

AD _____

GRANT NUMBER DAMD17-96-1-6316

TITLE: Routine Quality Assurance for Whole Breast Digital Mammography

PRINCIPAL INVESTIGATOR: Carolyn Kimme-Smith, Ph.D.

CONTRACTING ORGANIZATION: University of California
Los Angeles, CA 90095-1406

REPORT DATE: September 1998

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE September 1998	3. REPORT TYPE AND DATES COVERED Final (1 Sep 96 - 31 Aug 98)	
4. TITLE AND SUBTITLE Routine Quality Assurance for Whole Breast Digital Mammography			5. FUNDING NUMBERS DAMD17-96-1-6316	
6. AUTHOR(S) Carolyn Kimme-Smith, Ph.D.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of California Los Angeles, California 90095-1406			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Commander U.S. Army Medical Research and Materiel Command Fort Detrick, Frederick, Maryland 21702-5012			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES			19981214 104	
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200) <p>At the present time, three whole breast digital mammography units are undergoing clinical trials for FDA 510 (k) approval. During the two years of this grant period, the TREX system has been tested extensively over an 18-month period with a phantom designed for digital use. The GE and Fischer units have been tested with similar phantoms. A semi-automatic analysis program has been written in C++. The program can be used by a trained technologist for weekly quality control. For the TREX application, all 12 CCD's are tested independently for signal to noise ratio, contrast to noise ratio, calcification visibility, and dynamic range. Mean, standard deviation, and coefficient of variation are reported across all CCD's and over time. For the GE digital unit, a similar phantom and analysis program is used because of variations across the amorphous silicon. For the Fischer unit, the 4 CCD's in the slot are tested.</p> <p>As a result of this research, there have been three presentations at national meetings, a peer reviewed article has been accepted for publication, and a poster was presented at SPIE in San Diego. A simple method to monitor digital mammography quality control has been devised and tested. Modifications in the phantom and analysis program for specific systems has been shown to be feasible and all current manufacturers or the physicists at clinical sites have been supplied with a phantom and analysis program for their systems.</p>				
14. SUBJECT TERMS Breast Cancer, digital mammography, quality control			15. NUMBER OF PAGES 30	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

___ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

___ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

___ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

___ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

___ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Carolyn Kemme-Smith *Sept 23, 1998*
PI - Signature Date

TABLE OF CONTENTS

Foreword	1
Table of Contents	2
Introduction	3
Body	3
Conclusions	4
References	5
Appendix	6
Medical physics preprint	6

FINAL REPORT OF "ROUTINE QUALITY ASSURANCE FOR WHOLE BREAST DIGITAL MAMMOGRAPHY"

Introduction

Many Medical Physicists believe that the recent improvements in the death rates from breast cancer are in part the result of mammography quality control started by the ACR nine years ago. Combined with the mammography Quality Standards Act (MQSA) and the wide use of uniform diagnostic codes (BIRADSTM), fewer missed cancers and fewer "lost" patients result. Facilities currently testing whole breast digital image receptors have been concerned about maintaining these high quality standards for the new equipment. However, screen/film tests are not always appropriate for these digital systems, and optical density is no longer a criterion for correct exposure. More sophisticated testing is needed, and some test configurations must be modified for different systems, so that technologist testing may not be possible. This research concerns developing quality control tests for this new equipment, evaluating it, and modifying it for all the units currently being clinically tested.

Body

Experimental methods, as outlined in the statement of work, were delayed six months because the TREX system was not delivered to UCLA until April, 1997. As a result, two graduate students were employed each summer (instead of the budgeted one student) in order to prepare the computer programs needed. Thus, the following tasks were performed.

TASK 1. Establish an exposure technique chart, test cadaveric breasts and design a phantom with the minimum and maximum sized targets needed. Because UCLA had not received the TREX unit, Dr. Kimme-Smith traveled to GE at Schenectady, NY, to Bennett (TREX) in Copiague, NY, to the University of Virginia where a second TREX unit was in clinical trials, and to Fischer Imaging in Denver, CO. At each facility, unfixed cadaveric breasts with 150 to 200 μm simulated calcification clusters were imaged at a variety of techniques. As a result of these tests, we recommended:

1. That the screen/film uniformity test be adopted for digital image receptors. Since white fields are required for all three units, the plastic used to form the white fields can be imaged to determine if any fixed artifacts are present. As a result of signal to noise measurements made on these "flood" images, we developed the following phantom.

2. A phantom was developed that covered the entire image receptor and originally had the same type of simulated calcifications that were used for the cadaveric breast studies. These were later changed to the aluminum oxide crystals used in the ACR mammography phantom. In addition, aluminum step wedges and a 20 lp/mm gold phantom were arranged on interfaces between CCD's (Figure 1, from Bibliographic reference 5). Very low contrast pads were placed on each CCD position so that contrast to noise ratios (CNR) and signal to noise ratios (SNR) could be calculated reproducibly. The contrast in the pads was designed to be barely discernible in screen/film images of the phantom.

TASK 2. Write a C language QA program. This was done during the spring of the first year. The code used was IDL, rather than C++, in order to get the program up and running on our mammography workstation so that phantom measurements could be evaluated immediately after acquiring the digital images. This program was rewritten in C++ by the summer of the first year of the grant.

TASK 3. The QA program was used every day for six weeks and thereafter weekly. No significant change occurred until April, 1998, when images were found to have fixed

pattern noise on the "flood" film and low CNR, SNR on phantom imaging. This problem has occurred intermittently since April so that we now acquire a flood film on a daily basis before acquiring clinical images. This has allowed us to perform task 4.

TASK 4. Compare QA results with equipment failure. Until April, 1998, the TREX failures were entirely in the software and manifest themselves by freezing the system so that no images could be acquired. The present intermittent failures show that the phantom images predict hardware failure reliably.

TASK 5. Adopt the phantom for other digital mammography systems. The GE system at the University of Pennsylvania was tested this spring, while the GE system at the University of Colorado was not tested until July of this year. Both trips allowed us to modify the dimensions of the phantom so it was entirely contained on the smaller DMR image receptor. In May, the Fischer unit at the University of California, San Francisco, imaged another phantom similar in philosophy to the GE and TREX phantoms, but consisting of a strip covering the CCD slot that stretched from the chest wall to beyond the nipple of the breast (when a breast rather than a phantom was imaged).

TASK 6. Revise and test the QA analysis program for the other digital mammography units. This summer, the program was modified for the GE and Fischer systems. In particular, they both reversed grey level values from those of the TREX system. Therefore grey level was reversed to match the TREX custom. In addition, the GE system retained 12 bits while the TREX system had 14 bits. Image sizes also differed, since pixel size in each unit varied.

TASK 7. Because of confidentiality, a peer reviewed manuscript comparing the results for all three systems is not possible. However, the results of using our QA system on the TREX system has been described and accepted for publication by Medical Physics; a preprint of this paper is included in the appendix.

Conclusions

Three TREX phantoms, two GE phantoms, and one Fischer phantom have been distributed along with their analysis programs. They are simple to make from available, off the shelf items, so more can be fabricated inexpensively if users find them helpful. Technologists at UCLA have had no problem using the analysis program, although technologists at other facilities have not used it. It is particularly useful in calibrating one manufacturer's unit against a similar unit at another site. Values of SNR and CNR vary between manufacturers' digital units, as would be expected from differences in their designs. All three units vary across the image receptor from right to left and from chest wall to nipple. While the latter can be attributed to the heel effect, the former attests to the need to test the entire image receptor surface. Recent problems with the TREX system have shown that the phantom is affected quantitatively when noise increases in the digital receptor.

The phantom is also useful at ensuring a minimum standard in digital systems. The calcifications are the same size as are found in the third and fourth speck groups of the ACR mammography accreditation phantom. They are therefore a uniform size in each CCD grouping so that resolution change across the CCD's can be identified. The phantom can also be used with additional acrylic to determine a technique table that will ensure similar contrast and resolution over a range of breast sizes.

This grant has begun the quality control research that will be needed to establish MQSA guidelines for digital mammography. It is encouraging that simple measurements, partially automated, can be performed by a QA technologist in less than twenty minutes each week. Refinements in the phantom and analysis program will require the use of this QA method by a variety of facilities who operate different digital systems.

Bibliography of publications and meeting abstracts

1. Kimme-Smith C, Williams MB, Terwilliger BS, Fajardo LL, Bassett LW. Clinical quality control for whole-breast digital mammography. *Radiology* 1996; 102(P):369-370
2. Kimme-Smith C, Yang L, Arellano A, Beifuss M. Clinical QA measurements of prototype whole breast digital mammography equipment. 1997 AAPM National Meeting, Abstract in *Medical Physics* 1997; 24:1017-1018
3. Lewis CD, Kimme-Smith C, Beifuss M, Yang L, Bassett LW. Quality control and correct exposure for a whole breast digital mammography system. *SPIE* 1998; Volume 3336 poster and transaction report.
4. Kimme-Smith C, Lewis C, Yang L, Bassett W. An automated quality control program for whole breast digital image receptors. Fourth International Workshop on Digital Mammography. Nijmegen, 1998.
5. Kimme-Smith C, Lewis C, Williams MB, Bassett LW. Establishing minimum performance standards, calibration intervals, and optimal exposure values for a whole breast digital mammography unit. *Medical Physics*, 1998, in press.

Personnel receiving pay:

1. C Kimme-Smith Ph.D., PI
2. Graduate Students
 - a) Alonso Arellano MS, programmer
 - b) Limin Yang, programmer
 - c) Candace Lewis, medical physicist
 - d) David MacElroy, medical physicist

**Establishing Minimum Performance Standards, Calibration
Intervals, and Optimal Exposure Values for a Whole Breast
Digital Mammography Unit**

C. Kimme-Smith ¹

Candace Lewis ¹

Manuel Beifuss ²

Mark B. Williams ³

L.W. Bassett ¹

¹Iris Cantor Center for Breast Imaging, University of California, Los Angeles CA

²FH Aachen, Abt Juelich, Germany

³University of Virginia, Department of Radiology, Charlottesville VA

ABSTRACT

Methods are developed to establish minimum performance standards, calibration intervals, and criteria for exposure control for a whole breast digital mammography system. A prototype phantom was designed, and an automatic method programmed, to analyze CNR, resolution, and dynamic range between CCD components in the image receptor and over time. The phantom was imaged over a 5 month period and the results are analyzed to predict future performance. White field recalibration was analyzed by subtracting white fields obtained at different intervals. Exposure effects were compared by imaging the prototype phantom at different kVp, filtration (Mo vs. Rh) and mAs. Calcification detection tests showed that phantom images, obtained at 28 kVp with a Mo/Mo anode/filter and low mAs technique, often could not depict Al_2O_3 specks 0.24 mm in diameter, while a 28 kVp Mo/Rh, higher mAs technique usually could. Stability of the system tested suggests that monthly phantom imaging may suffice. Differences between CCD performance are greater (12%) than differences in a single CCD over time (6%). White field recalibration is needed weekly because of pixel variations in sensitivity which occur if longer intervals between recalibration occur. When mean glandular dose is matched, Rh filtration gives better phantom performance at 28 kVp than Mo filtration at 26 kVp and is recommended for clinical exposures. An aluminum step wedge shows markedly increased dynamic range when exit exposure is increased by using a higher energy spectrum beam. Phantoms for digital mammography units should cover the entire image receptor, should test intersections between components of the receptor, and should be automatically analyzed.

Key words: mammography, digital, phantoms, test objects, quality control

INTRODUCTION

The promise of commercially available whole breast digital mammography has fueled many research projects during the last five years. This research has provided design tools for manufacturers of digital mammography equipment and has encouraged medical physicists to predict preliminary criteria for its performance (1-4). Digital stereotactic breast biopsy equipment calibration procedures have caused medical physicists to realize that sufficient exposure is needed for this equipment to produce an acceptable contrast to noise ratio (CNR). The CNR, rather than optical density, ensures good calcification detection (5,6).

Because digital mammography display contrast can be increased after image acquisition, contrast is only limited by noise and subject contrast in the image. For this reason, it has been predicted that beam spectra and kVp can be higher than those used for screen-film mammography (7). These values will depend on the conspicuity of masses in their surround, and will be expected to vary with breast composition and compressed breast thickness (8). Experiments with contrast/detail phantoms (9) and phantoms designed specifically for digital mammography (3) have shown the dependence of image performance, contrast to noise, and calcification visibility with exposure.

Current clinical trials of whole breast digital mammography equipment have modeled quality control procedures on Mammography Quality Standards Act guidelines (10,11). The FDA is encouraging QA of these digital systems in a document entitled, "Information for Manufacturers Seeking Marketing Clearance of Digital Mammography Systems" (12). Clinical exposures in these trials usually produce the same mean glandular dose required for screen-film, although beam spectra is sometimes harder and/or kVp higher, depending on the manufacturer. These preliminary standards will be replaced with standards developed as a result of clinical experience with digital mammography equipment. At our institution, we have tested a variety of quality control procedures on a TREX digital mammography unit over a five month period. During this time, we have

developed methods for recommending quality control pass-fail criteria, frequency of testing and calibration, and criteria for exposure control which may be useful guidelines for other whole breast digital systems.

METHODS

The TREX (Danbury CT) whole breast digital mammography system consists of an array of 12 CCDs arranged in a 3 x 4 grid to produce an array 19.2 x 25.6 cm in size with 4800 x 6400 14-bit pixels. Each pixel is 40 μm at the image plane. Each CCD pixel is in contact with a 2:1 fiber optic taper which connects it to the 300 μm thick CsI screen. The CsI is also in discrete packages whose boundaries are not aligned with CCD boundaries (Figure 1). The TREX system has a permanently installed reciprocating grid with a 5:1 grid ratio and 35 strips per cm. The digital image receptor is mounted on a LORAD (Danbury, CT) M4 mammography unit supplying either Mo or Rh filters to a Mo anode. The non-digital portion of this unit was calibrated according to MQSA regulations.

Each mammography view is displayed on an acquisition workstation, archived to an optical disk, and transmitted to PACS storage. A diagnostic workstation is currently being clinically tested; routine clinical diagnosis is performed from films produced by a prototype Agfa Scopix 5200 laser printer with a 40 μ pixel size (Ridgefield Park, NJ) so that full resolution for a 19 x 25 cm image is possible. The diagnostic workstation also supports a quality control program designed at UCLA for the TREX system. Here the 2K by 2.5K monitor can display full resolution images of individual CCDs.

Because of the 12 independent CCDs which make up the digital image, a phantom was designed with 12 modules so that each CCD could be tested independently. The 12 CCDs are numbered as shown in Figure 2. Each module consists of a low contrast 1 cm^2 pad, formed from 3 thicknesses of gummed paper, and eight aluminum oxide calcifications, four of which are 0.32 mm in diameter and four are 0.24 mm in diameter (the size of the third and fourth speck groups of the American College of Radiology Mammography Accreditation phantom). In addition, a 16 lp/mm gold phantom (CIRS, Norfolk VA), with single increments from 9 to 16 lp/mm, lies on the intersection between

two CCDs, and four aluminum wedges (14 steps, each 400 μm high) lie on intersections between CCDs. The line pair phantom and step wedges are primarily for acceptance testing the digital unit, while the contrast pads and calcifications calibrate routine quality control. Two exposures were made whenever the phantom was imaged: the Mo anode/Mo filter (Mo/Mo) imaged the phantom at 28 kVp, 65 mAs, to yield a mean glandular dose of 186 mrad, and a Mo anode/Rh filter (Mo/Rh) exposure was made at 28 kVp, 85 mAs (the mR/mAs is lower with this anode/filter combination) to give a MGD of 177 mrad. Exit exposures (i.e., exposure to the grid) were measured with the phantom elevated above the ion chamber with a 10 cm air gap between the two objects. The Mo/Mo exit exposure was 28 mR while the Mo/Rh exit exposure was 33 mR, corresponding to the different absorption in tissue of the different energy spectra.

A computer program calculates quality measures from each contrast pad. The calculations found most useful are the mean divided by the standard deviation within a 50 pixel on a side square in the pad, as well as a contrast to noise ratio (CNR) consisting of the contrast of the pad divided by the standard deviation (noise) of the background in quadrature:

$$\text{CNR} = \frac{|\mu_{\text{in}} - 1/2 (\mu_{\text{out}_1} + \mu_{\text{out}_2})|}{\sqrt{\sigma_{\text{in}}^2 + (\sigma_{\text{out}_1}^2 + \sigma_{\text{out}_2}^2) / 2}}$$

with μ representing the mean and σ the pixel to pixel standard deviation and where the subscript "in" references data collected from the 50 x 50 pixels within the contrast pad and "out 1" and "out 2" reference data collected from two 50 x 50 pixel regions located adjacent to the contrast pad. Because manufacturers use μ/σ for a region as a measure of CNR, this is also calculated. In addition, the region of the fourth CCD module contains two additional contrast pads so that intra-CCD variation can be studied. Twelve groups of 8 specks each

are also magnified to full resolution by the analysis computer program so that they can be visually graded. The targets are fixed on a 3 mm thick plate of acrylic which is placed on 3.6 cm of acrylic. Care must be taken to position the phantom precisely using the light field of the unit so that wedges and line pair phantoms are aligned with CCD boundaries. With extreme window and leveling, individual CCDs can be distinguished from their neighbors. The thickness of the added acrylic is the same as that of the acrylic block supplied by the manufacturer to calibrate and correct CCD sensitivity variations.

In order to recalibrate CCD sensitivity, this 3.6 cm thick sheet of acrylic is imaged eight times once a week and a flood field (white field) is formed from the pixel by pixel average of the 8 samples. This white field is then used to normalize CCD response for all digital images acquired by the system. Two of these calibration files are acquired: one for Mo/Mo exposures and one for Mo/Rh exposures. Just as Nuclear Medicine gamma cameras are periodically recalibrated to normalize the uniformity of sensitivity in different detectors, digital systems must be periodically recalibrated to normalize individual CCD pixel sensitivity. The interval needed between this recalibration has not been established for whole breast digital mammography systems, although for stereotactic digital breast biopsy systems, different intervals are recommended by the two manufacturers. For the Fischer (Denver, CO) digital stereotactic biopsy unit, daily white field calibration is common, while the LORAD digital stereotactic unit usually requires biannual recalibration. The TREX system is currently recalibrated weekly according to the manufacturer's recommendations, but TREX has encouraged experiments to determine if this period can be extended. To determine if the weekly period is necessary, the white field acquired when the unit was acceptance tested was subtracted pixel by pixel from similar white fields obtained on the same day, at one week, two week, three week, six week, two month and three month intervals. The resulting differences were accumulated in a histogram; the mean, standard deviation, and maximum difference between pixels were tabulated. Because the white field consists of 6 linear strips on the CCD array (in the anode/cathode direction), corresponding

to different electronic partitions of the CCD data, comparisons were also made between these partitions to determine if a particular part of the image receptor was more subject to changes in sensitivity than other parts.

Finally, to determine the effects of energy spectrum and photon flux on digital receptor performance, a series of exposures at 28-34 kVp, with Mo and Rh filtration, were made with matching exit exposures. The performance of equipment to faithfully depict the phantom was compared for these techniques in order to make recommendations for future clinical trials.

RESULTS

Minimum performance standards.

At the present time, whole breast digital mammography performance is being compared to that of screen-film mammography. In order to justify the increased costs of these systems, improved performance should be expected once they are commercially available. Precursors of the three current digital mammography systems have been found to have superior performance to that of screen-film when images of phantoms were compared (3,9,13). In addition, the ACR accreditation procedure for digital stereotactic accreditation requires that these units be able to image the fourth speck group, rather than the third group, which is required for screen-film systems (14). Thus one minimum requirement for whole breast digital systems should be the ability to image aluminum oxide specks 0.24 mm in diameter. This standard varies less than the more quantitative performance criteria, of μ/σ and CNR, with varying technique (see "Calibration Intervals," below). Minimum subject contrast or contrast to noise is more difficult to assess. For the TREX system, using the prototype phantom, μ/σ varied from 43 to 90 depending on kVp, filtration, and photon fluence. Contrast to noise ratio, defined earlier, varied from 0.4 to 2.5 for the low contrast pads employed on our phantom. Note that for the TREX system, a high photon fluence yields a low gray level value, so for the exposure results reported here, all gray level values have been reversed so that high photon fluence corresponds to a high gray level value.

The line pair phantom performance represents the MTF at high contrast values (4% amplitude response or .04 MTF) and so does not reflect the performance at the 50% amplitude response or 0.5 MTF level where, because of the lower contrast afforded by glandular tissue, calcification detection occurs. As such, it is not as relevant a criterion for clinical performance as the aluminum oxide specks. The line pair phantom was consistently 9 lp/mm when magnified by resting on 3.9 cm of acrylic. When placed on the breast support system, between 10 and 11 lp/mm result.

Dynamic range, measured with the four step wedges shared by eight CCDs, is also dependent on exposure parameters. The number of adjacent steps visible, after window and level adjustments to increase contrast, represent the range between noise, where photon fluence is too random to visualize step differences, and between saturation of the CCD at a particular step. At the saturation point, no difference in pixel value between adjacent steps was observed. At the noise level, steps could not be distinguished from adjacent steps when the average pixel difference between these adjacent steps fell below 50 pixel values. Therefore, this was taken as the criterion for the number of steps visible. This criterion would change for a system with fewer than 14 bits, since pixel range for visible step differences will depend on amplifier gain and thus on analog to digital conversion. As exit exposure increases, the number of step pairs whose average pixel value differed by more than 50 varied from 6 to 12. Using this measure, dynamic range varied between CCD's by as much as two steps. Because the step wedge was linear rather than logarithmic, a measure of the linearity of the system required comparing the logarithm of the step size to the average pixel value within that step. By this measure, the system was not linear.

On acceptance testing, discontinuities at the CCD boundaries can best be measured by a fluoroscopy or mammography mesh placed diagonally across the intersections (Figure 3). As in Figure 1, assessment of the variations between CCDs requires higher contrast windowing than would typically be used for clinical images. In the prototype phantom, straight edges perpendicular to the CCD boundaries, such as the line pair phantom and step wedges (Figure 2), do not assess these discontinuities as well as a diagonally placed mesh image because the interpolation algorithm includes pixels perpendicular to the boundary and thus tends to smooth the artifact successfully when tested with targets aligned perpendicular to the boundary.

Calibration Intervals

For phantom imaging, the number of high contrast line pairs visible did not change over the five month period, while the number of perceived wedge steps rarely changed by even one step (or 50 gray levels). Variations between CNR measurements made on the same CCD exposed repeatedly using the same x-ray technique were less than .7%, while variations between different CCDs, which were exposed once, varied up to 12.4%. Phantom measurements were made daily for 6 weeks, and then three times a week for the remaining time. The largest variation for one of the CCDs was 6.3% over the five months of testing. Clearly, variations between CCDs was greater than variations in one CCD over time.

The CNRs, averaged over the 12 CCD values on any given day, are likely to vary from the preceding test day's values from 3 to 7% (see Table 1). While the coefficient of variation of the 12 CCDs on any given test day varies from 8 to 12%; CCDs which contribute to this variation recover their previous values on subsequent days. Over the 6 month period that the phantom was tested, CNR for individual CCDs had a coefficient of variation of 4% to 6.3% for the Mo/Mo images and 3.7% to 7.6% for Mo/Rh images (see Table 2). If CNR is used as a criterion for exposure control, similar to the way that optical density is used in screen/film automatic exposure control, variations in CNR between CCDs will need to meet minimum standards of uniformity to ensure consistent exposures across the image receptor, similar to digital receptor uniformity criterion described in ACR stereotactic biopsy recommendations (14). Variations over time were much less than variations between CCDs (Table 1). These variations may be partly due to the heel effect which varies 19% for Mo filtration and 14% for Rh filtration. While deterministic variations in photon flux are normalized by dividing each pixel by the matching white field pixel, photon noise is likely to be higher when photon flux is less. CCDs in the last row, closest to the anode end of the x-ray tube, are most affected by this phenomena. Thus CCDs 9, 10, 11, 12 have consistently lower CNR values (Table 1). However, variations

in the first row, that is, CCDs numbered from 1 to 4 (see Table 1) vary 21%. Since ion chamber measurements show that there is equal exposure to these two CCDs, difference in sensitivity between them must be assumed. Because the position of the contrast pad and the "outside" regions of interest where noise is calculated are fixed, variations due to the position of each 50 x 50 pixel region on the CCD are not responsible for these differences.

Calcification speck count varied by up to 4 specks over time, indicating that automatic counting rather than visibility should be used for this target (15). The 0.24 mm specks were less sharp for the Mo/Mo exposure compared to images made with the Rh filter. The phantom images made with the Rh filter always revealed from 1 to 4 of these smaller specks, while they were less visible in at least one CCD for 8 out of 12 exposures for Mo filter images (Figure 3). If we use the criterion that the 0.24 mm specks must be visible for digital imaging to ensure that its quality is equal to that of stereotactic breast biopsy imaging, then phantom imaging would fail more often with Mo filtration unless mean glandular dose or kVp was increased. Because MQSA will require weekly phantom imaging, digital mammography will be adequately monitored at this interval, although the image receptor tested in these experiments showed such stability that monthly testing would be more appropriate.

Because equipment failure in the whole breast digital system during the five month test period was usually computer related so that digital images were not formed during a failure, testing intervals for phantom imaging cannot be recommended based on our experience. The interval for testing during the last three and a half months has been three times a week. If testing occurred weekly, variations would be nearly identical.

White field calibration interval was compared by subtracting white fields obtained at different time intervals. To ensure that the six partitions of the white field did not vary in their performance from each other, each was analyzed independently for one week to three week intervals. The two partitions which varied most are documented in Table 3 for the Mo filter white field. The behavior of the Rh filter white field is similar. When all 6 partitions

of the white field are analyzed together over a longer interval, Table 4 reports the resulting statistics. Note that at a one week interval, mean and S.D. are similar to recalibration on the same day, but that for longer intervals these factors increase up to 60%. In addition, increasing pixel drop-out, similar to screen/film pick-off, occurs if recalibration is not performed weekly. Pixel drop-out represents pixels with lower sensitivity which, because of gray scale mapping algorithms, appear white. Once a new white field is formed, these pixels return to normal gray levels, since their lower sensitivity is corrected.

The most interesting results occurred when different techniques with varying energy spectra and photon fluence were compared using the prototype phantom. Table 5 reports the techniques tested, mean glandular dose, exit exposures, and phantom scores which resulted. While at 28 kVp, Rh filtered images have a slight advantage over Mo filtered ones, this changes as kVp increases. By 34 kVp, for both Mo and Rh filtration, calcification detection decreases. Note that μ/σ continues to increase with kVp, as does dynamic range, while CNR (which in our definition depends on contrast) decreases.

CONCLUSIONS

While the heel effect causes screen/film optical density to vary in the anode/cathode direction, variations across the screen/film cassette parallel to the chest wall are unusual. Variations in CNR across the digital image receptor (parallel to the chest wall) vary significantly, so the whole image receptor should be tested. Similarly, since CCDs may differ in resolution, the image receptor should be tested for calcification detection on each CCD. For digital systems employing discrete partitions, such as separate CsI elements, multiple CCDs, or white fields organized by electronic partitions, phantoms should test these partitions and their boundaries.

As was discovered when testing digital stereotactic systems, geometrical distortions can be evaluated with the mammography screen/film contact test mesh placed obliquely on the breast support platform with 4 cm of acrylic (5) (Figure 4). Because of increasing changes in CCD pixel sensitivity over time, weekly white field recalibration is recommended. Increased photon flux will improve digital mammography performance for these early prototype systems. These increases are still within MQSA dose recommendations, except in California, where a 2.00 mGray limitation would require the use of Mo/Rh anode/filter combinations at the higher mAs exposures. Medical physicists may wish to measure exit exposures for these systems in order to calibrate optimal exposures for varying sized breasts. Preserving subject contrast by limiting the kVp and filtration is apparently not as necessary for digital systems as it is for screen/film systems. However, contrast enhancement is limited by noise and image receptor artifacts so that kVp above 30 is not recommended.

MQSA regulations for digital mammography must be based on realistic tests of the equipment and not on historical criterion established for screen/film systems.

Acknowledgments

This work was supported by a U.S. Army Breast Cancer Research Grant # DAMD
17-96-1-6316.

Table 1. CNR for 12 CCDs over one week for Mo/Mo and Mo/Rh exposures.

CCD	1	2	3	4	5	6	7	8	9	10	11	12	μ	σ^*	cov**	
Mo/Mo																
Aug 6	1.18	1.08	1.08	0.97	1.03	0.98	1.02	1.02	0.88	0.84	1.00	0.89	0.998	.095	9.6%	
Aug 8	1.27	1.13	1.11	1.00	1.06	1.00	1.07	0.94	0.93	0.92	1.05	0.86	1.033	.126	12.2%	
Aug 11	1.31	1.06	1.12	1.08	1.01	0.99	1.07	1.05	0.84	0.89	1.11	0.89	1.035	.126	12.2%	
Mo/Rh																
Aug 6	1.51	1.32	1.23	1.31	1.18	1.24	1.01	1.17	1.10	1.02	1.00	1.05	1.178	.15	13%	
Aug 8	1.40	1.20	1.22	1.21	1.16	1.08	1.09	1.03	1.05	1.05	1.01	1.01	1.126	.12	10%	
Aug 11	1.42	1.28	1.32	1.31	1.22	1.16	1.21	1.06	1.12	1.11	1.16	1.14	1.209	.105	8.7%	

* σ = standard deviation

** cov = coefficient of variation = $\sigma/\mu * 100$

Table 2. Coefficient of variation for the 12 CCDs over a five month test period for CNR.

CCD:	1	2	3	4	5	6	7	8	9	10	11	12
Mo/Mo	4.81	4.37	4.85	3.99	5.46	4.43	4.25	4.81	5.55	5.60	6.27	5.34
Mo/Rh	4.37	5.92	3.95	3.69	6.42	7.48	5.24	3.80	5.67	7.63	6.71	6.70

Table 4. White field variations over a same day period to a three month period for a Mo/Mo anode/filter combination.

Interval	Mean difference	S.D.	Maximum difference
Same day	13.38	11.90	1973
1 week	12.95	11.78	3061
2 weeks	20.21	17.75	3256
3 weeks	20.21	17.75	3181
14 weeks	33.67	66.22	3362

Table 5. Effect of increasing energy spectra and photon fluence on mean glandular dose, calcification detection, μ/σ^{++} , dynamic range, and CNR in a 3.9 cm thick prototype phantom.

Mo Anode / Mo Filter							
kVp	mAs	MGD*	EE**	CNR***	μ/σ^+	Calcifications Detected++	Wedge Steps+++
28	60	157	17	.599	42.2	7	6
30	45	160	17	.780	45.2	6.8	6.1
32	32.5	143	17	1.249	45.2	6.4	7.1
34	26	139	16	1.866	47.3	6.2	8
Mo Anode / Rh Filter							
kVp	mAs	MGD*	EE**	CNR***	μ/σ^+	Calcifications Detected++	Wedge Steps+++
28	65	134	17	.99	44.3	7.3	6
30	45	119	16	.87	42.0	6.8	6.1
32	37.5	125	17	.90	44.7	6.5	6.9
34	30	122	16	.86	45.6	6.4	7

* MGD is mean glandular dose (in mGray).

** EE is exit exposure 5 cm above the grid (in mR).

*** CNR is the average contrast to noise ratio over all 12 CCDs.

+ μ/s is the average overall 12 CCDs of the mean pixel value ratio to the standard deviation between pixels in a 50 x 50 pixel area.

++ Calcifications detected are averaged over all 12 CCDs. Eight is the maximum present.

+++ Based on the number of 400 μm steps visible in an Al wedge, this represents Dynamic Range.

REFERENCES

1. Roehrig H, Yu T, Krupinski E. Image quality control for digital mammographic systems: Initial experience and outlook. *Journal of Digital Imaging* 1995; 8:52-66.
2. Henry J, Yaffe M, Turner T. Noise in hybrid photodiode array: CCD x-ray image detection for digital mammography. *SPIE* 1996; 2708:106-115.
3. Critten J, Emde K, Mawdsley G, Yaffe M. Digital mammography image correction and evaluation in digital mammography. Ed: Doi K, Giger M, Nishikawa R, Schmidt R. 1996, Elsevier Science.
4. Roehrig H, Fajardo L, Yu T, Schempp W. Signal noise and detective quantum efficiency in CCD-based x-ray imaging systems used in mammography. *SPIE* 1994; 2163:320-332.
5. Kimme-Smith C and Solberg T. Acceptance Testing Prone Stereotactic Breast Biopsy Units. *Medical Physics* 1994; 21:1197-1202.
6. Hendrick RE and Parker SH. Principles of Stereotactic Mammography and Quality Control. In: Percutaneous Breast Biopsy, editors: Parker SH and Jobe WE. Raven Press, NY 1993.
7. Fahrig R and Yaffe MJ. A model for optimization of spectral shape in digital mammography. *Medical Physics* 1994; 21:1463-1472.
8. Fahrig R and Yaffe MJ. Optimization of spectral shape in digital mammography: Dependence on anode material, breast thickness, and lesion type. *Medical Physics* 1994; 21:1473-1480.
9. Liu H, Fajardo LL, Baxter R. A Theoretical Model for Contrast-Detail Detectability Prediction in Digital Radiography. *SPIE* 1994; 2676:116-121.
10. Mammography Quality Standards Act of 1992. Public Law 102-539, October 27 1992, 102nd Congress.

11. Mammography Quality Control Manual, Revised Edition, 1994. American College of Radiology, American Cancer Society. ISBN1-55903-136-0.
12. Code of Federal Regulations 21CFR900.18. Alternative requirements for section 900.12 quality standards.
13. Siewerdsen JH, Antonuk LE, El-Mohri Y, Yorkston J, Huang W, Baudry JM, and Conningham IA. Empirical and theoretical investigation of the noise performance of indirect detection AMFPI's for diagnostic radiology. *Medical Physics* 1997; 24:71-89.
14. Stereotactic Breast Biopsy Quality Control Manual, 1998. American College of Radiology.
15. Chakraborty DP. Computer analysis of mammography phantom images (CAMPI): An application to the measurement of microcalcifications image quality of directly acquired images. *Medical Physics* 1997; 24:1269-1278.

Figure Captions

1. A defect in between two CsI screens covering one CCD is apparent in the white field image of the image receptor, a) before white field recalibration, b) after white field recalibration. Note that high contrast window and level were employed, particularly in image b, to show this defect, and that the artifact would not be visible at clinical window and level selection. Note also that the CsI partition boundary differs from the location of the boundary between CCDs, which is particularly visible in image b.
2. a) Schematic of 18 x 24 cm phantom designed to test all 12 CCDs for the TREX full field digital mammography unit. Multiple targets at the lower right hand side test intra-CCD performance. Numbers on each CCD correspond to results in Tables 1 and 2. b) Digital image showing phantom.
3. Magnified image of the 0.24 mm speck group in the phantom for a) Mo filtration, 28 kVp, 65 mAs and, b) Rh filtration, 28 kVp, 85 mAs.
4. Mesh used for mammography screen/film contact placed obliquely and imaged with 3.9 cm of added acrylic on the TREX system. Note that window and level are beyond clinical levels.

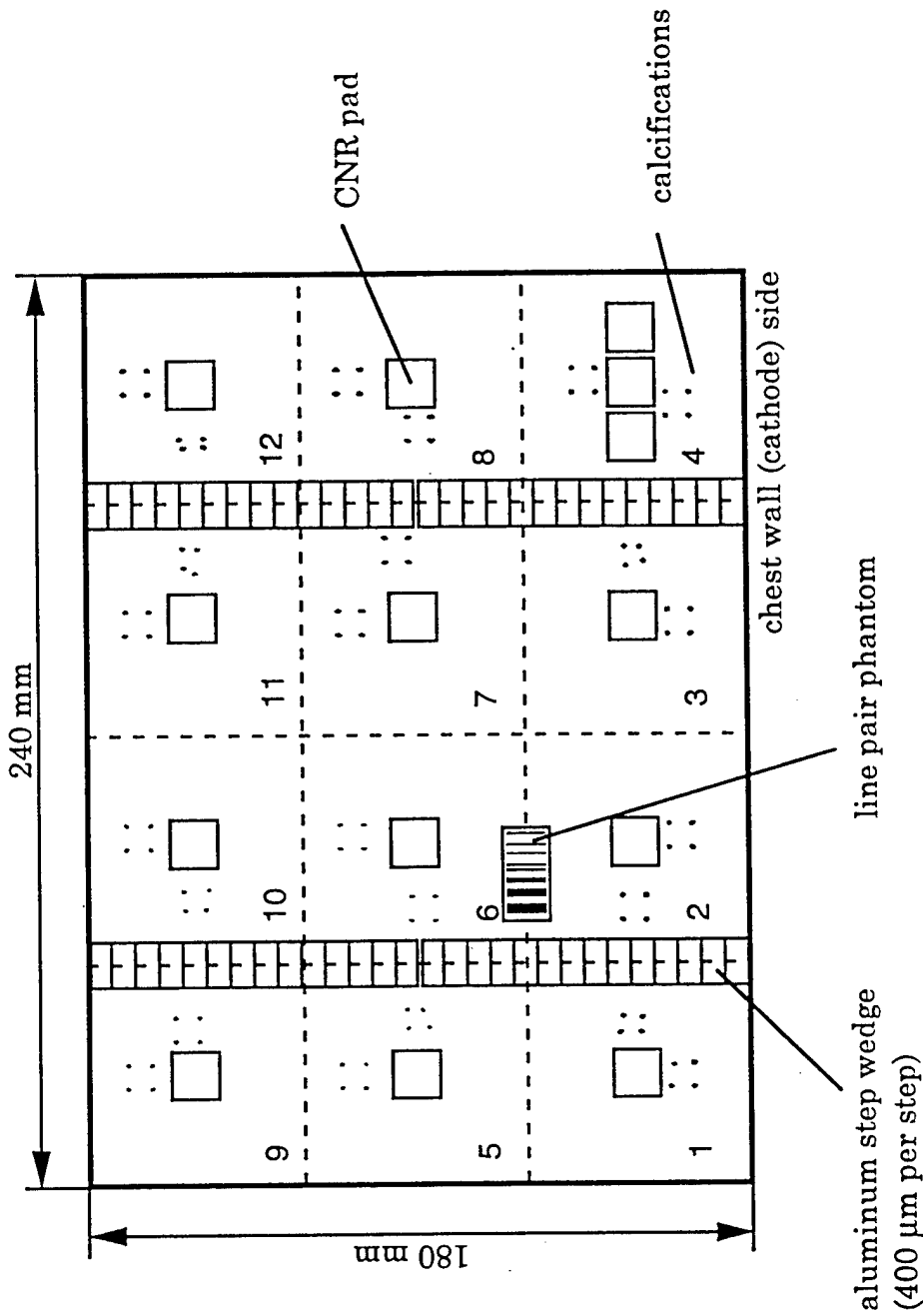


Figure 2a) Schematic of 18 x 24 cm phantom designed to test all 12 CCDs for the TREX full field digital mammography unit. Multiple targets at the lower right hand side (CCD #4) test intra-CCD performance.