Published in final edited form as:

Acad Radiol. 2009 September ; 16(9): 1064–1069. doi:10.1016/j.acra.2009.03.018.

Effect of the enhancement threshold on the computer-aided detection of breast cancer using MRI

JE Levman, P Causer, E Warner, and AL Martel

Sunnybrook Research Institute, Department of Medical Biophysics, University of Toronto, 2075 Bayview Avenue, Room S605, Toronto, ON, Canada

Abstract

Rationale and Objectives—To evaluate the effect that variations in the enhancement threshold have on the diagnostic accuracy of two computer-aided detection (CAD) systems for magnetic resonance based breast cancer screening.

Materials and Methods—Informed consent was obtained from all patients participating in cancer screening and this study was approved by the participating institution's review board. This retrospective study was nested in a prospective, single-institution high-risk, breast screening study involving dynamic contrast-enhanced magnetic resonance imaging. Only those screening examinations (n=223) for which a histopathological diagnosis was available were included. Two CAD methods were performed: the signal enhancement ratio (SER) and support vector machines (SVMs). Statistical analysis was performed by tracking changes in each CAD test's diagnostic accuracy (receiver operating characteristic (ROC) curve area, maximum possible sensitivity, etc.) with changes in the enhancement threshold.

Results—The enhancement threshold plays a significant role in affecting a CAD test's potential sensitivity, ROC curve area and number of assumed true and false positive predictions per cancerous examination. A high threshold can also limit the CAD based detection of the full size of a lesion.

Conclusion—Enhancement thresholds can limit a CAD test's ability to diagnose a lesion's full size and as such should not be raised above 60%. The clinically used SER method exhibits a high rate of false positives at low enhancement thresholds and as such the threshold should not be set lower than 50%. The SVM method yielded better results in our study than the SER method at clinically realistic enhancement thresholds.

Introduction

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) has been shown to be an effective modality for breast cancer screening [1–4], however, there is considerable inter-observer variability between radiologists in their interpretation of the large amounts of data acquired in a breast MRI examination [5]. Computer-aided detection (CAD) systems have the potential to further improve DCE-MRI based breast cancer screening by reducing inter-observer variability. A breast MR examination involves the injection of a contrast agent which causes some tissues to enhance (brighten). The amount of enhancement can be calculated as a percentage of increased signal intensity (brightness). Breast MR CAD systems identify suspected malignant regions and are routinely complemented by an

enhancement threshold which limits the number of false positive predictions by forcing the diagnosis of a specific region of tissue as being benign if the tissue's brightness does not reach the set enhancement threshold. The role of the enhancement threshold in breast MRI CAD systems is illustrated in figure 1 and example curves that exceed the enhancement threshold (A & B) as well as curves that do not exceed the threshold (C & D) are provided in figure 2. Considerable research has been conducted on the use of CAD tools for breast MRI that employ an enhancement threshold, however, justification for the threshold selected is typically limited and reported threshold values vary considerably from 40–150% [6–19]. The purpose of this study is to evaluate the effect that variations in the enhancement threshold have on the computer-aided diagnosis of breast cancer using MRI.

Materials and Methods

Patients and lesions

Between November 3, 1997 and August 21, 2008, 550 high risk women were recruited from familial cancer clinics in southern Ontario and Montreal, Canada. Participation in screening was offered to all eligible women in the context of genetic counseling. Informed consent was obtained from all participants. This retrospective study was approved by the institutional review board of the participating institution. Annual screening of the patient population has resulted in 1749 breast MR examinations. For the purposes of this retrospective study we included only the (223) breast MR screening examinations available with an enhancing lesion for which a gold standard diagnosis was obtained. Surgical biopsy of lesions was performed under MR guidance [20]. Gold standard was histopathology or at least one year of MRI follow-up for presumed benign lesions (n=223, 44 malignant, 179 benign).

Screening MRI Technique

The MRI prospective screening studies consisted of simultaneous bilateral dynamic imaging using a 1.5T magnet (GE Signa, version 11.4). Sagittal images were obtained using a phased-array coil arrangement and dual slab interleaved bilateral imaging method [21]. This provided 3D volume data over each breast obtained with an RF spoiled gradient recalled sequence (SPGR, scan parameters: TR/TE/angle=18.4/4.3/30°, 256×256×32 voxels, FOV: 18×18×6–8cm). Imaging was performed before and after a bolus injection of 0.1 mmol/kg of Gd-DTPA. Bilateral acquisitions were obtained in 2.8 minutes. Slice thickness was 2–3 mm. One pre-contrast and four post-contrast volumes were acquired. In order to compensate for any patient motion that takes place during the examination, a 3D non-rigid image registration technique for breast MRI was used [22].

Computer-Aided Detection

In order to evaluate the effect of varying the threshold on the diagnostic accuracy of CAD systems, two CAD methods were assessed retrospectively, the first being the signal enhancement ratio (SER) method. The software used for SER based CAD was developed at our research institute based on descriptions in the literature [18, 23]. CAD tools based on the SER method are commercially available and are currently in use by radiologists. The second CAD technique addressed in this study is support vector machines (SVMs) [24]. The SVM CAD software used in this study was also developed at our research institute based on

descriptions in the literature. Both our in-house techniques (SER and SVM) have been previously validated [25]. SVMs are an advanced statistical technique whose application to breast MR data has been an active area of research. Research has shown SVMs to outperform many alternative CAD techniques (signal enhancement ratio, k-nearest neighbour, tree methods, fisher discriminant analysis) in separating malignant and benign breast MR lesions [25, 26]. SVM based CAD for breast MRI has been implemented using the libsvm open source library [27].

Statistical Analysis

Both CAD techniques were evaluated through receiver operating characteristic (ROC) curve analysis [28], which is based on the trade-off between sensitivity and specificity for a given test. Here we are measuring the sensitivity and specificity of the CAD techniques on enhancing MR imaged breast tissue. Both CAD techniques were evaluated on the image slice where the suspect lesion was most visible. To evaluate the effect of variations in the enhancement threshold on the diagnostic accuracy of the two CAD systems, changes in the ROC curve area were tracked as threshold values were varied. Enhancement threshold values were varied from 0 to 350% in steps of 10%.

While ROC analysis provides a convenient way to evaluate the amount of separation between our malignant and benign lesions, it does not provide us with information regarding the number of pixels diagnosed correctly. Strictly speaking we do not have a gold standard at the pixel-by-pixel level, only a single histopathological diagnosis for each radiologically identified lesion. In order to provide a robust evaluation of these CAD systems we have also tracked the number of assumed true positive pixels (ATPP) and assumed false positive pixels (AFPP) diagnosed by the two CAD techniques. Ideally this ratio would be maximized, as reducing the number of assumed false positive pixels with respect to the number of assumed true positive pixels can help prevent a radiologist from making false positive diagnoses. This is meant to complement standard ROC analysis and so the ratio plot (ATPP/AFPP) has been placed alongside the ROC area plot for ease of evaluation.

We also tracked changes in the maximum possible CAD test sensitivity with variations in the enhancement threshold. This was obtained by simplifying the CAD system illustrated in figure 1 such that all brightness-time curves that exceed the enhancement threshold are labeled cancerous.

Classifier Visualization

In addition to ROC analysis we also created high-dimensional classifier visualization plots [25] to assist in the evaluation of the effects of threshold variations (see figure 4). A plot of our high-dimensional data was created using principal components analysis in a projected two-dimensional space. The decision boundary for a given prediction method was projected into this same space providing a visual interpretation of the boundary with respect to our data.

Results

Plots of the ROC area vs. enhancement thresholds for both the SER and SVM based CAD methods are provided in figure 3 (upper left). An ideal CAD system would maximize the ROC area. A plot of how the ratio of assumed true positive pixels to assumed false positive pixels (ATPP/AFPP) varies with enhancement thresholds for both the SER and SVM methods is provided in figure 3 (upper right). An ideal CAD system would maximize the ATPP/AFPP ratio. A plot of any test's maximum possible sensitivity is also provided in figure 3 (bottom).

The results show ROC areas peaking at low enhancement thresholds for both the SER and SVM methods. The SER method's ROC area begins to decline once the enhancement threshold is raised above 30%. The SVM method's ROC area begins to decline once the enhancement threshold is raised above 70% indicating that it is less affected by low enhancement threshold variations. The SER method is marked by a significant reduction in the ratio (ATPP/AFPP) for low enhancement thresholds indicating that very low enhancement threshold values can yield many false positive pixel predictions. The final plot (figure 3, bottom) demonstrates a consistent decline in potential test sensitivity as the enhancement threshold is increased.

To determine the effect enhancement thresholds are having on our CAD tests we have created classifier visualization plots [25]. Figure 4 provides plots of our breast MR data in 2D principal component projection space. A red X is representative of a single assumed malignant pixel's signal-intensity time curve and a green O is representative of a single assumed benign pixel's curve. Interpretation of these plots is simplified when we consider the lower left area of the plot as a prediction of cancer and the other plot regions as a non-cancerous prediction. The boundaries superimposed on the data indicate the border between a malignant and benign prediction for the given decision method. Solid blue contour lines represent the SER classifier boundaries and solid red contour lines represent the SVM classifier boundaries. Threshold boundaries were provided in black. The principal component projection plots (figure 4) assist in the interpretation of the plots from figure 3. Figure 4 (left) demonstrates how increases in the enhancement threshold will result in increases in the number of cancerous SI-time curves that are incorrectly classified as benign. Figure 4 (right) demonstrates the SER and SVM methods at more clinically realistic thresholds. We have also provided a CAD processed MR examination image in figure 5 to assist in the visual interpretation of what happens to CAD processed images at different enhancement thresholds. Figure 5 consists of an examination containing an invasive ductal carcinoma (IDC) diagnosed by an SVM based CAD system at a variety of thresholds, which can prevent the system from detecting the full size of the lesion. The IDC lesion in figure 5 has a CAD detected volume size of 18.25 mm³ at a 0% threshold, 9.13 mm³ at a 75% threshold and 3.04 mm³ at a 100% threshold (just 16.7% of the lesion volume at 0% threshold). Although we do not have ground truth as to what the correct volume should be, these results demonstrate how the CAD detected volume size diminishes as the enhancement threshold increases. The SER method, subjected to the same threshold limitations as has been demonstrated with the SVM CAD method (figure 5), would also be unable to correctly diagnose the full size of the lesion.

Discussion

We have demonstrated that the SER technique performs poorly at low enhancement threshold settings as it produces many assumed false positive pixels (AFPP) and this could lead a radiologist to ignore CAD processed results altogether. There is a tradeoff between maximizing the ROC area and maximizing the ATPP/AFPP ratio. It was observed that as the enhancement threshold increased, our AFPP decreased as more pixels representing benign tissue were correctly classified as benign. This observation is inline with published results from similar experiments which demonstrated a reduction in false positives at increasing enhancement thresholds [6, 7]. We have also determined that the enhancement threshold should not be set higher than 100% as this limits the CAD technique's maximum possible sensitivity (figure 3 bottom).

We have also shown how increases in the threshold can compromise a test's ability to correctly diagnose the full size of a lesion. This is significant as a reduction in the size of a CAD detected lesion may influence a radiologist's decision on a patient's care, and could result in a missed cancer. It can be seen that at a threshold of just 75%, much of the lesion tissue has been classified as non-cancerous. Although maximum potential test sensitivity (figure 3 bottom) is not compromised until the threshold is raised above 100%, it is clear that some of the less enhancing areas of our cancerous lesions are adversely affected by substantially lower thresholds. Enhancement thresholds of 0–60% are most appropriate for the SVM method as it is unaffected by low threshold variations. Raising the threshold above 60% will leave the SER method unable to diagnose the full size of the lesion (as with the SVM method). These recommendations are in contrast to numerous prior studies that have made use of higher enhancement thresholds [6–17]. Additionally, the SER method suffers from poor ratio values (ATPP/AFPP) for low enhancement thresholds (figure 3 upper right), indicating that its most appropriate threshold would be 50–60%. The SER method is well defined in the literature [18, 23] and so these enhancement threshold recommendations will generalize well to other SER based CAD systems, however, there is considerable flexibility in the use of SVMs in a breast MR CAD system and as such our recommendations may not generalize so reliably. Based on the results provided in this study, the SVM method yields higher ROC curve areas than the SER method at clinically realistic enhancement thresholds. It is hoped that the provided data will assist in the research and use of CAD systems for breast cancer screening by DCE-MRI. Future work may entail repeating the outlined experiments on images from additional breast MR acquisition protocols.

It should be noted that in this study we have not addressed the effect of having a radiologist interpret the final CAD processed images, but have limited our analysis to the evaluation of the diagnostic accuracy of the CAD tests alone. Future work will also address this issue by conducting a blinded reader study in order to draw firm conclusions regarding the appropriate selection of the enhancement threshold. However, it should be noted that it is important to have optimized technical criteria such as the enhancement threshold prior to committing to an expensive human observer study. It is also possible that more reliable results will be obtained when using a larger or more diverse MR data set.

Acknowledgments

The authors would like to thank Don Plewes and Elizabeth Ramsay for their assistance in image acquisition. The authors would also like to thank Mike Froh and Xiaogang Wang for their assistance in image registration. The authors would also like to thank funding support received from the Canadian Breast Cancer Foundation, the Canadian Breast Cancer Research Alliance and the Canadian Institute for Health Research.

References

1. Warner E, Plewes DB, Hill KA, et al. Surveillance of BRCA1 and BRCA2 Mutation Carriers With Magnetic Resonance Imaging, Ultrasound, Mammography, and Clinical Breast Examination. *J Am Med Assoc.* 2004; 292(11):1317–1325.
2. Lehman CD, Gatsonis C, Kuhl CK, et al. MRI Evaluation of the Contralateral Breast in Women with Recently Diagnosed Breast Cancer. *N Engl J Med.* 2007; 356(13):1295–1303. [PubMed: 17392300]
3. Pediconi F, Catalano C, Roselli A, et al. Contrast-enhanced MR Mammography for Evaluation of the Contralateral Breast in Patients with Diagnosed Unilateral Breast Cancer or High-Risk Lesions. *Radiology.* 2007; 243(3):670–680. [PubMed: 17446524]
4. Lee SG, Orel SG, Woo IJ, et al. MR Imaging Screening of the Contralateral Breast in Patients with Newly Diagnosed Breast Cancer: Preliminary Results. *Radiology.* 2003; 226(3):773–778. [PubMed: 12601182]
5. Warren R, Hayes C, Pointon L, et al. A test of performance of breast MRI interpretation in a multicentre screening study. *Magn Reson Imaging.* 2006; 24(7):917–929. [PubMed: 16916709]
6. Williams TC, DeMartini WB, Partridge SC, Peacock S, Lehman CD. Breast MR Imaging: Computer-aided Evaluation Program for Discriminating Benign from Malignant Lesions. *Radiology.* 2007; 244(1):94–103. [PubMed: 17507720]
7. Lehman CD, Peacock S, DeMartini WB, Chen X. A New Automated Software System to Evaluate Breast MR Examinations: Improved Specificity Without Decreased Sensitivity. *Am J Roentgenol.* 2006; 187(1):51–56. [PubMed: 16794155]
8. Joe, BN., Vahidi, K., Gibbs, J., Ying, Lu, Proctor, E., Hylton, N. Automated Breast MR Tumor Volume Measurements for Prediction of Outcome in Breast Cancer Patients on Neoadjuvant Chemotherapy. Proceedings of the Meeting of the Radiologic Society of North America; 2006.
9. Szabo BK, Aspelin P, Wiberg MK. Neural Network Approach to the Segmentation and Classification of Dynamic Magnetic Resonance Images of the Breast: Comparison With Empiric and Quantitative Kinetic Parameters. *Acad Radiol.* 2004; 11(12):1344–1354. [PubMed: 15596372]
10. DeMartini WB, Lehman CD, Peacock S, Russell M. Computer-Aided Detection Applied to Breast MRI: Assessment of CAD-Generated Enhancement and Tumor Sizes in Breast Cancers Before and After Neoadjuvant Chemotherapy. *Acad Radiol.* 2005; 12(7):806–814. [PubMed: 16039534]
11. Partridge SC, Gibbs JE, Lu Y, et al. MRI Measurements of Breast Tumor Volume Predict Response to Neoadjuvant Chemotherapy and Recurrence-Free Survival. *Am J Roentgenol.* 2005; 184(6): 1774–1781. [PubMed: 15908529]
12. Liu PF, Debatin JF, Caduff RF, Kael G, Garzoli E, Krestin GP. Improved diagnostic accuracy in dynamic contrast enhanced MRI of the breast by combined quantitative and qualitative analysis. *Br J Radiol.* 1998; 71(845):501–509. [PubMed: 9691895]
13. Penn A, Thompson S, Brem R, et al. Morphologic Blooming in Breast MRI as a Characterization of Margin for Discriminating Benign from Malignant Lesions. *Acad Radiol.* 2006; 13(11):1344–1354. [PubMed: 17070452]
14. Kuhl CK, Bieling HB, Gieseke J, et al. Healthy Premenopausal Breast Parenchyma in Dynamic Contrast-enhanced MR Imaging of the Breast: Normal Contrast Medium Enhancement and Cyclical-Phase Dependency. *Radiology.* 1997; 203(1):137–144. [PubMed: 9122382]
15. Mussurakis S, Gibbs P, Horsman A. Primary Breast Abnormalities: Selective Pixel Sampling on Dynamic Gadolinium-enhanced MR Images. *Radiology.* 1998; 206(2):465–573. [PubMed: 9457201]

16. Kuhl CK, Schild HH, Morakkabati N. Dynamic Bilateral Contrast-enhanced MR Imaging of the Breast: Trade-off between Spatial and Temporal Resolution. *Radiology*. 2005; 236(3):789–800. [PubMed: 16118161]
17. Fischer DR, Baltzer P, Malich A, et al. Is the ‘blooming sign’ a promising additional tool to determine malignancy in MR mammography? *Eur Radiol*. 2004; 14(3):394–401. [PubMed: 14517688]
18. Niemeyer TL, Wood C, Stegbauer KC, Smith JP. Comparison of automatic time curve selection methods for breast MR CAD. *Proceedings SPIE Medical Imaging*. 2004:785–790.
19. Tzacheva AA, Najarian K, Brockway JP. Breast Cancer Detection in Gadolinium-Enhanced MR Images by Static Region Descriptors and Neural Networks. *J Magn Reson Imaging*. 2003; 17(3): 337–342. [PubMed: 12594724]
20. Causer P, Piron C, Jong R, et al. MR Imaging-guided Breast Localization System with Medial or Lateral Access. *Radiology*. 2006; 240(2):369–379. [PubMed: 16775219]
21. Greenman RL, Lenkinski RE, Schnall MD. Bilateral imaging using separate interleaved 3D volumes and dynamically switched multiple receive coil arrays. *Magn Reson Med*. 1998; 39(1): 108–115. [PubMed: 9438444]
22. Martel AI, Froh MS., Brock KK., Plewes DB., Barber DC. Evaluating an optical-flow based registration algorithm for contrast-enhanced magnetic resonance imaging of the breast. *Phys Med Biol*. 2007; 52(13):3803–3816. [PubMed: 17664578]
23. Hylton N. Dynamic Contrast-Enhanced Magnetic Resonance Imaging As an Imaging Biomarker. *J Clin Oncol*. 2006; 24(20):3293–3298. [PubMed: 16829653]
24. Vapnik, VN. *The Nature of Statistical Learning Theory*. 2. New York, NY: Springer-Verlag; 1999. Support Vector (SV) Machines; p. 138-145.
25. Levman J, Leung T, Causer P, Plewes DB, Martel AL. Classification of dynamic contrast-enhanced magnetic resonance breast lesions by support vector machines. *IEEE Trans Med Imaging*. 2008; 27(5):688–696. [PubMed: 18450541]
26. Nattkemper TW, Arnrich B, Lichte O, et al. Evaluation of radiological features for breast tumour classification in clinical screening with machine learning methods. *Artif Intell Med*. 2005; 34(2): 129–139. [PubMed: 15894177]
27. Chang CC, Lin CJ. LIBSVM – A library for support vector machines [Online]. Accessed April, 2008 <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
28. Eng J. Receiver Operating Characteristic Analysis: A Primer. *Acad Radiol*. 2005; 12(7):909–916. [PubMed: 16039544]

Fig. 1. Block diagram illustrating the role of the enhancement threshold in breast MRI CAD systems.

Fig. 2.
Example signal intensity time curves with respect to the enhancement threshold. Curves C and D are below the threshold and therefore will both be labeled benign by the CAD system.

Fig. 3. Enhancement threshold vs. ROC area for SER & SVM (upper left), threshold vs. ratio (ATPP/AFPP) for SER & SVM (upper right) and threshold vs. maximum test sensitivity (bottom).

Fig. 4.

Principal component space plots for SER (solid blue line) and SVM (solid red line) classifiers along with threshold boundaries (solid black). The lower left area of each plot represents a malignant prediction, other regions are non-malignant predictions and lines mark boundaries between predictions for the given methods.

Fig. 5. A sagittal image with an invasive ductal carcinoma (A – see arrow) magnified and diagnosed by the SVM method at 0% (B), 75% (C) and 100% (D) thresholds. The images were acquired using the MR protocol described in the Methods.