Diagn Interv Radiol 2022; DOI: 10.5152/dir.2022.201075

BREAST IMAGING

ORIGINAL ARTICLE

Application of the Kaiser score by MRI in patients with breast lesions by ultrasound and mammography

Qibin Wang
Fengli Fu
Yao Chen
Di Yang
Jianjun Zhang
Huajun Yu
Lina Su

PURPOSE
This study aimed to verify whether the use of the Kaiser score can improve the diagnostic performance in breast magnetic resonance imaging (MRI) for suspicious lesions and avoid further invasive diagnostic approaches.

METHODS
This retrospective study enrolled 97 patients who underwent breast MRI before undergoing breast biopsy or surgery. Evaluations were conducted on all MRI images individually by 2 radiologists using the Kaiser score. Neither radiologist had the knowledge of the final histopathological diagnosis. The ability of the Kaiser score in diagnosis was established via a receiver performing characteristic (ROC) analysis, which was measured by the area under the ROC curve (AUC). Youden index was used to define the optimal cutoff value. Kaiser score categories were dichotomized into positive (cutoff score > 4) and negative scores (cutoff score ≤ 4). Cohen's kappa coefficient was used to analyze the inter-rater agreement.

RESULTS
Histopathology revealed 56 malignant and 41 benign lesions. The AUC for all lesions evaluated by the Kaiser score was 0.992 (95% CI: 0.981-1.0) and 0.958 (95% CI: 0.920-0.996) for 2 radiologists, respectively. Inter-rater agreement of the dichotomized Kaiser score was excellent ($\kappa = 0.894$, $P < .001$). A total of 20 lesions (33.8%) previously categorized as BI-RADS 4 were reduced to BI-RADS 2/3 (19 benign lesions and 1 malignant lesion).

CONCLUSION
The Kaiser score is a valuable auxiliary diagnostic tool for improving the diagnostic ability of radiologists, whose experiences in breast MRI are diverse. In some cases, the application of the Kaiser score could possibly avoid unnecessary breast biopsies.

Breast magnetic resonance imaging (MRI) has been proven as a considerable assistant method when clinical and routine imaging examinations (mammography and ultrasound) cannot completely solve the diagnostic task of confirming or excluding malignant breast lesions.[1-4] Solving the problem depends on the negative predictive value (NPV) of breast MRI.[5] A meta-analysis performed by Bennani-Baiti et al.[4] reported a sensitivity of 99% with a high NPV of 100% for the evaluation of non-calcified equivocal findings. For example, asymmetry in mammography, visible only in one view, in which a negative finding of breast MRI can effectively exclude malignancy.[6,7]

However, a variety of MRI protocol and different criteria for interpretation may result in an increase in the false-positive rate.[8-16] In order to facilitate the interpretation, the Breast Imaging Reporting and Data System (BI-RADS)[16] is developed to become a structured report for breast imaging (mammography, ultrasound, and breast MRI). It provides a lexicon of descriptors to ensure effective and unified communication between the radiologist and other physicians in the decision-making and management of breast lesions.[16-18] However, without a clear classification system such as a flowchart, the clinical use of BI-RADS to confirm or exclude malignant breast lesions on breast MRI is still challenging, especially for younger radiologists. Thus, with the intention of improving this issue, a definitive classification system combined with BI-RADS is needed.

From the Department of Radiology (L.S. ✉ suln13@163.com), Zhejiang Hospital, Hangzhou, China.

Received 31 January 2021; revision requested 23 March 2021; last revision received 1 July 2021; accepted 7 July 2021.

Available online: 11 June 2022.

DOI: 10.5152/dir.2022.201075

You may cite this article as: Wang Q, Fu F, Chen Y, et al. Application of the Kaiser score by MRI in patients with breast lesions by ultrasound and mammography. *Diagn Interv Radiol.* 2022;28(4):322-328.

Kaiser score is such a classification system in the form of a flowchart.[19-21] It is based on the lesion's morphology, enhancement pattern, apparent diffusion coefficient (ADC) value, as well as the presence of micro-calcifications (if available).[22] The final score can result in corresponding BI-RADS category with management recommendation. Previous studies[20-22,24] have investigated the role of the scoring system in distinguishing benign from malignant lesions on breast MRI; however, these studies did not include the optional moderators such as suspicious mammographic microcalcifications and high ADC value.

Therefore, this study aimed to verify whether the additional use of the Kaiser score (with optional moderators) can improve the diagnostic performance in breast MRI for suspicious lesions and avoid superfluous breast biopsies.

Methods

Study population and reference standard

The Ethics Committee at Zhejiang Hospital approved this study (approval no.:2019 11K), waiving informed consent because of its retrospective nature. From April 2018 to February 2020, 97 patients with suspicious lesions on mammography or ultrasonography (BI-RADS 4-5) on the institutional database were enrolled. These patients underwent breast MRI before undergoing percutaneous breast biopsy or surgery.

Histopathological diagnosis was carried out by professional and skilled breast pathologists by means of percutaneous biopsy or open surgery and it was considered the reference standard.

Breast MRI protocol

All breast MRI were carried out on a 3.0T Simens Skyra scanner. A bilateral, dedicated, 16-channel phased-array breast coil was used. Standard imaging was carried out as well, including 4 sequences as follows: (1) an axial, fast spin-echo T1-weighted imaging sequence, after generating the image of a bilateral lateral localizer; (2) an axial T2-weighted turbo inversion recovery magnitude (T2 TIRM) sequence; (3) an axial regular diffusion-weighted imaging sequence, with b values of 0 and 1000 s/mm^2; (4) 29 phases of dynamic contrast-enhanced T1-weighted images were obtained. A high-pressure syringe was employed to deliver gadolinium-diethylenetriaminepentaacetic acid (Gd-DTPA, Magnevist) through the antecubital vein using a high-pressure injector (0.2 mmol/kg of body weight at 3 mL/s), followed by a 20 mL saline flush. The detailed parameters are shown in Table 1.

Main points

- Kaiser score is a flowchart for diagnosing breast lesions on magnetic resonance imaging.
- Kaiser score shows excellent inter-rater agreement in the characterization of breast lesions.
- In some cases, applying the Kaiser score possibly could avoid unnecessary breast biopsies.

Image analysis

All MRI images were assessed retrospectively by 2 radiologists individually using the fifth edition of BI-RADS[16] and the Kaiser score,[19] which was previously called the tree flowchart. R1 represents the radiologist with 10 years of experience and R2 is the radiologist with 4 years of experience. Neither radiologists had the knowledge of the final histopathological diagnosis. The suspicious lesion with a histopathological analysis was noted in advance so that each radiologist analyzed the same lesion.

According to malignant probability, the final BI-RADS results were scored as 3, 4, or 5, and 3 represents that the lesion probably is benign, where a short-term follow-up is suggested, 4 represents that the lesion is suspected to be abnormal and biopsy is recommended, and 5 strongly suggests that the lesion is malignant and biopsy is recommended.[16] Five criteria of morphology and kinetics are edema, lesion margins, root sign, contrast enhancement kinetics, and internal enhancement patterns. The optional moderators should be considered: suspicious microcalcifications on mammography should upgrade the Kaiser scores by 2 points to avoid false-negative ductal carcinoma in situ (DCIS) diagnoses[19]; high ADC values which are greater than 1.4×10^{-3} mm^2/s should be regarded as an additional for benign lesions,[24] which may reduce the Kaiser score by 4 points. According to the above criteria, the Kaiser score ranges from 1 to 11, corresponding to the increase of malignant probability (Figure 1) and the final score can be translated into BI-RADS categories as follows: 1-4: minimal risk of breast cancer—BI-RADS 2/3; 5-7: intermediate risk of breast cancer—BI-RADS 4; 8-11: high risk of breast cancer—BI-RADS 5.[19]

Statistical analysis

All calculations are performed based on each lesion. With the intention of establishing the Kaiser score's overall diagnostic feasibility, a receiver operating characteristic (ROC) analysis was carried out by measuring the area under the ROC curve (AUC). Youden index was used to define the optimal cutoff value. Evaluations on the sensitivity, specificity, as well as positive and NPVs were conducted. The Kaiser score categories were dichotomized into positive (cutoff score $>$ 4) and negative scores (cutoff score $\leq$ 4). Inter-rater agreement between the 2 experienced breast radiologists was analyzed by Cohen's kappa coefficient. Values (κ) were interpreted as suggested: poor (κ is less than 0.20), fair (κ ranges from 0.21 to 0.40), moderate (κ ranges from 0.41 to 0.60), good (κ ranges from 0.61 to 0.80), and excellent (κ is greater than 0.81).[23] In our research, Statistical Package for the Social Sciences (version 22.0) was employed to perform the

Table 1. Protocols of MRI sequences

	T1WI	T2 TIRM	DWI	DCE-MRI
TR (ms)	260	3770	5410	4.5
TE (ms)	2.56	70	56	1.54
FOV (mm)	370	340	340	360
Slice thickness (mm)	4	4	5	4
Matrix size	512 × 512	448 × 448	156 × 240	448 × 448
NEX	1	2	2	1
Acquisition time (s)	58	160	261	494

MRI, magnetic resonance imaging; T1WI, T1-weighted imaging; T2 TIRM, T2-weighted turbo inversion recovery magnitude; DWI, diffusion-weighted imaging; DCE, dynamic contrast-enhanced; TR, repetition time; TE, echo time; FOV, field of view; NEX, number of excitations.

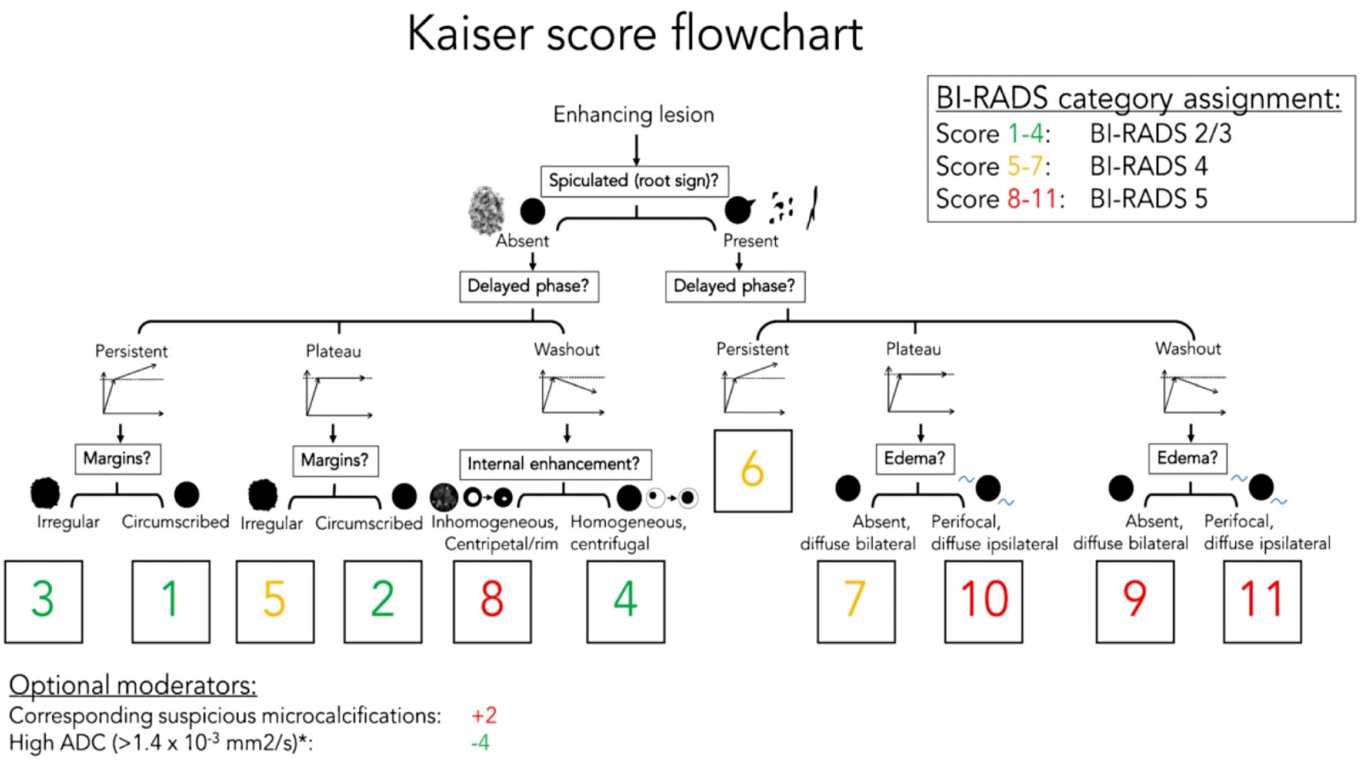

Figure 1. Kaiser score flowchart: The Kaiser score is assigned by following a simple flowchart from the top to the bottom, which lets the reader assign the presence or absence of 4 diagnostic criteria.

statistical analysis. The *P* value indicates the level of significance, if it is less than 0.05, the difference is significant.

Results

A total of 97 lesions (maximum diameter range, 0.5-8.3 cm) from 97 patients (mean age, 48 ± 10.7 years; age range, 25-74 years; median age, 48 years) were included in our study. Histopathology revealed 56 (57.7%) malignant lesions and 41 (42.3%) benign lesions. The malignant lesions consisted of 46 invasive ductal carcinomas, 3 invasive papillary carcinomas, 1 medullary carcinoma, and 6 DCIS. The benign lesions consisted of 23 fibroadenomas, 13 breast adenosis, and 5 intraductal papilloma.

The Kaiser scores of the 97 lesions were as follows: 1 (n = 18), 2 (n = 12), 3 (n = 3), 4 (n = 6), 5 (n = 3), 6 (n = 3), 7 (n = 7), 8 (n = 6), 9 (n = 17), 10 (n = 8), and 11 (n = 14). Kaiser score was >4 in 58 patients (59.8%) and ≤4 in 39 patients (40.2%).

Examples of the Kaiser score applications are shown in Figures 2-4.

The AUC for all lesions evaluated by the Kaiser score was 0.992 (95% CI: 0.981-1.0) and 0.958 (95% CI: 0.920-0.996) for R1 and R2, respectively (Figure 5). With a cutoff of >4 determined by the maximum Youden index (0.907), the Kaiser score achieved a sensitivity and specificity of 0.980 and 0.927 for R1, 0.940 and 0.805 for R2, respectively (Table 2).

Two readers independently read 97 consecutive cases and the value of k was 0.894, with $P < .001$, meaning that the inter-rater agreement between the 2 readers was excellent.

The application of the Kaiser scoring system changed the BI-RADS assignation of 53 lesions. Among them, 1 lesion previously classified as BI-RADS 3 was upgraded to BI-RADS 5 (final histology was invasive ductal carcinoma); 29 lesions previously classified as BI-RADS 4 were raised to BIRADS 5, all the final histopathological diagnoses showed malignancy; 20 lesions previously classified as BI-RADS 4 were relegated to BI-RADS 2/3, 19 of which proved to be benign and 1 was confirmed to be malignant; and 3 of them were relegated from previous BI-RADS 5 to BI-RADS 4 (all proven malignancy) (Table 3).

Discussion

The Kaiser score is a flowchart incorporating not only 5 criteria of morphology and kinetics mentioned above but also optional moderators (the presence of micro-calcifications and ADC value) to distinguish benign breast lesions from the malignant ones.[19-21] DCIS lesions presented suspicious microcalcifications on mammography may be false negative on breast MRI.[3] In this situation, the Kaiser scores should be upgraded by 2 points to avoid false-negative DCIS diagnoses.[19] In addition, high ADC values, which are greater than 1.4×10^{-3} mm^2/s, as an optional moderator may reduce 4 points of the Kaiser score.[24] The result of following the flowchart from the top to the bottom is a diagnostic score, ranging from 1 to 11. It reflects an increase in the possibility of malignancy. Then the corresponding BI-RADS category is given with a clinical decision.

According to ROC curve analysis, the results of our study demonstrated that the performance of Kaiser score was excellent in diagnosing suspicious breast lesions, even for the less-experienced radiologist (AUC was 0.958). The results of this study achieved higher diagnostic accuracy than the previous studies.[21-23,25] In the study by Woitek et al.,[25] different MRI scans from a variety of units and field strengths were employed, which is likely to limit the ability of Kaiser score in interpreting images and is one potential reason why the AUC was lower than our study. However, the difference in image quality of MRI scans shows the consistency of the Kaiser score in clinical

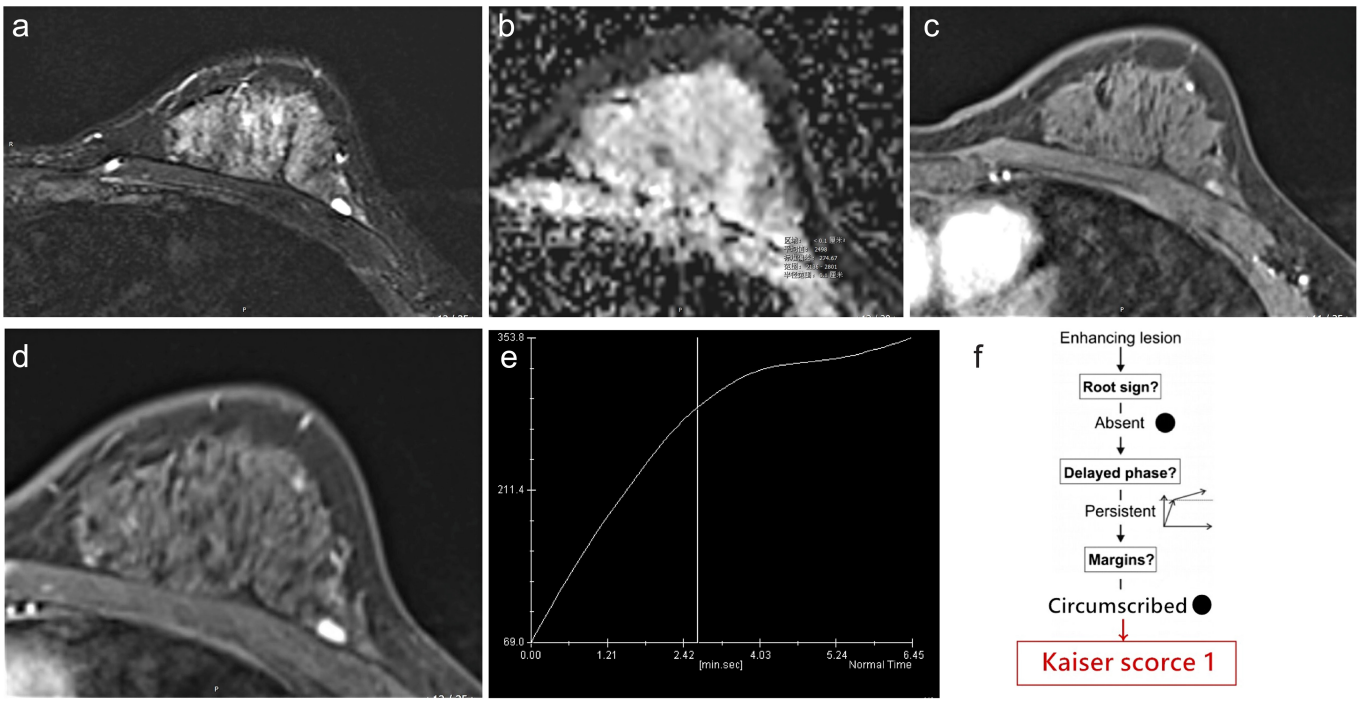

Figure 2. a-f. Representative axial slices of T2-TIRM sequence **(a),** apparent diffusion coefficient (ADC) image (b), early **(c)** and (**b**), delayed **(d)** post-contrast T1-weighted sequences, and signal-intensity time curve **(e)** and flowchart **(f)** are shown. A mass lesion presented with absent root sign, persistent enhancement, circumscribed margin, and high ADC (2.5×10^{-3} mm^2/s), classified as Kaiser score 1. Pathological diagnosis revealed a fibroadenoma.

applications. Wengert et al.[22] added a lot of evidence and proposed the role of the Kaiser score in solving BI-RADS 4 mammography calcifications, and Maria et al.[21] showed that tree flowchart is applicable not only to masses but also to non-mass lesions.

In our study, 20 lesions previously classified as BI-RADS 4 were relegated to BI-RADS 2/3, including 19 proven to be benign and 1 confirmed to be malignant. Based on this result, the Kaiser score could avoid some rate of unnecessary breast biopsies, thereby reducing medical expense, physically uncomfortable of patients, as well as the risk of negative influences resulting from the invasive diagnostic approaches. The false-negative result was further analyzed retrospectively. Because the lesion

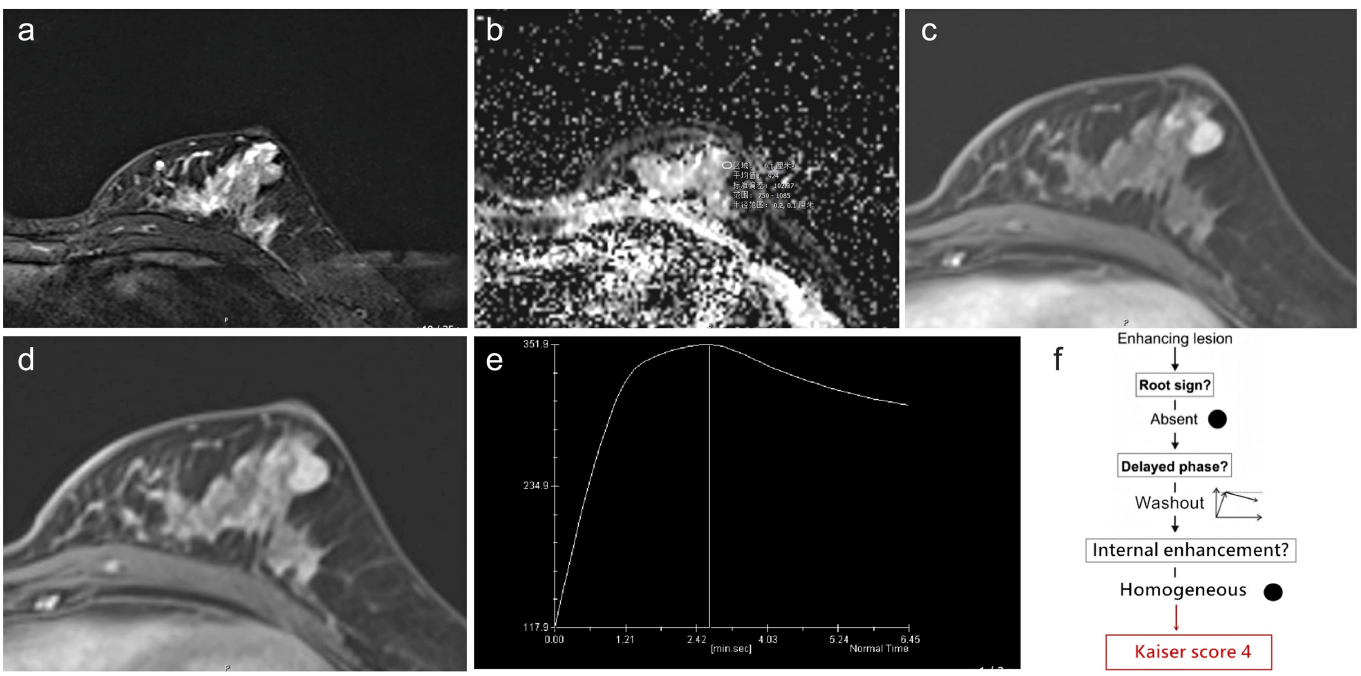

Figure 3. a-f. T2 TIRM sequence **(a),** ADC image (**b**), early **(c)** and delayed **(d)** post-contrast T1-weighted sequences and signal-intensity time curve **(e)** and flowchart **(f)** are shown. A mass lesion in the left breast of a 52-year-old female shows no root sign, a washout enhancement curve type, and homogeneous internal enhancement that corresponds to a Kaiser score of 4. The patient requested surgery that revealed an intraductal papilloma.

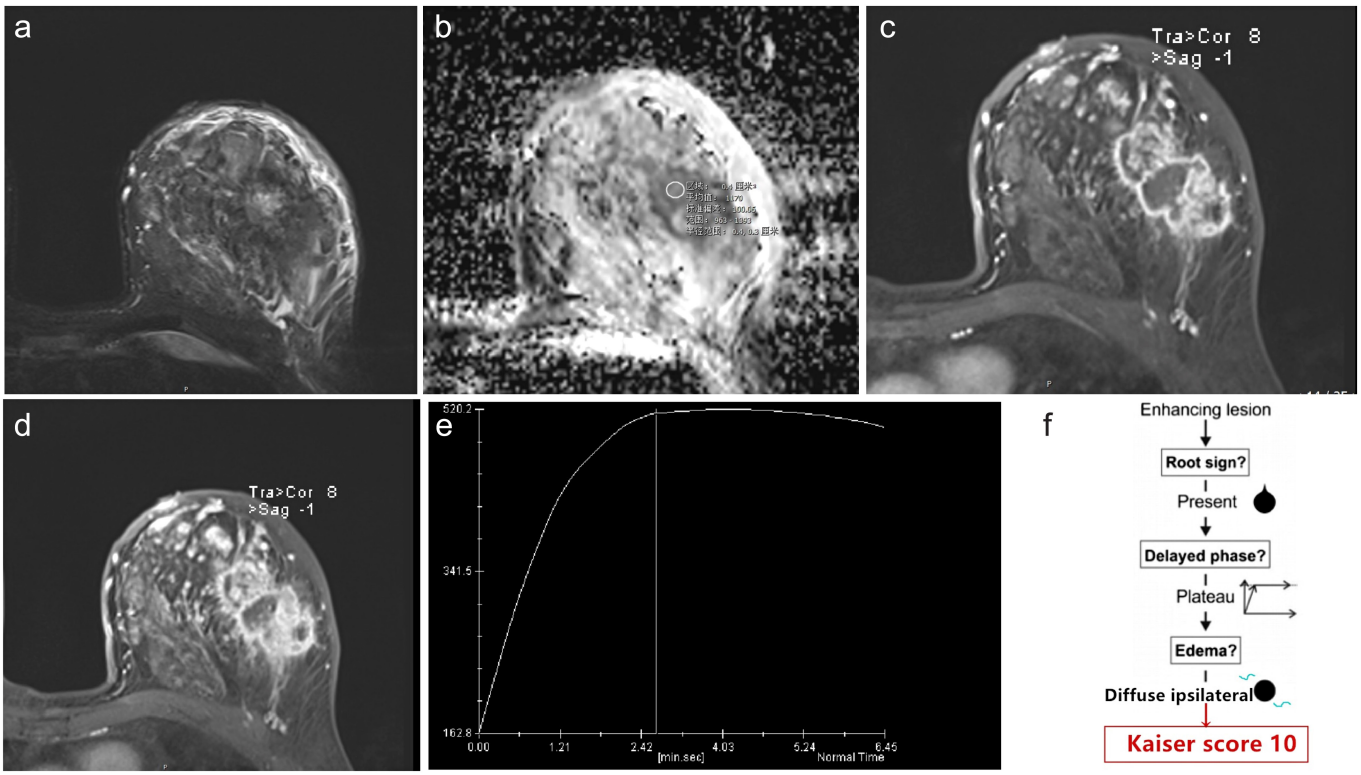

Figure 4. a-f. Representative axial slices of T2 TIRM sequence **(a),** ADC image (**b**), early **(c)** and delayed **(d)** post-contrast T1-weighted sequences, and signal-intensity time curve **(e)** and flowchart **(f)** are shown. A mass lesion presented with root sign, plateau enhancement, diffuse ipsilateral edema and was classified as Kaiser score 10. Pathological diagnosis revealed an invasive ductal carcinoma.

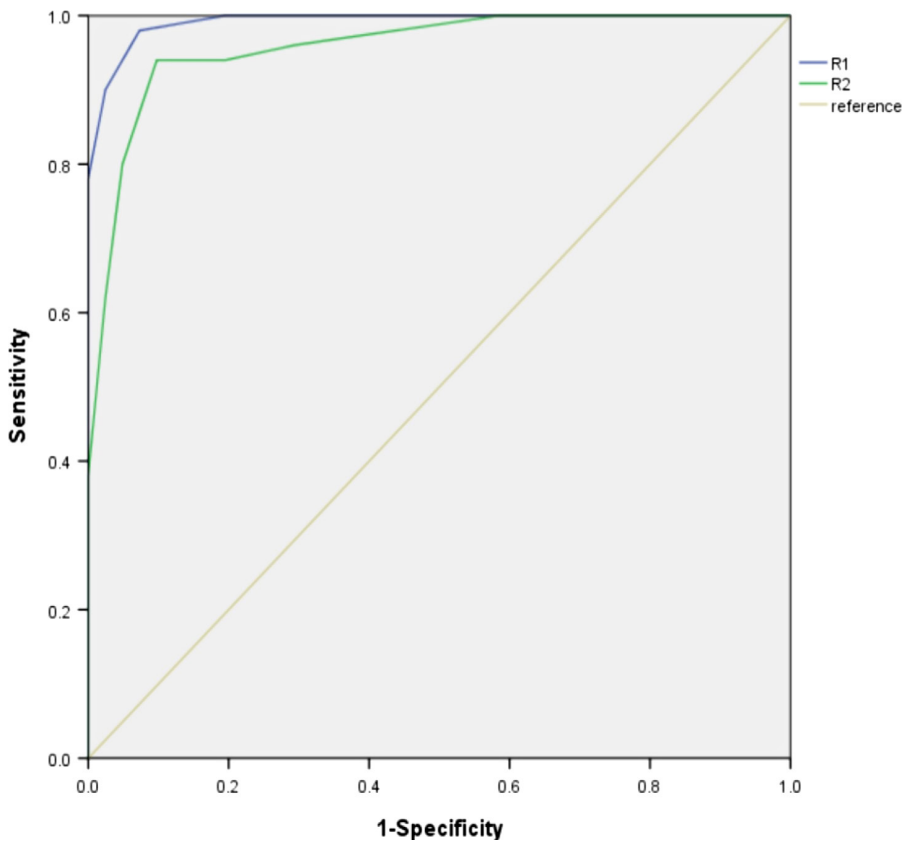

Figure 5. Receiver-operating characteristic (ROC) curves of all lesions the Kaiser score.

Table 2. Diagnostic parameters of the Kaiser score readings for each reader

	R1		R2	
Cutoff	Sensitivity	Specificity	Sensitivity	Specificity
>4	0.980	0.927	0.940	0.805

R1, radiologist with 10 years of experience; R2, radiologist with 4 years of experience.

Table 3. Summary of MRI reader BI-RADS and MRI Kaiser BI-RADS

BI-RADS allocations	MRI reader BI-RADS			MRI Kaiser BI-RADS		
	BI-RADS 3	BI-RADS 4	BI-RADS 5	BI-RADS 2/3	BI-RADS 4	BI-RADS 5
Total lesions	20	59	18	39	13	45

BI-RADS, Breast Imaging-Reporting and Data System; MRI, magnetic resonance imaging.

was too small (with a diameter of 0.5 cm), we had ignored its tiny root sign (Figure 6). It should be noted that when the size of the lesion is small, we need to observe more carefully. Our findings revealed that applying a cutoff value ≤4 can exclude malignancy to a large extent and produce a small number of false-negative results (i.e., all cases except 1 invasive ductal carcinoma, as we ignored a slight root sign). As the Kaiser score reflects an increased possibility of malignancy, a positive MRI result (Kaiser score >4) indicates that a repeat biopsy or surgical biopsy may be needed even if the first biopsy was negative, to make an early diagnosis, early treatment, and improve the prognosis.

Based on the Kaiser scoring system, evaluations on the lesions were conducted by 2 radiologists, whose experiences in breast MRI were different. The value of k was 0.894, with $P < .001$, meaning that the inter-rater agreement between the 2 readers was excellent. Our results demonstrated a reduction of inter-rater variability related to the experience of readers. That is, using a standardized diagnostic algorithm, the Kaiser score is easily applicable in breast MRI and helpful to radiologists in differentiating benign breast lesions from malignant ones, especially in less-experienced radiologists.

This research has several limitations as well. First of all, its retrospective nature shows that there is a certain degree of bias in choosing more suspicious cases. The potential effect tends to overestimate the sensitivity and underestimate the specificity. Consequently, our future research should reinforce both the possibility of breast MRI and Kaiser score in this setting. Second, the relatively small number of patients may affect the cutoff value differentiating benign from malignant breast lesions on breast MRI. Therefore, larger prospective studies would be needed to provide a more robust assessment of the clinical value of the Kaiser score. Third, since the influence of background parenchymal enhancement on the ability of breast MRI in diagnosis is still uncertain,[26,27] whether it may indeed have an influence on Kaiser score's diagnostic accuracy was not assessed.

In conclusion, the Kaiser score is a valuable auxiliary diagnostic tool for improving the diagnostic ability of radiologists with

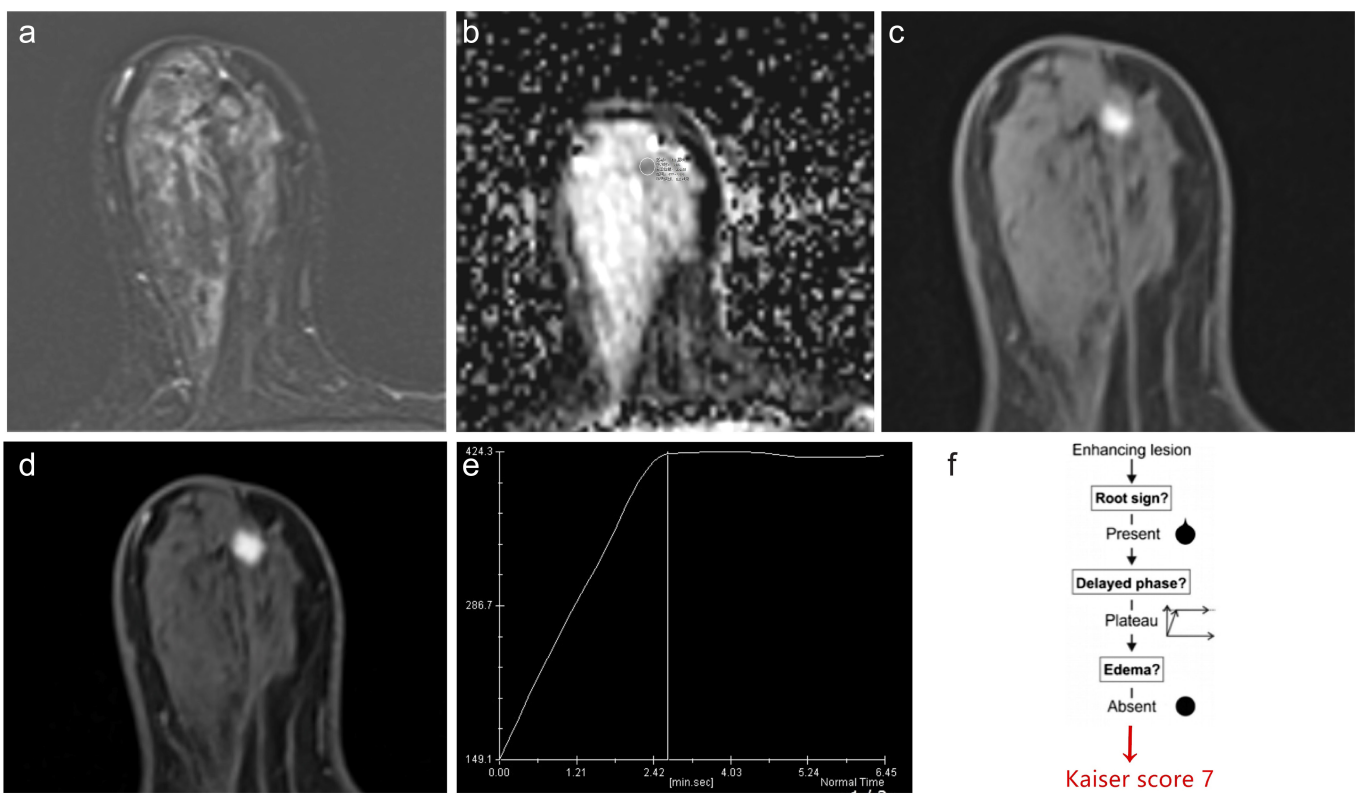

Figure 6. a-f. Representative axial slices of T2 TIRM sequence **(a),** ADC image (**b**), early **(c)** and delayed **(d)** post-contrast T1-weighted sequences, signal-intensity time curve **(e)** and flowchart **(f)** are shown. The false-negative lesion (with a diameter of 0.5 cm) was formerly classified as Kaiser score 4 because its tiny root sign had been ignored. Pathological diagnosis revealed an invasive ductal carcinoma.

different experiences in breast MRI. In some cases, the application of the Kaiser score could potentially avoid unnecessary breast biopsies.

Financial disclosure

This work was supported by Zhejiang Provincial Health Department [grant no.: 2019KY265] and Zhejiang Provincial Health Department [grant no.: 2020KY385].

Conflict of interest disclosure

The authors declared no conflicts of interest.

References

1. Kaiser WA. *MR Mammography (MRM)*. 1st ed. Berlin Heidelberg: Springer; 1994.
2. Mann RM, Kuhl CK, Kinkel K, Boetes C. Breast MRI: guidelines from the European Society of Breast Imaging. *Eur Radiol*. 2008;18(7):1307-1318. [CrossRef]
3. Bennani-Baiti B, Baltzer PA. MR imaging for diagnosis of malignancy in mammographic microcalcifications: a systematic review and meta-analysis. *Radiology*. 2017;283(3):692-701. [CrossRef]
4. Bennani-Baiti B, Bennani-Baiti N, Baltzer PA. Diagnostic performance of breast magnetic resonance imaging in non-calcified equivocal breast findings: results from a systematic review and meta-analysis. *PLoS ONE*. 2016;11(8):e0160346. [CrossRef]
5. Mann RM, Cho N, Moy L. Breast MRI: state of the art. *Radiology*. 2019;292(3):520-536. [CrossRef]
6. Giess CS, Chikarmane SA, Sippo DA, Birdwell RL. Clinical utility of breast MRI in the diagnosis of malignancy after inconclusive or equivocal mammographic diagnostic evaluation. *AJR Am J Roentgenol*. 2017;208(6):1378-1385. [CrossRef]
7. Moy L, Elias K, Patel V, et al. Is breast MRI helpful in the evaluation of inconclusive mammographic findings? *AJR Am J Roentgenol*. 2009;193(4):986-993. [CrossRef]
8. Choi BG, Kim HH, Kim EN, et al. New subtraction algorithms for evaluation of lesions on dynamic contrast-enhanced MR mammography. *Eur Radiol*. 2002;12(12):3018-3022. [CrossRef]
9. Trecate G, Tess JD, Vergnaghi D, et al. Breast microcalcifications studied with 3D contrast-enhanced high-field magnetic resonance imaging: more accuracy in the diagnosis of breast cancer. *Tumori*. 2002;88(3):224-233. [CrossRef]
10. Cecil KM, Schnall MD, Siegelman ES, Lenkinski RE. The evaluation of human breast lesions with magnetic resonance imaging and proton magnetic resonance spectroscopy. *Breast Cancer Res Treat*. 2001;68(1):45-54. [CrossRef]
11. Alamo L, Fischer U. Contrast-enhanced color Doppler ultrasound characteristics in hypervascular breast tumors: comparison with MRI. *Eur Radiol*. 2001;11(6):970-977. [CrossRef]
12. Huang W, Fisher PR, Dulaimy K, Tudorica LA, O'Hea B, Button TM. Detection of breast malignancy: diagnostic MR protocol for improved specificity. *Radiology*. 2004;232(2):585-591. [CrossRef]
13. Kristoffersen Wiberg M, Aspelin P, Perbeck L, Boné B. Value of MR imaging in clinical evaluation of breast lesions. *Acta Radiol*. 2002;43(3):275-281. [CrossRef]
14. Bluemke DA, Gatsonis CA, Chen MH, et al. Magnetic resonance imaging of the breast prior to biopsy. *JAMA*. 2004;292(22):2735-2742. [CrossRef]
15. Peters NH, Borel Rinkes IH, Zuithoff NP, Mali WP, Moons KG, Peeters PH. Meta-analysis of MR imaging in the diagnosis of breast lesions. *Radiology*. 2008;246(1):116-124. [CrossRef]
16. American College of Radiology. *ACR BI-RADS Atlas: Breast Imaging Reporting And Data System*. 5th ed. Virginia: Reston; 2013.
17. Burnside ES, Sickles EA, Bassett LW, et al. The ACR BI-RADS experience: learning from history. *J Am Coll Radiol*. 2009;6(12):851-860. [CrossRef]
18. Spak DA, Plaxco JS, Santiago L, Dryden MJ, Dogan BE. BI-RADS® fifth edition: a summary of changes. *Diagn Interv Imaging*. 5th ed. 2017;98(3):179-190. [CrossRef]
19. Dietzel M, Baltzer PAT. How to use the Kaiser score as a clinical decision rule for diagnosis in multiparametric breast MRI: a pictorial essay. *Insights Imaging*. 2018;9(3):325-335. [CrossRef]
20. Baltzer PA, Dietzel M, Kaiser WA. A simple and robust classification tree for differentiation between benign and malignant lesions in MR-mammography. *Eur Radiol*. 2013;23(8):2051-2060. [CrossRef]
21. Marino MA, Clauser P, Woitek R, et al. A simple scoring system for breast MRI interpretation: does it compensate for reader experience? *Eur Radiol*. 2016;26(8):2529-2537. [CrossRef]
22. Wengert GJ, Pipan F, Almohanna J, et al. Impact of the Kaiser score on clinical decision-making in BI-RADS 4 mammographic calcifications examined with breast MRI. *Eur Radiol*. 2020;30(3):1451-1459. [CrossRef]
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174. [CrossRef]
24. Woodhams R, Kakita S, Hata H, et al. Diffusion-weighted imaging of mucinous carcinoma of the breast: evaluation of apparent diffusion coefficient and signal intensity in correlation with histologic findings. *AJR Am J Roentgenol*. 2009;193(1):260-266. [CrossRef]
25. Woitek R, Spick C, Schernthaner M, et al. A simple classification system (the Tree flowchart) for breast MRI can reduce the number of unnecessary biopsies in MRI-only lesions. *Eur Radiol*. 2017;27(9):3799-3809. [CrossRef]
26. Baltzer PA, Benndorf M, Dietzel M, Gajda M, Runnebaum IB, Kaiser WA. False-positive findings at contrast-enhanced breast MRI: a BI-RADS descriptor study. *AJR Am J Roentgenol*. 2010;194(6):1658-1663. [CrossRef]
27. DeMartini WB, Liu F, Peacock S, Eby PR, Gutierrez RL, Lehman CD. Background parenchymal enhancement on breast MRI: impact on diagnostic performance. *AJR Am J Roentgenol*. 2012;198(4):W373-W380. [CrossRef]