specified below:

Supports RGB (WUXGA 1920x1200) or YPbPr (1080p) inputs

Triple 8bit ADCs (0.55~0.9V) with 500MHz bandwidth

188MHz HPLL with 64 steps phase adjust for each RGB channel

Auto offset for component video

Supports both non-interlaced and interlaced input signals

ADC bandwidth adjust : 500M, 450M, 400M, 350M, 300M, 250M, 150M, 75M

SOG clamp current adjustable

Digital Graphic and Video Inputs

DVI receiver up to 165MHz with HDCP with HDCP 280 bytes sram

Supports ITU-R BT.656 8-bit Input format

Video Processing

Zoom and shrink with non-linear scaling in horizontal direction for wide

The 3rd generation bright frame with adaptive contrast control, 24 color tones adjustment.

RGB real color engine and edge enhancement functions

Adjustable sharpness setting

Support DBC to save system operation power

Text Enhancement

Enhance ghost cancellation

Sync Processor

Support TTL Sync-On-Green (SOG) (including Sync Slicer)

Polarity detection

Frequency

Sync auto switching (including Sync Separator)

Sep /VSI level detection

VSI supports trigger level threshold adjustable

Programmable multi-color RAM font as well as a bitmapped graphical OSD are supported

Provide 1,2,3/4 bits/pixel RAM Fonts

Internal SRAM allows up to 2048 characters, with programmable OSD frame size. Width is 64 columns, and height is 32 rows.

Programmable shadow or border control for each character by each row

Programmable blinking effects for each character

Spacing control to avoid expansion distortion

Supports simultaneous display of up to 4 OSD windows
Maximum 4 times of global zoom for horizontal and vertical axis
Separate row zoom control
Support flexible FG or BG optional transparent, translucent, and opaque effects
256 palette with 64K color selectable.

The chassis supports up to 1920x1200 WUXGA resolution. Horizontal frequency range is from 15 to 250 KHz and vertical frequency range is from 45 to 80 Hz. The interface is Dual LVDS type. It has one DVI and one VGA input. The top view of the chassis layout is shown in Figure 2.3.1 The interface function description is given at Table 2.3.1.1 The chassis size and connector interface for inputs are shown below at Figure 2.3.1.1

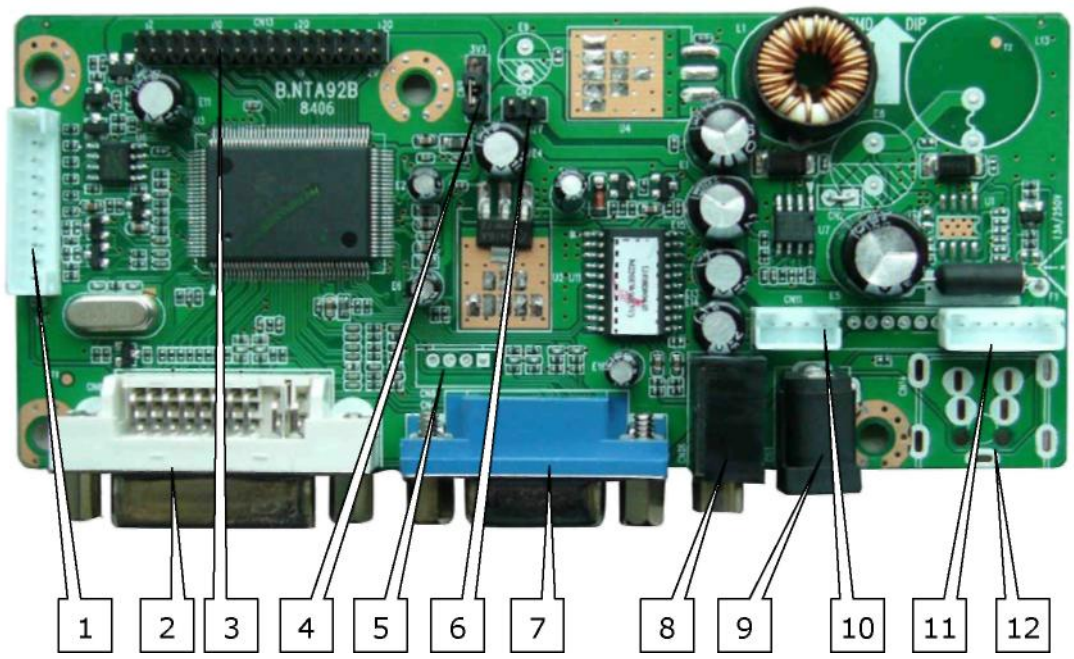

Figure 2.3.1.1 Top view of the chassis layout

Table 2.3.1.1

Interface Function Description

NO.	DESCRIPTION	NO.	DESCRIPTION
1(CN5)	Key board and LED Indicator Connector	7(CN4)	VGA Input
2(CN6)	DVI Input	8(CN18)	Earphone Input
3(CN13)	LVDS Panel Connector	9(CN1)	Power Input 1
4(CN9)	Power Jumper for Panel	10(CN11)	Speaker Connector
5(CN8)	Debugging Interface	11(CN3)	Inverter connector
6(CN7)	Power Jumper for Panel	12(CN14)	Power Input 2

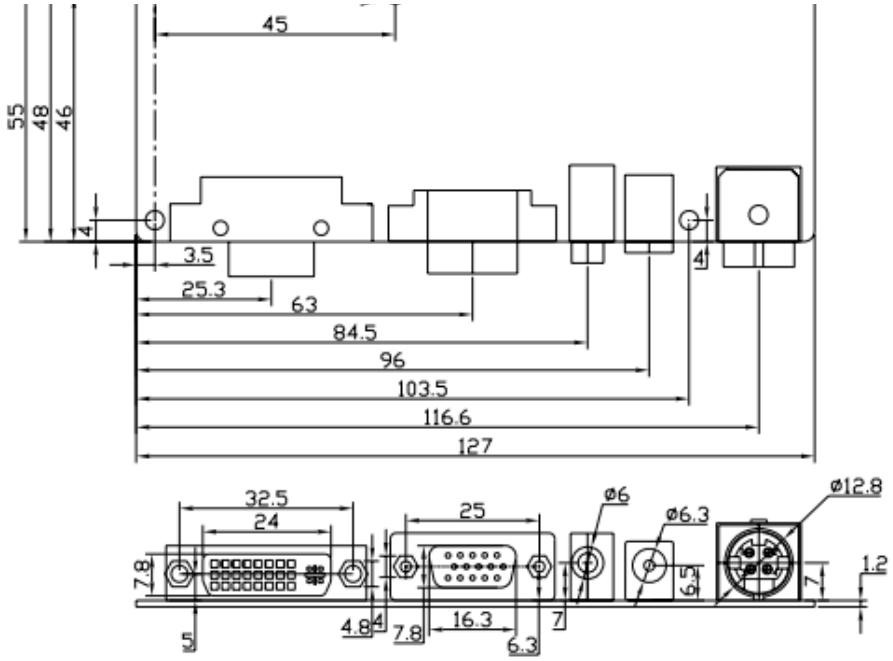

Figure 2.3.1.2 The chassis size and the connector interface

2.3.2 TFT LCD Module

31,5” Color TFT LCD Module from AU Optronics of Taiwan with part number of T315HW02 V0 was used in the design. The general characteristics of the panel module is presented in Table 2.3.2.1

Table 2.3.2.1
General Characteristics of T315HW02 V0 panel module

Items	Specification	Unit	Note
Active Screen Size	31.51 inches		
Display Area	698.4 (H) x 392.85 (V)	mm	
Outline Dimension	760.0(H) x 450.0(V) x 45(D)	mm	With inverter
Driver Element	a-Si TFT active matrix		
Display Colors	16.7M	Color	
Number of Pixels	1920x1080	Pixel	
Pixel Pitch	0.36375	mm	
Pixel Arrangement	RGB vertical stripe		
Display Mode	Normally Black		
Surface Treatment	AG, 3H		Haze = 11%

This is a TFT active matrix liquid crystal panel with 1920 x 1080 pixels. The module fully supports HDTV mode (Non-Interlace). Each pixel is divided in to red, green and blue sub-pixels which are arranged in vertical stripes. Gray scale or the brightness of each sub pixel is determined with a 8 bit gray scale signal for each dot. The contrast and brightness data of the panel is defined in Table 2.3.2.2

Table 2.3.2.2
The contrast and brightness data for T315HW02 V0 panel

	Minimum	Typical	Maximum
Contrast ratio	3200	4000	
Brightness	400 Cd/m²	500 Cd/m²	
Luminance variation			1,30 Cd/m²

2.3.3 The Monitor

The design of 32" TFT LCD PACS monitor with 2.07M pixel resolution has been finalized successfully. There is a photo of the design in Figure 2.3.3.1. The basic mechanical drawing is given at Figure 2.3.3.2

Fig. 2.3.3.1 Front photo of the design

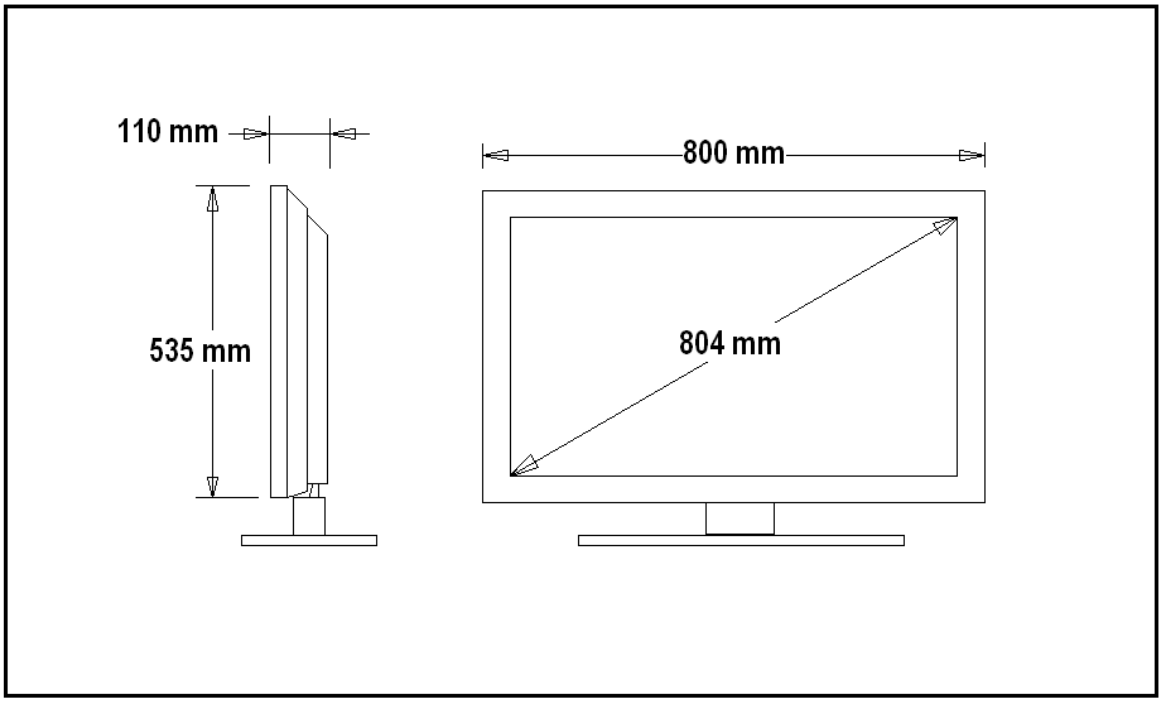

Fig. 2.3.3.2 The basic mechanical drawing of new design

3. RECEIVER OPERATING CHARACTERISTIC CURVES AND APPLICATIONS

Receiver Operating Characteristic (ROC) curves were developed in the 1950's as a by-product of research for understanding radio signals contaminated by noise. More recently it's become clear that they are remarkably useful in medical decision-making. It is an effective method of evaluating the quality or performance of diagnostic tests, and is widely used in radiology to evaluate the performance of many radiological tests.

The ROC curve can also be represented equivalently by plotting the fraction of true positives (TP = true positive rate) vs. the fraction of false positives (FP = false positive rate). With other definition, the receiver operating characteristic (ROC) curve is defined as a plot of test sensitivity as the y coordinate versus its 1-specificity as the x coordinate [3].

The sensitivity is how good the test is at picking out the cases with the related indication. It is simply the true positive fraction. In other words, sensitivity gives the proportion of cases picked out by the test, relative to all cases which actually have the indication. Specificity is the ability of the test to pick out cases which do not have the indication at all. This is synonymous with the true negative fraction.

So sensitivity and specificity, which are defined as the number of true positive decisions/the number of actually positive cases and the number of true negative decisions/the number of actually negative cases, respectively, constitute the basic measures of performance of diagnostic tests as shown in Table 3.1

Table 3.1
The Decision Matrix, Sensitivity and Specificity

	True Condition Status		
Test Result	**Positive**	**Negative**	**Total**
Positive	TP	FP	T+
Negative	FN	TN	T-
Total	D+	D-	

Note: TP: true positive = test positive in actually positive cases, FP: false positive = test positive in actually negative cases, FN: false negative = test negative in actually positive cases, TN: true negative = test negative in actually negative cases.

Sensitivity and Specificity of a test is defined as TP/D+ and TN/D, respectively. When the results of a test fall into one of two obviously defined categories, such as either the presence or absence of a disease, then the test has only one pair of sensitivity and specificity values. However, in many diagnostic situations, making a decision in a binary mode is both difficult and impractical. Image findings may not be obvious. There may be also considerable amount of variation in the diagnostic confidence levels between the radiologists who interpret the findings.

As a result, a single pair of sensitivity and specificity values is insufficient to describe the full range of diagnostic performance of a test [5]. For example, if there will be N number of patients with definite problem of pulmonary nodules who got chest radiography to determine whether the problem is benign or malignant. According to the biopsy results, some number of patients will perhaps appear to have actual malignancies and some patients with benign lesions. Chest radiographs could be interpreted according to a five-point scale: 1 (definitely benign), 2 (probably benign), 3 (possibly malignant), 4 (probably malignant), and 5 (definitely malignant). In this example, one can choose from four different cutoff levels to define a positive test for malignancy on the chest radiographs: X. >=2 (the most liberal criteria), >=3, >=4, and 5 (the most stringent criteria). Therefore, there will be four pairs of sensitivity and specificity values, one pair for each cutoff level, and the sensitivities and specificities depend on the cutoff levels which are used to define the positive and negative test results [5] As the cutoff level decreases, the sensitivity increases while the specificity decreases, and vice versa. To deal with these multiple pairs of sensitivity and specificity values, one can draw a graph using the sensitivities as the y coordinates and the 1-specificities or FPRs as the x coordinates. Each discrete point on the graph, called an operating point, is generated by using different cutoff levels for a positive test result.

To deal with these multiple pairs of sensitivity and specificity values, one can draw a graph using the sensitivities as the y coordinates and the 1-specificities or FPRs as the x coordinates as shown at Fig 3.1 Each discrete point on the graph, called an operating point, is generated by using different cutoff levels for a positive test result. An ROC curve can be estimated from these discrete points, by making the assumption that the test results follow a certain distribution. For this purpose, the assumption of a binormal distribution (i.e., two Gaussian distributions: one for the test results of those patients with benign solitary pulmonary nodules and the other for the test results of those patients with malignant solitary pulmonary nodules) is most commonly made [1, 6]. The resulting curve is called the fitted or smooth ROC curve as shown at Fig 3.2 The estimation of the smooth ROC curve based on a binormal distribution uses a statistical method called maximum

likelihood estimation (MLE) [1,9]. When a binormal distribution is used, the shape of the smooth ROC curve is entirely determined by two parameters. The first one, which is referred to as *a*, is the standardized difference in the means of the distributions of the test results for those subjects with and without the condition. The other parameter, which is referred to as *b*, is the ratio of the standard deviations of the distributions of the test results for those subjects without versus those with the condition [1, 6]. Another way to construct an ROC curve is to connect all the points obtained at all the possible cutoff levels and the two endpoints on the ROC curve are 0, 0 and 1, 1 with each pair of values corresponding to the FPR and sensitivity, respectively as shown at Fig 3.3 The resulting ROC curve is called the empirical ROC curve. The ROC curve illustrates the relationship between sensitivity and FPR. Because the ROC curve displays the sensitivities and FPRs at all possible cutoff levels, it can be used to assess the performance of a test without depending on the decision threshold.

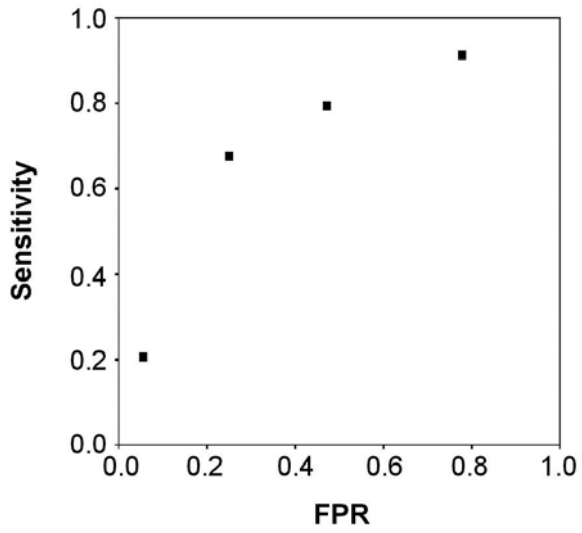

Figure 3.1 A plot of test sensitivity (y coordinate) versus its false positive rate (x coordinate) at each cutoff level.

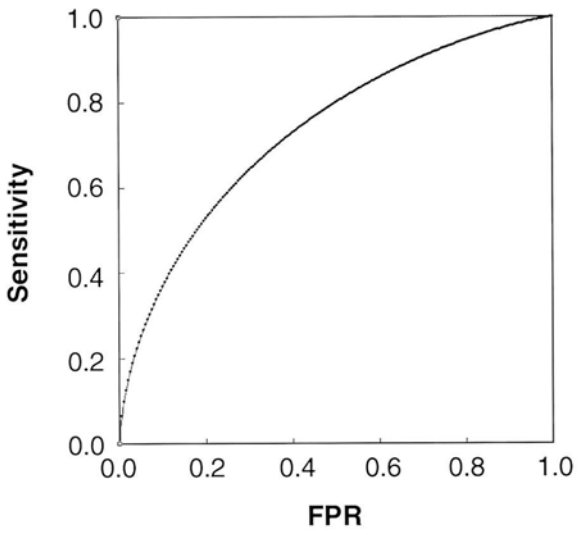

Figure 3.2 The fitted or smooth ROC curve that is obtained by the estimation with the assumption of a binormal distribution

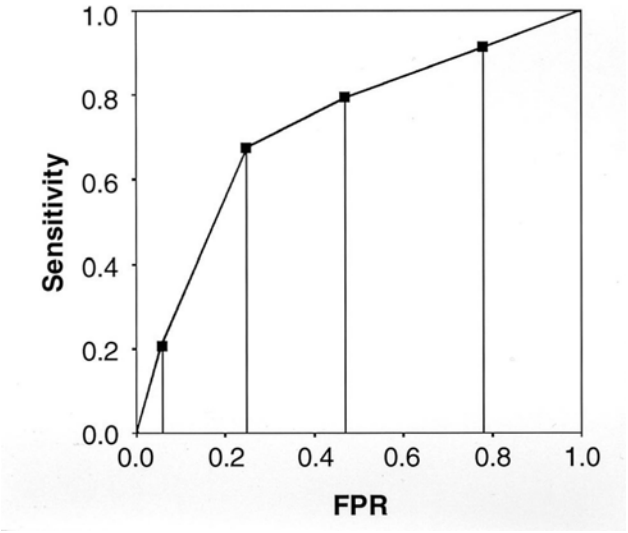

Figure 3.3 The empirical ROC curve. The discrete points on the empirical ROC curve are marked with dots.

Several summary indices are associated with the ROC curve. One of the most popular measures is the area under the ROC curve denoted as AUC. AUC is a combined measure of sensitivity and specificity. AUC is generally a measure of the overall performance of a diagnostic test and is interpreted as the average value of sensitivity for all possible values of specificity. It can take on any value between 0 and 1, since both the x and y axes have values ranging from 0 to 1. The closer AUC is to 1, the better the overall diagnostic performance of the test, and a test with an AUC value of 1 as shown by A curve on Fig 3.4 is considered to be perfectly accurate. The practical lower limit for the AUC of a diagnostic test is 0.5. The line segment from 0, 0 to 1, 1 has an area of 0,5 as shown by curve D on Fig. 3.4 If we were to rely on pure chance to distinguish those subjects with a particular disease from the ones without that disease, the resulting ROC curve would fall along this diagonal line, which is referred to as the chance diagonal. A diagnostic test with an AUC value greater than 0,5 is better than relying on pure chance and it has at least some ability to discriminate between subjects with and without a particular disease.

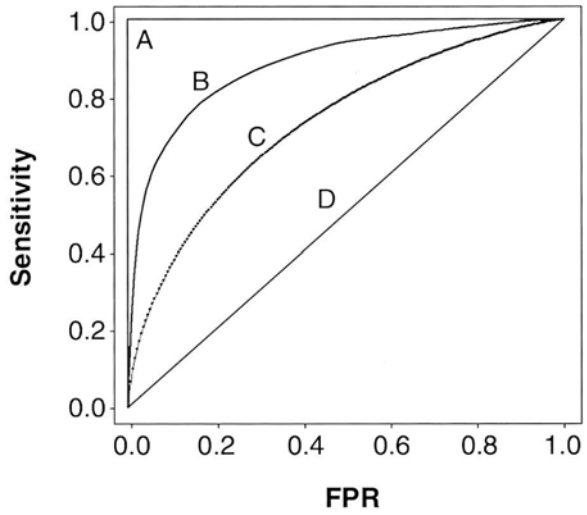

Figure 3.4 Four ROC curves with different values of the area under the ROC curve. A perfect test (A) has an area under the ROC curve of 1. The chance diagonal (D), the line segment from 0, 0 to 1, 1) has an area under the ROC curve of 0.5. ROC curves of tests with some ability to distinguish between those subjects with and without a disease (B, C) lie between these two extremes. Test B with the higher area under the ROC curve has a better overall diagnostic performance than test C

Because sensitivity and specificity are independent of disease prevalence, AUC is also independent of disease prevalence [4, 5]. AUC can be estimated both parametrically, with the assumption that either the test results themselves or some unknown monotonic transformation of the test results follows a binormal distribution, and non-parametrically from the empirical ROC curve without any distributional assumption of the test results [6,9]. Several nonparametric methods of estimating the area under the empirical ROC curve and its variance have been described [8, 9]. The nonparametric estimate of the area under the empirical ROC curve is the summation of the areas of the trapezoids formed by connecting the points on the ROC curve as shown at Fig. 3.3 [7, 9]. The nonparametric estimate of the area under the empirical ROC curve tends to underestimate AUC when discrete rating data (e.g., the five-point scale in the previous example) are collected, whereas the parametric estimate of AUC has negligible bias except when extremely small case samples are employed [1]. For discrete rating data , the parametric method is,

therefore, preferred [1]. However, when discrete rating data are collected, if the test results are not well distributed across the possible response categories (e.g., in the previous example, those patients with actually benign lesions and those patients with actually malignant lesions tend to be rated at each end of the scale, 1 = definitely benign and 5 = definitely malignant, respectively), the data may be degenerate and, consequently, the parametric method may not work well [1]. Using the nonparametric method is an option in this case, but may provide even more biased results than it normally would [1]. For continuous or quasi-continuous data (e.g., a percent-confidence scale from 0% to 100%), the parametric and nonparametric estimates of AUC will have very similar values and the bias is negligible [1]. Therefore, using either the parametric or nonparametric method is fine in these cases [1]. In most ROC analyses of radiological tests, discrete rating scales with five or six categories (e.g., definitely absent, probably absent, possibly present, probably present and definitely present) are used, for which the parametric method is recommended unless there is a problem with degenerate data.

AUC is often presented along with its 95% confidence interval. An AUC of a test obtained from a group of patients is not a fixed, true value, but a value from a sample that is subject to statistical error. Therefore, if one performs the same test on a different group of patients with the same characteristics, the AUC which is obtained may be different [4, 11]. Although it is not possible to specifically define a fixed value for the true AUC of a test, one can choose a range of values in which the true value of AUC lies with a certain degree of confidence. The 95% CI gives the range of values in which the true value lies and the associated degree of confidence. That is to say, one can be 95% sure that the 95% CI includes the true value of AUC [2]. In other words, if one believes that the true value of AUC is within the 95% CI, there is a 5% chance of its being wrong. Therefore, if the lower bound of the 95% CI of AUC for a test is greater than 0.5, then the test is statistically significantly better (with a 5% chance of being wrong or a significance level of 0.05) than making the diagnostic decision based on pure chance, which has an AUC of 0.5 [2, 3].

Since AUC is a measure of the overall performance of a diagnostic test, the overall diagnostic performance of different tests can be compared by comparing their AUCs. The bigger its AUC is, the better the overall performance of the diagnostic test is. When comparing the AUCs of two tests, equal AUC values mean that the two tests yield the same overall diagnostic performance, but this does not necessarily mean that the two ROC curves of the two tests are identical.. Figure 3.5 illustrates two ROC curves with equal AUCs. The curves are obviously not identical. Although the AUCs and, therefore, the overall performances of the two tests are the same, test B is better than test A in the high FPR range (or high sensitivity range), whereas test A is better than test B in the low FPR range (or low sensitivity range) a shown at Fig. 3.5 The equality of two ROC curves can be tested by using the two parameters, *a* and *b*, instead. Because the shape of a binormal smooth ROC curve can be completely specified by the two parameters, *a* and *b*, the equality of the two ROC curves under the binormal assumption can be assessed by testing the equality of the two sets of parameters, *a* and *b*, i.e. by comparing the two sets of values from the two ROC curves. The null hypothesis and alternative hypothesis of the test are H_0: $a_1 = a_2$ and $b_1 = b_2$ versus H_1: $a_1 a_2$ or $b_1 b_2$, respectively, where 1 and 2 denote the two different ROC curves [1, 2]. According to this method, the ROC curves and, consequently, the diagnostic performances of different tests are considered to be different, unless the ROC curves are identical [2, 5].

In some clinical settings, when comparing the performances of different diagnostic tests, one may be interested in only a small portion of the ROC curve and comparing the AUCs and the overall diagnostic performance may be misleading. When screening for a serious disease in a high-risk group (e.g., breast cancer screening), the cutoff range for a positive test should be chosen in such a way as to provide good sensitivity, even if the FPR is high, because false negative test results may have serious consequences [2, 5]

On the other hand, in screening for a certain disease, whose prevalence is very low and for which the subsequent confirmatory tests and/or treatments are very risky, a high specificity and low FPR is required. If the cutoff range for a positive test is not adjusted accordingly, almost all of the positive decisions will be false positive decisions, resulting

in many unnecessary, risky follow-up examinations and/or treatments.

In Figure 3.5, although the AUCs and overall performances of the two tests are the same, in the former diagnostic situation requiring high sensitivity, test B would be better than test A, whereas in the latter situation requiring a low FPR, test A would be better than test B. AUC, as a measure of the overall diagnostic performance, is not helpful in these specific diagnostic situations.

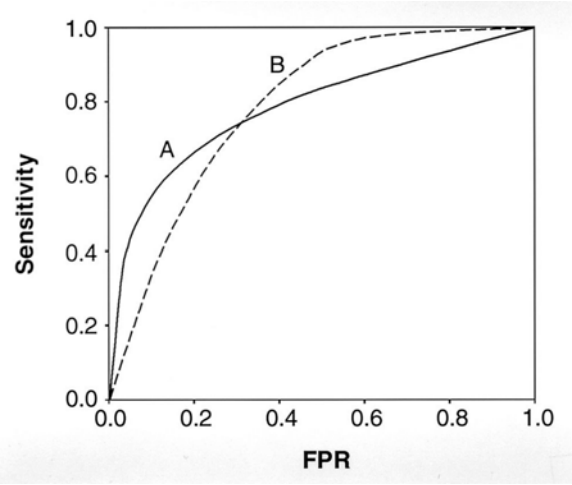

Figure 3.5 Two ROC curves (A and B) with equal area under the ROC curve. However, these two ROC curves are not identical. In the high false positive rate range (or high sensitivity range) test B is better than test A, whereas in the low false positive rate range (or low sensitivity range) test A is better than test B

The diagnostic performance of a test should be judged in the context of the diagnostic situation to which the test is applied. Depending on the specific diagnostic situation, only a portion of the overall ROC curve may need to be considered. One way to consider only a portion of an ROC curve is to use the ROC curve to estimate the sensitivity at a particular FPR, and to compare the sensitivities of different ROC curves at a particular FPR

4. THE STATISTICAL METHODS FOR ANALYZING ROC CURVES

Over the past 40 years investigators have proposed a number of indices to describe receiver operating characteristic. The general index used in the evaluation of ROC analysis is area under curve, AUC. Areas under ROC curves can be obtained in three ways: (i) by the trapezoidal rule; *(ii)* as output from the Dorfman and Alf maximum likelihood estimation program [5, 7]; or *(iii)* from the slope and intercept of the original data when plotted on binormal graph paper [5]. The trapezoidal approach systematically underestimates areas. Because the Dorfman and Alf approach is widely used and accepted method, the areas in this study will be calculated by using this approach.

In most of the ROC analyses of radiological tests which have been conducted to date, a discrete rating scale with five or six categories has been used. When discrete *rating* data are used (e.g., 5-point or 6-point rating scales for radiology imaging experiments), and when ROC curves are assumed to be based on two underlying Gaussian distributions, a maximum likelihood estimation program by Dorfman and Alf [7, 11] can be properly used to fit data points to a smooth curve and to derive the area under this fitted curve (AUC) and its associated standard error. This area is designated as AUC and ranges in value from 0.0 to 1.0.

When continuous data are available (as from chemistry laboratory tests, white cell counts, individualized predictions from logistic regression) and ROC curves are created, no assumptions on underlying distributions need be made to obtain area measurements [7,11] The Wilcoxon statistic follows immediate and direct calculation of the area [7]. Hanley and McNeil's derivation of a closed form approximate expression for the standard error associated with the Wilcoxon statistic can be used to approximate the standard error of the area [7,11]

The statistical method for comparison of the areas under two ROC curves derived from the same set of patients by taking into account the correlation between the areas is clearly specified by Hanley and McNeil [11]. The statistical comparison of the new design and existing medical monitor precisely matches with the content of this former study [5, 11].

The relationship of the area under the ROC curve to the Wilcoxon statistic could be used to derive its statistical properties, such as its standard error, SE and the sample sizes required to measure the area with a pre-specified degree of precision and to provide a desired level of statistical power in comparative experiments. So a statistical analysis to another large class of situations can be extended where the two or more ROC curves are generated using the same set of patients. The standard error, SE for a ROC study sample can be calculated as below specified again by Hanley and McNeil [7, 11]

$$SE = \frac{\sqrt{A(1-A) + (n_A - 1)(Q_1 - A^2) + (n_N - 1)(Q_2 - A^2)}}{n_A n_N} \tag{4.1}$$

Where A is the area under the curve (AUC), n_A and n_N are the number of abnormal and normal cases respectively, and factors Q_1 and Q_2 are calculated as below:

$$Q_1 = \frac{A}{2 - A} \tag{4.2}$$

$$Q_2 = \frac{2A^2}{1 + A} \tag{4.3}$$

In these situations, it is inappropriate to calculate the standard error of the difference between two areas $(\hat{Area}_1$ and $\hat{Area}_2)$ as

$$SE(\hat{Area}_1 - \hat{Area}_2) = \sqrt{SE^2(\hat{Area}_1) + SE^2(\hat{Area}_2)} \tag{4.4}$$

This is due to the fact that *Area₁* and *Area₂* are likely to be correlated. This correlation is likely to be positive; if the vagaries of random sampling of cases produce a higher/lower than expected accuracy index for one modality *(e.g.,* if the sample consisted of a larger than usual number of easy/difficult cases), then the accuracy of the second modality will probably also be correspondingly higher/lower than one would expect [11]. In other words, while the two indices may fluctuate independently by amounts SE_1 and SE_2 in separate samples, they will tend to fluctuate in parallel when derived from a single sample. Because of this correlation, instead of above formula, the below given formula for common standard error of two curves must be used.

$$SE(\hat{A}rea_1 - \hat{A}rea_2) = \sqrt{SE^2(\hat{A}rea_1) + SE^2(\hat{A}rea_2) - 2rSE(\hat{A}rea_1)SE(\hat{A}rea_2)} \quad (4.5)$$

where *r* is a quantity representing the correlation introduced between the two areas by studying the same sample of patients [11].

For comparing two ROC curves, critical ratio must be calculated which is called as z here. If z is above a critical level, then it is accepted that the two areas are statistically different. It is common to set this critical level at 1.96, as this will mean to make a type I error, one in twenty chance by rejecting the hypothesis that the two curves are similar. The value of 1.96 indicates that the areas of the two curves are two standard deviations apart, so there is only a ~5% chance that this occurred randomly and that the curves are in fact the same. This quantity z is then referred to tables of the normal distribution and values of z above some cutoff, *e.g., z* ≥ 1.96, are taken as evidence that the "true" ROC areas are different. The importance of introducing the $2rSE_1SE_2$ term in the above equation is obvious: failure to subtract out from the sampling variability those fluctuations that the paired design has already eliminated will leave the denominator of below equation too large and z too small, thereby reducing the chance of detecting a difference between two modalities [10, 11]. The general approach to assessing whether the difference in the areas under two ROC curves derived from the same set of patients is random or real is to calculate a critical ratio z, defined as

$$z = \frac{A_1 - A_2}{\sqrt{SE_1^2 + SE_2^2 - 2rSE_1SE_2}} \tag{4.6}$$

where A_1, A_2 and SE_1 and SE_2 refer to the observed area and estimated standard error of the ROC area associated with the related curves. [7, 11] The correlation coefficient r can be calculated by using Table 4.1 referenced again from Hanley and McNeil [11].

Table 4.1
The calculation of correlation coefficient for two discrete areas under curve

Average Correlation Between Ratings*	Average Area**											
	.700	.725	.750	.775	.800	.825	.850	.875	.900	.925	.950	.975
0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01
0.04	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.02
0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.03	0.02
0.08	0.07	0.07	0.07	0.07	0.07	0.06	0.06	0.06	0.06	0.05	0.04	0.03
0.10	0.09	0.09	0.09	0.09	0.08	0.08	0.08	0.07	0.07	0.06	0.06	0.04
0.12	0.11	0.11	0.11	0.10	0.10	0.10	0.09	0.09	0.08	0.08	0.07	0.05
0.14	0.13	0.12	0.12	0.12	0.12	0.11	0.11	0.11	0.10	0.09	0.08	0.06
0.16	0.14	0.14	0.14	0.14	0.13	0.13	0.13	0.12	0.11	0.11	0.09	0.07
0.18	0.16	0.16	0.16	0.16	0.15	0.15	0.14	0.14	0.13	0.12	0.11	0.09
0.20	0.18	0.18	0.18	0.17	0.17	0.17	0.16	0.15	0.15	0.14	0.12	0.10
0.22	0.20	0.20	0.19	0.19	0.19	0.18	0.18	0.17	0.16	0.15	0.14	0.11
0.24	0.22	0.22	0.21	0.21	0.21	0.20	0.19	0.19	0.18	0.17	0.15	0.12
0.26	0.24	0.23	0.23	0.23	0.22	0.22	0.21	0.20	0.19	0.18	0.16	0.13
0.28	0.26	0.25	0.25	0.25	0.24	0.24	0.23	0.22	0.21	0.20	0.18	0.15
0.30	0.27	0.27	0.27	0.26	0.26	0.25	0.25	0.24	0.23	0.21	0.19	0.16
0.32	0.29	0.29	0.29	0.28	0.28	0.27	0.26	0.26	0.24	0.23	0.21	0.18
0.34	0.31	0.31	0.31	0.30	0.30	0.29	0.28	0.27	0.26	0.25	0.23	0.19
0.36	0.33	0.33	0.32	0.32	0.31	0.31	0.30	0.29	0.28	0.26	0.24	0.21
0.38	0.35	0.35	0.34	0.34	0.33	0.33	0.32	0.31	0.30	0.28	0.26	0.22
0.40	0.37	0.37	0.36	0.36	0.35	0.35	0.34	0.33	0.32	0.30	0.28	0.24
0.42	0.39	0.39	0.38	0.38	0.37	0.36	0.36	0.35	0.33	0.32	0.29	0.25
0.44	0.41	0.40	0.40	0.40	0.39	0.38	0.38	0.37	0.35	0.34	0.31	0.27
0.46	0.43	0.42	0.42	0.42	0.41	0.40	0.39	0.38	0.37	0.35	0.33	0.29
0.48	0.45	0.44	0.44	0.43	0.43	0.42	0.41	0.40	0.39	0.37	0.35	0.30
0.50	0.47	0.46	0.46	0.45	0.45	0.44	0.43	0.42	0.41	0.39	0.37	0.32
0.52	0.49	0.48	0.48	0.47	0.47	0.46	0.45	0.44	0.43	0.41	0.39	0.34
0.54	0.51	0.50	0.50	0.49	0.49	0.48	0.47	0.46	0.45	0.43	0.41	0.36
0.56	0.53	0.52	0.52	0.51	0.51	0.50	0.49	0.48	0.47	0.45	0.43	0.38
0.58	0.55	0.54	0.54	0.53	0.53	0.52	0.51	0.50	0.49	0.47	0.45	0.40
0.60	0.57	0.56	0.56	0.55	0.55	0.54	0.53	0.52	0.51	0.49	0.47	0.42
0.62	0.59	0.58	0.58	0.57	0.57	0.56	0.55	0.54	0.53	0.51	0.49	0.45
0.64	0.61	0.60	0.60	0.59	0.59	0.58	0.58	0.57	0.55	0.54	0.51	0.47
0.66	0.63	0.62	0.62	0.62	0.61	0.60	0.60	0.59	0.57	0.56	0.53	0.49
0.68	0.65	0.64	0.64	0.64	0.63	0.62	0.62	0.61	0.60	0.58	0.56	0.51
0.70	0.67	0.66	0.66	0.66	0.65	0.65	0.64	0.63	0.62	0.60	0.58	0.54
0.72	0.69	0.69	0.68	0.68	0.67	0.67	0.66	0.65	0.64	0.63	0.60	0.56
0.74	0.71	0.71	0.70	0.70	0.69	0.69	0.68	0.67	0.66	0.65	0.63	0.59
0.76	0.73	0.73	0.72	0.72	0.72	0.71	0.71	0.70	0.69	0.67	0.65	0.61
0.78	0.75	0.75	0.75	0.74	0.74	0.73	0.73	0.72	0.71	0.70	0.68	0.64
0.80	0.77	0.77	0.77	0.76	0.76	0.76	0.75	0.74	0.73	0.72	0.70	0.67
0.82	0.79	0.79	0.79	0.79	0.78	0.78	0.77	0.77	0.76	0.75	0.73	0.70
0.84	0.82	0.81	0.81	0.81	0.81	0.80	0.80	0.79	0.78	0.77	0.76	0.73
0.86	0.84	0.84	0.83	0.83	0.83	0.82	0.82	0.81	0.81	0.80	0.78	0.75
0.88	0.86	0.86	0.86	0.85	0.85	0.85	0.84	0.84	0.83	0.82	0.81	0.79
0.90	0.88	0.88	0.88	0.88	0.87	0.87	0.87	0.86	0.86	0.85	0.84	0.82

Correlation coefficient *r* between two ROC areas A_1 and A_2 *is determined* as a function of average correlation between ratings (rows) and average area (columns).

* $(r_N + r_A) / 2$

** $(A_1+A_2) / 2$

5. DESIGN OF EXPERIMENT AND CLINICAL TEST RESULTS WITH STATISTICAL EVALUATION

The data set consisting of 60 chest radiographs with nodules on half of them have been obtained from the authors of a previous study [5] These 60 normal digital adult chest radiographs were said to be randomly selected from a wide database of patients without pulmonary disease. Expert consensus by two experienced chest radiologists was said to be used as the gold standard for normality. After selection, all patient-related information was digitally obscured. Lung nodules with different subject-contrasts and diameters were said to be simulated by digitally superimposing circular Gaussian profiles on 30 normal radiographs with Matlab software (Mathworks Inc, Natick) [5]. Two nodule diameters of 5 mm and 10 mm had been considered by adjusting the profile full width. As a result, a set of 60 radiographs was obtained that consisted of 30 normal radiographs and 30 solitary-nodule radiographs containing simulated nodules placed on various locations within the lung fields [5].

20.8” size, 3 Mega-Pixel medical monitor has been selected as the benchmark for the new design. The basic technical comparison of the reference monitor model and the new design is specified in Table.5.1

Table 5.1
Features of Reference Monitor and New Design

Monitor type	20.8”- 3MP	32”- New Design
Colour tone	11-bit grey scale	32-bit colour
Resolution	1.536×2.048	1.080×1.920
Brightness	400 cd/m^2	500 cd/m^2
Contrast ratio	900:1 typical	4.000:1 typical
Input Signals	DVI Signal Link	DVI Signal Link
Orientation	Portrait or landscape	Portrait or landscape
Use time	> 18 months	<1 month
Dicom viewer	DicomWorks	DicomWorks
Ambient lighting	20–30 lux	20–30 lux

Three radiologists from Acıbadem Hospital, Bakırköy Istanbul , all experienced in digital chest radiography participated in the study. They independently evaluated the set of images on Reference monitor and the new design. The reviewers were told that 60 chest radiographs were to be shown randomly and 50% of them contained one nodule with different sizes and with different coordinates. They were asked to rank their level of confidence in the presence or absence of a nodule by using a continuous rating scale (0-definitely no nodule, 100-nodule definitely present). They were also asked to record the position of each suspected nodule to verify that the identified nodule corresponded to the location of the simulated nodule. Prior to the reading sessions, a training session was carried out to get the observers used with the scoring forms and the display systems, including the controls for brightness, contrast and zoom. The interval between sessions was at least 3 weeks to eliminate the learning effects. The radiologists were allowed to change the window level and width on the monitors and viewing time was unrestricted. The observations were carried out in the same location; ambient light in the room was controlled to be at the same level for each observer. As dicom viewer, DicomWorks software tool has been used in the evaluation of both monitors. The detailed content of the data set with nodule coordinates is given in Table 5.2 and Table 5.3

The results of Radiologists' evaluations for 60 radiographs on both Reference and New Design are presented in Tables 5.4 - 5.15 The correlation coefficients for all radiologist' ratings were calculated by using PAST statistics software tool. Kendall tau was used to determine the correlation coefficient values as suggested by Hanley and Mc Neil. The correlation coefficients are presented in Tables 5.16 – 5.18

ROC curve analysis has been done through JROCFIT 1.02.2 which is Johns Hopkins University's specific edition for ROCFIT software package from Chicago University. Through JROCFIT software tool, the receiver operating characteristics curves for all radiologist - monitor pairs are specified in Figures 5.1-5.6 After all calculations and analysis of full data through mentioned software tool, the statistical comparison of the data are presented in Tables 5.19 - 5.22

Table 5.2
The data set with each radiograph's nodule types

	nodule FWHM (mm)
s = subtle	5 and 10
k = small	5
g = large	10
b = no nodule	

Image No	Nodule Type
1	k9
2	g7
3	k13
4	g20
5	s9
6	b21
7	g23
8	b4
9	b13
10	s8
11	b5
12	s4
13	b2
14	k27
15	b18
16	b29
17	b20
18	k1
19	b6
20	k22
21	b15
22	k28
23	b26
24	b27
25	k16
26	k25
27	g17
28	b12
29	s7
30	b24

Image No	Nodule Type
31	b7
32	b1
33	b8
34	b28
35	b10
36	k20
37	b22
38	k3
39	b11
40	b9
41	k4
42	k15
43	k14
44	s5
45	b19
46	b17
47	b14
48	b3
49	b25
50	g27
51	g16
52	k12
53	b16
54	s2
55	g21
56	k2
57	b29
58	b23
59	s3
60	s6

Table 5.3
Nodule Coordinates

		coordinates	
	nodule (mm)	X	Y
s1	5	404	1032
s2	5	1164	1016
s3	5	1419	542
s4	5	276	1368
s5	5	663	540
s6	10	1111	544
s7	5	1239	1099
s8	5	625	536
s9	5	682	1035
s10			
s11	5	352	618

		coordinates	
		X	y
k1	5	521	470
k2	5	1335	928
k3	5	1128	241
k4	5	252	1115
k5	5	1369	532
k6	5	1312	712
k7	5	675	363
k8	5	1320	880
k9	5	1367	530
k10	5	454	675
k11	5	375	1040
k12	5	385	1028
k13	5	1302	451
k14	5	502	939
k15	5	1217	1307
k16	5	1213	1142
k17	5	1270	845
k18	5	667	482
k19	5	1438	776
k20	5	1259	490
k21	5	1107	241
k22	5	467	654
k23	5	1238	1277
k24	5	301	1137

		coordinates	
		X	Y
k25	5	1392	529
k26	5	506	477
k27	5	1341	481
k28	5	1148	974
k29	5	1459	1150
		coordinates	
		X	Y
g1	10	360	908
g2	10	560	429
g3	10	1283	618
g4	10	1165	422
g5	10	1564	1057
g6	10	1246	554
g7	10	1414	105
g8	10	460	580
g9	10	1385	1094
g10	10	1225	597
g11	10	1200	497
g12	10	1245	411
g13	10	1383	699
g14	10	367	1001
g15	10	522	1278
g16	10	1315	1126
g17	10	1495	1372
g18	10	516	1131
g19	10	1182	345
g20	10	488	599
g21	10	474	554
g22	10	344	1386
g23	10	1285	495
g24	10	563	507
g25	10	459	544
g26	10	477	995
g27	10	1349	632
g28	10	571	465
g29	10	1137	587

Table 5.4
Radiologist 1's evaluation on Reference monitor

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 1 (0-100)	Rating (1-5)
1	k9	1	100	5
2	g7	1	100	5
3	k13	1	0	1
4	g20	1	100	5
5	s9	1	0	1
6	b21	0	0	1
7	g23	1	30	3
8	b4	0	20	2
9	b13	0	0	1
10	s8	1	0	1
11	b5	0	50	3
12	s4	1	80	4
13	b2	0	90	3
14	k27	1	0	1
15	b18	0	80	4
16	b29	0	90	4
17	b20	0	30	3
18	k1	1	100	5
19	b6	0	0	1
20	k22	1	100	5
21	b15	0	0	1
22	k28	1	0	1
23	b26	0	0	1
24	b27	0	0	1
25	k16	1	30	3
26	k25	1	0	1
27	g17	1	50	3
28	b12	0	30	3
29	s7	1	0	1
30	b24	0	0	1
31	b7	0	0	1
32	b1	0	0	1
33	b8	0	40	3
34	b28	0	50	3
35	b10	0	0	1
36	k20	1	100	5
37	b22	0	0	1
38	k3	1	100	5
39	b11	0	0	1

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 1 (0-100)	Rating (1-5)
40	b9	0	0	1
41	k4	1	100	5
42	k15	1	0	1
43	k14	1	0	1
44	s5	1	30	3
45	b19	0	0	1
46	b17	0	0	1
47	b14	0	0	1
48	b3	0	0	1
49	b25	0	30	3
50	g27	1	0	1
51	g16	1	30	3
52	k12	1	100	5
53	b16	0	10	2
54	s2	1	100	5
55	g21	1	0	1
56	k2	1	100	5
57	b29	0	80	4
58	b23	0	0	1
59	s3	1	50	3
60	s6	1	0	1

Table 5.5
Results from Radiologist 1's evaluation on Reference Monitor

	1	2	3	4	5	
	definetely no nodule present	probably no nodule present	possibly nodule present	probably nodule present	definitely nodule present	
Nodule Present	12	0	6	1	11	30
No Nodule present	18	2	7	3	0	30

Table 5.6
Radiologist 1's evaluation on New Design Monitor

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 1 (0-100)	Rating (1-5)
1	k9	1	100	5
2	g7	1	100	5
3	k13	1	40	3
4	g20	1	90	5
5	s9	1	0	1
6	b21	0	0	1
7	g23	1	30	3
8	b4	0	0	1
9	b13	0	0	1
10	s8	1	0	1
11	b5	0	100	5
12	s4	1	0	1
13	b2	0	30	3
14	k27	1	50	3
15	b18	0	40	3
16	b29	0	60	4
17	b20	0	0	1
18	k1	1	100	5
19	b6	0	0	1
20	k22	1	60	4
21	b15	0	0	1
22	k28	1	20	2
23	b26	0	0	1
24	b27	0	0	1
25	k16	1	20	2
26	k25	1	0	1
27	g17	1	60	4
28	b12	0	0	1
29	s7	1	0	1
30	b24	0	0	1
31	b7	0	30	3
32	b1	0	0	1
33	b8	0	80	4
34	b28	0	0	1
35	b10	0	0	1
36	k20	1	100	5
37	b22	0	0	1
38	k3	1	100	5
39	b11	0	0	1

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 1 (0-100)	Rating (1-5)
40	b9	0	0	1
41	k4	1	50	3
42	k15	1	0	1
43	k14	1	40	3
44	s5	1	50	3
45	b19	0	0	1
46	b17	0	0	1
47	b14	0	0	1
48	b3	0	0	1
49	b25	0	20	2
50	g27	1	30	3
51	g16	1	80	4
52	k12	1	80	4
53	b16	0	0	1
54	s2	1	60	4
55	g21	1	0	1
56	k2	1	60	4
57	b29	0	20	2
58	b23	0	0	1
59	s3	1	100	5
60	s6	1	0	1

Table 5.7
Results from Radiologist 1's evaluation on New Design Monitor

	1	2	3	4	5	
	definitely no nodule present	probably no nodule present	possibly nodule present	probably nodule present	definitely nodule present	
Nodule Present	8	2	7	6	7	30
No Nodule present	22	2	3	2	1	30

Table 5.8
Radiologist 2's evaluation on Reference monitor

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 2 (0-100)	Rating (1-5)
1	k9	1	100	5
2	g7	1	90	5
3	k13	1	0	1
4	g20	1	100	5
5	s9	1	0	1
6	b21	0	0	1
7	g23	1	80	4
8	b4	0	0	1
9	b13	0	40	3
10	s8	1	40	3
11	b5	0	0	1
12	s4	1	0	1
13	b2	0	50	3
14	k27	1	50	3
15	b18	0	0	1
16	b29	0	0	1
17	b20	0	30	3
18	k1	1	100	5
19	b6	0	0	1
20	k22	1	100	5
21	b15	0	0	1
22	k28	1	80	4
23	b26	0	0	1
24	b27	0	20	2
25	k16	1	0	1
26	k25	1	50	3
27	g17	1	20	2
28	b12	0	10	2
29	s7	1	0	1
30	b24	0	0	1
31	b7	0	0	1
32	b1	0	0	1
33	b8	0	40	3
34	b28	0	40	3
35	b10	0	0	1
36	k20	1	100	5
37	b22	0	30	3
38	k3	1	100	5
39	b11	0	0	1

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 2 (0-100)	Rating (1-5)
40	b9	0	80	4
41	k4	1	60	4
42	k15	1	0	1
43	k14	1	0	1
44	s5	1	30	3
45	b19	0	0	1
46	b17	0	0	1
47	b14	0	0	1
48	b3	0	0	1
49	b25	0	0	1
50	g27	1	0	1
51	g16	1	70	4
52	k12	1	100	5
53	b16	0	50	3
54	s2	1	70	4
55	g21	1	0	1
56	k2	1	60	4
57	b29	0	0	1
58	b23	0	0	1
59	s3	1	90	5
60	s6	1	0	1

Table 5.9
Results from Radiologist 2's evaluation on Reference Monitor

	1	2	3	4	5	
	definitely no nodule present	probably no nodule present	possibly nodule present	probably nodule present	definitely nodule present	
Nodule Present	10	1	4	6	9	30
No Nodule present	20	2	7	1	0	30

Table 5.10
Radiologist 2's evaluation on New Design Monitor

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 2 (0-100)	Rating (1-5)
1	k9	1	100	5
2	g7	1	100	5
3	k13	1	0	1
4	g20	1	100	5
5	s9	1	0	1
6	b21	0	0	1
7	g23	1	0	1
8	b4	0	20	2
9	b13	0	0	1
10	s8	1	70	4
11	b5	0	40	3
12	s4	1	0	1
13	b2	0	60	4
14	k27	1	50	3
15	b18	0	30	3
16	b29	0	20	2
17	b20	0	50	3
18	k1	1	100	5
19	b6	0	0	1
20	k22	1	100	5
21	b15	0	30	3
22	k28	1	0	1
23	b26	0	0	1
24	b27	0	0	1
25	k16	1	20	2
26	k25	1	90	5
27	g17	1	90	5
28	b12	0	0	1
29	s7	1	0	1
30	b24	0	0	1
31	b7	0	0	1
32	b1	0	0	1
33	b8	0	30	3
34	b28	0	0	1
35	b10	0	0	1
36	k20	1	100	5
37	b22	0	90	5
38	k3	1	100	5
39	b11	0	0	1

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 2 (0-100)	Rating (1-5)
40	b9	0	0	1
41	k4	1	90	5
42	k15	1	60	4
43	k14	1	0	1
44	s5	1	20	2
45	b19	0	0	1
46	b17	0	0	1
47	b14	0	0	1
48	b3	0	0	1
49	b25	0	0	1
50	g27	1	20	2
51	g16	1	90	5
52	k12	1	100	5
53	b16	0	0	1
54	s2	1	90	5
55	g21	1	0	1
56	k2	1	90	5
57	b29	0	0	1
58	b23	0	0	1
59	s3	1	90	5
60	s6	1	0	1

Table 5.11
Results from Radiologist 2's evaluation on New Design Monitor

	1	2	3	4	5	
	definitely no nodule present	probably no nodule present	possibly nodule present	probably nodule present	definitely nodule present	
Nodule Present	9	3	1	2	15	30
No Nodule present	21	2	5	1	1	30

Table 5.12
Radiologist 3's evaluation on Reference monitor

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 3 (0-100)	Rating (1-5)
1	k9	1	100	5
2	g7	1	90	5
3	k13	1	0	1
4	g20	1	100	5
5	s9	1	0	1
6	b21	0	30	3
7	g23	1	70	4
8	b4	0	20	2
9	b13	0	30	3
10	s8	1	0	1
11	b5	0	30	3
12	s4	1	30	3
13	b2	0	0	1
14	k27	1	60	4
15	b18	0	70	4
16	b29	0	0	1
17	b20	0	0	1
18	k1	1	100	5
19	b6	0	0	1
20	k22	1	100	5
21	b15	0	20	2
22	k28	1	50	3
23	b26	0	0	1
24	b27	0	0	1
25	k16	1	50	3
26	k25	1	0	1
27	g17	1	50	3
28	b12	0	40	3
29	s7	1	0	1
30	b24	0	0	1
31	b7	0	0	1
32	b1	0	0	1
33	b8	0	30	3
34	b28	0	0	1
35	b10	0	0	1
36	k20	1	100	5
37	b22	0	20	2
38	k3	1	100	5
39	b11	0	0	1

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 3 (0-100)	Rating (1-5)
40	b9	0	70	4
41	k4	1	80	4
42	k15	1	20	2
43	k14	1	0	1
44	s5	1	30	2
45	b19	0	0	1
46	b17	0	0	1
47	b14	0	0	1
48	b3	0	0	1
49	b25	0	0	1
50	g27	1	60	4
51	g16	1	60	4
52	k12	1	90	5
53	b16	0	40	3
54	s2	1	70	4
55	g21	1	0	1
56	k2	1	50	3
57	b29	0	0	1
58	b23	0	0	1
59	s3	1	90	5
60	s6	1	0	1

Table 5.13
Results from Radiologist 3's evaluation on Reference Monitor

	1	2	3	4	5	
	definitely no nodule present	probably no nodule present	possibly nodule present	probably nodule present	definitely nodule present	
Nodule Present	8	2	5	6	9	30
No Nodule present	19	3	6	2	0	30

Table 5.14
Radiologist 3's evaluation on New Design Monitor

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 3 (0-100)	Rating (1-5)
1	k9	1	100	5
2	g7	1	100	5
3	k13	1	20	2
4	g20	1	100	5
5	s9	1	0	1
6	b21	0	40	3
7	g23	1	60	4
8	b4	0	30	3
9	b13	0	0	1
10	s8	1	20	2
11	b5	0	50	3
12	s4	1	50	3
13	b2	0	0	1
14	k27	1	70	4
15	b18	0	30	3
16	b29	0	30	3
17	b20	0	20	2
18	k1	1	100	5
19	b6	0	0	1
20	k22	1	90	5
21	b15	0	0	1
22	k28	1	30	3
23	b26	0	0	1
24	b27	0	0	1
25	k16	1	30	3
26	k25	1	0	1
27	g17	1	80	4
28	b12	0	30	3
29	s7	1	0	1
30	b24	0	0	1
31	b7	0	0	1
32	b1	0	0	1
33	b8	0	20	2
34	b28	0	0	1
35	b10	0	0	1
36	k20	1	100	5
37	b22	0	60	4
38	k3	1	100	5
39	b11	0	0	1

Case Number	Nodule Type	Nodule Existence in Reality	Radiolog 3 (0-100)	Rating (1-5)
40	b9	0	0	1
41	k4	1	50	3
42	k15	1	0	1
43	k14	1	0	1
44	s5	1	20	2
45	b19	0	0	1
46	b17	0	0	1
47	b14	0	0	1
48	b3	0	0	1
49	b25	0	0	1
50	g27	1	50	3
51	g16	1	80	4
52	k12	1	100	5
53	b16	0	60	4
54	s2	1	50	3
55	g21	1	30	3
56	k2	1	70	4
57	b29	0	0	1
58	b23	0	0	1
59	s3	1	100	5
60	s6	1	0	1

Table 5.15
Results from Radiologist 3's evaluation on New Design Monitor

	1	2	3	4	5	
	definitely no nodule present	probably no nodule present	possibly nodule present	probably nodule present	definitely nodule present	
Nodule Present	6	3	7	5	9	30
No Nodule present	20	2	6	2	0	30

Table 5.16
The correlation coefficients for Radiologist 1's observations on both monitors

r_{A1}		r_{N1}	
5	5	1	1
5	5	2	1
1	3	1	1
5	5	3	5
1	1	3	3
3	3	4	3
1	1	4	4
4	1	3	1
1	3	1	1
5	5	1	1
5	4	1	1
1	2	1	1
3	2	3	1
1	1	1	1
3	4	1	3
1	1	1	1
5	5	3	4
5	5	3	1
5	3	1	1
1	1	1	1
1	3	1	1
3	3	1	1
1	3	1	1
3	4	1	1
5	4	1	1
5	4	1	1
1	1	3	2
5	4	2	1
3	5	4	2
1	1	1	1
0,67795		0,59695	

r_{A1} : The correlation coefficient of the ratings between Reference and New Design observations for Radiologist 1 when the nodule is actually present

r_{N1} : The correlation coefficient of the ratings between Reference and New Design observations for Radiologist 1 when there is no nodule

Table 5.17
The correlation coefficients for Radiologist 2's observations on both monitors

r_{A2}		r_{N2}	
5	5	1	1
5	5	1	2
1	1	3	1
5	5	1	3
1	1	3	4
4	1	1	3
3	4	1	2
1	1	3	3
3	3	1	1
5	5	1	3
5	5	1	1
4	1	2	1
1	2	2	1
3	5	1	1
2	5	1	1
1	1	1	1
5	5	3	3
5	5	3	1
4	5	1	1
1	4	3	5
1	1	1	1
3	2	4	1
1	2	1	1
4	5	1	1
5	5	1	1
4	5	1	1
1	1	1	1
4	5	3	1
5	5	1	1
1	1	1	1
0,67899		0,21973	

r_{A2} : The correlation coefficient of the ratings between Reference and New Design observations for Radiologist 2 when the nodule is actually present

r_{N2} : The correlation coefficient of the ratings between Reference and New Design observations for Radiologist 2 when there is no nodule

Table 5.18
The correlation coefficients for Radiologist 3's observations on both monitors

r_{A3}		r_{N3}	
5	5	3	3
5	5	2	3
1	2	3	1
5	5	3	3
1	1	1	1
4	4	4	3
1	2	1	3
3	3	1	2
4	4	1	1
5	5	2	1
5	5	1	1
3	3	1	1
3	3	3	3
1	1	1	1
3	4	1	1
1	1	1	1
5	5	3	2
5	5	1	1
4	3	1	1
2	1	2	4
1	1	1	1
2	2	4	1
4	3	1	1
4	4	1	1
5	5	1	1
4	3	1	1
1	3	1	1
3	4	3	4
5	5	1	1
1	1	1	1
0,8547		0,55427	

r_{A3} : The correlation coefficient of the ratings between Reference and New Design observations for Radiologist 3 when the nodule is actually present

r_{N3} : The correlation coefficient of the ratings between Reference and New Design observations for Radiologist 3 when there is no nodule

Table 5.19
Statistical Comparison of Radiologist 1's evaluations on both monitors

Reference Monitor-Radiolog1		**New Design Monitor-Radiolog 1**	
Number of Cases	60	Number of Cases	60
Number Correct	38	Number Correct	44
Accuracy	63,30%	Accuracy	73,30%
Sensitivity	60,00%	Sensitivity	66,70%
Specificity	66,70%	Specificity	80,00%
Pos Cases Missed	12	Pos Cases Missed	10
Neg Cases Missed	10	Neg Cases Missed	6
Fitted ROC Area	0,634	Fitted ROC Area	0,811
Empiric ROC Area	0,658	Empiric ROC Area	0,763
A_1	0,634	A_2	0,811
SE_1	0,072	SE_2	0,056
Q_1	0,464	Q_1	0,682
Q_2	0,492	Q_2	0,726
SE	0,059		
r_{N1}	0,678	r_{N1}	
r_{A1}	0,597	r_{A1}	
r	0,637	r	
n_A	30	n_A	30
n_N	30	n_N	30
A	0,723		
z	3,002		
r_x	0,600		

For Radiologist 1's observations, z value was calculated as 3,002 which is far bigger than 1,96 as the critical barrier. So it is convenient to say that new design monitor is superior to Reference in Radiologist 1's evaluations with 95% confidence interval.

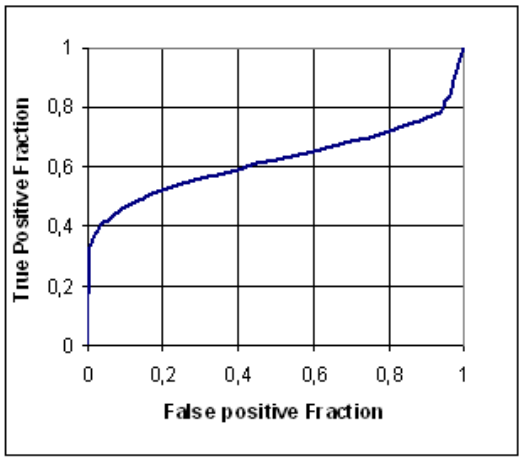

Figure 5.1 ROC Curve for Radiologist 1 with Reference Monitor

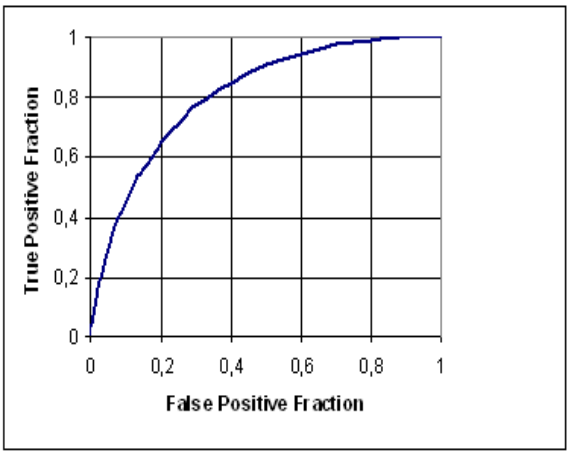

Figure 5.2 ROC Curve for Radiologist 1 with New Design Monitor

Table 5.20
Statistical Comparison of Radiologist 2's evaluations on both monitors

Reference Monitor-Radiolog2		**New Design Monitor-Radiolog 2**	
Number of Cases	60	Number of Cases	60
Number Correct	41	Number Correct	41
Accuracy	68,30%	Accuracy	68,30%
Sensitivity	63,30%	Sensitivity	60,00%
Specificity	73,30%	Specificity	76,70%
Pos. Cases Missed	11	Pos. Cases Missed	12
Neg. Cases Missed	8	Neg. Cases Missed	7
Fitted ROC Area	0,703	Fitted ROC Area	0,746
Empiric ROC Area	0,744	Empiric ROC Area	0,762
A_1	0,703	A_2	0,746
SE_1	0,067	SE_2	0,064
Q_1	0,542	Q_1	0,595
Q_2	0,580	Q_2	0,637
SE	0,071		
r_{N2}	0,679	r_{N2}	
r_{A2}	0,220	r_{A2}	
r	0,449	r	
n_A	30	n_A	30
n_N	30	n_N	30
A	0,725	A	
z	0,603	Z	
r_x	0,410	r_x	

For Radiologist 2's observations, although AUC value for New Design monitor is bigger than Reference monitor, z value was calculated as 0,603 which is less than 1,96 as the critical barrier. So it is not possible to say that new design monitor is superior to Reference one in Radiologist 2's evaluations with 95% confidence interval.

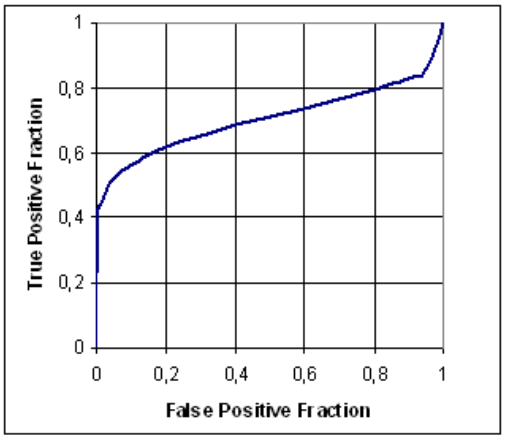

Figure 5.3 ROC Curve for Radiologist 2 with Reference Monitor

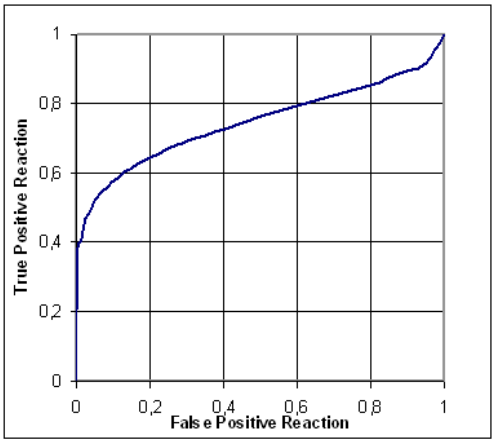

Figure 5.4 ROC Curve for Radiologist 2 with New Design Monitor

Table 5.21
Statistical Comparison of Radiologist 3's evaluations on both monitors

Reference Monitor-Radiolog3		**New Design Monitor-Radiolog 3**	
Number of Cases	60	Number of Cases	60
Number Correct	42	Number Correct	43
Accuracy	70,00%	Accuracy	71,70%
Sensitivity	66,70%	Sensitivity	70,00%
Specificity	73,30%	Specificity	73,30%
Pos Cases Missed	10	Pos Cases Missed	9
Neg Cases Missed	8	Neg Cases Missed	8
Fitted ROC Area	0,755	Fitted ROC Area	0,811
Empiric ROC Area	0,762	Empiric ROC Area	0,792
A_1	0,755	A_2	0,811
SE_1	0,063	SE_2	0,056
Q_1	0,606	Q_1	0,682
Q_2	0,650	Q_2	0,726
SE	0,049		
r_{N3}	0,855	r_{N3}	
r_{A3}	0,554	r_{A3}	
r	0,704	r	
n_A	30	n_A	30
n_N	30	n_N	30
A	0,783	A	
z	1,132	z	
r_x	0,660	r_x	

For Radiologist 3's observations, although AUC value for New Design monitor is bigger than Reference monitor, z value was calculated as 1,132 which is still less than 1,96 as the critical barrier. So it is not possible to say that new design monitor is superior to Reference one in Radiologist 3's evaluations with 95% confidence interval.

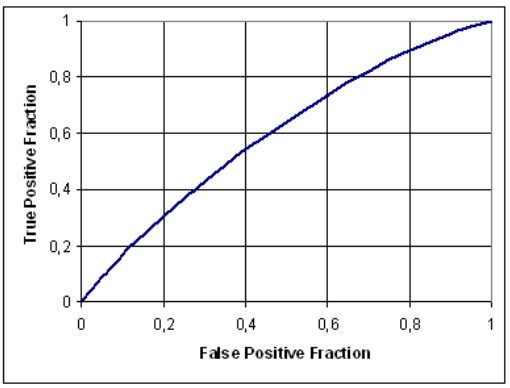

Figure 5.5 ROC Curve for Radiologist 3 with Reference Monitor

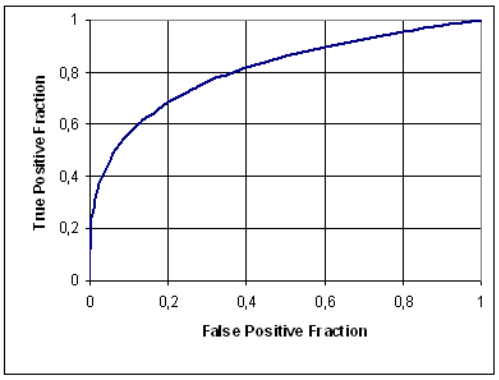

Figure 5.6 ROC Curve for Radiologist 3 with New Design Monitor

6. CONCLUSION

6.1 Discussion and Concrete Results

After the analysis of all ROC curves and related statistical data for all radiologist-monitor pairs, it is possible to say that new design monitor is convenient to be used as a PACS monitor without losing any diagnostic value. As specified in Table 6.1, for radiologist 1, AUC value for Reference monitor is 0.634 which is far less than AUC value of 0.811 for the new design. So with 95% confidence interval, for radiologist 1, new design got better diagnostic performance than Reference monitor. For radiologist 2 and 3, although AUC values and naturally z values were far better on behalf of the new design, no significant statistical difference could be proven.

Table 6.1
Comparison of radiologists' evaluation on Reference and New Design monitors

	Area Under the ROC Curve (AUC)		z value
Radiologist	Reference Monitor	New Design	Reference-New Design
1	0.634	0.811	3.002
2	0.703	0.746	0.603
3	0.755	0.811	1.132

As a summary, with the advent of new consumer grade TFT LCD panels with Full HD resolution, high contrast ratio, high brightness and with new generation chipsets for monitor applications, it becomes possible to use consumer grade monitors with enhanced features as PACS monitors.

6.2 Future Work

A new design can be targeted which will have the ability to divide the screen in to two halves so as to show two separate images on each half simultaneously. This is mainly a software work. Furthermore, as available in most of medical monitors, a photo sensor could be mounted on the front plastic cabinet which will sense the ambient light level to drive the back light of the panel to be in line with room lighting conditions. As the final step, the new design monitor could be also tested through the criteria heralded by AAPM, "The American Association of Physicists in Medicine" AAPM criteria for testing medical monitors is accepted as the gold standard before the launch of a medical monitor in to commercial field. This criteria relies on some well-defined physical parameters' measurement under special pattern usage. Routine visual evaluations of performance are conveniently done using a single comprehensive test pattern. A new pattern designed by the AAPM Task Group 18 committee, referred to in this thesis as the TG18-QC pattern, is recommended for overall display quality assessment.

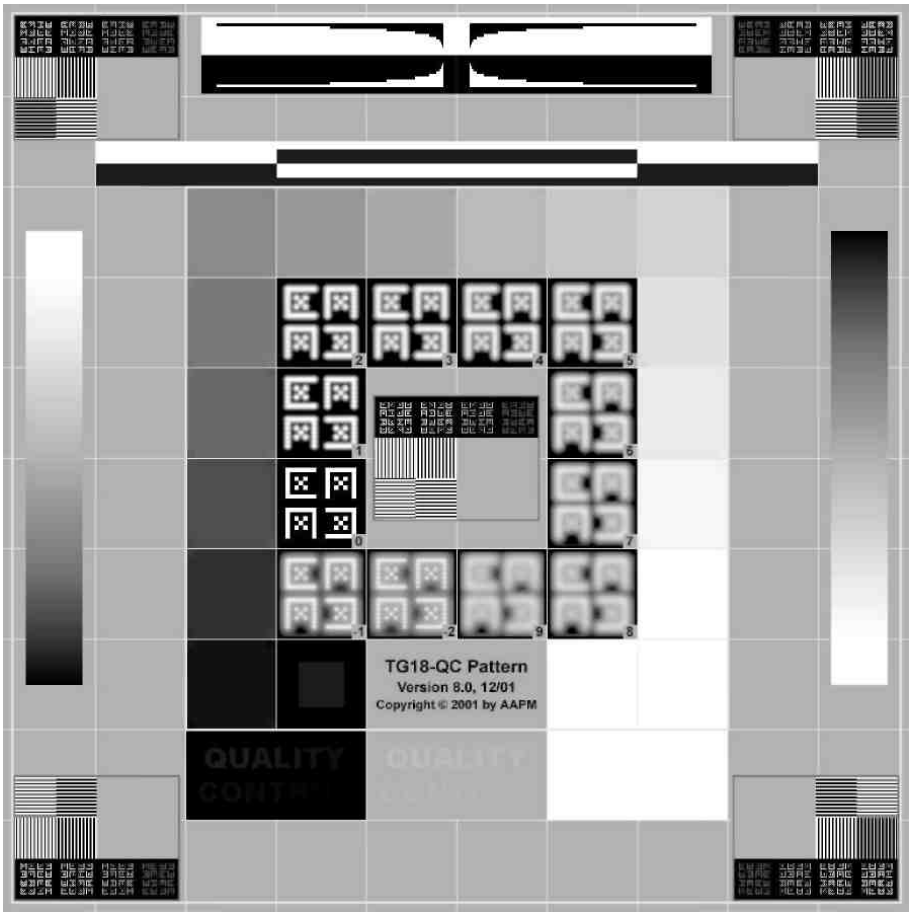

Figure 6.1. The TG18-QC Test Pattern.

The TG18-QC test pattern is shown in Figure 6.6 consists of multiple inserts embedded in a mid-pixel value background. The inserts include the following:

1. Grid lines (one pixel) with thicker lines (three pixels) along periphery and around central region, for the evaluation of geometric distortions.
2. Sixteen 102 × 102 (1k version) luminance patches with pixel values varying from 8 to 248 (in 8-bit version) [128 to 3968 in 12-bit version] for luminance response evaluation. Each patch contains four small 10 × 10 corner patches (1k version) at ±4 [±64] of pixel value difference from the background, +4 [+64] in upper left and lower right, –4 [–64] in lower left and upper right. The small patches

are used for visual assessment of luminance response. Additionally, two patches with minimum and maximum pixel value are embedded containing 13 [205], and 242 [3890] pixel value internal patches, similar to 5% and 95% areas in the SMPTE test pattern.

3. Line-pair patterns at the center and four corners at Nyquist and half-Nyquist frequencies for resolution evaluation, having pixel values at 0–255 [0–4095] and 128–130 [2048–2088].
4. "Cx" patterns at the center and four corners with pixel values of 100, 75, 50, and 25% of maximum pixel values against a zero pixel value background, for resolution evaluation in reference to a set of 12 embedded scoring references with various amounts of Gaussian blurring applied.
5. Contrast-detail "QUALITY CONTROL" letters with various contrasts at minimum, midpoint, and maximum pixel values for user-friendly low-contrast detectability at three luminance levels.
6. Two vertical bars with continuous pixel value variation for evaluating bit depth and contouring artifacts.
7. White and black bars for evaluating video signal artifacts, similar to those in the SMPTE pattern.
8. A horizontal area at the top center of the pattern for visual characterization of cross talk in flat-panel displays.
9. A border around the outside of the pattern, similar to SMPTE's.

The evaluation of QC pattern is visual. When TG18 QC pattern is displayed on the new design monitor, the image quality of new design monitor from resolution, luminance, distortion, artifacts point of view also seems in the acceptable region. The tests must be done in an accredited laboratory for international certification.

REFERENCES

1. Zhou XH, Obuchowski Nancy A, McClish Donna K. *Statistical Methods in Diagnostic Medicine*, 1st ed. New York: John Wiley & Sons, 2002:10-393

2. Pepe Margaret Sullivan, *The Statistical Evaluation of Medical Tests for Classification and Prediction*, 2nd ed. New York: Oxford University Press, 2004:1-279

3. Gönen Mithat, *Analyzing Receiver Operating Characteristics Curves with SAS*, 1st ed. Cary, NC, SAS Institute Inc, 2007:1-97

4. Geijer Hakan, Geijer Mats, Forsberg Lillemor, Kheddache Susanne, Sund Patrik, *Comparison of Color LCD and Medical Grade Monochrome LCD Displays in Diagnostic Radiology*, Journal of Digital Imaging, Vol.20, No.2, 2007:114-121

5. Buls N, Shabana W, Verbeek P, Pevenage P, Mey J De, *Influence of Display Quality on Radiologists' Performance in the Detection of Lung Nodules on Radiographs,* British Institute of Radiology, Vol.80, 2007:738-743

6. Obuchowski NA, *Receiver operating characteristic curves and their use in radiology, Radiology* 2003, 229:3-8

7. Hanley JA, McNeil BJ, *The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology* 1982, 143:29-36

8. DeLong ER, DeLong DM, Clarke-Pearson DL, *Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach*, *Biometrics* 1988;44:837-844

9. Bamber D. *The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. JMath Psychology* 1975, 12:387-415

10. Usami H, Ikeda M, Ishigakil T, Fukushima H, Shimamoto K, . *The influence of liquid crystal display (LCD) monitors on observer performance for the detection of nodular lesions on chest radiographs, European Radiology* 2006,16:726-32

11. Hanley J A, McNeil BJ, *A method for comparing the area under two ROC curves derived from the same cases*, *Radiology*1983, 148:839-843

12. AU Optronics, *TFT LCD Introduction*. http://auo.com/auoDEV/technology.php?sec=tftIntro&ls=en