BI-RADS breast density classification - an international standard?

Poster No.: C-1762  
Congress: ECR 2013  
Type: Scientific Exhibit  
Authors: N. Sauber, A. Chan, R. Highnam; Wellington/NZ  
Keywords: Quality assurance, Screening, Mammography, Breast  
DOI: 10.1594/ecr2013/C-1762

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.

As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method is strictly prohibited.

You agree to defend, indemnify, and hold ECR harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.

Please note: Links to movies, ppt slideshows and any other multimedia files are not available in the pdf version of presentations.

www.myESR.org
Purpose

Mammographic breast density describes the proportion of radiopaque, fibroglandular tissue in the breast. Higher breast densities are strongly associated with an increased risk of breast cancer, with women reported to have a four- to six-fold increase in breast cancer risk if they have extremely dense breasts, compared to women with predominantly fatty breasts [1-3]. Higher densities are also correlated with a decrease in the sensitivity of mammographic screening, due to the potential masking of tumours in higher-attenuating areas of fibroglandular tissue [2].

The American College of Radiology (ACR)’s Breast Imaging Reporting and Data System (BI-RADS) aims to provide a standardized classification system for reporting mammographic breast densities. BI-RADS 4th Edition classifies breast density into four categories based on a visual assessment, by a radiologist, of the proportion of fibroglandular tissue (see Figure 1): (1) Almost entirely fatty (<25% glandular); (2) Scattered fibroglandular densities (approximately 25-50% glandular); (3) Heterogeneously dense (approximately 51-75% glandular); and (4) Extremely dense (>75% glandular). Several states in the US (i.e. Connecticut, Texas, Virginia, New York and California) now have legislation in place requiring the mandatory reporting of breast densities. As additional states and countries follow suit, a more objective method for determining breast density will become increasingly important.

Although widely used worldwide, the assessment of breast density by eye is known to be subjective. Previous studies have highlighted the inter- and intra-observer variability amongst radiologists, when classifying breast densities according to the BI-RADS system [4, 5]. In particular, the inter-reader agreement for the intermediate categories appears to cause the most problems in terms of consistency, with Ciatto et al. (2005) reporting #-values of 0.25 and 0.28, for BI-RADS 2 and 3 categories, respectively. The BI-RADS system originated in the US, and while no difference in the inter-observer agreement was observed for studies based in the US versus outside the US [6], it is possible that inter-region or inter-country differences in the visual assessment of mammographic densities also exist. We conducted a short study to investigate country-to-country variability in BI-RADS density assessment, and assess the agreement between radiologists' BI-RADS scores compared to an automated method, Volpara™. Volpara™ currently has FDA approval to give out a BI-RADS density assessment.

Images for this section:
Fig. 1: Examples of mammographic images (CC views) from each of the four BI-RADS density categories. (A) BI-RADS 1/almost entirely fatty; (B) BI-RADS 2/scattered fibroglandular densities; (C) BI-RADS 3/heterogeneously dense; and (D) BI-RADS 4/extremely dense. Fibroglandular tissue is more highly-attenuating than fatty tissue, and appears brighter on a mammogram.
Methods and Materials

Mammographic images

The "For Processing" and "For Presentation" mammographic images were obtained for 12 cases from a US site. Craniocaudal (CC) and medio-lateral oblique (MLO) mammographic views were taken of each breast, for each case. Each woman had given their prior consent to their mammograms being used for research purposes.

BI-RADS assessment by radiologists

An experienced radiologist from South Korea, the Netherlands and the US, reviewed the "For Presentation" mammograms from each of the 12 cases. Using the ACR guidelines for reporting breast density, and without any additional training prior to viewing the randomized mammograms, each radiologist was asked to categorize each case according to the four BI-RADS density scores (see Figure 1), by making a visual assessment by eye. Each reader was blinded to the other readers' assessments.

Automated assessment of BI-RADS using Volpara™

The "For Processing" data was also run through Volpara™, an automated volumetric breast density assessment system to obtain an objective BI-RADS score (Volpara Density Grade; VDG). The thresholds for assigning VDG/BI-RADS scores differ to the ACR BI-RADS system, because the software uses quantitative volumetric measurements of fibroglandular and fatty tissue to calculate volumetric percent density, and hence, VDG (i.e. 1 = <4.5%; 2 = 4.5 - 7.5%; 3 = 7.6 - 15.5%; 4 = >15.5%) rather than an area-based visual assessment.

Analyses

The average BI-RADS score over all the cases was determined for each individual radiologist, as well as for Volpara™. The percentage agreement between each of the radiologists and between each of the radiologists and Volpara™ was also determined for each BI-RADS category.

Images for this section:
Fig. 1: Examples of mammographic images (CC views) from each of the four BI-RADS density categories. (A) BI-RADS 1/almost entirely fatty; (B) BI-RADS 2/scattered fibroglandular densities; (C) BI-RADS 3/heterogeneously dense; and (D) BI-RADS 4/extremely dense. Fibro glandular tissue is more highly-attenuating than fatty tissue, and appears brighter on a mammogram.
Results

Country-to-country variability by BI-RADS categories

The average BI-RADS scores (over the 12 cases; see Table 1) by the South Korean and Netherlands radiologists were 2.33 and 2.17, respectively. The average BI-RADS score, determined by both the US radiologist and Volpara™, was 2.67. Compared to the South Korean radiologist, the US radiologist assigned a higher BI-RADS density score in five out of the 12 cases, and a lower BI-RADS score in only one case. Similarly, Volpara™ gave a higher BI-RADS score for four of the cases compared to the South Korean radiologist. The South Korean and Netherlands radiologists gave lower BI-RADS scores in five and six cases, respectively, compared to the US radiologist.

The level of agreement between the various assessments was quite varied. The percentage agreement (see Table 2) between the US radiologist and both the South Korean and Netherlands radiologists was only 50%. The agreement between Volpara’s™ BI-RADS scores and those of the three radiologists was between 50-67%. The highest level of agreement (83%) was observed between the South Korean and Netherlands radiologists. Examples of mammograms (i.e. case #3, #5 from Table 1) that showed a poor agreement between the BI-RADS scores are shown in Figure 2.

Images for this section:
Table 1: BI-RADS density scores, as assessed visually by radiologists from different countries, are shown for each of the 12 cases. For comparison, the cases were also run through Volpara™ to generate an objective BI-RADS density score.

<table>
<thead>
<tr>
<th>Case</th>
<th>South Korea</th>
<th>Netherlands</th>
<th>US</th>
<th>Volpara™</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Average</td>
<td>2.33</td>
<td>2.17</td>
<td>2.67</td>
<td>2.67</td>
</tr>
</tbody>
</table>

Table 2: The percentage agreement of BI-RADS density scores, as assessed by the three radiologists (from South Korea, the Netherlands or the US) or by Volpara™.

<table>
<thead>
<tr>
<th></th>
<th>South Korea</th>
<th>Netherlands</th>
<th>US</th>
<th>Volpara™</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Korea</td>
<td>83.3%</td>
<td>50%</td>
<td></td>
<td>66.7%</td>
</tr>
<tr>
<td>Netherlands</td>
<td></td>
<td>50%</td>
<td></td>
<td>50%</td>
</tr>
<tr>
<td>US</td>
<td></td>
<td></td>
<td></td>
<td>66.7%</td>
</tr>
<tr>
<td>Volpara™</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 2: Examples of mammographic images (CC views) from cases where there was not a 100% agreement in the assignment of BI-RADS density scores. See Table 1 for the actual BI-RADS scores as determined by each radiologist or Volpara™. For case #3, all assessors said 3 apart from the US radiologist who said 4; for case #5, South Korea and Netherlands said 2, Volpara and US said 3.
Conclusion

The fact that the average BI-RADS score for both Volpara™ and the US radiologist was the same (2.67) is not unexpected, as Volpara™ was set based on a US radiologist’s assessment of BI-RADS density. The highest level of agreement was between the South Korean and Netherlands radiologists (83%). However, it should be noted that when there were discrepancies between the US radiologist’s scores compared to the other radiologists’ BI-RADS readings, both the South Korean and Netherlands radiologists tended to give lower BI-RADS scores. As further validation studies are carried out comparing how well the Volpara™ BI-RADS scores correlate with those of international radiologists, it will be important to take into consideration any potential country bias in how they judge BI-RADS.

BI-RADS breast density assessment is known to vary markedly between readers, but there appears to also potentially be a systematic bias country-by-country. Moving towards a more standardized approach to breast density assessments will aid in interpreting and comparing breast cancer risk studies that are carried out internationally. The use of objective automated software, such as Volpara™, should allow for this.

References

Personal Information

Natascha Sauber, Matakina Technology Ltd, Wellington, New Zealand; natascha.sauber@matakina.com

Ariane Chan PhD, Matakina Technology Ltd, Wellington, New Zealand; ariane.chan@volparadensity.com

Ralph Highnam PhD, Matakina Technology Ltd, Wellington, New Zealand; ralph.highnam@volparadensity.com