



Breast MRI: does a clinical decision algorithm outweigh reader experience?

Nina Pötsch¹ · Aida Korajac¹ · Philipp Stelzer¹ · Panagiotis Kapetas¹ · Ruxandra-Iulia Milos¹ · Matthias Dietzel² · Thomas H. Helbich¹ · Paola Clauser¹ · Pascal A. T. Baltzer¹ 

Received: 11 February 2022 / Revised: 30 May 2022 / Accepted: 2 July 2022 / Published online: 19 July 2022
© The Author(s) 2022

Abstract

Objectives Due to its high sensitivity, DCE MRI of the breast (MRIb) is increasingly used for both screening and assessment purposes. The Kaiser score (KS) is a clinical decision algorithm, which formalizes and guides diagnosis in breast MRI and is expected to compensate for lesser reader experience. The aim was to evaluate the diagnostic performance of untrained residents using the KS compared to off-site radiologists experienced in breast imaging using only MR BI-RADS.

Methods Three off-site, board-certified radiologists, experienced in breast imaging, interpreted MRIb according to the MR BI-RADS scale. The same studies were read by three residents in radiology without prior training in breast imaging using the KS. All readers were blinded to clinical information. Histology was used as the gold standard. Statistical analysis was conducted by comparing the AUC of the ROC curves.

Results A total of 80 women (median age 52 years) with 93 lesions (32 benign, 61 malignant) were included. The individual within-group performance of the three expert readers (AUC 0.723–0.742) as well as the three residents was equal (AUC 0.842–0.928), $p > 0.05$, respectively. But, the rating of each resident using the KS significantly outperformed the experts' ratings using the MR BI-RADS scale ($p \leq 0.05$).

Conclusion The KS helped residents to achieve better results in reaching correct diagnoses than experienced radiologists empirically assigning MR BI-RADS categories in a clinical “problem solving MRI” setting. These results support that reporting breast MRI benefits more from using a diagnostic algorithm rather than expert experience.

Key Points

- Reporting breast MRI benefits more from using a diagnostic algorithm rather than expert experience in a clinical “problem solving MRI” setting.
- The Kaiser score, which provides a clinical decision algorithm for structured reporting, helps residents to reach an expert level in breast MRI reporting and to even outperform experienced radiologists using MR BI-RADS without further formal guidance.

Keywords Algorithms · MRI · BI-RADS · Breast neoplasms · Radiology

Abbreviations

ACR	American College of Radiology
BI-RADS	Breast Imaging Reporting and Data System
EUSOBI	European Society of Breast Imaging

KS	Kaiser score
MRIb	Breast MRI

Introduction

Contrast-enhanced magnetic resonance imaging of the breast (MRIb) is well established in breast imaging for various indications. Standard applications include screening in high-risk patients, diagnostic work-up of equivocal lesions in mammography and ultrasound, or treatment planning and monitoring in breast cancer patients [1, 2]. Moreover, recent data are promising regarding MRIb for intermediate risk screening in women with extremely dense breasts [3–5]. Therefore, an increase

✉ Pascal A. T. Baltzer
pascal.baltzer@meduniwien.ac.at

¹ Division of General and Pediatric Radiology, Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna and General Hospital, Waehringer Guertel 18-20, A-1090 Vienna, Austria

² Institute of Radiology, Erlangen University Hospital, Maximiliansplatz 2, 91054 Erlangen, Germany

in the demand for MRIB examinations is envisioned in the near future.

Interpretation of MRIB is considered a task for well-trained radiologists. Since the interpretation of MRIB is challenging and inter-reader agreement is variable [6, 7], improvement in standardized image interpretation is an active research topic [1, 8–14]. The Breast Imaging Reporting and Data System (BI-RADS) lexicon of the American College of Radiology (ACR) in the current 5th edition undoubtedly serves as the most widely accepted standard for reporting MRIB findings [10]. Nevertheless, inter-observer variability and diagnostic accuracy vary using the standardized BI-RADS lexicon [15–19]. This may be explained by the fact that the BI-RADS lexicon does not include a clinical decision algorithm that formally combines diagnostic criteria for structured reporting of MRIB [20].

The Kaiser score (KS) provides such clinical decision rules that uses machine learning methodology to combine five independent MR BI-RADS criteria: (i) spiculations; (ii) SI-time curve type; (iii) margins of the lesion; (iv) internal enhancement; and (v) presence of edema. These criteria are weighted differently in their combination and result in a score from 1 “lowest risk of breast cancer” to 11 “highest risk of breast cancer.” This score can then be directly translated into a BI-RADS category rating (KS 1–4 = BI-RADS 2/3, KS 5–7 = BI-RADS 4, KS 8–11 = BI-RADS 5) with corresponding clinical implications [21]. Thereby, the KS has shown to facilitate the characterization of enhancing breast lesions, potentially avoiding misdiagnosis and unnecessary biopsies [22–24]. In addition, initial results have suggested that the KS compensates for reader experience and improves inter-reader agreement [25]. To investigate the influence of reader experience, however, independent, off-site readers experienced in breast MRI, but not previously exposed to the KS, should be compared to readers inexperienced in breast MRI who simply apply the KS as a diagnostic algorithm in the same patients.

Therefore, the aim of our study was to evaluate the diagnostic performance of residents using the KS for structured MRIB reporting compared to radiologists with experience in breast imaging using only MR BI-RADS.

Methods

Study design

The local ethics committee approved this retrospective, single-center study, waiving the need for informed consent (IS 1917/2016). The patient-related data were de-identified and handled in accordance with standards of good scientific practice. All data generated and analyzed during this study are available from the corresponding author by request. The same

dataset has already been described in a previous analysis with a completely different study design and aims [26].

Patients

We included a subset of patients, prospectively recruited in the context of another study [26]. The inclusion criteria were defined as follows: female patients > 21 years undergoing MRIB for inconclusive or suspicious findings (BI-RADS 0, 4, and 5) on conventional imaging (mammography/tomosynthesis, ultrasound, clinical examination) [27]. Pregnant or lactating women, as well as women with breast implants or a history of breast cancer and related treatment, were excluded. Moreover, women unable to give written, informed consent or with contraindications for MRI or gadolinium-based contrast agents were excluded. Histopathology, acquired through stereotactic, US-, or MRI-guided biopsy, was used as the standard of reference. Patients not undergoing biopsy, due to distinct benign MRI findings that did not require biopsy, were excluded from further analysis.

MRI acquisition

The MRIB studies included were mainly performed at our institution ($n = 48$, 60%) or at outside facilities ($n = 32$, 40%) using either 1.5-T or 3-T scanners, with dedicated eight-channel (at least) breast coils. All images were acquired with the patient in a prone position using a standard protocol in line with European Society of Breast Imaging (EUSOBI) recommendations [27], including, at least, a T2-weighted sequence and a gradient echo T1-weighted sequence, prior to and after i.v. injection of a gadolinium-based contrast agent, in the axial plane. A minimum of three T1-weighted post-contrast sequences was acquired. For improved evaluation of contrast enhancement, subtracted images were available. Details on the acquired series are given as Supplemental Material (Supplemental Table 1).

Image interpretation

Image interpretation was performed by three, independent, off-site radiologists with more than 15 years of experience in MRIB. Comparable to clinical routine, the expert readers rated the following: (i) presence and location of a lesion (breast quadrant); (ii) lesion type (mass or non-mass enhancement); and (iii) lesion size (mm); and assigned an MR BI-RADS [10] score for each lesion. The readers were instructed to rate only the most suspicious lesion ($n = 1$) per breast.

The same studies were analyzed independently by three, off-site radiology residents using the Kaiser score (KS), prior to their basic (ESR level I) rotation in “breast radiology,” with

little to no experience in MRIb. Briefly, the KS is a simple classification system based on a total of five kinetic and morphologic criteria (I: contrast enhancement kinetics, II: internal enhancement, III: presence or absence of spiculations, IV: lesion margins, and V: edema) resulting in assignment categories from 1 (= lowest risk of breast cancer) to 11 (= highest risk of breast cancer). Generally, a KS rating higher than 4 corresponds to a BI-RADS 4/5 and is, therefore, considered suspicious with a consequent indication for biopsy [21]. The residents used the publicly available KS online tool (accessible via: <http://www.meduniwien.ac.at/kaiser-score/>), which queries the individual criteria one by one and automatically provides a KS rating and the corresponding BI-RADS category. Residents were aware of the lesion's location (side and quadrant) and size, but were blinded to the experts' MR BI-RADS rating to ensure the exact same lesion was rated by experts and trainees. All readers were blinded to all clinical information and patient history. No prior imaging studies were available to the readers.

Statistical analysis

Statistical analysis was done using SPSS 22 for MAC (IBM). Relevant clinical and demographic data are presented by descriptive analyses. Metric continuous data are given as median and ranges. The individual diagnostic performance differentiating benign and malignant breast lesion was assessed by comparing the AUC of the ROC curves. Inter-reader agreement for expert and resident readers was evaluated using Cohen's kappa. A two-sided p value of $p < 0.05$ was considered significant.

Results

A total of 80 women (median age 52 years, range 34–83 years) with 93 lesions, 32 benign and 61 malignant, were included. Twenty lesions were rated as non-mass lesions, 13 of which were malignant. Median lesion size was 18 mm (range 4–80 mm). Histologic details are given in Table 1.

Using MR BI-RADS only, the performance between the three expert readers (AUC 0.723–0.742, $p > 0.72$) did not differ significantly. Using the KS, the difference in the performance between the three residents (AUC 0.842–0.928, $p = 0.03$, $p > 0.08$, respectively; Table 2) varied between significance and borderline significance. The rating of each single resident using the KS significantly outperformed the experts' ratings using the MR BI-RADS scale ($p < 0.01$ – 0.05 ; Fig. 1). Moreover, we found a tendency toward better inter-reader agreement using the KS for structured reporting compared to MR BI-RADS (KS: kappa 0.579–0.710, MR BI-RADS: kappa 0.531–0.624). Exemplary images from the study

Table 1 Histologic details of the 93 lesions included in this analysis

Histology	Number (%)
Benign	32
Fibrocystic changes	15 (47%)
Papilloma	6 (19%)
Inflammatory changes	6 (19%)
Fibroadenoma	5 (15%)
Malignant	61
Invasive carcinoma NST	19 (31%)
Invasive lobular carcinoma	2 (3)
Invasive carcinoma with DCIS	25 (41%)
DCIS	15 (25%)

NST, nonspecific type; DCIS, ductal carcinoma in situ

population are shown in Figs. 2 and 3. Figure 4 shows concordant findings in expert and resident readers depicted against ground truth.

Discussion

Our study shows that the KS, which provides a clinical decision algorithm for structured reporting, helps residents to reach an expert level in MRIb reporting. Residents who applied the KS for MRIb reading achieved even better results than expert readers using only MR BI-RADS. Moreover, there was also a tendency for better inter-reader agreement using the KS. These results support that successful structured reporting of MRIb using the KS does not require expert training.

Prior studies have demonstrated and validated the diagnostic value of the KS in several clinical MRIb scenarios [22–25, 28–30]. It is applicable in high-risk, intermediate-risk, and average-risk patients alike; can be used for breast cancer diagnosis independent of mammographic appearance; and has a substantial potential to reduce unnecessary biopsies in a broad variety of indications [22–25, 28, 29]. The conceivable impact on inexperienced readers has been suggested previously [22, 25]. Marino et al demonstrated better reader performance for a

Table 2 AUC values for each individual reader

	AUC	95% CI
Expert 1	0.742	0.641–0.827
Expert 2	0.727	0.625–0.814
Expert 3	0.723	0.621–0.811
Resident 1	0.928	0.856–0.971
Resident 2	0.842	0.752–0.909
Resident 3	0.875	0.790–0.935

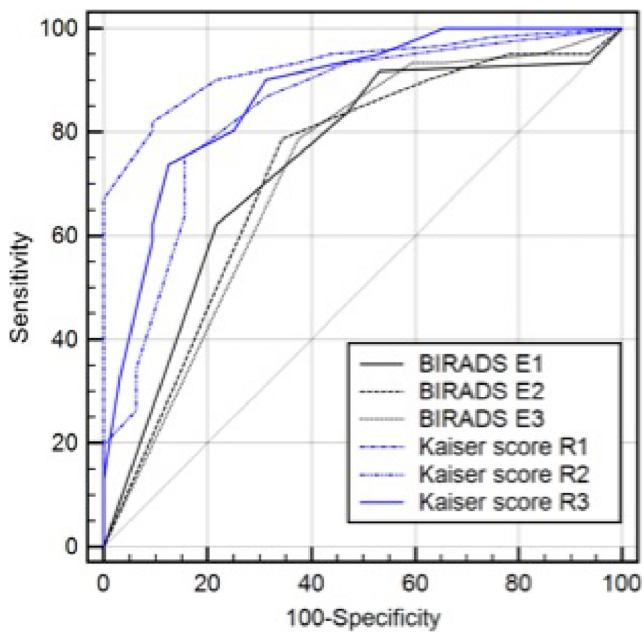


Fig. 1 ROC curves demonstrating the performance of the residents (depicted in blue) and of the experts (depicted in black)

less experienced reader in an intra-individual comparison study using MR BI-RADS with and without the KS. Overall, they reported that using empirical MR BI-RADS interpretation led to significant differences between readers based on their individual level of experience, while using the KS did not. Our findings go beyond this, as they do not only corroborate the suggestion that the application of the KS could compensate for reader experience, but show that inexperienced readers outperform even experts who did not use algorithmic guidance for structured reporting. The independent reader approach used in this study is more valid, as the prior study allowed the same readers to use the KS after the initial read without any attempt to alleviate a possible recognition bias by a randomized reading approach. In addition, the readers were trained in using the KS before the beginning of the study, and thus, may have subconsciously been influenced by the KS. Our approach avoided such a potential influence and the experienced readers were neither specifically exposed to nor did they apply the KS in their routine clinical practice. And, our results show that MRIb interpretation, relying on the

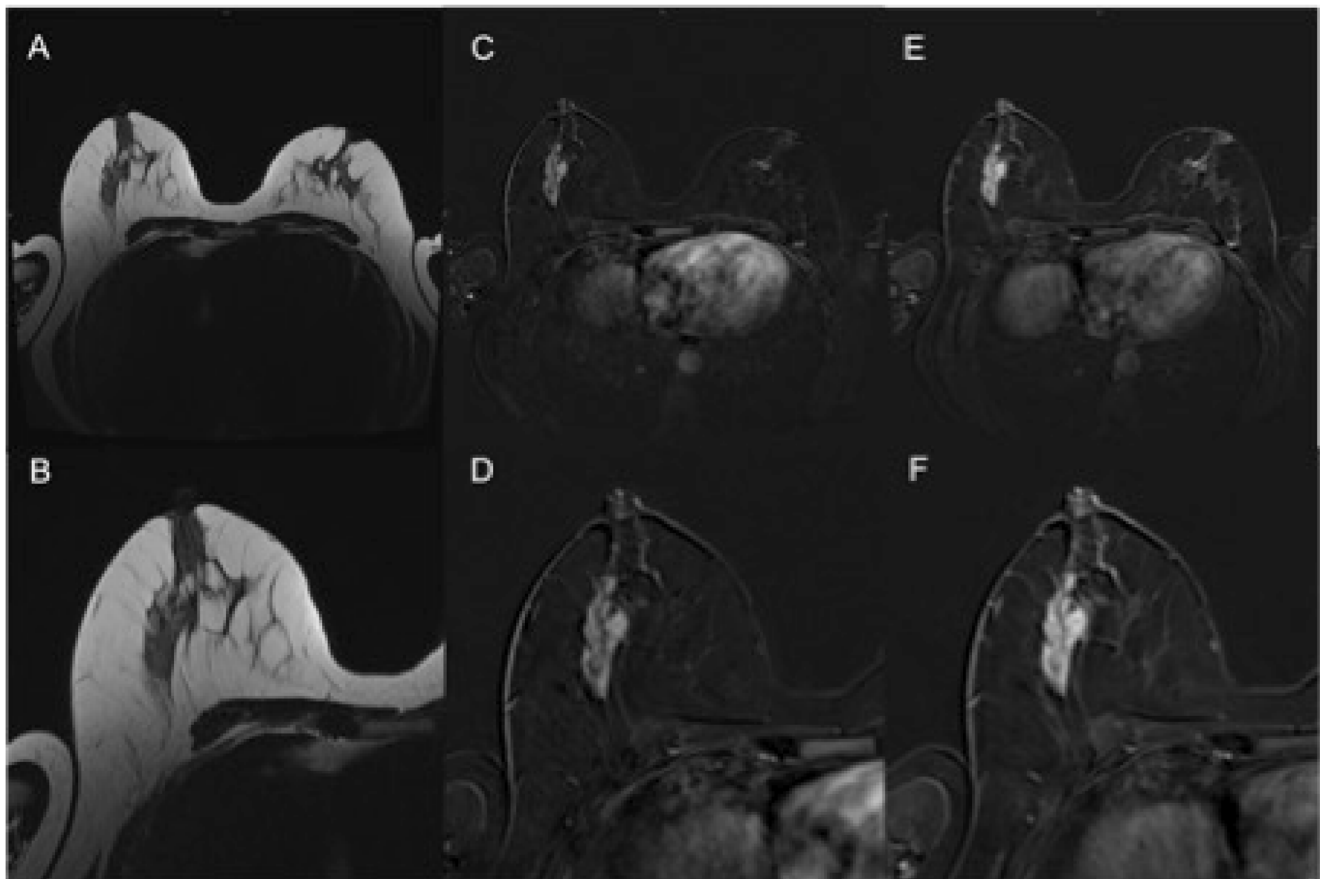


Fig. 2 B2 lesion rated BI-RADS 4/5 by experts and benign using the KS. T2 TSE (A), early (C) and late (E) post-contrast T1 with fat suppression, as well as corresponding magnifications of the right breast (B, D, F). This lesion was rated as MR BI-RADS 4 or 5 by the experts. Applying the Kaiser score to this mass lesion without spiculations, persistent

enhancement, and irregular margins gives a Kaiser score of 3, which is a benign result corresponding to a BI-RADS 2/3. Histologic verification showed tumor-free mammary parenchyma with low-grade pseudoangiomatous stromal changes, B2

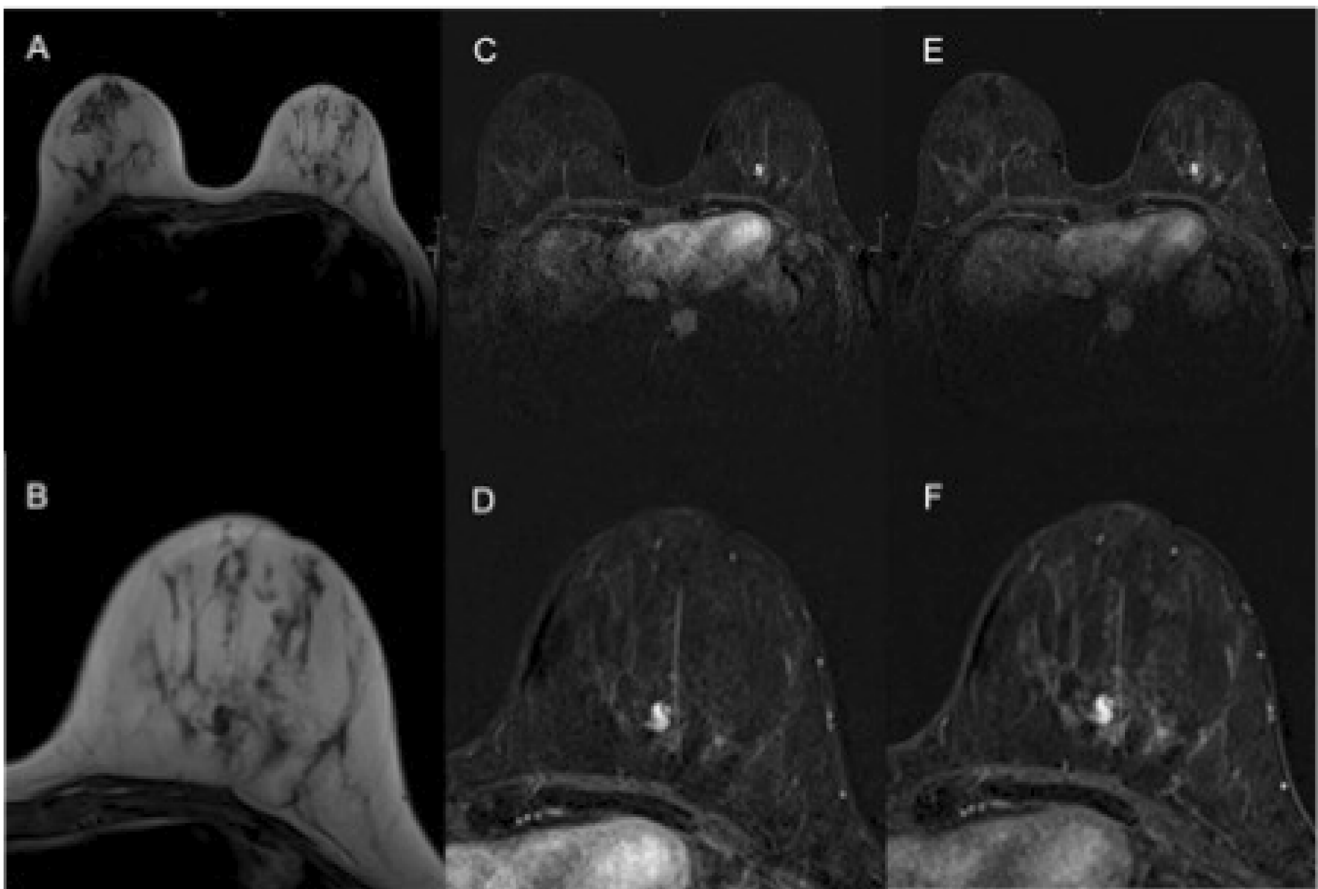


Fig. 3 Fibroadenoma rated BI-RADS 4/5 by experts and benign using the KS. T2 TSE (A), early (C) and late (E) post-contrast T1 with fat suppression, as well as corresponding magnifications of the left breast (B, D, F). This lesion was rated as MR BI-RADS 4 or 5 by the experts. Applying the

Kaiser score to this mass lesion without spiculations, plateau enhancement, and circumscribed margins gives a Kaiser score of 2, which is a benign result corresponding to a BI-RADS 2/3. Histologic verification showed a fibroadenoma, B2

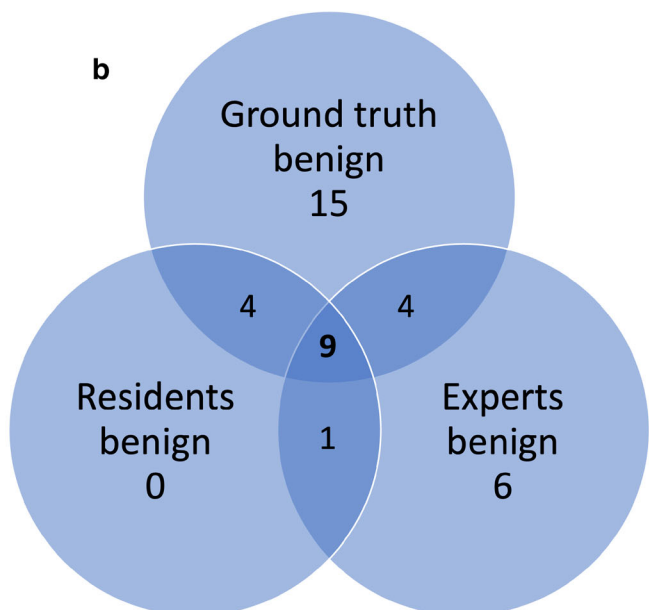
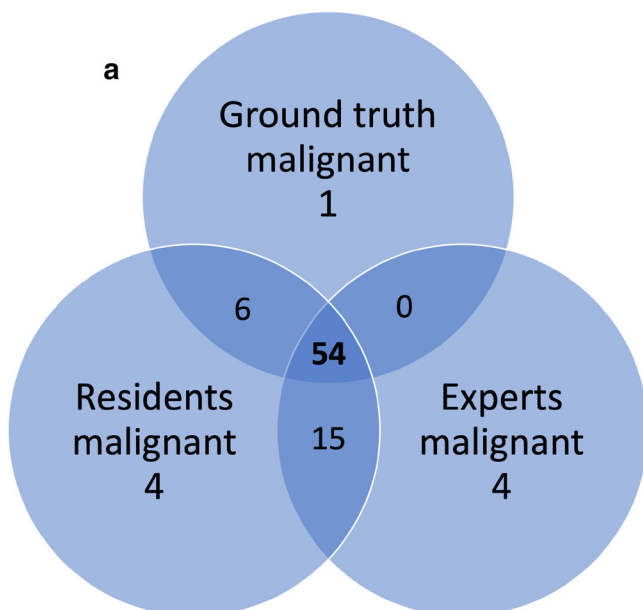


Fig. 4 Venn diagram depicting concordant findings in expert and resident readers together with histologic ground truth for (a) malignant lesions and (b) benign lesions. For the Kaiser score, an exploratory cut-off was set at a

value of 3. Thus, lesions with a rating < 3 were considered benign and lesions with a rating ≥ 3 were considered malignant

KS as a decision algorithm, seemingly does not require expert training. This is important considering the increasing demand on MRIb examinations because of the medical and economic evidence in favor of extended MRIb indications [3, 31–36].

Image interpretation in MRIb is considered a task for experienced and specifically trained radiologists [9]. Why? The main reason is complexity: a number of different parameters in various MRI sequences have to be considered and need to be combined in a final BI-RADS rating, which is not based on a formal decision rule, but—rather annoyingly put—on gut feeling. In addition, the physiological background enhancement, which varies inter-individually, might lead to diagnostic difficulties [37]. These factors might contribute to the variability of inter-reader agreement in MRIb image interpretation [6, 7]. For instance, Grimm et al [17] described only a limited inter-reader agreement between experts who reviewed BI-RADS 3 lesions using the current 5th edition of the MR BI-RADS lexicon [10]. Complexity and lack of formal guidance for structured reporting within the diagnostic process, with an unclear weighting of the individual criteria, may also explain why some authors describe comparable diagnostic performance metrics between MRIb with full compared to abbreviated protocols [38]. The KS, as a clinical decision rule based on predefined and statistically weighted MR BI-RADS criteria, does address these needs [22]. In line with this reasoning, several studies have demonstrated that if the Kaiser score as a clinical decision rule is used for breast MRI interpretation, the additional value of DWI is negligible [24, 30, 39].

Our study has some limitations, which should be addressed. First of all, initial data were collected in the framework of another study, including only patients with a final BI-RADS category 4 or 5 assessment who received invasive management [26]. These selection criteria imply a study population enriched with clinically challenging cases. This resulted in a correspondingly higher proportion of carcinomas, leading to a possible overestimation of sensitivity and underestimation of specificity. Therefore, the results strictly apply to the investigated setting, and the impact of the KS may be less pronounced in, e.g., a screening population. In addition, the inexperienced readers classified only indicated lesions and did not perform a detection task. Even though lesion detection in breast MRI is not considered a problem in the literature, that study did not answer whether high experience is required to identify lesions upfront [8, 20, 38]. However, the results are rather striking, as all the aforementioned potential biases would instead be in favor of a higher diagnostic performance for the experienced radiologists.

In conclusion, structured MRIb reporting, using the KS to reach an objective diagnostic category, helped residents to achieve better results than experienced radiologists who empirically assigned MR BI-RADS categories in a clinical “problem solving MRI” setting. These results support that

reporting breast MRI benefits more from using a diagnostic algorithm rather than expert experience.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00330-022-09015-8>.

Funding Open access funding provided by Medical University of Vienna. The authors state that this work has not received any funding.

Declarations

Guarantor The scientific guarantor of this publication is Pascal A.T. Baltzer.

Conflict of interest N.P. And P.C. are Associate Editors of *European Radiology*. As such, they had no role in the review and decision process underlying this manuscript. The authors of this manuscript declare no other relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise. No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Study subjects or cohorts overlap Some study subjects or cohorts have been previously reported in Clauser et al *Magn. Reson. Imaging JMRI*. 2020;52(2):589–595.

Methodology

- retrospective
- diagnostic study
- multicenter study

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Sardanelli F, Boetes C, Borisch B et al (2010) Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. *Eur J Cancer* 46:1296–1316. <https://doi.org/10.1016/j.ejca.2010.02.015>
2. Clauser P, Mann R, Athanasiou A et al (2018) A survey by the European Society of Breast Imaging on the utilisation of breast

- MRI in clinical practice. *Eur Radiol* 28:1909–1918. <https://doi.org/10.1007/s00330-017-5121-4>
3. Bakker MF, de Lange SV, Pijnappel RM et al (2019) Supplemental MRI screening for women with extremely dense breast tissue. *N Engl J Med* 381:2091–2102. <https://doi.org/10.1056/NEJMoa1903986>
 4. Verburg E, van Gils C, Bakker M et al (2020) Computer-aided diagnosis in multiparametric magnetic resonance imaging screening of women with extremely dense breasts to reduce false-positive diagnoses. *Invest Radiol* 55:438–444. <https://doi.org/10.1097/RLI.0000000000000656>
 5. Comstock CE, Gatsonis C, Newstead GM et al (2020) Comparison of abbreviated breast MRI vs digital breast tomosynthesis for breast cancer detection among women with dense breasts undergoing screening. *JAMA* 323:746–756. <https://doi.org/10.1001/jama.2020.0572>
 6. Stoutjesdijk MJ, Fütterer JJ, Boetes C et al (2005) Variability in the description of morphologic and contrast enhancement characteristics of breast lesions on magnetic resonance imaging. *Invest Radiol* 40:355–362. <https://doi.org/10.1097/01.rli.0000163741.16718.3e>
 7. Kinkel K, Helbich TH, Esserman LJ et al (2000) Dynamic high-spatial-resolution MR imaging of suspicious breast lesions: diagnostic criteria and interobserver variability. *AJR Am J Roentgenol* 175:35–43. <https://doi.org/10.2214/ajr.175.1.1750035>
 8. Kuhl C (2007) The current status of breast MR imaging. Part I. Choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice. *Radiology* 244:356–378. <https://doi.org/10.1148/radiol.2442051620>
 9. Mann RM, Balleyguier C, Baltzer PA et al (2015) Breast MRI: EUSOBI recommendations for women’s information. *Eur Radiol* 25:3669–3678. <https://doi.org/10.1007/s00330-015-3807-z>
 10. Morris EA, Comstock C, Lee C et al (2013) ACR BI-RADS® magnetic resonance imaging. In: ACR BI-RADS® Atlas, Breast imaging reporting and data system, 5th edn. American College of Radiology, Reston
 11. Baum F, Fischer U, Vosschenrich R, Grabbe E (2002) Classification of hypervascularized lesions in CE MR imaging of the breast. *Eur Radiol* 12:1087–1092. <https://doi.org/10.1007/s00330-001-1213-1>
 12. Malich A, Fischer DR, Wurdinger S et al (2005) Potential MRI interpretation model: differentiation of benign from malignant breast masses. *AJR Am J Roentgenol* 185:964–970. <https://doi.org/10.2214/AJR.04.1073>
 13. Pinker K, Bickel H, Helbich TH et al (2013) Combined contrast-enhanced magnetic resonance and diffusion-weighted imaging reading adapted to the “Breast Imaging Reporting and Data System” for multiparametric 3-T imaging of breast lesions. *Eur Radiol* 23:1791–1802. <https://doi.org/10.1007/s00330-013-2771-8>
 14. Pinker K, Bogner W, Baltzer P et al (2014) Improved differentiation of benign and malignant breast tumors with multiparametric 18F-fluorodeoxyglucose positron emission tomography magnetic resonance imaging: a feasibility study. *Clin Cancer Res* 20:3540–3549. <https://doi.org/10.1158/1078-0432.CCR-13-2810>
 15. 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. <https://doi.org/10.1002/jmri.1127>
 16. Kim SJ, Morris EA, 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* 177:551–557. <https://doi.org/10.2214/ajr.177.3.1770551>
 17. Grimm LJ, Anderson AL, Baker JA et al (2015) Interobserver variability between breast imagers using the fifth edition of the BI-RADS MRI lexicon. *AJR Am J Roentgenol* 204:1120–1124. <https://doi.org/10.2214/AJR.14.13047>
 18. Lunkiewicz M, Forte S, Freiwald B et al (2020) Interobserver variability and likelihood of malignancy for fifth edition BI-RADS MRI descriptors in non-mass breast lesions. *Eur Radiol* 30:77–86. <https://doi.org/10.1007/s00330-019-06312-7>
 19. El Khoury M, Lalonde L, David J et al (2015) Breast Imaging Reporting and Data System (BI-RADS) lexicon for breast MRI: interobserver variability in the description and assignment of BI-RADS category. *Eur J Radiol* 84:71–76. <https://doi.org/10.1016/j.ejrad.2014.10.003>
 20. Dietzel M, Baltzer PAT (2018) How to use the Kaiser score as a clinical decision rule for diagnosis in multiparametric breast MRI: a pictorial essay. *Insights Imaging* 9:325–335. <https://doi.org/10.1007/s13244-018-0611-8>
 21. Baltzer PAT, Dietzel M, Kaiser WA (2013) A simple and robust classification tree for differentiation between benign and malignant lesions in MR-mammography. *Eur Radiol* 23:2051–2060. <https://doi.org/10.1007/s00330-013-2804-3>
 22. Wengert GJ, Pipan F, Almohanna J et al (2019) Impact of the Kaiser score on clinical decision-making in BI-RADS 4 mammographic calcifications examined with breast MRI. *Eur Radiol* 30:1451–1459. <https://doi.org/10.1007/s00330-019-06444-w>
 23. Milos RI, 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* 30:6052–6061. <https://doi.org/10.1007/s00330-020-06945-z>
 24. Dietzel M, Krug B, Clauser P et al (2021) A multicentric comparison of apparent diffusion coefficient mapping and the Kaiser score in the assessment of breast lesions. *Invest Radiol* 56:274–282. <https://doi.org/10.1097/RLI.0000000000000739>
 25. Marino MA, Clauser P, Woitek R et al (2016) A simple scoring system for breast MRI interpretation: does it compensate for reader experience? *Eur Radiol* 26:2529–2537. <https://doi.org/10.1007/s00330-015-4075-7>
 26. Clauser P, Baltzer PAT, Kapetas P et al (2020) Low-dose, contrast-enhanced mammography compared to contrast-enhanced breast MRI: a feasibility study. *J Magn Reson Imaging* 52:589–595. <https://doi.org/10.1002/jmri.27079>
 27. Mann RM, Kuhl CK, Kinkel K, Boetes C (2008) Breast MRI: guidelines from the European Society of Breast Imaging. *Eur Radiol* 18:1307–1318. <https://doi.org/10.1007/s00330-008-0863-7>
 28. Jajodia A, Sindhvani G, Pasricha S et al (2020) Application of the Kaiser score to increase diagnostic accuracy in equivocal lesions on diagnostic mammograms referred for MR mammography. *Eur J Radiol*:109413. <https://doi.org/10.1016/j.ejrad.2020.109413>
 29. Zhang B, Feng L, Wang L, et al (2020) Kaiser score for diagnosis of breast lesions presenting as non-mass enhancement on MRI. *Nan Fang Yi Ke Xue Xue Bao* 40:562–566. <https://doi.org/10.12122/j.issn.1673-4254.2020.04.18>
 30. An Y, Mao G, Ao W et al (2022) Can DWI provide additional value to Kaiser score in evaluation of breast lesions. *Eur Radiol*. <https://doi.org/10.1007/s00330-022-08674-x>
 31. Gommers JJ, Voogd AC, Broeders MJ et al (2021) Breast magnetic resonance imaging as a problem solving tool in women recalled at biennial screening mammography: a population-based study in the Netherlands. *Breast* 60:279–286. <https://doi.org/10.1016/j.breast.2021.11.014>
 32. Fueger BJ, Clauser P, Kapetas P et al (2021) Can supplementary contrast-enhanced MRI of the breast avoid needle biopsies in suspicious microcalcifications seen on mammography? A systematic review and meta-analysis. *Breast* 56:53–60. <https://doi.org/10.1016/j.breast.2021.02.002>
 33. Baltzer PAT (2021) Supplemental screening using breast MRI in women with mammographically dense breasts. *Eur J Radiol* 136:109513. <https://doi.org/10.1016/j.ejrad.2020.109513>
 34. Kaiser CG, Dietzel M, Vag T, Froelich MF (2020) Cost-effectiveness of MR-mammography vs. conventional mammography in screening patients at intermediate risk of breast cancer - a

- model-based economic evaluation. *Eur J Radiol*:109355. <https://doi.org/10.1016/j.ejrad.2020.109355>
35. Tollens F, Baltzer PAT, Dietzel M et al (2021) Cost-effectiveness of digital breast tomosynthesis vs. abbreviated breast MRI for screening women with intermediate risk of breast cancer-how low-cost must MRI be? *Cancers* 13:1241. <https://doi.org/10.3390/cancers13061241>
 36. Geuzinge HA, Bakker MF, Heijnsdijk EAM et al (2021) Cost-effectiveness of magnetic resonance imaging screening for women with extremely dense breast tissue. *J Natl Cancer Inst* 113:1476–1483. <https://doi.org/10.1093/jnci/djab119>
 37. Hambly NM, Liberman L, Dershaw DD et al (2011) Background parenchymal enhancement on baseline screening breast MRI: impact on biopsy rate and short-interval follow-up. *AJR Am J Roentgenol* 196:218–224. <https://doi.org/10.2214/AJR.10.4550>
 38. Kuhl CK, Schrading S, Strobel K et al (2014) Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection—a novel approach to breast cancer screening with MRI. *J Clin Oncol* 32:2304–2310. <https://doi.org/10.1200/JCO.2013.52.5386>
 39. Meng L, Zhao X, Lu L et al (2021) A comparative assessment of MR BI-RADS 4 breast lesions with Kaiser score and apparent diffusion coefficient value. *Front Oncol* 11:779642. <https://doi.org/10.3389/fonc.2021.779642>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.