Award Number: DAMD17-98-1-8194

TITLE: Computerized Analysis of MR and Ultrasound Images of Breast Lesions

PRINCIPAL INVESTIGATOR: Maryellen Giger, Ph.D.

CONTRACTING ORGANIZATION: The University of Chicago
Chicago, Illinois  60637

REPORT DATE: July 2000

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, 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.
4. TITLE AND SUBTITLE
Computerized Analysis of MR and Ultrasound Images of Breast Lesions

6. AUTHOR(S)
Maryellen Giger, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of Chicago
Chicago, Illinois 60637

E-MAIL:
m-giger@uchicago.edu

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

13. ABSTRACT (Maximum 200 Words)
Although general rules for the differentiation between benign and malignant mammographically identified breast lesions exist, considerable misclassification of lesions occurs with the current methods. The main goal of the proposed research is to develop noninvasive, computerized methods for analyzing ultrasound and MR (magnetic resonance) images of breast lesions to aid radiologists in their workup of suspect lesions.

We currently have retrospectively collected over 400 ultrasound cases of mass lesions, all that had gone on to either biopsy or cyst aspiration. We extracted and calculated features related to lesion margin, shape, homogeneity (texture) and the nature of the posterior acoustic attenuation pattern. Linear discriminant analysis round-robin runs yielded $A_Z$ values of 0.94 and 0.87 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively. The "equivocal" database contained lesions that had been proven to be benign or malignant by either cyst aspiration or biopsy.

Dynamic MR data was analyzed with the computerized classification method, which includes temporal features of normalized speed and inhomogeneity of uptake, and spatial features of margin descriptors such as circularity and irregularity. Results indicate that classification based on temporal and spatial features combined can yield a positive predictive value of 94%, and has the potential to reduce the number of unnecessary biopsies by approximately 92%. We also have developed a new method for automatically extracting the lesion from the 3D image set of the breast. ROC analysis yielded $A_Z$ values of 0.90 when the manual segmentation was used in the classification and 0.93 when automatic segmentation was included.

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT
Unlimited
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.

N/A 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).

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

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

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

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRONT COVER</td>
<td>1</td>
</tr>
<tr>
<td>STANDARD FORM (SF 298)</td>
<td>2</td>
</tr>
<tr>
<td>FOREWORD</td>
<td>3</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>5</td>
</tr>
<tr>
<td>BODY</td>
<td>5</td>
</tr>
<tr>
<td>KEY RESEARCH ACCOMPLISHMENTS</td>
<td>10</td>
</tr>
<tr>
<td>REPORTABLE OUTCOMES</td>
<td>10</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>10</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>10</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
</tr>
</tbody>
</table>
INTRODUCTION

The main hypothesis to be tested is that, computerized analysis of breast ultrasound and MR images should yield new methods for distinguishing between malignant and benign lesions and thus, reduce the number of unnecessary biopsies. In addition, even higher performance is expected when a combination of features from mammographic, MR, and ultrasound images is used as an aid to radiologists in the task of distinguishing between malignant and benign lesions. The main goal of the proposed research is to develop noninvasive, computerized methods for analyzing ultrasound and MR (magnetic resonance) images of breast lesions to aid radiologists in their workup of suspect lesions. The specific objectives of the research to be addressed are: 1. Create a database of ultrasound and MR images including malignant lesions, benign solid masses, and complex cysts; 2. Develop noninvasive, computerized methods for characterizing the lesions to yield an output related to the probability of malignancy; and 3. Evaluate the efficacies of the new image analysis methods in the task of distinguishing between malignant and benign lesions. It is expected that the results from this research will aid radiologists in determining the likelihood of malignancy and in reducing the number of benign cases sent to biopsy. Computerized image analysis techniques that can objectively and reliably classify lesions based upon reported sonographic and/or MR characteristics of benign and malignant masses, especially if combined with their mammographic features, could significantly improve the specificity of breast imaging and the evaluation of breast masses. The proposed work is novel in that computer-aided diagnosis techniques applied to gray-scale sonographic images has not yet been reported. In addition, computerized analysis of MR images of the breast has mainly been limited to only temporal analysis using contrast media.

BODY

Breast cancer is a leading cause of death in women, causing an estimated 46,000 deaths per year (1). Mammography is the most effective method for the early detection of breast cancer, and it has been shown that periodic screening of asymptomatic women does reduce mortality (2-4). Many breast cancers are detected and referred for surgical biopsy on the basis of a radiographically detected mass lesion or cluster of microcalcifications. Although general rules for the differentiation between benign and malignant mammographically identified breast lesions exist (5, 6), considerable misclassification of lesions occurs with the current methods. On average, less than 30% of masses referred for surgical breast biopsy are actually malignant (7).

Breast sonography is used as an important adjunct to diagnostic mammography and is typically performed to evaluate palpable and mammographically identified masses in order to determine their cystic vs. solid natures. The accuracy of ultrasound has been reported to be 96-100% in the diagnosis of simple benign cysts (8). Masses so characterized do not require further evaluation; however, 75% of masses prove to be indeterminate or solid on sonography and are candidates for further intervention (9). With the advent of modern high-frequency transducers that have improved spatial and contrast resolution, a number of sonographic features have emerged as potential indicators of malignancy, while other features are typical for benign masses (10,11). Benign features include hyperechogenicity, ellipsoid shape, mild lobulation, and a thin, echogenic pseudocapsule. Malignant features include spiculation, angular margins, marked hypoechoegenicity, posterior acoustic shadowing, and a depth:width ratio greater than 0.8. Recently, Stavros, et al., used these and other features to characterize masses as benign, indeterminate, and malignant (12). Their classification scheme had a sensitivity of 98.4% and a negative predictive value of 99.5%. However, the sonographic evaluation described by these investigators is much more extensive and complex than is usually performed at most breast imaging centers.

Breast MR imaging as an adjunct to mammography and sonography reveals breast cancer with a higher sensitivity than do mammography and sonography only (13). However, using all three methods in the human interpretation process yielded a lower specificity. It also has been shown that temporal analysis from dynamic MR correlates with intensity of fibrosis in fibroadenomas (14).
Some computerized analyses of spatial features are being performed. Adams et al. achieved a separation between malignant and benign lesions using a statistical analysis, however, their database consisted of only 16 cases (15).

Computerized image analysis techniques that can objectively and reliably classify lesions based upon reported sonographic and/or MR characteristics of benign and malignant masses, especially if combined with their mammographic features, could significantly improve the specificity of breast imaging and the evaluation of breast masses. Computer-aided techniques have been applied to the color Doppler evaluation of breast masses with promising results (16). However, color Doppler imaging is a technique which focuses only upon the vascularity of lesions. Since not all sonographically visible cancers have demonstrable neovascularity, this technique is inherently somewhat limited. On the other hand, computer-aided diagnosis techniques applied to gray-scale sonographic images has not yet been reported. In addition, computerized analysis of MR images of the breast has mainly been limited to only temporal analysis using contrast media.

Comprehensive summaries of investigations in the field of mammography CAD have been published by the co-P.I. (17, 18). In the 1960s and 70s, several investigators attempted to analyze mammographic abnormalities with computers. These previous studies demonstrated the potential capability of using a computer in the detection of mammographic abnormalities. Gale et al. (19) and Getty et al. (20) are both developing computer-based classifiers, which take as input diagnostically-relevant features obtained from radiologists' readings of breast images. Getty et al. found that with the aid of the classifier, community radiologists performed as well as unaided expert mammographers in making benign-malignant decisions. Swett et al. (21) are developing an expert system to provide visual and cognitive feedback to the radiologist using a critiquing approach combined with an expert system. At the University of Chicago, we have shown that the computerized analysis of mass lesions (22) and clustered microcalcifications (23) on digitized mammograms yields performances similar to an expert mammographer and significantly better than average radiologists in the task of distinguishing between malignant and benign lesions.

The proposed work is novel in that computer-aided diagnosis techniques have not yet been applied to gray-scale breast ultrasound and/or MR images. In addition, future research involving the use computers to merge features from mammographic, MR, and ultrasound images, as an aid to radiologists, has not yet been investigated.

The main goal of the proposed research is to develop noninvasive, computerized methods for analyzing ultrasound and MR (magnetic resonance) images of breast lesions to aid radiologists in their workup of suspect lesions. The specific objectives of the research to be addressed are: 1. Create a database of ultrasound and MR images including malignant lesions, benign solid masses, and complex cysts; 2. Develop noninvasive, computerized methods for characterizing the lesions to yield an output related to the probability of malignancy; and 3. Evaluate the efficacies of the new image analysis methods in the task of distinguishing between malignant and benign lesions. It is expected that the results from this research will aid radiologists in determining the likelihood of malignancy and in reducing the number of benign cases sent to biopsy.

1. Establishment of a database of ultrasound and MR images

**Methods**

Approximately 500 sonographically demonstrated lesions will be collected which will include aspirated complex cysts, and biopsied solid benign and malignant masses. The database of these collected cases will include the MR, sonographic, and mammographic images as well as the lesions' ultimate dispositions and diagnoses. (Note that funding already exists for the computerized analysis of the mammographic lesions). Based upon our current case load, we estimate that approximately 30% of the lesions will be complex cysts which required aspiration to prove their cystic nature, 40% will be benign solid masses, and 30% will be cancers. Palpable and mammographically identified
masses are evaluated sonographically by representative images in orthogonal planes, obtaining measurements in these same planes, and most masses are also evaluated with color Doppler imaging. Although the preliminary studies on ultrasound images involved the digitization of ultrasound films, the ultrasound images in this new database will be obtained directly from an ATL ultrasound machine, which produces digital image data. In addition, approximately 50 cases of MR images of the breast will be collected with a T1-weighted sequence, using coronal slices. After injection of GD contrast, 4 to 6 scans of both breasts will be obtained at 90 sec. time intervals. Biopsy results will be used to determine truth regarding malignancy.

Results to Date

We currently have retrospectively collected over 400 ultrasound cases of mass lesions, all that had gone on to either biopsy or cyst aspiration. The images are obtained from University of Chicago and Northwestern University. The images are transferred in digital format from the ATL unit. The digital images within the ATL unit are obtained by screen capture. For each case we have at least two views of the lesion. We are currently collecting the corresponding mammograms for the study.

We currently have retrospectively collected 35 coronal MR cases from University of Muenster, 362 sagittal MR cases of the breast from University of Pennsylvania, and 90 cases from the University of Berlin (which follow a protocol similar to University of Muenster). These are all volume datasets. Of the 362 sagittal cases, 253 are focal (192 malignant, 51 benign, 10 normal), 74 are diffuse lesions (48 malignant and 25 benign), 10 are ductal (9 malignant and 1 benign), and 25 showed no enhancement (3 malignant, 19 benign, 3 normal).

2. Development of computerized method for the classification of lesions

The computerized method will include the image analysis of the texture within the lesion, the analysis of the margin of the lesion, and a comparison of the lesion with its surrounding tissue. Computerized analysis of the texture pattern in the lesion will be based on various texture analysis methods we have been investigating in our laboratory including Fourier spectra analysis and artificial neural networks. We note that it is extremely important to understand the relationship between the mathematical texture measures and the physical nature of the breast parenchyma.

The computerized analysis of the margin characteristics (edge definition) will involve feature extraction using radial edge-gradient analysis. We have done similar analysis on radiographic masses in determining their margin characteristics (spiculated and ill defined) (22). Two promising measures are the FWHM and the average radial gradient which correspond to the degree of spiculation and how ill-defined is the margin, respectively. From the computer-extracted margin, we will also determine the shape and irregularity of each lesion.

Specifically for the ultrasound images, comparison of the "density" and the texture patterns of the lesion with neighboring regions, including those deep to the lesion, will be performed in order to quantify its echogenicity and the amount of any posterior acoustic shadowing or enhancement. This will be performed by comparing feature values "below" the lesion to those obtained along side and below the lesion.

Temporal features will be determined from analyzing the MR image data over time. The contrast medium uptake curve will be analyzed at various spatial locations within and around the suspect lesion. Temporal operators include the maximum uptake, mean gradient of uptake, and rms variation. Both two-dimensional and three-dimensional features will be calculated, e.g., irregularity and margin gradient characteristics. In addition, the spatial features will be investigated as a function of time.

We plan to use artificial neural networks along with other measures of the mass in question to obtain an estimate of the likelihood of malignancy. We will investigate merging the ultrasound image features and MR features with those from mammographic images of the same lesion. We already have funding support for the investigation involving radiographic imaging of masses.
The various features will serve as input data and will be supplied to the input units of the artificial neural network. Prior to input to the ANN, the features will be normalized between 0 and 1. The output data from the neural network are then obtained through successive nonlinear calculations in the hidden and output layers. The calculation at each unit in a layer includes a weighted summation of all entry numbers, an addition of a certain offset number, and a conversion into a number ranging from 0 to 1 using a sigmoid-shape function such as a logistic function. The neural network will be trained by a back-propagation algorithm using pairs of training input data and desired output data. The desired output data will be initially 1 if features of a malignant lesion are input and 0 otherwise. Once trained, the neural network will accept features of a lesion and will output a value that will be related to a likelihood of malignancy. Feature selection will be performed by analyzing the average and standard deviation of the various features for both malignant and benign lesions. Az values will be calculated for each feature as well as for the output of the ANNs. In addition, genetic algorithms, which we have used, in a pilot study, for optimizing feature selection for the task of distinguishing true-positive and false-positive mass detections, will also be used.

Results to Date: Ultrasound
We are developing computerized analyses of breast lesions in ultrasound images to aid in the discrimination between malignant and benign lesions (24). We extracted and calculated features related to lesion margin, shape, homogeneity (texture) and the nature of the posterior acoustic attenuation pattern in ultrasound images of the breast. Our database contained 184 digitized ultrasound images from 58 patients with 78 lesions. Benign lesions were confirmed by biopsy, cyst aspiration, or image interpretation alone, while malignant lesions were confirmed by biopsy. ROC analysis was used to study the performance of the various individual features and the output from linear discriminant analysis in distinguishing benign from malignant lesions. From ROC analysis, the feature characterizing the margin yielded $A_z$ values of 0.85 and 0.75, in the task of distinguishing between benign and malignant lesions in the entire database and in an equivocal database, respectively. The "equivocal" database contained lesions that had been proven to be benign or malignant by either cyst aspiration or biopsy. Linear discriminant analysis round-robin runs yielded $A_z$ values of 0.94 and 0.87 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively.

We are currently evaluating the method on a database of ultrasound images from Northwestern University. The database of over 400 cases includes pathology truth as well as radiologists BI-RADS ratings with all cases having gone to biopsy or aspiration. Our previous method required radiologists' manually-drawn lesion contours as input to the computerized classification scheme. The current method, however, involves automatic segmentation of the lesion contour from the ultrasound image data. Of the 410 cases, 126 were complex cysts, 186 were benign solid lesions, and 98 were malignant lesions. Features related to lesion margin, shape, echogenicity (texture) and posterior acoustic attenuation were automatically extracted. To evaluate the performance of the computer alone, the entire database was divided into training and testing groups. The independent linear discriminant analysis yielded a validation result of an $A_z$ of 0.89 and a partial $A_z$ value at 0.90 sensitivity of 0.52. In addition, in order to evaluate the performance of the computer relative to that of the radiologists, 125 cases were assessed for suspicion by an expert sonographer. Round-robin analysis in the task of distinguishing malignant from benign lesions yielded $A_z$ values of 0.88 and 0.92 for the computer and the radiologist, respectively.

Results to Date: MRI
We are developing computerized analyses of breast lesions in MR images to aid in the discrimination between malignant and benign lesions (25). Dynamic MR data was obtained from 27 patients by a T1-weighted sequence, using 64 coronal slices, a typical slice thickness of 2 mm, and a pixel size of 1.25 mm. After injection of GDTPA contrast, 4 to 6 scans of both breasts were obtained at 90 sec. time intervals. The database contained 13 benign and 15 malignant lesions. Our
computerized classification method includes temporal features of normalized speed and inhomogeneity of uptake, and spatial features of margin descriptors such as circularity and irregularity. Our results indicate that classification based on temporal and spatial features combined can yield a positive predictive value of 94%, and has the potential to reduce the number of unnecessary biopsies by approximately 92%.

We have developed a new method for automatically extracting the lesion from the 3D image set of the breast. Our previous results were based on the use of manually-drawn lesion contours in the various slices of the MR data. The new segmentation method involves the use of an encompassing shell to limit the region for local thresholding. ROC analysis yielded $A_z$ values of 0.90 when the manual segmentation was used in the classification and 0.93 when automatic segmentation was included.

We are currently evaluating the method on 362 cases from the University of Pennsylvania as well as the cases from the University of Muenster and University of Berlin. The UPENN images differ from our initial database in that these cases are sagittal and had fat suppression applied. Also, the UPENN dataset uses fat suppression and thus a modification in the automatic lesion extraction method was made. For the evaluation, we developed a new interface for the human delineation of the lesion margin in 3D to serve as “margin truth”. While outlining the margin in a slice, the observer can see their outline in other views. One performance of index is an overlap calculation in which, in terms of voxels, we calculate the intersection of the human and computer margins divided by the union. We now have this margin truth for roughly 200 cases and we are now running the overlap comparison to determine if the computer outlines similar to the human. We will also do the comparison in terms of the performance of the features extracted from the lesion in the task of distinguishing malignant and benign lesions.

3. Evaluation in the task of distinguishing between malignant and benign lesions.

In order to test the capability of the neural networks to learn the features of malignant and benign lesions, a consistency test will be conducted in which the network is first trained with all the cases in the database and then tested with the same cases used in the training. A consistency test indicates that the network is able to "remember" all of the input types that were used for training. However, it is more important to test if the network can learn a generalized set of inputs from the examples provided and if it can then make a correct prediction for new cases that were not included in the training. Thus, a round-robin method will be employed to test the network's generalizing ability. With the jack-knife method, all but one of the cases are selected randomly from the database for training of the network, and the remaining one case is used for testing the network. The output values are then compared to the "truth" data. Various combinations of training and testing pairs will be selected by using a random number generator and the results will be analyzed using ROC analysis. ROC curves will be obtained by fitting continuous output data from the neural networks using the LABROC4 program (26). The area under the ROC curve ($A_z$) will be used as an indicator of performance. In order to determine the structure of the neural network as well as the necessary number of training iterations, we will analyze the consistency results and the round-robin results in terms of $A_z$ as a function of number of iterations, momentum, learning rate and number of hidden units. We use $A_z$ as an indicator of performance since it includes information on both the sensitivity and specificity of the measures.

The proposed techniques are expected to yield measures about the likelihood of malignancy. Receiver Operating Characteristic (ROC) analysis (26) will be employed in evaluating the performance of the measures. We have used ROC analysis successfully in both evaluating the performance of human observers as well as that of computerized schemes. The task in which the image features will be evaluated will be in their ability to determine an estimate of the likelihood of malignancy. The decision variable for the ROC analysis will be each individual feature as well as combined measures within a modality and combined measures from multiple modalities (x-ray, MR, and ultrasound).
We expect that 500 lesions and their ultrasound images will be available for testing. Note that here the measure of performance will be the A2 value (from ROC analysis) obtained in the task of distinguishing between malignant and benign lesions. To obtain an estimate of the number of lesions needed for adequate statistical power in testing differences in A2 values, we assume only a correlation of 0.60 between the estimates of A2 that are found for our current method involving the computerized analysis of mammographic lesions (A2=0.87) and that for the expected improved method (A2=0.92). With Npos patients who have a malignant lesion and Nneg patients who have a benign lesion, the standard error of the resulting estimate can be approximated (Eqn. 9 in Ref. 27) by the expression 
\[
[(2A2-(1-f)(1-A2))(1-A2)/3Npos)]^{1/2},
\]
where f represents the ratio Npos/Nneg. Thus, with f = 1, the statistical power at a critical significance level of α = 0.05 for 500 mass lesions is 94%.

Results to Date
The results from the evaluation of the methods is described in the preliminary studies described above.

KEY RESEARCH ACCOMPLISHMENTS

- Development of robust features for characterizing lesions in ultrasound images of the breast.
- Development of robust features for characterizing lesions in MRI images of the breast.
- Investigation and development of methods for segmentation in 2D for ultrasound images and in 2D and 3D for MRI image datasets.

REPORTABLE OUTCOMES


CONCLUSIONS
We have made great strides in the development of methods for the classification of lesions in ultrasound and MR images of the breast. We are retrospectively collecting large datasets of ultrasound and MR cases with solid pathology truth and radiologists’ ratings. These cases include malignant lesions, benign solid masses, and complex cysts. We are developing noninvasive, computerized methods for characterizing the lesions to yield an output related to the probability of malignancy and plan to evaluate the efficacies of the new image analysis methods in the task of distinguishing between malignant and benign lesions. It is expected that the results from this research
will aid radiologists in determining the likelihood of malignancy and in reducing the number of benign cases sent to biopsy.

REFERENCES


