RESEARCH

Diagnostic value of combined use of abbreviated breast MRI and diffusion-weighted MRI in the prediction of breast cancer

Meme kanseri tahmininde kısaltılmış meme MRG ve diffüzyon ağırlıklı MRG’nin beraber kullanımının tanısal değeri

Şehnaz Tezcan1, Ülkü Bekar1, Funda Ulu Öztürk2
1Koru Hospital, Radiology Department, Ankara, Turkey
2Baskent University Hospital, Radiology Department, Ankara, Turkey

Abstract

**Purpose:** The aim of this study was to compare the diagnostic performances of abbreviated protocol (AP) magnetic resonance imaging (MRI), AP combined with diffusion-weighted imaging (DWI), and full dynamic protocol (FDP) in the differentiation of breast cancers from benign breast diseases.

**Materials and Methods:** The total study population consisted of 68 patients who underwent breast MRI (1.5 Tesla) between January 2016 and December 2021 for the evaluation of suspicious findings on mammography or ultrasonography. All lesions were evaluated by 2 radiologists using AP, AP+DWI, and FDP. The reader sensitivity, specificity, and accuracy were analyzed using the “Chi-squared” test. The inter-observer agreement (IOA) between the Breast Imaging Reporting and Data System (BI-RADS) category assessments of the two readers was evaluated using the “Kappa statistics”.

**Results:** Sixty-eight patients with 72 lesions (31 malignant and 41 benign) were analyzed. The sensitivity/specificity for AP and AP+DWI for reader 1 was 67.7/90.2% and 80.6/87.8%, respectively, and for reader 2 was 67.7/92.6% and 70.9/90.2%, respectively. The sensitivity/specificity for FDP for reader 1 was 83.7/85.3% and for reader 2 was 80.6/90.2%. The IOA in the BI-RADS category assessment was almost perfect in all models between two readers (the kappa value was 0.907, 0.825, and 0.858 in AP, AP+DWI, and FDP, respectively).

**Conclusion:** FDP showed greater diagnostic efficiency in the characterization of tumor biology as compared to AP and AP+DWI. Combining AP with DWI improved the diagnostic performance of MRI for the determination of malignancy.

**Keywords:** Diffusion-weighted imaging, abbreviated protocol breast MRI, breast cancer.

**Address for Correspondence:** Sehnaz Tezcan, Koru Hospital, Radiology Department, Ankara, Turkey
E-mail: sehnanze@yahoo.com
Received: 08.10.2022 Accepted: 30.01.2023
INTRODUCTION

Breast magnetic resonance imaging (MRI) has a high sensitivity for malignancy. MRI is a widely used method for screening women at high risk for breast cancer (BC) and staging, preoperative evaluation, and post-treatment follow-up of patients. The breast MRI acquisition time is approximately 35 to 40 minutes. The long acquisition time required for the full dynamic protocol (FDP) breast MRI increases costs and extends interpretation time, and it may also cause patient discomfort. Kuhl et al. developed an abbreviated protocol (AP) breast MRI consisting of precontrast T1-weighted sequence with fat saturation, single early postcontrast sequence, and maximum intensity projection (MIP) images. In the study by Kuhl et al., AP and FDP were found to have equivalent diagnostic accuracy. Its ability to reduce healthcare costs and interpretation time increased the popularity of AP. A recent meta-analysis showed that the diagnostic performances of the AP and FDP used in either screening or enriched cohorts were comparable. The studies included in that meta-analysis used the AP developed by Kuhl et al. or a combination of AP with T2-weighted (T2W) sequences. A study by Park et al. revealed an equivalent accuracy for AP with T2W and FDP in the MRI surveillance of patients with a history of BC. The sequences which have to be included in AP are still controversial. The majority of the studies in the relevant literature analyzed the performance of AP with or without T2W; however, there are only a few studies analyzing the impact of combining AP with diffusion-weighted imaging (DWI). DWI is a non-invasive and cheap method that evaluates the micro-structural attributes of water diffusion in tissues. DWI gives quantitative information about water diffusion by using apparent diffusion coefficient (ADC) maps which show the cellularity of the lesions. When tissue cellularity increases, the diffusion of water molecules reduces. Two previous meta-analyses revealed that a combined model of dynamic contrast-enhanced (DCE) MRI and DWI was more valuable in the differentiation of malignancy compared to DWI and DCE-MRI. The previous studies evaluating the impact of combining AP with DWI generally focused on the comparison of the diagnostic performances of DWI and contrast-enhanced sequences. However, using the combination of DWI and AP was reported in a few studies. A recent study demonstrated that combining AP with DWI provided a diagnostic performance similar to FDP in the determination of BC. Additionally, Chen et al. revealed that adding DWI to AP improved specificity in the diagnosis of BC.

Although DWI is an inexpensive and relatively fast imaging modality, there was no consensus in the literature regarding the integration of DWI into AP. This study aimed to compare the modalities of AP, AP with DWI, and FDP in terms of accuracy and determine whether the integration of DWI into AP could improve the diagnostic performance of breast MRI.

MATERIALS AND METHODS

Study population

Our retrospective study was approved by the Institutional Review Board of Dr. Abdurrahman Yurtaslan Ankara Oncology Training and Research Hospital (Decision number 2022-04/179 dated 20/04/2022). Informed consent was obtained from all participants. A total of 340 patients who underwent breast MRI (with a 1.5 Tesla scanner) between January 2016 and December 2021 at Koru Hospital, Ankara for the evaluation of suspicious findings on mammography or ultrasonography were included in this study. The research data including, MR images, histopathological results, and demographics were obtained from the electronic registration system of Koru Hospital. Out of 340 patients, 271 were excluded from the study due to the unavailability of histopathological results at our institution. One patient was excluded from the study due to the inadequate quality of DCE sequences. Finally, 68 patients with histopathological results at our hospital were included in the study. Tru-cut biopsy or excisional biopsy was used for the histopathologic diagnoses of all lesions. All specimens were evaluated by an experienced pathologist.

MRI Protocol

All MRI examinations were performed with a 1.5 T MRI scanner (GE Optima 360, USA) by using the breast coil while the patient was in the prone position. A routine protocol was performed including pre-contrast axial T1-weighted (TR/TE, 443/10.2; matrix, 320x224; NEX, 2; slice thickness, 5 mm) and
non-fat suppressed, T2-weighted fast spin-echo (TR/TE, 5800/68; matrix, 288x244; NEX, 2; slice thickness, 5 mm). Both before and after intravenous injection of the contrast material, 6 sequential fat-suppressed LAVA sequences were obtained, and subtraction was performed. DCE images were obtained after the administration of 0.1 mmol/kg of gadoteric acid. The scanning parameters for dynamic contrast-enhanced MRI were TR/TE, 3.35/1.6; matrix, 256x256; NEX, 1; slice thickness, 2 mm; flip angle, 10°; FOV, 3.6x3.2 cm; acquisition time (56 secs), respectively. Prior to the dynamic analysis, 2D spin-echo echo-planar images (EPI) were obtained with diffusion gradients in the x, y, and z planes at b values of 0 and 1000 s/mm2. The DWI sequences were obtained with the following parameters: TR/TE, 5500/78; matrix, 96x128; NEX, 2; slice thickness, 5 mm). The ADC maps and MIP images were created automatically.

**Image analysis**

All MRI examinations were evaluated by two experienced radiologists (11 and 6 years of experience) who were blinded to patient data. We established 3 models which were AP, AP+DWI, and FDP. While AP consisted of a pre-contrast and the first post-contrast T1W sequence, and MIP images, AP+DWI was the combination of AP with the axial DWI findings. FDP included all breast MRI sequences which were defined in the MRI protocol section. First, the readers indicated a category according to the American College of Radiology (ACR) Breast Imaging Reporting and Data System (BI-RADS) 5th edition MR imaging criteria for the AP of each patient. Two weeks later from the first interpretation, the readers determined the BI-RADS category for the model of AP+DWI. Two weeks later from the second interpretation, the BI-RADS category was assessed for the FDP of each patient. The BI-RADS classifications were separated into two groups as positive and negative to evaluate the diagnostic accuracy: BI-RADS 1, 2, and 3 were considered negative, and BI-RADS 4A, 4B, 4C, and 5 were considered positive. According to these findings, the sensitivity, specificity, and accuracy of 3 models were calculated for each reader using the histopathological diagnosis as the gold standard.

**Statistical analysis**

The Kolmogorov-Smirnov test was used to analyze the normality of the distribution. The reader sensitivity, specificity, and accuracy were analyzed using the “Chi-squared” test. Continuous variables were presented as mean ± standard deviation. Patients were compared in terms of differences in age by using the “Student-t” test. The inter-observer agreement (IOA) between readers was evaluated using Cohen’s weighted kappa statistics, considering categories according to Landis and Koch recommendations (Kappa (K) value; <0 Poor; 0.00-0.20 Slight; 0.21-0.40 Fair; 0.41-0.60 Moderate; 0.61-0.80 Substantial; 0.81-1.00 Almost perfect) with 95% confidence interval (CI). Analyses were performed with SPSS (version 22). The "p" value of less than 0.05 was considered statistically significant.

**RESULTS**

The total study population consisted of 68 patients with 72 lesions including 31 (43%) malignant lesions and 41 (56.9%) benign lesions. Four patients had bilateral breast lesions. The lesion was counted as one lesion for each side of the breast. In the malignant group, there were 27 patients with invasive carcinoma (grade 1 in 19 patients, grade 2 in 5 patients, and grade 3 in 3 patients) and 4 patients with ductal carcinoma in situ (DCIS). The mean age of the patients with malignant lesions was 54.1±10.7 years and that of the benign group was 49.9±12.3 years. The difference between groups in terms of mean age was not statistically significant (p: 0.14).

The distributions of BI-RADS category assessments for each reader are demonstrated in Figures 1 and 2. In benign lesions, the rates of BI-RADS 1 to 3 category assignment of reader 1 were 90.2%, 87.8%, and 85.3% by AP, AP+DWI, and FDP respectively. Reader 2 assigned BI-RADS 1 to 3 category in 92.6%, 90.2%, and 90.2% of patients with benign pathology by AP, AP+DWI, and FDP respectively. For the malignant group, the rates of BI-RADS 4 or 5 category assignment were 67.7%, 80.6 and 83.7 for AP, AP+DWI, and FDP respectively by reader 1. Reader 2 assigned BI-RADS 1 to 3 category in 92.6%, 90.2%, and 90.2% of patients with benign pathology by AP, AP+DWI, and FDP respectively. For the malignant group, the rates of BI-RADS 4 or 5 category assignment were 67.7%, 80.6 and 83.7 for AP, AP+DWI, and FDP respectively by reader 1. Reader 2 assigned BI-RADS 4 or 5 category in 67.7%, 70.9, and 80.6% of malignant patients by AP, AP+DWI, and FDP respectively.

The sensitivity, specificity, and accuracy of the three models for each reader are demonstrated in Table 1. For both readers, the sensitivity of AP was 67.7% which was lower than that of AP+DWI (whose sensitivity was 80.6% and 70.9% for reader 1 and reader 2, respectively) and than FDP sensitivity (which was 83.7% and 80.6% for reader 1 and 2,
respectively, Table 1). When AP was combined with DWI, higher sensitivity and lower specificity were observed compared to AP alone (Table 1).

When AP was combined with DWI, higher sensitivity and lower specificity were observed compared to AP alone (Table 1).

Table 1. Diagnostic performance of established models

<table>
<thead>
<tr>
<th></th>
<th>AP</th>
<th>AP+DWI</th>
<th>FDP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reader 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>67.7</td>
<td>&lt;0.001</td>
<td>60.6</td>
</tr>
<tr>
<td>(48.6-83.3)</td>
<td></td>
<td>(62.5-92.5)</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>90.2</td>
<td>&lt;0.001</td>
<td>87.8</td>
</tr>
<tr>
<td>(76.8-97.2)</td>
<td></td>
<td>(73.8-95.9)</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>80.5</td>
<td>&lt;0.001</td>
<td>84.7</td>
</tr>
<tr>
<td>(69.5-88.9)</td>
<td></td>
<td>(74.3-92.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Reader 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>67.7</td>
<td>&lt;0.001</td>
<td>70.9</td>
</tr>
<tr>
<td>(48.6-83.3)</td>
<td></td>
<td>(51.9-85.7)</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>92.6</td>
<td>&lt;0.001</td>
<td>90.2</td>
</tr>
<tr>
<td>(80-98.4)</td>
<td></td>
<td>(76.8-97.2)</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>81.9</td>
<td>&lt;0.001</td>
<td>81.9</td>
</tr>
<tr>
<td>(71.1-90)</td>
<td></td>
<td>(71.1-90)</td>
<td></td>
</tr>
</tbody>
</table>

(*) Data are percentages of patients with a 95% confidence interval in parentheses. AP, abbreviated protocol; DWI, diffusion-weighted imaging; FDP; full dynamic protocol.

A representative image of the lesions which were upgraded from BI-RADS 3 to BI-RADS 4 with AP+DWI was demonstrated in Figure 3. For both readers, the accuracy of FDP was higher compared to the other models (accuracy was 84.7% for reader 1 and 86.1% for reader 2).
Five patients (2 with DCIS; 2 with grade 2 invasive carcinoma; 1 with grade 3 invasive carcinoma) with malignancies were labeled as benign by both readers in all models. One patient with DCIS was missed out by both readers when they used AP and AP+DWI. In that specific patient, the malignancy was diagnosed with FDP by reader 1. One patient with DCIS was diagnosed correctly in all models by both readers. While the malignancy was missed out with AP in 2 patients (grade 1 and grade 3 invasive carcinoma), the correct diagnosis was achieved with FDP by both readers. In those patients, the malignancy was diagnosed correctly with AP+DWI by reader 1. In 20 malignant patients (14 with grade 2; 3 with grade 3; 2 with grade 2 invasive carcinoma and 1 with DCIS), patients were assigned with the BI-RADS categories of 4 or above in all models by both readers.

Table 2. Inter-observer agreement of established models

<table>
<thead>
<tr>
<th>Model</th>
<th>Kappa value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>0.907</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>AP+DWI</td>
<td>0.825</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FDP</td>
<td>0.858</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

AP, abbreviated protocol; DWI, diffusion-weighted imaging; FDP, full dynamic protocol.

The agreement in the BI-RADS category assessments was almost perfect in all models between two readers (k value, 0.907, 0.825, and 0.858 in AP, AP+DWI, and FDP, respectively; Table 2). The agreement was higher in AP as compared to the other models. The lowest agreement was found in AP+DWI. There was disagreement in 3 patients (2 with grade 2 carcinoma; 1 with fibroadenoma) by AP. By AP+DWI, no agreement was achieved in 6 patients (3 with fibroadenoma; 1 with fibrocystic disease; 1 with grade 2 carcinoma; 1 with grade 3 carcinoma) between readers. By FDP, there was disagreement in 5 patients (3 with fibroadenoma; 1 with fibrocystic disease; 1 with DCIS).

DISCUSSION

Breast screening is substantially based on bilateral mammography which has been proven to reduce mortality by 16-40% in the literature. Nevertheless, heterogeneous and dense breast pattern decreases the sensitivity of mammography. The ACR has recommended various additional imaging techniques including breast tomosynthesis, breast ultrasound, and MRI with FDP or different APs, especially in patients with a history of BC, dense tissue, and BC diagnosed before 50 years of age. In recent years, among these techniques, AP has gained popularity in breast MRI due to the advantages of patient comfort, cost-effectiveness, and minimization of interpretation time. While achieving the purpose of the APs, the risk of missing out on the malignancy is the main problem of APs in breast MRI. Our study demonstrated that FDP had a higher sensitivity than the other models. Nevertheless, AP provided a more accurate diagnosis in benign patients with higher specificity. A recent meta-analysis evaluating the diagnostic performance of AP, it has not been able to replace FDP in several settings, yet. Besides its use in screening, the problem-solving role of FDP including the assessment of disease extension or response to chemotherapy limits the widespread use of AP against FDP in breast MRI. Our study demonstrated that FDP had a higher sensitivity than the other models. Nevertheless, AP provided a more accurate diagnosis in benign patients with higher specificity.
performance of AP and FDP in either screening studies or enriched cohorts revealed that the diagnostic performance of AP was equivalent to FDP in enriched cohorts but lower than that of FDP in screening studies. However, the heterogeneity in APs investigated in the studies included in that meta-analysis might limit the evaluation of the performance of AP compared to FDP. Contrary to the literature, our study indicated better overall accuracy with FDP than with AP and AP+DWI. Despite its advantages, the risk of missing out the malignancy limits the use of AP. Additionally, not having a well-defined single protocol of AP leads to confusion and hinders the universal use of the protocol.

Most of the previous studies evaluated the performance of abbreviated protocols similar to the one reported by Kuhl et al. Some studies analyzed APs with T2-weighted images. In the literature, there was no consensus about the sequences which have to be included in AP. At this point, DWI has gained popularity due to its being non-invasive and cost-effective. As mentioned in previous reports, DCE-MRI had high sensitivity but lower to moderate specificity which may result in unnecessary biopsy procedures. The majority of previous studies showed that the specificity of FDP improved with the addition of DWI. However, only a few of the previous studies evaluated the addition of DWI to AP. Shao et al. suggested that AP combined with DWI had an equivalent diagnostic performance to FDP and better sensitivity and accuracy than AP, with no decrease in specificity. Chen et al. also showed better sensitivity and specificity with the integration of DWI to AP compared to the use of AP alone. Similar to previous studies, in our study, the sensitivity increased when DWI was added to AP. However, we found a mild decrease in specificity in AP+DWI compared to AP alone which was different from the outcomes of previous studies. These different results may be related to the characteristics of our patient population who had MRI due to the suspicious findings on mammography or ultrasonography instead of as a part of cancer screening.

Although APs were generally based on shortening the duration of scanning and interpretation, the use of contrast agents is one of the limitations of AP. Studies comparing DWI with contrast-enhanced techniques in AP have been also reported in the literature. Yamada et al. found that unenhanced abbreviated MRI based on DWI was comparable to AP based on postcontrast MRI including MIP images of the second early phase (60–120 s) on DCE-MRI and fat-suppressed T2-weighted imaging in the lesions ≤2 cm in diameter. In the study of Yamada et al., the sensitivity of AP including contrast-enhanced sequence and T2 image was higher than the sensitivity we obtained for AP in our study. This may be related to the diagnostic impact of the T2 image which was not included in AP in our study. The study of Bickelhaupt et al. analyzing the efficacy of AP combined with DWI with background suppression (DWIBS) and T2 images, showed better sensitivity and specificity in DWIBS protocol than in postcontrast AP (the sensitivity/specificity for DWIBS protocol and postcontrast AP was 92%/94% and 85%/90%, respectively). A previous study reported similar sensitivity but improved specificity for the unenhanced protocol which included DWI compared to DCE-MRI. In our study, the overall accuracy of AP combined with DWI was higher than that of AP alone. Although the diagnostic performance of FDP was better than that of AP and AP+DWI, the potential value of DWI in the diagnosis should be primarily considered while determining the components of AP. Additionally, the DWIBS protocol, reported in the study of Bickelhaupt et al., may enhance the diagnostic performance of DWI in AP for the prediction of malignancy.

We observed that in most of the cases, invasive carcinomas were diagnosed correctly in all models. In our study, among the patients with DCIS, only one patient was diagnosed as malignant (BI-RADS 4 or above category) by two readers with all models. Shao et al. indicated that the probability of missing out on the malignancy was fewer in AP+DWI compared to AP in patients with invasive carcinoma. However, Shao et al. suggested that the risk of missing out on DCIS was comparable between the protocols in the study. In our study, the impact of AP on the diagnosis of DCIS was not evaluated adequately due to the limited number of patients with DCIS.

We found almost perfect agreement between the readers in AP, AP+DWI, and FDP (k value > 0.80). The highest agreement was observed in AP (k value, 0.907) followed by FDP (k value, 0.885) in our study. Even though the lowest agreement was in AP+DWI (k value, 0.825), the agreement category of the combined AP+DWI was still almost perfect.
Although, there was a lower agreement among the readers in DWI-combined AP compared to the other models; the high k value of AP+DWI represented the high-reliability potential of DWI. Shao et al. demonstrated high agreement in AP, AP+DWI, and FDP which is consistent with our study. Additionally, they reported the agreement between AP+DWI and FDP to be higher than that between AP and FDP. In the present study, disagreement was observed mostly in the cases of fibroadenoma by AP+DWI and FDP. There was no disagreement between the readers when it came to invasive carcinomas diagnosed by FDP.

Our study has some limitations that need to be pointed out here. First, the study design was retrospective and our sample size was small to generalize the results. Second, our study evaluated the agreement between the two readers without considering the experience levels of readers which may have influenced the agreement. Third, all examinations were performed at a single center by using the same protocol which might have induced similar approaches by readers. Tru-cut biopsy or excisional biopsy was used for the histopathologic diagnoses in this study; the Tru-cut biopsy might have impacted the accuracy of histopathological results due to the lack of evaluation of all parts of the specimens. The present study included patients who underwent MRI due to suspicious lesions on the specimens. The present study included patients who underwent MRI due to suspicious findings on mammography or ultrasonography. For this reason, the diagnostic performance of APs in the screening of women with high risk was not taken into consideration in our analyses due to the characteristics of the patients.

In conclusion, we found AP had a slightly lower diagnostic performance than FDP did, which is contrary to the results in the literature. However, the agreement was higher in AP, AP or AP combined with DWI was found to be efficient in the diagnosis of invasive carcinoma. Adding DWI to AP decreased the possibility of missing out on malignancies when compared to AP alone. While determining the components of AP, DWI should be given priority as it has higher accuracy potential in the diagnosis.

REFERENCES


9. Zhu CR, Chen KY, Li P, Xia ZY, Wang B. Accuracy of multiparametric MRI in distinguishing the breast...