BREAST

Carla Meeuwis
Stephanie M. van de Ven
Gerard Stapper
Arancha M. Fernandez Gallardo
Maurice A. A. J. van den Bosch
Willem P. Th. M. Mali
Wouter B. Veldhuis

Computer-aided detection (CAD) for breast MRI: evaluation of efficacy at 3.0 T

C. Meeuwis
Department of Radiology,
Alysis Zorggroep, Rijnstate Hospital,
Wagnerlaan 55,
6815 AD Arnhem, The Netherlands

S. M. van de Ven
Department of Radiology,
Stanford University,
1201 Welch Road,
Stanford, CA, 94305, USA

G. Stapper · A. M. Fernandez Gallardo ·
M. A. A. J. van den Bosch ·
W. P. T. M. Mali · W. B. Veldhuis
Department of Radiology,
University Medical Center Utrecht,
Heidelberglaan 100, E.01.132,
3584 CX Utrecht, The Netherlands

C. Meeuwis (✉)
Department of Radiology,
Alysis Zorggroep, Rijnstate Hospital,
Postbus 9555, Intern postnr. 1530,
6800 TA Arnhem, The Netherlands
e-mail: CMeeuwis@alysis.nl
Tel.: +31-26-3788888
Fax: +31-26-3787381

Abstract *Objective:* The purpose of the study was to evaluate the accuracy of 3.0-T breast MRI interpretation using manual and fully automated kinetic analyses. *Material and methods:* Manual MRI interpretation was done on an Advantage Workstation. Retrospectively, all examinations were processed with a computer-aided detection (CAD) system. CAD data sets were interpreted by two experienced breast radiologists and two residents. For each lesion automated analysis of enhancement kinetics was evaluated at 50% and 100% thresholds. Forty-nine malignant and 22 benign lesions were evaluated. *Results:* Using threshold enhancement alone, the sensitivity and specificity of CAD were 97.9% and 86.4%, respectively, for the 50% threshold, and 97.9% and 90%, respectively, for the 100% threshold. Manual interpretation by two breast radiologists showed a sensitivity of 84.6% and a specificity of 68.8%. For the same two radiologists the mean sensitivity and specificity for CAD-based interpretation was 90.4% (not significant) and 81.3% (significant at $p < 0.05$), respectively. With one-way ANOVA no significant differences were found between the two breast radiologists and the two residents together, or between any two readers separately. *Conclusion:* CAD-based analysis improved the specificity compared with manual analysis of enhancement. Automated analysis at 50% and 100% thresholds showed a high sensitivity and specificity for readers with varying levels of experience.

Keywords Breast · Magnetic resonance imaging · Diagnosis · Computer-assisted detection · High field · Neoplasm

Abbreviations MRI: magnetic resonance imaging · BI-RADS: Breast Imaging Reporting and Data System · DCIS: ductal carcinoma in situ · CAD: computer-aided detection

Introduction

The use of MRI for the detection and evaluation of breast lesions continues to increase. With a sensitivity of around 90% it is a more sensitive method than conventional mammography or ultrasound for the detection of breast cancer [1]. A recent meta-analysis showed the pooled-weighted specificity to be 72%, with a 95% confidence interval of 67–77% [1]. An added drawback of breast MRI is the significant time needed for image processing and interpretation. Also, the experience level and number of radiologists trained in breast MRI is not yet as high as for conventional mammography. Computer-aided detection (CAD) algorithms have been developed that automate processing and analysis functions usually performed manually by MRI technologists and radiologists. These

systems have the potential to improve efficiency of breast MRI and to reduce the number of false-positive diagnoses [2]. Automation may improve consistency and detection rate, but also provides new analysis methods, such as kinetic curve threshold, that are not available with manual interpretation.

Compared with 1.5 T, current MRI systems operating at 3.0 T offer a higher signal-to-noise ratio and allow higher spatial resolution imaging without affecting overall image acquisition time or the temporal resolution of the dynamic contrast-enhanced series [3].

This study was performed to compare the consistency and accuracy of 3.0-T breast MRI interpretation, between manual and fully automated kinetic analyses.

Materials and methods

Retrospective inclusion of patients and lesions

We retrospectively analysed the data from a total of 426 consecutive women who underwent contrast-enhanced high-spatial-resolution 3.0-T breast MRI at our hospital, between May 2005 and December 2006. We chose this time period because at that time our hospital replaced the manual kinetic analysis system with a fully automated computer-aided detection (CAD)-based kinetic analysis. In this way, the data from all patients included in this study were analysed on both systems. No approval by the ethics board or informed consent was needed because breast MRI was not performed in a study setting, but in the clinical setting.

Women who underwent breast MRI not for clinical indications but for research purposes were excluded from this study ($n=71$). In addition, 286 patients were excluded because histology was not obtained. Of the 69 patients in whom histology was obtained, 4 patients with 5 lesions were excluded for technical reasons: severe patient motion ($n=2$), failed fat suppression ($n=1$) and a technical error with contrast material injection ($n=2$), which caused an incomplete study. Finally 65 patients were included with a total of 71 breast lesions proven surgically or by core biopsy. The evaluation of the accuracy of CAD threshold enhancement was based on these 65 patients. The evaluation of the diagnostic accuracy of the four different readers was based on the same data set, with the exclusion of all 29 BI-RADS category 6 known cancers. Exclusion of known cancers was necessary because readers were blinded to the pathological results but not to patient history (Fig. 1).

Breast MRI: technique

MRI was performed on a 3.0-T MR system (Achieva, Philips Medical Systems, Best, the Netherlands). Patients were placed in the prone position on a dedicated phased-array bilateral breast coil (MRI devices, Würzburg, Germany). Transverse high-resolution T1-weighted fast field echo (T1-FFE) images were obtained (TE/TR 1.7/4.5 ms; inversion delay SPAIR 130 ms; flip angle 10°; FOV 340×340 mm^2, acquired voxel size $0.66 \times 0.66 \times 1.6$ mm^3, reconstructed voxel size $0.66 \times 0.66 \times 0.80$ mm^3), followed by transverse T2-weighted fat-suppressed images (TE/TR 120/9,022 ms; inversion delay SPAIR 125 ms; flip

Fig. 1 Flow chart of patient inclusion, with reasons for exclusion and total number of patients for each analysis method
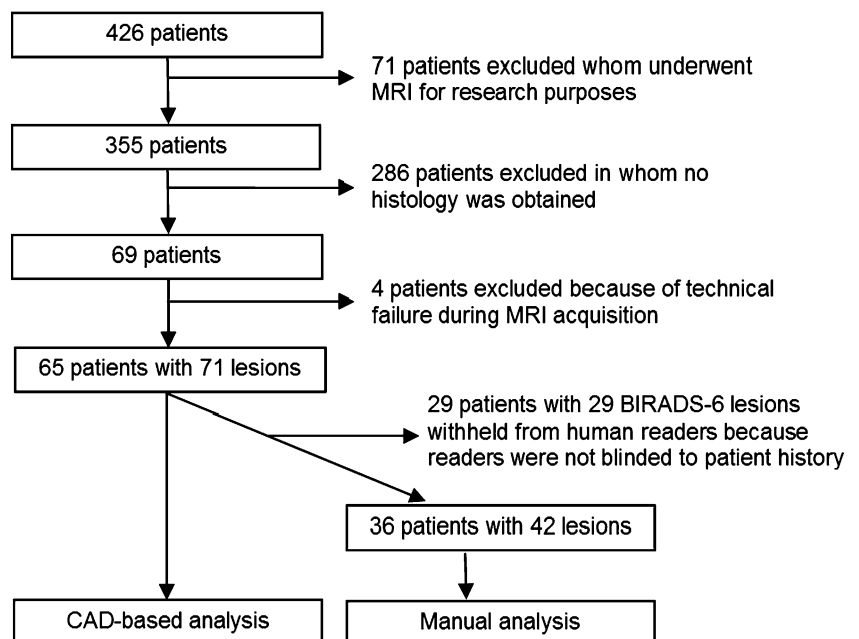
angle 90°; FOV 340×340 mm^2, acquired voxel size 1.01 × 1.31 × 2.0 mm^3, reconstructed voxel size 0.66 × 0.66 × 2.00 mm^3). Dynamic fat-suppressed T1-weighted gradient echo images were acquired before and immediately after intravenous administration of 0.1 mmol/kg gadopentate (Magnevist, Schering, Germany) (TE/TR 1.3/3.4 ms; flip angle 10°; FOV 320 × 320 mm^2, acquired voxel size 0.91 × 0.91 × 2.00 mm^3, reconstructed voxel size 0.83 × 0.83 × 1.00 mm^3; temporal resolution of 60 s per dynamic acquisition, with a total of 6 dynamic acquisitions; finally, a post-gadolinium T1-weighted gradient echo series (TE/TR 1.7/4.5 ms; inversion delay SPAIR 130 ms; flip angle 10°; FOV 340×340 mm^2, acquired voxel size 0.66 × 0.66 × 1.60 mm^3, reconstructed voxel size 0.66 × 0.66 × 0.80 mm^3).

Breast MRI interpretation and data collection

Breast MR interpretation using manual kinetics analysis was done on an Advantage Workstation (Philips Viewforum, Best, the Netherlands), which allows assessment of enhancement kinetics by manual region-of-interest (ROI) placement. Studies were read by one of two experienced breast radiologists who were blinded to the pathological results. For each MRI-detected lesion, a separate American College of Radiology (ACR) BI-RADS-MRI Lexicon Classification Form [4] was filled out. In addition, for each lesion the maximum size in two orthogonal directions was measured.

All MRI examinations were subsequently processed by CADstream (Confirma, Inc., Kirkland, WA), a commercially available computer-aided detection system. With CADstream, areas of enhancement that meet a user-defined minimum threshold for initial enhancement are automatically identified by colour overlays on all MRI slices. This initial enhancement threshold refers to the minimum increase in signal intensity on the early post-contrast-images over the pre-contrast images, within the same dynamic series. In addition to indicating the degree of initial enhancement, the colour overlay allows differentiation between persistent-, plateau- and washout-type enhancement in the late phase after contrast injection.

CADstream data sets were interpreted by the same two breast radiologists. In addition the data sets were interpreted by two residents. The two breast radiologists are both registered breast radiologists and have more then 15 years' experience in breast ultrasound and conventional mammography. Both radiologist also have more then 5 years' experience in breast MRI. The first residents was in his third year and had 6 months' experience in conventional mammography, breast ultrasound and breast MRI, by the time this study took place in our hospital. The second resident was in her first year and only had 2 months of experience in conventional mammography and breast ultrasound. She had only little experience in breast MRI, by following a 4-day course in breast MRI. All of the readers were blinded to the patient's history, treatment and pathological outcome, and to the results of the initial interpretation using manual enhancement kinetics. CAD reading was done 6 months after manual analysis of the same data set. The interval was felt to be long enough to prevent the readers from remembering the details of the initial manual kinetic analysis. All readers were familiar with the BI-RADS classification system. Separate BI-RADS-MRI score sheets were again filled out for all MRI-detected lesions, by all individual readers.

Lastly, lesions were scored based on CAD threshold analysis alone, first on the basis of a 50% threshold of initial enhancement and subsequently on the basis of a 100% threshold of initial enhancement.

Statistical analysis

Statistical analysis was done with SPSS 12.0. Significance was assumed at a level of $p<0.05$. The interobserver differences in the evaluation of quality criteria were evaluated by the paired Student's t test. One-way analysis of variance (ANOVA) was used to test for differences among the two experienced breast radiologists and the two residents together, or between any two readers separately with the use of CAD.

According to the American College of Radiology ACR guidelines, which suggest follow-up for "probably benign" lesions (BI-RADS category 3) and consideration of biopsy for "suspicious" lesions (BI-RADS category 4), sensitivity and specificity according to BI-RADS score were defined as follows:

$$\text{Sensitivity BI} - \text{RADS} = \frac{\text{Numbers of PA}-\text{proven malignant lesions with BI}-\text{RADS score 4 or 5}}{\text{Number of PA}-\text{proven malignant lesions}}$$

$$\text{Specificity BI} - \text{RADS} = \frac{\text{Numbers of PA}-\text{proven benign lesions with BI}-\text{RADS rating} \leq 3}{\text{Number of PA}-\text{proven benign lesions}}$$

where PA refers to pathology.

Table 1 Indication for MRI examination in 65 patients

	Number
High-risk screening	13 (20%)
Positive family history	7
History of high-risk lesion or cancer	6
Problem solving	30 (46.2%)
Palpable abnormality	6
Questionable lesion mammogram/US	23
Positive axillary node, unknown primary	1
Staging of known breast cancer	29 (44.6%)

Results

Patients and lesions

The mean age of the included patients was 49 years (range, 29–71 years). Indications for undergoing MRI (Table 1) were: evaluation of indeterminate mammographic and/or ultrasound findings (46%); staging of recently diagnosed breast cancer (45%); and high-risk screening (20%). Five patients had two indications and one patient had three indications.

Lesion characteristics

Lesions consisted of a focal (8.5%), mass (77.7%), or non-mass-like enhancement (13.8%). Size ranged from 4 to 49 mm. Histopathological evaluation after core-needle biopsy or surgical excision revealed 22 benign (including

Table 2 Histopathologic findings in 71 breast lesions

	Number
Malignant lesions	49 (69%)
Invasive ductal carcinoma	15
Infiltrating lobular carcinoma	7
Invasive ducto-lobular carcinoma	2
Pure ductal carcinoma in situ (DCIS)	7
Combination of DCIS and invasive carcinoma	14
Small cell carcinoma	4
High-risk lesions	5 (7%)
Lobular carcinoma in situ	5
Benign lesions	17 (23.9%)
Fibrocystic change	6
Fibroadenoma	9
Lymph node	1
Benign cyst	1
Total	71

Table 3 Sensitivity and specificity of 42 lesions (benign $n=16$, malignant $n=26$)

	Sensitivity	Specificity
Radiologist; manual	84.6% ($n=22$)	68.8% ($n=11$)
Radiologist 1; CAD	88.5% ($n=23$)	75.0% ($n=12$)
Radiologist 2; CAD	92.3% ($n=24$)	87.5% ($n=14$)
Residents 1; CAD	88.5% ($n=23$)	93.8% ($n=15$)
Resident 2; CAD	84.6% ($n=22$)	81.3% ($n=13$)

high-risk) lesions and 49 malignant lesions. Of the 22 benign lesions, 9 (13%) were fibroadenoma, 6 (8%) showed focal fibrocystic change, 1 (1%) was a benign complex cyst, 1 (1%) was a lymph node, and 5 (7%) were lobular carcinoma in situ (LCIS) (Table 2).

Of the 49 malignant lesions, 24 (34%) were invasive carcinoma (15 invasive ductal carcinoma, 7 invasive lobular carcinoma and 2 invasive ducto-lobular carcinoma), 7 (10%) were pure ductal carcinoma in situ (DCIS), 14 (20%) were a combination of in situ and infiltrating carcinoma, and 4 separate lesions in a single patient turned out to be metastases from small cell lung cancer (Table 2).

Table 3 summarises the sensitivity and specificity for interpretation with manual analysis of enhancement kinetics and for interpretation with CAD-based analysis enhancement kinetics. Sensitivity and specificity based on CAD threshold analysis alone are summarised in Table 4.
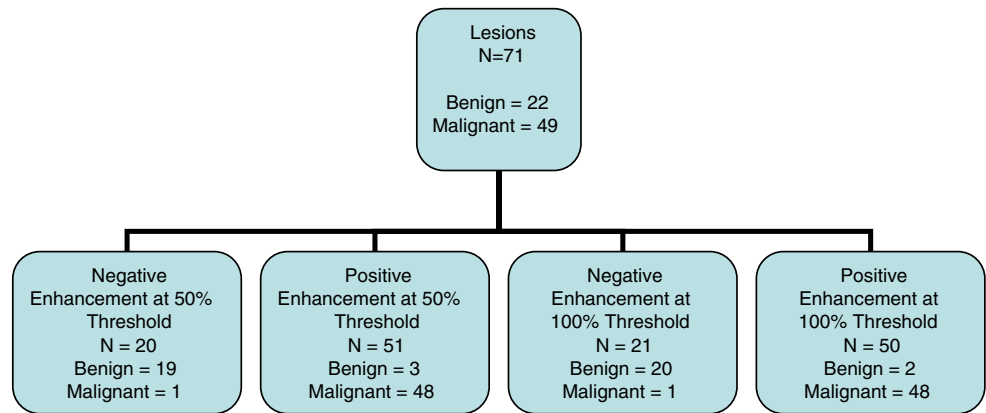
Manual interpretation showed a sensitivity of 84.6% and a specificity of 68.8%. For the same two radiologists the mean sensitivity and specificity for CAD-based interpretation was 90.4% and 81.3%, respectively. This difference in specificity was significant at $p<0.05$. Between all four CAD readers the sensitivity varied from 84.6% to 92.3% and the specificity varied from 75.0% to 81.3%. By one-way ANOVA no statistical significant difference was found in our study between the two experienced breast radiologists and the two residents together, or between any two readers separately with the use of CAD (Table 3).

Using threshold enhancement alone, the sensitivity and specificity of CAD were 97.9% and 86.4% for the 50% threshold, respectively, and 97.9% and 90% for the 100% threshold, respectively (Fig. 2 and Table 4).

Table 4 Sensitivity and specificity of 71 lesions at threshold enhancement at 50% and 100% (benign $n=22$, malignant $n=49$)

	Sensitivity	Specificity
CAD 50%	97.9%	86.4%
CAD 100%	97.9%	90.0%

```
                        Lesions
                         N=71

                     Benign = 22
                     Malignant = 49
```

| Negative Enhancement at 50% Threshold N = 20 Benign = 19 Malignant = 1 | Positive Enhancement at 50% Threshold N = 51 Benign = 3 Malignant = 48 | Negative Enhancement at 100% Threshold N = 21 Benign = 20 Malignant = 1 | Positive Enhancement at 100% Threshold N = 50 Benign = 2 Malignant = 48 |

Discussion

We found that interpretation with CAD-based analysis for enhancement kinetics significantly improved the discrimination of benign from malignant lesions compared with interpretation with manual analysis of enhancement kinetics. In our study initial enhancement was below 50% in 86% of benign lesions. These results are in keeping with the findings of Williams et al. [5]. Other authors also described an increase in specificity of using CAD that is based on excluding lesions with low threshold enhancement [6–11].

Difference in specificity between MR interpretation on an Advantage Workstation with manual assessment of enhancement kinetics and CAD may also partly be explained by the fact that CAD provides enhancement information for all pixels in a lesion rather than for a portion of a lesion measured by using manual region-of-interest placement [5].

Using threshold enhancement alone, the sensitivity of CAD is high for both the 50% threshold and 100%

threshold. There was only one malignant lesion that was false negative at CAD and did not demonstrate enhancement at the 50% and 100% thresholds. This lesion was described by three readers as a lobular-shaped mass with irregular margin and heterogeneous enhancement. One reader described this lesion as a round-shaped mass with regular margin and heterogeneous enhancement. Placing a manual ROI showed a malignant curve (Fig. 3). Three out of the four readers classified this lesion as a BI-RADS 5 lesion, highly suggestive of malignancy, and one reader as a BI-RADS 4 lesion, suspicious. Histopathological evaluation revealed an invasive ductal carcinoma. There are some technical limitations to the CAD program that can cause negative enhancement at malignant lesions which underscore the importance of using CAD as a complement to but not as a replacement for the radiologist's assessment [5].

Three benign lesions had positive enhancement at a 50% threshold and two of these lesions had positive enhancement at 100%. The lesion that had positive enhancement at 50% threshold and negative enhancement at 100% thresh-
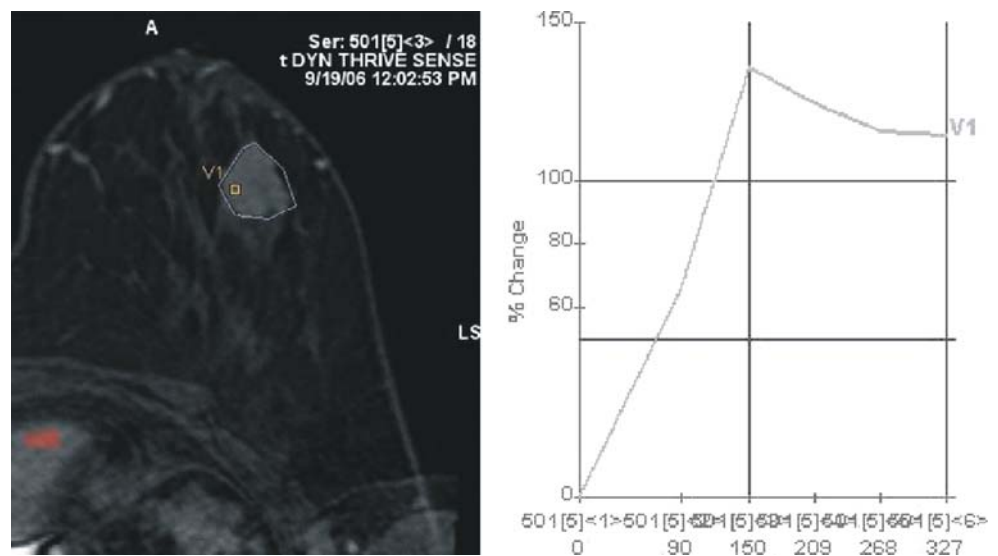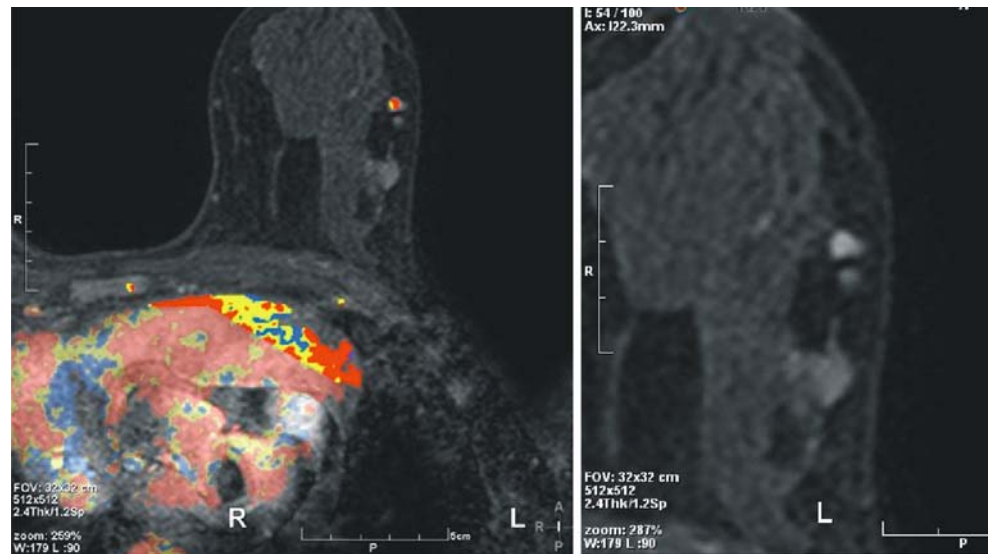
Fig. 4 Example of a false-positive finding. CAD analysis of this lesion (the lesion in *red* and *yellow* on the *left* image) indicated positive enhancement at the 50% and 100% thresholds. The image on the *right* is the same lesion in the same slice without the colour-coded overlay. Pathology revealed a benign fibroadenoma

old consisted of normal tissue at histopathological evaluation. Two readers classified this lesion as a BI-RADS 3 lesion, probably benign (one resident and one radiologist), and the other two readers classified this lesion as a BI-RADS 2 lesion, benign. The two false-positive lesions with positive enhancement at the 50% and 100% thresholds revealed one lobular carcinoma in situ and one fibroadenoma. Two readers classified the fibroadenoma (Fig. 4) as a BI-RADS 3 lesion, one as a BI-RADS 2 lesion and one as an BI-RADS 4 lesion. Lobular carcinoma in situ (Fig. 5), a high-risk lesion, was twice classified as a BI-RADS 3 lesion and twice as a BI-RADS 4 lesion.
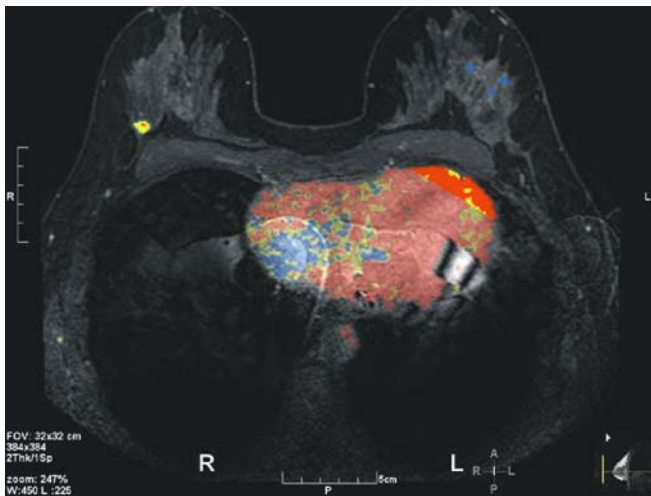

Fig. 5 Example of a false-positive finding. CAD analysis of this lesion (the *yellow* lesion in the right breast) indicated positive enhancement at the 50% and 100% thresholds. Pathology revealed a lobular carcinoma in situ, a high-risk lesion

In our study the readers were not blinded to the results of other imaging examinations such as mammography, ultrasound and previous MRI findings. All readers had access to patient history and clinical information, i.e. similar to the normal clinical setting. In the study by Kurz et al. who compared three different software systems in the evaluation of breast MRI, all three investigators were blinded to this information [10].

Our study had several limitations. Time needed to fully assess one MRI examination of the breast was not objectified as in the study of Kurz et al. [10]. They describe an average time needed to evaluate each exam of 7.0 min. Wiener et al. reported that most cases were interpreted with the use of an automatic post-processing program which generated colour parametric maps as well in less than 5 min [9]. Although not objectified, the two breast radiologists reported shorter interpretation times with CAD assistance than with manual analysis.

Another limitation is that manual interpretation was not repeated by the two residents. Meinel et al. showed that inexperienced readers without the use of CAD assistance performed as expected for their level of experience and that these differences became less pronounced with the use of CAD assistance [11]. In our study, no significant differences were found between the two experienced breast radiologists and the two residents together, or between any two readers separately with the use of CAD.

With respect to interpreting the very high sensitivities reported here for CAD-based analysis based solely on the thresholding of enhancement kinetics, it should be noted that a selection bias was introduced by only including data from patients with lesions proven by core or excision biopsy.

In summary, our findings suggest that CAD has the potential to improve the discrimination of benign from

malignant breast lesions at 3.0-T MRI. Additionally CAD may decrease the heterogeneity of interpretation across radiologists of varying levels of experience in breast MR interpretation.

References

1. Peters NH, Borel Rinkes IH, Zuithoff NP, Mali WP, Moons KG, Peeters PH (2008) Meta-analysis of MR imaging in the diagnosis of breast lesions. Radiology 246:116–124
2. Lehman CD, Peacock S, Demartini WB, Chen X (2006) A new automated software system to evaluate breast MR examinations: improved specificity without decreased sensitivity. AJR Am J Roentgenol 187:51–56
3. Kuhl CK, Jost P, Morakkabati N, Zivanovic O, Schild HH, Gieseke J (2006) Contrast-enhanced MR imaging of the breast at 3.0 and 1.5 T in the same patients: initial experience. Radiology 239:666–676
4. Ikeda DM, Hylton NM, Kinkel K et al (2001) Development, standardization, and testing of a lexicon for reporting contrast-enhanced breast magnetic resonance imaging studies. J Magn Reson Imaging 13:889–895
5. Williams TC, DeMartini WB, Partridge SC, Peacock S, Lehman CD (2007) Breast MR imaging: computer-aided evaluation program for discriminating benign from malignant lesions. Radiology 244:94–103
6. Leong CS, Daniel BL, Herfkens RJ et al (2000) Characterization of breast lesion morphology with delayed 3DSSMT: an adjunct to dynamic breast MRI. J Magn Reson Imaging 11:87–96
7. Deurloo EE, Muller SH, Peterse JL, Besnard AP, Gilhuijs KG (2005) Clinically and mammographically occult breast lesions on MR images: potential effect of computerized assessment on clinical reading. Radiology 234:693–701
8. Deurloo EE, Peterse JL, Rutgers EJ, Besnard AP, Muller SH, Gilhuijs KG (2005) Additional breast lesions in patients eligible for breast-conserving therapy by MRI: impact on preoperative management and potential benefit of computerised analysis. Eur J Cancer 41:1393–1401
9. Wiener JI, Schilling KJ, Adami C, Obuchowski NA (2005) Assessment of suspected breast cancer by MRI: a prospective clinical trial using a combined kinetic and morphologic analysis. Am J Roentgenol 184:878–886
10. Kurz KD, Steinhaus D, Klar V, Cohnen M, Wittsack HJ, Saleh A, Mödder U, Blondin D (2009) Assessment of three different software systems in the evaluation of dynamic MRI of the breast. Eur J Radiol 69(2):300–307
11. Meinel LA, Stolpen AH, Berbaum KS, Fajardo LL, Reinhardt JM (2007) Breast MRI lesion classification: improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. J Magn Reson Imaging 25:89–95