Impact of Magnification Views on the Characterization of Microcalcifications in Digital Mammography

Zusammenfassung


Ergebnisse: Die Area under curve (AUC) betrug für alle Reader kombiniert 0,664 ± 0,052 für MG und 0,813 ± 0,042 für MG + MAG, einer signifikanten Verbesserung von 0,148 ± 0,120 entsprechend. Jeder Reader erreichte eine höhere AUC für MG + MAG als für MG, diese Verbesserung war bei vier Readern signifikant. In 76,34 % der Fälle zeigte MG + MAG eine bessere Sichtbarkeit der Kalzifikationen. Bei 33 % waren geringer mehr, bei 39 % deutlich mehr Verkalkungen abgrenzbar.


Abstract

Purpose: To evaluate the additional benefit of true geometric (air-gap) magnification views for the characterization of microcalcifications in digital mammography.

Materials and Methods: After ethical approval, we retrospectively reviewed patient records to identify 100 patients with suspicious microcalcifications (35 malignant, 65 benign) who had a standard digital mammography and an additional digital magnification view in the same projection within three months. All images were obtained using an amorphous silicon-based full-field digital system (Senographe 2000 D, GE Healthcare, Chalfont St. Giles, UK). Images were independently analyzed by six board-certified radiologists. The probability of malignancy was estimated using first standard contact mammography alone (MG) and then mammography in combination with the magnification view (MG + MAG) using a modified Breast Imaging Reporting and Data System (BI-RADS) classification system and a percentage scale. Results were compared using receiver operating characteristic (ROC) analysis. In addition, readers assessed the subjective visibility of the calcifications.

Results: For all six readers combined, the area under the curve (AUC) was 0.664 ± 0.052 for MG and 0.813 ± 0.042 for MG + MAG, resulting in a statistically significant improvement of 0.148 ± 0.120. Each reader had a higher AUC for MG + MAG than MG, with the improvement being statistically significant in four of the six readers. In 76.34 % of the cases, MG + MAG resulted in better visibility of calcifications compared with mammography alone. In 33 % slightly more and in 39 % significantly more calcifications were found.

Conclusion: Even in digital mammography with the option of using electronic magnification (zoom) at the viewing workstation, true geometric (air-gap) magnification views remain impor-
Introduction

Microcalcifications are, along with the mass-like appearance of breast lesions, the most important sign of malignancy and are the key finding to detect in situ and early invasive cancers in screening mammography [1]. Although some microcalcifications may also be seen with modern high-resolution ultrasound [2], many, especially smaller, areas of tiny microcalcifications are only detected by mammography. Despite the nominally lower spatial resolution of digital mammography compared to film-screen mammography, it has been shown that digital mammography screening improves the detection of ductal carcinoma in situ (DCIS) associated with microcalcifications [3–7]. Whereas the detection of microcalcifications on mammography is usually straightforward and may be assisted by employing highly sensitive systems for computer-aided diagnosis (CAD), distinguishing between benign and malignant microcalcifications may be very difficult and, in some situations, impossible without obtaining a biopsy for histological workup [8]. Of microcalcification cases sent to biopsy, usually only 30–50% of cases represent cancer [9]. Several criteria such as number of individual calcifications in a cluster, spatial distribution, morphology of individual calcifications, and change over time can help in estimating the probability of malignancy for a given microcalcification cluster. In film-screen mammography, air-gap magnification views have traditionally been used to improve the characterization of microcalcifications by better visualizing the morphology of the individual calcifications as well as by showing additional smaller calcifications [10–14].

According to Fischer et al., air-gap magnification views may no longer be necessary in digital mammography and can be replaced by digitally zooming the images [15]. However, their initial study was limited by a small number of cases and the fact that images were read on hard-copy, which does not allow one to exploit the full potential of digital images through interactive manipulation on a reading workstation. The purpose of our study therefore was to test whether additional air-gap magnification views may improve the characterization of microcalcifications in digital mammography using soft-copy reading and a larger number of cases.

Materials and Method

The institutional ethics committee approved this retrospective study. The inclusion criteria of the study were defined as follows:

- Complete diagnostic workup in our university hospital
- Digital standard mammography (MG) and digital magnification views in the same plane and on the same system (GE Senographe 2000 D)
- Microcalcifications as dominant suspicious findings
- Magnification views within three months of MG

Study Population

From January 2000 through December 2007, 42,147 patients underwent mammography at our department for the workup of breast symptoms, follow-up after breast cancer, or early detection of breast cancer. A total of 14,914 of these patients had a digital mammography. 741 patients in whom two-view digital mammograms demonstrated suspicious microcalcifications underwent additional digital magnification views no more than three months after the initial examination and on the same digital mammography machine and in the same plane as the initial mammography.

35 of these 741 cases were diagnosed with breast cancer (in situ or invasive) and fulfilled the inclusion criteria for our study (complete diagnostic workup including biopsy as well as final surgery performed at our hospital). To create a realistic sample covering the full spectrum of abnormalities, these 35 malignant cases were supplemented by 65 benign cases, which were randomly selected from the original dataset of 741 cases with a valid combination of digital mammography and digital magnification views. Roughly half of the benign cases went to biopsy after the magnification and were confirmed to be benign by histology (n = 33; 25 vacuum-assisted biopsies, 8 open surgical biopsies). The other half of the benign cases, which after full assessment did not go to biopsy, were confirmed to be benign by a follow-up of at least 24 months (n = 32).

The malignant cases were categorized by the World Health Organization (WHO) classification of breast cancers. Based on the classification published by Elston and Ellis, tumors were graded as well (grade 1), moderately (grade 2), or poorly (grade 3) differentiated [16].

Image Acquisition

All mammographies were acquired using an amorphous silicone-based full-field digital system with a 19 × 24 cm detector and 100 µm pixel size (Senographe 2000 D, GE Healthcare, Chalfont St. Giles, UK).

The magnification views were obtained with a magnification factor of 1.75 with a standard round spot compression paddle with a diameter of 7 cm. The resulting effective pixel size of the magnification views was 0.057 µm.

Reader Study

All images were analyzed using a dedicated digital mammography workstation (MammaReport, Siemens Medical Solutions, Erlangen, Germany) with two 5 megapixel thin film transistor monitors (TFT monitors) and with all viewing features including windowing and electronic zoom available to the readers. Six radiologists specialized in breast imaging with an average experience in digital mammography of 8.6 years (SD: 6.2 years; range: 1.5–19 years) participated in the reader study. The cases were collected and prepared by an additional experienced breast radiologist (U.B.). All six readers independently evaluated all 100 cases. Benign and malignant cases were read in random order. All
images were anonymized prior to the reading, and readers were blinded to patient age, clinical history, and histopathological outcome. During the reader study, only the single contact mammography view corresponding to the plane of the magnification view and no prior films were available to the readers.

If more than one microcalcification area was present in the image, the reader was asked to base the assessment on the most suspicious microcalcification cluster in the image.

All readers first rated the digital contact mammography alone (MG) and then in combination with the digital magnification view (MG+MAG) using a modified 7-point BI-RADS® scale (1, 2, 3, 4a, 4b, 4c, 5) as well as a probability of malignancy scale from 0 % to 100 % [17].

In addition, they were asked to assess the presence of additional subtle microcalcifications on the magnification views on a three-point scale (none, slightly more without diagnostic relevance, and considerably more with diagnostic relevance) and to compare the subjective visibility of the microcalcifications (MG much better with diagnostic relevance, MG better without diagnostic relevance, no difference, MAG better without diagnostic relevance, MAG much better with diagnostic relevance).

Statistical Analysis
A receiver operating characteristic (ROC) analysis was performed using the LABMRMC program provided by the University of Chicago [18].

The area under the curve (AUC) was calculated for each reader individually and for all readers combined using the seven-point modified BI-RADS scale and the probability of malignancy with histology or two-year follow-up results as the reference standard. The sensitivity and specificity were calculated using a four-fold table. A threshold of a ranking BI-RADS 3 and BI-RADS 4a was used for the BI-RADS scale. This statistical analysis was performed using Microsoft Excel 2007 for Windows.

Power calculation estimated a minimum of six readers for the 100 cases to reach significance [19].

Results
The mean age of all study patients was 57.3 years (SD, 8.9 years) with a range of 37 to 76 years.

The patients with malignant lesions had a mean age of 58.37 years (SD, 9.27 years; range, 40 – 76 years). The subjects with histologically verified benign lesions had a mean age of 58.7 years (SD, 9.3 years; range, 42 – 76 years); the mean age of patients with benign lesions confirmed by follow-up mammographic examinations was 55.3 years (SD 7.9 years; range, 37 – 72 years).

There were 12 invasive carcinomas and 23 DCIS. The histology of the malignant lesions was as follows: 8 invasive ductal carcinomas (3 of grade 2 and 2 of grade 1, 3 not further specified), 1 invasive mucinous carcinoma (grade 2), 1 invasive lobular carcinoma (grade 1), and 2 invasive carcinomas not otherwise specified (NOS).

The preinvasive lesions included 11 high-grade, 6 intermediate-grade, and 6 low-grade DCIS.

The histology of the benign lesions was nodular fibrosis in 11 cases, fibrocystic mastopathy in 16 cases, one radial scar, one dysplastic hyaline scar, papilloma in 3 cases, and fatty tissue in one case.

The distribution of mammographic breast density categories (ACR) is presented in Table 1.

There was no significant difference in the distribution of ACR breast density categories between patients with benign breast lesions and those with breast malignancy (p = 0.262).

Table 1

<table>
<thead>
<tr>
<th>ACR1 [%]</th>
<th>ACR2 [%]</th>
<th>ACR3 [%]</th>
<th>ACR4 [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>benign</td>
<td>1.53</td>
<td>36.92</td>
<td>49.23</td>
</tr>
<tr>
<td>malignant</td>
<td>1.0</td>
<td>51.42</td>
<td>45.23</td>
</tr>
<tr>
<td>all</td>
<td>1.1</td>
<td>42</td>
<td>48</td>
</tr>
</tbody>
</table>

Fig. 1 Columns demonstrating the visibility of microcalcifications in the magnification views compared to the zoomed images ("much better" is defined as diagnostically relevant, "better" does not reach diagnostic relevance). Summation does not exactly reach 100 % due to rounding.

Abb. 1 Das Säulendiagramm stellt die Sichtbarkeit der Mikroverkalkungen in den Vergrößerungsaufnahmen im Vergleich zu den gezoomten Aufnahmen dar. „Much better“ ist als diagnostisch relevant definiert; „better“ ist nicht diagnostisch relevant. In der Summe werden aufgrund von Rundung nicht exact 100 % erreicht.

Additional microcalcifications were detected on the magnification views in more than two thirds of cases (see Fig. 1). Furthermore, additional microcalcifications were detected on the magnification views in more than two thirds of cases (see Fig. 2).

Regarding the modified BI-RADS scale results of all six readers, the AUC was 0.6643 (SD = 0.0518) for MG and 0.8127 for MG + MAG (SD = 0.0416) with a 95 % confidence interval for the population mean difference of 0.1904 to 0.2681. Individually, each of the six observers showed a higher diagnostic accuracy (AUC) for the combination MG + MAG than MG, with the improvement being statistically significant in four of the six readers (see Table 2, Fig. 3) and a tendency in favor of the magnification views in one reader.

Regarding the modified probability of malignancy scale, the AUC was 0.6916 (SD = 0.0513) for MG and 0.7885 (SD = 0.0473) for MG + MAG with a 95 % confidence interval for the population mean difference of -0.0285 to -0.0285. Individually, each of the six readers revealed a higher diagnostic accuracy (AUC) for MG + MAG than MG, with the improvement being statistically significant in three readers (see Table 2, Fig. 4).

Analyzing the influence of the experience of the readers, the time of experience in mammography did not influence the results, as there have been “significant” and “not significant” results in readers with comparable times of experience (see Table 2).
If the observers are divided into three experienced observers and three less experienced observers, there are two "significant" results and one "not significant" result regarding BI-RADS scores in each group.

The distribution of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for the cut-offs of BI-RADS 3 and BI-RADS 4a is presented in Table 2.

There was an improvement of sensitivity independently of the breast density from 69% in MG to 73.8% in ACR 1 and 2 dense breast as well as from 72.2% to 77.8% in ACR 3 or 4 dense breasts.

Discussion

Our data clearly confirm that digital air-gap magnification views improve the visibility of microcalcifications, in terms of both number of calcifications and morphology of the individual calcifications. The physical basis for this is mainly the higher spatial resolution of the magnification views. However, other factors such as improved local compression due to the smaller paddle...
may also play a role. What is more important, however, is how much this improved visibility of microcalcifications in magnification views will turn out to be beneficial to patients (see Fig. 5).

Patients would benefit, for example, if a more confident characterization of microcalcifications could spare patients unnecessary biopsies or help for instance to avoid delays in cancer diagnosis.

In 2002, Fischer et al. did not find a significant benefit of additional magnification views regarding the overall diagnostic accuracy of digital mammography. Four readers (two with more than 15 years, one with one and one with three years experience) evaluated 55 patients with 57 microcalcification-associated lesions (21 malignant, 36 benign) using the BI-RADS scale. They found overall better results for the magnification views, but no statistically significant differences and therefore concluded that most problems can be solved with digital zooming and that omission of additional magnification views has the potential to improve workflow, to lower radiation exposure, and to reduce the recall rate in a digital mammography-based screening program. [15] However, the major limitation of this study is the use of hard copy reading alone.

Recently, Kim et al. found in an analysis of 120 clusters with three readers (1, 2 and 10 years of experience, 6.3 years average) that magnification mammograms were superior to contact mammography when using a zoom factor of 1.8 [20]. Another recent study published by Kim et al. compared a magnification factor of 1.8 and a zoom factor of 2.0 for the assessment of microcalcifications. They read 185 clusters with three radiologists (4, 5 and 12 years experience, 7 years average). As a major result, the authors did not find any significant differences between zooming and geometrical magnification regarding sensitivity, specificity, and AUC analysis of a probability of malignancy score that was comparable to a modified BI-RADS scale [21].

The most recent study was published by Moraux-Wallyn et al., who found geometrical magnification to be better than electronic zooming in terms of sensitivity, specificity, PPV, and NPV, but the difference was not statistically significant. This study included 88 clusters of microcalcifications analyzed by two readers [22]. These results are inline with ours showing an improvement of sensitivity for all readers in magnification views.

In our study, we used the maximum geometrical magnification of 1.8 and a comparable zooming factor of 2.0 to identify a possible benefit of true magnification views using soft-copy reading or to show if there is an influence due to the possibility to use digital zooming.

With these parameters, the diagnostic performance of geometrical magnification was found to be superior to zooming. Furthermore, these results were statistically significant. $A_s$ values of

<table>
<thead>
<tr>
<th>reader</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>cutoff $\geq$ BI-RADS 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sensitivity Mx</td>
<td>100.0 %</td>
<td>94.3 %</td>
<td>65.7 %</td>
<td>71.4 %</td>
<td>88.6 %</td>
<td>71.4 %</td>
</tr>
<tr>
<td>specificity Mx</td>
<td>1.5 %</td>
<td>30.8 %</td>
<td>49.2 %</td>
<td>50.8 %</td>
<td>23.1 %</td>
<td>50.8 %</td>
</tr>
<tr>
<td>PPV</td>
<td>0.354</td>
<td>0.423</td>
<td>0.411</td>
<td>0.439</td>
<td>0.383</td>
<td>0.439</td>
</tr>
<tr>
<td>NPV</td>
<td>1.000</td>
<td>0.909</td>
<td>0.727</td>
<td>0.767</td>
<td>0.789</td>
<td>0.767</td>
</tr>
<tr>
<td>cutoff $\geq$ BI-RADS 4a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sensitivity Mx</td>
<td>68.6 %</td>
<td>71.4 %</td>
<td>54.3 %</td>
<td>51.4 %</td>
<td>51.4 %</td>
<td>57.1 %</td>
</tr>
<tr>
<td>specificity Mx</td>
<td>40.0 %</td>
<td>60.0 %</td>
<td>69.2 %</td>
<td>80.0 %</td>
<td>44.6 %</td>
<td>53.8 %</td>
</tr>
<tr>
<td>PPV</td>
<td>0.381</td>
<td>0.490</td>
<td>0.487</td>
<td>0.581</td>
<td>0.333</td>
<td>0.400</td>
</tr>
<tr>
<td>NPV</td>
<td>0.703</td>
<td>0.796</td>
<td>0.738</td>
<td>0.754</td>
<td>0.630</td>
<td>0.700</td>
</tr>
<tr>
<td>sensitivity Mag</td>
<td>80.0 %</td>
<td>100.0 %</td>
<td>74.3 %</td>
<td>71.4 %</td>
<td>97.1 %</td>
<td>85.7 %</td>
</tr>
<tr>
<td>specificity Mag</td>
<td>29.2 %</td>
<td>47.7 %</td>
<td>73.8 %</td>
<td>76.9 %</td>
<td>61.5 %</td>
<td>53.8 %</td>
</tr>
<tr>
<td>PPV</td>
<td>0.378</td>
<td>0.507</td>
<td>0.605</td>
<td>0.625</td>
<td>0.576</td>
<td>0.500</td>
</tr>
<tr>
<td>NPV</td>
<td>0.731</td>
<td>1.000</td>
<td>0.842</td>
<td>0.833</td>
<td>0.976</td>
<td>0.875</td>
</tr>
</tbody>
</table>
MAG were at a higher level in all readers regarding BIRADS. Ours is the first study showing a statistically significant difference for all readers combined as well as for four individual readers with experience of 1.5, 4, 8 and 19 years regarding a BI-RADS scale. A possible explanation for the lack of statistical significance in earlier studies is the smaller number of readers, resulting in a lower statistical power for the ROC analysis. For two of the readers in our study, one with 8 years of experience and one with 5 years of experience, the difference was not statistically significant as well, but there was a tendency in favor of true magnification in the last one. The statistical power analysis for our data collection of 100 cases estimated a minimum number of readers necessary to reach statistical significance to be 6. None of the previous studies reached such a combined number of cases or readers. Another major benefit of our study was the huge range of experience among the readers (1.5 years to 19 years) resulting in significant and not significant outcomes in experienced and less experienced examiners. Regarding the distribution of reader experience in previously published studies, they contain more inexperienced readers with less than 5 years of experience than we included in our study. Kimme-Smith et al. [23] could show that experience in interpreting mammography positively influences the performance in judging microcalcifications in digitized enhanced and truly magnified images in favor of the more experienced readers. This could lead to the assumption that the experience time influenced the results in our study as well. If only readers with minimal experience would benefit from an additional magnification view and the resulting additional radiation dose would only equalize a lack of experience, it would be unacceptable in a diagnostic and screening mammography setting. Especially in the screening of potentially healthy women, the dose has to be as low as possible to avoid any harm for the patient particularly if digital mammography offers the opportunity to reduce the radiation dose [24, 25]. Nevertheless, a lack of experience was not the explanation for the difference in significance, as the "not significant" results were achieved by quite experienced readers with 5 and 8 years of experience. Even if we divide our readers into two groups, experienced (11.7 years average) and less experienced (3.5 years average), two of three results were significant in each group. There is a major difference between our study and earlier reports. We included histologically proven benign and malignant lesions as well as benign cases verified by follow-up mammograms. Thus, we covered the whole spectrum of microcalcifications that may be encountered in the clinical setting. The only other studies assessing the same range of microcalcification types included a much smaller number of cases than our study [15, 20]. All other studies only assessed microcalcifications classified as suspicious enough to warrant biopsy. But how did magnification views improve diagnostic performance in our study? The resulting decrease in pixel size improved the delineation of the shape and contour of calcifications. This effect has been shown by Ruschin et al. [26]. More very tiny calcifications were visible, which has an impact on the assessment of the extent of the affected region and therefore allows individual surgery planning to achieve free margins resulting in the fewest possible surgeries. [26] Another limitation on making things visible is the capability of the human eye. It is known from several studies that calcifications as small as 130 µm are visible to the human eye [27, 28]. With the decrease in pixel size in magnifications and the less attenuated dose (magnification views are obtained without a grid), there is much more information/photons per pixel in more pixels [29]. This results in a higher effective spatial resolution, which may improve visibility by sharpening object delineation and facilitating the differentiation of calcifications from overlying structures. Spot views with focal compression reduce overlying structures. With the use of a small spot compression plate and magnification, we reduced the overlying tissue even more than with magnification views alone. It has been shown that overlying tissue/structures significantly reduce the detection rate of microcalcifications [29]. The magnification and reduction of overlying tissue by compression may jointly have contributed to the better diagnostic performance in our study.

Limitations

One limitation of this study is that we only tested one detector. Hence, our results may not be transferable to other digital mammography systems using different pixel sizes, focal spot sizes, image noise, and magnification factors. Another limitation of the study is the fact that the readers knew whether they were judging a normal mammography or a magnification. Finally, magnification views are obtained with a higher dose than normal mammograms. Hence, it is not clear if the higher radiation dose contributed to the improved visibility.

Conclusion

Even in digital mammography with the option of using electronic magnification (zoom) at the viewing workstation, true geometric (air-gap) magnification views remain important for visibility and correct classification and the assessment of the extent of microcalcifications.

Affiliations

1 Clinic of Radiology, Campus Virchow-Klinikum, Charité – Universitätsmedizin, Berlin
2 Institute of diagnostic and interventional Radiology, DRK-Kliniken Berlin-Mitte, Berlin
3 Department of Medical Imaging, St. Joseph-Stift, Bremen
4 Neuroimaging for Research, Berlin
5 Reference-Centre of Mammography-Screening, Berlin
6 Department of Radiology, Campus Charité Mitte, Charité – Universitätsmedizin, Berlin
7 Breast Center, Department of Gynecology, Campus Charité Mitte, Charité – Universitätsmedizin, Berlin

Acknowledgments

"This manuscript is dedicated to Professor Bernd Hamm for his 60th birthday." Without his support and continuous affords in providing a high end equipped Institution research would not be possible. We are thankful to Professor Marc Dewey for discussions about this paper.
The authors thank Bettina Herwig for editorial support in preparing the manuscript.

References
8. Rominger M, Wisicki C, Timmesfeld N. Breast microcalcifications as type descriptors to stratify risk of malignancy: a systematic review and meta-analysis of 10665 cases with special focus on round/punctate microcalcifications. Fortschr Röntgenstr 2012; 184: 1144–1152