A Margin Sharpness Measurement for the Diagnosis of Breast Cancer from Magnetic Resonance Imaging Examinations

Jacob E. D. Levman, PhD, Anne L. Martel, PhD

Rationale and Objectives: Cancer screening by magnetic resonance imaging (MRI) has been shown to be one of the most sensitive methods available for the early detection of breast cancer. There is high variability in the diagnostic accuracy of radiologists analyzing the large amounts of data acquired in a breast MRI examination, and this has motivated substantial research toward the development of computer-aided detection and diagnosis systems. Most computer-aided diagnosis systems for breast MRI focus on dynamic information (how a lesion's brightness changes over the course of an examination after the injection of a contrast agent). The inclusion of lesion margin measurements is much less common. One characteristic of malignant tumors is that they grow into neighboring tissues. This growth creates tumor margins that are variably fuzzy or diffuse (ie, they are not sharp).

Materials and Methods: In this short report, the authors present a new method for measuring a tumor's margin from breast MRI examinations and compare it with an existing mathematical technique for margin measurements.

Results: The proposed method can yield a test with sensitivity of 77% (specificity, 65%) on screening data, outperforming existing mathematical lesion margin measurement methods. Furthermore, when the presented margin measurement is combined with existing dynamic features, there is a statistically significant improvement in computer-aided diagnosis test performance ($P < .0014$).

Conclusions: The proposed method for measuring a tumor's margin outperforms existing mathematical methods on an extremely challenging data set containing many small lesions. The technique presented may be useful in discriminating between malignant and benign lesions in the context of the computer-aided diagnosis of breast cancer from MRI.

Key Words: Margin; computer aided diagnosis; breast cancer; MRI.

Genetic mutations on the BRCA1/2 genes can result in an 85% lifetime risk for developing breast cancer (1). Regular screening has been identified as key for improving survival rates (2). Dynamic contrast-enhanced magnetic resonance imaging (MRI) has been shown to be the most sensitive screening methodology for detecting breast cancer when compared with x-ray mammography, ultrasound, and clinical breast examination (3). It has also been demonstrated that in 3% of women diagnosed with unilateral breast cancer by mammography, pretreatment MRI examinations detect mammographically occult cancer in the contralateral breast (4–6). These studies suggest that MRI-based breast screening is likely to play a significant clinical role in the future. Furthermore, the American Cancer Society has recommended that women with a lifetime risk for developing breast cancer of 20% to 25% or greater should undergo MRI-based screening (7). It has been shown that there is a high degree of variability among trained radiologists in their ability to correctly diagnose lesions from breast MRI examinations (8). Breast MRI examinations typically involve the acquisition of hundreds of images, compared to just four images for typical x-ray mammographic screening. The data set used in this study is also abnormally challenging, containing lesions as small as 2 to 3 mm. This provides motivation for the research, design, and development of computer-aided detection and diagnosis systems to assist radiologists in identifying malignancies from these images.

When analyzing a breast magnetic resonance image set, a radiologist will visually inspect the examination for a number of signs of malignancy. During the examination, the patient is imaged and then injected with a contrast agent, followed by more imaging. It is known that certain patterns in the changes in lesion brightness over time can be indicative of cancer, and this forms one of the main features that radiologists look for

Acad Radiol 2011; 18:1577–1581

From the Department of Medical Biophysics, Sunnybrook Health Sciences Centre, University of Toronto, 2075 Bayview Avenue, Room S605, Toronto, ON M4N 3M5, Canada. Received February 7, 2011; accepted August 9, 2011. This research project was supported by the Canadian Breast Cancer Foundation (Toronto, ON, Canada) and the Canadian Institute for Health Research (Ottawa, ON, Canada). **Address correspondence to:** J.E.D.L. e-mail: jacob.levman@sri.utoronto.ca

doi:10.1016/j.acra.2011.08.004

when making diagnoses from breast MRI examinations. Radiologists also look for spiculated lesions (or generally irregularly shaped lesions), heterogeneous tissue vascularization, and diffuse tumor edges, all of which are indicative of cancer and can affect their final diagnoses on the basis of the Breast Imaging Reporting and Data System reporting lexicon. Assessing tumor characteristics on the basis of the visual assessment of a radiologist is susceptible to human error, highlighting the need for automated methods for characterizing potentially malignant lesions. Although many methods have been presented in the literature for extracting various measurements of potential malignancy, relatively few methods have been developed for measuring the sharpness or diffuseness of a tumor's margins, resulting in a gap in the literature on computer-aided diagnosis (CAD) for breast MRI. In this paper, we present a new method for measuring the sharpness and variability of a tumor's margin.

CAD systems for breast MRI tend to focus on dynamic information (how a lesion's brightness changes over the course of the examination), and the inclusion of measurements of the sharpness of a lesion's margin is much less common. Cancerous tumors are characterized by their growth into neighboring tissues. This growth creates tumor margins that are not sharp; the margins typically present as variably fuzzy or diffuse. In this short paper, we present a new method for measuring a tumor's margin from breast MRI examinations and compare the technique with preexisting mathematical breast MRI lesion margin measurements (9,10). Margin information can also be visually assessed by a radiologist, with the results fed directly into a CAD system, but this approach does not support automatic CAD systems and is subject to human error.

The main competing heuristic-based technique presented in the literature is known as the blooming sign (11). The blooming sign is an image feature that a CAD system can identify in the diagnosis of potential malignancies. In the early phase of an examination, a malignant lesion's edges are sharp, but as the examination continues, its edges become more diffuse. This is described as being analogous to a pill in liquid whose edges dissolve over time, thus becoming fuzzy. Unfortunately, heuristic approaches are more challenging to reproduce than purely mathematical approaches. Furthermore, heuristic approaches tend to have variable test performance with respect to changes in the heuristic settings of the algorithm. Because of these shortcomings, we have elected to compare our approach exclusively with the quantitative mathematical measures presented by Gilhuijs et al (9,10).

MATERIALS AND METHODS

Image Acquisition and Preprocessing

Between 1997 and 2009, 550 women at high risk 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. Ethics approval for this retrospective study was obtained from the institutional review board of the participating institution. Annual screening of the patient population has resulted in 1749 breast MRI examinations, from which 200 radiologically suspicious lesions were included in this retrospective analysis.

The screening protocol used was as follows. Simultaneous bilateral MRI was performed using a 1.5-T magnet (GE Signa version 11.4; GE Healthcare, Milwaukee, WI). Sagittal images were obtained with a phased-array coil arrangement using a dual-slab interleaved bilateral imaging method (12). This provided three-dimensional volume data over each breast obtained with a radiofrequency spoiled gradient-recalled sequence (repetition time, 18.4 ms; echo time, 4.3 ms; flip angle, 30°; 256 × 256 × 32 voxels; field of view, 18 × 18 × 6–8 cm). Imaging was performed before and after a bolus injection of 0.1 mmol/kg of contrast agent (gadolinium diethylenetriamine penta-acetic acid). Each bilateral acquisition was obtained in 2 minutes 48 seconds. Slice thickness was 2 to 3 mm. A total of 200 dynamic contrast-enhanced MRI breast lesions from patients at high risk were obtained. This retrospective analysis analyzed lesions pathologically proven to be malignant (43 cases) or benign (157 cases). Ground truth for malignant lesions was based on the analysis of tissue biopsies by a histopathologist. When the histopathologist determined a tissue sample to be noncancerous, a benign diagnosis was accepted. In cases in which a suspicious mass did not receive a biopsy but returned to screening without observed changes to the lesion for >1 year, a benign diagnosis was also accepted.

Image registration is the process of aligning images that vary in position over time. This is performed to compensate for any patient motion that occurs during the examination. For this study, we used a three-dimensional nonrigid registration technique for magnetic resonance breast images that is applied globally on the breast examination (13). To perform measurements of the sharpness of the edge of each radiologically identified lesion, segmentation was performed to create a region of interest from which to extract these edge sharpness measurements. This segmentation was accomplished by the enhancement threshold method, whereby any voxel whose signal intensity increased by more than a set threshold over the precontrast signal intensity was included as part of our lesion. Each lesion was segmented semiautomatically with the enhancement threshold by a researcher with 7 years of experience analyzing breast MRI examinations. The threshold used for segmentation was variable on the basis of the lesion and set by the researcher performing the segmentations, although the majority of the lesions could be successfully segmented at 60% enhancement (14). Necrotic cores of tissues that did not enhance past the provided threshold were not included as part of the initial lesion segmentation. Each segmentation was visually inspected to confirm that an acceptable region of interest was produced.

Proposed Edge Sharpness Measurement

The proposed method presented here is computed both from those voxels that are close to the edge of the lesion (but still in the lesion) and those that immediately neighbor the lesion (but are outside the lesion). This is illustrated in Figure 1, which demonstrates the 1-pixel rings on which these measurements are made, colored blue and red for outside and inside the lesion, respectively. In the proposed approach, the set of voxels in the inner ring (1-voxel-thick shell; marked red in Fig 1) are represented by the set r_i. Each member of r_i yields a set of gradient values with neighboring voxels from the outer (blue) ring, which form a set of voxel pairs that are averaged in the numerator of the equation

$$\frac{\overline{\mathrm{RSI}_{\mathrm{in}}(r_i, t) - \mathrm{RSI}_{\mathrm{out}}(n(r_i), t)}}{d}, \quad (1)$$

where $\mathrm{RSI}_{\mathrm{group}}$ is the relative signal intensity (normalized by the precontrast signal intensity) of a voxel in the inner or outer group, a function of position (r_i) and time (t); r_i is a voxel position marker in three-dimensional coordinates (the inner ring); $n(r_i)$ is the six-connect three-dimensional neighborhood operator, which provides a set of voxel positions that neighbor r_i but are outside the lesion region of interest; t is the time point on which the margin measurement is being made (ie, a user parameter); t_f is the final time point of the examination; and d is the normalization term $\overline{\mathrm{RSI}_{\mathrm{in}}(r_i, t) - \mathrm{RSI}_{\mathrm{out}}(n(r_i), t_f)}$.

Equation 1 provides the definition for the proposed margin measurement. The averaging operator in the numerator of the equation can also be changed to be a standard deviation spread operator, which changes the measurement from a simple margin sharpness measurement to a margin variability measurement.

Evaluation Technique: Existing Edge Sharpness Measurements

The most commonly used mathematical methods for measuring a breast MRI lesion's edge sharpness were first presented by Gilhuijs et al (9,10). Their technique is computed by examining the 3-voxel shell that marks the outer contour of our lesion (as seen in green in Fig 1). Gilhuijs et al's two margin measurement methods are based on the average and the variance of the spatial gradient of the subtracted enhancement data from the green shell in Figure 1. Detailed definitions of the techniques are available in the literature (9,10). Descriptions of the similarities and differences between our proposed method and Gilhuijs et al's methods are provided in the "Discussion" section.

Validation and Statistical Analyses

The two methods proposed by Gilhuijs et al (9,10) and the method proposed here were evaluated using receiver-operating characteristic (ROC) curve analysis, which measures the robustness of a particular test by evaluating the trade-off between the test's sensitivity and specificity (15). Wilcoxon's signed-rank test was performed to evaluate the benefits of including the proposed margin measurement in a CAD system using established dynamic feature measurements.

RESULTS

The ROC curve area results for each of the methods addressed are provided in Figure 2. The proposed method at a time point t of 2 minutes 48 seconds yielded the most robust test with the mean operation in the numerator of the equation (ROC curve area, 0.72). This outperformed Gilhuijs et al's (9,10) best margin measurement, which yielded an ROC curve area of 0.62. The proposed method can yield test sensitivity and specificity of 77% and 65%, respectively. Sensitivity and specificity for Gilhuijs et al's best-performing method were 60% and 62%. Thus, the proposed method yields improvements in both the test's sensitivity and the test's specificity over the best-performing existing mathematical margin measurement. When the proposed margin measurement is added to three established dynamic feature measurements (enhancement, time to peak enhancement, and washout) (16), a statistically significant improvement in test performance is observed with Wilcoxon's signed-rank test ($P < .0014$) when comparing bootstrap validation distributions of ROC curve areas with and without the proposed margin measurement. Example malignant and benign lesions exhibiting variable and nonvariable margins are provided in Figure 3.

DISCUSSION

The main similarity between the proposed technique and those presented by Gilhuijs et al (9,10) is that all attempt to quantify a malignant lesion's margin by measuring the differences between neighboring voxels. Gilhuijs et al's method looks at all the neighboring differences within the 3-voxel ring that marks a lesion's outer shell, whereas our proposed technique measures neighboring differences only between the two single-voxel shells that mark the edge of the lesion. This is illustrated in Figure 1, which demonstrates the proposed comparison being performed between the red voxel shell within the lesion and the blue voxel shell immediately outside of the lesion. The data set used in this research paper contains many very small breast cancers (range, 2 mm–6 cm across). This helps explain why the 3-voxel-thick method underperforms (a 3-voxel-thick shell on a small lesion often includes the whole lesion). Thus, using 1-voxel-thick shells may be a superior method for extracting a margin measurement from small malignancies.

The main difference between the two proposed approaches is the method for normalizing the spatial gradient information. Gilhuijs et al's (9,10) technique combines all of the time point information available (rather than having a time point parameter) but in doing so appears to blur together

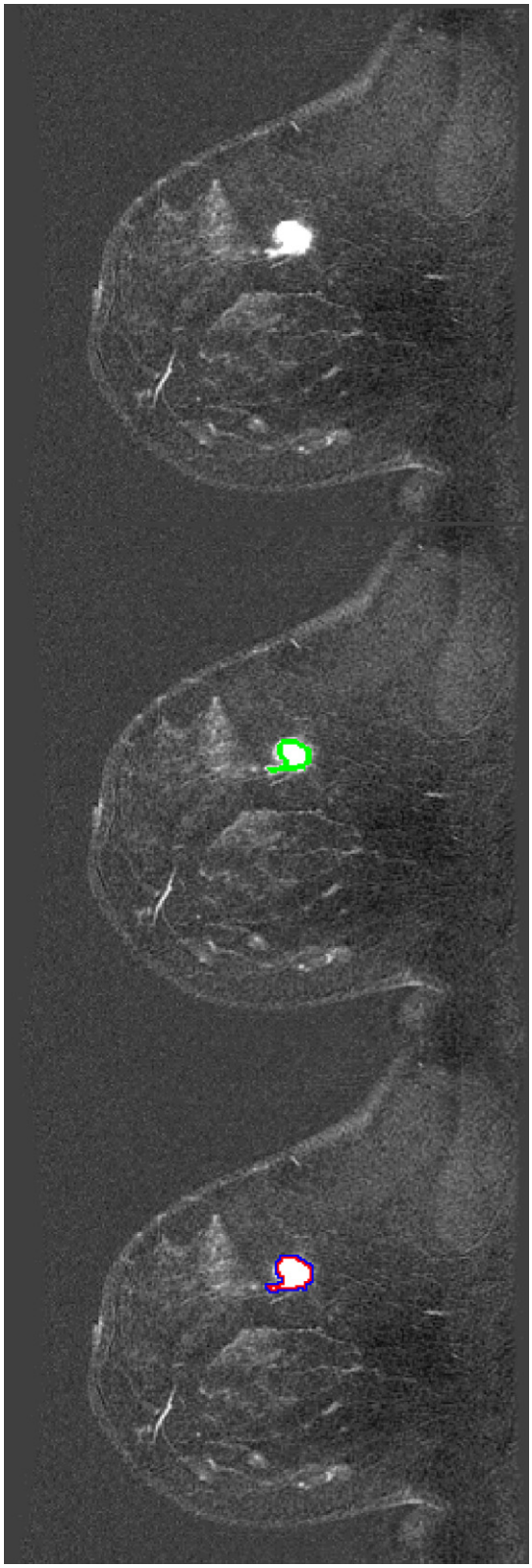

Figure 1. Example malignant lesion (*top*) with a 3-pixel-thick shell marking where the measurement is performed for the existing technique (*middle image, green contour*) and two 1-pixel-thick shells marking where margin measurements are performed for the proposed technique (*bottom image, red and blue contours*).

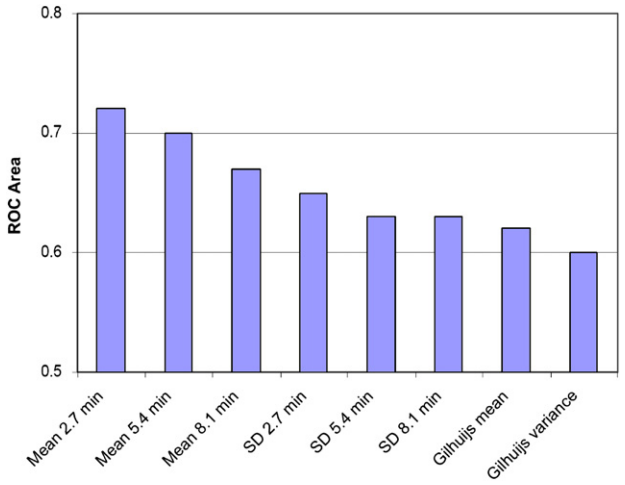

Figure 2. Receiver-operating characteristic (ROC) curve areas for the various margin measurements addressed in this paper. SD, standard deviation.

some of the discriminatory information available, whereas our approach normalizes by the final time point's margin gradient. This normalization was selected because it yielded the most separation between our malignant and benign lesions.

The results presented depend on the time point at which the margin measurement is performed. The results indicate that quantifying the sharpness or variability of the tumor's margin is best performed earlier in the examination and that results degrade as the time point parameter is increased. It is expected that this occurs because more of the contrast agent in cancerous lesions diffuses out toward neighboring tissues as the exam progresses (thus reducing tumor margin variability). It is possible that the proposed approach (or a similar one) might benefit from having the measurement taken before 2 minutes 48 seconds. The temporal resolution of our screening protocol is large (2 minutes 48 seconds), so better discrimination between malignant and benign lesions may be possible at earlier time points. Future work will address this issue by testing the proposed technique on breast MRI examinations with a much finer temporal resolution.

Because both the proposed technique and Gilhuijs et al's (9,10) technique are based on spatial gradients, it is expected that both methods are dependent on the spatial resolution of the MRI acquisition protocol. Future work will look at testing these methods on breast MRI examinations with widely varying spatial resolutions.

Ultimately, we are interested in whether adding the proposed margin measurement helps improve CAD of breast cancer from MRI examinations. To this end, we performed bootstrap validation trials comparing a breast MRI CAD system with three dynamic features to a breast MRI CAD system with three dynamic features plus the proposed margin feature. Bootstrap validation trials produced a statistically significant improvement for the CAD system incorporating the proposed margin measurement compared with a CAD system based purely on dynamic features as evaluated by

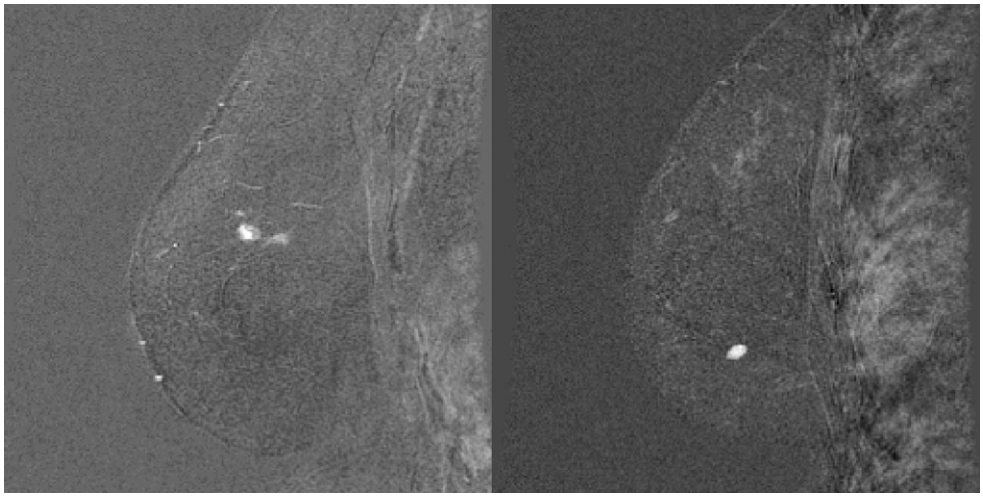

Figure 3. Example malignant lesion exhibiting diffuse and variable margin sharpness (*left*) and an example benign lesion exhibiting sharper margins with lower variability (*right*).

Wilcoxon's signed-rank test ($P < .0014$). This analysis indicates that the proposed margin measurement has the potential to provide significant improvements in breast MRI CAD systems.

CONCLUSIONS

In this short paper, we have presented a new mathematical method for measuring the margins of a lesion from breast MRI examinations. We have shown that our technique can produce sensitivity and specificity of 77% and 65%, respectively, compared to sensitivity and specificity of 60% and 62% for preexisting mathematical margin sharpness measurements. These results indicate that the proposed method may be useful for helping discriminate between malignant and benign lesions in the context of a computer-aided detection and diagnosis system. Future work will further test this approach on independent breast MRI data sets.

ACKNOWLEDGMENTS

We would like to thank Gary Wang for his assistance in image registration. We would like to thank Kenneth Gilhuijs for correspondence, which helped clarify the methods used in existing techniques. The MRI data were provided by Dr Ellen Warner and were acquired using funding from the Canadian Breast Cancer Research Alliance.

REFERENCES

1. Ford D, Easton DF, Stratton M, et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. Am J Hum Genet 1998; 62:676–689.
2. Curry SJ. Fulfilling the Potential of Cancer Prevention and Early Detection. Washington, DC: National Academies Press, 2003.
3. Warner E, Plewes DP, Hill KA, et al. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. JAMA 2004; 292: 1317–1325.
4. Lehman C, Gatsonis C, Kuhl CK, et al. MRI evaluation of the contralateral breast in women with recently diagnosed breast cancer. N Engl J Med 2007; 356:1295–1303.
5. 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: 670–680.
6. 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; 226:773–778.
7. Saslow D, Boetes C, Burke W, et al., American Cancer Society Breast Cancer Advisory Group. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. Cancer J Clin 2007; 57:75–89.
8. 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; 24:917–929.
9. Gilhuijs K, Giger ML, Bick U. Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging. Med Phys 1998; 25:1647–1654.
10. Gilhuijs K, Giger ML, Bick U. Automated feature extraction and classification of breast lesions in magnetic resonance images. Proc SPIE 1998; 3338:294–300.
11. 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; 13:1344–1354.
12. Greenman RL, Lenkinski RE, Schnall MD, et al. Bilateral imaging using separate interleaved 3D volumes and dynamically switched multiple receive coil arrays. Magn Reson Med 1998; 39:108–115.
13. Martel AL, Froh MS, Brock KK, et al. Evaluating an optical-flow-based registration algorithm for contrast-enhanced magnetic resonance imaging of the breast. Phys Med Biol 2007; 52:3803–3816.
14. Levman J, Causer P, Warner E, et al. Effect of the enhancement threshold on the computer-aided detection of breast cancer using MRI. Acad Radiol 2009; 16:1064–1069.
15. Eng J. Receiver operating characteristic analysis: a primer. Acad Radiol 2005; 12:909–916.
16. Levman J, Leung T, Causer P, et al. Classification of dynamic contrast-enhanced magnetic resonance breast lesions by support vector machines. IEEE Trans Med Imaging 2008; 27:688–696.