

Softcopy Display Requirements for Digital Mammography

Bradley M. Hemminger

School of Information and Library Science, University of North Carolina at Chapel Hill

ABSTRACT

Purpose One of the advantages of digital mammography is to display mammograms on softcopy (electronic displays). Softcopy display of mammography is challenging because of the spatial and contrast resolution demands present in mammograms. We have designed and developed a softcopy mammography display application, Mammoview©, which is capable of allowing radiologists to read mammograms as quickly and as accurately as they can on film alternators. We review the studies using Mammoview to elucidate the requirements of a successful softcopy display station.

Methods The design and development of the Mammoview softcopy display station are described. Results of several studies using Mammoview are reported, including subjective feedback from RSNA demonstrations, and clinical studies measuring performance in terms of speed and accuracy. Additional analysis of user interactions, and user feedback is used to study the successes and shortcomings of mammography display stations like Mammoview.

Results and Conclusions Overall radiologist readings using Mammoview have been shown to be as fast and as accurate as readings using mammography film alternators. However, certain parts of the softcopy interface were more successful than their film counterparts, while other parts of Mammoview were less successful than their film counterparts. Data analysis of the recorded human-computer interactions for the softcopy component of the clinical trial indicate statistically significant correlations between the difference in review time of softcopy versus alternator readings and three factors: the number of interactions, the reader, and the size of the image being reviewed. The first (number of interactions) suggests that simpler interfaces require less time to use; the second, reader, supports previous findings, that radiologists vary in how fast they read screening mammography studies; the third, size of image, suggests that the speed of softcopy review is increased relative to film readings when images are significantly larger than the display size. Feedback from radiologists using the system in clinical trials and at demonstration exhibits at RSNA indicated good acceptance of the interface and easy adaptation. Radiologists indicated that they felt comfortable using the interface, and would use such a softcopy interface clinically. Finally, preliminary work suggests that the addition of a simple interaction to incorporate computer aided detection (CAD) results would be improve reading accuracy without significantly increasing reader times.

Keywords: Image Display, Human Computer Interaction, Digital Mammography, Image Processing, Image Presentation, Softcopy, Image Review, Pan, Zoom.

1. BACKGROUND AND SIGNIFICANCE

1.1 Digital mammography overview

Breast cancer is often quite similar in density to surrounding normal dense breast tissue. Up to 20% of breast cancers are mammographically occult.^{1,2} Digital mammography has the potential to improve breast cancer detection.³ A major limitation of film-screen mammography is the film itself. The film serves as the medium of image acquisition, storage and display, and thus cannot be optimized for each purpose independently. With digital systems the image acquisition, storage, and display are decoupled, and each system is optimized, yielding optimal overall performance. Digital detectors offer improved detection because of improved efficiency of absorption of the incident x-ray photons, a linear response over a wide-range of incident radiation intensities, and low system noise.^{4,5} In addition, once a film-screen mammogram is obtained, it cannot be significantly altered. Radiologists cannot manipulate the image directly. With film, improvements in image display involve acquiring more images with magnification or focal compression (and thus exposing the patient to more radiation), or looking at the images with a hot light and/or magnifying glass.

Digital acquisition systems directly quantify x-ray photons and de-couple the process of x-ray photon detection from image display. Digital images can be processed by a computer and displayed in multiple formats and processing conditions on an electronic display such as a cathode ray tube (CRT) monitor. Lesion conspicuity can be affected by these contrast manipulations. Image processing has been shown to improve visualization of details within medical images in at least one other application.⁶ Since the steps of image acquisition and display are separated, each can be optimized. In addition, image storage, transmission and retrieval can be improved. Computer aided detection algorithms can be utilized to improve the accuracy of the radiologists interpreting the images. Electronic markup of the images can facilitate image review, and the creation and dissemination of teaching materials.

In order to fully take advantage of these advances in digital mammography acquisition, image processing, and CAD, softcopy display of mammograms is required.⁷ The acceptance of digital mammography has been slow, in part due to the lack of capable softcopy display systems for mammography.

1.2 Digital mammography display

Digital mammograms can be printed to film or displayed on a monitor. Radiologists are less experienced with reading mammograms on monitors, and are more comfortable with printed film images viewed on a lightbox. Typically, current laser-printed films can display 4000x5000 pixels at 12 bits of grey scale. The disadvantages of film display for digital mammography are obvious. Once an image is printed, it can no longer be manipulated. The full information available in the digital data might not be evident in the printed image. Recent laboratory studies at UNC have shown that the optimal mammography image processing for presentation is dependent on the lesion type being detected (masses versus microcalcifications).⁸⁻¹² Printing multiple presentations of the same image data with different processing makes it too cumbersome for the radiologists to view all the images conveniently.

High luminance (100ftL+), high resolution monitors (2000x2560 pixels) are available for soft-copy interpretation.¹³ With currently available monitors (5Megapixels), only a portion of a breast imaged on a 50 micron full field digital detector (12 megapixels) can be displayed at one time at full resolution. Panning and zooming with the computer monitor, while possible, is not as natural for radiologists, and can be inefficient and time-consuming. In addition, comparing old and new and left and right images requires developing a new paradigm for viewing the images. In order for readings on monitors to take place, well-designed and carefully thought out computer-human interfaces are required to accomplish short, clinically acceptable review times for each case, including the display of previous studies for comparison. Without a usable computer-human interface to the display system, digital mammography cannot reach its full potential. A final concern is that images displayed on monitors must be processed to appear similar to the film appearance to which mammographers are accustomed, both for the digitized film images, and direct digital mammography images.

2. METHODS

2.1 Digital mammography display requirements

Mammoview was designed to provide a softcopy display application that would support reading of mammography exams as well as, or better than, film readings. It was specifically designed to accommodate screening mammography readings, which have a well defined structure, and a high throughput. The high throughput is necessary because of the large volume of screening mammograms that occur, and the low reimbursement rate. Screening mammograms tend to be read in batches, with the average time spent per case being about 1 minute, including dictation. The vast majority of screening cases are negative, and a negative case with no areas of interest may take only about 15 seconds to read and dictate. Thus, the workstation that supports digital mammography must be able to load studies nearly instantaneously, and all human-computer interactions on the workstation must also occur instantaneously. A screening mammography study consists of a new exam and an old exam (prior). Each exam in the US includes two projections of each breast for a total of 4 standard views for an imaging study: right cranial caudal (RCC), left cranial caudal (LCC), right mediolateral oblique (RMLO), and left mediolateral oblique (LMLO). Thus, a screening read with a prior and current exam consists of 8 images total. Figure 1 shows an example softcopy display with all 8 views depicted in a standard viewing format.

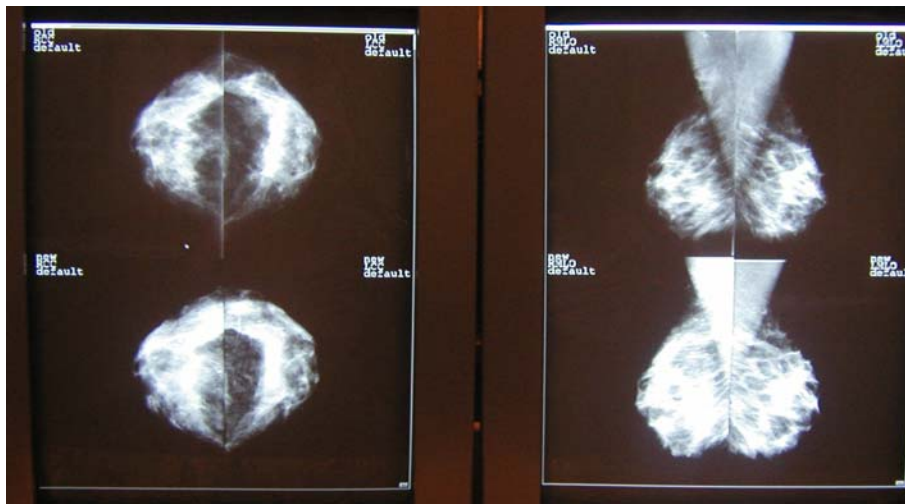


Figure 1. This shows a standard hanging protocol for a screening mammography study. The new study images are on the bottom row, and the prior study images are on the top row (identical positions). The CC views are on the left and the MLO views are on the right. The right views in this protocol are displayed on the left side, so for instance, the current RCC is the image in the bottommost left corner.

From interviews, videotaping and eye tracking¹⁴ we have ascertained that there are three primary comparisons made by radiologists in the US when viewing screening mammography studies on film. They compare the right and left views of the same projection (for example the new RCC compared with the new LCC); the new and old views of the same breast and same projection (for example, the new RMLO and old RMLO); and the two projections of the same breast for the same study (for example the new RCC and new RMLO). Figure 2 depicts these three comparisons on the reading layout shown in figure 1. Further, our analysis of radiologist's reading patterns also indicated that they initially looked at an overview (all the images) to get a gestalt, and then decided what, if anything, needs further investigation. Areas that warranted further study were viewed with a magnifying glass, comparing one image to another. How the radiologists proceeded with analyzing the images in more detail differed between radiologists; however, what they viewed was generally the overall set of images, and the pair-wise comparisons of the three comparison pairs described above. Another aspect they differed in was how they initially looked at the images on the lightbox, or the "hanging protocol". A different way to hang the images is shown in Figure 3. All hanging protocols we saw were a variation of those shown in Figures 1 and 3. The variations were the position of the CC and MLOs (possibly reversed from figure 1), the position of the new versus the old (possibly reversed from Figure 1), and whether the left breast images appeared on the left or the right. The latter is due to mammography film being single emulsion and having a shiny and flat appearing side. Some radiologists, in an effort to minimize reflections when viewing the films would flip the images to the "flat" side (reversing the right and left orientation). With laser printed digital mammograms or softcopy display there is no difference in display, except to make the radiologist comfortable by supporting the way they are accustomed to reading.

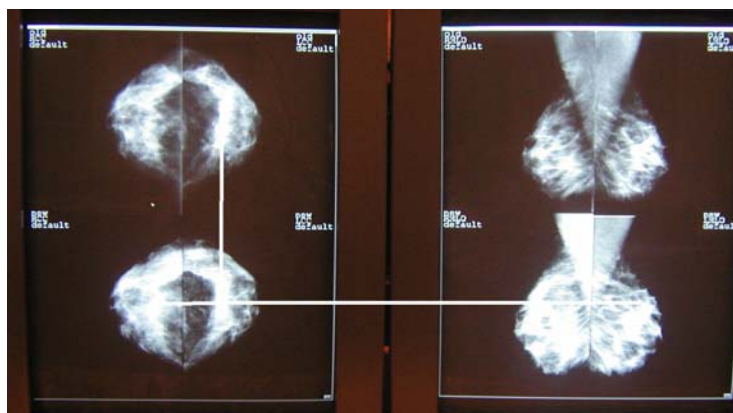


Figure 2. Shows the same display as figure 1, but depicts the links between the three comparison views for the LCC image (which is in the bottom right quadrant of the 4 images on the left screen). The images that are compared (shown by white lines) with this image are the RCC (on the left), the prior LCC (above), and the LMLO (far right).

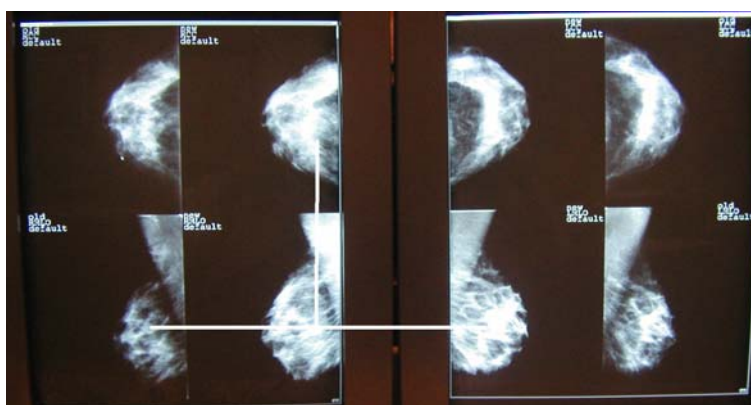


Figure 3. The same study as in figures 1 and 2, displayed with a different hanging protocol. This protocol displays the current study in the center (in the case of the softcopy display this has the disadvantage of splitting the study across the two screens). An advantage of this hanging protocol is that all comparisons are made to adjacent images. Thus, the eye has to travel less to make the visual comparison of the LCC to the LMLO, as compared to the hanging protocol in figure 2.

2.2 Design of Mammoview

Mammoview was designed to support the requirements described above. It supports any hanging protocol layout. It supports the three side by side paired comparison views. The paired comparison views are displayed in turn, controlled by clicking on one of the mouse buttons. With film readings the radiologist changes between viewing the gestalt (all images) and focusing on individual images at full resolution (with the aid of a magnifying lens). Mammoview provides an electronic equivalent to this interaction by displaying two zoom levels, one showing the overview set of images, and a second showing the individual images at full resolution. On the overview presentation the images are interpolated down so that the images fit on the two screens (equivalent to 200 micron resolution). Full resolution images are displayed at their original acquired resolution (50 microns per pixel for our clinical trial). This operation is controlled via a single mouse click. The last operation supported by Mammoview is the display of two different image processing presentations. The first presentation is the “default screening” presentation, and is what comes up initially. The second presentation is a processing optimized to show contrast detail in the dense areas of the breast. This processing choice is based on laboratory experiments which showed improved feature detection of masses and calcifications when images were viewed with this processing compared to the default film screen presentation.¹⁵ As with the other interactions, the user can toggle between the two presentations via a single mouse click. Interactive intensity windowing was intentionally not provided so that radiologists would not spend extra time trying

to window and level the studies. Our experience with chest CT,¹⁶ Xray, and mammography has shown that providing interactive intensity windowing in addition to appropriate presets does not significantly increase performance; however, it does increase reading times. The choice of optimal or good preset processing conditions is important to allow the radiologists to perform as accurately and quickly as when reading film. Our choice for Mammoview was to provide a default screening presentation similar to what they are familiar with, and to provide as a second option the algorithm that had the best mass and calcifications detection performance in our clinical and laboratory trials evaluating processing methods.¹⁷⁻¹⁸

3. RESULTS AND DISCUSSION

Mammoview has been evaluated in three different conditions. First was laboratory testing at UNC-CH, where radiologists read digital and digitized mammography cases and provided feedback. Second, was in educational exhibits in the InfoRAD section of the Radiological Society of North America (RSNA) conference. Third was in clinical trials at UNC-CH. We will summarize the main results from the use of the current version of Mammoview, which was used in clinical trials at UNC from 2000 to 2001, and displayed at RSNA in InfoRAD exhibits from 1998-2001.

3.1 Clinical trial main results: same speed and accuracy

The most important results came from a clinical trial testing the performance of softcopy readings using Mammoview versus film readings.¹⁹ The clinical trial found that the speed and accuracy of readings using the Mammoview softcopy display were not significantly different from readings of the same cases by the same observers using film display. The softcopy reading times, while not statistically significantly different, were slightly faster than those for film. Accuracy was measured by area under the ROC curve. The area under the ROC curve and sensitivity were slightly higher for film display, while specificity was slightly higher for softcopy display. Again, the accuracy results were not statistically significantly different for the two presentations.

3.2 Clinical trial related results: successful automatic preparation of images for viewing

The clinical trial tested not only the softcopy display interface itself, by also many related processing steps that are necessary for a digital mammography display application to function effectively. Important steps included in this trial were the automatic recognition of the breast tissue portion of the image, the cropping of the image to just the breast tissue, the automatic alignment of paired images, and the automatic pre-processing of displayed images into both of the two presentations (default screening and dense breast). The significance of these steps is that they can be automated, and thus performed by the computer. This will allow the images to be fully ready for the radiologist to view them without the need for any manual work by technologists, residents or other staff. This has the potential for reducing labor in two areas. First, with films, work such as re-intensity windowing, re-processing, or even re-shooting of the films may be required if the contrast is not appropriate on the films (or films printed from digital). With softcopy display automatic preprocessing to provide good presentation(s) can eliminate the additional personnel time or resources required to do this. The second savings is in the hanging of the images. With film display manual labor is required to hang the films according to the appropriate protocol (with different protocols for different radiologists). With softcopy display, the appropriate protocol for a radiologist can easily be automatically applied, and changed at any time to a different viewing protocol with the click of a button.

When the clinical trial was performed the best version of these automated methods was included, and no corrections were made when they failed. This was to allow us to test whether a complete softcopy display system could perform adequately. Thus, if parts of the processing performed poorly (for instance automatic recognition of the breast, or display processing) these could potentially significantly impair the performance of the softcopy display system.

Automatic recognition is required in order to perform the cropping and the alignment steps. Many fairly simple mechanisms can be used to automatically identify the breast tissue; the main trick is to avoid false positives due to markers on the films. In this work we create a blurred version of the image and do image analysis to recognize the single large signal present that consists of the breast tissue. The portion of the detector image (or digitized film) that contains breast tissue may range from 20% to nearly 100%. Cropping the original image to one that only contains the breast tissue allows us to maximize the spatial resolution of the image presented in the overview presentation. This is because the smaller the matrix size of the image to start with, the less it has to be shrunk (reducing information content) to spatially fit into one quarter of the 2048x2560 screen size. Note that this also implies that images of different sizes may be interpolated different amounts resulting in different spatial resolutions in the overview presentation. For our clinical trial this was true between studies; however, within the same study a consistent interpolation scale is used to maintain a consistency between the images,

because the radiologist is accustomed to this. Without this it would be difficult to visually compare the size of lesions etc. Some training was required, because not all radiologists had experience reading non-true size images.

Automatic alignment was performed just as with films hung on a lightbox, to aid the radiologist in making cross comparisons between like images (for instance right and left same projection views, or current and prior of the same view). Alignment was done by matching the top and bottom of the breast on the image (known from the recognition of the breast tissue). Alignment worked well (in softcopy display the images were hung as well as the films would be hung manually) about 95% of the time. The cases where the alignment was not perfect were usually only slightly offset. These were not manually corrected, but read as they were.

Lastly, each image was pre-processed for two different presentations, default screening and dense breast. The default screening presentation was used to match as closely as possible what the radiologists were familiar with clinically. The processing was derived from matching the H and D curve from the film display parameters. The dense breast presentation was taken from the best performing processing condition discovered in our laboratory^{8-12,15} and clinical experiments.¹⁷⁻¹⁸ The specific processing condition used was the histogram based intensity windowing technique developed at UNC and described more completely in Hemminger.¹⁵

The recognition of the image was 100% successful. This is necessary as many of the other steps depend on this step. The cropping was also 100% successful. The two areas that were not perfect were the alignment and the processing. Alignment errors usually occurred when there were significant differences between the comparison images, for instance due to surgical procedures performed on one breast but not the other. An example is shown in figure 4. Most misalignments were not large. While some of these failures were noted by the radiologists, in the subjective evaluations the radiologists did not report that they thought their performance was impaired by this factor. It is more difficult to rate the success of the processing for presentation of images. Several radiologists and authors (developers of the processing algorithm) felt that about 10% of the cases were not processed as well as they might have been (compared to manual fine-tuning). The differences in most cases did not appear large. Again, our reasoning was that this tradeoff is acceptable in order to accrue the advantages of fully automated processing. While the radiologists did not report significant problems due to misalignments or poor processing, it is possible that these had a detrimental effect, and that improvements in these two areas could result in improved overall softcopy display performance.

One additional pre-processing step that was recognized as important was to standardize the appearance between the current and prior presentations. In our clinical trial the prior study was digitized film, and the current study was Fischer full field digital acquisitions. In earlier versions of Mammoview, we received comments from the radiologists that comparisons between the current and prior images were more difficult because they didn't look the same. This was due to differences in the acquisition processes. To address this, we computed the equivalent processing of the digital cases to match them to the standard H and D curve presentation for film images, and incorporated this into the preparation of the digital cases. This reduced the dissimilarities between the two studies, and the radiologists using this later version of Mammoview did not report this problem.

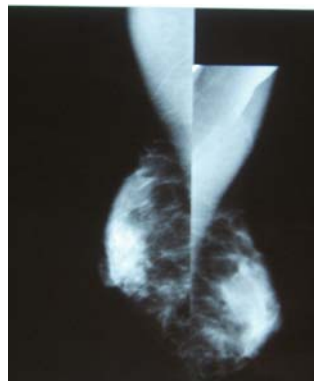


Figure 4. The MLO images are not properly vertically aligned. This can be seen by comparing the tops of the two images, with the image on the right being significantly lower than the one on the left. The bottoms of the images are fairly closely aligned; however, a better alignment would have had the image on the right raised to better match the left image.

3.3 Clinical trial secondary analysis plan

Secondary analyses were planned for the clinical trial based on the videotaping, and user interaction logging that occurred during the clinical trial. In the pilot work developing Mammoview, several weaknesses were identified and fixed, including supporting different hanging protocols, linked panning of comparison images, and matching the pre-processing of digitized film presentation to digital acquisition presentations.

The major weakness that remained in the softcopy display application was how to perform pan and zoom operations to see the mammogram image at full acquired resolution. Based on observations and videotape, radiologists clearly spent more time interacting with the mouse trying to perform pan and zoom operations in order to study the entire image, than they did when just moving their eye around (while using the magnifying glass as needed). While the same information can be seen on the softcopy display, most users reported that it was more cumbersome to use pan and zoom interactions on the softcopy display compared with raising and lowering their magnifying glass.

While panning and zooming is one particular operation that potentially made the softcopy display application more difficult and time-consuming to use, we were also concerned about keeping the interactions with the user interface to a minimum, since the computer human interaction involved with film display is very small (magnifying glass, and forward/backward scrolling motion of film alternator). This was the reason for keeping the mental model simple, and limiting the interactions to a few easily understood ones. To test whether we accomplished this in the interface, and how significant the issues of panning and zooming interactions were, we performed an analysis on the user's interactions recorded from their use of Mammoview in the clinical trial. We tested whether elements of the computer human interaction, the reader, the case, and the image size affected the review time of the radiologist. We were explicitly interested in whether larger images would cause increased review times on the softcopy presentation relative to the film presentation.

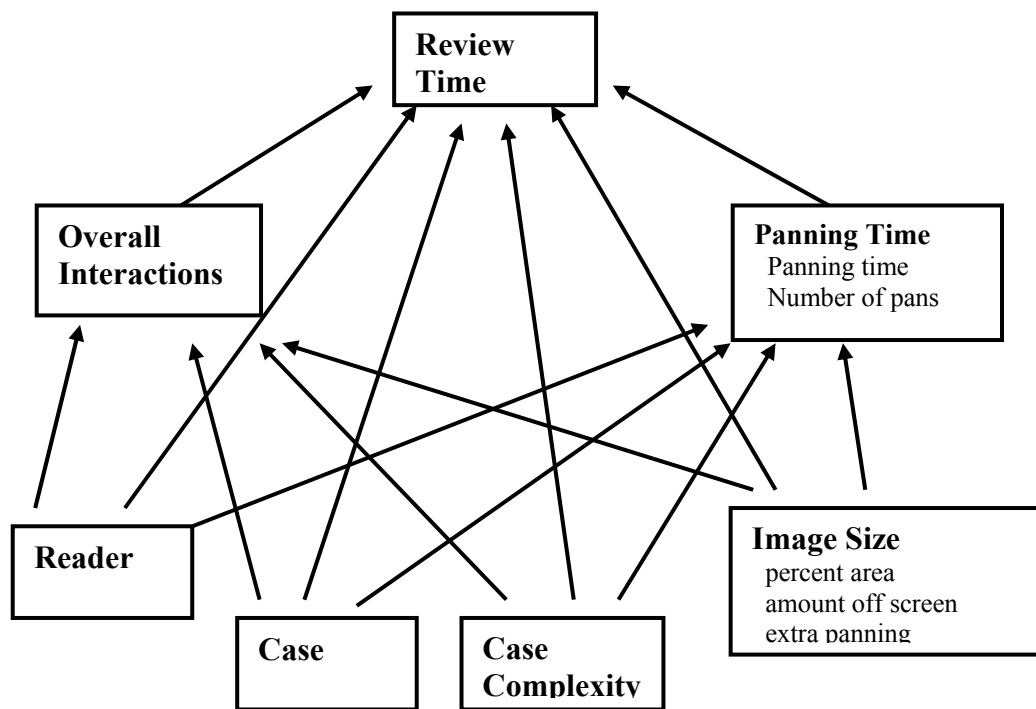


Figure 5. The regression analysis plan for the secondary analyses of the clinical trial. The independent variables are shown on the bottom row. The mediator variables are shown in the middle, and the dependent variable, review time, is shown on the top. The arrows indicate the possible effects that the analysis investigated. The primary investigation was of the effects on review time.

Figure 5 shows the multiple regression analysis performed to study what variables effect the speed of the radiologist's readings. The radiologist's review time was the measure of speed. Review time was the dependent variable in the regression analysis. Independent variables were observer, case, case complexity and image size. Mediator variables were the

number of overall interactions (all human computer interactions) and the amount of panning (panning when zoomed to full resolution). Review time is the time used by the radiologist to make their determination on the case, and does not include the dictation time. This is because dictation times vary significantly based on factors independent of reviewing the case (for instance the radiologist makes a dictation error and rewinds and rerecords, increasing the dictation time). Thus, not including the dictation times reduces the variance due to unrelated factors and improves our sensitivity. The main factor investigated for its effect on review time was image size. Larger image sizes mean more of the image was off-screen and more panning and zooming would be required. We tested three different measures of the concept of image size, shown pictorially in figure 6, to see which measure captured it the best:

total percentage area: measured the area of the “cropped” breast versus the screen size,
amount offscreen: measured the area of the cropped breast area that was off-screen,
extra panning: measured the linear amount of extra panning required to view all areas initially off-screen.

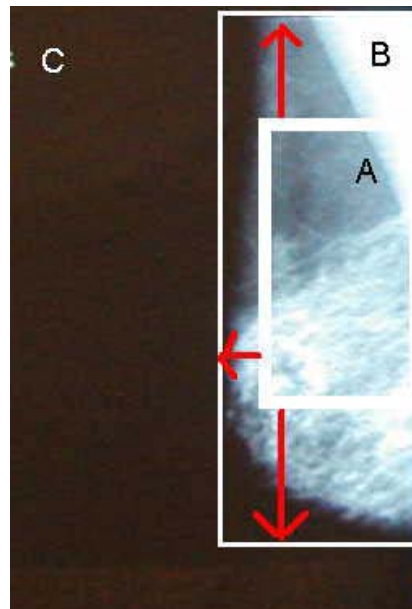


Figure 6. Shows the definitions of the three image size measure concepts. C is the entire acquisition image. B is the entire breast as imaged. A is the size of the display mapped onto the image. Because the display size (A) is smaller than the breast, some panning will be required. Three different specifications of the amount of panning required are given: *total percentage area* = $\text{area } B / \text{area } A$; *amount offscreen* = $\text{area } B - \text{area } A$; *extra panning* = *length of the arrows, which is equivalent to* $\text{length } B - \text{length } A + \text{width } B - \text{width } A$.

Case and observer were descriptive variables. They were included to see if there were differences based on the readers or the case. Case complexity was a variable defined by the number of features reported for that case (averaged across all observers). Larger values of case complexity corresponded to more mammographic features being visible in the case. Thus, increased case complexity may imply longer review times. The mediator variable of interactions was the number of computer human interactions recorded for the review. It includes operations like pan, zoom, change processing preset, and change viewed image comparison. This measure was intended to capture whether the number of interactions had an effect on the review time. We had two choices for a measure of the second mediator variable, panning time: the total time spent panning during a review, or the total number of pan operations that occurred during a review.

3.4 Results

Choice of measures to use for Panning Amount

The measure of total panning time was clearly superior to that of the total number of pans. Across the different regression analyses performed, the values for $\text{Pr} > \{t\}$ were in the range 0.20-0.30 for total panning time, while they were in the range 0.50-0.80 for the total number of pans in the corresponding regression analyses (with all other variables the same). This seems reasonable since the total amount of time spent panning should reflect more accurately how much time the radiologist

spent in the panning operation. For instance in one case the observer only had two pan operations, with the total pan time being 34 seconds out of a total review time of 64 seconds, while in another case with two pan operations the total pan time was 1.5 seconds out of 84 seconds. For our main analysis we choose to use total panning as the measure of Panning Amount.

Choice of measure for Image Size

There was not a large difference between the three measures tested. In general, the extra pan measure was always slightly more strongly correlated with review time. It also was barely statistically significantly correlated with review time ($Pr > \{t\} = 0.052$) while the other two were not ($Pr > \{t\} = 0.058$ for total percent area, $Pr > \{t\} = 0.066$ for amount off screen). Thus, for our main analysis we used the extra pan measure for the concept of Image Size.

Main Analysis

In order to determine which effects were significant, multiple regression analyses were performed for all combinations of the variables (included and not included) in figure 6. The regression analysis was performed using REG (regression) procedure in the SAS statistical package. Three variables showed a statistically significant correlation with review time: reader, interactions, and image size. The results with all variables included is shown in table 1. The overall adjusted R² correlation was 0.162 for this regression analysis. When a regression analysis is done including only Interactions, the adjusted R² correlation is 0.153. Thus, the overall correlation is not overly strong, and further most of the correlation is due to the Interactions variable, with Reader and ExtraPan contributing only a small part of the correlation.

Variable	Parameter Estimate	Standard Error	T Value	Pr > {t}
Complexity	-0.001090	0.01091	-1.00	0.3183
Reader	0.01360	0.0055	2.45	0.0143
Case	-0.00035	0.0003	-1.20	0.2309
Pan time	7.079	6.229	1.14	0.2561
Interactions	0.01051	0.00087	12.04	<.0001
Extra pan	-0.00003	0.00002	-1.94	0.0521

Table 1. The output of the SAS regression procedure, analyzing the effect on review time. The variables of Reader, Interactions, and Extra Pan were statistically significantly correlated with ($Pr > t$) values less or equal to 0.05, shown in bold.

The correlation of reader with review time is expected, as inter-reader differences in review time have been demonstrated previously. The correlation of interactions with review time, with the larger number of interactions causing longer review times would logically be expected. This supports the choice of simple and easy to use interfaces, as well as simplifying mental modes and operations. Thus, the choice not to allow interactive intensity windowing and support only fixed presets that have been proven to be beneficial seems to be supported. The correlation of larger image sizes with longer review times indicates an area of potential improvement for softcopy displays. This could be addressed by improvements in the spatial resolution supported by monitors (for instance newly demonstrated technology, 8 megapixel CRTs and 9 megapixel flatscreens, would eliminate most of the panning because the entire image would fit on one screen). Alternatively, improvements in pan and zoom human computer interactions might reduce the time spent panning as well.

3.5 Qualitative evaluations

In this section we summarize the results from the structured interviews as well as the radiologists' subjective comments recorded during the clinical study. The comments given by radiologists using Mammoview during the RSNA InfoRAD demonstrations (1998 and 1999) are included in this section as well. At the end comments from the combined Mammoview+CAD tool (RSNA 2000 and 2001 InfoRAD demonstrations) are summarized.

Structured questionnaires and interviews were conducted with each radiologist observer at the end of the clinical study. Comments were also recorded during the study. Videotape of the observers using the system, and log files indicating operations and timing were utilized as well to help understand the advantages and disadvantages of the softcopy display system. The responses have been grouped into four categories:

Advantages of Softcopy/Disadvantages of Film

- Second preset (dense breasts) is helpful.
- Easier to do comparisons. Faster and less interactions required.
- Could focus better on images; images were brought to your attention instead of looking all around to find and move them to your attention.
- Automatic preparation of images speeds up reading (alignment, hanging).
- Magnification ability or tool (click of a button) was better.
- Were able to see images better because no masking of light box was required (done automatically by softcopy display).
- Interface was well designed, easy to pickup.
- Very clean image view (only breast and labels).

Advantages of Film/Disadvantages of Softcopy

- Sometimes the views were not perfectly aligned on softcopy.
- Sometimes the contrast (image presentation) seemed poorer on the digital images. (However, it was also reported as sometimes poorer on the film presentations).
- The edge of the monitors gets in the way. Radiologists like to have comparison images right next to each other; the softcopy interface with two CRTs causes a gap of 9cm in the middle between the images.
- Like the feel of shuffling the films. Feels more involved than pressing buttons on softcopy.
- Like the ability to move films around and flip them as desired (to arbitrarily change layout).
- Felt like there was more eye strain when using the CRT displays for long hours as compared to film.
- Wanted to be able to have more complete adjustment of height of monitors versus sitting height (ergonomics).
- Want ability to markup images on softcopy display.
- Variable sizing on digital images is sometimes confusing; need to have a digital ruler overlay capability so they can be sure of scale.
- Panning interface is slightly slower than desirable (sometimes slightly jerking/swimming). Would be less bothersome, and give faster review times if the pan were completely real time (no lag).
- When clicking through comparison views, would like to maintain position. I.e. if have panned half way down, would like the comparison views to show the same anatomy location.
- Panning felt more time-consuming on the softcopy display than on film.
- Load time for next study was too slow.
- Would like to have a full screen image mode (a zoom size between overview and full acquisition resolution) to make it easier to see the whole image at once but at higher quality.
- Easier to miss things on large breast images when zoomed in—because easier to lose track of where you are and what you've seen.

What would you like to have on the Softcopy Display that was not provided?

- Ability to invert images.
- Ability to interactively intensity window cases with potential pathology, i.e. screening cases that become diagnostic.
- Ability to have the third comparison (current CC with current MLO of same breast) available on the comparison mouse button. (This is a standard feature of Mammoview and was present for the RSNA demonstrations and CAD experiment; however it was removed from the clinical experiment due to technical reasons, which caused the load time between studies to be too long to be practical).

Would you be willing to use the softcopy interface as provided clinically?

Seven out of eight radiologists in the clinical study stated they would be willing to use it clinically now. The one who did not, responded negatively because they felt the image quality of the digital images was not good enough yet. They indicated that this was due to the prototype scanner and less well developed acquisition protocols, and not due to the user interface for the softcopy display system.

Unrelated to the softcopy display, some radiologists commented on the quality and technique of the digital images. This was due to the sometimes poorer acquisition technique in acquiring the digital cases, caused by difficulties in imaging close to the chest, and less experienced technologists.

All of the problems listed above for softcopy display have been resolved in the production version of Mammoview (September 2001), which incorporates improved preprocessing, electronic markup, interactive intensity windowing, inverting, arbitrary placement and flipping of images, and three stage zoom (overview, fit to screen, full resolution). The remaining issues could be addressed by technological changes:

- The edge of the monitors gets in the way. *Could be addressed by new technology such as the 9 megapixel flat panel displays.*
- Felt like there was more eyestrain when using the CRT displays for long hours as compared to film. *This has been noted by other researchers,²⁰ and may be an issue with CRTs refresh. Possibly addressed by flat panel displays.*
- Panning update is slightly slow. *Depends on technology (main bottleneck is PCI bus on PCs; the newer 66Mhz PCI bus and cards may alleviate this). Using larger spatial resolution displays would eliminate this issue.*
- Pan&zoom operation is slower/worse. Easier to miss things on large breast images when zoomed in—because easier to lose track of where you are and what you’ve seen. *Best solution is again larger monitors, such as the 8 megapixel CRTs or the 9 Mega pixel flat panels.*

The important consideration is to have the spatial resolution of a single monitor be as large as the acquisition resolution. If it is less, then the pan and zoom issues become problems, and softcopy display review times will increase relative to film review times as the image is larger relative to the display system resolution. Table 2 shows the spatial resolutions of the current full field digital mammography acquisition systems, and the high end CRT and flat panel displays. Note that the highest end display systems have sufficient resolution for some acquisition systems; however, the highest resolution display commonly available is the 5 megapixel CRT. Some manufacturers of acquisition systems with lower spatial resolution benefit because their image size is less than 5 megapixels, and thus an image fits onto one screen.

Display -or- Scanner	GE	5 MPixel CRT Displays	8 Mpixel CRT Displays	9 Mpixel flat panel Display	Hologic Se Plate	Fischer	Fuji	Hologic CCD
Spatial Resolution (matrix size)	1800x2304	2048x2560	2560x3072	2700x3400	2560x3584 3584x4096	3072x4800	3540x4740 4728x5928	4800x6400

Table 2. Comparison of spatial matrix sizes of display systems and acquisition systems for digital mammography.

The advantages of the softcopy display are increased when considering a full-blown softcopy interface, such as the production version of Mammoview. These advantages have been demonstrated in the production version of Mammoview in pilot experiments and at RSNA demonstrations, and include

- **Electronic Markup.** The electronic grease pen markup ability was rated as significantly superior to film grease pens, because you could write directly on the image, but also use graphics and text annotations. Further you could tie the marks to the dictated report, which better pinpoints features than quadrants or clock face descriptions (the two commonly used location description methods). Lastly, the graphics and annotations can be toggled on and off so that they do not have to obscure the image.
- **Faster.** Except for pan/zoom, the softcopy display interface was faster.
- **Faster still when loading counted.** If preparation and handling of the images is included, then softcopy is substantially faster than film (due to loading the images, and hanging them, which take no time in softcopy).
- **Standard advantages of PACS systems,** i.e. multiple copies available, telemedicine possible, no lost films, decreased enterprise costs from filmless hospitals.
- **Improved incorporation of CAD.** CAD is available via a toggling the CAD marks on and off, instead of having multiple fixed views of the images with and without CAD applied.

Support for standardized presentation of some of these features, including the image and associated graphics and text markup, is supported by the DICOM medical imaging standard²¹. Additionally the standardized exchange of CAD reports has been defined, and a DICOM standard for mammography reports is nearing completion.²¹

3.6 CAD combined with softcopy display

A pilot experiment and three RSNA demonstrations have evaluated using CAD in conjunction with softcopy display. The number of observers, and the experimental conditions (small numbers of cases in different conditions) limit the results to subjective impressions recorded from the observers, and experimenter's observations. In the initial version, different CAD states could be toggled on and off, including presenting marks for CAD identified masses only, calcifications only, or both. Additionally, control was provided over the whether the calcifications were presented individually or clustered. After the pilot work we choose to include only a single toggle button that toggled on or off the complete CAD markup (which marked each mass, and each calcification cluster, with solid or dashed line rectangles, respectively). An example is seen in figure 7. The simplification was done because we found that allowing more detailed control resulted in increasing the reading times for radiologists, which negatively impacts screening readings which are almost all negative. This version was tested during RSNA 2000 and RSNA 2001, with positive responses. Reading times were on par with those reported in our clinical study. However, because the cases were not the same, and the observers were not as well trained in using Mammoview, it is difficult to draw comparisons. Accuracy seemed to improve, especially for less experienced readers. This is similar to the majority of literature evaluating similar academic and commercial CAD devices.²²⁻²⁵ The major area of concern was whether adding the display of CAD results would increase reading times. By having an additional presentation mode where the views are all same but with CAD results overlaid, there is potential for increasing the reading times. This potential cause of an increase in viewing times may have been offset by the radiologists being more directed, spending less time identifying all calcifications, due to the CAD markings indicating areas of interest. Through the use of a simple interaction (toggling on and off the CAD markers) there seemed to be little additional overhead for the radiologists using the CAD.

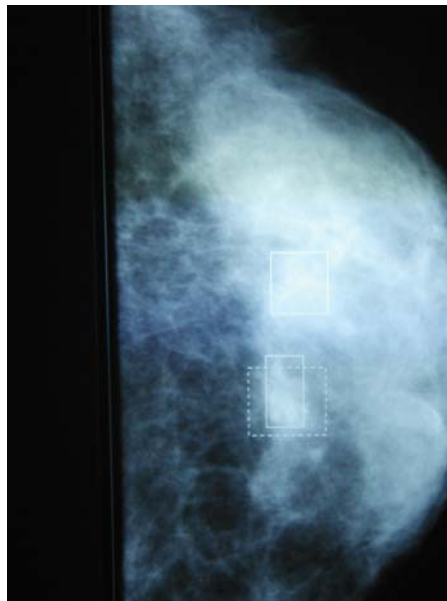


Figure 7. Mammoview with CAD. The CAD markers are turned on, and two suspected masses (solid line rectangles) and one suspected microcalcification cluster (dashed line rectangle) are shown on this single view.

4. CONCLUSIONS

A softcopy display, Mammoview, has been demonstrated to be as fast and as accurate as film reading for radiologists performing screening mammography. When image preparation and hanging times are included, the softcopy display is significantly faster than the film display. There are several areas where softcopy has demonstrated further improvements, and could provide significant advantages compared to film display, including better markup and annotations, fully automated reading without the need for manual processing or hanging of films, and better support for CAD and filmless display

environments. There are, however, some areas where softcopy is worse than hardcopy, and these are mainly due to constraints of the physical displays themselves. Most all of these disadvantages can be addressed by utilizing very high resolution displays, so that the display has spatial resolution greater than or equal to the original acquired resolution of the individual images. Otherwise, as the acquired resolution becomes significantly larger than the display resolution, the review times of the softcopy display can begin to increase relative to the corresponding film reading times.

ACKNOWLEDGMENTS

This clinical trial evaluating Mammoview versus film display was conducting in the Radiology Research Laboratory and supported by the mammography research group in the Department of Radiology at UNC-CH, directed by Dr. Etta Pisano. Original acquisition of the Fischer digital mammography cases was supported by a grant from Fischer Digital Imaging. Work evaluating the incorporation of CAD results into softcopy display was supported by and done in collaboration with Qualia Computing, developers of the CADx computer aided detection product for digital mammography. Printing of hardcopy films was made possible by the Eastman Kodak Corporation. Anne Bauers helped with the secondary analyses. The Odum Institute for Research in Social Sciences at UNC-CH, especially Jun Koo and Chris Wiesen, aided us in performing the statistical regression analysis. In addition, the following grants partially supported this work:

Grey Scale Image Processing for Digital Mammography, National Cancer Institute, National Institutes of Health, Grant # RO1 CA60193-05.

Evaluation of Digital Mammography Display, U.S. Army Medical Research and Material Command, Grant # DAMD 17-94-J-4345.

REFERENCES

1. Hollingsworth AB, Taylor LD, Rhodes DC. "Establishing a histologic basis for false negative mammograms." *American Journal of Surgery*. 166(6): 643-647, 1993.
2. Baker LH. Breast cancer detection demonstration project: five-year summary report. *CA* 32(4), 194-225, 1982.
3. Shtern, F. Digital mammography and related technologies: a perspective from the National Cancer Institute. *Radiology*. 183: 629-630, 1992.
4. Pisano ED, Yaffe M. Digital Mammography. *Breast Disease*. 1998; 10(3,4): 127-136.
5. Feig SA, Yaffe MJ. Current Status of Digital Mammography. *Seminars in US, CT and MRI*. 17(5): 424-443, 1996.
6. Rosenman J, Roe CA, Cromartie R, et al. Portal film enhancement: Technique and clinical utility. *Int J Radiat Oncol Biol Physics* 1993;25(2):333-338.
7. Shtern F, Winfield DW, Behlen F, Blume H, Flynn MJ, Hemminger BM, Huang HK, Krupinski EA, Kundel H, Roehrig H, Shile PE, Sickles E, "Report of Working Group on Digital Mammography: Digital Displays and Workstation Design," *Academic Radiology*, Vol 6, Supplement 4, April 1999, pp 197-218.
8. Puff DT, Cromartie R, Pisano ED. Evaluation and optimization of contrast enhancement methods for medical images. *Proceedings of the SPIE Visualization in Biomedical Computing Conference* 1992; 1808: 336-346.
9. Puff DT, Pisano ED, Muller KE, et al. A method for determination of optimal image enhancement for the detection of mammographic abnormalities. *J Dig Imaging* 1994; 7:161-171.
10. Pisano ED, Garrett W, Hemminger BM, Johnston RE, Muller K, Pizer S. The effect of Contrast Limited Adaptive Histogram Equalization on the detection of simulated spiculations embedded in digitized mammograms. *Journal of Digital Imaging*. (In Press).

11. Pisano ED, Chandramouli J, Hemminger BM, DeLuca M, Glueck D, Johnston RE, Muller K, Braeuning MP, Pizer S. Does Intensity Windowing Improve the Detection of Simulated Calcifications in Dense Mammograms? *Journal of Digital Imaging*. 1997; 10(2): 79-84.
12. Pisano ED, Chandramouli J, Hemminger BM, Johnston RE, Muller K, Pizer S. The effect of intensity windowing as an image processing tool in the detection of simulated masses embedded in digitized mammograms. *Journal of Digital Imaging*. 1997; 10(4): 174-182.
13. Weibrecht M, Spekowius G, Quadflieg P, Blume H. Image quality assessment of monochrome monitors for medical soft copy display. *SPIE* 1997; 3031: 232-244.
14. Beard DV. Bream P. Pisano ED. Conroy P. Johnston RE. Braeuning P. McLelland R. Clark R. A pilot study of eye movement during mammography interpretation: eyetracker results and workstation design implications. *Journal of Digital Imaging*. 10(1):14-20, 1997 Feb.
15. Hemminger BM, Zong S, Muller KE, Coffey CS, DeLuca MC, Johnston RE, Pisano ED, "Improving the Detection of Simulated Masses in Mammograms through two different Image Processing Techniques", *Academic Radiology*, Sept 2001.
16. Beard DV. Hemminger BM. Perry JR. Mauro MA. Muller KE. Warshawer DM. Smith MA. Zito AJ. Interpretation of CT studies: single-screen workstation versus film alternator. *Radiology*. 187(2):565-9, 1993 May.
17. Pisano ED. Cole EB. Major S. Zong S. Hemminger BM. Muller KE. Johnston RE. Walsh R. Conant E. Fajardo LL. Feig SA. Nishikawa RM. Yaffe MJ. Williams MB. Aylward SR. Radiologists' preferences for digital mammographic display. *Radiology*. 216(3):820-30, 2000 Sep.
18. Pisano ED. Cole EB. Hemminger BM. Yaffe MJ. Aylward SR. Maidment AD. Johnston RE. Williams MB. Niklason LT. Conant EF. Fajardo LL. Kopans DB. Brown ME. Pizer SM. Image processing algorithms for digital mammography: a pictorial essay. *Radiographics*. 20(5):1479-91, 2000 Sep-Oct.
19. Pisano ED. Cole EB, Kistner E, Muller KE, Hemminger BM, Brown M, Johnston RE, Kuzmiak C, Freimanis R, Soo MS, Baker J, Walsh R, "Digital Mammography Interpretation-Comparison of Speed and Accuracy of Softcopy versus Printed Film Display", *Radiology*, Feb, 2002.
20. Krupinski E. Roehrig H. Furukawa T. Influence of film and monitor display luminance on observer performance and visual search. *Academic Radiology*. 6(7):411-8, 1999 July.
21. Digital Imaging and Communications (DICOM) international standard, published by the National Electrical Manufacturers Association, Washington DC, 2002.
22. Birdwell RL. Ikeda DM. O'Shaughnessy KF. Sickles EA. Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. *Radiology*. 219(1):192-202, 2001 Apr.
23. Leichter I. Fields S. Nirel R. Bamberger P. Novak B. Lederman R. Buchbinder S. Improved mammographic interpretation of masses using computer-aided diagnosis. *European Radiology*. 10(2):377-83, 2000.
24. Chan HP. Sahiner B. Helvie MA. Petrick N. Roubidoux MA. Wilson TE. Adler DD. Paramagul C. Newman JS. Sanjay-Gopal S. Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: an ROC study. *Radiology*. 212(3):817-27, 1999 Sep.
25. Freer TW. Ullissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology*. 220(3):781-6, 2001 Sep.

Correspondence: Bradley Hemminger, bmh@rad.unc.edu, phone (919) 966-2998, fax (919) 962-8071, School of Information and Library Science, 206A Manning Hall, University of North Carolina at Chapel Hill, 27599-3360.