Mammographic quantitative image analysis and biologic image composition for breast lesion characterization and classification

Karen Drukker[a)]
Department of Radiology, University of Chicago, Chicago, Illinois 60637

Fred Duewer
Radiology Department, University of California, San Francisco, California 94143

Maryellen L. Giger
Department of Radiology, University of Chicago, Chicago, Illinois 60637

Serghei Malkov
Radiology Department, University of California, San Francisco, California 94143

Chris I. Flowers
Department of Radiology, University of South Florida, Tampa, Florida 33612

Bonnie Joe and Karla Kerlikowske
Radiology Department, University of California, San Francisco, California 94143

Jennifer S. Drukteinis
Department of Radiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida 33612

Hui Li
Department of Radiology, University of Chicago, Chicago, Illinois 60637

John A. Shepherd
Radiology Department, University of California, San Francisco, California 94143

(Received 16 August 2013; revised 3 February 2014; accepted for publication 4 February 2014; published 27 February 2014)

Purpose: To investigate whether biologic image composition of mammographic lesions can improve upon existing mammographic quantitative image analysis (QIA) in estimating the probability of malignancy.
Methods: The study population consisted of 45 breast lesions imaged with dual-energy mammography prior to breast biopsy with final diagnosis resulting in 10 invasive ductal carcinomas, 5 ductal carcinoma *in situ*, 11 fibroadenomas, and 19 other benign diagnoses. Analysis was threefold: (1) The raw low-energy mammographic images were analyzed with an established in-house QIA method, "QIA alone," (2) the three-compartment breast (3CB) composition measure—derived from the dual-energy mammography—of water, lipid, and protein thickness were assessed, "3CB alone", and (3) information from QIA and 3CB was combined, "QIA + 3CB." Analysis was initiated from radiologist-indicated lesion centers and was otherwise fully automated. Steps of the QIA and 3CB methods were lesion segmentation, characterization, and subsequent classification for malignancy in leave-one-case-out cross-validation. Performance assessment included box plots, Bland–Altman plots, and Receiver Operating Characteristic (ROC) analysis.
Results: The area under the ROC curve (AUC) for distinguishing between benign and malignant lesions (invasive and DCIS) was 0.81 (standard error 0.07) for the "QIA alone" method, 0.72 (0.07) for "3CB alone" method, and 0.86 (0.04) for "QIA+3CB" combined. The difference in AUC was 0.043 between "QIA + 3CB" and "QIA alone" but failed to reach statistical significance (95% confidence interval [–0.17 to + 0.26]).
Conclusions: In this pilot study analyzing the new 3CB imaging modality, knowledge of the composition of breast lesions and their periphery appeared additive in combination with existing mammographic QIA methods for the distinction between different benign and malignant lesion types.
 [http://dx.doi.org/10.1118/1.4866221]

Key words: 4 compartment breast imaging, dual-energy mammography, breast cancer, quantitative image analysis, computer-aided diagnosis

1. INTRODUCTION

The fundamental information in mammography, i.e., the single-energy attenuation of low-energy x-rays, has remained the same since the inception of breast imaging in 1913.[1] Replacement of film with digital detectors[2] has reduced dose and increased dynamic range but with reduced spatial resolution. However, despite large changes in imaging characteristics,

the sensitivity and specificity of digital and film mammography have remained comparable,[3,4] indicating that the technique's sensitivity and specificity are likely limited by physics fundamentals.

Recently it was reported that up to 60% of women undergoing annual screening mammography over a period of 10 years will have a false-positive result.[5,6] Thus, the modest accuracy of mammography can subject many women to unnecessary, and not necessarily harmless, invasive work-up. For example, it is well established that only about 25% of biopsies are breast cancer among women aged 40–79 years undergoing screening mammography.[7] Thus, new techniques are needed to reduce the number of lesions selected for biopsy.

Computer-aided detection for mammography is a method in clinical practice providing radiologists with a "second opinion" on the presence of suspicious lesions that may require biopsy.[8] Computer-aided diagnosis, which provides radiologists with the likelihood that an identified lesion is malignant and should be biopsied, has been limited to research settings. In reader studies it has been shown that computer-aided diagnosis, or more generally, quantitative image analysis (QIA) for mammography may improve diagnostic accuracy.[9,10] QIA methods for mammography are based on information available in 2D single-energy x-ray attenuation images and are hence bound by limitations inherent in mammography.

We developed an extension of x-ray mammography, called Three-Compartment Breast (3CB) composition and based on dual-energy mammography, to quantify the absolute molecular tissue compartments of lipid, protein, and water by introducing two new independent sources of information: high-energy x-ray attenuation and breast tissue thickness distribution. With these three pieces of information (low-energy attenuation, high-energy attenuation, thickness) for every pixel in a breast lesion and its periphery, we created images that contain single molecular compartments of lipid, protein, or water.[11] Different lesion types are expected to exhibit distinct atomic compositions because they typically consist of different proportions of tissue components. Both malignant and benign lesions, e.g., are known to have higher water content than normal breast tissue,[12] and invasive cancers are expected to show high degrees of vascularity. The underlying hypothesis is that 3CB "signatures," i.e., image-based biomarkers, are unique for each lesion type and hence can be used to better identify which lesions require biopsy. In past studies, we developed a method for dual-energy x-ray mammography using a calibration phantom to measure breast density.[13] In parallel, we developed a highly accurate method for modeling breast thickness using a geometric phantom attached to the compression paddle and combined the thickness measure with a screening mammogram to estimate breast density.[14] By combining the two methods, we developed a method to estimate the local composition of imaged breast tissue in terms of its water, lipid, and protein content.[11] In order to estimate water, lipid, and protein content separately, we measured the x-ray attenuation in a standard screening mammography image, the x-ray attenuation in a high-energy/low-dose mammography image, and the thickness using the model derived from our geometric phantom. We then modeled the individual compartment thicknesses from the high and low energy attenuations using a priori calibration data from a previously scanned calibration phantom with 26 different combinations of lipid (wax), water (solid water), and protein (Delrin).[11]

The long-term goal of our effort in 3CB imaging is to determine whether biological diagnostic measures of breast lesions, obtained using widely available standard full-field digital mammography equipment, can help in the diagnosis of breast cancer. The technique may be implemented during both screening and diagnostic mammography with only small changes to the x-ray dose or procedures. In this paper, we present the first in vivo results of biologic image composition of mammographic lesions on women with abnormal breast findings on diagnostic mammography. In addition, we examine whether inclusion of 3CB measures improve the performance of mammographic QIA methods to differentiate benign from malignant breast lesions.

2. METHODS

2.A. Patient population

Images were acquired in a prospective clinical trial under IRB-approved HIPAA-compliant protocols involving patient consent. Fifty women with suspicious findings at mammography (BIRADS 4 and 5 assessment) were recruited to be imaged with dual-energy mammography in order to assess 3CB prior to biopsy. The women were recruited consecutively by a part-time recruiter from January 2010 through December 2011 at a Same Day Imaging Clinic (diagnostic imaging and biopsy on the same day). Exclusion criteria for the study were failure to receive a biopsy, known prior breast cancer, known prior biopsies, protocol failure, inability of our radiologist to identify the breast lesion on the low-energy x-ray attenuation image, and inability to schedule additional imaging prior to biopsy.

Breast biopsies were clinically reviewed by the Pathology Department at the University of California San Francisco. Of the 50 women, twelve were excluded for reasons including; the finding not identified on mammography, lack of an available pathology report, reclassification of the finding as benign (BI-RADS 2) or probably benign (BI-RADS 3) postimaging, or protocol failures including significant movement between the high and low energy mammography images, imaging the wrong breast, imaging a breast where the lesion was not visible in either craniocaudal (CC) or mediolateral-oblique (MLO) view, and an unidentified malfunction in our file transfer protocol. For the 39 remaining women, there were 45 distinct findings of which all but 2 were visible in both CC and MLO views. There were 10 invasive ductal carcinoma (IDC), 5 ductal carcinoma *in situ* (DCIS), 11 fibroadenoma (FA), and 19 benign—other than fibroadenoma—(BN) pathologies. There were five women with more than one finding: Two women with a combination of IDC and DCIS, one woman with a FA and a BN lesion, one woman with two BN lesions, and one woman with three FA lesions.

2.B. Imaging system

A single Hologic Selenia full-field digital mammography system (Hologic, Inc.) was used to image women with 3CB. This system configuration has a molybdenum x-ray anode and two internal x-ray filters of either molybdenum or rhodium. Two mammographic images were acquired on each woman's affected breast using a single compression. The first exposure was made to mimic the clinical screening or diagnostic mammogram conditions such that Selenia's internal software chooses the voltage and current settings based on breast thickness (usually below 30 kVp). The second mammographic image was acquired at a fixed voltage of 39 kVp (the highest attainable voltage on the Selenia unit) and fixed current for all participants. This high-energy exposure was made using an additional 3-mm plate of aluminum in the beam to increase the average energy of the high-energy image. We limited the total dose of this procedure to be approximately 110% of the mean-glandular dose of an average screening mammogram. The calibration standards and 3CB algorithms are described in full elsewhere.[11]

An expert board-certified radiologist with 26 years of experience manually delineated all findings on the low-energy mammography images. The geometric center of these delineations served as the input (together with the image data) to out automated image analysis. Since increased noise is a known limitation of dual-energy decomposition techniques,[11] 2×2 pixel binning was used prior to analysis to reduce noise and increase the signal to noise ratio. This resulted in an image resolution (pixel size) of 140 μm.

2.C. Computerized image analysis

The computerized image analysis was performed retrospectively on deidentified data and consisted of three main steps: (1) QIA using the raw low-energy mammography images alone, "QIA alone," (2) analysis of the 3CB (water, lipid, and protein) images alone, "3CB alone," and (3) a combined analysis of low-energy mammography and 3CB images, "QIA+3CB." The general steps in the analysis were lesion segmentation, lesion characterization through extraction of mathematical descriptors, i.e., features, and subsequent classification for malignancy.

The "QIA alone" analysis served as the reference method and was performed using "established" in-house research methodology. Computerized segmentation was performed on the low-energy mammography images, using a slightly modified version of a previously published method,[15] and subsequently used to characterize findings in those images as well as in the corresponding water, lipid and, protein images (Fig. 1). The features extracted for the "QIA alone" analysis have been described elsewhere[16–18] and included features describing morphology (such as lesion shape and margin characteristics) and texture. It is important to note that often multiple (correlated) mathematical features describe a single physical characteristic. For the analyses including 3CB information, the following features were extracted from the water, lipid, and protein thickness images: the mean, median, standard deviation, and skewness within the segmented region as well as within a surrounding 5-mm-thick border, i.e., the periphery.

Feature selection was performed for "QIA alone," "3CB alone," and "QIA+3CB" features using stepwise multilinear regression for the task of distinguishing between malignant and benign findings. Selected features were subsequently used as input to a neural net classifier to obtain the estimated probability of malignancy (PM) for each finding in leave-one-case-out cross validation. We performed two types of cross validations: The primary approach was leave-one-*lesion*-out (eliminating all data pertaining to a lesion from classifier training during one training/testing step), and the secondary approach was leave-one-*patient*-out (eliminating all data pertaining to a given patient). The latter was performed due to concern about correlation between multiple findings within one patient. In both approaches, the PM of a lesion was calculated as the average of the estimates obtained from the CC and MLO views (when visible in both views).[19] All reported leave-one-case-out results were obtained using the primary approach, unless otherwise noted.

2.D. Performance evaluation

The ability of individual computer-extracted features to distinguish between different finding-types was assessed by comparing characteristics of the four types of findings (IDC versus DCIS versus FA versus BN), and by comparing one finding type to all other finding types combined [e.g., malignant (IDC+DCIS) versus benign (BN+FA) or IDC versus (DCIS+FA+BN)]. For this purpose, we used box plots and Receiver Operating Characteristic (ROC) Analysis[20,21] with the area under the ROC curve (AUC) as the figure of merit. ROC analysis was performed using the *lesion*-based computer-estimated probabilities of malignancy as decision variable. Analysis at this level was done separately for "QIA" and "3CB" features and was intended to uncover trends rather than to perform extensive comparisons.

The performance of the selected features merged by the neural net classifier was assessed on a detailed level comparing the estimated PMs, and on a comprehensive level through ROC analysis. The former analysis included comparison of PMs obtained with "QIA alone" versus "3CB alone" (scatter plot), and "QIA+3CB" versus the reference method "QIA alone," both in the absolute sense (Bland–Altman plot[22]) and in the relative sense (plot of the relative shift in PM). ROC analysis was performed for "QIA alone," "3CB alone," and "QIA+3CB" using the proper binormal model[20] to obtain AUC values. Statistical significance of the difference in AUC between the reference method "QIA alone" and the combined method "QIA+3CB" was determined using the 95% confidence interval obtained through bootstrapping (1000 iterations).

3. RESULTS

Suspicious mammographic findings were found to have highly noticeable compositional changes relative to surrounding tissue in the 3CB compositional images (Fig. 1).

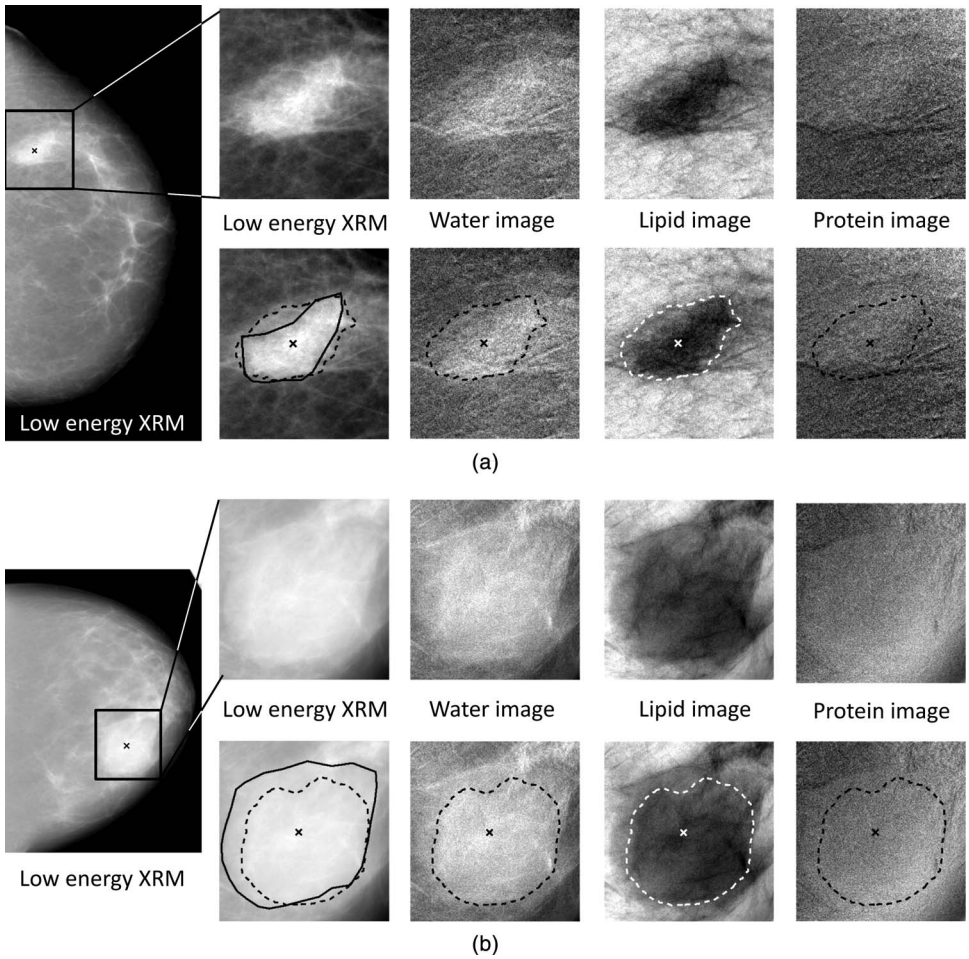

FIG. 1. Example images of (a) an invasive ductal carcinoma (IDC) and (b) a fibroadenoma (FA): Full field CC view (low-energy mammogram) and regions-of-interest of the (1) low-energy x-ray mammogram, (2) water, (3) lipid, and (4) protein images without (top) and with (bottom) the computer-determined lesion boundaries (dashed line), radiologist manual outline (solid line, on low-energy mammogram only), and radiologist-indicated lesion center ("x"). Computer-determined lesion boundaries were used in all calculations presented here.

3.A. Individual features

Features showing the largest overall difference in the median value between the four types of findings—IDC, DCIS, FA, and BN—were entropy for "QIA," and the median water thickness within the periphery surrounding the computer-segmented finding for "3CB," respectively (Fig. 2). Entropy (within the segmented lesion on low-energy mammography)—a measure of disorganization—was on average highest for IDC and lowest for fibroadenomas. The water thickness in tissue surrounding the findings was on average highest for IDC and lowest for the other-type benign lesions (other than fibroadenomas).

Features obtained from "QIA" (low-energy mammography) and "3CB," that individually offered the best pair-wise discrimination between the four lesion types, obtained similar AUC values (Table I). For "QIA" features, both morphological and textural features were amongst the best performing. The main radiologic identifying factor for IDC lesions was spiculation, i.e., IDC lesions tended to be more spiculated than all other lesions. In the distinction between DCIS and benign lesions, a combination of margin sharpness and lesion shape was the most important, with DCIS lesions appearing less circular with less distinct margins (a lower value for the radial gradient measures). Fibroadenomas had on average a smoother appearance than other-type benign lesions (a lower value for the texture measure). For "3CB" features, the best individual performers included features from all components (water, lipid, and protein thickness images). IDC lesions tended to have a less asymmetrical lipid distribution (lower skewness) than DCIS, a larger difference in water thickness between lesion and periphery than fibroadenomas, and more water in their periphery than all other lesions (Fig. 2). DCIS tended to have a larger difference in lipid thickness between the segmented region and its periphery than fibroadenomas, and a more asymmetrical distribution of water in its periphery (higher skewness) than other-type benign lesions. Fibroadenomas had on average higher protein content relative to that in the periphery than other-type benign lesions.

Features from "QIA" and "3CB," that individually offered the best discrimination between one type of finding and all other finding types combined (Table II), obtained AUC values similar to each other and were similar in nature to those in the pair-wise assessment (Table I) but with slightly lower

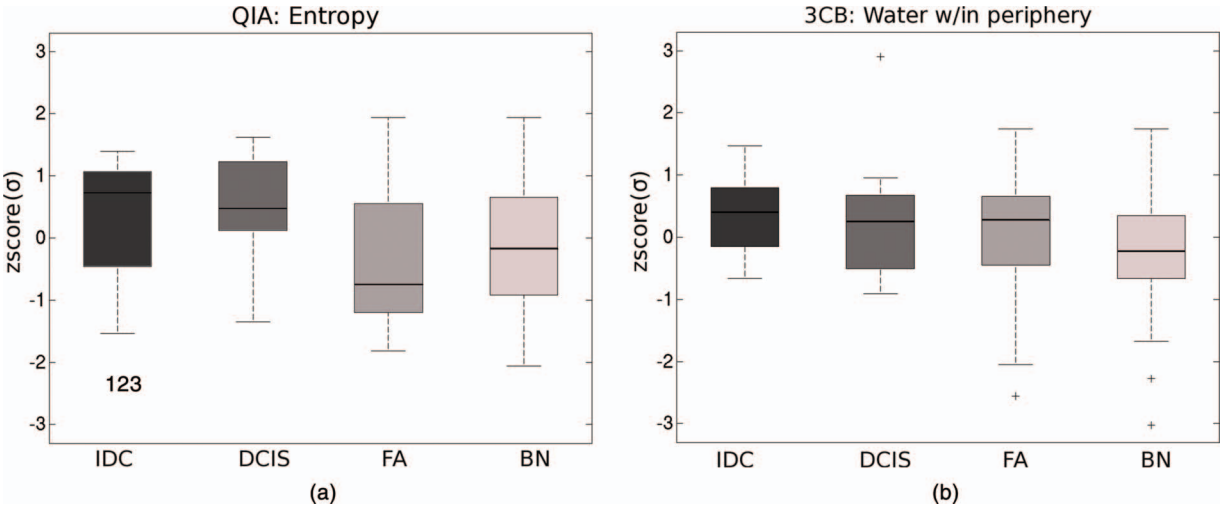

FIG. 2. Boxplots of (a) a QIA feature and (b) a 3CB feature [standardized ("zscore") to zero mean and unit standard deviation σ] selected in stepwise feature selection in the task of distinguishing between malignant and benign lesions. A horizontal line in each box indicates the median value, the box boundaries denote the 25th and 75th percentiles, the whiskers indicate the range excluding outliers, and "+"s mark individual outliers.

performance. In the distinction between all malignant (IDC+DCIS) and all benign (FA+BN) findings, malignant findings tended to have a less homogeneous appearance than benign ones on the low-energy mammography images (higher entropy) with a more asymmetrical water distribution in their periphery (higher skewness). In feature selection, five features were selected for the task of distinguishing between malignant and benign findings, of which four were described above (Tables I and II). The additional selected feature was the "QIA" feature of margin sharpness (Fig. 4), with the malignant findings having more ill-defined margins than benign findings.

3.B. Merged features: Estimated probability of malignancy

The probability of malignancy for all findings, as estimated by the neural net classifier within leave-one-case-out cross-validation, obtained for "QIA alone" and "3CB alone" analyses differed substantially for some cases while agreeing for other ones [Fig. 3(a)]. The obtained values were weakly correlated at best with a correlation coefficient between the PM obtained in the "QIA alone" and "3CB alone" analyses of 0.23 ($p = 0.12$). Combined analysis, "QIA+3CB," appeared beneficial with respect to the reference "QIA alone" [Figs. 3(b) and 3(c)], obtaining on average a higher PM for malignant findings and a lower PM for benign ones. The improvement in PM of the "QIA+3CB" analysis with respect to the "QIA alone" analysis seemed more pronounced for malignant findings than for benign findings [Fig. 3(c)]. The "QIA+3CB" computer-estimated probabilities of malignancy were substantially correlated with those from "QIA alone" and those from "3CB alone" with correlation coefficients of 0.80 ($p \sim 10^{-11}$) and 0.73 ($p \sim 10^{-8}$), respectively.

3.C. Merged features: ROC analysis

The area under the ROC curve (AUC) for the task of distinguishing between benign and malignant lesions (both invasive and *in situ*) was 0.81 (standard error 0.07) for the reference

TABLE I. Pair-wise discrimination between the four different lesion types for "QIA alone" features and "3CB alone" features: Best performing individual features with their AUC value. Note that this table combines information from two symmetrical tables: One table for the best performing QIA features and one for the best performing 3CB features. In the table presented here, entries above the diagonal pertain to QIA features and entries below the diagonal to 3CB features. For example, the best performing 3CB feature for the distinction between BN and IDC was "water: median in periphery" with an AUC value of 0.71, and the best QIA feature for this task was "FWHM lesion" with an AUC value of 0.65 (features selected by stepwise feature selection for the task of distinguishing between malignant and benign lesions indicated in bold font, IDC = invasive ductal carcinoma, DCIS = ductal carcinoma *in situ*, FA = fibroadenoma, BN = benign other than FA).

	N	IDC	DCIS	FA	BN
N		10	5	11	19
IDC	10	...	QIA FWHM[a] border 0.76	QIA FWHM[a] lesion 0.69	QIA FWHM[a] lesion 0.65
DCIS	5	3CB Lipid: Skewness w/in lesion 0.71	...	**QIA Radial gradient margin 0.78**	QIA Radial gradient 0.70
FA	11	3CB Water: Relative[b] σ 0.75	**3CB Lipid: Relative[b] mean 0.80**	...	QIA Texture (σ[c]) 0.69
BN	19	**3CB Water: Median in periphery 0.71**	3CB Water: Skewness in periphery 0.72	3CB Protein: Relative[b] mean 0.73	...

[a]FWHM = full width at half maximum of gray value histogram.
[b]"Relative" indicates the value w/in the lesion with respect to that in the lesion periphery.
[c]σ = standard deviation.

TABLE II. Individual features with the highest AUC in the task of discriminating between N lesions of given pathology and those of all other pathologies (features selected by stepwise feature selection for the task of distinguishing between malignant and benign lesions are indicated in bold font, IDC = invasive ductal carcinoma, DCIS = ductal carcinoma *in situ*, FA = fibroadenoma, BN = benign other than FA).

		QIA		3CB	
	N	**Feature**	**AUC**	**Feature**	**AUC**
IDC	**10**	FWHM[a] border	0.66	**Water: Median w/in periphery**	0.66
DCIS	**5**	Circularity	0.71	Water: Skewness w/in lesion	0.71
FA	**11**	**Radial gradient margin**	0.65	Protein: Relative[b] median	0.70
BN	**19**	Texture (σ[c])	0.64	Protein: Relative mean lesion	0.65
IDC+DCIS	**15**	**Entropy**	0.66	Water: Skewness w/in periphery	0.68

[a]FWHM = full width at half maximum of gray value histogram.
[b]"Relative" indicates the value w/in the lesion with respect to that in the lesion periphery.
[c]σ = standard deviation.

method "QIA alone" [Fig. 4(a)], 0.72 (0.07) for "3CB alone" [Fig. 4(b)], and 0.86 (0.04) for "QIA+3CB" (Fig. 5). The median difference in AUC was 0.043 between "QIA+3CB" and "QIA alone" but the difference in AUC failed to reach statistical significance (95% confidence interval [–0.17 to + 0.26]). The corresponding AUC values obtained in the secondary leave-one-patient-out analysis, were 0.80 (0.07), 0.70 (0.07), and 0.84 (0.05) for "QIA alone, "3CB alone," and "QIA+3CB," respectively.

4. DISCUSSION

This pilot study was based on the first *in vivo* application of three-compartment breast imaging and investigated the utility of biological measures of local breast tissue composition in the distinction between different lesion types and as an adjunct to QIA for the diagnosis of breast cancer. Analysis of lesions and their periphery using the 3CB images of water, lipid, and protein thickness showed promising performance. It was interesting to note that many of the 3CB features important in the distinction between different lesion types pertained to the 5-mm-wide lesion periphery, i.e., the surrounding parenchyma (within the accuracy of the computerized lesion segmentation method). These features appear to reflect angiogenesis for the malignant lesions (water content), the presence of less fat in malignant lesions, and lesion-specific changes in protein content.

Several studies have shown differences in composition for different lesion types. Raptopoulos *et al.* measured the computed tomography (CT) attenuation of 44 surgical biopsy specimens using x-ray CT and found that fibroadenomas showed significantly higher CT attenuation than malignant tissue.[23] The high lesion water content and low lipid content measured for fibroadenomas is consistent with this work. Cerussi *et al.* estimated the composition of 58 malignant breast lesions relative to normal breast tissue using broadband infrared diffuse optical spectroscopy.[24] They found that malignant breast tissues showed reduced lipid content and increased water content relative to normal breast tissue. In a later study, Kukreti *et al.* measured differences between benign and malignant lesions using broadband infrared diffuse optical spectroscopy.[25] They were able to differentiate between malignant lesions and fibroadenomas with sensitivities and specificities of 91% and 95%, respectively. The authors of the study speculated that the differences in spectrum were attributable to changes in lipid metabolism. We found, similarly, that malignant tissue had higher water and lower lipid content than normal (background tissue). However, it is not clear how our results may be interpreted in terms of lipid metabolism.

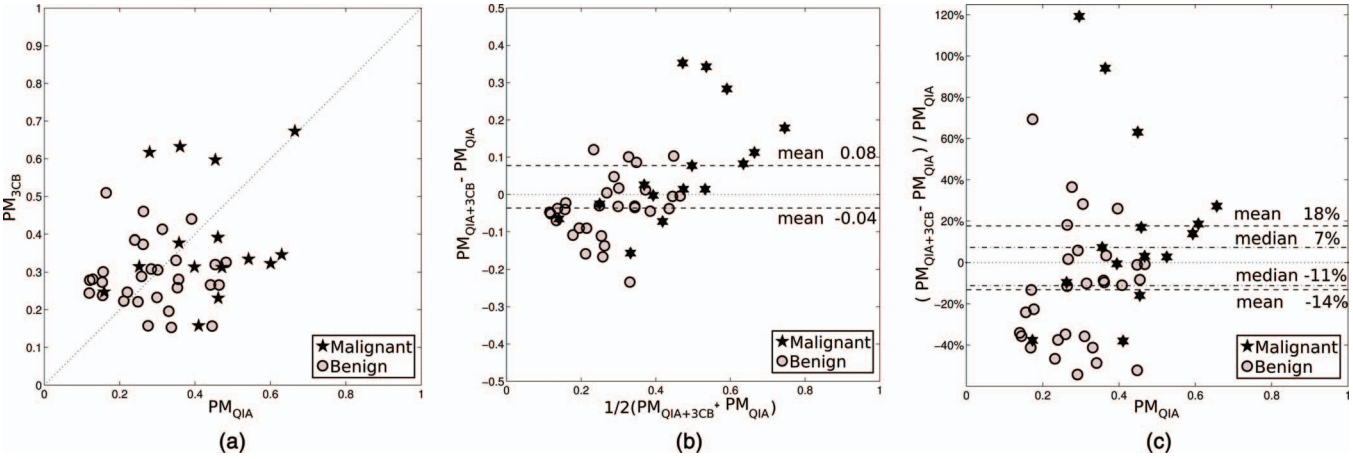

FIG. 3. The computer-estimated probability of malignancy in leave-one-case-out cross-validation (a) obtained by "3CB alone" vs "QIA alone" analyses, (b) in a Bland–Altman plot of "QIA+3CB" and "QIA alone" analyses, and (c) as the relative shift of "QIA+3CB" with respect to "QIA alone."

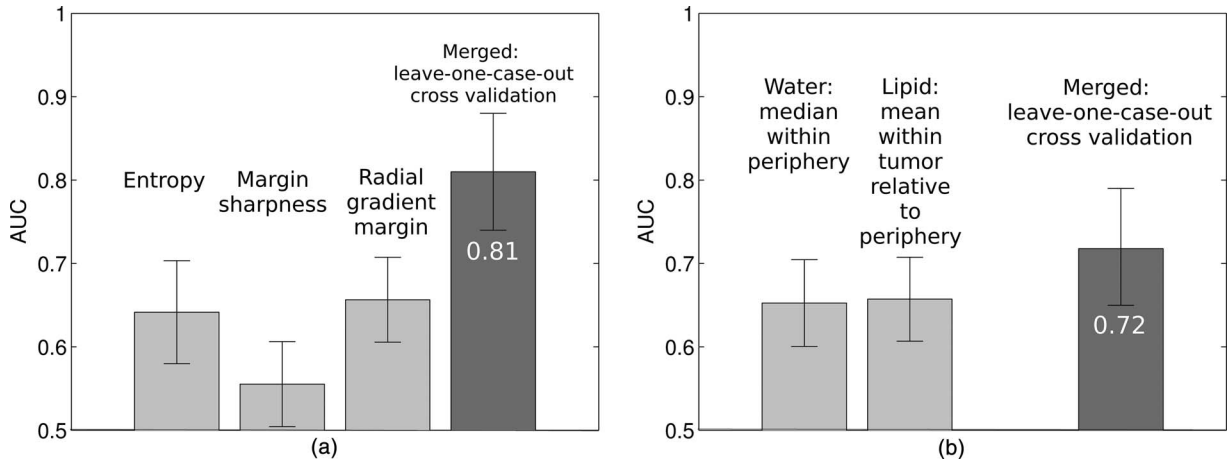

FIG. 4. Areas under the ROC curve (AUC) for individual features selected in stepwise feature selection and features merged by a neural net in a leave-one-case-out cross-validation for (a) "QIA alone" and (b) "3CB alone."

A limitation of this study was the modest size of the dataset, and hence trends were observed rather than statistically significant improvements regarding the analysis including the new 3CB imaging modality with respect to the reference method of "QIA alone." The performance for the task of distinguishing between benign and malignant lesions in the leave-one-patient-out analysis appeared to be slightly lower than in the leave-one-lesion-out analysis, but the trend towards improved performance when combining 3CB and QIA was the same. It is also important to note that the weak correlation between the computer-estimated probability of malignancy obtained in the "QIA alone" analysis and that obtained in the "3CB alone" analysis, indicates that QIA measures and localized 3CB measures were largely independent and provided different predictors of breast cancer. Hence, there should be ample opportunity for synergy when combining these two methods.

Future plans include robustness testing with respect to the computerized lesion segmentation method and the impact of the radiologist-determined lesion centers.[19] It may be possible to further improve our overall method by adapting and fine-tuning the segmentation method (which was currently not feasible due to the limited size of the dataset). It is worth noting, however, that similar trends were observed in analysis (not presented here) based on the expert radiologist manually determined lesion boundaries. Patient recruitment for this study will resume shortly and in the coming years a large dataset will be collected allowing verification and improvement of our methodology, as well as more extensive statistical testing.

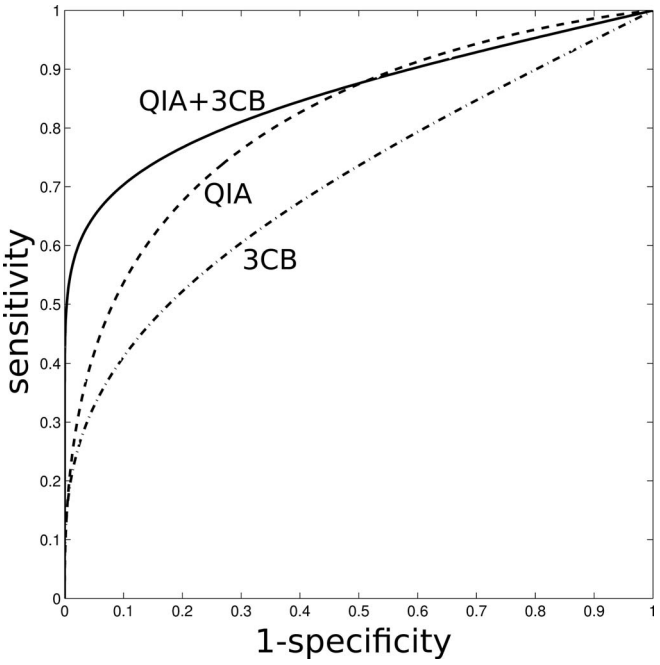

FIG. 5. ROC curves for leave-one-case-out cross-validation for "QIA alone" (AUC = 0.81 ± 0.07), "3CB alone" (AUC = 0.72 ± 0.07), and "QIA+3CB" (AUC = 0.86 ± 0.04).

ACKNOWLEDGMENTS

The authors are grateful for the recruitment assistance of S. Hwang, S. Eder, and protocol execution of J. Sherman. The authors would like to acknowledge the following funding sources: National Cancer Institute No. R01 CA166945, and the California Breast Cancer Research Program No. 18IB-0042. M.L.G. is a stockholder in R2 Technology/Hologic, cofounder of and stockholder in Quantitative Insights, Inc., and receives royalties from Hologic, GE Medical Systems, MEDIAN Technologies, Riverain Medical, Mitsubishi, and Toshiba. K.D. and H.L. received royalties from Hologic. It is the University of Chicago Conflict of Interest Policy that investigators disclose publicly actual or potential significant financial interest that would reasonably appear to be directly and significantly affected by the research activities.

[a]Author to whom correspondence should be addressed. Electronic mail: kdrukker@uchicago.edu

[1]A. Salomon, "Beiträge zur Pathologie und Klinik der Mammakarzinome," Arch. Klin. Chir. **101**, 573–668 (1913).
[2]F. Shtern, "Digital mammography and related technologies: A perspective from the National Cancer Institute," Radiology **183**, 629–630 (1992).
[3]K. Kerlikowske, R. A. Hubbard, D. L. Miglioretti, B. M. Geller, B. C. Yankaskas, C. D. Lehman, S. H. Taplin, and E. A. Sickles, "Comparative effectiveness of digital versus film-screen mammography in

community practice in the United States: A cohort study," Ann. Intern. Med. **155**, 493–502 (2011).
[4]E. D. Pisano, C. A. Gatsonis, M. J. Yaffe, R. E. Hendrick, A. N. Tosteson, D. G. Fryback, L. W. Bassett, J. K. Baum, E. F. Conant, R. A. Jong, M. Rebner, and C. J. D'Orsi, "American College of Radiology Imaging Network digital mammographic imaging screening trial: Objectives and methodology," Radiology **236**, 404–412 (2005).
[5]K. Kerlikowske, W. Zhu, R. A. Hubbard, B. Geller, K. Dittus, D. Braithwaite, K. J. Wernli, D. L. Miglioretti, and E. S. O'Meara, "Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy," JAMA Intern. Med. **173**, 807–816 (2013).
[6]R. A. Hubbard, K. Kerlikowske, C. I. Flowers, B. C. Yankaskas, W. Zhu, and D. L. Miglioretti, "Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: A cohort study," Ann. Intern. Med. **155**, 481–492 (2011).
[7]H. D. Nelson, K. Tyne, A. Naik, C. Bougatsos, B. K. Chan, and L. Humphrey, "Screening for breast cancer: An update for the U.S. Preventive Services Task Force," Ann. Intern. Med. **151**, 727–737 (2009).
[8]E. Azavedo, S. Zackrisson, I. Mejare, and M. H. Arnlind, "Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review," BMC Med. Imaging **12**, 22 (2012).
[9]K. Horsch, M. L. Giger, C. J. Vyborny, L. Lan, E. B. Mendelson, and R. E. Hendrick, "Classification of breast lesions with multimodality computer-aided diagnosis: Observer study results on an independent clinical data set," Radiology **240**, 357–368 (2006).
[10]Z. Huo, M. L. Giger, C. J. Vyborny, and C. E. Metz, "Breast cancer: Effectiveness of computer-aided diagnosis observer study with independent database of mammograms," Radiology **224**, 560–568 (2002).
[11]A. D. Laidevant, S. Malkov, C. I. Flowers, K. Kerlikowske, and J. A. Shepherd, "Compositional breast imaging using a dual-energy mammography protocol," Med. Phys. **37**, 164–174 (2010).
[12]S. Liewei, E. R. Ward, and B. Stroy, "A review of dielectric properties of normal and malignant breast tissue," *Proceedings of IEEE SoutheastCon*, April 5–7, 2002, 2002SECon-SEC108, page 1–7.
[13]J. A. Shepherd, K. M. Kerlikowske, R. Smith-Bindman, H. K. Genant, and S. R. Cummings, "Measurement of breast density with dual X-ray absorptiometry: Feasibility," Radiology **223**, 554–557 (2002).
[14]S. Malkov, J. Wang, K. Kerlikowske, S. R. Cummings, and J. A. Shepherd, "Single x-ray absorptiometry method for the quantitative mammographic measure of fibroglandular tissue volume," Med. Phys. **36**, 5525–5536 (2009).
[15]Y. Yuan, M. L. Giger, H. Li, K. Suzuki, and C. Sennett, "A dual-stage method for lesion segmentation on digital mammograms," Med. Phys. **34**, 4180–4193 (2007).
[16]Y. Yuan, M. L. Giger, H. Li, and C. Sennett, "Correlative feature analysis on FFDM," Med. Phys. **35**, 5490–5500 (2008).
[17]H. Li, M. L. Giger, Y. Yuan, W. Chen, K. Horsch, L. Lan, A. R. Jamieson, C. A. Sennett, and S. A. Jansen, "Evaluation of computer-aided diagnosis on a large clinical full-field digital mammographic dataset," Acad. Radiol. **15**, 1437–1445 (2008).
[18]Z. Huo, M. L. Giger, C. J. Vyborny, D. E. Wolverton, and C. E. Metz, "Computerized classification of benign and malignant masses on digitized mammograms: a study of robustness," Acad. Radiol. **7**, 1077–1084 (2000).
[19]N. P. Gruszauskas, K. Drukker, M. L. Giger, C. A. Sennett, and L. L. Pesce, "Performance of breast ultrasound computer-aided diagnosis: Dependence on image selection," Acad. Radiol. **15**, 1234–1245 (2008).
[20]L. L. Pesce and C. E. Metz, "Reliable and computationally efficient maximum-likelihood estimation of "proper" binormal ROC curves," Acad. Radiol. **14**, 814–829 (2007).
[21]C. E. Metz, "ROC methodology in radiologic imaging," Invest Radiol. **21**, 720–733 (1986).
[22]J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," Lancet **327**, 307–310 (1986).
[23]V. Raptopoulos, J. K. Baum, M. Hochman, A. Karellas, M. J. Houlihan, and C. J. D'Orsi, "High resolution CT mammography of surgical biopsy specimens," J. Comput. Assist. Tomogr. **20**, 179–184 (1996).
[24]A. Cerussi, N. Shah, D. Hsiang, A. Durkin, J. Butler, and B. J. Tromberg, "*In vivo* absorption, scattering, and physiologic properties of 58 malignant breast tumors determined by broadband diffuse optical spectroscopy," J. Biomed. Opt. **11**, 044005 (2006).
[25]S. Kukreti, A. E. Cerussi, W. Tanamai, D. Hsiang, B. J. Tromberg, and E. Gratton, "Characterization of metabolic differences between benign and malignant tumors: High-spectral-resolution diffuse optical spectroscopy," Radiology **254**, 277–284 (2010).