Journal Pre-proof

A multiparametric classification system for lesions detected by breast magnetic resonance imaging

A. Istomin (Conceptualization) (Validation) (Formal analysis) (Investigation) (Data curation) (Writing - original draft) (Writing - review and editing) (Visualization), A. Masarwah (Formal analysis) (Data curation) (Writing - original draft) (Writing - review and editing) (Visualization), H. Okuma (Formal analysis) (Investigation) (Data curation) (Writing - original draft) (Writing - review and editing) (Visualization), A. Sutela (Conceptualization) (Methodology) (Investigation) (Writing - review and editing) (Visualization), R. Vanninen (Conceptualization) (Methodology) (Resources) (Supervision) (Project administration) (Funding acquisition) (Writing - review and editing) (Visualization), M. Sudah (Conceptualization) (Methodology) (Validation) (Investigation) (Resources) (Data curation) (Writing - original draft) (Writing - review and editing) (Supervision) (Project administration) (Visualization)

PII: S0720-048X(20)30511-8

DOI: https://doi.org/10.1016/j.ejrad.2020.109322

Reference: EURR 109322

To appear in: *European Journal of Radiology*

Received Date: 21 August 2020

Revised Date: 19 September 2020

Accepted Date: 24 September 2020

Please cite this article as: Istomin A, Masarwah A, Okuma H, Sutela A, Vanninen R, Sudah M, A multiparametric classification system for lesions detected by breast magnetic resonance imaging, *European Journal of Radiology* (2020), doi: https://doi.org/10.1016/j.ejrad.2020.109322

A Multiparametric classification system for lesions detected by breast magnetic resonance imaging

A. Istomin[1*], MD; A. Masarwah[1*], MD PhD; H. Okuma[1], MD PhD; A. Sutela [1,2], MD PhD; R. Vanninen [1,2,3], MD PhD; M. Sudah [1,2], MD PhD

[1] Kuopio University Hospital, Diagnostic Imaging Center, Department of Clinical Radiology, Puijonlaaksontie 2, P.O. Box 100, FI-70029 Kuopio, Finland.
[2] University of Eastern Finland, Cancer Center of Eastern Finland, Yliopistonranta 1, 70210 Kuopio, Finland.
[3] University of Eastern Finland, Institute of Clinical Medicine, School of Medicine, Yliopistonranta 1, 70210 Kuopio, Finland.

Correspondence to:
Mazen Sudah, MD, PhD
Department of Clinical Radiology, Breast Unit
Kuopio University Hospital
Puijonlaaksontie 2, P.O. Box 100, FI-70029 Kuopio, Finland.
E-mail: mazen.sudah@kuh.fi

*A. Istomin and A. Masarwah contributed equally to this work.

Highlights:

- A multiparemetric evaluation of suspicious lesions is needed in breast MRI
- Our proposed categorisation of findings is comparable to the BI-RADS recommendations
- The proposed system helps in avoiding unnecessary follow-ups or biopsies

ABSTRACT

Background

To validate a new categorisation scheme for suspicious breast lesions according to the well-defined Breast Imaging Reporting and Data System (BI-RADS) magnetic resonance imaging (MRI) lexicon descriptors, apparent diffusion coefficients (ADC), T2-weighted signal intensity (SI), and kinetic curve assessment categories.

Methods

The MRI descriptors and ADC were analysed in 697 lesions detected in 499 subjects. The descriptors were classified into Minor, Intermediate, and Major findings, and were divided into the BI-RADS subcategories 3, 4A, 4B, 4C, and 5 according to the number of descriptors. Positive predictive values (PPV) were calculated for each descriptor. The descriptors were then fitted into a multinomial logistic regression model to determine the odds ratio for a malignant diagnosis. The PPV were measured for the new categories and compared with the assigned PPV of the BI-RADS descriptors.

Results

The PPV for MRI descriptors ranged from 17.9% to 100%. Of the 697 lesions assessed, 19 (2.7%) were categorised as BI-RADS 3, 27 (3.9%) as 4A, 53 (7.6%) as 4B, 174 (25.0%) as 4C, and 424 (60.8%) as 5. None of the subjects in BI-RADS category 3 had a malignant diagnosis. The PPV for malignancy increased progressively with increasing BI-RADS category (4A, 11.1%; 4B, 28.3%; 4C, 64.4%; 5, 94.8%). All descriptor groups were significant in the logistic regression model.

Conclusions

This study shows that using BI-RADS MRI descriptors together with ADC and T2-weighted SI in a multiparametric classification system can yield an applicable categorisation of lesions with PPV values within the recommended ranges for BI-RADS categories.

Abbreviations:

ADC (Apparent diffusion coefficient), BI-RADS (Breast Imaging Reporting and Data System), CI (Confidence intervals), DWI (Diffusion-weighted imaging), EUSOMA (European Society of Breast Cancer Specialists working group), MRI (Magnetic resonance imaging), NPV (Negative predictive values), OR (Odds ratio), ROI (Regions of interest), PPV (Positive predictive values), SI (Signal intensity), US (Ultrasound).

Keywords:

Magnetic resonance imaging; Breast cancer; BI-RADS; Diffusion-weighted MRI; Multiparametric MRI.

Introduction

Mammography and ultrasound (US) are considered to be the first-line imaging modalities for breast cancer diagnosis, but magnetic resonance imaging (MRI) shows superior sensitivity as a diagnostic method and has become a valuable tool in clinical practice. However, the main disadvantage of MRI is its lower specificity and, therefore, any additional lesion detected by MRI that might change the treatment pathway should be histologically confirmed by US- or MRI-guided biopsy [1].

The Breast Imaging Reporting and Data System (BI-RADS) was established by the American College of Radiology to standardise the description of mammography, US, and MRI findings and reporting [2]. This system was also designed to help communication between clinicians and radiologists, and decrease the risk and complications of unnecessary biopsy. For mammography and US, the BI-RADS subcategories 4A–C have been used in routine clinical and

radiological practice for many years. The 5th edition of the BI-RADS lexicon includes seven assessment categories. Of note, there is a wide range in the risk malignancy (2%–95%) in lesions assigned to category 4 [2]. To date, no MRI subcategories have been implemented within the BI-RADS lexicon, mostly due to the scarcity of data.

Breast MRI comprises functional and morphological, quantitative, and qualitative parameters that are combined by the reader to provide a recommendation for the subsequent management of the patient [3, 4]. Previous studies have evaluated the positive predictive values (PPV) of individual BI-RADS MRI descriptors. While some descriptors, such as clumped, ring-enhancement or spiculated margins have a high PPV and therefore a high probability of malignancy, others are nonspecific with differing, low PPV values, which makes it difficult to assign these lesions [5-9]. Difficulties in the decision-making process are particularly noticeable for lesions of BI-RADS category 3, where the reported frequency typically exceeds the optimal rate of 1%–2% [2, 10].

Diffusion-weighted imaging (DWI) and T2-weighted sequences are an essential part of the MRI protocol, but the information gained from these techniques has not been incorporated into the BI-RADS classification. As a non-contrast imaging modality, DWI provides information about the local diffusivity of water in tissue that is typically restricted in malignancies. This can be quantified as the apparent diffusion coefficient (ADC) [11]. Numerous studies have shown the benefit of using ADC to differentiate between malignant and benign breast lesions, and to improve diagnostic specificity [12-14]. Furthermore, incorporating T2-weighted tumour signals together with DWI improved specificity without compromising sensitivity [12]. Therefore, the interpretation of breast MRI should involve a multiparametric evaluation system that is validated in a clinical reporting setup. Logically, the ideal results are achieved when the system complies with the assigned PPV of the BI-RADS and mirrors the values obtained by mammography and US.

Several scoring systems have been developed with the aim of helping radiologists better differentiate malignant from benign breast lesions [7-9, 15-17]. However, none has achieved universal acceptance. The Kaiser scoring system was recently introduced as a system with high diagnostic accuracy and low interobserver variability [17]. The structured algorithm combines five independent diagnostic BI-RADS criteria in an intuitive flowchart. Maltez de Almeida et al. [9] recently proposed a simple points-based classification system for category 4 lesions that is based on the BI-RADS 5th edition descriptors and ADC. Their system is easily reproducible and shows good interobserver agreement. However, none of the published algorithms provides a classification of suspicious findings that complies with all of the assessment categories and descriptors of the BI-RADS.

The aim of this study was to validate a new categorisation scheme for suspicious breast lesions according to the well-defined BI-RADS MRI lexicon descriptors, ADC, T2-weighted signal intensity (SI), and kinetic curve assessment categories. In clinical practice, simple, adequate, and reliable scoring systems are a key element of the decision-making algorithm, and are needed to increase the specificity of MRI and decrease unnecessary follow-up or biopsy.

Methods

This study was approved by the ethics committee of Kuopio University Hospital, which waived the need for written informed consent due to the retrospective nature of the study. All clinical investigations were conducted according to the relevant guidelines and the principles expressed in the Declaration of Helsinki.

Patient selection

All consecutive 3T breast MRI examinations performed on patients at our university hospital between January 2011 and December 2015 were retrospectively retrieved from the regional PACS system. Clinical data were retrieved from the local digital archives and transferred to a database.

A total of 829 consecutive breast MRI examinations were performed on 613 patients during this period. In our institution, breast MRI is performed for selected patients according to national guidelines, which are in concordance with the European Society of Breast Cancer Specialists working group (EUSOMA) guidelines [1]. All lesions were carefully evaluated and histological verification was obtained from suspicious findings, including additional lesions detected by MRI whenever the MRI findings necessitated a change in the treatment plan. The exclusion criteria for this study were as follows: screening examinations with no findings, examinations with only BI-RADS 2 findings and repeated examinations without new suspicious findings, bringing the final number of examinations included to 697 histologically confirmed lesions found in 499 patients.

MRI

Bilateral breast MRI was performed using a 3T MRI scanner (Philips Achieva TX, Philips N.V., Eindhoven, The Netherlands) with a dedicated seven-element phased-array bilateral breast coil. The MRI protocol is presented in Table 1.

Image analysis

The MRI images were first analysed upon diagnosis by specialised breast radiologists, and then retrospectively and blindly by one of two experienced breast radiologists with over 10 and 25 years of multimodal breast imaging experience each. Suspicious findings were strictly described according to the BI-RADS MRI lexicon. In cases with discrepancies with the primary report in the BI-RADS descriptor classifications, concordance was achieved by consensus reading by both radiologists. In our analysis, we used a modified lesion descriptors algorithm, as originally described by Maltez de Almeida et al. [9]. In their study, the authors stated a PPV value of 84% for BI-RADS 4C, that included one major descriptor. Therefore, to evaluate the categorisation scheme for BI-RADS 5 lesions and before the beginning of our analyses, our unit's most senior breast consultant, with over 25 years of experience in multimodal breast imaging, analysed 100 random samples of breast MRI examinations. The statistical results evaluated at different quartiles,

separately and combined, showed that the presence of 3 major findings consistently gave a PPV value of 100%, while the presence of 2 major findings consistently gave PPV values ranging between 92.4% and 100%. Therefore, and for the purpose of this study, 2 major findings were considered sufficient to assign a lesion to the BIRADS 5 category. Using the BI-RADS 5th edition lexicon terminology, we analysed the internal enhancement patterns, types of curves, shapes, and margins of masses, as well as the T2-weighted SI. The algorithm was combined with the ADC threshold values and with the modified categories of kinetic curve assessment.

The ADC were measured by positioning three small regions of interest (ROI; 4 pixels) on visually assessed regions with the lowest SI inside the solid tumour on the ADC map. This measurement method was defined by Wielema et al. as the lowest diffusion breast tumor tissue selection method [18]. The lowest measurement value was then documented. Haemorrhagic, fatty, cystic, and necrotic areas were carefully avoided. An optimal cut-off threshold of 0.69×10^{-3} mm^2/s was applied to differentiate between malignant and benign lesions. This value was based on a previous study from our working group [19], which suggested a lower ADC to establish higher sensitivity and specificity.

Definition of new categories

The findings were classified as Minor, Intermediate, or Major findings. For mass lesions, we classified oval, circumscribed and homogeneous enhancement descriptors as well as dark internal septations and high T2-weighted SI as Minor findings, because these findings are mostly observed in benign lesions. Round shape, irregular margins and "not circumscribed" (irregular), as well as heterogenous enhancement and intermediate or low T2-weighted SI were classified as Intermediate findings. The most suspicious characteristics (i.e. spiculated margin and rim enhancement) were classified as Major lesions. Kinetic curve assessments types 1, 2, and 3 were classified as Minor, Intermediate, and Major, respectively. For non-mass lesions, we classified those with linear, regional, or heterogenous enhancement, and intermediate/mixed or low T2-

weighted SI as Intermediate findings. Segmental, clumped, or clustered ring enhancement were categorised as Major findings. Lesions were then assigned to a BI-RADS category according to the presence and extent of Intermediate or Major findings. Clinical examples of interpretation are presented in Figures (1-5). The classification system and category assignment algorithm are summarised in Table 2.

In the final histopathological diagnosis, both invasive carcinomas and ductal carcinoma in situ were classified as malignant. Lobular neoplasia and atypical ductal hyperplasia were categorised as high-risk lesions. Papillomas, fibroadenomas, and intraductal proliferations without atypia were collectively classified as benign.

Statistical analysis

The PPV with 95% confidence intervals (CI) were calculated using VasarStats for statistical computation available online (http://www.vasarstats.net). Other statistical analyses were performed using IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp, Armonk, NY: IBM Corp.). Correlations between the new BI-RADS categorisation and the final histopathological diagnosis were calculated using Spearman rank correlation coefficient. We fitted a univariate multinomial logistic regression model with malignant diagnosis as the predictor to calculate the odds ratio (OR) of a malignant diagnosis based on all MRI descriptors. Because the nominal predictors of the internal enhancement characteristics of masses and non-mass distributions were categorical but non-ascending in a continuous risk scale, we grouped the descriptors into low-, intermediate-, or high-risk groups according to the risk level originally defined to each descriptor in the 5th edition of the BI-RADS. All of the variables were nominal, except for ADC. Therefore, we dichotomised ADC into ≤ 0.69 and $>0.69 \times 10^{-3}$ mm^2/s. All statistical tests were two-sided with a predetermined cut-off for statistical significance set at $p \leq 0.05$.

Results

The mean age of the patients was 57 years (range 24–88 years). A total of 697 lesions were analysed in 499 patients. The mean size of all lesions was 24 mm (range 4–130 mm), while the mean sizes of mass and non-mass lesions were 19 and 42 mm, respectively. The final histopathologic diagnosis was malignant for 531 (76.2%), benign for 158 (22.7%), and high-risk for 8 (1.1%) lesions.

After assessing the lesions according to the proposed MRI criteria, the BI-RADS subcategory was 4A in 27 (3.9%) lesions, 4B in 53 (7.6%) lesions, and 4C in 174 (25.0%) lesions. A total of 19 (2.7%) lesions had only Minor findings and were classified as BI-RADS category 3; none of these were malignant. The PPV for malignancy in our designated BI-RADS categories 4A, 4B, and 4C were 11.1%, 28.3%, and 64.4%, respectively. The findings are presented in Table 3.

The malignant lesions included 336 invasive ductal carcinomas (63.3%), 96 invasive lobular carcinomas (18.1%), 59 ductal carcinomas in situ (11.1%), and 40 (7.5%) other malignant lesions. Univariate associations among mass lesions and non-mass lesions, as opposed to the final histopathology, and their BI-RADS descriptors and their PPV are presented in Table 4.

The individual mass descriptors with the highest PPV were spiculated margin (99.6%), irregular shape (96.5%), rim enhancement (95.9%), and heterogeneous internal enhancement (87.8%). The individual non-mass descriptors with the highest PPV were clustered ring internal enhancement (94.7%), clumped (87.5%), and segmental distribution (83.8%). In 134 patients, the DWI sequences were missing or deemed nondiagnostic because some lesions were not visible or could not be delineated with certainty or due to technical issues. For evaluable cases, the ADC ranged from 0.13 to 1.30 × 10^{-3} mm^2/s (mean 0.52 × 10^{-3} mm^2/s) for malignant lesions and from 0.31 to 1.37 × 10^{-3} mm^2/s (mean 0.75 × 10^{-3} mm^2/s) for non-malignant lesions.

The OR of our various MRI descriptors are presented in Table 5. All descriptor groups were statistically significant. The descriptors with the highest OR for malignancy were irregular shape, spiculated margins, and rim enhancement. An ADC < 0.69×10^{-3} mm^2/s was associated with an OR of 4.812.

Discussion

This study shows that the proposed multiparametric classification system categorises breast MRI lesions efficiently, practically and with values consistent with the latest edition of BI-RADS recommendations. The PPV of BI-RADS category 3 as well as the category 4 subclassifications, excluding 4A, fall within the same ranges for mammographic and ultrasonographic reports. However, the algorithm seems to slightly overrate lesions in BI-RADS category 5 and slightly underrate lesions in BI-RADS subcategory 4A. Additionally, the assignment of only 2.7% of patients to BI-RADS category 3 was well within the desired frequency of < 10%, and is very close to the optimal rate of 1%–2% [2]. This is particularly important considering that breast MRI is the most sensitive imaging modality and is usually reserved for high-risk patients, in the preoperative setting, or to resolve unclear mammographic or US findings. Previously published multiparametric scoring systems and analytical strategies have incorporated kinetic and morphological parameters to achieve high accuracy, but none of these integrated all suspicious findings (for both mass- and non-mass lesions) or achieved an optimal recommended PPV threshold of BI-RADS categories 3–5, which also include the subcategories 4A–C. A summary of recently published studies is presented in Table 6.

Some of the preliminary systems were over-represented by kinetic features [7, 16]. However, multivariable models involving feature assessment showed superior diagnostic accuracy to those involving qualitative characterisation of the dynamic enhancement patterns [33]. Nevertheless, the

wash-out curve has a high PPV for malignancy, as demonstrated in our study, and it was therefore considered as a Major criterion, together with the morphological features with high PPV.

Fischer et al. were amongst the first to propose a scoring system to categorise findings, with an initial sensitivity of 93% and specificity of 65% [16]. Sarica et al. [22] reported a BI-RADS classification with sensitivity and specificity of 94.2% and 56.1%, respectively. Lui et al. [21] used the ADC and Fischer score to subdivide the BI-RADS category 4 lesions with PPV of 7.69%, 52.38%, and 89.29% for BI-RADS subcategories 4A, 4B, and 4C, respectively. Unfortunately, most of the studies based on Fischer scores were limited by the fact that the MRI findings do not strictly conform to all of the descriptors in the 5th edition of the BI-RADS, causing confusion when implemented into clinical practice.

Tozaki et al. [5, 15] evaluated the PPV and negative predictive values (NPV) of BI-RADS MRI descriptors to develop a model based on kinetic and morphological parameters. For mass lesions, the sensitivity, specificity, PPV, and NPV were 99%, 89%, 96%, and 98%, respectively, and the PPV for non-mass lesions was 94% [34]. Fujiwara et al. [25] investigated the contributions of morphological and kinetic features of mass lesions to the probability of malignancy using a grading system of Minor to Major descriptors with high PPV of 27.8%, 79.2%, and 98.4% for BI-RADS subcategories 4A, 4B, and 4C, respectively. The authors acknowledged, however, that the inclusion of ADC may be used to downgrade benign lesions to categories 3, 4A, and 4B. The Kaiser score, first introduced in 2013 [17], combines five BI-RADS criteria into a flowchart (margins, SI–time curve type, internal enhancement and presence of oedema) to determine the probability of detecting a malignant lesion, and can be used to downgrade suspicious category 4 findings [35].

This study assessed a subcategorisation scheme for BI-RADS MRI lesions that was based on an earlier classification described by Maltez de Almeida et al. in 2015 [9]. Their system subcategorised BI-RADS category 4 lesions into three subcategories with PPV comparable with the BI-RADS recommendations for mammography and US (4A, 15%; 4B, 37%; and 4C, 84%).

Our classification departs from theirs in a few ways. First, we sought to introduce classification criteria that would also classify lesions into BI-RADS categories 3 and 5. We also changed the predictive order of the kinetic curve enhancement descriptors, and added some strong descriptors from the BI-RADS lexicon, which have high OR for malignancy, but were missing from their system. We also used a different method to measure ADC. Furthermore, the PPV in our study are in accordance with those reported by Tozaki et al. [15, 34], but differ slightly from those reported by de Almeida et al. [36], which may be due to the small number of patients in that study.

Low ADC are well-established markers of malignancy that are independent of the morphological and kinetic features of dynamic contrast MRI [37, 38]. Numerous studies have shown a benefit of ADC for differentiating between malignant and benign lesions of the breast and improving diagnostic specificity [12, 38]. In accordance with an earlier study from our group, we measured the ADC in small ROI [19]. This method was superior to, and differentiated malignant and benign lesions more accurately than whole-lesion ROI. The ADC threshold used in this study (0.69×10^{-3} mm^2/s) is lower than the values used in other studies ($0.91–1.25 \times 10^{-3}$ mm^2/s). This is mostly due to a lack of standardisation of ADC measurement methods, where different centres use varying protocols, varying b values, and DWI sequences are collected either before or after contrast administration. The small ROI resulted in lower ADC and, while the ROI analysis methods remain unstandardized, the recent guidelines recommend using a small ROI to document the worst area [11], as in the present study. Therefore, we recommend that each centre should use a locally validated cut-off value best suited to their imaging and measurement protocols.

This study has some limitations. Our database comprised a high percentage of malignancies, which relates to the fact that our hospital is a tertiary teaching centre that receives referrals from a wide catchment compromising the entirety of eastern Finland. Not all patients in our hospital undergo preoperative MRI and therefore these results represent a selected population of patients complying to the EUSOMA guidelines. Furthermore, this is a retrospective, single-centre study.

Nevertheless, this study constitutes a prospectively gathered database with consistent imaging and interpretation protocols.

In conclusion, using a large database analysis of 697 breast lesions, we showed that incorporating BI-RADS MRI-descriptors coupled with ADC and T2-weighted SI into a multiparametric classification system, an applicable categorization of lesions can be achieved with PPV values within the recommended BI-RADS categorization range.

Funding:
This work was supported in part by a grant (AI) from Kuopio University Hospital (VTR grant 5063542), the Radiological Society of Finland (MS) and the Mauri and Sirkka Wiljasalo fund (AI, HO). The authors declare no relationships with any companies whose products or services may be related to the subject matter of the article. The funding sources were not involved in study design, data collection or analysis, preparation of the manuscript, or the decision to submit the manuscript.

Disclosure:

All authors declare that they have no competing interests.

CRediT authorship contribution statement

Conceptualization	MS; RV; AI;AS
Methodology	MS; AS; RV;
Validation	MS; AI
Formal analysis	AM; HO; AI
Investigation	AI; MS; AM; HO; AS
Resources	MS; RV
Data Curation Management	AI; HO; AM; MS
Writing - Original Draft	MS; AI; AM; HO
Writing - Review & Editing	All authors
Visualization	All authors
Supervision	MS; RV
Project administration	MS; RV
Funding acquisition	RV

References

1. F. Sardanelli, C. Boetes, B. Borisch, T. Decker, M. Federico, F.J. Gilbert, et al, Magnetic resonance imaging of the breast: Recommendations from the EUSOMA working group, Eur J Cancer. 46 (2010) 1296-1316. Doi:10.1016/j.ejca.2010.02.015.
2. C.J. D'Orsi, E.A. Sickles, E.B. Mendelson, E.A Morris, ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. Reston, VA, American College of Radiology; 2013.
3. R.M. Mann, N. Cho, L. Moy, Breast MRI: State of the art, Radiology. 292 (2019) 520-536. Doi:10.1148/radiol.2019182947.

4. G.J. Wengert, F. Pipan, J. Almohanna, H. Bickel, S. Polanec, P. Kapetas, et al, Impact of the kaiser score on clinical decision-making in BI-RADS 4 mammographic calcifications examined with breast MRI, Eur Radiol.30 (2020) 1451-1459. Doi:10.1007/s00330-019-06444-w.
5. M. Tozaki, K. Fukuda, High-spatial-resolution MRI of non-masslike breast lesions: Interpretation model based on BI-RADS MRI descriptors, AJR Am J Roentgenol. 187 (2020) 330-337. Doi:187/2/330.
6. W.B. Demartini, B.F. Kurland, R.L. Gutierrez, C.C. Blackmore, S. Peacock, C.D. Lehman, Probability of malignancy for lesions detected on breast MRI: A predictive model incorporating BI-RADS imaging features and patient characteristics, Eur Radiol. 21 (2011) 1609-1617. Doi:10.1007/s00330-011-2094-6
7. M.O Leach, C.R. Boggis, A.K. Dixon, D.F. Easton, R.A. Eeles, D.G. Evans, et al, Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: A prospective multicentre cohort study (MARIBS), Lancet. 365 (2005) 1769-1778. Doi:S0140-6736(05)66481-1
8. D.M. Ikeda, N.M. Hylton, K. Kinkel, M.G. Hochman, C.K. Kuhl, W.A. Kaiser, et al, Development, standardization, and testing of a lexicon for reporting contrast-enhanced breast magnetic resonance imaging studies, J Magn Reson Imaging. 13 (2001) 889-895. Doi:10.1002/jmri.1127
9. J.R. Maltez de Almeida, A.B. Gomes, T.P. Barros, P.E. Fahel, M. de Seixas Rocha, Subcategorization of suspicious breast lesions (BI-RADS category 4) according to MRI criteria: Role of dynamic contrast-enhanced and diffusion-weighted imaging, AJR Am J Roentgenol. 205 (2015) 222-231. Doi:10.2214/AJR.14.13834
10. C. Spick, H. Bickel, S.H. Polanec, P.A. Baltzer, Breast lesions classified as probably benign (BI-RADS 3) on magnetic resonance imaging: A systematic review and meta-analysis, Eur Radiol 28 (2018) 1919-1928. Doi:10.1007/s00330-017-5127-y
11. P. Baltzer, R.M Mann, M. Iima, E.E. Sigmund, P. Clauser, F.J. Gilbert, et al, Diffusion-weighted imaging of the breast-a consensus and mission statement from the EUSOBI international breast diffusion-weighted imaging working group, Eur Radiol. 30 (2020) 1436-1450. Doi:10.1007/s00330-019-06510-3
12. O. Arponen, A. Masarwah, A. Sutela, M. Taina, M. Könönen, R. Sironen R, et al, Incidentally detected enhancing lesions found in breast MRI: Analysis of apparent diffusion coefficient and T2 signal intensity significantly improves specificity, Eur Radiol. 26 (2016) 4361-4370. Doi:10.1007/s00330-016-4326-2
13. S.C. Partridge, W.B. Demartini, B.F. Kurland, P.R. Eby, S.W. White, C.D. Lehman, Differential diagnosis of mammographically and clinically occult breast lesions on diffusion-weighted MRI, J Magn Reson Imaging. 31 (2010) 562-570. Doi:10.1002/jmri.22078
14. D. Leithner, G. Wengert, T. Helbich, E. Morris, K. Pinker, MRI in the assessment of BI-RADS(R) 4 lesions, Top Magn Reson Imaging. 26 (2017) 191-199. Doi:10.1097/RMR.0000000000000138.
15. M. Tozaki, T. Igarashi, S. Matsushima, K. Fukuda, High-spatial-resolution MR imaging of focal breast masses: Interpretation model based on kinetic and morphological parameters, Radiat Med. 23 (2005) 43-50.
16. U. Fischer, L. Kopka, E. Grabbe, Breast carcinoma: Effect of preoperative contrast-enhanced MR imaging on the therapeutic approach, Radiology. 213 (1999) 881-888. Doi:10.1148/radiology.213.3.r99dc01881
17. P.A. Baltzer, M. Dietzel, W.A. Kaiser, A simple and robust classification tree for differentiation between benign and malignant lesions in MR-mammography, Eur Radiol. 23 (2013) 2051-2060. Doi:10.1007/s00330-013-2804-3
18. Wielema M, Dorrius MD, Pijnappel RM, et al., Diagnostic performance of breast tumor tissue selection in diffusion weighted imaging: A systematic review and meta-analysis. PLoS One 2020;15(5):e0232856.
19. O. Arponen, M. Sudah, A. Masarwah, M. Taina, S. Rautiainen, M. Könönen, et al, Diffusion-Weighted Imaging in 3.0 Tesla Breast MRI: Diagnostic Performance and Tumor Characterization Using Small Subregions vs. Whole Tumor Regions of Interest, PLoS One. Oct (2015) 12;10(10):e0138702. doi: 10.1371/journal.pone.0138702. eCollection 2015.
20. Mahoney MC, Gatsonis C, Hanna L, DeMartini WB, Lehman C, Positive predictive value of BI-RADS MR imaging. Radiology 2012;264(1):51-8.
21. Strigel RM, Burnside ES, Elezaby M, et al., Utility of BI-RADS Assessment Category 4 Subdivisions for Screening Breast MRI. AJR Am J Roentgenol 2017;208(6):1392-9.

22. O. Sarica, G. Kaymaz, N. Ucar, Effectiveness of additional diagnostic parameters in magnetic resonance mammography: A comparative study with the BI-RADS classification and scoring system, J Comput Assist Tomogr. 38 (2014) 985-991. Doi:10.1097/RCT.0000000000000132
23. Pinker K, Bogner W, Baltzer P, et al., Improved diagnostic accuracy with multiparametric magnetic resonance imaging of the breast using dynamic contrast-enhanced magnetic resonance imaging, diffusion-weighted imaging, and 3-dimensional proton magnetic resonance spectroscopic imaging. Invest Radiol 2014;49(6):421-30.
24. Kawai M, Kataoka M, Kanao S, et al., The Value of Lesion Size as an Adjunct to the BI-RADS-MRI 2013 Descriptors in the Diagnosis of Solitary Breast Masses. Magn Reson Med Sci 2018;17(3):203-10.
25. K. Fujiwara, T. Yamada, Y. Kanemaki, S. Okamoto S, Y. Kojima, K. Tsugawa, et al, Grading system to categorize breast MRI in BI-RADS 5th edition: A multivariate study of breast mass descriptors in terms of probability of malignancy, AJR Am J Roentgenol. 210 (2018) W118-W127. Doi:10.2214/AJR.17.17926
26. Asada T, Yamada T, Kanemaki Y, Fujiwara K, Okamoto S, Nakajima Y, Grading system to categorize breast MRI using BI-RADS 5th edition: a statistical study of non-mass enhancement descriptors in terms of probability of malignancy. Jpn J Radiol 2018;36(3):200-8.
27. D. Liu, Z. Ba, X. Ni, L. Wang, D. Yu, X. Ma, Apparent diffusion coefficient to subdivide breast imaging reporting and data system magnetic resonance imaging (BI-RADS-MRI) category 4 lesions, Med Sci Monit. 24 (2018) 2180-2188. Doi:907000

28. Ellmann S, Wenkel E, Dietzel M, et al., Implementation of machine learning into clinical breast MRI: Potential for objective and accurate decision-making in suspicious breast masses. PLoS One 2020;15(1):e0228446.
29. 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):2529-37.
30. 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):3799-809.
31. Cloete DJ, Minne C, Schoub PK, Becker JHR, Magnetic resonance imaging of fibroadenoma-like lesions and correlation with Breast Imaging-Reporting and Data System and Kaiser scoring system. SA J Radiol 2018;22(2):1532.
32. Milos RI, Pipan F, Kalovidouri A, et al., 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 2020. doi: 10.1007/s00330-020-06945-z. Online ahead of print.
33. M.D. Schnall MD, J. Blume, D.A. Bluemke, G.A. DeAngelis, N. DeBruhl, S. Harms, et al, Diagnostic architectural and dynamic features at breast MR imaging: Multicenter study, Radiology. 238 (2006) 42-53. Doi:238/1/42
34. M. Tozaki, T. Igarashi, K. Fukuda, Positive and negative predictive values of BI-RADS-MRI descriptors for focal breast masses, Magn Reson Med Sci. 5 (2006) 7-15. Doi:JST.JSTAGE/mrms/5.7
35. M. Dietzel, P.A.T. Baltzer, How to use the kaiser score as a clinical decision rule for diagnosis in multiparametric breast MRI: A pictorial essay, Insights Imaging. 9 (2018) 325-335. Doi:10.1007/s13244-018-0611-8.
36. J.R. de Almeida, A.B. Gomes, T.P. Barros, P.E. Fahel, S. Rocha Mde, Predictive performance of BI-RADS magnetic resonance imaging descriptors in the context of suspicious (category 4) findings, Radiol Bras. 49 (2016) 137-143. Doi:10.1590/0100-3984.2015.0021
37. M. Dietzel, S. Ellmann, R. Schulz-Wendtland, P. Clauser, E. Wenkel, M. Uder, et al, Breast MRI in the era of diffusion weighted imaging: Do we still need signal-intensity time curves?, Eur Radiol. 30 (2020) 47-56. Doi:10.1007/s00330-019-06346-x
38. G.C. Baxter, M.J. Graves, F.J. Gilbert, A.J. Patterson, A meta-analysis of the diagnostic performance of diffusion MRI for breast lesion characterization, Radiology. 291 (2019) 632-641. Doi:10.1148/radiol.2019182510

Figure 1: A true negative mass lesion categorization in a 41-year-old female. An oval circumscribed mass lesion with high signal intensity in T2w imaging (A; STIR) is seen in the medial aspect of the retroglandular area of the right breast (arrows). The lesion shows homogeneous enhancement and dark internal septations in contrast-enhanced T1W sequence (B) and in subtraction image (C). The colour map image of enhancement (D) shows green-coded persistent (type 1) enhancement curve (E). Diffusion weighted imaging shows a "shine through" effect (F; DWI, b=800) with high ADC values (G; ADC map). Due to the presence of only minor descriptors, the mass was categorized as BI-RADS 3 and histopathology revealed a benign fibroadenoma.

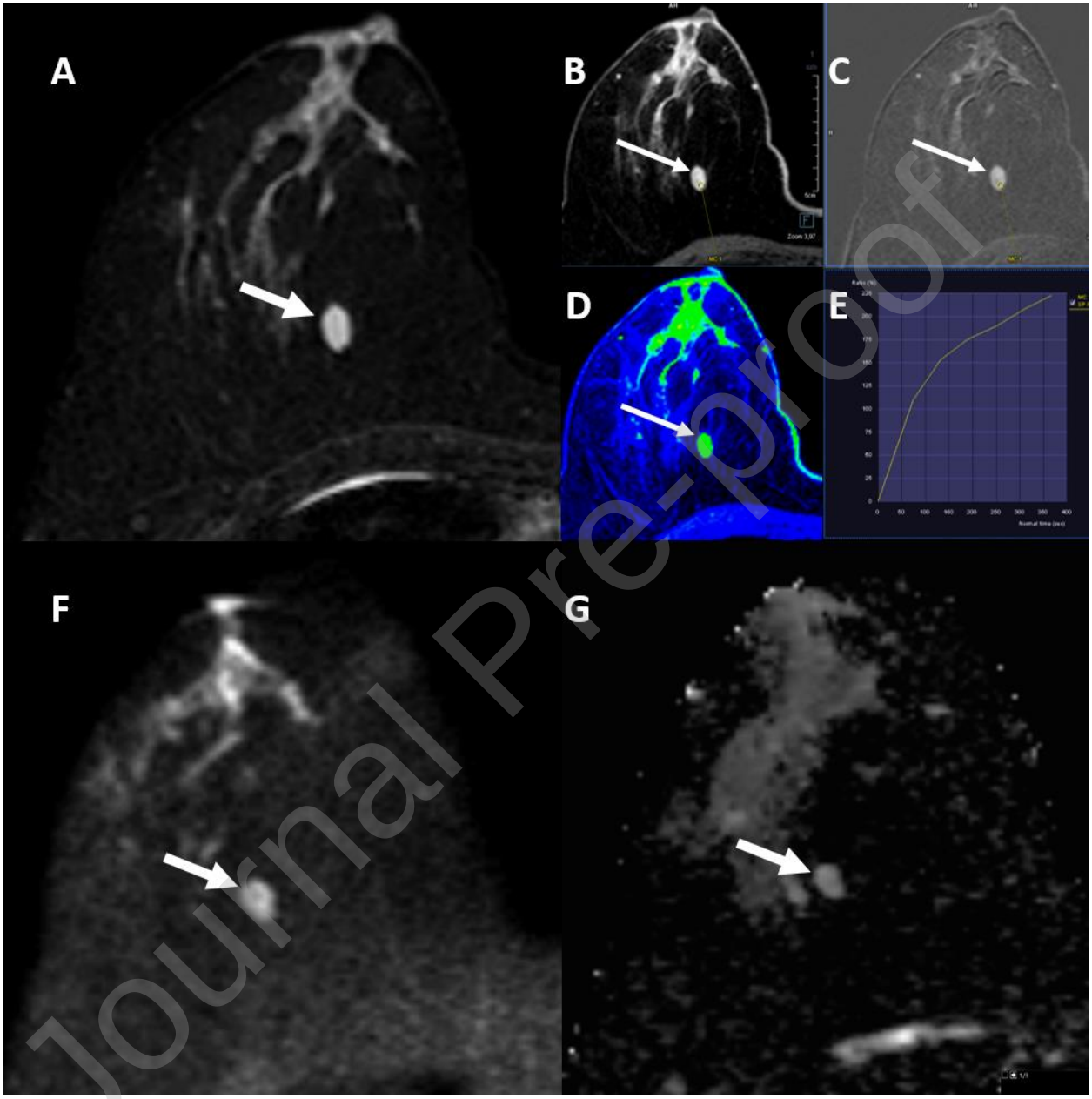

Figure 2: A true positive mass lesion categorisation in a 48-year-old female. A non-circumscribed mass lesion (arrows), with intermediate T2 signal intensity (A; STIR), rim-enhancement in contrast enhanced T1w sequence (B), restricted diffusion (C; DWI: b=800), low ADC values (D; ADC-map: $<0.69\times10^{-3}\,mm^2/s^{-1}$) and a wash-out (type 3) enhancement curve (E). The lesion was categorized as BI-RADS 5 due to the presence of 3 major descriptors. Histopathology revealed invasive ductal carcinoma.

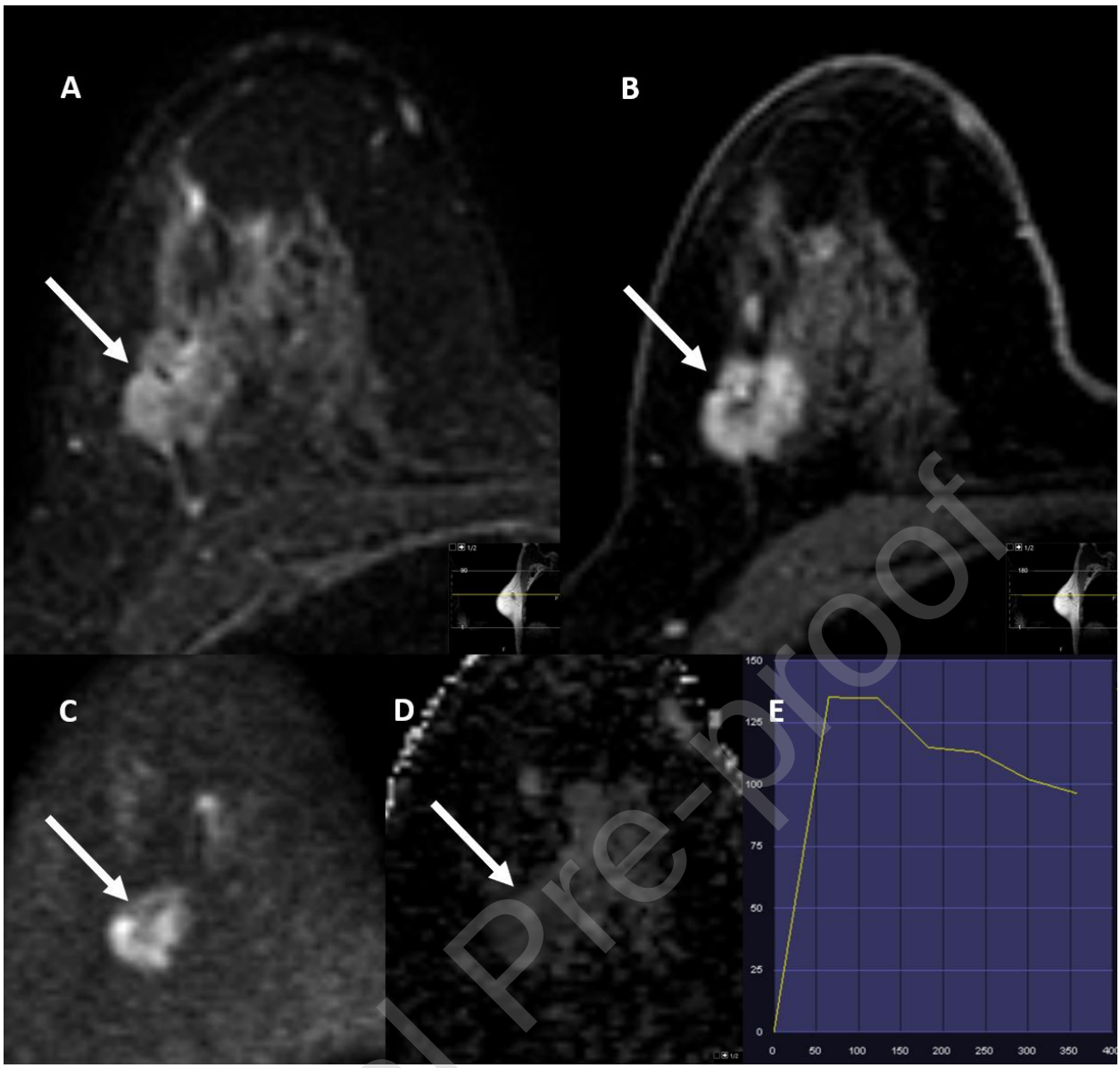

Figure 3: A false positive mass lesion categorisation in a 41-year-old female. A round circumscribed mass lesion mostly with high signal intensity in T2w imaging (A; STIR), rim-enhancement in contrast-enhanced T1W sequence (B), areas of restricted diffusion with low ADC values (C; ADC-map: <0.69×10−3 mm2/s−1 and a wash-out (type 3) enhancement curve (D). It was categorized as BI-RADS 5 due to the presence of 3 major descriptors. Histopathology revealed a benign tubular adenoma with areas of adenosis.

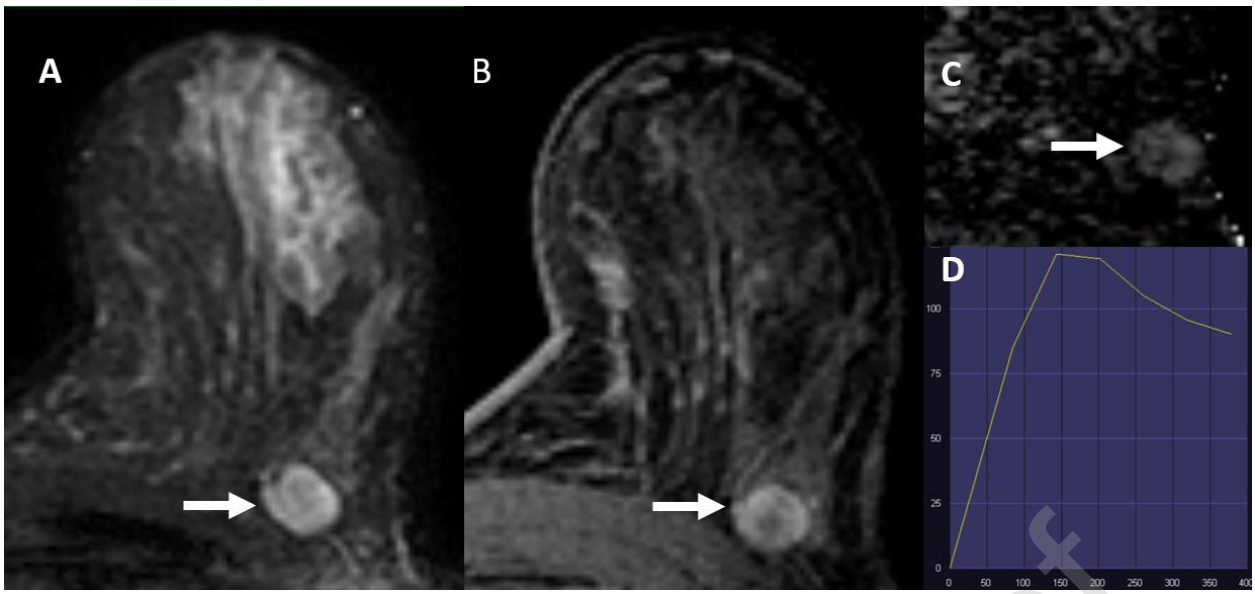

Figure 4: A true positive non-mass lesion categorisation in a 32-year-old female. An asymmetrical diffuse non-mass enhancement (arrows) with mainly clustered ring enhancement pattern is visualised in maximum intensity projections of the left breast (arrows) from contrast enhanced T1w sequence in axial (A) and sagittal (B) orientations. The colour map image of enhancement (C) showed small red-coded areas of wash-out (type 3) enhancement curves (D). It was categorized as BI-RADS 5 due to the presence of 2 major descriptors. Histopathology revealed global DCIS with multiple foci of invasive ductal carcinoma.

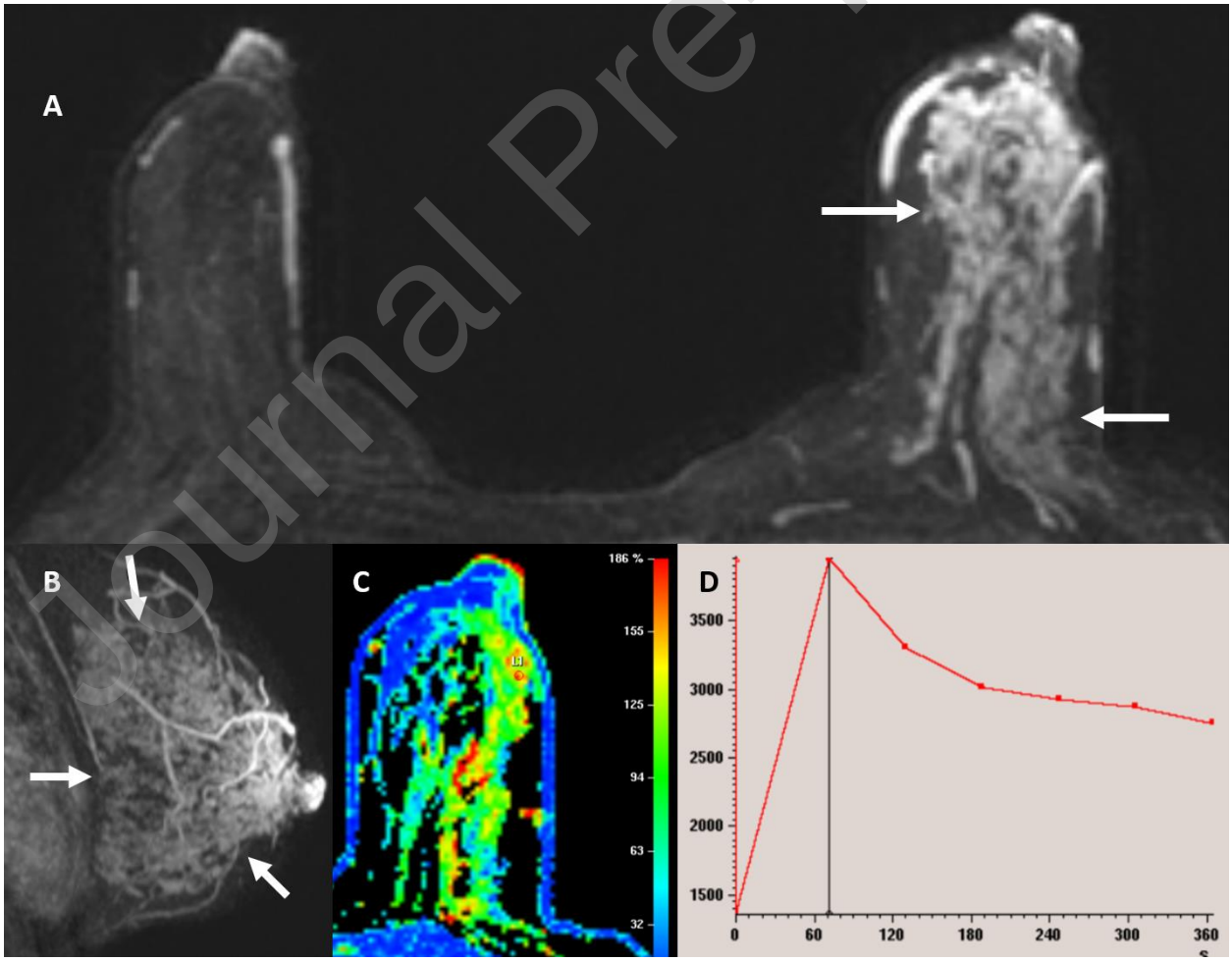

Figure 5: A false positive non-mass lesion categorisation in a 38-year-old female. A maximum intensity projection of both breasts in axial (A) and sagittal (B) orientations shows an asymmetrical segmental, clumped (C; contrast enhanced T1w) non-mass enhancement in the left breast. The colour map image of enhancement (D) shows small areas of yellow-coded plateau enhancement curve (E; type 2). It was categorized as BI-RADS 5 due to the presence of 2 major descriptors. Histopathology revealed non-specific inflammatory changes.

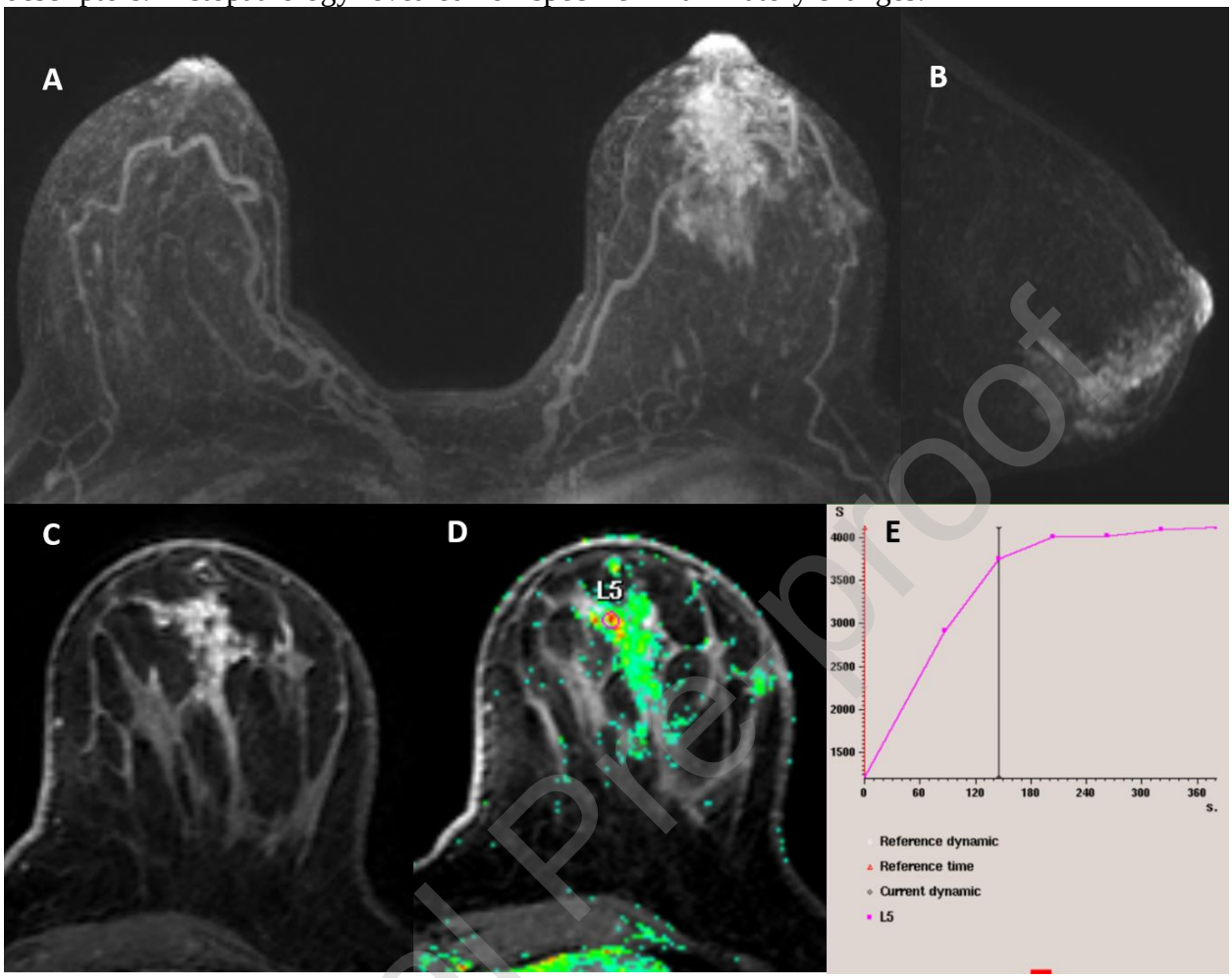

Table 1. Breast MRI protocol

Sequence	TR/TE (ms)	in-plane resolution mm	Slice thickness (mm)	Scanning time
T1-FFE¤	4.57 /2.3	0.48 × 0.48	0.7	6 min 11 s
T2-TSE※	5000/120	0.6 × 0.6	2	3 min 20 s
STIR□	5000 /60	1 × 1	2	5 min 40 s
T1 dynamic*	4.67 /2.31	0.96 × 0.96	1	58.5 s
DWI#	7168 /95	1.15 × 1.15	4	4 min 8 s

¤FFE = fast field echo
※TSE = turbo spin echo
□STIR = Short tau inversion recovery
*eTHRIVE spectrally adiabatic inversion recovery (SPAIR) fat suppression; pre-contrast and six phases after gadoterate meglumine (0.1 ml/kg, 3 ml/s) administration followed by a saline chaser
#DWI = diffusion-weighted echo planar imaging with five b factors (0, 200, 400, 600 and 800 s/mm^2) performed after contrast administration. The apparent diffusion coefficients maps were automatically calculated linearly using the manufacturer's method

Table 2. Proposed MRI criteria and interpretation algorithm for Minor, Intermediate and Major findings

Main Pattern	Minor	Intermediate	Major
Mass	-Oval -Circumscribed -Dark internal septations -Homogeneous enhancement T2 signal intensity: -High	-Round -Irregular -Not circumscribed (irregular) -Heterogeneous enhancement T2 signal intensity: -Intermediate or low	-Rim enhancement -Not circumscribed (Spiculated)
Non-mass	Enhancement: -Focal -Multiple regions -Diffuse -Homogeneous T2 signal intensity: -High	Enhancement: -Linear -Regional -Heterogeneous T2 signal intensity: -Intermediate or low	Enhancement: -Segmental -Clumped -Clustered ring
Kinetic curve	Type 1	Type 2	Type 3
DWI ADC			$\leq 0{,}69 \times 10^{-3}$ mm ^{2}/s
BI-RADS category			
BI-RADS 3	only Minor findings		
BI-RADS 4A	one Intermediate, otherwise Minor findings		
BI-RADS 4B	two or three Intermediate findings		
BI-RADS 4C	one Major and BIRADS 4B findings, or four Intermediate findings		
BI-RADS 5	two or more Major findings		

DWI = diffusion-weighted imaging

ADC = apparent diffusion coefficient
BI-RADS = Breast Imaging Reporting and Data System

Table 3. The BI-RADS categories of detected lesions and positive predictive values (PPV) according to the final histopathological diagnoses

BI-RADS*	Lesions n(%)	Malignant n(%)	Benign n(%)	Risk n(%)	p	PPV[#] (95% CI[¤])
3	19(2.7)	0	19(100)	0	.000	0.0
4a	27(3.9)	3(11.1)	24(88.9)	0	.000	11.1 (2.9 – 30.3)
4b	53(7.6)	15(28.3)	36(67.9)	2(3.8)	.000	28.3 (17.2 – 42.6)
4c	174(25.0)	112(64.4)	60(34.5)	2(1.1)	.000	64.4 (56.7 – 71.4)
5	424(60.8)	401(94.6)	19(4.5)	4(0.9)	.000	94.6 (91.6 – 96.5)
TOTAL	697(100)	531(76.2)	158(22.7)	8(1.1)		

*BI-RADS = Breast Imaging Reporting and Data System
[#]PPV = positive predictive value
[¤]CI = confidence interval

Table 4. Positive predictive values (PPV) and associations of BI-RADS lexicon-descriptors with the final histopathological diagnosis

Factors	Descriptors	Total (%)	Benign lesions n (%)	Malignant + risk Lesions n (%)	PPV* (%)	95% CI[#]
MASS-lesions		555 (79,6)	111(20)	444(80)	**80**	76,4 - 83,2
Shape	Oval	169(30.5)	69(40.8)	100(59.2)	**59.2**	51,3 – 66,6
	Round	129(23.2)	33(25.6)	96(74.4)	**74.4**	65,8 – 81,5
	Irregular	257(46.3)	9(3.5)	248(96.5)	**96.5**	93,2 – 98,3
Margin	Circumscribed	152(27.4)	96(63.2)	56(36.8)	**36.8**	29,3 – 45,1
	Irregular	138(24.9)	14(10.1)	124(89.9)	**89.9**	83,3 - 94,1
	Spiculated	265(47.7)	1(0.4)	264(99.6)	**99.6**	97,6 – 100
Internal enhancement	Homogenous	25(4.5)	19(76)	6(24)	**24.0**	10.2 – 45.5
	Heterogeneous	329(59.3)	40(12.2)	289(87.8)	**87.8**	83.7 – 91.1
	Rim enhancement	145(26.1)	6(4.1)	139(95.9)	**95.9**	90.8 – 98.3
	Dark internal septations	56(10.1)	46(82.1)	10(17.9)	**17.9**	9.3 – 30.8
NON-MASS-lesions		142(20.4)	47(33.1)	95(66.9)	**66.9**	58.4 – 74.4
Enhancement Distribution	Focal	37(26.1)	19(51.4)	18(48.6)	**48.6**	32.2 – 65.3
	Linear	16(11.3)	12(75)	4(25)	**25.0**	8.3 – 52.6
	Segmental	69(48.6)	11(15.9)	58(84.1)	**84.1**	72.8 – 91.4
	Regional	13(9.2)	4(30.8)	9(69.2)	**69.2**	38.9 – 89.6
	Multiple region	6(4.2)	1(16,7)	5(83.3)	**83.3**	36.5 – 99.1
	Diffuse	1(0.7)	0	1(100)	**100**	5.5 – 100
Internal enhancement characteristics	Homogenous	11(7.7)	8(72.7)	3(27.3)	**27.3**	7.3 – 60.7
	Heterogenous	64(45.1)	32(50)	32(50)	**50.0**	37.4 – 62.6
	Clumped	48(33.8)	6(12.5)	42(87.5)	**87.5**	74.1 – 94.8

	Clustered ring	19(13.4)	1(5.3)	18(94.7)	**94.7**	71.9 – 99.7
MASS and NON-MASS-lesions						
T2 signal intensity	High	76(10.9)	46(60.5)	30(39.5)	**39.5**	28.7 – 51.4
	Intermediate/ Mixed	600(86.1)	106(17.7)	494(82.3)	**82.3**	79.0 – 85.3
	Low	21(3.0)	6(28.6)	15(71.4)	**71.4**	47.7 – 87.8
Kinetic curve assessment/ Delayed Phase	Persistent	114(16.4)	64(56.1)	50(43.9)	**43.9**	34.7 – 53.5
	Plateau	299(42.9)	65(21.7)	234(78.3)	**78.3**	73.1 – 82.7
	Washout	284(40.7)	29(10.2)	255(89.8)	**89.8**	85.5 – 92.9

*PPV = positive predictive value
[#]CI = confidence interval

Table 5. Multinomial logistic regression model showing odds ratios (OR) for MRI descriptors with 95% confidence intervals and p-values

Variable	**Exp (B)**	**95% CI[¤]**		**p**
		Lower	***Upper***	
Mass Shape				<0.001
Oval*				
Round	1.502	0.885	2.551	
Irregular	14.002	6.597	29.723	
Mass Margins				<0.001
Circumscribed*				
Irregular	4.501	2.336	8.670	
Spiculated	135.835	18.475	998.684	
Mass T2 Signal Intensity				<0.001
High*				
Intermediate	7.030	4.240	11.657	
Low	3.833	1.338	10.982	
Mass Internal Enhancement Characteristics				<0.001
Homogenous and Dark Internal Septations*				
Heterogenous	28.945	15.273	54.856	
Rim Enhancement	94.115	35.202	251.618	
Non mass Distribution				<0.001
Focal, multiple regions and diffuse*				
Linear and Regional	0.652	0.250	1.701	
Segmental	4.427	1.834	10.688	
Non-Mass Internal Enhancement Characteristics				<0.001
Homogenous*				
Heterogenous	0.244	0.143	0.418	
Clumped	1.682	0.696	4.067	
Clustered Ring	4.541	0.600	34.382	
Kinetic Curve Assessment Delayed				<0.001
Persistent*				
Plateau	4.839	3.036	7.715	
Washout	11.880	6.937	20.344	
ADC[#] Values				<0.001

>0.69*				
≤0.69	4.812	3.059	7.570	

*Reference value
¤CI = confidence interval
#ADC = apparent diffusion coefficient

Table 6: Recently published studies on Breast Imaging-Reporting and Data System (BI-RADS) lesion classification and interpretation models.

Author	Year	Purpose	Focus	Included features	Results
Demartini et al. [6]	2011	Multivariate model incorporating patient and lesion characteristics for MRI-detected breast lesions	Malignant vs. Benign	Lesion size, clinical indication, age, kinetics	Strongest associations: washout (OR 4.2); clinical indication OR 3.0.
Mahoney et al. [20]	2012	PPV of BI-RADS and identify the most predictive lesion features for malignancy	Malignant vs. Benign	BI-RADS descriptors	PPV for BI-RADS 3, 0.009; BI-RADS 4, 0.205; BI-RADS 5, 0.714
Strigel et al. [21]	2017	Subcategorization of BI-RADS 4 by calculating PPV	Subcategorization of BI-RADS 4	According to radiologists experience + PPV	PPV for 4A, 0.025; 4B, 0.276; 4C, 0.833
Sarica et al. [22]	2014	Combining Fischer scoring system and BI-RADS descriptors	Malignant vs. Benign	BI-RADS descriptors and adjacent vessel sign, focal edema, hook sign	sensitivity 89.9%, specificity 88.5%
Pinker et al. [23]	2014	Compare 2 vs 3 parameters in cancer diagnosis	Malignant vs. Benign	Dynamic, DWI, proton MR spectroscopic imaging	sensitivity 100%, specificity 87.2%
Almeida et al. [9]	2015	Subcategorization of BI-RADS 4 and impact of ADC on diagnostic performance.	Subcategorization of BI-RADS 4	BI-RADS+ADC	AUC 0.89; PPV for 4A, 0.15; 4B, 0.37; 4C, 0.84
Kawai et al. [24]	2018	Appropriate categorization BI-RADS for solitary masses with the focus on lesion size	Categorization (BI-RADS 3; 4A; 4B; 4C; 5)	BI-RADS descriptors, size, ADC	PPV for 4A, 0.006; 4B, 0.31; 4C, 0.62-0.71; 5, 1.0
Fujiwara et al. [25]	2018	Grading system to categorize mass lesion	Categorization (BI-RADS 3; 4A; 4B; 4C; 5)	shape, margin, internal enhancement, delayed phase, ADC	PPV for BI-RADS 3, 0.05; 4A, 0.38; 4B, 0.79; 4C, 0.98; 5, 0.99
Asada et. al. [26]	2018	Grading system to categorize non-mass lesions	Categorization (BI-RADS 3; 4A; 4B; 4C; 5)	internal enhancement, distribution	PPV for BI-RADS 3, 0; 4A, 0.56; 4B, 0.73; 4C, 0.81; 5, 0.94
Liu et al. [27]	2018	Subcategorization of BI-RADS 4 and evaluate Fischer´s scoring system+ADC	Subcategorization of BI-RADS 4	shape, margin, initial and postinitial enhancement, ADC	PPV for 4A, 0.08, 4B, 0.52; 4C, 0.89
Ellmann et al. [28]	2020	Integration of machine learning into MRI interpretation improves the management of suspicious masses.	Malignant vs. Benign, especially in BI-RADS 4	Age, lesion size, diffusion restriction, T2W intensity, lesion vascularity.	AUC 90.1%, sensitivity 92.5%, specificity 76.8%
Baltzer et al. [17]	2013	Simple classification tree for differential diagnosis in breast MRI	Malignant vs. Benign	”Kaiser score” features: Root sign delayed enhancement margins internal enhancement edema	88,40 %
Marino et al. [29]	2016	Impact of "Kaiser score" on inter-reader agreement and diagnostic performance	Malignant vs. Benign		AUC 0.889-0.943
Woitek et al. [30]	2017	"Kaiser score" obviates unnecessary MRI-guided biopsy in lesions only visible on MRI.	Malignant vs. Benign, especially in BI-RADS 4		AUC 0.873
Cloete et al. [31]	2018	Characterise lesions resembling fibroadenomas	Malignant vs. Benign		sensitivity 50%, specificity 84.6%, PPV 6.25%, NPV 98.8%
Wengert et al. [4]	2020	"Kaiser score" downgrades breast lesions with calcifications	Malignant vs. Benign		AUC 0.859-0.889
Milos et al. [32]	2020	"Kaiser score" excludes malignancy in BI-RADS 4 classified lesions	Malignant vs. Benign, especially in BI-RADS 4		AUC 0.865-0.902

Abbreviations: OR, odds ratio; PPV, positive predictive value; AUC, area under curve; ADC, apparent diffusion coefficient; NPV, negative predictive value.