The Egyptian Journal of Hospital Medicine (July 2022) Vol. 88, Page 4068- 4075

MRI Diagnostic Performance and Inter-Observer Agreement of Kaiser Score in The Assessment of Different Breast Lesions

***Marwa Makboul, Shimaa Farghaly**
Department of Radio Diagnosis, Faculty of Medicine, Assiut University, Assiut, Egypt
***Corresponding author:** Marwa Makboul, Mobile: 01006541595,
Email: makboul@aun.edu.eg,, **ORCID ID**: https://orchid.org/0000-0001-5001-2622

ABSTRACT

Introduction: to evaluate Kaiser Score's diagnostic efficacy and inter-reader reliability and contrast it with the typical BI-RADS Lexicon.**Materials and methods:** This retrospective study included 100 participants with a total of 109 breast lesions, at Assiut University hospital in the period from January 2021 to June 2021. The fifth edition of the MRI BI-RADS Lexicon was used to determine the BI-RADS category for each lesion after two radiologists reviewed all MRI scans. Following the flowchart for the Kaiser Score, they were then categorized according to BI-RADS category assignments. Finally, a comparison of the Kaiser BI-RADS and score results with the MRI BI-RADS and histological data was made to assess the diagnostic accuracy and inter-observer agreement.

Results: There was no noticeable difference between the Kaiser score and the BI-RADS, MRI BI-RADS, or the BI-RADS as the p-values for the diagnosis of all breast lesions, whether mass or NMEL, were 0.597, 0.84, and 0.495 respectively. The inter-observer agreement between Readers 1 and 2 in the diagnosis of all breast lesions, regardless of whether they were mass lesions or NMELs, using both Kaiser BI-RADS and MRI BI-RADS, was also significantly higher since the p-value was less than 0.001. **Conclusion:** Kaiser score and MRI BI-RADS can be combined to enhance reader agreement and reduce experience-related variability. For readers with less experience, it can also be very helpful in making the diagnosis of doubtful and suspicious breast lesions.

Keywords: Kaiser BI-RADS; Kaiser score; MRI BI-RADS; NMEL.

INTRODUCTION

Worldwide, magnetic resonance imaging (MRI) of the breast is regarded as the most sensitive imaging technique for detecting breast cancer [1,2]. Although MRI has a very high sensitivity for detecting breast lesions, characterizing those lesions requires the interpretation of multiple images with a variety of contrasts [3]. Additionally, without a planned, consistent interpretation approach, the number of false-positive MRI results would rise, lowering the specificity of the test [4].

The American College of Radiology (ACR) Breast Imaging Reporting and Data System (BI-RADS) Lexicon is regarded as the most widely used structured reporting since it provides a consistent vocabulary for common lesion descriptions across clinicians [5].

Since the BI-RADS Lexicon doesn't offer instructions on how to translate imaging signals into diagnostic categories, diagnostic accuracy is very variable and inter-reader agreement is only modest. As a result, needless biopsies of benign lesions are carried out [6]. To help radiologists characterize breast MRI results and increase its specificity for the detection of malignant tumors, Baltzer et al. presented a grading system [7].

This scoring system, known as the Kaiser score, is a tree flowchart that translates imaging findings into a numerical score ranging from 1 to 11, for differentiating between malignant and benign breast lesions. A biopsy is required if the score is higher than 4, and this score allows readers with less medical training to give breast lesions that are suspicious or uncertain a score [8,9].

By increasing inter-reader agreement, the Kaiser rating method is anticipated to reduce experience-related variability [10]. This study compares Kaiser Score with the conventional BI-RADS Lexicon to explore the diagnostic accuracy and inter-reader agreement of Kaiser Score.

MATERIALS AND METHODS:

This retrospective study included 100 patients with a total of 109 lesions who had contrast-enhanced breast MRIs between January 2021 and June 2021.

Ethical consideration:

The study was approved by the Ethics Board of Assiut University and informed written consent was taken from each participant in the study. This work was performed in full accordance with the code of Ethics of the World Medical Association (Declaration of Helsinki) for studies involving humans.

MRI imaging protocol:

A four-channel breast coil and a 1.5 Tesla system (Avanto Siemens Healthcare) were used for every MRI examination.

Following a survey sequence, there was an axial 3D T1 weighted image TR/TE of 8.6/7.4 with a field of view of 400 mm and a slice thickness of 1 mm, an axial T2 fat suppression TR/TE of 5250/60 with a field of view of 380 mm and a slice thickness of 4 mm, and an axial diffusion-weighted imaging TR/TE of 5300/91 with a field of view of 460 mm.

Gadolinium dimeglumine (Gd-DTPA) (Magnevist, Schering AG Berlin, Germany) is supplied intravenously at a dose of 0.1 mmol/kg at a rate of 2 ml/s using a

Received: 17/04/2022
Accepted: 14/06/2022

power injector. Next, a 20 ml saline flush is given using an automatic injector.

Data analysis

Two radiologists with between 10 and 15 years of experience blinded to the preliminary radiological reports and the final histopathology results reviewed all MRI scans.

Using the fifth version of the MRI BI-RADS lexicon as a guide, the readers assigned the BI-RADS category for each lesion from BI-RADS 2 to BI-RADS 5 based on the suspicion of malignancy [(11)].

Following the Kaiser Score flowchart, the two readers categorized all observed lesions by translating the Kaiser Score into BI-RADS category designations. This classification system was based on morphologic and kinetic criteria (root sign, lesion margin, enhancement kinetics, pattern of internal enhancement, and edema), as described in (Table 1), and this straightforward flowchart has 11 assignment categories regarding the probability of malignancy, with Score (1-4) being considered BI-RADS2/3, Score (5-7) being BI-RADS4, and Score 8-11 being BI-RADS5 **(Fig. 1).**

Finally, a comparison of the Kaiser BI-RADS and score results with the MRI-BIRADS and histological data was made to assess the diagnostic accuracy and interobserver agreement.

Statistical analysis

The researcher validated, coded, and ran an SPSS version 24* analysis on the data. Calculated using descriptive statistics are means, standard deviations, medians, ranges, and percentages. The total diagnostic performance was assessed using a receiver operating characteristic (ROC) analysis and the area under the ROC curve. At a cut-off value of > 4, which denoted malignancy, the sensitivity, specificity, LR positive, and negative were calculated. The dichotomized (Kaiser scores 1-4 were considered benign, 5-11 malignant) had a high inter-reader agreement. Kappa statistics were used to evaluate Kaiser score readings. A p-value of 0.05 or less was regarded as significant.

RESULTS

The study included 100 individuals with 109 lesions altogether, with a median age of 46 years, 83.5% (91/109) mass lesions and 16.5% (18/109) non-mass enhanced lesions (NMEL) on Breast MRI.

According to MRI results, uneven border and edema were both observed in 76.1 % of lesions and 69 (63.3 %) of them had root signs, while enhancement curve type III was the most prevalent in all lesions at 54.2 % **(Table 1).** According to the MRI BI-RADS and Kaiser score classification systems: BI-RADS V was the most prevalent in all lesions as well as Kaiser BI-RADSV by 56 % and 70 % respectively, followed by BI-RADS II and IV and Kaiser BI-RADS (II-III) by 18.3 % and 28 % respectively, while BI-RADS III and Kaiser BI-RADS IV were found in only 8 and 11 lesions by (7.4 %) and (10.1 %) respectively **(Table 2).**

Invasive ductal carcinoma was the most prevalent pathology, accounting for 50.5 % of the lesions, followed by invasive lobular carcinoma (11.9 %), Paget's disease (3.7 %), intra-ductal carcinoma (0.9 %), atypical hyperplasia (0.9 %), and apocrine metaplasia (0.9 %), In contrast, only benign pathology was discovered in 31.2 % of them (34/109); the most frequent pathology was fibro-adenoma, which was followed by granulomatous mastitis and fat necrosis by (5.5 % and 3.7 %, respectively), abscess and fibrocystic disease by 1.8 % for each, and focal adenosis and hamartoma by 0.9 % for each.

While fibro-adenoma was the most common pathology in 28 benign lesions when mass lesions were discovered, granulomatous mastitis was the most prevalent pathology in only 6 benign lesions where NMEL was found **(Fig 1).** Invasive ductal carcinoma and ductal carcinoma in situ, which are the most frequent pathologies in both, were determined to be the mass lesion in 63 malignant lesions and NMEL in only 12 lesions, respectively **(Fig 2).**

In our study, we discovered that there is no significant difference between Kaiser score, Kaiser BI-RADS, and MRI BI-RADS in the diagnosis of all suspicious and malignant breast lesions either mass lesion or NMEL as the p-value was 0.597, 0.84, and 0.495 correspondingly. However, we found that the Kaiser score is the most useful one because its AUC (for all lesions, mass lesions, and NMEL, respectively) was 0.985, 0.997, and 0.813, which is a little higher than theirs **(Table 3, Fig.2).** Additionally, because the p-value was less than 0.001, there was a significant inter-observer agreement between Readers 1 and 2 for the diagnosis of all suspicious and malignant breast lesions, whether they were mass lesions or NMEL, using both Kaiser BI-RADS and MRI BI-RADS **(Table 4, Fig.3).**

Additionally, we discovered that the MRI BI-RADS and Kaiser BI-RADS are both extremely good positive tests for the diagnosis of all breast lesions, whether mass lesion or NMEL, as they both had 100% sensitivity and a high +LR ratio (>1) by both readers **(Table 4, Fig.4).** However, both readers agree that both negative tests have low-LR ratios (1) and are less successful in terms of specificity. With 82.4 %, 91.2 %, 96.4 %, and 96.4 % respectively for all lesions and mass lesions, Kaiser BI-RADS was deemed more accurate than MRI BI-RADS, which had 58.8 %, 64.8 %, and 71.4 %, 78.6 % for readers 1 and 2, respectively. While Kaiser BI-RADS had a 16.7% and a 66.7 % specificity rate for the diagnosis of NMEL, MRI BI-RADS only demonstrated 100% specificity for both readers **(Table 4, Fig.4).**

TABLES

Table (1): Morphologic and kinetics criteria of Kaiser score.

Root sign	This means speculation is seen at the lesion margin and can be single or multiple speculations….And this denotes high suspicion of malignancy.
Margins	Smooth…Not suspicious. Irregular…Suspicious.
Washout curve	Increase signal in the early phase of enhancement then decreases in signal in delayed phases which means suspicious of malignancy.
Plateau curve	Increase signal in the early phase of enhancement then no further rise in signal in delayed phases and the lesion is considered equivocal.
Persistent enhancement	Progressive rise in signal in early and delayed phases of enhancement and lesion is mostly considered benign lesion.
A pattern of internal enhancement	Rim enhancement ….Suspicious Homogenous….Not suspicious
Edema	High signal on T2WI and STIR which may be surrounding the lesion or diffuse in breast parenchyma ….Suggestive of breast cancer

Table 2: Baseline characteristics of the studied sample

Variable	Category	n = 109
Mass	• **No**	18 (16.5%)
	• **Yes**	91 (83.5%)
MRI BI-RADS	• **II**	20 (18.3%)
	• **III**	8 (7.4%)
	• **IV**	20 (18.3%)
	• **V**	61 (56%)
Root Sign	• **Yes**	69 (63.3%)
Enhancement Curve	• **I**	30 (27.5%)
	• **II**	20 (18.3%)
	• **III**	59 (54.2%)
Border	• **Regular**	26 (23.9%)
	• **Irregular**	83 (76.1%)
Edema	• **Yes**	57 (52.3%)
Kaiser score	• **Mean ± SD**	7.48 ± 3.9
	• **Median (Range)**	9 (1 – 11)
Kaiser BI-RADS	• **II-III**	28 (25.7%)
	• **IV**	11 (10.1%)
	• **V**	70 (64.2%)
Parents' Consanguinity	• **Positive**	115 (65%)
Pathology	• **Benign**	34 (31.2%)
	• **Malignant**	75 (68.8%)

Received: 17/04/2022
Accepted: 14/06/2022

Table 3: ROC curve Comparison for both Lesions

Reading	All Lesions	Mass	NMEL
	AUC (95% CI)		
MRI BI-RADS ≥ 3	0.984 (0.966-1.000)	0.998 (0.993-1.000)	0.757 (0.531-0.983)
Kaiser Score ≥ 4	0.985 (0.969-1.000)	0.997 (0.991-1.000)	0.813 (0.601-1.000)
Kaiser BI-RADS ≥ 3	0.946 (0.887-1.000)	0.981 (0.939-1.000)	0.729 (0.462-0.996)
• P-value	= 0.597	= 0.824	= 0.495

Table 4: Lesion Diagnostic Criterion for MRI and Kaiser BI-RADS by the Two Readers

Criterion	Weighted Kappa	AUC (95% CI)	Sensitivity	Specificity	+ LR	- LR
All Lesions						
• MRI BI-RADS R1	0.941 (p < 0.001)	0.977 (0.955-0.999)	100% (75/75)	58.8% (20/34)	1.70	<0.01
• MRI BI-RADS R2		0.984 (0.966-1.000)	100% (75/75)	64.7% (22/34)	1.55	<0.01
• Kaiser BI-RADS R1	0.947 (p < 0.001)	0.946 (0.886-1.000)	100% (75/75)	82.4% (28/34)	1.21	<0.01
• Kaiser BI-RADS R2		0.946 (0.887-1.000)	100% (75/75)	91.2% (31/34)	1.10	<0.01
NMEL						
• MRI BI-RADS R1	0.772 (p < 0.001)	0.681 (0.431-0.930)	100% (12/12)	100% (6/6)	1	<0.01
• MRI BI-RADS R2		0.757 (0.531-0.983)	100% (12/12)	100% (6/6)	1	<0.01
• Kaiser BI-RADS R1	0.797 (p < 0.001)	0.722 (0.447-0.998)	100% (12/12)	16.7% (1/6)	5.89	<0.01
• Kaiser BI-RADS R2		0.729 (0.762-0.996)	100% (12/12)	66.7% (4/6)	1.50	<0.01
Mass Lesion						
• MRI BI-RADS R1	0.962 (p < 0.001)	0.997 (0.991-1.000)	100% (63/63)	71.4% (20/28)	1.35	<0.01
• MRI BI-RADS R2		0.998 (0.993-1.000)	100% (63/63)	78.6% (22/28)	1.27	<0.01
• Kaiser BI-RADS R1	0.987 (p < 0.001)	0.980 (0.938-1.000)	100% (63/63)	96.4% (27/28)	1.04	<0.01
• Kaiser BI-RADS R2		0.981 (0.939-1.000)	100% (63/63)	96.4% (27/28)	1.04	<0.01

FIGURES

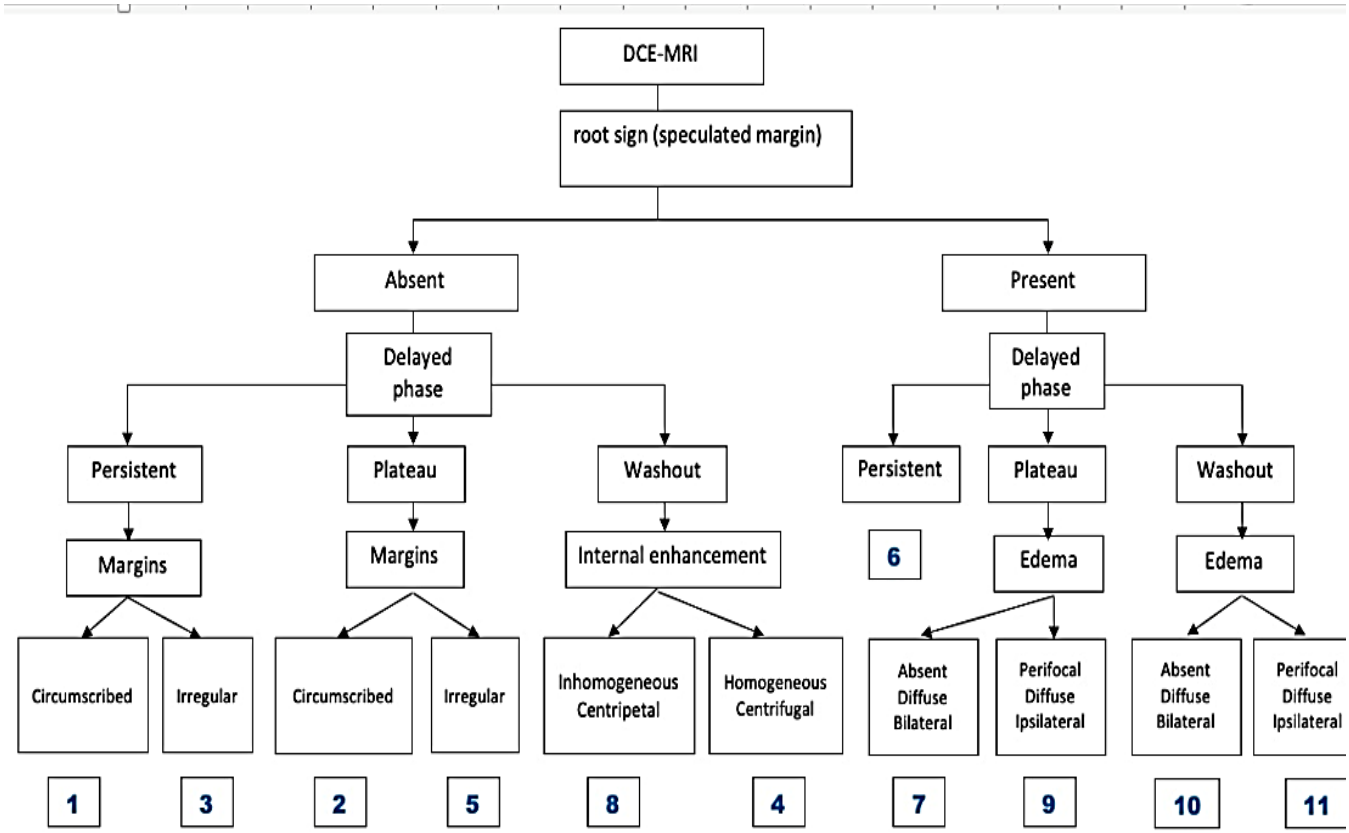

Figure 1: Kaiser score flow chart illustration.

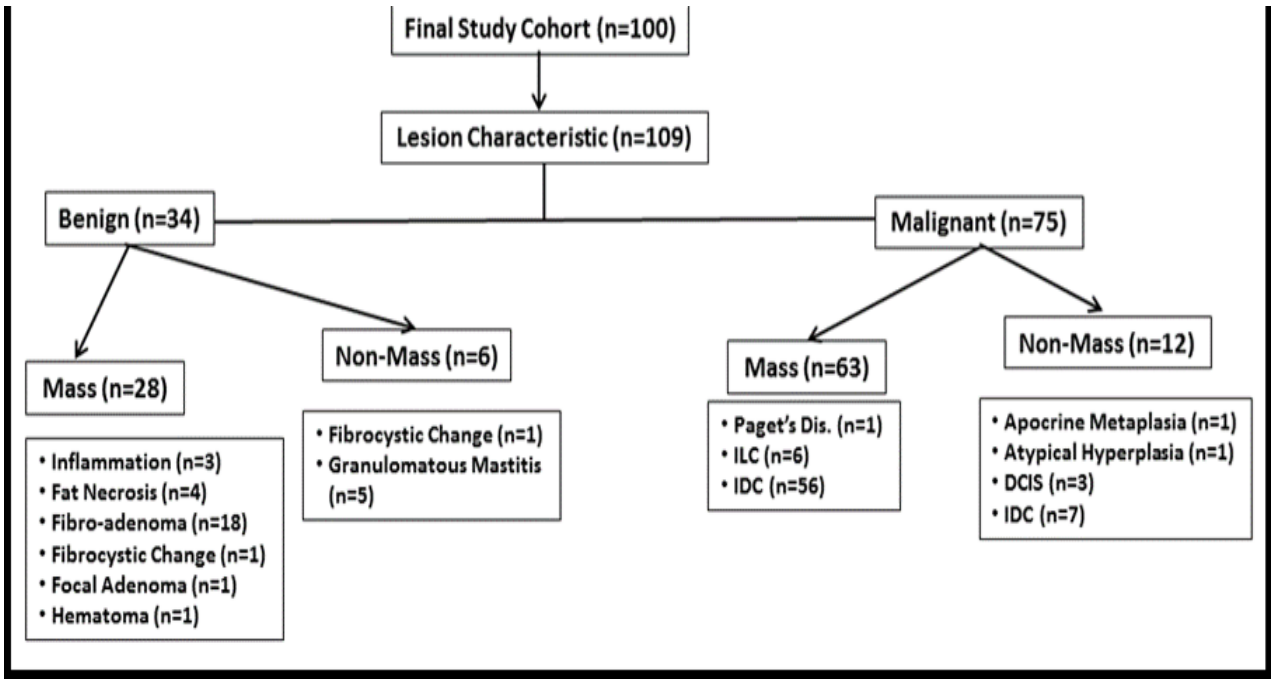

Figure 2: Flow Chart of the studied cohort.

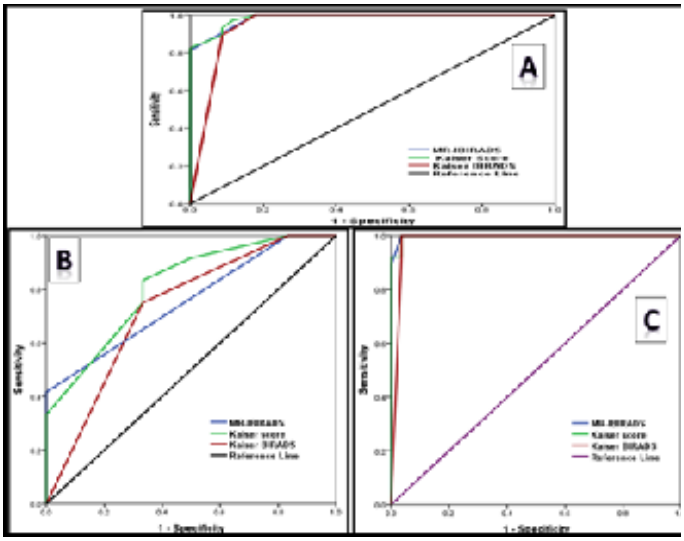

Figure 3: ROC curve for MRI BI-RADS, Kaiser Score, and Kaiser BI-RADS: A: All Lesions B: NMEL and C: Mass.

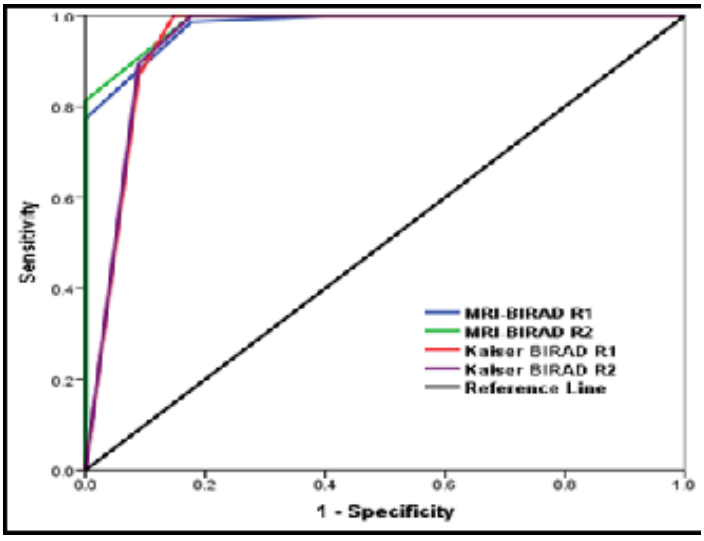

Figure 4: ROC curve for MRI BI-RADS and Kaiser BI-RADS for the Two Readers.

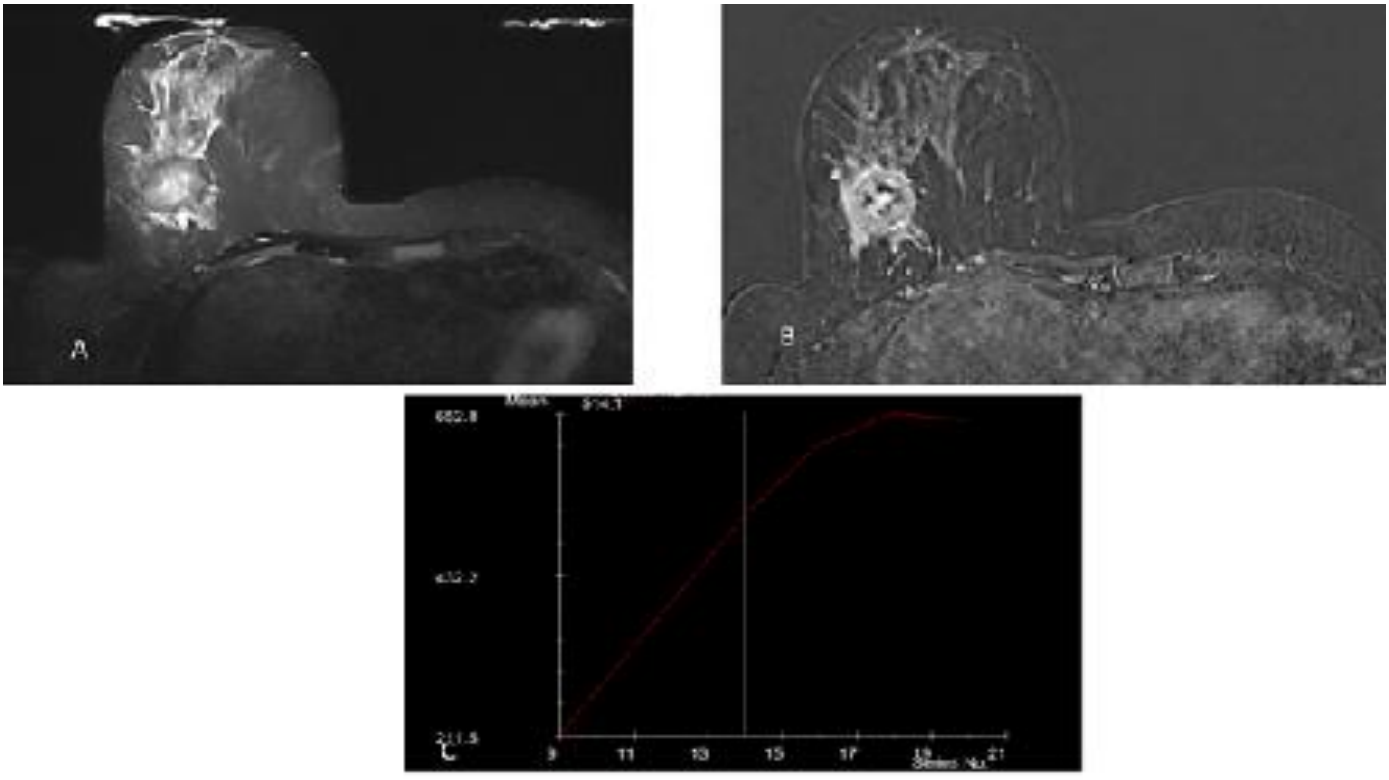

Figure 5: Female patient with a right breast mass. **(A):** The mass shows root-sign and surrounding perifocal edema, **(B):** With heterogeneous Post-GD enhancement. **(C):** The mass shows the washout curve (type III) in dynamic sequences. This is consistent with Kaiser score 11, Kaiser BI-RADS 5, and MRI BI-RADS 5 for both readers, and it was proved histopathologically to be invasive duct carcinoma.

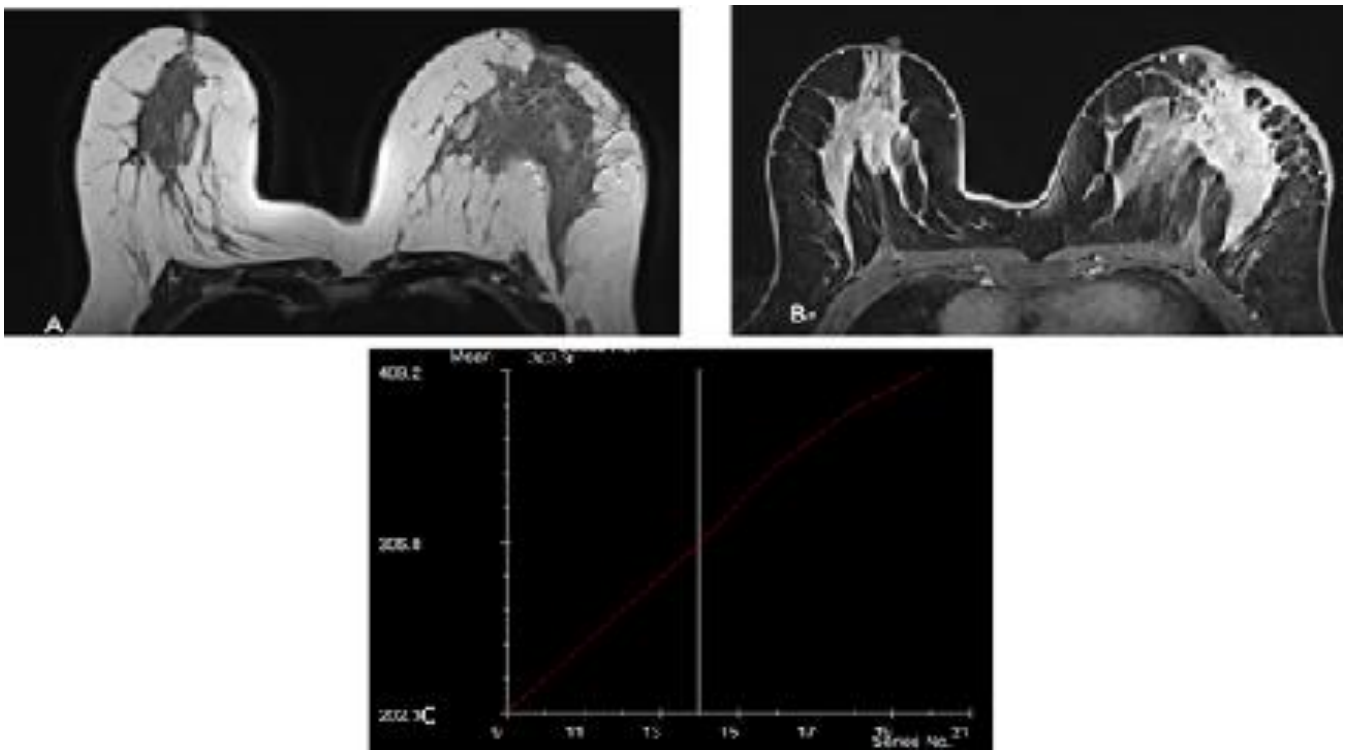

Figure 6: Non-mass lesion within the left breast seen at the retro-areolar region. **(A):** the lesion shows an irregular outline with no surrounding edema, **(B):** With heterogeneous Post-GD enhancement. **(C):** This lesion shows a persistent increase enhancement (type I) curve in dynamic sequences. This was consistent with Kaiser score 6, Kaiser BI-RADS 4, and assigned as MRI BI-RADS 5 by one reader and BI-RADS 4 by the other reader, and it was proved histopathologically to be granulomatous mastitis.

DISCUSSION

Mammography is a two-dimensional technique, therefore it has low sensitivity and specificity in identifying various breast lesions. Additionally, operator-dependent picture collection and overlapped glandular tissue shadows limit its ability to diagnose breast abnormalities. Additionally, there may be a gap in the clinical translation of suspect lesions that may be challenging to find on additional views or targets when employing stereotactic techniques. Due to its higher sensitivity as a crucial diagnostic tool for various breast lesions and also for screening in high-risk patients, Breast MRI has been developed as a problem-solving method for these dubious lesions in modern times [12,13].

In this study, we investigated the Kaiser score's diagnostic capabilities and inter-reader consistency as a scoring system for breast MRI.

As the Kaiser score uses small and simple morphologic and dynamic features (root sign or speculation, perilesional edema, lesion margin, enhancement kinetics, and pattern of enhancement), it may help to consolidate diagnosis and formulate management. It also simplifies the process of lesion interpretation and leads the radiologist step-by-step toward the final diagnosis of the breast lesion [9].

The Kaiser score was determined to be the most effective one in this study because its AUC was 0.985, 0.997, and 0.813 (for all lesions, mass, and NMEL respectively) which is slightly higher compared to theirs. However, there was no significant difference in accuracy between the Kaiser score, Kaiser BI-RADS, and MRI BI-RADS in the diagnosis of all suspicious and malignant breast lesions. This closely resembles the findings of the study of Ruxandra Iulia Milos et al [8].

In terms of the inter-observer agreement, we discovered that there was a strong inter-observer agreement between the two readers for Kaiser BI-RADS in the diagnosis of all suspicious and malignant breast lesions, whether they were mass lesions or NMEL, as the p-value was 0.001. These findings are nearly identical to those of Maria Adele Marino, et al. study's [9], which discovered that Kaiser BI-RADS and score improved inter-reader agreement.

Regarding sensitivity and specificity, we discovered that both the MRI BI-RADS and the Kaiser BI-RADS are very good positive tests for the diagnosis of breast lesions, either mass or NMEL, with 100% sensitivity and a high +LR ratio (>1). However, the Kaiser BI-RADS was found to be more specific in the diagnosis of the mass lesion by 91.2 percent and 96.4 percent for readers 1 and 2

respectively, while the MRI BI-RADS showed more specificity

We highlighted that Kaiser BI-RADS may not be able to identify the margin type and may have trouble identifying the pattern of enhancement in smaller non-mass lesions, which may account for its reduced specificity in the identification of non-mass-enhancing lesions.

Additionally, the Kaiser score application does not require post-processing software, MR spectroscopy, or diffusion sequences because it uses normal MRI procedures and does not require any additional sequences [(14)]. Additionally, it is simple to apply to all breast lesions, regardless of whether they are mass, non-mass, or focused [(15)], and adding this score to the standard MRI BI-RADS, aids in the exclusion of unnecessary biopsies, which lowers healthcare costs, patient discomfort, and the adverse effects of invasive procedures [(14,16)].

CONCLUSION

Finally, in this study, we concluded that the Kaiser score, which achieves high diagnostic accuracy comparable to that of MRI BI-RADS, can be used in conjugation with MRI BI-RADS to improve inter-reader agreement and decrease experience-related variability. Additionally, it can be quite helpful in making the diagnosis of doubtful and suspicious breast lesions for readers with less experience.

Acknowledgments: Not applicable.
Financial support and sponsorship: Nil.
Conflict of interest: Nil.

List of abbreviations:

MRI: Magnetic resonance imaging.
ACR: American college of radiology.
BIRADS: Breast imaging reporting and data system.

REFERENCES

1. **Warner E, Messersmith H, Causer P *et al.* (2008):** Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. Ann Intern Med., 148(9):671-9.
2. **Sarica O, Uluc F (2014):** Additional diagnostic value of MRI in patients with suspicious breast lesions based on ultrasound. Br J Radiol., 87(1041):20140009.
3. **Dietzel M, Baltzer P (2018):** How to use the Kaiser score as a clinical decision rule for diagnosis in multiparametric breast MRI: a pictorial essay. Insights Imaging, 9(3):325-35.
4. **Houssami N, Turner R, Morrow M (2013)**: Preoperative magnetic resonance imaging in breast cancer: a meta-analysis of surgical outcomes. Ann Surg., 257(2):249-55.
5. **Marino M, Riedl C, Bernathova M *et al.* (2018):** Imaging Phenotypes in Women at High Risk for Breast Cancer on Mammography, Ultrasound, and Magnetic Resonance Imaging Using the Fifth Edition of the Breast Imaging Reporting and Data System. Eur J Radiol., 106:150-9.
6. **Benndorf M, Baltzer P, Kaiser W (2011):** Assessing the degree of collinearity among the lesion features of the MRI BI-RADS lexicon. Eur J Radiol., 80(3):e322-4.
7. **Baltzer P, Dietzel M, Kaiser W (2013):** A simple and robust classification tree for differentiation between benign and malignant lesions in MR-mammography. Eur Radiol., 23(8):2051-60.
8. **Milos R, Pipan F, Kalovidouri A *et al.* (2020):** The Kaiser score reliably excludes malignancy in benign contrast-enhancing lesions classified as BI-RADS 4 on breast MRI high-risk screening exams. Eur Radiol., 30(11):6052-61.
9. **Marino M, Clauser P, Woitek R *et al.* (2016):** A simple scoring system for breast MRI interpretation: does it compensate for reader experience? Eur Radiol.,26(8):2529-37.
10. **Pinker K, Helbich T, Morris E (2017):** The potential of multiparametric MRI of the breast. Br J Radiol., 90(1069):20160715.
11. **Magny S, Shikhman R, Keppke A (2021):** Breast Imaging Reporting and Data System. StatPearls. Treasure Island (FL).
12. **Mann R, Balleyguier C, Baltzer P *et al.* (2015):** Breast MRI: EUSOBI recommendations for women's information. Eur Radiol., 25(12):3669-78.
13. **Sardanelli F, Boetes C, Borisch B *et al.* (2010)**: Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. Eur J Cancer, 46(8):1296-316.
14. **Woitek R, Spick C, Schernthaner M *et al.* (2017)**: A simple classification system (the Tree flowchart) for breast MRI can reduce the number of unnecessary biopsies in MRI-only lesions. Eur Radiol., 27(9):3799-809.
15. **Wengert G, Pipan F, Almohanna J *et al.* (2020):** Impact of the Kaiser score on clinical decision-making in BI-RADS 4 mammographic calcifications examined with breast MRI. Eur Radiol., 30(3):1451-9.
16. **Kim S, Morris E, Liberman L *et al.* (2001):** Observer variability and applicability of BI-RADS terminology for breast MR imaging: invasive carcinomas as focal masses. AJR Am J Roentgenol., 177(3):551-7.