Breast Ultrasound in Women With Familial Risk of Breast Cancer
LSJ Sim,1 MBBS, FRCR, JHCL Hendriks,2 MD, PhD, SMC Fook-Chong,3 MSc, CStat

Introduction
Mammography is the modality of choice to screen for breast cancer in asymptomatic women. However, it is known that about 10% to 12% of breast cancers are mammographically occult.1 It has been reported that for younger women, the sensitivity of screening mammography is reduced and this has been chiefly attributed to the inherent dense breast parenchyma.2 Bird et al2 also stated that breast cancer is less likely to be detected in women with dense breasts and in women <50 years old. Another postulation for the reduced benefit of mammographic screening in young women is the shorter sojourn time of breast tumours occurring at this age versus the relatively longer screening interval.4

In women with a hereditary risk of breast cancer, these limiting factors are further multiplied. First, women with germ-line mutations, such as BRCA1 and BRCA2, have an increased lifetime risk of breast cancer of between 60% to 80%.3 Secondly, >50% of these women are afflicted with breast cancer before the age of 50.4 It would be logical to begin screening at an earlier age for this population of females with a familial risk of breast cancer. However, if mammography were employed for this group of women, a sensitivity of much lower than 85% would not be unexpected for the reasons stated earlier. In addition, there is a theoretical risk of radiation-induced carcinogenesis in these genetically susceptible women.

Magnetic resonance imaging (MRI) of the breast has been shown to provide a higher sensitivity in diagnosing invasive breast cancer in a general population.7 A recent study by Stoutjesdijk et al8 suggests that MRI performed better than mammography in annual breast cancer

1 Department of Diagnostic Radiology
2 Department of Clinical Research
Singapore General Hospital, Singapore
3 Department of Radiology
University Medical Center St Radboud, Nijmegen, The Netherlands
Address for Reprints: Dr LSJ Sim, Department of Diagnostic Radiology, Singapore General Hospital, Outram Road, Singapore 169608.
surveillance of women with hereditary risk of breast cancer. Being an expensive tool that requires intravenous contrast injection and not being widely available, the usefulness of MRI in breast cancer screening is limited. Its specificity is also relatively low, which may lead to unnecessary biopsies, particularly in pre-menopausal women with spontaneous hormone-induced enhancement. Currently, MRI-guided biopsy is also problematic.

A cheaper, practical and more prevalent imaging alternative, which is not impaired by dense breast parenchyma, is sonography; but what role has breast ultrasound to play in this special group of women? Few reports have been published focusing on this aspect or on the comparison of the different imaging modalities in screening for breast cancer in this high-risk population. This study assesses the performance and value of breast ultrasound in women with familial risk of breast cancer.

**Materials and Methods**

The study data were obtained from the radiological records, pathology reports and human genetics data of the University Medical Centre St Radboud, Nijmegen, The Netherlands. Since 1994, it has been institutional practice to perform annual breast MRI for women with familial risk of breast cancer.

An initial dataset of 245 women used in a prior study by Stoutjesdijk et al., in which they selected all breast cancer surveillance reports utilising mammography or MRI from November 1994 till February 2001, was examined. There were several criteria for selection into that study based on the woman’s lifetime risk of breast cancer and the availability of validation of the radiological interpretation. The former had to exceed 15% judging from the family history of breast or ovarian cancer, or the finding of a germline mutation in the BRCA1 or BRCA2 gene. The latter required that either adequate follow-up imaging with an interval of at least 2 years or histopathological correlation was available. Further details of that study are described elsewhere.

The age of the patient and her category of lifetime risk of breast cancer were recorded. The model of Claus et al. was used to estimate this risk and subdivisions were made as follows: category 1 (mutation carrier) with 50% to 85% lifetime risk; category 2 (very high risk) with 30% to 50% lifetime risk, and category 3 (high risk) with 15% to 30% lifetime risk.

Sixty-six women were excluded from that study due to inadequate follow-up or lack of histopathological correlation. The remaining 179 women formed the basis of this study. Of these, 84 had an ultrasound and MRI.

Histopathological correlation was available from 48 percutaneous core needle or surgical excision breast biopsies performed in 42 of these women. The frequency at which a suspicious lesion warranting biopsy was visible on ultrasound was sought and compared with mammography and MRI. This rates how practical a tool breast sonography can be compared with the other 2 imaging modalities.

The reports on each imaging modality in these women were examined and classified as positive or negative for malignancy. To standardise mammographic interpretations, the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) was employed. Similarly for ultrasound and MRI interpretations, a score of 4 or 5 indicates a positive result and a score of 1, 2 or 3 shows a negative result.

The records of the women who had undergone genetic counselling and testing were cross-matched with the study cohort. Data from the patients’ case notes were also used to establish the genetic risk of the women. To test for BRCA1 and BRCA2 mutations, 5 mL of blood was drawn from patients with at least 30% risk. The assay included protein truncation testing of the large exons and direct sequencing of the remaining exons.

Mammography was performed utilising a Mammomat 3000 (Siemens, Erlangen, Germany) or a Senographe 2000D (GE Medical Systems, Milwaukee, WI, USA) mammography system. Standard medio-lateral oblique and cranio-caudal projections were obtained; magnification or coned compression views were performed when indicated. The procedure was conducted during the second week of the menstrual cycle to avoid mastalgia and irradiating an unsuspected foetus.

Breast ultrasound was performed using a high-frequency probe (10 MHz) on a Toshiba SSA 380 ultrasound unit (Toshiba America Medical Systems, Tustin, CA, USA). In most cases, a directed examination was performed to look for and to characterise a breast lesion that could explain the clinical or imaging (mammography or MRI) abnormality. It also guided percutaneous core needle biopsy of suspicious lesions. Other indications for employing breast ultrasound included poorly interpretable mammograms secondary to dense breast parenchyma or the presence of breast prostheses.

Breast MRI was carried out on a 1.5T system (Magnetom Vision; Siemens, Erlangen, Germany) with a dedicated double-breast coil (CP Breast Array; Siemens, Erlangen, Germany) during the second week of the menstrual cycle to minimise glandular tissue enhancement. The protocol employed was a dynamic contrast-enhanced FLASH 3D sequence, with a repetition time of 8.1 milliseconds, an echo time of 4 milliseconds and a flip angle of 20 degrees. Following a plain series of images, 0.2 mmol/L/kg bodyweight of gadopentetate dimeglumine contrast agent (Magnevist®);
Schering, Berlin, Germany) was injected intravenously via an automated injector (Powerinjector®; Medrad, Pittsburgh, PA, USA). A series of 5 post-contrast images were then obtained. The scanning plane for all images was axial prior to November 1999, with an acquisition time of 80 seconds per series. After November 1999, coronal scanning was introduced to reduce artefacts in the axilla arising from cardiac pulsations. The revised acquisition time for each series was 87 seconds.

For radiological interpretation, all the acquired images were displayed in the axial format. Subtracted images were employed to detect early and late contrast enhancement of tumours. From 1995 onwards, enhancement kinetics curves were drawn and analysed for suspicious breast lesions. Maximum intensity projections were also introduced in early 1999 to improve visualisation of suspicious breast lesions.

Confirmation of reported imaging results was obtained by histology or by subsequent annual examinations. Histopathological correlation was derived from core needle biopsy or excisional biopsy of suspicious lesions. If the histological examination revealed a malignant lesion, all imaging studies from the preceding 2 years were evaluated retrospectively to identify any possible false-negative reports. Unfortunately, for ultrasound, only still images acquired and printed on film could be reviewed. Negative or benign reports were considered true-negative result if the imaging results for the next 2 years were negative. They were also considered true-negative if the malignancy could not be identified on the original images. A false-negative result was defined as an examination originally reported as BI-RADS 1 or 2, but a malignancy was actually visible or demonstrated on another modality. The malignancy must have been detected during the 2 years of follow-up.

All lesions were classified at histopathologic examination. For in situ and invasive carcinomas, the tumour size, histological type and differentiation grade were determined. Ductal carcinoma in situ was graded I, II or III according to the classification of Holland et al. For invasive carcinomas, examination of conventional haematoxylin eosin-stained slides of the axillary lymph nodes was done to look for metastases. Invasive carcinomas were graded using the Elston method and their oestrogen and progesterone receptor status were determined by immunohistochemistry.

The various imaging modalities were compared with the histological examination on sensitivity, specificity, diagnostic accuracy, positive predictive value (PPV) and negative predictive value (NPV). These were each calculated as a percentage for mammography, MRI and ultrasound.

A result was classified as false-negative when a diagnostic test was negative for a histologically confirmed benign lesion. The performance of each imaging modality was compared individually and with the combined test of mammography and ultrasound. To obtain a parameter for diagnostic performance, receiver operator characteristic (ROC) curves were plotted using the BI-RADS scores for each modality and the area under the curve (AUC) was compared. Statistical analysis was obtained for all variables with Fisher’s exact test and Pearson’s χ² test.

Results

There were 48 breast biopsies involving 42 women. Six women had 2 biopsies for different lesions. Five of them had 1 lesion in each breast (occurring at different times) and the last woman had 2 lesions in the same breast. The biopsies were done 4 years apart and the findings indicated that they were not the same lesion.

The mean age of the 42 women was 42.4 years (range, 25 to 58 years). Six (14.3%) women had a personal history of carcinoma of the breast prior to biopsy, while the rest had no such history.

The risk profile of the subjects was classified according to Claus et al. 11.9% were in category 1, 38.1% in category 2, 31% in category 3 and 19% in the unknown risk category.

Eighteen (42.9%) women had mammographic evidence of increased breast density, presumably due to underlying benign breast change. Seventeen (40.5%) women did not show dense breast parenchyma. The remaining 7 (16.7%) women did not have a mammogram.

A total of 40 procedures were performed under imaging guidance and 6 (12.5%) biopsies were done without any imaging guidance. Two (4.2%) cases had a mastectomy without preoperative biopsy. Thirty (62.5%) procedures utilised ultrasound guidance and 7 (14.6%) had mammographic guidance for either localisation excisional biopsy or core needle biopsy. Another 3 (6.3%) procedures were performed under MRI guidance.

Benign lesions accounted for 32 (66.7%) cases, cancer was diagnosed in 15 (31.3%) cases and there was 1 (2.1%) case of unknown pathology.

Twenty-four (50%) lesions were in the right breast, 21 (43.8%) in the left breast and 3 (6.3%) were unspecified. Of the 15 malignant lesions, 8 were in the right breast and 6 in the left breast. One lesion was unspecified. The cancers were detected in 13 women, with 2 women having bilateral but asynchronous lesions occurring at least 1 year apart.

A positive mammogram was seen in 11 (22.9%) cases, a negative mammogram in 30 (62.5%) cases and 7 (14.6%) did not have a mammogram. Twenty-six (54.2%) women tested positive on MRI and the rest were negative.

Ultrasound was not performed in 7 (14.6%) cases. A positive sonographic result was recorded in 20 (41.7%)
cases and a negative result in 21 (43.7%) cases. Two cases had MRI done with no mammogram or ultrasound.

If mammography and ultrasound results were combined, with a positive result in either test implying a combined positive score, there were 25 (54.4%) positive cases and 21 (45.6%) negative cases. In the reverse scenario, there were 12 (26.1%) positive cases and 34 (73.9%) negative cases.

The receiver operator characteristic (ROC) curves plotted from the BI-RADS score for each imaging modality are illustrated in Figure 1. The AUC for MRI, ultrasound, mammography and combined mammography and ultrasound were 0.844 (95% confidence interval [CI], 0.74 to 0.95), 0.712 (95% CI, 0.55 to 0.87), 0.586 (95% CI, 0.40 to 0.77) and 0.761 (95% CI, 0.61 to 0.91), respectively. The differences in the AUC of MRI and combined mammography and ultrasound as well as the other 2 imaging modalities, were statistically significant. The null hypothesis for all imaging modalities was AUC = 0.5 (that is, the modality is not a good diagnostic test). The AUC for MRI was the largest ($P < 0.0005$), indicating that it was the best test, followed by combined mammography and ultrasound ($P = 0.004$) and ultrasound alone ($P = 0.02$). Mammography had the worst AUC ($P = 0.344$).

The AUC, sensitivity, specificity, PPV, NPV, diagnostic accuracy and cancer detection rate of mammography, ultrasound, MRI and combined mammography and ultrasound are summarised in Table 1.

The sensitivity, specificity, PPV, NPV, diagnostic accuracy and cancer detection rate of combined mammography and ultrasound were 92.9%, 62.5%, 52%, 95.2%, 71.7% and 0.28%, respectively, when a positive test in either test was taken as a positive combined test score. The results of the Fisher’s exact test were not significantly different from those obtained by MRI ($P = 1.00$). If a negative result in either test was recorded as a negative combined test score, the respective values are 50%, 84.4%, 58.3%, 79.4%, 73.9% and 0.15%.

Discussion

To assess the role of breast ultrasound, a few assumptions were made in this study. First, the interpretation of the ultrasound findings could be influenced by the availability of ancillary investigations. The radiologist is, therefore, not an unbiased reader. Secondly, breast sonography is heavily operator-dependent. The radiological report is assumed to be the de facto assessment of the breast by a competent sonographer. Thirdly, breast ultrasound as a screening tool is not practised in this institution; usually, only 1 breast is scanned. No information is available on the health of the contralateral breast (unless it was specifically stated in the radiology report). An important premise is that a lesion that is visible on a diagnostic ultrasound can be seen on a screening ultrasound. Lastly, a follow-up radiological examination of the breast by mammography, sonography or MRI 2 years later confirms that the initial report was truly negative.

The mean age of the 42 women with hereditary risk of breast cancer when they first underwent a breast radiological examination was 42.4 years, which is relatively young since the Dutch national screening programme invites women aged from 50 to 75 years. This underscores the need to screen these women earlier than the rest of the female population.

| AUC | 0.586 | 0.712 | 0.844 | 0.761 |
| Sensitivity (%) | 53.9 | 83.3 | 93.3 | 92.9 |
| Specificity (%) | 85.7 | 65.5 | 63.6 | 62.5 |
| PPV (%) | 63.6 | 50 | 53.9 | 52 |
| NPV (%) | 80 | 90.5 | 95.5 | 95.2 |
| Diagnostic accuracy (%) | 75.6 | 70.7 | 72.9 | 71.7 |
| Cancer detection rate (%) | 0.17 | 0.24 | 0.29 | 0.28 |

Table 1. Comparison of the Performance of Mammography, Ultrasound, MRI and Combined Mammography and Ultrasound

AUC: area under the curve; MRI: magnetic resonance imaging; NPV: negative predictive value; PPV: positive predictive value
The risk stratification model of Claus et al\textsuperscript{9} is based on family history and may underestimate the risk in the presence of a highly penetrant susceptibility gene.\textsuperscript{11} Hence, it is not surprising that the correlation with BRCA1 or BRCA2 germ-line mutation is poor. Perhaps, the use of other specific models, such as Couch\textsuperscript{12} or Frank,\textsuperscript{13} which quantitatively estimate the risk of an individual carrying a predisposing mutation in BRCA1 or BRCA2 may improve the correlation.

Breast density was determined by the radiologist reporting on the mammograms, who would comment on mammary dysplasia, sclerosing mastopathy, fibrocystic disease or benign breast change if it had a significant impact on the interpretation of the mammograms. Though subjective, it was a practical method to assess for the presence, or absence, of higher breast parenchymal density. At least 43% of the subjects had mammographically dense breasts. No further extrapolation can be made from the data.

The prevalence of dense breast parenchyma (P2 and DY patterns in the Wolfe classification) in a general population with no known genetic risk participating in a screening programme in Nijmegen, The Netherlands, was 33% at the first mammographic examination.\textsuperscript{14} This is significantly lower than the proportion of high-risk women with dense breasts documented in this study. The higher prevalence of dense breast parenchyma in these women is probably related to their younger mean age. It also explains the reduced sensitivity of mammography. As mammographically dense breasts can obscure an underlying lesion, the data support efforts to devise a screening modality for such women, which is not impaired by dense fibroglandular breast tissues.

Almost two-thirds of the women who had breast biopsy did so under ultrasound guidance. It is a testament to the usefulness and practicality of using ultrasound to detect and biopsy suspicious lesions. As only 3 biopsy procedures were performed under MRI guidance, it highlights the current difficulties associated with MRI-guided procedures and emphasises the need to make them user-friendly and practical in order to gain widespread acceptance.

The benign to malignant biopsy ratio in this study was approximately 2 to 1, a respectable ratio considering the subjects were asymptomatic and carried a genetic risk.

Our results show that MRI was the most sensitive modality, but had the lowest specificity. Ultrasound had intermediate sensitivity, specificity, NPV and cancer detection rate. Mammography had the highest PPV and diagnostic accuracy, while ultrasound had the lowest rate for both parameters. In terms of usefulness in screening, the cancer detection rate of MRI was the highest, followed by ultrasound and mammography in that order.

One of the two published studies comparing magnetic resonance imaging with conventional modalities in screening high-risk women was that by Kuhl et al.\textsuperscript{15} This was a prospective trial where 15 breast cancers were identified; 9 in 192 asymptomatic women and 6 in 6 symptomatic women. In 105 asymptomatic women with validation of the first year screening results, the sensitivities of mammography, MRI and ultrasound were 33%, 100% and 33% (mammography and ultrasound combined was 44%) respectively. The specificities were 93%, 95% and 80% respectively. The sensitivities of mammography and ultrasound reported appear lower compared to our study, while the specificities seem high. This may be related to our smaller sample size. Our results were closer to those obtained from assessing general populations.\textsuperscript{16,17} In particular, the performance of breast ultrasound compares favourably despite the directed approach employed in our study.

Kuhl et al\textsuperscript{15} admitted that breast MRI screening in this young high-risk cohort was difficult due to the atypical imaging features of BRCA-induced breast cancers and the frequent occurrence of spontaneous contrast enhancement. One of her recommendations was that only radiologists with specific expertise with MRI in pre-menopausal women should be reading such examinations. Expertise in performing and interpreting MRI breast examinations of high-risk women is not universally available. Ultrasound being more prevalent and by virtue of its longer existence and hence greater end-user experience, affords wider applicability for screening such individuals.

The other study is by Warner et al\textsuperscript{18} who identified 6 invasive cancers and 1 non-invasive cancer among 196 high-risk women. The sensitivities of mammography, ultrasound and MRI were 33%, 60% and 100% respectively. The corresponding specificities were 99.5%, 93% and 91%. The sensitivity of ultrasound was lower than that of our study, but its specificity was higher. Compared with Kuhl et al\textsuperscript{15} and Warner et al,\textsuperscript{18} the statistical performance of ultrasound had the greatest discrepancy. This can be explained by the fact that ultrasound is heavily operator-dependent.

MRI is not as widely available as ultrasound. The latter commands 25% of the worldwide diagnostic imaging market compared to MRI’s 16%.\textsuperscript{19} The main reason for this discrepancy is cost. A high-end ultrasound machine retails for US$150,000 and some hand-held devices cost only US$20,000. For breast diagnostic work, a mid-range unit with a high frequency transducer is the minimum requirement. In any case, the cost of an ultrasound machine pales in comparison with that of an MRI unit. It is useful to know how ultrasound performs in relation to MRI in visualising lesions detected by the latter. The problems...
encountered in performing percutaneous MRI-guided biopsies have also not been fully resolved. Often, these procedures are time-consuming and technically difficult. The absence of a means to document lesion removal under MRI guidance, other than repeating the MRI scan immediately post-biopsy, is another hindrance to its utility in everyday practice. The lack of real-time visualisation of the biopsy procedure is another disadvantage.

Ultrasound, on the other hand, allows real-time imaging guidance; in experienced hands, it can document accurate lesion biopsy or removal simultaneously. Although ultrasound has not been advocated for screening purposes, a sonographic correlate is recommended for an MRI-detected lesion with a view to biopsy.

A prospective study to evaluate the role of ultrasound screening in high familial risk women and to compare its diagnostic accuracy with that of screening mammography and MRI would be ideal. This is the next step in our research. However, in view of the ethical issues and costs of false-positive ultrasounds leading to negative biopsies, this study has demonstrated the value of ultrasound in a high-risk population, and that it does not lead to unnecessary biopsies.

The results are useful for geneticists and clinicians involved in the management of such women as it suggests that combining mammography with ultrasound increases the sensitivity of breast cancer detection in this group.

Several studies have suggested that ultrasound has a role in screening for breast cancer in women with mammographically dense breasts. A cancer detection rate of approximately 0.3% to 0.4% in an asymptomatic population is expected with ultrasound. This figure would be higher if a prevalent screen with ultrasound was performed since the women involved in the studies already had routine annual or biannual mammography, which would have detected some cancers. The incidence would also be higher in an at-risk population and in Asian women, where there is a higher prevalence of dense breast parenchyma that obscures cancer on mammography.

Although this study does not involve ultrasound screening per se, ultrasound performance can be extrapolated to screening a subgroup of women with a high familial risk of breast cancer where a similar result should not be unexpected. This would have better specificity than MRI, and a decrease of 10% in sensitivity, but at a fraction of the cost of MRI screening. A breast ultrasound examination would typically cost one-tenth of that of a contrast-enhanced MRI study. In many centres, ultrasound is the only alternative to mammography and should not be neglected, particularly for screening high-risk women for breast cancer.

Often, ultrasound is used to supplement mammography in the diagnostic work-up of patients with clinical breast problems. It is also used to reassess the breasts in the presence of an equivocal or suspicious mammographic finding. Ultrasound is increasingly employed to screen poorly interpretable mammograms or mammographically dense breasts. The latter is common in Asian women and explains the importance of ultrasound in routine breast imaging and its widespread acceptance in Asia as compared to America or Europe. It would be reasonable to extend ultrasound to screen high-risk individuals following mammography. The combined mammography and ultrasound test would match MRI in sensitivity, specificity, PPV, NPV and accuracy at a fraction of the costs and should not be ignored in centres that do not have breast MRI expertise.

REFERENCES


