



Can supplementary contrast-enhanced MRI of the breast avoid needle biopsies in suspicious microcalcifications seen on mammography? A systematic review and meta-analysis



Barbara J. Fueger, Paola Clauser, Panagiotis Kapetas, Nina Pötsch, Thomas H. Helbich, Pascal A.T. Baltzer*

Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Waehringer Guertel 18-20, A-1090, Vienna, Austria

ARTICLE INFO

Article history:

Received 31 December 2020
Received in revised form
5 February 2021
Accepted 6 February 2021
Available online 15 February 2021

Keywords:

Breast
Sensitivity and specificity
Biopsy
Microcalcifications
Magnetic resonance imaging
Mammography

ABSTRACT

Purpose: To analyze the rate of potentially avoidable needle biopsies in mammographically suspicious calcifications if supplementary Contrast-Enhanced MRI (CE-MRI) is negative.

Methods: Using predefined criteria, a systematic review was performed. Studies investigating the use of supplemental CE-MRI in the setting of mammographically suspicious calcifications undergoing stereotactic biopsy and published between 2000 and 2020 were eligible. Two reviewers extracted study characteristics and true positives (TP), false positives, true negatives and false negatives (FN). Specificity, in this setting equaling the number of avoidable biopsies and FN rates were calculated. The maximum pre-test probability at which post-test probabilities of a negative CE-MRI met with BI-RADS benchmarks was determined by a Fagan nomogram. Random-effects models, I^2 -statistics, Deek's funnel plot testing and meta-regression were employed. P-values <0.05 were considered significant.

Results: Thirteen studies investigating 1414 lesions with a cancer prevalence of 43.6% (range: 22.7–66.9%) were included. No publication bias was found ($P = 0.91$). CE-MRI performed better in pure microcalcification studies compared to those also including associate findings ($P < 0.001$). In the first group, the pooled rate of avoidable biopsies was 80.6% (95%-CI: 64.6–90.5%) while the overall and invasive cancer FN rates were 3.7% (95%-CI: 1.2–6.2%) and 1.6% (95%-CI 0–3.6%), respectively. Up to a pre-test probability of 22%, the post-test probability did not exceed 2%.

Conclusion: A negative supplementary CE-MRI could potentially avoid 80.6% of unnecessary stereotactic biopsies in BI-RADS 4 microcalcifications at a cost of 3.7% missed breast cancers, 1.6% invasive. BI-RADS benchmarks for downgrading mammographic calcifications would be met up to a pretest probability of 22%.

© 2021 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Suspicious calcifications are a common finding in patients undergoing mammography [1]. These calcifications may indicate the presence of cancer and are often the only finding associated with one type of breast cancer, namely ductal carcinoma in situ (DCIS) [2]. A systematic review and meta-analysis reported varying malignancy rates in mammographic calcifications ranging between 6 and 82%, depending on their appearance on mammography [3]. Based on the varying numbers of malignancy most institutions

usually perform invasive procedures rather than a short-term imaging follow-up. Stereotactic vacuum-assisted breast biopsy (VABB) is considered safe and cost-effective to diagnose malignancy in suspicious microcalcifications as compared to open surgery [4,5]. In the United States of America annually about 1.6 Million women undergo breast biopsies with an estimated cost of US\$ 3.07 billion [6]. The majority of calcifications is benign [3]. Besides costs, other concerns are the psychological impact on patients and possible complications of radiographic evaluation of future mammograms [7].

An alternative diagnostic test providing better specificity and greater diagnostic confidence that can lower the health care expenditures and shorten the patient's diagnostic work-up is highly desired. Contrast enhanced magnetic resonance imaging (CE-MRI)

* Corresponding author.

E-mail address: pascal.baltzer@meduniwien.ac.at (P.A.T. Baltzer).

Abbreviations

BI-RADS	Breast Imaging – Reporting And Data System
CE-MRI	Contrast-Enhanced MRI
DCIS	Ductal Carcinoma In Situ
LR	Likelihood Ratio
MRI	Magnetic Resonance Imaging
TP, FN, FP, TN	true positive, false negative, false positive, true negative
VABB	Vacuum-Assisted Breast Biopsy

demonstrates high sensitivity and specificity for diagnosing breast cancer, in particular in non-calcified-lesions [8,9]. CE-MRI provides complementary diagnostic information due to the visualization of tissue vascularization on the grounds of neoangiogenesis, a hallmark of cancer that is present in invasive cancer as well as in DCIS [10,11]. Consequently, it may be assumed that mammographic calcifications with corresponding contrast-enhancement on MRI should be biopsied while biopsy may be omitted in lesions that do not demonstrate suspicious correlates on CE-MRI. Bennani-Baiti et al. demonstrated that CE-MRI of the breast is suitable for diagnosis of malignancy in suspicious mammographic calcifications [9]. Baltzer et al. demonstrated that CE-MRI of the breast can accurately distinguish between benign and malignant calcifications and may thus be helpful to reduce unnecessary breast biopsies [12]. Another study also demonstrated that CE-MRI is an accurate tool to further diagnose BI-RADS 4a and 4 b calcifications and may be helpful in avoiding unnecessary biopsies [13]. In clinical practice, it would be desirable to be able to potentially downgrade suspicious mammographic calcifications to BI-RADS 2 or 3 to omit needle guided breast biopsies.

The purpose of this study was to analyze the rate of potentially avoidable needle biopsies in women presenting with mammographically suspicious calcifications (BIRADS 4) if supplementary CE-MRI is negative by means of a systematic review and meta-analysis.

2. Materials and methods

2.1. Study design and eligibility criteria for study selection

This was a systematic review and meta-analysis of research studies on women in which dynamic CE-MRI of the breast was used for further characterization of mammographic calcifications suspicious for malignancy at screening mammography or in symptomatic patients that had a biopsy indication. Histopathologic sampling, preferably but not necessarily with additional imaging follow-up of at least 12 months was defined as the required reference standard for CE-MRI results. Raw data to construct a 2×2 table of true-positive, true-negative, false-positive and false-negative findings had to be provided within eligible articles. Studies in which authors focused on the exclusion of malignancy in biopsy-proven lesions with uncertain malignant potential for which open surgery was recommended were not included. Also, mixed cohorts including mammographic calcifications classified BI-RADS 3 in which the biopsy indication was questionable were not included as the research question was to determine the rate of potentially avoidable biopsies. For the same reason, articles reporting on pure BI-RADS 3 or BI-RADS 5 calcifications were considered not representative for the research question and excluded. Furthermore, studies performed before the year 2000

were not considered eligible to avoid a bias due to outdated MRI technology.

2.2. Search strategy

Two independent readers (BF, with 15 years of experience in oncologic imaging and PATB, with 17 years of experience in breast imaging) performed a systematic review of articles in the PubMed and Scopus databases published between January 1st 2000 to March 1st, 2020. The review protocol was defined before the start of the study and all relevant data are reported within this manuscript. Search terms were predefined as follows: “breast MRI microcalcifications”. A second search using a broader definition of “calcifications mammography” was also performed. Titles and abstracts of search results were analyzed for eligibility. The full texts of the eligible studies were retrieved. Research articles and review articles cited in selected articles were further reviewed for additional eligible studies (“backward snowballing”). A third reader (a breast imaging specialist with more than 10 years of experience) was appointed as an arbitrator to solve discrepancies in consensus.

2.3. Study selection, data collection, study quality and calculation of diagnostic estimates

Two independent readers (BF, PATB) selected eligible studies, extracted the data, assessed risk of bias and applicability concerns by means of the QUADAS-2 tool. Finally, calculations of the diagnostic estimates were performed. The extracted data included publication year, study, design (retrospective or prospective), indication for CE-MRI imaging, field strength, diagnostic criteria, presence of associated findings vs. calcifications only, reference standard (histopathology and follow-up examinations), and invasiveness of false-negative lesions. Furthermore, extraction of imaging results (true-positive, false-positive, true-negative, and false-negative) was performed. As all included mammographic calcification cases had a biopsy indication, all (including those referred to as e.g. BI-RADS 5 in the original publications) were considered BI-RADS 4 in line with the current BI-RADS lexicon. The BI-RADS classification as reported in the original studies is reported in the study characteristics tables for sake of transparency. If diagnostic performance indices were reported for different diagnostic criteria within a single study, we used “presence of enhancement” that has been described as more sensitive [9]. If no consensus on extracted criteria was reached, a third reader (PC) was assigned as an arbitrator.

2.4. Statistical analysis

Analysis were performed by using commercial (STATA, Statcorp, MIDAS plugin) and open source software (OpenMeta Analyst, <http://www.cebm.brown.edu>). Data from individual cross-tabulations were used to calculate pooled estimates of sensitivity, specificity, and negative predictive value. Data were pooled using both a bivariate random effects model of combined sensitivity and specificity and a hierarchical summary receiver operating characteristics (HSROC) model to calculate a summary Receiver Operating Characteristics (sROC) curve. The false negative rates for both overall malignancy and invasive cancers only were also pooled using random effects models. Between-study heterogeneity was investigated using Chi-Square statistics and the inconsistency I^2 was calculated and interpreted according to Ref. [14]. Publication bias was investigated using Deek’s test and a funnel plot of Diagnostic Odds Ratios. By random effects meta-regression, the influence of study characteristics on the diagnostic performance metrics

sensitivity, specificity and diagnostic odds ratio was investigated. Likelihood ratios derived from sensitivity and specificity at the summary operating point were used in a Fagan nomogram to provide post-test probabilities based on variable pre-test probabilities for clinical decision making. P-values below <0.05 were deemed to characterize significant findings for all calculations.

3. Results

3.1. Overall sensitivity and specificity of CE-MRI

The study selection procedure is given as a flowchart in supplemental figure 1. The literature search yielded 3774 non-duplicate results. Thirteen studies investigating 1414 (616 malignant and 798 benign) lesions with a cancer prevalence of 43.6% (range: 22.7–66.9%) matched the criteria and were included into the analysis [12,13,15–25]. Study specific characteristics are given in Table 1 and extracted data are summarized in Table 2. QUADAS-2 assessment is summarized in Fig. 1. Unclear risk of bias was present in >50% of the studies regarding flow and timing between CE-MRI and stereotactic biopsy and the majority of studies were assigned an unclear risk of bias due to lack of data regarding patient selection. There were no concerns regarding the applicability of the included studies to the research questions. No disagreement was found between the results of both readers.

The pooled sensitivity of CE-MRI for correctly detecting malignancy (DCIS and invasive cancer) in BI-RADS 4 lesions was 0.92, 95%-CI: 0.88–0.95 (bivariate model) and 0.91, 95%-CI: 0.84–0.95 (HSROC model). Sensitivities ranged from 0.75 [22] to 1.00 [16,24].

In this BI-RADS 4 setting, where all lesions were biopsied, specificity equals the rate of avoidable biopsies. Pooled specificity was 0.77, 95%-CI: 0.69–0.84 (bivariate model) and 0.74, 95%-CI: 0.66–0.82 (HSROC model). Specificity ranged from 0.25 [20]–0.95 [26]. The sROC plane containing the sROC curve and both single study and summary sensitivity and specificity is given in Fig. 2.

Table 1
Study characteristics.

First Author	Year	No. of Patients	Age (mean/range)	MR Imaging Field Strength (T)	Contrast Agent	Dose (mmol/kg)	Coil
Hrkac Pustahija	2018	113	55 (36–71)	1.5	Gadopentetate dimeglumine	0.1	Dedicated double breast coil
Eun NL	2018	108	50 (31–78)	1.5–3	Gadopentetate dimeglumine or gadobutrol	0.1	Dedicated double breast coil
Shimauchi A	2018	138	51.7 (33–74)	1.5	Gadopentetate dimeglumine	0.1	Dedicated double breast coil
Baltzer P	2018	152	57 (32–89)	1.5–3	Gadolinium chelate	0.1	NA
Bennani-Baiti	2017	248	60 (31–82)	1.5	Gadopentetate dimeglumine	0.1	Dedicated double breast coil
Strobel K	2015	78 ^a	53 (23–81)	1.5	NA	NA	Dedicated double breast coil
Li E	2014	84	46 (25–76)	3	Gadodiamide	0.1	Dedicated double breast coil
Zhu J	2007	52	NA (30–74)	1.5	Gadodiamide	0.1	Microscopy coil
Bazzocchi M	2006	112	NA	1–1.5	Gadoteritol	0.1	Dedicated double breast coil
Kneeshaw PJ	2006	88	58 (50–75)	1.5	Gadopentetate dimeglumine	0.1	Dedicated double breast coil
Bluemke D	2004	300 ^a	53 (18–80)	1.5	Gadolinium chelate	0.1	Dedicated double breast coil
Trecate G	2002	28	NA (33–65)	1.5	Gadopentetate dimeglumine	0.1	Dedicated double breast coil
Nakahara H	2001	40	49.5 (27–76)	0.5	Gadopentetate dimeglumine	0.1	Single breast coil

NA = not available.

[#] Data is mean ± standard deviation.

^o Data is median, with the range in parenthesis.

^a Subpopulation of the original study population fulfilling the eligibility criteria of the current study.

3.2. Threshold effect and publication bias

A negative correlation of logit (sensitivity) and logit (specificity) indicated a threshold effect (0.666, $P < 0.05$, see Fig. 2). Funnel plot analysis Deek’s testing did not reveal any Funnel plot asymmetry and did thus not point out a significant risk of publication bias ($p = 0.91$, Fig. 3).

3.3. Investigation of sources of heterogeneity

There was statistically significant ($P = 0.013$) moderate heterogeneity (I^2 53.6%) between the included studies. Multivariable meta-regression revealed “presence of associated findings” as the only factor significantly influencing the diagnostic performance of CE-MRI (diagnostic odds ratio, $P < 0.001$) while magnetic field strength, prevalence and diagnostic criteria were not associated with CE-MRI diagnostic performance ($P > 0.1$, respectively). Studies including pure mammographic calcifications [12,17,21,25,27,28] showed higher sensitivity (93.7%, 95%-CI: 89.3–96.3% versus 87.8%, 95%-CI: 79.3–93.1%), specificity (80.5%, 95%-CI: 64.9–90.2% versus 61.0%, 95%-CI: 41.6–77.5%), odds ratio (71.9, 95%-CI: 40.2–128.7 versus 16.1, 95%-CI: 10.0–25.7) and negative likelihood ratio (LR-0.07, 95%-CI: 0.04–0.12 versus 0.1, 95%-CI: 0.06–0.16) compared to those allowing also associated findings in their study population [18,23,24,27,29]. Heterogeneity in the subgroup investigating only mammographic calcifications was not significant ($P = 0.267$) and inconsistency was low (I^2 21.7%) and while it was not significant ($P = 0.155$) in studies allowing also associated findings while I^2 of 42.9% indicated moderate inconsistency. Notably, no further details on the nature of associated findings could be extracted from the original studies.

Neither MRI field strength, publication date (as a surrogate for MRI technical status) or prevalence of malignancy showed a significant association with diagnostic performance indices ($P > 0.05$, respectively).

Table 2
Key parameters extracted from the eligible studies.

First Author	Year	Diagnostic Criteria	BIRADS-Score	Associated Lesions	Reference Standard	Total Cases	TP	FP	FN	TN	FN _i	total _i
Hrkac Pustahija	2018	Enhancement	4,5	no	HE	125	48	23	0	54	0	17
Eun NL	2018	Enhancement	4,5	no	HE	108	37	8	4	59	0	10
Shimauchi A	2018	Enhancement	4	no	HE	32	15	5	3	9	0	3
Baltzer P	2018	BI-RADS	4,5	no	HE + FU	152	69	49	2	32	0	30
Bennani-Baiti	2017	BI-RADS	4,5	no	HE + FU	248	103	25	4	116	1	84
Strobel K	2015	BI-RADS	4,5	no	HE/FU	78	22	8	3	45	NA	NA
Li E	2014	BI-RADS	4,5	yes	HE	51	22	21	1	7	NA	NA
Zhu J	2007	BI-RADS	3–5	no	HE	52	23	2	3	24	1	9
Bazzocchi M	2006	Enhancement	4,5	yes	HE	112	69	12	6	25	2	42
Kneeshaw PJ	2006	BI-RADS	3–5	NA	HE	88	15	7	5	61	2	10
Bluemke D	2004	BI-RADS	3–5	yes	HE + FU	300	106	42	21	131	NA	NA
Trecate G	2002	Enhancement	4,5	yes	HE	28	15	5	0	8	0	7
Nakahara H	2001	Enhancement	4,5	no	HE	40	19	4	1	16	0	6

FU = follow up, Enhancement = presence of enhancement (positive if not absent or diffuse), FN = false negative result, FP = false positive result, HE = histopathologic examination, ;: invasive cancer, NA = not available, TN = true negative result, TP = true positive result.

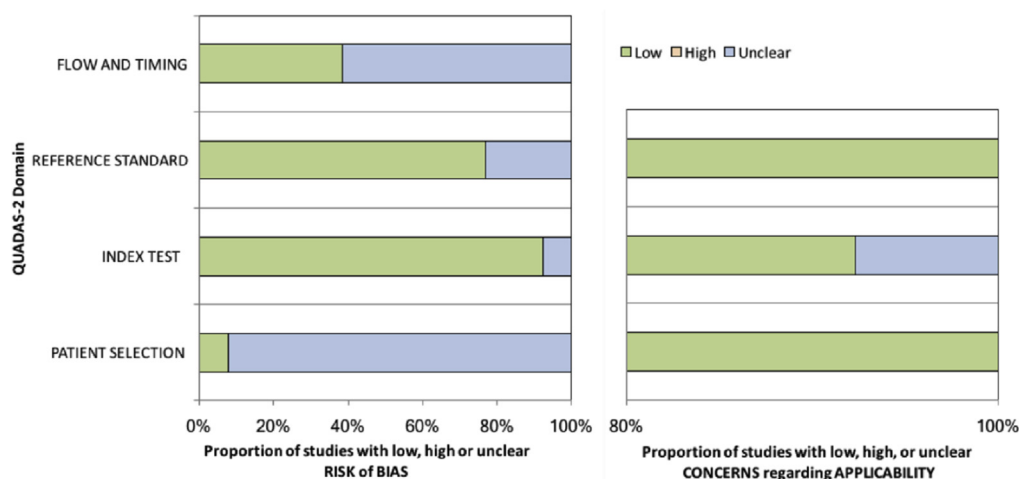


Fig. 1. QUADAS-2 study assessment.

3.4. Details of false-negative findings among invasive cancers and all malignant lesions

Ten studies provided data on the numbers of invasive cancer and DCIS subgroups [12,13,15–18,21,22,24,25]. The overall pooled rate of false-negative findings was 4.9% (95%-CI: 2.2%–7.7%). Among 441 malignant lesions, 28 were not characterized as breast cancer at CE-MRI (Fig. 4).

For invasive cancers, the pooled rate of false-negative findings was 2.1% (95%-CI: 0.2%–3.9%). Out of 218 invasive breast cancers, six were missed (Fig. 5).

In the subset of six studies investigating mammographic calcifications without associated findings [12,17,21,25,27,28], the overall pooled FN rate was lower with 3.7%, 95%-CI 1.2–6.2% and the FN rate for invasive cancers was 1.6%, 95%-CI 0–3.6%. Out of 159 invasive breast cancers, two were missed. No further subtype information on FN findings was available from the original studies.

3.5. Pretest probability of CE-MRI for ruling out malignancy (invasive cancer and DCIS)

To estimate the probability of CE-MRI of the breast and to rule out cancer in BI-RADS 4 lesions we applied the pooled positive and negative likelihood ratios to a Fagan nomogram.

Up to a pre-test probability of 17% the negative likelihood ratio

(0.10) of a negative CE-MRI results in a post-test probability of 2%, allowing to downgrade the lesion while meeting BI-RADS benchmarks. In the subset of pure mammographic calcifications without associated findings, the same goal was achieved up to a pre-test probability of 22% (LR- 0.07, see Fig. 6).

4. Discussion

Our results support the role of CE-MRI of the breast to distinguish benign from malignant BI-RADS 4 mammographic microcalcifications. Our level of evidence grade 1 results are based on a series of diagnostic studies with consistently applied reference standard (level of evidence grade 2) identified by systematic review. In summary, a negative CE-MRI result could downgrade a pooled rate of 76% of mammographic BI-RADS 4 lesions. By switching management to imaging surveillance rather than invasive diagnostic procedures, a significant number of biopsies could be avoided. Though VABB procedures are considered safe [30], minor complications including pain and scarring on the site of biopsy as well as psychological and physical pain due to the procedure could be averted [7]. Finally, costs for VABB in many western countries exceed those of CE-MRI, pointing out the opportunity to save costs of breast patient management [31], e.g. the current U.S. medicare (medicare.gov) shows that VABB costs (item 19,081) are four to five times higher than breast MRI costs (items 77,047 or

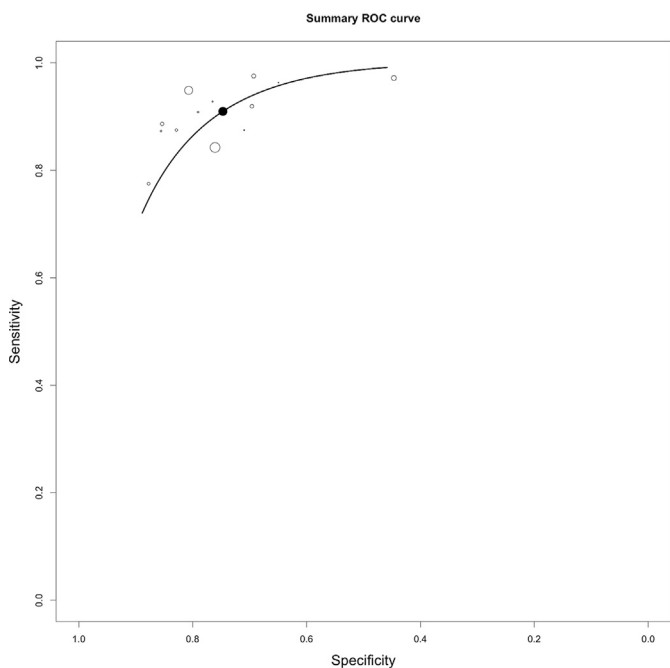


Fig. 2. Summary Receiver Operating Characteristics curve (black line) calculated by the HSROC model. White circles represent single study data while the black circle represents the sensitivity/specificity pair at the summary operating point.

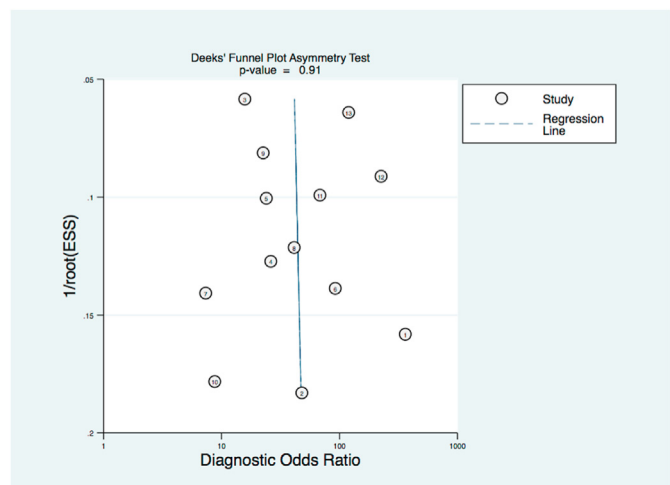


Fig. 3. Funnel plot of CE-MRI results in mammographic BI-RADS 4 microcalcifications.

C8908).

In diagnostic tests, there is a trade-off between sensitivity and specificity: increasing specificity to avoid unnecessary biopsies may decrease sensitivity and thereby increase the risk of missing cancers. According to BI-RADS, a rate of up to 2% malignant cases can be followed-up rather than referred to biopsy. The overall false negative rate in this study exceeded 2% and thus did not meet BI-RADS 3 requirements. However, the 2.1% rate of false-negative invasive cancers was close to the BI-RADS 3 benchmark of 2% [32,33]. These numbers improved when considering the patient subgroup presenting with pure mammographic calcifications without additional findings. The progression rate of DCIS to invasive breast cancer has been reported between 18 and 50% within the next five to ten years. In addition, the sensitivity profile of breast MRI as compared to mammography is towards biologically

significant cancers, meaning that cancers missed by CE-MRI show a less aggressive biological behavior [34]. Considering these facts, we consider a follow-up exam oncologically acceptable. Mammographic follow-up could detect a possible underlying malignant disease in MRI-negative cases within a short to intermediate follow-up time frame (6–12 months) before a significant progression occurs. This approach does have additional downsides: short-term mammography follow-up has been associated with psychosocial harms [35]. Though women are willing to accept even the harms of invasive additional diagnostic procedures for the sake of earlier breast cancer diagnosis [36] it is possible that many women would choose immediate diagnostic certainty by invasive diagnostic procedures over the non-invasive active surveillance approach. A longer follow-up interval of 12 or even 24 months could be considered safer in terms of psychological harm (as it implies a higher diagnostic certainty) but its oncological safety needs to be proven empirically.

The BI-RADS 4 category covers a wide range of likelihood of malignancy, from more than 2% to less than 95%. Since 2003, BI-RADS has suggested the use of BI-RADS 4 category subdivisions to provide improved stratification of likelihood of malignancy [32]. While this suggests the substantial number of potentially avoidable biopsies also found in this study, it further provides us with a possibility to stratify the risk of malignancy in mammographic calcifications using BI-RADS lexicon criteria [3,32]. Thereby, we could predict whether a negative breast MRI result would result in probabilities of malignancy low enough to formally downgrade a lesion to BI-RADS 3 and suggest mammographic follow-up rather than stereotactic biopsy. To further investigate this topic, a Fagan nomogram was used to clarify whether CE-MRI of the breast as an additional test to rule-out malignancy could be applied to BI-RADS 4 lesions with low, intermediate or high likelihood of malignancy. According to our results, a negative CE-MRI could be used to downgrade mammographic BI-RADS 4 lesions with a pretest probability up to 17%. This pretest probability could be determined by malignancy rates associated with BI-RADS features as suggested by the literature [3] which have also been combined into multi-variable risk scores [37]. Recently, also machine-learning methods have been employed for risk stratification of breast calcifications [38]. The lack of validation studies and the encouraging results reported here suggest a prospective trial on the clinical use of using breast MRI as an additional test in mammographic BI-RADS 4 microcalcifications to avoid unnecessary biopsies.

A prior meta-analysis on the use of CE-MRI for diagnosis of malignancy in mammographic microcalcifications was published in 2016 [9]. The authors investigated general diagnostic performance indices in 20 studies using CE-MRI of the breast as an additional test in mammographic microcalcifications classified BI-RADS 3–5. In a subset of four studies investigating BI-RADS 4 calcifications, an NPV for malignant lesions of 92% was reported and the rate of avoidable biopsies was not assessed. Our more specific study selection led to the inclusion of 13 studies, five of them published after 2016 [12,17,18,27,28].

The main limitation of our analysis is high heterogeneity between the analyzed studies. As a main factor, diagnostic outcomes were influenced by the presence of associated findings. Notably, the FN rates were lower in cases of pure mammographic microcalcifications without any associated findings. In this subgroup, heterogeneity was low and insignificant, thereby stressing the clinical applicability of our findings. Still, we observed sparse information regarding two further factors on CE-MRI diagnostic performance: first, details on actual lesion size and BI-RADS lexicon characteristics of the investigated microcalcifications were not provided within the included studies. Differences in these criteria could point out heterogeneous population characteristics between

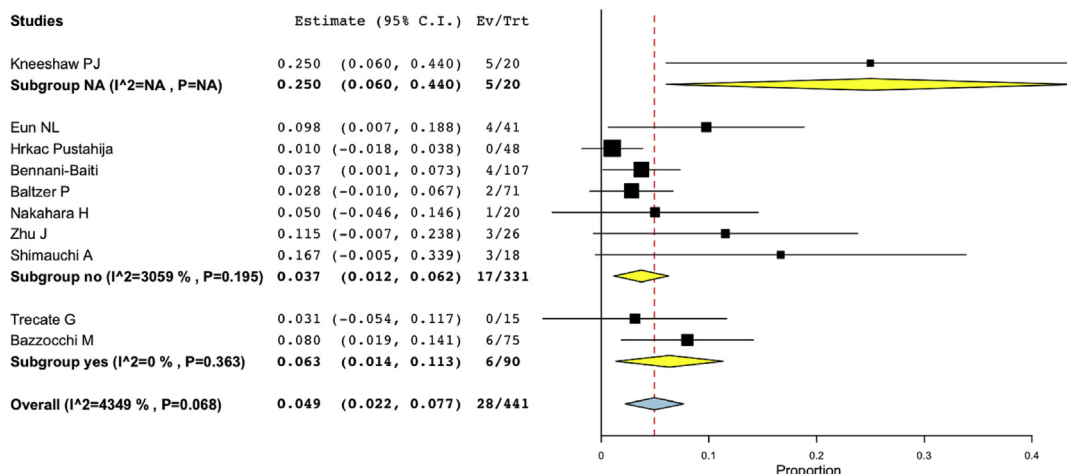


Fig. 4. Rate of false-negative findings in all malignant lesions (invasive cancer and DCIS) in the subset of ten CE-MRI studies providing details on invasive cancer and DCIS subgroups stratified by associated findings (NA, no, yes). Note that I²-numbers are given as percentage (first two digits) and decimals (last two digits), e.g. 4090 is 40.9%.

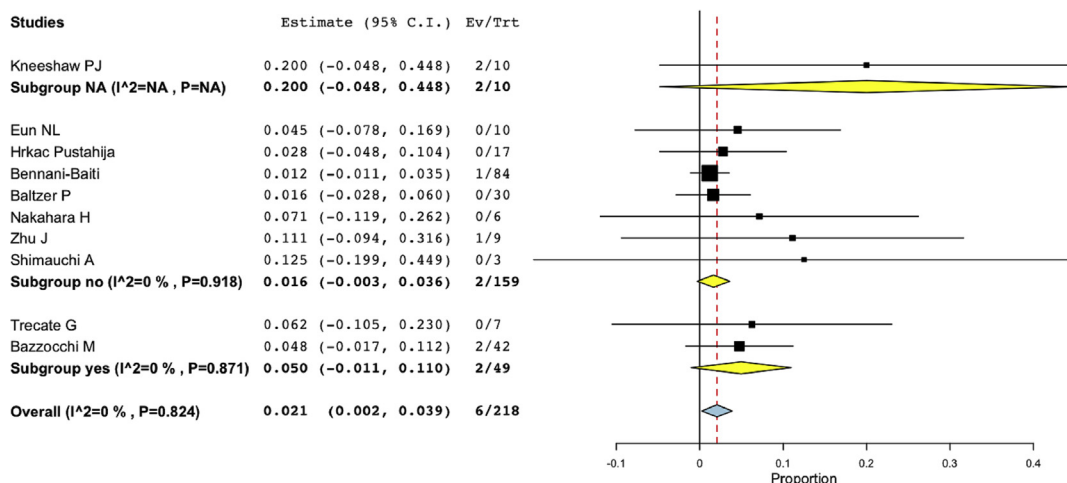


Fig. 5. Rate of false-negative invasive cancers in the subset of ten CE-MRI studies providing details on invasive cancer and DCIS subgroups. Results are stratified by associated findings (NA, no, yes). Note that I²-numbers are given as percentage (first two digits) and decimals (last two digits), e.g. 4090 is 40.9%.

the included studies. Consequently, we could not calculate subgroup results for BI-RADS 4a, b or c calcifications. A further limitation is the lack of subtype information on the FN cases that limit the assessment of their biological aggressiveness. Though no systematic selection bias leading to a publication bias was found, patient selection bias with a tendency towards clinically more suspicious calcifications is likely as evidenced by the rather high prevalence of malignancy in the investigated studies. As predictive values depend on cancer prevalence, our NPV is possibly underestimated and we suggest to use the calculated likelihood ratio and the Fagan nomogram to assess the clinical applicability of the reported findings. Second, CE-MRI can be performed in many ways using different equipment, protocols and diagnostic criteria as another potential source of heterogeneity [39]. Again, the lack of systematic bias corroborates the general validity and applicability of our findings. However, to fully understand which specific patients would profit from additional CE-MRI workup and which would not, prospective studies are required. The results presented here strongly support the conduct of such research. Such research could also include a formal cost-comparison or cost-effectiveness analysis which was not done in this paper. Such analysis should also comprise an analysis of incidental lesions aside from the

mammographic calcifications that require additional follow-up or biopsy procedures and thus may reduce the rate of potentially avoidable biopsies. Also, adding a highly sensitive test such as MRI yields the risk of diagnosing additional, though biologically insignificant cancers that cannot be estimated based on currently available data. Comparative research with conventional methods suggests a higher inherent risk of MRI-detected false positive findings [40] but no solid data regarding clinical outcomes is available. Finally, there have been concerns against the use of gadolinium-based contrast media due to potential gadolinium deposits in the brain. Recent research in women undergoing repeated exposition of currently recommended macrocyclic contrast media for the purpose of high-risk breast cancer screening has refuted the hypothesis that otherwise healthy women are at risk of such deposits [41].

In conclusion, this meta-analysis generally supports the feasibility of CE-MRI of the breast to downgrade BI-RADS 4 mammographic microcalcifications and therefore avoid unnecessary stereotactic biopsies. Further prospective studies to identify which patients would profit from additional CE-MRI workup and which would not.

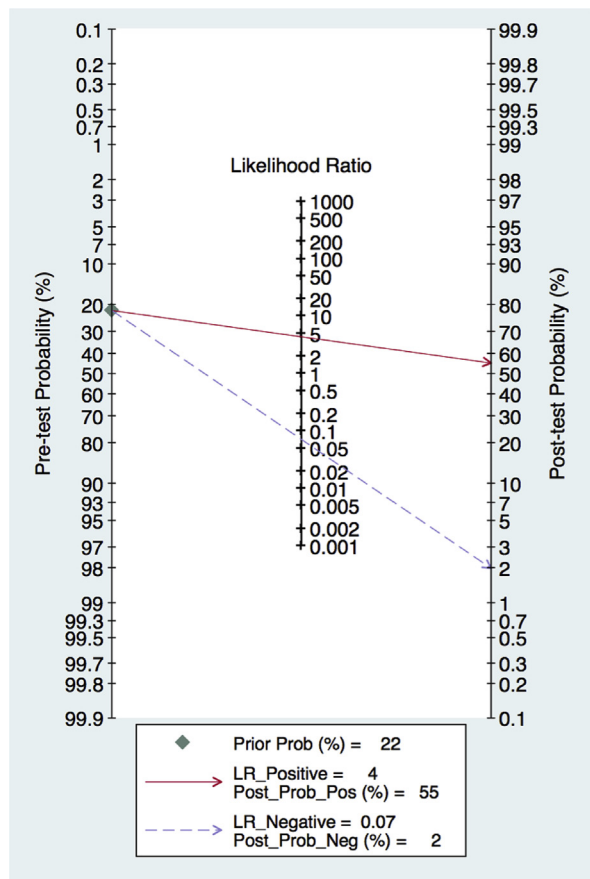


Fig. 6. Fagan nomogram providing post-test probabilities for both positive (red line) and negative (dashed blue line) CE-MRI results in the subgroup of pure mammographic BI-RADS 4 calcifications. The pre-test probability of 22% is the maximum pre-test probability at which a negative CE-MRI result yields a post-test probability that does not exceed the 2% cancer rate of BI-RADS 3, therefore allowing to formally downgrade the investigated mammographic microcalcification. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Declaration of competing interest

None of the authors have any conflicts of interest to declare.

Acknowledgement

This research was conducted in association with a funded project from the Austrian FWF (Project number: I 4240, PI Pascal A.T. Baltzer).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.breast.2021.02.002>.

References

[1] Bluekens AMJ, Holland R, Karssemeijer N, Broeders MJM, den Heeten GJ. Comparison of digital screening mammography and screen-film mammography in the early detection of clinically relevant cancers: a multicenter study. *Radiology* 2012;265(3):707–14. <https://doi.org/10.1148/radiol.12111461>.
 [2] Castronovo V, Bellahcene A. Evidence that breast cancer associated microcalcifications are mineralized malignant cells. *Int J Oncol* 1998;12(2):305–8. <https://doi.org/10.3892/ijo.12.2.305>.
 [3] Rominger M, Wisgickl C, Timmesfeld N. Breast microcalcifications as type descriptors to stratify risk of malignancy: a systematic review and meta-

analysis of 10665 cases with special focus on round/punctate microcalcifications. *RöFo - Fortschritte Auf Dem Geb Röntgenstrahlen Bildgeb Verfahr* 2012;184(12):1144–52. <https://doi.org/10.1055/s-0032-1313102>.
 [4] Liberman L, Sama MP. Cost-effectiveness of stereotactic 11-gauge directional vacuum-assisted breast biopsy. *Am J Roentgenol. American Roentgen Ray Society* 2000;175(1):53–8. <https://doi.org/10.2214/ajr.175.1.1750053>.
 [5] Tsai H-Y, Huang S-T, Chao M-F, et al. Cost-effectiveness of stereotactic vacuum-assisted biopsy for nonpalpable breast lesions. *Eur J Radiol* 2020;127:108982. <https://doi.org/10.1016/j.ejrad.2020.108982>.
 [6] Vlahiotis A, Griffin B, Stavros AT, Margolis J. Analysis of utilization patterns and associated costs of the breast imaging and diagnostic procedures after screening mammography. *Clin. Outcomes Res. Dove Press*; 2018. p. 157–67. <https://doi.org/10.2147/CEOR.S150260>.
 [7] Nelson HD, O'Meara ES, Kerlikowske K, Balch S, Miglioretti D. Factors associated with rates of false-positive and false-negative results from digital mammography screening: an analysis of registry data. *Ann Intern Med.* American College of Physicians 2016;164(4):226–35. <https://doi.org/10.7326/M15-0971>.
 [8] Bennani-Baiti B, Bennani-Baiti N, Baltzer PA. Diagnostic performance of breast magnetic resonance imaging in non-calcified equivocal breast findings: results from a systematic review and meta-analysis. *PLoS One* 2016;11(8):e0160346. <https://doi.org/10.1371/journal.pone.0160346>.
 [9] Bennani-Baiti B, Baltzer PA. MR imaging for diagnosis of malignancy in mammographic microcalcifications: a systematic review and meta-analysis. *Radiology* 2016;283(3):692–701. <https://doi.org/10.1148/radiol.2016161106>.
 [10] Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100(1):57–70. [https://doi.org/10.1016/S0092-8674\(00\)81683-9](https://doi.org/10.1016/S0092-8674(00)81683-9).
 [11] Engels K, Fox SB, Whitehouse RM, Gatter KC, Harris AL. Distinct Angiogenic patterns are associated with high-grade in situ ductal carcinomas of the breast. *J Pathol* 1997;181(2):207–12. [https://doi.org/10.1002/\(SICI\)1096-9896\(199702\)181:2<207::AID-PATH758>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1096-9896(199702)181:2<207::AID-PATH758>3.0.CO;2-4).
 [12] Baltzer PAT, Bennani-Baiti B, Stöttinger A, Bumberger A, Kapetas P, Clauser P. Is breast MRI a helpful additional diagnostic test in suspicious mammographic microcalcifications? *Magn Reson Imaging* 2018;46:70–4. <https://doi.org/10.1016/j.mri.2017.10.012>.
 [13] Bennani-Baiti B, Dietzel M, Baltzer PA. MRI for the assessment of malignancy in BI-RADS 4 mammographic microcalcifications. *PLoS One* 2017;12(11):e0188679. <https://doi.org/10.1371/journal.pone.0188679>.
 [14] Cochrane handbook for systematic reviews of interventions. /handbook/current. Accessed January 25, 2021.
 [15] Bazzocchi M, Zuiani C, Panizza P, et al. Contrast-enhanced breast MRI in patients with suspicious microcalcifications on mammography: results of a multicenter trial. *AJR Am J Roentgenol* 2006;186(6):1723–32. <https://doi.org/10.2214/AJR.04.1898>.
 [16] Hrkac Pustahija A, Ivanac G, Brkljacic B. US and MRI in the evaluation of mammographic BI-RADS 4 and 5 microcalcifications. *Diagn Interv Radiol Ank Turk* 2018;24(4):187–94. <https://doi.org/10.5152/dir.2018.17414>.
 [17] Eun NL, Son EJ, Gweon HM, Youk JH, Kim J-A. The value of breast MRI for BI-RADS category 4B mammographic microcalcification: based on the 5th edition of BI-RADS. *Clin Radiol* 2018;73(8):750–5. <https://doi.org/10.1016/j.crad.2018.04.014>.
 [18] Shimauchi A, Machida Y, Maeda I, Fukuma E, Hoshi K, Tozaki M. Breast MRI as a problem-solving study in the evaluation of BI-RADS categories 3 and 4 microcalcifications: is it worth performing? *Acad Radiol* 2018;25(3):288–96. <https://doi.org/10.1016/j.acra.2017.10.003>.
 [19] Strobel K, Schrading S, Hansen NL, Barabasch A, Kuhl CK. Assessment of BI-RADS category 4 lesions detected with screening mammography and screening US: utility of MR imaging. *Radiology* 2015;274(2):343–51. <https://doi.org/10.1148/radiol.14140645>.
 [20] Li E, Li J, Song Y, Xue M, Zhou C. A comparative study of the diagnostic value of contrast-enhanced breast MR imaging and mammography on patients with BI-RADS 3-5 microcalcifications. *PLoS One* 2014;9(11):e111217. <https://doi.org/10.1371/journal.pone.0111217>.
 [21] Zhu J, Kurihara Y, Kanemaki Y, et al. Diagnostic accuracy of high-resolution MRI using a microscopy coil for patients with presumed DCIS following mammography screening. *J Magn Reson Imaging JMRI* 2007;25(1):96–103. <https://doi.org/10.1002/jmri.20809>.
 [22] Kneeshaw PJ, Lowry M, Manton D, Hubbard A, Drew PJ, Turnbull LW. Differentiation of benign from malignant breast disease associated with screening detected microcalcifications using dynamic contrast enhanced magnetic resonance imaging. *Breast Edinb Scotl* 2006;15(1):29–38. <https://doi.org/10.1016/j.breast.2005.05.002>.
 [23] Bluemke DA, Gatsonis CA, Chen MH, et al. Magnetic resonance imaging of the breast prior to biopsy. *J Am Med Assoc* 2004;292(22):2735–42. <https://doi.org/10.1001/jama.292.22.2735>.
 [24] Trecate G, Tess JD, Vergnaghi D, et al. Breast microcalcifications studied with 3D contrast-enhanced high-field magnetic resonance imaging: more accuracy in the diagnosis of breast cancer. *Tumori* 2002;88(3):224–33.
 [25] Nakahara H, Namba K, Fukami A, et al. Three-dimensional MR imaging of mammographically detected suspicious microcalcifications. *Breast Cancer Tokyo Jpn* 2001;8(2):116–24.
 [26] Nakahara H, Namba K, Fukami A, et al. Three-dimensional MR imaging of mammographically detected suspicious microcalcifications. *Breast Cancer Tokyo Jpn* 2001;8(2):116–24.
 [27] Bennani-Baiti B, Dietzel M, Baltzer PA. MRI for the assessment of malignancy

- in BI-RADS 4 mammographic microcalcifications. *PloS One* 2017;12(11): e0188679. <https://doi.org/10.1371/journal.pone.0188679>.
- [28] Hrkac Pustahija A, Ivanac G, Brkljacic B. US and MRI in the evaluation of mammographic BI-RADS 4 and 5 microcalcifications. *Diagn Interv Radiol Ank Turk* 2018;24(4):187–94. <https://doi.org/10.5152/dir.2018.17414>.
- [29] Bazzocchi M, Zuiani C, Panizza P, et al. Contrast-enhanced breast MRI in patients with suspicious microcalcifications on mammography: results of a multicenter trial. *AJR Am J Roentgenol* 2006;186(6):1723–32. <https://doi.org/10.2214/AJR.04.1898>.
- [30] Lin LLY, Gao Y, Lewin AA, Toth HK, Heller SL, Moy L. Overstated harms of breast cancer screening? A large outcomes analysis of complications associated with 9-gauge stereotactic vacuum-assisted breast biopsy. *AJR Am J Roentgenol* 2019;212(4):925–32. <https://doi.org/10.2214/AJR.18.20421>.
- [31] Wengert GJ, Pipan F, Almohanna J, et al. Impact of the Kaiser score on clinical decision-making in BI-RADS 4 mammographic calcifications examined with breast MRI. *Eur Radiol* 2020;30(3):1451–9. <https://doi.org/10.1007/s00330-019-06444-w>.
- [32] D'Orsi C, Sickles E, Mendelson E, Morris E. *ACR BI-RADS® atlas, breast imaging reporting and data System*. American College of Radiology; 2013.
- [33] Spick C, Bickel H, Polanec SH, Baltzer PA. Breast lesions classified as probably benign (BI-RADS 3) on magnetic resonance imaging: a systematic review and meta-analysis. *Eur Radiol* 2018;28(5):1919–28. <https://doi.org/10.1007/s00330-017-5127-y>.
- [34] Sung JS, Stamler S, Brooks J, et al. Breast cancers detected at screening MR imaging and mammography in patients at high risk: method of detection reflects tumor histopathologic results. *Radiology* 2016;280(3):716–22. <https://doi.org/10.1148/radiol.2016151419>.
- [35] Bolejko A, Hagell P, Wann-Hansson C, Zackrisson S. Prevalence, long-term development, and predictors of psychosocial consequences of false-positive mammography among women attending population-based screening. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 2015;24(9):1388–97. <https://doi.org/10.1158/1055-9965.EPI-15-0060>.
- [36] Mathioudakis AG, Salakari M, Pylkkanen L, et al. Systematic review on women's values and preferences concerning breast cancer screening and diagnostic services. *Psycho Oncol* 2019;28(5):939–47. <https://doi.org/10.1002/pon.5041>.
- [37] Youk JH, Gweon HM, Son EJ, Eun NL, Choi EJ, Kim JA. Scoring System to stratify malignancy risks for mammographic microcalcifications based on breast imaging reporting and data System 5th edition descriptors. *Korean J Radiol* 2019;20(12):1646–52. <https://doi.org/10.3348/kjr.2019.0262>.
- [38] Stelzer PD, Steding O, Raudner MW, Euller G, Clauser P, Baltzer PaT. Combined texture analysis and machine learning in suspicious calcifications detected by mammography: potential to avoid unnecessary stereotactical biopsies. *Eur J Radiol* 2020;132:109309. <https://doi.org/10.1016/j.ejrad.2020.109309>.
- [39] Clauser P, Mann R, Athanasiou A, et al. A survey by the European Society of Breast Imaging on the utilisation of breast MRI in clinical practice. *Eur Radiol* 2018;28(5):1909–18. <https://doi.org/10.1007/s00330-017-5121-4>.
- [40] Kuhl CK, Keulers A, Strobel K, Schneider H, Gaisa N, Schrading S. Not all false positive diagnoses are equal: on the prognostic implications of false-positive diagnoses made in breast MRI versus in mammography/digital tomosynthesis screening. *Breast Cancer Res* 2018;20(1):13. <https://doi.org/10.1186/s13058-018-0937-7>.
- [41] Bennani-Baiti B, Krug B, Giese D, et al. Evaluation of 3.0-T MRI brain signal after exposure to gadoterate meglumine in women with high breast cancer risk and screening breast MRI. *Radiology* 2019;293(3):523–30. <https://doi.org/10.1148/radiol.2019190847>.