Effect of independent double and multiple reading of screening mammograms by breast density

Poster No.: C-0358
Congress: ECR 2014
Type: Scientific Exhibit
Authors: N. Karssemeijer¹, R. Hupse¹, G. J. den Heeten¹, M. Lobbes², R. M. Mann¹; ¹Nijmegen/NL, ²Maastricht, LB/NL
Keywords: Breast, Mammography, Observer performance, Cancer
DOI: 10.1594/ecr2014/C-0358

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
Aims and objectives

The degree of mammographic breast density (i.e. the proportion of fibroglandular or "dense" tissue in the breast) has a considerable impact on the sensitivity of screening mammography. In women with very fatty breasts, sensitivity is often reported between 75% and 98%, whereas in dense breasts, sensitivity has been reported as low as 48% [1-4]. A large-scale trial of digital mammography (DMIST) noted sensitivity was less than 60% in younger, pre- and perimenopausal women with dense breasts, compared to 86% in women with fatty breasts [5].

In breast cancer screening programs, it is necessary to achieve an optimal balance between sensitivity and recall rates. The interpretation of screening mammograms is often challenging and highly variable, but the addition of multiple readers can improve cancer detection rates. Compared to single reading, double reading increases cancer detection rates by 5-17% [6] and is, therefore, an important consideration when designing screening strategies for a given population.

While numerous variations in screening protocols have been studied, they have largely sought to apply one protocol across entire populations. Given the substantial effects of breast density on mammographic sensitivity, it is not unreasonable to assume that tailoring screening protocols by breast density may improve on current strategies. In this study, our aim was to investigate the effect of increasing the number of independent readers on mammography sensitivity, for dense versus non-dense cases.
Methods and materials

Image Dataset

200 full-field digital mammography cases (all from Hologic Selenia Systems), acquired between 2003 and 2008 from the Preventicon screening centre in Utrecht (the Netherlands), were used in this retrospective reader study. All images were anonymized and institutional review board approval was waived. Only cases with bilateral craniocaudal (CC) and mediolateral oblique (MLO) views were selected. The final dataset consisted of 63 randomly selected screen-detected cancers, 17 cancers that were retrospectively judged to be visible, 20 false-positive referrals, and 100 non-cancer cases. Interpretation of this dataset was more challenging than for a typical screening population, as obvious lesions and obvious benign abnormalities were excluded, as well as cases with only microcalcifications present.

Mammogram Interpretation

12 readers participated in the study; nine were certified breast screening radiologists with 1-24 years of experience and 3 were residents. Readers were informed about the approximate proportion of abnormal cases and that no microcalcifications were included in the dataset, and also had access to prior mammograms for comparison (digital or film). After a short training session on the features of the display station (e.g. zooming and contrast), readers were asked to identify all potential abnormalities. Readers marked the location of each suspicious finding, assigned a suspiciousness score (0-100), and noted if they would refer the case or not in normal screening practice.

Analysis of sensitivity by number of readers and breast density

Breast cancer localization performance was determined by correlating and averaging reader scores corresponding to the same findings. Based on the suspiciousness scores given for each lesion, "positive" thresholds were set, such that 10-20% of non-cancer cases were referred as positive by the reader(s) (i.e. fixed false positive rate of 0.1-0.2). Cancer location was determined to be correct if the distance to the center of the mass was within 2 cm. The mean True Positive Fraction (TPF) for single, double and multiple readers, was determined and stratified by volumetric breast density (VBD). VBD was determined for each case using automated software (VolparaDensity™, version 1.4, Matakina Technology, Wellington NZ), with cases designated as dense (>10% VBD; 60 cases) or non-dense (< 10% VBD; 140 cases). The 10% threshold is approximately a mid-BIRADS 3 breast density category. It was chosen in order to obtain sufficient numbers of dense cases for the analysis.
Results

Examples of screening studies categorized as non-dense and dense are shown in Figures 1 and 2, respectively. Increasing the number of readers improved cancer detection rates. For non-dense cases, the mean TPF increased from 0.68 for a single reader, to 0.74, 0.77, and 0.78 for two, three and four readers, respectively (see Figure 3). Similarly, in dense cases, the mean TPF increased from 0.57 for a single reader to 0.64, 0.67 and 0.69 for two, three and four readers, respectively.

Interpretation of dense mammograms by three readers, increased cancer detection rates to levels that were comparable with a single reader in non-dense breasts. Notably, the incremental gain in sensitivity that was observed for up to 9 readers in dense breasts, was not observed for non-dense mammograms (see Table 1). There was very little improvement in detection performance for more than four readers in non-dense cases, indicating a plateau in sensitivity gain. Compared to a single reader, double reading increased sensitivity by 9.6% and 12.2% for non-dense and dense cases, respectively, whereas, the maximum increase in sensitivity was 29.7% (12 readers) and 15.5% (six readers), respectively.
Fig. 1: Example of a mammographic study that was categorized as "non-dense" in this study (i.e. average volumetric breast density <10%). For presentation images (A) and the corresponding density results from VolparaDensity (B) are shown.

© Radiology, Radboud University Nijmegen Medical Center - Nijmegen/NL
Fig. 2: Example of a mammographic study that was categorized as "dense" in this study (i.e. average volumetric breast density >10%). For presentation images (A) and the corresponding density results from VolparaDensity (B) are shown.

© Radiology, Radboud University Nijmegen Medical Center - Nijmegen/NL
**Fig. 3:** Location receiver operating characteristic curves showing the effect of multiple readers on mammogram sensitivity, for dense versus non-dense cases. The mean true positive fractions were determined at a fixed false positive rate (0.1-0.2).

© Radiology, Radboud University Nijmegen Medical Center - Nijmegen/NL

<table>
<thead>
<tr>
<th># Readers</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-dense</strong></td>
<td>TPF</td>
<td>0.679</td>
<td>0.745</td>
<td>0.770</td>
<td>0.777</td>
<td>0.782</td>
<td>0.785</td>
<td>0.783</td>
<td>0.783</td>
<td>0.783</td>
<td>0.784</td>
<td>0.779</td>
</tr>
<tr>
<td></td>
<td>% increase</td>
<td>ref</td>
<td>9.6</td>
<td>13.3</td>
<td>14.3</td>
<td>15.1</td>
<td>15.5</td>
<td>15.3</td>
<td>15.3</td>
<td>15.3</td>
<td>15.4</td>
<td>14.7</td>
</tr>
<tr>
<td><strong>Dense</strong></td>
<td>TPF</td>
<td>0.569</td>
<td>0.638</td>
<td>0.666</td>
<td>0.690</td>
<td>0.707</td>
<td>0.713</td>
<td>0.724</td>
<td>0.730</td>
<td>0.734</td>
<td>0.735</td>
<td>0.737</td>
</tr>
<tr>
<td></td>
<td>% increase</td>
<td>ref</td>
<td>12.2</td>
<td>17.1</td>
<td>21.3</td>
<td>24.2</td>
<td>25.3</td>
<td>27.3</td>
<td>28.3</td>
<td>29.0</td>
<td>29.1</td>
<td>29.5</td>
</tr>
</tbody>
</table>

**Table 1:** Table showing the percentage increase in mean TPF for increasing numbers of readers (compared to a single reader), in dense versus non-dense breasts.

© Radiology, Radboud University Nijmegen Medical Center - Nijmegen/NL
Conclusion

In the United States, breast cancer screening typically comprises annual mammography and single reading with CAD, whereas European guidelines recommend bi-ennial mammography with independent double reading. Double interpretation of mammograms is a proven, effective strategy to improve cancer detection rates [6]. The significant improvement in sensitivity that was observed in this study (for up to 12 readers) is in agreement with previous work [7], and indicates that there is wide scope for improving detection rates.

The benefits of additional readers on sensitivity was more pronounced in dense cases, suggesting that tailored screening based on breast density has the potential to improve breast cancer detection rates for this population of women. The allocation of multiple readers to mammograms that may be difficult to interpret offers a relatively simple and effective approach for dealing with dense breasts in screening mammography; especially with automated density assessment systems and density-based workflow tools. It should be noted though, that this work was restricted to cancers that are retrospectively visible. Therefore the proposed method can only partly solve the problem. To detect cancers in dense breasts that are mammographically occult, additional screening modalities are required.
Personal information

Nico Karssemeijer, PhD, Department of Radiology, Radboud University Nijmegen Medical Centre, Netherlands; n.karssemeijer@radboudumc.nl

Rianne Hupse, PhD, Department of Radiology, Radboud University Nijmegen Medical Centre, Netherlands; rianneswart@upcmail.nl

Gerard J den Heeten, MD, PhD# Academic Medical Centre, Amsterdam, Netherlands and National Expert and Training Centre for Breast Cancer Screening, Nijmegen, Netherlands; g.denheeten@lrcb.nl

Marc B Lobbes, MD, PhD# Maastricht University Medical Center, Maastricht, Netherlands; marc.lobbes@mumc.nl

Ritse M Mann, MD, PhD; Department of Radiology, Radboud University Nijmegen Medical Centre, Netherlands; ritse.mann@radboudumc.nl
References