Comparison of Mammography in Combination with Breast Ultrasonography Versus Mammography Alone for Breast Cancer Screening in Asymptomatic Women

Sarawan Boonlikit

Abstract

Aim: To compare the agreement of screening breast mammography plus ultrasound and reviewed mammography alone in asymptomatic women. Materials and Methods: All breast imaging data were obtained for women who presented for routine medical checkup at National Cancer Institute (NCI), Thailand from January 2010 to June 2013. A radiologist performed masked interpretations of selected mammographic images retrieved from the computer imaging database. Previous mammography, ultrasound reports and clinical data were blinded before film re-interpretation. Kappa values were calculated to assess the agreement between BI-RADS assessment category and BI-RADS classification of density obtained from the mammography with ultrasound in imaging database and reviewed mammography alone. Results: Regarding BI-RADS assessment category, concordance between the two interpretations were good. Observed agreement was 96.1%. There was moderate agreement in which the Kappa value was 0.58% (95%CI; 0.45, 0.87). The agreement of BI-RADS classification of density was substantial, with a Kappa value of 0.60 (95% CI; 0.54, 0.66). Different results were obtained when a subgroup of patients aged ≥60 years were analyzed. In women in this group, observed agreement was 97.6%. There was also substantial agreement in which the Kappa value was 0.74% (95% CI; 0.49, 0.98). Conclusions: The present study revealed that concordance between mammography plus ultrasound and reviewed mammography alone in asymptomatic women is good. However, there is just moderate agreement which can be enhanced if age-targeted breast imaging is performed. Substantial agreement can be achieved in women aged ≥60. Adjunctive breast ultrasound is less important in women in this group.

Keywords: Mammography plus ultrasound - adjunctive breast ultrasound - supplement breast ultrasound

Introduction

Breast cancer is the most common neoplasm of the Thai female population (Attasara and Buasom, 2011). The incidence of breast cancer in Asian and Thai women increased significantly each year (Bhothisuwan, 2004; Shin et al., 2010). Breast cancer is one clear example of excellent survival statistics when early-stage disease is treated using current therapies. Currently, numerous clinical methods are used in breast cancer screening and diagnosis. The most effective screening technique at the present time is mammography. In fact, it is the only widely used imaging modality for breast cancer screening. Several large randomized clinical trials have shown that mammography reduces mortality from breast cancer (Shapiro et al., 1971; Thrurfjell and Lindgren, 1996; Hendrick et al., 1997; Tabar et al., 2001). However, it is not yet a perfect technique. The overall sensitivity of mammography for breast cancer detection is moderate (67-78%) (Kolb et al., 2002; Oestreicher et al., 2005; Jensen et al., 2006; Mo et al., 2013). The method cannot accurately distinguish between benign and malignant tumors. In addition, sensitivity of mammography is reduced in women with dense breasts because dense fibroglandular tissue may obscure calcifications and masses. In those women, the normal breast tissue and the tumor are difficult to distinguish from each other on mammogram (Kolb et al., 2002; American College of Radiology, 2003; Carney et al., 2003; Devolli-Disha et al., 2009). Moreover, dense breast tissue itself is a marker of increased risk of breast cancer on the order of 2- to 6-fold (Warner et al., 1992; Boyd et al., 1995; Boyd et al., 2011). Techniques such as magnetic resonance imaging (MRI) and ultrasound are used as the adjunctive tools to mammography. The addition of ultrasonography screening may detect those tumors that are missed by mammography alone. It has been shown to be effective in distinguishing cystic lesions from solid lesions and to further differentiate benign solid masses from malignant solid masses (Stavros et al., 1995; Skaane and Engedal, 1998; Rahbar et al., 1999; Bhothisuwan, 2004). Prevalence studies in women with radiographic dense breasts have shown that approximate...
3-4 cancers per 1,000 women are detected by ultrasound only (Gordon and Goldenberg, 1995; Kolb et al., 1998; Buchberger et al., 1999; Berg et al., 2008). Ultrasound has become a valuable tool to use with mammograms because it is widely available, noninvasive, and less costs than other tools. Kolb et al. (2002) showed improved sensitivity (97% versus 74%) when adjunctively used with mammography compared to physical examination with mammography (Kolb et al., 2002). However, there was also a substantial decrease in specificity and positive predictive value with mammography plus ultrasound compared to mammography alone (Berg et al., 2008). The limitations of breast ultrasound as a screening tool are well known. For the example, ultrasound test value depends on the operator’s level of skill and experience. It needs a real time interpretation which is not therefore available for subsequent audit and review. In addition, ultrasound needs skilled operator and it is labor intensive, the on-site radiologist must be available. There is currently no standardized examination technique or interpretation criterion for breast ultrasound. It has a high false positive rate and is less sensitive than breast MRI (Kolb et al., 1998; Buchberger et al., 1999; Gordon, 2002) which might increase the number of biopsies and unnecessary treatment. Based on systematic review, little evidence-based support is found to confirm the well-recognized value of ultrasonography as an adjunct to mammography in the detection of breast cancer in clinical practice and routine use of ultrasonography as an adjunct screening tool in women at average risk for breast cancer is not justified (Flobbe et al., 2002; Gartlehner et al., 2013). In Thailand, the limitations to implement widespread screening ultrasound include a shortage of qualified personnel to perform and interpret the examination and lack of standardized scanning protocols. So far, performing and interpreting of breast ultrasound are limited to radiologists only and the development in training of qualified technician capable of this procedure is delayed. Many rural areas have geographic and economic disadvantage for breast screening by radiologist. These problems have hampered widespread use of breast ultrasound as a adjunctive of mammography in this country.

In the present study, the author investigated the performance of screening ultrasound in conjunction with mammography, using a standard protocol and criteria. The objective is to compare the intraobserver agreement of screening breast mammography plus ultrasound and reviewed mammography alone in asymptomatic women. We also assessed intraobserver agreement of the mammographic interpretation of density.

**Materials and Methods**

The present study was conducted with institutional review board approval. The study protocol was approved in September 2013. All breast imaging data retrieved for the present study were taken from women who presented for routine medical checkup at National cancer institute (NCI), Thailand from January 2010 to June 2013. Breast imaging data were eligible for inclusion criteria under the following circumstances: (1) cases previously interpreted by the assigned radiologist (the author), (2) imaging data included of mammography and breast ultrasonography (3) adequate imaging and clinical data. Exclusion criteria were (1) women who had previously diagnosed or concurrent breast cancer, (2) cases who had any breast symptoms, (3) cases with palpable breast mass , (4) cases with breast implant and (5) inadequate data. Mammography was performed in craniocaudal and mediolateral oblique views using a digital mammography machine (Lorad™ Selenia and after 2012; Dimension) (Lorad coporation, Danbury, USA). Ultrasound machine was Aixplorer multiwave (Supersonic Imagine, Aix en Province, France) with the high resolution 8 MHz. linear probe. Visually estimated overall breast density pattern on mammograms was categorized into four categories using a 4-level density scale of BI-RADS classification of density; type 1. entirely fat, type 2. scattered fibro glandular, type 3. heterogeneously dense and type 4. extremely dense (American College of Radiology, 2003). In the event of discordance in the density of the right and left breast, the woman was classified according to the higher density classification. Assessments for each lesion and for each breast overall were recorded on the BI-RADS assessment categories: 1. negative; 2. benign; 3. probably benign; 4. suspicion; and 5. highly suggestive of malignancy (American College of Radiology, 2003). Because of the mammography interpretation largely depends on experience of radiologists. In the present study, the selected radiologist had special interest in this area and 14 years of experience interpreting mammograms before entering the study. Routinely, she has been interpreting an average of 20 mammograms per week. After the breast imaging data were retrieved from computer database, the radiologist then performed separate, masked interpretations of selected mammographic images retrieved form computer imaging database. She was blinded to her own previous mammographic and ultrasound reports and was not aware of a woman’s age or clinical data before film interpretation. The mammogram images of the contralateral breast of the same case was then displayed and reviewed in the same way. The entire process was repeated for each case until all eligible breast imaging data were completely evaluated. In the present study, for the purpose of analysis, the categories of the BI-RADS scale were collapsed into two-tiered category in terms of whether additional evaluation was required. In general, BIRADS category 1-3 will be responded by follow-up and therefore grouped together while BIRADS category 4-5 will be responded by tissue biopsy and grouped in the same category. Kappa values were calculated to assess the agreement between BIRADS assessment category and BIRADS classification of density obtained from the mammography with ultrasound and mammography alone. A Kappa value of 0% indicates that the agreement between two values is no greater than would be expected by chance. Perfect agreement between two tests is reflected by a value of 0.8-1, substantial agreement; 0.6-0.79, moderate agreement; 0.4-0.59 and fair agreement; 0.2-0.39. The data were analyzed using the statistical package for the social sciences (SPSS) (SPSS Inc, Illinois, USA) version 17.0.
Results

During the study period, 635 Thai women were eligible. The average age was 54.7±8.92 years (range 30-87 years). Four hundred and forty-seven (70.4%) and 167 (26.3%) women were aged ≥50 and ≥60 years, respectively. According to a density scale of BI-RADS classification in breast imaging data, it was almost entirely fat (type 1) in 23 women (3.6%), scattered fibroglandular (type 2) in 72 women (11.3%), heterogeneously dense (type 3) in 475 women (74.8%) and extremely dense (type 4) in 65 women (10.2%). Regarding the agreement of BI-RADS classification of density, observed agreement was 83% and the Kappa value for the agreement was 0.60 (95% CI; 0.54, 0.66). The data were shown in Table 1. Regarding BI-RADS assessment category, concordance between the two interpretations were good. Observed agreement was 96.1%. Of the 635 mammography with ultrasound, 591(93%) were assigned the same BI-RADS (1-3) category as reviewed mammography alone and a 19 (2.9%) were assigned the same BI-RADS (4-5) category as reviewed mammography alone. Eleven cases with BI-RADS (1-3) category by mammography with ultrasound (1.73 %) were over assigned as BI-RADS (4-5) category by reviewed mammography alone. Fourteen cases with BI-RADS (4-5) category by mammography with ultrasound (2.20%) were under assigned as BI-RADS (1-3) category by reviewed mammography alone. Table 2 displayed the observed agreement, agreement between mammography with ultrasound and reviewed mammography alone. There was moderate agreement in which the Kappa value was 0.58% (95% CI; 0.45, 0.87).

The different results were obtained when subgroup of patients aged ≥60 years were analyzed (Table 3). In women in this group, observed agreement was 97.6%. There was also substantial agreement in which the Kappa value was 0.74% (95% CI; 0.49, 0.98). As expected, a distribution of each density scale in BI-RADS classification was changed. It was almost entirely fat (type 1) in 13 women (7.78%), scattered fibroglandular (type 2) in 36 women (21.6%), heterogeneously dense (type 3) in 111 women (66.5%) and extremely dense (type 4) in 7 women (4.19%).

Discussion

The data presented in this report indicated that concordance between mammography with ultrasound and reviewed mammography alone in Thai women was good. Observed agreement was 96.1%. The Kappa value was 0.58% in the present study, which was moderate agreement. Our concern is the problem in the cases with under-interpreting. Fourteen cases (2.20%) with BI-RADS (4-5) category by mammography with ultrasound were under assigned as BI-RADS (1-3) category by reviewed mammography alone. The result of under-interpreting may be an important barrier to use mammography alone as a screening because it can result in delayed diagnosis, the advanced stage of breast cancer at the time of diagnosis and delayed treatment. In these 14 women, 11 women had dense breast as they were in type 3 of BI-RADS classification of density. That is possible explanation and confirms the disadvantage of using mammography in dense breast tissue. Mammographic sensitivity declined significantly with increasing breast density (Kolb et al., 2002; Carney et al., 2003; Titus-Ernstoff et al., 2002; Carney et al., 2003; Devolli-Disha et al., 2009). As known, dense breast tissue is common: approximately half of women younger than 50 years and a third of older women have dense breast parenchyma (Stomper et al., 1996). Compared to European breasts, Thai breasts differ in the composition of breast tissue. There is a study in this country showing that heterogeneously dense and extremely dense are very common (59% and 14%, respectively) (Bhothisuwan, 2004). Data from the present study (74.8% and 10.2%) also supported a previous report. It is well accepted that mammographic density declines progressively with age, especially after menopause and because the breast tissue becomes less dense and more fatty in women in this age, mammography might be easier to be interpreted (Stomper et al., 1996; Boyd et al., 2002; Carney et al., 2003; Titus-Ernstoff et al., 2006). Mammographic sensitivity declined significantly with increasing breast density and in younger women with dense breasts (Kerlikowske et al., 1996; Kolb et al., 2002; Carney et al., 2003). There is a study reported a

Table 1. Agreement of BI-RADS Classification of Density between Reviewed Mammography and Original Mammography by the same Radiologist

<table>
<thead>
<tr>
<th>Original Mammography</th>
<th>Total</th>
<th>Kappa (95% CI)</th>
<th>Observed agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reviewed</td>
<td>1</td>
<td>6</td>
<td>101</td>
</tr>
<tr>
<td>Mammography</td>
<td>2</td>
<td>37</td>
<td>66 (0.54, 0.66)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>72</td>
<td>475</td>
</tr>
</tbody>
</table>

*Boldface values indicate exact agreement

Table 2. Agreement of BI-RADS Assessment scale* between Reviewed Mammography and Mammography with Ultrasound by the same Radiologist

<table>
<thead>
<tr>
<th>Mammography with ultrasound</th>
<th>Total</th>
<th>Kappa (95% CI)</th>
<th>Observed agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reviewed Mammography alone</td>
<td>1</td>
<td>6</td>
<td>101</td>
</tr>
<tr>
<td>Birads 1-3</td>
<td>591</td>
<td>14</td>
<td>605</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Bisrad 4-5</td>
<td>11</td>
<td>19</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td>602</td>
<td>33</td>
<td>635</td>
</tr>
</tbody>
</table>

*Presented in two-tiered modification

Table 3. Agreement of BI-RADS Assessment scale* between Reviewed Mammography and Mammography with Ultrasound by the same Radiologist in women aged ≥60

<table>
<thead>
<tr>
<th>Mammography with ultrasound</th>
<th>Total</th>
<th>Kappa (95% CI)</th>
<th>Observed agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reviewed Mammography alone</td>
<td>1</td>
<td>6</td>
<td>101</td>
</tr>
<tr>
<td>Birads 1-3</td>
<td>157</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>159</td>
<td>8</td>
<td>167</td>
</tr>
</tbody>
</table>

*Presented in two-tiered modification
The goal of an organized mammography screening program is to offer high level medical care to every woman at a reasonable expense. Previous studies have shown barriers to entry into the health system for low income women and women in an area with inadequate mammography capacity, resulting in delays in diagnosis (Elkin et al., 2010; Maly et al., 2011). Providing mobile mammography services with community organizations, can be effective in increasing access and decreasing barriers to screening hard-to-reach populations. On-site mammography at community-based sites where women gather is an effective method for increasing breast cancer screening rates among underserved women (Brooks et al., 2013; Fontenoy et al., 2013). In our institute, National Cancer Institute, we have mobile mammography service team which includes advanced practice staffs, workers and technical support staffs. However, commonly raised concerns about mobile mammography include quality control, cost-effectiveness (Brooks et al., 2013), and in our country; lacking of radiologists. As mentioned earlier, to do mammography plus ultrasound, an abnormality must be perceived while ultrasound scanning for it to be documented. On-site radiologists are mandatory and this is unlikely to achieve in this country. The other barrier to implementing screening ultrasound is the risk of false positive results. (ie, biopsy with benign results or short interval follow-up). Another minor concern is regarding patient anxiety and discomfort induced by addition of screening ultrasound and the more time consuming in breast screening. The clinical implication of the present study is that the data show the good efficacy of mammography alone in the aged women and supplemental radiologist-performed ultrasound may be omitted safely among women in this age group. This concept seems to be very promising. Nevertheless, this issue regarding performing mammography in postmenopausal women alone without the adjunctive breast ultrasound is just our proposal. These are many issues worthy of future study. Nationally, government policy should encourage specialized training of technologists to counter a current shortage of qualified radiologist and technologist personnel. In addition to less sensitivity, the other disadvantage of omitting the adjunctive breast ultrasound is that BIRADS category 4 cannot be subcategorized (4a, 4b and 4c). BI-RADS category 4 indicates that a lesion has been recommended for biopsy but it provides no frame of reference for either the referring physician or the patient as to the risk for malignancy. Dividing category 4 lesions into those with a small (category 4a), moderate (category 4b), or substantial (category 4c) (American College of Radiology, 2003) likelihood of malignancy better informs the physician and patient as to the level of concern regarding the lesion and prepares both the physician and the patient for the likely biopsy findings and the potential need for follow-up.

The present study has several strengths. First, study population represents women seen in clinical practice because data were selected from a screening population. Therefore, the results closely approximate that expected in actual practice. Secondly, literature has shown that interobserver agreement for the BI-RADS assessment is only fair to moderate (Elmore et al., 1994; Beam et al., 1996; Kerlikowske et al., 1998; Redondo et al., 2012). To warrant consistent interpreting, the study was designed to exclude the factor of interobserver variation that all breast imaging were interpreted by the same experienced radiologist. Consequently, the certain limitation of the study is that it was based on the performance of one radiologist. Therefore, results may not be generalizable to all practicing radiologists, in particular, those who have less experienced with BI-RADS or read a low volume of breast screening. In addition, another factors not considered and reported in this analysis was the detail of clinical characteristics such as hormonal replacement therapy, body mass index (BMI), menstrual cycle phase or parity. There were not data regarding those parameters in the eligible women that they will help to represent the detail of the study population. Another limitation is the possibility of recognizing the cases by the radiologist as the cases were interpreted by the same radiologist prior to this study. Therefore, some later cases close to the start of the study may influence breast imaging categorization. In summary, the present study reveals that concordance between mammography with ultrasound and reviewed mammography alone in asymptomatic women is good. However, there is just moderate agreement which can be enhanced if age-targeted breast imaging was performed. Substantial agreement can be achieved in women aged ≥60. Adjunctive breast ultrasound is less important in women in this group.

Acknowledgements

The author wishes to thank associate professor Dr.
Sathone Boonlikit as the statistical consultant and Miss Anothai Intakawin for her help in data collection

References


Sarawan Boonlikit

