Published in final edited form as:

Breast J. 2014 ; 20(3): 249–257. doi:10.1111/tbj.12255.

Association between Computed Tissue Density Asymmetry in Bilateral Mammograms and Near-term Breast Cancer Risk

Bin Zheng¹, Maxine Tan¹, Pandiyarajan Ramalingam², and David Gur²

¹School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK 73019

²Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15213

Abstract

This study investigated association between bilateral mammographic density asymmetry and near-term breast cancer risk. A database of digital mammograms acquired from 690 women was retrospectively collected. All images were originally interpreted as negative by radiologists. During the next subsequent screening examinations (between 12 and 36 months later), 230 women were diagnosed positive for cancer, 230 were recalled for additional diagnostic workups and proved to be benign, and 230 remained negative (not-recalled). We applied a computerized scheme to compute the differences of five image features between the left and right mammograms, and trained an artificial neural network (ANN) to compute a bilateral mammographic density asymmetry score. Odds ratios (ORs) were used to assess associations between the ANN-generated scores and risk of women having detectable cancers during the next screening examinations. A logistic regression method was applied to test for trend as a function of the increase in ANN-generated scores. The results were also compared with ORs computed using other existing cancer risk factors. The ORs showed an increasing risk trend with the increase of ANN-generated scores (from 1.00 to 9.07 between positive and negative case groups). The regression analysis also showed a significant increase trend in slope ($p < 0.05$). No significant increase trends of the ORs were found when using woman's age, subjectively rated breast density, or family history of breast cancer. This study demonstrated that the computed bilateral mammographic density asymmetry had potential to be used as a new risk factor to improve discriminatory power in predicting near-term risk of women developing breast cancer.

Keywords

Breast cancer risk; Breast density; Bilateral mammographic density asymmetry; Cancer risk prediction; Mammography screening

I. INTRODUCTION

Since the majority of breast cancers are detected in women without known risk factors [1, 2], a uniformly applied mammography screening protocol is currently recommended for all women who qualify for screening (e.g., annual screening for women over 40 years old in the

United States). Early detection combined with improved treatment strategies have incrementally and significantly reduced patients' mortality and morbidity rates over the last four decades [3, 4]. However, interpreting mammograms is a difficult and time-consuming task in a screening environment due to large variability in the depicted breast abnormalities, overlapping dense fibro-glandular tissue, and low cancer prevalence (i.e., 3 to 5 cancers in 1000 non-baseline screening examinations) [5]. These factors substantially reduce detection sensitivity and specificity of mammography [6], in particular in younger women with dense breast tissue as well as in other "high risk" groups [7, 8]. The specificity of mammography is also low [9]. One study reported that during a 10-year period, more than half of screened women would receive at least one false-positive recall and 7% to 9% would have at least one benign biopsy [10]. In addition to women's anxiety, which can cause long-term psychosocial consequences [11], and potential harmful effects due to cumulative radiation exposure [12] and unnecessary biopsies [10], limited healthcare resources and associated high costs are also issues that will have to be addressed in current breast cancer screening practices [13]. As a result, the efficacy of current mammography screening remains quite controversial [14, 15]. To overcome the limitations of current population-based mammography screening practices, it is desirable to make personalized screening recommendations based on individualized risk assessment [16], and this concept has been recently attracting significant research interest [17–20]. The ultimate goal of developing and implementing a personalized cancer screening paradigm is to enable the identification of a small fraction of women with significantly higher than average near-term risk of developing breast cancer. As a result, this small fraction of high-risk women should be more frequently screened (e.g., annually or perhaps even more frequently), while the vast majority of women at average or lower risk of developing cancer in the near-term could be screened at longer intervals (e.g., every two to five years) until their near-term cancer risk has significantly increased.

Although a number of epidemiology-based breast cancer risk prediction models (e.g., Gail, Tyrer-Cuzick, Claus, and others [2, 21]) have been previously developed and tested, all of which rely primarily on a combination of several risk factors, such as women's age, family cancer history, genotyping based information, and breast density, these risk models typically estimate a long-term or a lifetime relative risk of developing cancer in women as compared to an average general population risk. However, these risk prediction models, in most instances, do not have a clinically acceptable discriminatory power (e.g., positive predictive value) when applied to the individual woman [22]. In addition, the predicted risk levels resulting from these risk models remain constant for the individual woman throughout her lifetime. Since cancer development is a progressive process, the establishment of an optimal and personalized cancer screening paradigm requires a new breast cancer risk prediction model in which the risk factors could produce variable values as a decreasing time lag between negative and positive (cancer being detectable) screening examinations. Hence, each woman could have an adjustable screening interval as a function of her near-term predicted cancer risk level.

In our group, we have preliminarily investigated a computerized method to detect bilateral mammographic density asymmetry between left and right breasts [23], and the potential of using the computed bilateral mammographic density asymmetry scores to predict the risk of

a woman developing breast cancer after a negative screening examination of interest [24]. In this study, we aim to assess the feasibility of applying a machine learning method to develop a near-term breast cancer risk prediction model using an expanded image dataset and a new analytical method. Our hypothesis is that the bilateral mammographic tissue or density asymmetry between the left and right breasts is an important radiographic image phenotype related to abnormal biological processes that may lead to cancer development [25], and an increase in bilateral mammographic density asymmetry could be an important indicator of developing breast abnormality or cancer. As an example, Figure 1 shows four sets of bilateral mammograms acquired from two women during two repeated (sequential) screening examinations. The “baseline” examinations of interest were interpreted clinically as “negative” by the radiologists. During the next subsequent examination, cancer was detected in the woman showing a higher degree of bilateral mammographic density asymmetry on the baseline examination, while another woman remained “negative” although she had overall higher but more symmetrically mammographic density on the baseline examination. However, subjectively rating mammographic density is difficult and inconsistent due to the large intra- and inter-observer variability [26]. Since studies have demonstrated that using a computerized scheme enabled achievement of more consistent results in assessing mammographic tissue density on individual images [27, 28], we will also develop and apply a computerized scheme to detect bilateral mammographic tissue or density asymmetry.

In this study, we built an artificial neural network (ANN) based on a set of image features computed from negative digital mammograms prior to the detection of the abnormalities in question, in order to generate a new classification score that represents bilateral mammographic density asymmetry. We then applied the ANN-generated bilateral mammographic density asymmetry scores to predict the likelihood of a woman developing a “detectable” breast cancer or a high risk lesion on the subsequent mammographic screening examination acquired between 12 to 36 months following the negative screening examination of interest. The risk factor or model performance was also compared with several other frequently used known breast cancer risk factors (including women’s age, subjectively rated breast tissue density, and family history of breast cancer) that have been routinely used and incorporated into many existing breast cancer risk assessment or prediction models.

II. MATERIALS AND METHODS

Under an institutional review board (IRB) approved image data collection protocol, we asked an IRB-certified research staff (“honest broker”) to retrospectively search for and collect the full-field digital mammography (FFDM) images from our clinical database. From this ongoing data collection process, we established an initial testing dataset for this study, which includes screening examinations acquired from 690 women. Each case includes FFDM images acquired during two sequential mammographic screening examinations within a time lag of 12 to 36 months. The first (or “baseline”) examination in the sequence was interpreted by radiologists as “negative” (screening BIRADS 1) or “definitely benign” (screening BIRADS 2), and had not been recalled for additional workup. These “baseline” examinations were used for computing mammographic image features and for building the

breast cancer risk prediction model. The second examination in the sequence was used to determine if the case is to be considered “positive” (depicting an abnormality) or negative (remained negative). In this preliminary study, we only focused on predicting the likelihood of women having or developing “detectable” breast cancers in the next sequential screening examinations after one (“baseline”) negative screening of interest. Hence, whether the “negative” and “benign” cases remained in the “cancer-free” status in the third sequential mammographic screening examinations was not validated in this study.

Using this approach, all cases were assigned to a case status and the dataset was divided into three categories or subgroups namely, positive, benign and negative. The “positive” subgroup included 202 cases in which the pathology verified breast cancers were detected as well as 28 high risk pre-cancer cases in which the identified lesions were surgically excised during the second examination. The “benign” subgroup included 230 cases in which the suspicious abnormalities were detected on mammograms and the woman was recalled; however, during the diagnostic workup that followed (imaging and/or biopsy), the finding proved to be benign. The “negative” subgroup included 230 cases that remained negative (not recalled by the radiologists). In addition, several known risk factors including the woman’s age, subjectively rated breast or mammographic density (density BIRADS), and reported breast cancer family history were also ascertained and saved in our computer database. Table 1 summarizes the distribution of three known risk factors among three groups of positive, benign and negative cases. Table 2 shows the distribution of time differences (intervals) between the “verification” (second) examination and the prior “negative” examination used in this study.

Although a FFDM examination typically includes four images of craniocaudal (CC) and mediolateral oblique (MLO) view of left and right breasts, applying a computerized scheme to detect mammographic tissue density using CC view images is more accurate and reliable than using MLO view images, as the procedure of automatically detecting the chest wall and removing other non-breast regions depicted on MLO view images is difficult and often unreliable. Hence, we selected a pair of bilateral CC view FFDM images of left and right breasts from each of the “baseline” negative examinations in this study. We first applied a computerized scheme that uses an iterative threshold method to segment breast area from the air background depicted on each CC view image [29], and then computed five mammographic tissue composition and density related image features using previously reported methods directly from the original FFDM images [30]. In brief, these include four pixel value based statistical features, namely (1) mean pixel value of the whole segmented breast area depicted on one image, (2) standard deviation, (2) skewness and (4) kurtosis of all pixel values. The fifth feature represents the local variation (or local maximum difference in pixel values) within a region of interest (R) with a radius r . For each pixel (x, y) inside the breast area, the local variation is
$$V(x, y, R) = \max_{(i,j) \in R} I(i, j) - \min_{(i,j) \in R} I(i, j),$$
 whereby (i, j) is the coordinate of the neighboring pixel. The total local variation as the function of r is computed as
$$V(r) = \sum_{(x,y) \in R} V(x, y, R(r)).$$
 Then, the scheme estimates fractal dimension (FD) from the relationship $V(r) \propto r^{2-FD}$ by fitting a straight line to the $V(r)$ versus r function. The slope of the fitted line is used as the fifth image feature. The computerized scheme was

independently applied to each CC view image of the left and right breasts in order to segment the breast area and compute the five image features. Finally, five feature differences were computed by subtracting matched features computed from the two bilateral CC view images, $\Delta F_i = |F_i^L - F_i^R|$, where $i = 1, 2, \dots, 5$.

To generate the bilateral mammographic density asymmetry score by combining these five computed image feature differences, we built a simple three layer artificial neural network (ANN) [31]. The ANN has five input neurons (represented by the five computed image feature differences) in the first (input) layer, two hidden neurons in the second layer, and one decision neuron in the third (output) layer. To minimize the training/testing bias when using the ANN, we used a leave-one-case-out (LOCO) method [32] to compute and obtain a bilateral mammographic density asymmetry score for each case in our testing dataset. For example, when we compared the risk prediction performance between the 230 positive and 230 negative cases (total 460 cases), the ANN was first trained using 459 cases, and the trained ANN was then applied to the one remaining (left out) case to obtain a bilateral mammographic density asymmetry score (ranging from 0 to 1). The higher the score, the higher the bilateral mammographic density asymmetry level is. This process was repeated 460 times whereby each case was used in the training sample in 459 cycles and as a test sample once. The same ANN training/testing protocol reported and used in our previous study [33] was applied in all 460 training/testing computations. The LOCO method was also applied to train and test the other two sets of ANNs for classification between the positive and benign cases as well as between the benign and negative cases.

The data was then analyzed using odds ratios (ORs) as summary measures (or as a performance index) in assessing the associations, if any, between several risk factors and the detection of breast cancer or high risk lesions 12 to 36 months after a “baseline” negative screening examination of interest in this study. The investigated and compared risk factors including the ANN-generated bilateral mammographic density asymmetry score, women’s age, subjectively rated breast density (BIRADS), and family history of breast cancer were all evaluated for this purpose. To test for trend in ORs, we used a regression method. We divided all training/testing cases into four or five subgroups (bins) based on the values and/or categories of each of these risk factors. All data analysis was performed using a publicly available software package of statistical computing (R version 2.1.1, <http://www.r-project.org>). The results were then tabulated and compared.

III. RESULTS

Table 3 shows the distribution of five computed image feature differences in three subgroups of positive, benign and negative cases. The results show a general trend in that (1) the positive (cancer) cases have larger mean and median values than the recalled benign cases, and (2) the benign cases have larger mean and median values than the screening negative (not-recalled) cases for all five feature differences. Table 4 summarizes the correlation coefficients among all combinations of the computed values of the five image feature differences. The results of the relatively low correlation coefficients indicate that these features are not highly redundant. The low correlation of these features enables us to

develop a machine learning based classifier (e.g., an ANN) that combines the features to further improve the performance of the classifier in assigning the cases into the appropriate case subgroups as well as in the cancer risk prediction task.

The ANN-generated scores are able to detect bilateral mammographic density asymmetry. An example is shown in Figure 1 whereby the ANN-generated score (0.728) for the negative image set acquired from a case with an invasive mammary carcinoma detected in the subsequent (“second”) screening examination (upper row of Figure 1) is higher than the ANN-generated score (0.484) for another negative image set of a case that remained negative during the subsequent examination (lower row of Figure 1). Table 5 summarizes the ORs and the corresponding 95% confidence intervals (CIs) computed for five subgroups (bins) of three comparisons including (1) positive versus negative case subgroups, (2) positive versus benign case subgroups, and (3) benign versus negative case subgroups. A set of thresholds was automatically determined for division of the cases into five subgroups (bins) with an approximately equal number of cases in each subgroup. The ANN-generated density asymmetrical scores gradually increased from subgroup 1 to 5. Using the cases in subgroup 1 as a baseline (reference), ORs and 95% CIs for subgroups 2 to 5 were computed and compared in Table 5. An increasing trend of ORs is observed as a function of increasing ANN-generated bilateral mammographic density asymmetry scores in each of the three comparison sets. Using regression analysis, the slopes of all three regression lines are significantly different from the zero slope ($p < 0.05$), which indicates a trend in increasing ORs with the increase of the ANN-generated bilateral mammographic density asymmetry scores. Comparing the positive and negative case groups, the ORs systematically increase from 1.0 to 9.07 with 95% CI of [4.64, 17.7] in subgroup (bin) #5, which has the highest range of ANN-generated bilateral mammographic density asymmetry scores. The results also show that the ORs between the positive and benign case groups are in general greater than the ORs between the benign and negative case groups, which indicate that bilateral mammographic density asymmetry levels of the recalled benign cases are more similar to the negative cases than to the positive (cancer) cases.

Table 6 summarizes the ORs (including 95% CIs) and regression trend analysis when applying three other known risk factors (namely women’s age, breast density as subjectively rated by radiologists, and reported family history of breast cancer) to predict the near-term cancer risk in this testing dataset. Compared with the ANN-generated bilateral mammographic density asymmetry scores (Table 5), the computed ORs are substantially lower (ORs = 2.74). Regression based trend analysis also showed that the slopes were not statistically significantly different from a zero slope indicating that these risk factors have little discriminatory power (if any) in predicting near-term risk of women developing a “detectable” breast abnormality or cancer during the next subsequent examination as compared to the “negative” cases.

IV. DISCUSSION

In this study, we preliminarily investigated the association between the computed bilateral mammographic density asymmetry related features in negative mammograms and near-term risk of a woman having a detectable breast abnormality that may lead to developing breast

cancer in the next subsequent examination. The study demonstrated an increasing trend in the odds ratios (ORs) with the increase of the computed bilateral mammographic density asymmetry level. This study has several unique characteristics. Firstly, although many epidemiology-based breast cancer risk assessment models have been established and applied [2], these models typically show statistically significant associations with lifetime cancer risk of a woman as compared with the general population, albeit with relatively low positive predictive value (discriminatory power) at the individual level that is required for routine personalized clinical practice [22]. Our approach is different as we focus on investigating a new concept of assessing near-term risk of developing breast cancer from the bilateral mammographic density asymmetry based information from the images themselves. Therefore, the prediction results of this study are not directly comparable with the lifetime breast cancer risk assessment results when using the conventional epidemiology based risk models. It should provide supplementary information to existing breast cancer risk prediction models.

Secondly, bilateral mammographic density asymmetry has little association with the breast (or mammographic) density. In this study, the results showed that subjectively rated mammographic density and ANN-generated bilateral mammographic density asymmetry scores were largely independent (i.e., correlation coefficient $r = -0.06$ in the positive case group). Although subjectively rated breast or mammographic density by the radiologists has been widely practiced in screening mammography and is commonly considered as one of strongest breast cancer risk factors [2, 34], the bilateral mammographic density asymmetry has never been investigated as the basis for breast cancer risk assessment models. Using the computerized bilateral mammographic density asymmetry may also have an important advantage as it is well known that subjective assessment of breast density is prone to large intra and inter observer variability [26, 27]. Breast density could be affected by several uncontrollable factors (i.e., losing or gaining weight, using hormonal agents, smoking and drinking alcohol) [35], and is also difficult to be accurately assessed or computed from mammograms without knowing the exact X-ray exposure values and breast thickness during imaging [36]. However, computing bilateral mammographic density asymmetry may be more robust in this regard as many of these factors should similarly affect both breasts. Therefore, the features that are based on the subtracted values of two bilateral mammograms are more likely to be valid for this intended purpose.

Thirdly, in this study we divided our testing dataset into three subgroups of positive, benign, and negative cases, following which we conducted three comparisons involving different combinations of the paired subgroups. The data analysis results (Table 3) indicate that as the risk levels increase (from recalled suspicious abnormalities to verified cancers), the computed bilateral mammographic tissue composition and/or density asymmetrical levels also increase in the “baseline” negative mammograms. Hence, the obtained consistent trends strengthen the likelihood that our observations are valid.

Lastly, the differences of odds ratios and their increasing trends in the three data comparison sets (as shown in Table 5) indicate that this new risk factor or model that we are exploring is also time sensitive, namely, the near-term breast cancer risk status of a woman may change if her bilateral mammographic density asymmetry changes. Hence, the likelihood increases

that this type of cancer risk assessment approach could eventually lead to helping in the establishment of a new and personalized breast cancer screening paradigm with adaptively adjustable screening interval recommendations.

We also recognize that this is a preliminary study with a number of limitations. Firstly, this is a laboratory type retrospective study with a selected and enriched group of positive examinations. Although all cases were randomly selected by the independent research staff (“honest brokers”) who was not involved in any of the data analysis tasks to minimize case selection bias, the dataset is not an adequate representative of the ensemble of negative mammograms. In our future study, we will continue to collect new cases to expand the size of our database, in particular to increase the number of negative cases to better represent the cancer prevalence ratio in a mammographic screening environment. Secondly, although a large number of computed image pixel value based features and schemes have been previously developed and tested to represent or assess mammographic tissue density and/or patterns, defining an optimal and widely accepted computerized method that can replace radiologists’ subjective density ratings remains under intensive investigation. In this study, we tested a simple ANN approach using five image features. Despite the encouraging experimental results, these features and the structure of the ANN based classifier may not be optimal for the task. Thirdly, unlike the majority of existing breast cancer risk models [2] that estimate primarily a long-term and/or a lifetime risk using primary risk factors that rarely, if ever, change over time (e.g., family history, genetic disposition etc.), the bilateral mammographic density asymmetry can change over time, hence a woman’s status in terms of computed/assigned near-term or short-term risk may change as well due to the development of breast abnormalities that may lead to the cancer detection. How such changes over a series of examinations may affect predictive values has not been investigated to date. Lastly, in this study we only tracked the detection status change between the two sequential mammographic screening examinations, in which the first examination of interest was interpreted by the radiologist as negative and the woman was not recalled. Whether an early cancer was missed or overlooked by the radiologist in the first negative examination was not retrospectively reassessed in this study. Hence, the risk factor or model developed in this study has only been tested as an attempt to identify women with significantly higher risk of having image “detectable” breast cancers in the next sequential examination not regarding whether the cancer has been missed in the prior “negative” examinations, so that the high risk women assessed using this new risk factor should be recommended for more frequent screening in order to increase the likelihood of the cancers being detected at an early stage (thus avoiding further delay).

In summary, this study demonstrates that automatically computed bilateral mammographic density asymmetry levels combined with a simple machine learning classifier enables the assessment of near-term risk of developing an image “detectable” abnormality that requires intervention. Our proposed approach has substantially higher discriminatory power than other known risk factors in terms of predicting the likelihood of individual women having or developing the image detectable high-risk breast abnormalities and/or cancer in the near-term. This general approach could lead to the development of an optimal and adaptively adjustable personalized breast cancer screening paradigm. Clearly, this preliminary work the

reported risk prediction performance needs further validation in a large prospective (or cohort) study.

Acknowledgments

This work was supported in part by the National Cancer Institute, National Institutes of Health, under Grant R01CA160205.

REFERENCES

1. Madigan MP, Ziegler RG, Benichou J, et al. Proportion of breast cancer cases in the United States explained by well-established risk factors. *J Natl Cancer Inst.* 1995; 87:681–1685.
2. Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. *J Natl Cancer Inst.* 2010; 102:680–691. [PubMed: 20427433]
3. Tabar L, Vitak B, Chen HH, et al. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. *Cancer.* 2001; 91:1724–1731. [PubMed: 11335897]
4. Smith RA, Cokkindes V, Brooks D D, et al. Cancer screening in the United States, 2011. *CA Cancer J Clin.* 2011; 61:8–30. [PubMed: 21205832]
5. Sickles EA, Wolverton DE, Dee KE. Performance parameters for screening and diagnostic mammography: specialist and general radiologists. *Radiology.* 2002; 224:861–869. [PubMed: 12202726]
6. Fenton JJ, Wheeler J, Carney PA, et al. Reality check: perceived versus actual performance of community mammographers. *Am J. Rotentgenol.* 2006; 187:42–46.
7. Kriege M, Brekelmans CT, Boetes C, et al. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. *N Engl J Med.* 2004; 351:427–437. [PubMed: 15282350]
8. Leach MO, Boggis CR, Dixon AK, et al. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: a prospective multicentre cohort study (MARIBS). *Lancet.* 2005; 365:1769–1778.
9. Kopans DB. The positive predictive value of mammography. *Am J. Rotentgenol.* 1991; 158:521–526.
10. Hubbard RA, Kerlikowske K, Flowers CI, et al. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: A cohort study. *Ann Intern Med.* 2011; 155:481–492. [PubMed: 22007042]
11. Brodersen J, Siersma VD. Long-term psychosocial consequences of false-positive screening mammography. *Ann Family Medicine.* 2013; 11:106–115.
12. Yaffe MJ, Mainprize JG. Risk of radiation-induced breast cancer from mammographic screening. *Radiology.* 2011; 258:98–105. [PubMed: 21081671]
13. Buist DS, Anderson ML, Haneuse SJ, et al. Influence of annual interpretive volume on screening mammography performance in the United States. *Radiology.* 2011; 259:72–84. [PubMed: 21343539]
14. Nelson HD, Tyne K, Naik A, et al. U.S. Preventive Services Task Force, Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2009; 151:727–737. [PubMed: 19920273]
15. Berlin L, Hall FM. More mammography muddle: emotions, politics, science, costs and polarization. *Radiology.* 2010; 255:311–316. [PubMed: 20413746]
16. Brawley OW. Risk-based mammography screening: an effort to maximize the benefits and minimize the harms. *Ann Intern Med.* 2012; 156:662–663. [PubMed: 22547477]
17. Tice JA, Cummings SR, Smith-Bindman R, et al. Using clinical features and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med.* 2008; 148:337–347. [PubMed: 18316752]

18. Schousboe JT, Kerlikowske K, Loh A, et al. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Ann Intern Med.* 2011; 155:10–20. [PubMed: 21727289]
19. Van Ravesteyn NT, Miglioretti DL, Stout NK, et al. Tipping the balance of benefits and harms to favor screening mammography starting at age 40 years: A comparative modeling study of risk. *Ann Intern Med.* 2012; 156:609–617. [PubMed: 22547470]
20. Nelson HD, Zakher B, Cantor A, et al. Risk factors for breast cancer for women aged 40 to 49 years: A systematic review and meta-analysis. *Ann Intern Med.* 2012; 156:635–648. [PubMed: 22547473]
21. Martin LJ, Melnichouk O, Guo H, et al. Family history, mammographic density, and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2010; 19:456–463. [PubMed: 20142244]
22. Gail MH, Mai PL. Comparing breast cancer risk assessment models. *J Natl Cancer Inst.* 2010; 102:665–668. [PubMed: 20427429]
23. Wang X, Lederman D, Tan J, Zheng B. Computerized prediction of risk for developing breast cancer based on bilateral mammographic breast tissue asymmetry. *Med Eng Phys.* 2011; 33:934–942. [PubMed: 21482168]
24. Zheng B, Sumkin JH, Zuley ML, Wang X, Klym AH, Gur D. Bilateral mammographic density asymmetry and breast cancer risk: A preliminary assessment. *Euro J Radiology.* 2012; 81:3222–3228.
25. Scutt D, Manning JT, Whitehouse GH, et al. The relationship between breast asymmetry, breast size and the occurrence of breast cancer. *Br J Radiol.* 1997; 70:1017–1021. [PubMed: 9404205]
26. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast imaging reporting and data system: Inter- and intra-observer variability in feature analysis and final assessment. *Am J Roentgenol.* 2000; 174:1769–1777. [PubMed: 10845521]
27. Glide-Hurst CK, Duric N, Littrup P. A new method for quantitative analysis of mammographic density. *Med Phys.* 2007; 34:4491–4498. [PubMed: 18072514]
28. Wei J, Chang HP, Wu Y, et al. Association of computerized mammographic parenchymal pattern measure with breast cancer risk: a pilot case-control study. *Radiology.* 2011; 260:42–49. [PubMed: 21406634]
29. Zheng B, Leader JK, Abrams GS, et al. Multiview based computer-aided detection scheme for breast masses. *Med Phys.* 2006; 33:3135–3143. [PubMed: 17022205]
30. Chang Y, Wang X, Hardesty LA, et al. Computerized assessment of tissue composition on digitized mammograms. *Acad Radiol.* 2002; 9:899–905. [PubMed: 12186438]
31. Mitchell, TM. *Machine Learning.* Boston, MA: WCB/McGraw-Hill; 1997.
32. Li Q, Doi K. Reduction of bias and variance for evaluation of computer-aided diagnosis schemes. *Med Phys.* 2006; 33:868–875. [PubMed: 16696462]
33. Zheng B, Chang Y, Good WF, Gur D. Performance gain in computer-assisted detection schemes by averaging scores generated from artificial neural networks with adaptive filtering. *Med Phys.* 2001; 28:2302–2308. [PubMed: 11764037]
34. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Eng J Med.* 2007; 356:670–675.
35. Harvey J, Bovbjerg VE. Quantitative assessment of mammographic breast density: relationship with breast cancer risk. *Radiology.* 2004; 230:29–41. [PubMed: 14617762]
36. Kopans DB. Basic physics and doubts about relationship between mammographically determined tissue density and breast cancer risk. *Radiology.* 2008; 246:348–353. [PubMed: 18227535]

Figure 1.

An example of bilateral mammographic density asymmetry of a positive and a negative case. The negative “baseline” bilateral CC view images (left) are shown side-by-side with the corresponding subsequent mammogram (right). The woman whose images are displayed on the top row was diagnosed as having invasive mammary carcinoma (arrow) in the subsequent examination while the one displayed on the bottom row remained negative.

Table 1

Distributions of women's age, subjectively rated breast density (BIRADS rating) and family history of breast cancer in the three groups of positive, benign and negative cases

Risk factor	Category	Positive	Benign	Negative
	Total Cases	230	230	230
Age (Years old)	< 45	22 (9.6%)	46 (20.0%)	34 (14.8%)
	45 – 55	81 (35.2%)	88 (38.3%)	95 (41.3%)
	55 – 65	56 (24.3%)	73 (31.7%)	61 (26.5%)
	> 65	71 (30.9%)	23 (10.0%)	40 (17.4%)
	Mean $\pm$ standard deviation	59.0 $\pm$ 12.2	59.9 $\pm$ 8.5	54.4 $\pm$ 9.9
	Median	58.5	53.0	53.0
	Inter-quartile range	49 – 67	46 – 58	47 – 62
Density	Almost all fatty tissue (1)	8 (3.5%)	4 (1.7%)	10 (4.3%)
BIRADS	Scattered fibroglandular densities (2)	58 (25.2%)	71 (30.9%)	75 (32.6%)
	Heterogeneously dense (3)	156 (67.8%)	147 (63.9%)	135 (58.7%)
	Extremely dense (4)	8 (3.5%)	8 (3.5%)	10 (4.3%)
Family	No family history known	130 (56.5%)	118 (51.3%)	122 (53.0%)
History	Cancers in the 1 st degree relatives	47 (20.4%)	49 (21.3%)	47 (20.4%)
	Cancers in the 2 nd degree relatives	40 (17.4%)	44 (19.1%)	42 (18.3%)
	Cancers in the 3 rd degree relatives	13 (5.7%)	19 (8.3%)	19 (8.3%)

Table 2

Distributions of time lag between the latest examination and the “prior” negative FFDM examinations used for development and testing of the classifier.

Time Gap (Months)	Positive cases	Recalled benign cases	Negative cases
12 – 18	190	186	216
18 – 24	34	44	13
24 – 36	6	0	1

Table 3

Distributions of the five computed bilateral image feature differences in the three case groups (positive, benign and negative).

Feature	Case group	Mean $\pm$ SD	Median	Inter-quartile range
1. Average pixel value	Positive	68.9 $\pm$ 63.7	50.0	26.0 – 90.8
	Benign	50.5 $\pm$ 42.5	39.5	20.3 – 68.8
	Negative	41.8 $\pm$ 36.5	33.0	16.3 – 55.0
2. Standard deviation	Positive	51.1 $\pm$ 43.3	41.2	21.0 – 69.8
	Benign	43.0 $\pm$ 39.8	33.9	14.4 – 69.0
	Negative	41.3 $\pm$ 33.2	34.2	16.2 – 58.4
3. Skewness	Positive	0.194 $\pm$ 0.189	0.134	0.073 – 0.260
	Benign	0.153 $\pm$ 0.130	0.122	0.061 – 0.217
	Negative	0.147 $\pm$ 0.119	0.124	0.069 – 0.198
4. Kurtosis	Positive	0.482 $\pm$ 0.485	0.337	0.164 – 0.652
	Benign	0.417 $\pm$ 0.468	0.301	0.145 – 0.532
	Negative	0.333 $\pm$ 0.320	0.224	0.115 – 0.464
5. FD slope	Positive	0.481 $\pm$ 0.500	0.330	0.140 – 0.678
	Benign	0.394 $\pm$ 0.342	0.285	0.160 – 0.548
	Negative	0.388 $\pm$ 0.353	0.300	0.150 – 0.537

Table 4

Correlation coefficients between paired values of computed image feature differences.

Feature difference	1	2	3	4
2	0.311			
3	0.215	0.099		
4	0.121	0.103	0.563	
5	0.197	0.119	0.029	-0.086

Relative odds ratios (OR) and 95% confidence intervals (CI) with increasing levels of computed bilateral mammographic density asymmetry scores in three paired comparisons of different case groups

Table 5

Subgroup	Positive vs. Negative			Positive vs. Benign			Benign vs. Negative		
	Case No.	OR	95%CI	Case No.	OR	95%CI	Case No.	OR	95%CI
1	25 – 68	1.00	baseline	32 – 59	1.00	baseline	33 – 59	1.00	baseline
2	42 – 50	2.28	[1.26, 4.23]	29 – 64	0.84	[0.45, 1.54]	41 – 51	1.43	[0.80, 2.60]
3	49 – 43	3.10	[1.68, 5.73]	46 – 46	1.84	[1.02, 3.34]	46 – 46	1.78	[0.99, 3.23]
4	44 – 48	2.49	[1.35, 4.61]	55 – 37	2.74	[1.51, 4.99]	52 – 40	2.32	[1.28, 4.20]
5	70 – 21	9.07	[4.64, 17.7]	68 – 24	5.22	[2.77, 9.85]	58 – 34	3.05	[1.67, 5.56]
<i>P</i> value ^{<i>1</i>}			0.038			0.014			< 0.001

¹ *P* value for F-test of zero slope of a regression between mammographic density asymmetrical score levels and adjusted odds ratios (in log scale).

Table 6

Relative odds ratios (OR) and 95% confidence intervals (CI) for three known risk factors in predicting near-term cancer risk between positive and negative case groups

Risk factor	Threshold level	Case Number	Odds ratio	95% Confidence interval
Age (years old)	< 45	22 – 34	1.00	baseline
	45 – 55	81 – 95	1.32	[0.71, 2.43]
	55 – 65	56 – 61	1.42	[0.74, 2.71]
	> 65	71 – 40	2.74	[1.42, 5.32]
	<i>P</i> value ^{<i>I</i>}			0.064
Density BIRADS	1	8 – 10	1.00	baseline
	2	58 – 75	0.97	[0.36, 2.60]
	3	156 – 135	1.44	[0.55, 3.76]
	4	8 – 10	1.00	[0.27, 3.72]
	<i>P</i> value ^{<i>I</i>}			0.727
Family history	No	130 – 122	1.00	baseline
	3 rd degree relatives	13 – 19	0.64	[0.30, 1.36]
	2 nd degree relatives	40 – 42	0.89	[0.54, 1.47]
	1 st degree relatives	47 – 47	0.94	[0.58, 1.51]
	<i>P</i> value ^{<i>I</i>}			0.908

^{*I*} *P* value for F-test of zero slope of a regression between the involved risk factor levels and adjusted ORs (in log scale).