J Digit Imaging (2014) 27:679–686
DOI 10.1007/s10278-014-9695-y

# Evaluation of Low-Cost Telemammography Screening Configurations: A Comparison with Film-Screen Readings in Vulnerable Areas

**Antonio J. Salazar · Javier Romero · Oscar Bernal · Angela Moreno · Sofía Velasco · Xavier Díaz**

Published online: 7 May 2014


**Abstract** The aim of this study was to evaluate the diagnostic accuracy for detecting breast cancer using different telemammography configurations, including combinations of both low-cost capture devices and consumer-grade color displays. At the same time, we compared each of these configurations to film-screen readings. This study used a treatment-by-reader-by-case factorial design. The sample included 70 mammograms with 34 malignant cases. The readers consisted of four radiologists who classified all of the cases according to the categories defined by the Breast Imaging Reporting and Data System (BI-RADS). The evaluated capture devices included a specialized film digitizer and a digital camera, and the evaluated displays included liquid crystal display (LCD) and light-emitting diode (LED) consumer-grade color displays. Receiver operating characteristic curves, diagnostic accuracy (measured as the area under these curves), accuracy of the composition classification, sensitivity, specificity, and the degree of agreement between readers in the detection of malignant cases were also evaluated. Comparisons of diagnostic accuracy between film-screen and the different combinations of digital configurations showed no significant differences for nodules, calcifications, and asymmetries. In addition, no differences were observed in terms of sensibility or specificity when the degree of malignancy using the film-screen method was compared to that provided with digital configurations. Similar results were observed for the classification of breast composition. Furthermore, all observed reader agreements of malignant detection between film-screen and digital configurations were substantial. These findings indicate that the evaluated digital devices showed comparable diagnostic accuracy to the reference treatment (film-screen).

**Keywords** Mammography · Teleradiology · Display device · Observer performance · ROC-based analysis · Sensitivity and specificity

A. J. Salazar (✉) · X. Díaz
Department of Electrical and Electronic Engineering, Laboratory of Telemedicine, University of Los Andes, Carrera 1 Este No. 19A-40, Bogotá, Colombia
e-mail: ant-sala@uniandes.edu.co

X. Díaz
e-mail: xa.diaz32@uniandes.edu.co

J. Romero · A. Moreno
Department of Diagnostic Imaging, Fundacion Santa Fe de Bogota, University Hospital, Calle 119 No. 7-75, Bogotá, Colombia

J. Romero
e-mail: jaromero@uniandes.edu.co

A. Moreno
e-mail: angela__m@hotmail.com

O. Bernal · S. Velasco
School of Government, University of Los Andes, Cra. 1E No. 19A-40, Bogotá, Colombia

O. Bernal
e-mail: obernal@uniandes.edu.co

S. Velasco
e-mail: sc.velasco62@uniandes.edu.co

J. Romero · O. Bernal
School of Medicine, University of Los Andes, Cra. 1E No. 19A-40, Bogotá, Colombia

## Introduction

To reduce the mortality rates associated with breast cancer, screening programs including routine mammograms have been developed [1–5]. For rural or underserved areas, telemedicine may provide a cost-effective solution for screening mammography programs and computed radiography (CR) and full-field digital mammography (FFDM) are useful in the implementation of telemammography services. Although

previous studies have reported no significant differences between film-screen mammography and digital modalities [6–8], this latter technology may be unaffordable for vulnerable populations, where only conventional film-screen mammograms may be available. Therefore, the digitization of films is required to provide telemammography services.

Specialized equipment is available for digitizing mammogram films, and several studies have confirmed their clinical equivalence to film-screen mammography [9–11]. However, similar to CR, these specialized digitizers also remain unaffordable for rural areas in developing countries, and digital cameras are therefore being utilized for teleradiology services. Such equipment may reduce costs, although diagnostic accuracy should be evaluated prior to adopting these new medical services. In addition, for teleradiology diagnostic delivery by radiologists at homes or in small offices, specialized medical-grade grayscale displays also remain unaffordable. As a result, it is important to evaluate standard color consumer-grade displays that may be used in teleradiology applications with images originating from digitized plain films. However, few studies have evaluated the diagnostic accuracy of such displays [12, 13], and the possibility of using color LCD medical- or consumer-grade displays instead of medical-grade grayscale displays remains to be determined [14, 15].

The aim of this study was to evaluate the diagnostic accuracy for detecting breast cancer using different telemammography configurations, including combinations of low-cost capture devices (specialized digitizer and digital cameras, which are less expensive than CR or FFDM) and low-cost displays (consumer-grade color displays, which are less expensive than medical-grade grayscale displays). At the same time, we sought to compare these configurations to film-screen readings. In particular, we evaluated the detection of nodules, calcifications and asymmetries using receiver operating characteristic curve (ROC) analysis. The sensitivity, specificity, and reader agreement for the degree of malignancy observed by the radiologist, according to the categories defined by the Breast Imaging Reporting and Data System (BI-RADS) [16], were evaluated, as well as the accuracy of the composition classification.

The current evaluation represents part of a larger study, in which we included mammography images obtained using CR and assessed visualization using medical-grade displays. When a medical-grade device was used at one end of the configuration (i.e., capture or display) in combination with a low-cost alternative, no significant differences were observed in comparison to film-screen readings [17].

The capture devices and displays evaluated in this study were also evaluated in our previous studies for chest radiography, and no significant differences in performance were noted in terms of detection between the capture devices [18–20] and displays evaluated [21–23]. In the current study, we evaluated whether the same devices could be used for mammography using low-cost solutions at both ends of the teleradiology chain, including capture with a specialized digitizer or digital camera and display with consumer-grade color LCD or LED displays.

## Methods

This retrospective study was approved by the ethics committee of our institution, and informed consent was not required. The current evaluation was part of a larger factorial design study that included 70 cases, four radiologists, three displays (one medical- and two consumer-grade color displays), two display calibration methods, three capture devices, and the reference readings (i.e., film-screen), for a total of 5,320 observations for each variable. However, the current analysis only included film-screen observations and those that were obtained using low-cost displays (with factory and with Digital Imaging and Communications in Medicine (DICOM) calibration) and two low-cost capture devices, for a total of 2,520 observations for each variable.

### Sample, Standard, and Readers

The standard for a positive case was a malignant lesion confirmed by biopsy within 2 years of the initial mammography screening, corresponding to BI-RADS category 4 or 5 [6, 10, 24]. Negative cases were defined as those without any lesions confirmed by biopsy or cases with normal follow-up mammograms for 1 year, corresponding to BI-RADS category 2 or 3. Mammography studies from patients who attended mammography screenings between 2008 and 2012 at the Fundación Santa Fe de Bogotá (FSFB) were selected without repetition. Each case was required to include four views: mediolateral oblique, craniocaudal, left, and right. Cases with calcifications, nodules, asymmetries, and distortions were included. Conditions in cases were not exclusive, i.e., multiple types of lesions per case were allowed; to avoid sensitivity bias, cases with obvious lesions were not included.

The first analysis in this evaluation was based on receiver operating characteristics (ROC) curves, and the diagnostic accuracy was calculated as the area under the ROC curve (AUC). To determine the sample size for comparisons of AUCs, we used the table proposed by Obuchowski [25], with the following criteria: (1) four observers, (2) low variability between radiologists and high accuracy, (3) moderate differences among AUCs (i.e., 0.1) and (4) a 1:1 ratio between malignant and benign cases. Using these criteria, 67 cases were required. The sample size was set at 70 cases. The distribution of cases according to type of nodule and calcification, according to the BI-RADS classification, is shown in Table 1. To achieve a sample with a 1:1 distribution of benign

cases and malignant cases (as shown in Table 2), we chose 36 benign cases and 34 malign cases. As BI-RADS 4B are more frequent in our database than BI-RADS 4A, 4C and 5, we chose more cases from BI-RADS 4B. In a first step, we filtered in the database cases that belonged to a specific BI-RADS category, and then we chose a case at random from the screen list. If this case had any exclusion criterion (e.g., large sizes, incomplete views) or was already included in the sample, this case was excluded, otherwise it was included. The same procedure was repeated until the previously established number of cases per BI-RADS category was achieved. This procedure was repeated for each BI-RADS category. Patients ranged in age from 41 to 84 years, with a mean of 62.1 years.

The distribution composition of the cases was as follows: 17 almost entirely fatty, 32 scattered fibroglandular, 11 heterogeneously dense, and 10 extremely dense. Four radiologists, who read about 8,000 mammograms per year at our hospital, and with experiences between 2 and 10 years after board certification, served as observers.

### Observed Variables

For each interpretation, the radiologist examined the following variables and classified them according to the BI-RADS: (1) classification of pathological findings, (2) breast composition, and (3) degree of malignancy. Based on these variables, several assessment variables were defined and are presented in the following data analysis section.

### Mammograms and Capture Devices

Routine mammograms on 18×24 cm films were digitized using the following two capture devices: (1) an iCR 612SL specialized digitizer (iCR Company, Torrance, CA), with a maximum spatial resolution of 875 dpi, a pixel spot of 29 μm, 16 bits per pixel, an optical density (OD) of 3.6 and a cost of $15,000 (hereafter referred to as ICR), and (2) a Lumix DMC-FZ28 digital camera (Panasonic Corporation, Secaucus, NJ), with a 10-megapixel resolution, focal length of 4.8 to 86.4 mm, 1/2.33″ charge-coupled device (CCD), ISO 100–6400 and cost of $450 (plus $400 for a support system and light box). The digital camera is hereafter referred to as LUMIX.

For each patient (case), the following three case-studies were used in this evaluation: (1) the printed film, hereafter referred to as FILM, (2) images digitized with ICR (2,436×3,636 pixel matrix and 8-bit grayscale), and (3) images digitized with LUMIX (2,538×3,463 pixel matrix and 8-bit grayscale). This procedure was completed for each of the 70 sample mammograms, which produced 210 case studies.

**Table 1** Detailed classification of the tested conditions

| Condition | Classification | BI-RADS[a] | Cases |
|---|---|---|---|
| Nodules | Normal | 1 | 44 |
| | Well defined | 3 | 7 |
| | Obscured | 4B | 10 |
| | Poorly defined | 4C | 4 |
| | Spiculated | 5 | 5 |
| Calcifications | Normal | 1 | 13 |
| | Benign | 2 | 33 |
| | Grouped punctuate | 3 | 4 |
| | Coarse heterogeneous | 4A | 8 |
| | Amorphous | 4B | 7 |
| | Fine pleomorphic | 4C | 4 |
| | Pleomorphic ductal pattern | 5 | 1 |

[a] Classification according to the American College of Radiology15

### Visualization Displays

For visualization, two consumer-grade color displays were selected: (1) a Dell UltraSharp U2711 LCD (Dell Computer Corporation, Round Rock, TX, USA), hereafter referred to as LCD, with a dot pitch of 0.23 mm, a spatial resolution of 2,560×1,440 pixels, a maximum luminance of 350 cd/m$^2$, and a cost of $862; and (2) the LED display of a Dell Vostro 3750 laptop computer, hereafter referred to as LED, with a dot pitch of 0.24 mm, spatial resolution of 1,600×900 pixels, maximum luminance of 220 cd/m$^2$, and a cost of $780. These two displays were used with both the factory calibration and with the "Part 14: Grayscale Standard Display Function" (GSDF) DICOM standard calibration, as presented in our previous study [17].

### Data Analysis

To compare the AUCs for the detection of nodules, calcifications, and asymmetries, IBM SPSS Statistics 19 (SPSS Inc.,

**Table 2** Overall distribution of cases in the sample according to the degree of malignancy

| BI-RADS[a] | Malignant | Cases | |
|---|---|---|---|
| 2—Benign finding | No | 18 | 36 |
| 3—Most likely benign finding | No | 18 | |
| 4A—Low suspicion | Yes | 6 | 34 |
| 4B—Intermediate suspicion | Yes | 16 | |
| 4C—Moderate suspicion | Yes | 6 | |
| 5—Highly suggestive of malignancy | Yes | 6 | |
| Total | | 70 | 70 |

[a] Classification according to the American College of Radiology15

USA) was used to compare the analyses of variance (ANOVAs) of pseudovalues from the AUC ROC curves. The pseudovalues were obtained using the DBM-MRMC 2.3 software [26] developed by Dorfman-Berbaum-Metz [26, 27].

According to the degree of malignancy observed by radiologist, we calculated the sensitivity and specificity using generalized estimating equations (GEEs) in SPSS. Cases that were classified as BI-RADS 2 and 3 were considered negative, and cases that were classified as 4a, 4b, 4c, and 5 were considered positive, i.e., malignant.

For the tissue composition type variable, the proportion of correct readings compared to the gold standard was calculated using the GEE function in SPSS. Finally, agreements on malignant detection between the film and digitizer readings were calculated using STATA 12 software (Stata Corp., College Station, TX, USA).

## Procedure

The software features included image manipulation functions (e.g., filters, zoom, window/level, and negative/positive) that could be used at the observer's discretion and case blinding and data integrity to ensure that all case readings had been completed. Each radiologist evaluated the readings over a 6-month period in 4-hour sessions in a room were the ambient luminosity for all readings was set to 6 lx. For each observer and for a particular display, the viewing order was randomly assigned for all the possible combinations of the 70 cases, capture device, and calibration method. For the next display, the viewing order was modified, assuring a 70-case interval between two observations of the same case by a radiologist. The time interval between sessions was 1 week and the time interval between displays was about 4 weeks. During each

**Table 3** Comparison of area under the ROC curve (AUC) values for film and digital configuration readings

| Display calibration | Capture device | Display | AUC[a] | SE | 95 % CI | | *p* value |
|---|---|---|---|---|---|---|---|
| | | | | | LB | UB | |
| Nodules | | | | | | | |
| | FILM | | 0.89 | 0.042 | 0.805 | 0.974 | 0.943 |
| Factory | LED | ICR | 0.87 | 0.040 | 0.793 | 0.955 | |
| | | LUMIX | 0.88 | 0.035 | 0.805 | 0.945 | |
| | LCD | ICR | 0.87 | 0.036 | 0.800 | 0.944 | |
| | | LUMIX | 0.90 | 0.031 | 0.838 | 0.962 | |
| DICOM | LED | ICR | 0.87 | 0.038 | 0.796 | 0.948 | |
| | | LUMIX | 0.88 | 0.037 | 0.803 | 0.951 | |
| | LCD | ICR | 0.87 | 0.040 | 0.788 | 0.948 | |
| | | LUMIX | 0.89 | 0.039 | 0.817 | 0.972 | |
| Calcifications | | | | | | | |
| | FILM | | 0.69 | 0.026 | 0.639 | 0.741 | 0.74 |
| Factory | LED | ICR | 0.67 | 0.028 | 0.612 | 0.722 | |
| | | LUMIX | 0.63 | 0.032 | 0.565 | 0.692 | |
| | LCD | ICR | 0.66 | 0.025 | 0.610 | 0.709 | |
| | | LUMIX | 0.70 | 0.054 | 0.593 | 0.810 | |
| DICOM | LED | ICR | 0.66 | 0.029 | 0.602 | 0.718 | |
| | | LUMIX | 0.76 | 0.028 | 0.707 | 0.819 | |
| | LCD | ICR | 0.78 | 0.051 | 0.674 | 0.878 | |
| | | LUMIX | 0.68 | 0.023 | 0.637 | 0.727 | |
| Asymmetries | | | | | | | |
| | FILM | | 0.80 | 0.042 | 0.715 | 0.884 | 0.114 |
| Factory | LED | ICR | 0.79 | 0.037 | 0.720 | 0.867 | |
| | | LUMIX | 0.73 | 0.045 | 0.635 | 0.814 | |
| | LCD | ICR | 0.67 | 0.048 | 0.576 | 0.766 | |
| | | LUMIX | 0.72 | 0.080 | 0.558 | 0.878 | |
| DICOM | LED | ICR | 0.70 | 0.059 | 0.579 | 0.816 | |
| | | LUMIX | 0.80 | 0.053 | 0.694 | 0.908 | |
| | LCD | ICR | 0.67 | 0.063 | 0.549 | 0.799 | |
| | | LUMIX | 0.73 | 0.050 | 0.633 | 0.832 | |

*ROC* receiver operating characteristic, *SE* standard error of the mean, *CI* confidence interval, *LB* lower bound, *UB* upper bound

[a] Each mean AUC was calculated for 280 observations

session, the display was the same for all radiologists, but the order of display assigned to radiologists was different.

## Results

The AUC values that were obtained for the detection of breast abnormalities, with their respective 95 % confidence intervals (CI), are shown in Table 3. For nodules, the AUC value observed for FILM was 0.89, while the values for the digital configurations ranged from 0.87–0.90; for calcifications, the AUC value observed for FILM was 0.69, while the values for the digital configurations ranged from 0.63–0.78; for asymmetries, the AUC value observed for FILM was 0.80, while the values for the digital configurations ranged from 0.67–0.80. The result of the omnibus statistical comparison for each condition ($p$ value) is shown in the last column of Table 3. For a sample size that was designed to detect differences of 0.1 in the AUC with a power of 80 %, none of the comparisons exhibited a significant difference; the observed $p$ values were 0.943, 0.74, and 0.114 for nodules, calcifications, and asymmetries, respectively.

The values for sensitivity and specificity and their corresponding confidence intervals are shown in Table 4. The observed sensitivities ranged from 0.70–0.74, with a value of 0.71 for FILM, and there were no significant differences between the nine evaluated methods ($p$=0.841). The observed specificities ranged from 0.83–0.90, with 0.90 for FILM, and there were no significant differences noted ($p$=0.175).

The proportions of readings with compositions that were correctly classified ranged from 0.53–0.65, with a value of 0.60 observed for FILM (see Table 5). In the overall comparison, there were significant differences between the examined methods ($p$=0.03). In addition, when FILM was compared to the digital methods grouped by calibration, no differences were observed ($p$=0.818 for factory calibration; $p$=0.422 for GSDF calibration).

The agreement on malignant detection between FILM and each of the configurations is shown in Table 6. The observed agreement values ranged from 81.8–87.5 %, with kappa values ranging from 0.62–0.74. According to the Landis and Koch classification [28], agreement on malignant detection between FILM and each of the configurations was substantial.

## Discussion

The comparisons of diagnostic accuracy (calculated as the area under ROC curves) between FILM and the different combinations of digital configurations (ICR, LUMIX, LED, LCD, factory calibration, and DICOM calibration) showed no significant differences for nodules, calcifications, and

**Table 4** Sensitivity and specificity values for film and digital configuration readings

| Display calibration | Capture device | Display | Mean[a] | SE | 95 % CI | | $p$ value |
|---|---|---|---|---|---|---|---|
| | | | | | LB | UB | |
| Sensitivity | | | | | | | |
| | FILM | | 0.71 | 0.06 | 0.60 | 0.83 | 0.841 |
| Factory | LUMIX | LED | 0.74 | 0.05 | 0.64 | 0.83 | |
| | | LCD | 0.74 | 0.05 | 0.64 | 0.84 | |
| | ICR | LED | 0.73 | 0.05 | 0.63 | 0.83 | |
| | | LCD | 0.74 | 0.05 | 0.64 | 0.84 | |
| DICOM | LUMIX | LED | 0.72 | 0.05 | 0.62 | 0.82 | |
| | | LCD | 0.70 | 0.05 | 0.59 | 0.81 | |
| | ICR | LED | 0.71 | 0.06 | 0.60 | 0.83 | |
| | | LCD | 0.71 | 0.05 | 0.60 | 0.81 | |
| Specificity | | | | | | | |
| | FILM | | 0.90 | 0.03 | 0.84 | 0.95 | 0.175 |
| Factory | LUMIX | LED | 0.85 | 0.03 | 0.80 | 0.91 | |
| | | LCD | 0.85 | 0.03 | 0.78 | 0.91 | |
| | ICR | LED | 0.83 | 0.04 | 0.76 | 0.91 | |
| | | LCD | 0.83 | 0.03 | 0.77 | 0.89 | |
| DICOM | LUMIX | LED | 0.86 | 0.03 | 0.80 | 0.92 | |
| | | LCD | 0.85 | 0.03 | 0.79 | 0.91 | |
| | ICR | LED | 0.86 | 0.03 | 0.80 | 0.92 | |
| | | LCD | 0.86 | 0.03 | 0.80 | 0.93 | |

*ROC* receiver operating characteristic, *SE* standard error of the mean, *CI* confidence interval, *LB* lower bound, *UB* upper bound

[a] Each mean value was calculated for 280 observations

**Table 5** Proportion of cases with correctly classified composition

| Display calibration | Capture device | Display | Mean[a] | SE | 95 % CI | | *p* value |
|---|---|---|---|---|---|---|---|
| | | | | | LB | UB | |
| Factory | FILM | | 0.60 | 0.04 | 0.53 | 0.68 | 0.03 |
| | LUMIX | LED | 0.56 | 0.04 | 0.47 | 0.65 | |
| | | LCD | 0.53 | 0.05 | 0.43 | 0.62 | |
| | ICR | LED | 0.61 | 0.04 | 0.52 | 0.70 | |
| | | LCD | 0.61 | 0.05 | 0.52 | 0.71 | |
| DICOM | LUMIX | LED | 0.63 | 0.04 | 0.55 | 0.72 | |
| | | LCD | 0.63 | 0.05 | 0.54 | 0.72 | |
| | ICR | LED | 0.60 | 0.05 | 0.51 | 0.69 | |
| | | LCD | 0.65 | 0.05 | 0.55 | 0.74 | |

*SE* standard error of the mean, *CI* confidence interval, *LB* lower bound, *UB* upper bound

[a] Each mean value was calculated for 280 observations

asymmetries. In addition, no differences were observed in sensitivity or specificity when the degree of malignancy using FILM was compared to that obtained using the digital configurations. The same result was observed for the classification of breast composition when FILM was compared to digital configurations grouped by calibration factor. Furthermore, all observed reader agreement on malignant detection between FILM and digital configurations was substantial. These findings indicate that the digital devices provided comparable diagnostic accuracy to the reference treatment (FILM). In addition, these results are in agreement with our previous evaluation of displays (also including a medical-grade grayscale display) using images obtained with CR [17], as well as the evaluation of capture devices (including standard CR) visualized using a medical-grade grayscale display.

Previous studies have compared film-screen mammography to digitized films [9–11], and similar to the present study, these reports found no significant differences in diagnostic accuracy. In our study, we obtained AUC mean values that were similar to the values reported by Powell [9], Gitlin et al. [10], and Pisano et al. [11]. However, these previous studies did not evaluate digital cameras and low-cost displays. Additionally, to our knowledge, no previous study has evaluated the performance of observers reading mammograms captured with a digital camera, for which the present study obtained AUC values that were similar to the values for FILM obtained with all displays (and calibrations). The study of Chen and Gale [29] evaluated the accuracy in mammography with low cost displays—including an iPAD—compared with a medical-grade display, finding as in the present study, comparable results between the two displays, while poor performance was detected for the iPAD. The study of Krupinski [30] found also comparable result between medical-grade displays and consumer-grade displays, with more appreciated differences in the evaluation of subtle lesions, especially when the monitors have been used more than a year. In addition, in part of the our overall study when images from CR were included and all the images were reading using a medical-grade display, in addition to the consumer-grade displays of this evaluation, no differences were observed in detection of nodules and calcifications in screening mammograms when GSDF calibration or factory calibration was used [17].

With the capability of detecting differences of 0.1 with diagnostic accuracy at a power of 80 %, our results indicate that

**Table 6** Agreements on malignant classification between film and digital configuration readings

| Display calibration | Capture device | Display | OA (%) | EA (%) | Kappa[a] | SE | *Z*-test | *p* value | Agreement[b] |
|---|---|---|---|---|---|---|---|---|---|
| Factory | LUMIX | LED | 83.2 | 51.4 | 0.65 | 0.0596 | 10.98 | <0.0001 | Substantial |
| | | LCD | 85.0 | 51.3 | 0.69 | 0.0596 | 11.61 | <0.0001 | Substantial |
| | ICR | LED | 85.4 | 51.2 | 0.70 | 0.0596 | 11.75 | <0.0001 | Substantial |
| | | LCD | 84.3 | 51.1 | 0.68 | 0.0595 | 11.40 | <0.0001 | Substantial |
| DICOM | LUMIX | LED | 82.4 | 51.6 | 0.65 | 0.0597 | 10.82 | <0.0001 | Substantial |
| | | LCD | 81.8 | 51.6 | 0.62 | 0.0597 | 10.44 | <0.0001 | Substantial |
| | ICR | LED | 87.5 | 51.6 | 0.74 | 0.0597 | 12.42 | <0.0001 | Substantial |
| | | LCD | 82.9 | 51.7 | 0.65 | 0.0597 | 10.80 | <0.0001 | Substantial |

*OA* observed agreement, *EA* expected agreement, *SE* standard error

[a] Each agreement level was calculated from 560 readings

[b] According the Landis and Koch classification28

FILM, ICR, and LUMIX may all be suitable for detecting malignant or benign mammograms in screening programs when consumer-grade color displays are used. This finding suggests that high-quality medical imaging services, including telemammography screening programs, can be provided to underserved populations at a low cost, which would extend the coverage of health services related to prevention and early detection.

When the statistical analysis was performed for all of the observations (i.e., the 5,320 overall readings), the power of the tests showing no significant results were larger than that noted in the present evaluation. Therefore, to definitively establish the equivalence of these low-cost alternatives to film readings, further study with a larger sample size is required.

Usually, in deserved rural areas only film-screen mammograms may be available. However, a database with real rural film-screen mammograms was not available at the time of this study. For this reason, we used CR images from the database of radiologic imaging studies of our hospital, which before the end of this study, were the routine mammograms and the best quality images available in our hospital—until the installation of the tomosynthesis mammography in the last months. However, as printed CR images may be quite different from conventional film-screen mammograms, the results of this study should be revisited using mammography images obtained at rural hospitals with different qualities of equipment and technician standards. Another limitation of our study was the way in which DICOM GSDF calibration was implemented, as the luminance transformed characteristic curve obtained after calibration is an approximation of the GSDF curve, because the calibrated values are selected from a 8-bit look-up table (LUT), producing "quantization errors" [31]. These errors may be reduced using a LUT from the hardware driver of the display, using more complex methods, as the International Color Consortium (ICC) color profile [32]; nevertheless, these profiles are not always available for all low cost display monitors.

**Acknowledgments** We thank the radiologists Bibiana Pinzón, Gloria Palazuelos, Nelson Bedoya, and Sandra Ochoa, who carried out the readings. We also thank our institutions and the National Department of Science, Technology and Innovation (Departamento Administrativo de Ciencia, Tecnología e Innovación - Colciencias) for funding this study.

We thank the University of Los Andes and the Fundación Santa Fe de Bogotá - University Hospital (the author's institutions) for funding this study.

**Disclosure Statement** No competing financial interests exist.

## References

1. Smith RA, Caleffi M, Albert U-S, Chen THH, Duffy SW, Franceschi D, Nyström L: Breast Cancer in Limited-Resource Countries: Early Detection and Access to Care. Breast J 12: S16–S26, 2006
2. Anderson BO, Braun S, Lim S, Smith RA, Taplin S, Thomas DB: Early Detection of Breast Cancer in Countries with Limited Resources. Breast J 9:S51–S59, 2003
3. Smith R: IARC Handbooks of Cancer Prevention, Volume 7: Breast Cancer Screening. Breast Cancer Res 5:216–217, 2003
4. Humphrey LL, Helfand M, Chan BKS, Woolf SH: Breast Cancer Screening: A Summary of the Evidence for the U.S. Preventive Services Task Force. Ann Intern Med 137:347–360, 2002
5. Fletcher SW, Elmore JG: Mammographic Screening for Breast Cancer. N Engl J Med 348:1672–1680, 2003
6. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, Conant EF, Fajardo LL, Bassett L, D'Orsi C, Jong R, Rebner M: Diagnostic Performance of Digital versus Film Mammography for Breast-Cancer Screening. N Engl J Med 353:1773–1783, 2005
7. Kamitani T, Yabuuchi H, Soeda H, Matsuo Y, Okafuji T, Sakai S, Furuya A, Hatakenaka M, Ishii N, Honda H: Detection of masses and microcalcifications of breast cancer on digital mammograms: comparison among hard-copy film, 3-megapixel liquid crystal display (LCD) monitors and 5-megapixel LCD monitors: an observer performance study. Eur Radiol 17:1365–1371, 2007
8. Obenauer S, Hermann K-P, Marten K, Luftner-Nagel S, Heyden D, Skaane P, Grabbe E: Soft Copy versus Hard Copy Reading in Digital Mammography. J Digit Imaging 16:341–344, 2003
9. Powell K, Obuchowski N, Chilcote W, Barry M, Ganobcik S, Cardenosa G: Film-screen versus digitized mammography: assessment of clinical equivalence. AJR 173:889–894, 1999
10. Gitlin J, Narayan A, Mitchell C, Akmal A, Eisner D, Peterson L, Nie D, McClintock T: A Comparative Study of Conventional Mammography Film Interpretations with Soft Copy Readings of the Same Examinations. J Digit Imaging 20:42–52, 2007
11. Pisano ED, Cole EB, Kistner EO, Muller KE, Hemminger BM, Brown ML, Johnston RE, Kuzmiak CM, Braeuning MP, Freimanis RI, Soo MS, Baker JA, Walsh R: Interpretation of digital mammograms: Comparison of speed and accuracy of soft-copy versus printed-film display. Radiology 223:483–488, 2002
12. Hwang SA, Seo JB, Choi BK, Do KH, Ko SM, Lee SH, Lee JS, Song JW, Song KS, Lim TH: Liquid-crystal display monitors and cathode-ray tube monitors: A comparison of observer performance in the detection of small solitary pulmonary nodules. Korean J Radiol 4: 153–156, 2003
13. Park CM, Lee HJ, Goo JM, Han DH, Kim JH, Lim KY, Kim SH, Kang JJ, Kim KG, Lee CH, Chun EJ, Im JG: Comparison of observer performance on soft-copy reading of digital chest radiographs: High resolution liquid-crystal display monitors versus cathode-ray tube monitors. Eur J Radiol 66:13–18, 2008
14. Langer S, Fetterly K, Mandrekar J, Harmsen S, Bartholmai B, Patton C, Bishop A, McCannel C: ROC Study of Four LCD Displays Under Typical Medical Center Lighting Conditions. J Digit Imaging 19:30–40, 2006
15. Geijer H, Geijer M, Forsberg L, Kheddache S, Sund P: Comparison of Color LCD and Medical-grade Monochrome LCD Displays in Diagnostic Radiology. J Digit Imaging 20:114–121, 2007
16. ACR: American College of Radiology Breast Imaging Reporting and Data System Atlas (BI-RADS Atlas), Reston, Va: ACR, 2003
17. Salazar AJ, Romero J, Bernal O, Moreno A, Velasco S, Díaz X: Effects of the DICOM grayscale standard display function on the accuracy of medical-grade grayscale and consumer-grade color displays for telemammography screening. Proc. IX International Seminar on Medical Information Processing and Analysis: Mexico, November 19, 8922:89220R-89220R-89210, 2013
18. Salazar AJ, Camacho JC, Aguirre DA: Comparison between different cost devices for digital capture of X-ray films with Computed Tomography (CT) correlation. Telemed J Health 14:275–282, 2011
19. Salazar AJ, Camacho JC, Aguirre DA: Agreement and reading-time assessment of differently priced devices for digital capture of X-ray films. J Telemed Telecare 18:82–85, 2012

20. Salazar AJ, Camacho JC, Aguirre DA: Comparison between different cost devices for digital capture of X-ray films: an image characteristics detection approach. J Digit Imaging 25:91–100, 2012. doi:10.1007/s10278-011-9391-0
21. Salazar AJ, Camacho JC, Aguirre DA, Ocampo J, Diaz XA: Clinical Diagnostic Accuracy on chest X-ray of the DICOM Grayscale Standard Display Function in Medical-grade Grayscale and Consumer-grade Color Displays. AJR In press, 2014
22. Salazar AJ, Camacho JC, Aguirre DA, Ocampo J, Diaz XA: Diagnostic accuracy of digitized chest X-rays using consumer-grade color displays for low-cost teleradiology services. Telemed J Health 20:304–311, 2013
23. Salazar A, Aguirre D, Ocampo J, Camacho J, Díaz X: 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 1–7, 2013
24. Lewin JM, D'Orsi CJ, Hendrick RE, Moss LJ, Isaacs PK, Karellas A, Cutter GR: Clinical Comparison of Full-Field Digital Mammography and Screen-Film Mammography for Detection of Breast Cancer. AJR 179:671–677, 2002
25. Obuchowski NA: Sample Size Tables For Receiver Operating Characteristic Studies. AJR 175:603–608, 2000
26. Dorfman DD, Berbaum KS, Metz CE: Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method. Investig Radiol 27:723–731, 1992
27. Roe CA, Metz CE: Dorfman-Berbaum-Metz method for statistical analysis of multireader, multimodality receiver operating characteristic data: validation with computer simulation. Acad Radiol 4:298–303, 1997
28. Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics 33:159–174, 1977
29. Chen Y, Gale A: Mammographic Interpretation Training: How Could Low Cost Display Devices Help? Springer Berlin, Heidelberg, 2010
30. Krupinski EA: Visual search of mammographic images: Influence of lesion subtlety. Acad Radiol 12:965–969, 2005
31. Kimpe T, Tuytschaever T: Increasing the Number of Gray Shades in Medical Display Systems—How Much is Enough? J Digit Imaging 20:422–432, 2007
32. Fetterly K, Blume H, Flynn M, Samei E: Introduction to Grayscale Calibration and Related Aspects of Medical Imaging Grade Liquid Crystal Displays. J Digit Imaging 21:193–207, 2008