



Diagnostic Performance of Breast Magnetic Resonance Imaging in Non-Calcified Equivocal Breast Findings: Results from a Systematic Review and Meta-Analysis

Barbara Bennani-Baiti^{1,2©}*, Nabila Bennani-Baiti³, Pascal A. Baltzer^{2©}*

- 1 Department of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria, 2 Department of Biomedical Imaging and Image-guided Therapy, Vienna General Hospital (AKH), Medical University of Vienna, Vienna, Austria, 3 Division of Hematology, Mayo Clinic, Rochester, Minnesota, United States of America
- These authors contributed equally to this work.
- * barbara.bennani-baiti@meduniwien.ac.at (BBB); pascal.baltzer@meduniwien.ac.at (PAB)

Abstract

Objectives

To evaluate the performance of MRI for diagnosis of breast cancer in non-calcified equivocal breast findings.

Materials and Methods

We performed a systematic review and meta-analysis of peer-reviewed studies in PubMed from 01/01/1986 until 06/15/2015. Eligible were studies applying dynamic contrast-enhanced breast MRI as an adjunct to conventional imaging (mammography, ultrasound) to clarify equivocal findings without microcalcifications. Reference standard for MRI findings had to be established by histopathological sampling or imaging follow-up of at least 12 months. Number of true or false positives and negatives and other characteristics were extracted, and possible bias was determined using the QUADAS-2 applet. Statistical analyses included data pooling and heterogeneity testing.

Results

Fourteen out of 514 studies comprising 2,316 lesions met our inclusion criteria. Pooled diagnostic parameters were: sensitivity (99%, 95%-CI: 93–100%), specificity (89%, 95%-CI: 85–92%), PPV (56%, 95%-CI: 42–70%) and NPV (100%, 95%-CI: 99–100%). These estimates displayed significant heterogeneity (P<0.001).

Conclusions

Breast MRI demonstrates an excellent diagnostic performance in case of non-calcified equivocal breast findings detected in conventional imaging. However, considering the





Citation: Bennani-Baiti B, Bennani-Baiti N, Baltzer PA (2016) Diagnostic Performance of Breast Magnetic Resonance Imaging in Non-Calcified Equivocal Breast Findings: Results from a Systematic Review and Meta-Analysis. PLoS ONE 11(8): e0160346. doi:10.1371/journal.pone.0160346

Editor: Kenneth G.A. Gilhuijs, Universitair Medisch Centrum Utrecht, NETHERLANDS

Received: April 3, 2016

Accepted: July 18, 2016

Published: August 2, 2016

Copyright: © 2016 Bennani-Baiti et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This article was supported by the Open Access Publishing Fund of the University of Vienna.

Competing Interests: The authors have declared that no competing interests exist.



substantial heterogeneity with regard to prevalence of malignancy, problem solving criteria need to be better defined.

Introduction

Breast is the most frequently affected organ by cancer in women [1]. Imaging plays a major role in secondary and tertiary prevention of breast cancer. Depending on whether healthy women are screened for breast cancer or whether assessment of symptomatic patients or screening findings is performed, mammography, breast ultrasound and percutaneous imageguided biopsies play a major role in diagnosis and to rule-out cancer [2-4]. However, these methods individually or in combination can yield inconclusive results, whereby the presence or absence of breast cancer is not clearly ascertained. Not everyone agrees on what qualifies as an equivocal finding. In clinical practice, a variety of results are usually classified as such: asymmetry without associated microcalcifications, architectural distortions and other ambiguous abnormalities such as multiple lesions, discrepancy between clinical symptoms and imaging findings, benign biopsy results with insufficient radiological-pathological concordance, lesions that could not be sufficiently localized during biopsy attempts, as well as scars. In these instances, an additional imaging-based diagnostic test would be most welcome. Microcalcifications are considered less problematic, since these lesions can be visualized by mammography and the workup of these lesions either by biopsy or follow-up imaging does usually not require additional imaging modalities.

MRI is considered by most to be one of the most sensitive imaging modalities for the detection of breast cancer [5,6]. Thus, breast MRI has for instance been shown to be a good imaging modality to exclude advanced nodal disease and to be helpful in the differential diagnosis of architectural distortions [7,8]. However, the effectiveness of breast MRI as a problem-solving tool remains controversial. Regular concerns are that the high sensitivity of MRI may not be high enough to rule-out malignancy and that MRI may associate with a high number of false positive findings as reflected by a low positive predictive value. As a result, there currently is a lack of clear recommendations on the application of MRI to resolving breast-imaging equivocality.

Consequently, the purpose of this systematic review and meta-analysis was to evaluate the performance of breast MRI for diagnosis of breast cancer in non-calcified equivocal breast findings. Or, put otherwise: can breast MRI rule-in or rule-out malignancy in non-calcified equivocal breast findings?

Materials and Methods

Search strategy

Two authors (BBB, PAB), one of them with 13 years of clinical experience in breast MRI, independently performed a systematic query of all full-text articles in the openly accessible PubMed database from 01/01/1986 up to 06/15/2015 (www.ncbi.nlm.nih.gov/pubmed/). Search terms were predefined as "breast MRI BI-RADS 0", "breast MRI BI-RADS 3", "breast MRI problem-solving" and "breast MRI equivocal". A separate search was performed for each search term combination as indicated by the quotation marks above. Resulting titles/abstracts were analyzed for eligibility and full texts were retrieved, accordingly. Since no specific MeSH terms for this systematic literature study were identified, additional results were obtained by backward



snowballing [9]. Results at every step were compared and discrepancies solved in a consensus review. If no consensus was reached, a third reader (NBB) served as an arbitrator.

Eligibility criteria for study selection

Eligible were peer-reviewed studies applying the index test, dynamic contrast-enhanced breast MRI, as an adjunct to conventional imaging (mammography, ultrasound) to clarify unequivocal findings without microcalcifications in at least 20 human subjects. The reference standard for index test findings was defined as presence of histopathological sampling or imaging follow-up of at least 12 months. A diagnosis of cancer by the reference standard was considered a positive finding, absence of cancer as a negative finding. No language restrictions were applied.

Data extraction and quality assessment

Two authors (BBB, PAB) independently extracted the following data: publication year, study design (retrospective/prospective), patient number and demographics, whether patient recruitment was consecutive, and indications for MRI examinations. Moreover, technical MRI parameters (field strength, coil, contrast medium dosage, and whether fat saturation was applied) were also collected. Index test (breast MRI) and reference standard (histopathology, follow-up) data were retrieved to fit a cross-tabulation with true or false positives (TP, FP) and negatives (TN, FN). Imaging results were called positive in case of BI-RADS 4 or 5 and negative in case of BI-RADS 1, 2 or 3. In case only a subpopulation of a study fulfilled the eligibility criteria, these specific data were extracted. Quality of studies and likelihood of bias were independently (BBB, PAB) evaluated, using *Quality Assessment of Diagnostic Accuracy Assessment* (QUADAS-2), an applet that assesses risk and bias in patient spectrum, reference standard, disease progression, verification, clinical review, incorporation, test execution, study withdrawals, and indeterminate results [10]. Any disparities in the findings were resolved by consensus. If no consensus was reached, a third reader (NBB) served as an arbitrator.

Study outcome

Diagnostic parameters of breast MRI were defined as the study outcome. These were: sensitivity (TP/(TP+FN)), specificity (TN/(TN+FP)), positive predictive value, PPV (TP/(TP+FP)) and negative predictive value, NPV (TN/(TN+FN)). The influence of a series of covariates was determined by subgroup analysis.

Statistical analysis

Analyses were performed using STATA 13.0 (StataCorp, College Station, TX, USA) and Open-MetaAnalyst 12.11.14 (http://www.cebm.brown.edu/open_meta/download). The possible presence of publication bias was further assessed by use of Begg's funnel scatterplot and calculated by Egger's test. In this setup, the plot allowed us to probe both for bias and systematic heterogeneity as a function of study size, whereby a symmetrical inverted funnel shape denotes a largely unbiased dataset [11].

Once our concerns of data bias were addressed, we used data from individual cross-tabulations to construct forest plots for sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Data heterogeneity was analyzed using Cochran's Q and I-squared statistics. We calculated to that effect Q as the sum of squared differences between the effects of individual studies and those pooled across all studies. Since Q can be affected by sampling size, we also calculated I^2 , the latter being thought to represent the percent of variability that it is relatively independent from sampling errors, and defined as $I^2 = \left(\frac{Q-df}{Q}\right)x$ 100%,



wherein *df* represents Q's degrees of freedom [12]. Pooled estimates for breast cancer prevalence, sensitivity, specificity, PPV, NPV, and likelihood ratios were calculated by applying random effects models or bivariate analyses using maximum likelihood estimates, as appropriate. In this study setting, the likelihood ratio can be used to calculate the post-test odds from the pre-test odds of breast cancer. In addition, a bivariate summary Receiver Operating Characteristics (sROC) curve was calculated and meta-regression using random effects models was used to investigate sources of heterogeneity. To put the results into context with Bayes' theorem, a Fagan's nomogram and probability-modifying plots for positive and negative breast MRI results were constructed.

Results

Study design and reporting

Out of 514 peer-reviewed studies, we identified fourteen eligible reports comprising 2,316 cases [13–26]. In five of these fourteen studies, we extracted only data of the eligible subgroups [17,18,21,22,26]. This meta-analysis adheres to the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*, and Fig 1 depicts a PRISMA flow chart summarizing the selection process. The completed PRISMA checklist is given in S1 Table. Three of the included fourteen studies were prospective [14,20,26]. The remaining studies focused on retrospectively reviewing the diagnostic accuracy of breast MRI in a problem-solving population. Indications for MRI were described as required for further evaluation of unclear or suspicious findings, with or without questionable correlates in conventional imaging. Studies included primarily unclear BI-RADS 0 or 3 finding analogs, whereas some studies also included lesions labeled as BI-R-ADS 4. In the latter, however, the lesions could not be localized for biopsy and as such qualify as equivocal lesions [13,18,23]. Four studies indicated detailed reasons for referral to breast MRI (e.g. the exact type of clinical or imaging findings) [15,23,25,26], two of which stratified breast MRI results by indication [25,26]. The remaining ten studies did not specify detailed indications for breast MRI referral [13,16–22,24].

Nineteen studies were deemed not eligible due to various reasons, including lack of peer review, lack of extractable raw data, studies with too small of a sample-size (n<20), those focusing on the exclusion of malignancy in biopsy-proven lesions with uncertain malignant potential (B3) and recommended open surgery, those only aiming at diagnosing suspicious mammographic microcalcifications, or those limited to certain risk groups or specific clinical questions [27–45] (Fig 1).

All MRI technical data are summarized in <u>S2 Table</u>. QUADAS-2 analysis of risk of bias and applicability assessment identified a likely patient selection related bias in four studies: one recruited patients conditional on palpable findings [22], one on lesion size [20], and two considered solitary conventional BI-RADS categories only [14,26]. An unclear patient selection bias was assigned in another study due to restriction to two BI-RADS categories [17], and an unclear bias was also assigned in two studies since a proportion of the patients were lost to imaging follow-up [19,24]. Regarding applicability of the analyzed studies to the research question, bias was deemed low in the majority of studies. An assignment of unclear applicability bias was assigned to the two studies with patients lost to follow-up [19,23], and to one study investigating only BI-RADS 4 findings [26]. Detailed QUADAS-2 assessment results are given in S1 Fig.

Finally, a Begg's funnel plot analysis revealed a symmetrical distribution, indicating a lack of publication bias ($\underline{\text{Fig 2}}$). The complimentary Egger's test also did not show a significant risk of publication bias (p = 0.12).

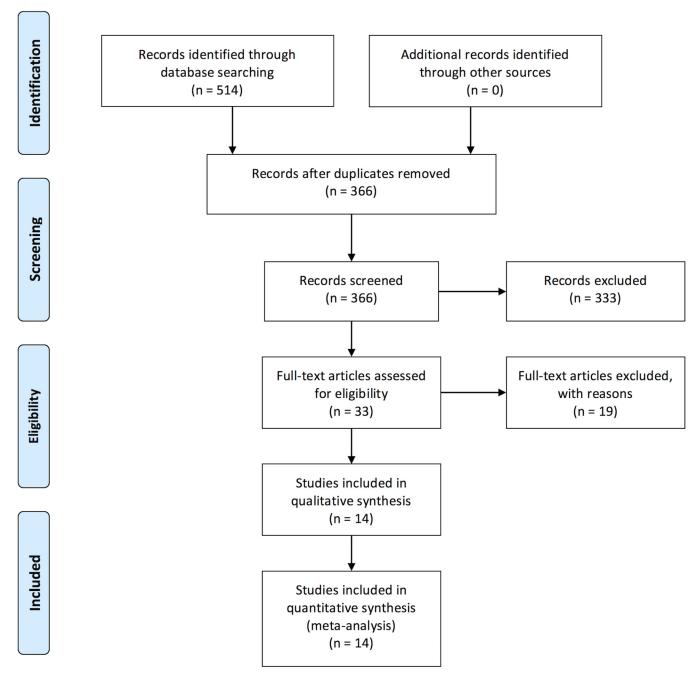


Fig 1. Flowchart depicting the selection process during systematic literature review.

Prevalence of malignancy and diagnostic performance of breast MRI

Analysis of prevalence showed substantial heterogeneity ranging from 1.8 to 56.7% (Q 162.2, I^2 92%; p<0.0001, <u>Table 1</u>). Using a random effects model, pooled prevalence was 14.3% (95% CI: 9.8–18.8%). The number of true positives, false positives, true negatives, and false negatives in each study are listed in <u>Table 2</u>. A bivariate analysis sROC curve revealed an area under the curve (AUC) of 96% (95%-CI 94–98%, <u>S2 Fig</u>). We then run bivariate analyses of breast MRI-



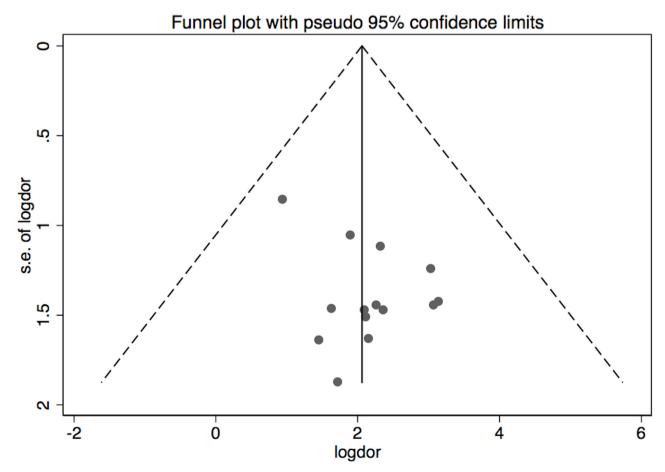


Fig 2. Begg's funnel scatterplot analysis of selected studies. The graph plots the logarithmic values of diagnostic odds ratios (logDOR) of considered studies in the abscissae axis against the standard error of *logDOR* in the ordinate axis. The two lines delimiting the inversed funnel denote pseudo 95% confidence intervals. Note the absence of any funnel plot asymmetry (confirmed by Egger's testing).

associated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The tested parameters also showed high heterogeneity as reflected by the corresponding I² and Q values. These were 82.2% and 73.1 (p < 0.001) for sensitivity, 83.2% and 73.5 (p < 0.001) for specificity, 84.7% and 84.9 (p < 0.001) for the PPV, and 92.5% and 174 (p = 0.001) for the NPV, respectively. Pooled sensitivity and specificity were high, reaching 99% (95% CI: 0.93-1) for sensitivity, and 89% (95% CI: 0.85-0.92) for specificity (Fig 3). PPV varied the most, ranging from 25% to 96.4% and yielding a pooled PPV of 56% (95% CI: 0.42-0.7; bivariate analysis; Fig 4). Based on pooled data, the likelihood ratio of a positive MRI scan was 9 (95% CI: 6.6–12.4, bivariate analysis; Fig 5), meaning that a positive MRI result increases the post-test odds of breast cancer by a factor of 9. Finally, with the exception of one outlier [18], the NPV ranged from 97.6 to 100%. Including this outlier, the pooled NPV was 99.9% (95% CI: 0.99-1.0, bivariate analysis; Fig 4). Altogether, seven false negative findings were described, four of which in the outlier study [18]. Three other studies reported a single false negative each [16,17,19]. There was only one false negative case, that of a Paget's disease of the breast, which showed no abnormality in mammography, ultrasound, and MRI [17]. Five of the remaining six false negatives were reader misinterpretations of abnormal MRI enhancements [16,18]. No specific MRI information was provided on the last false negative [19]. The pooled negative likelihood ratio was 0.01 (95% CI: 0.0-0.08; Fig 5), meaning that a negative MRI result



Table 1. Characteristics of patients and lesions considered in meta-analysis.

Author, year	QUADAS2 bias risk	Applicability concerns	MRI Indications**	Patients	Age	Ref. standard***	Lesions with FU only	FU months	Lesion size (mm)	Total cases	Malignancy prevalence
Lee, 1999	Р	No	3	86	53 (30– 80)	1	54	19 (5– 66)	12.4	98	0.10
Gökalp, 2006	Р	No	2	43	49.7 (37– 68)	1	46	>24	9.2 (3– 21)	56	0.02
Moy, 2009	Р	No	3	115	54 (32– 72)	1	100	34 (26– 72)	16 (10– 29)	115	0.05
Pediconi, 2009	Р	No	3	97	47.5 (16– 77)	1	NA	>18	NA	97	0.56
Benndorf, 2010	Р	No	1, 2	113*	54.4 ±11,8	1	NA	>12	NA	113	0.12
El- Barhoun, 2011	Р	No	3	41*	NA	1	8	>6	NA	41	0.12
Yau, 2011	P, R	No	3	204	NA	1	61°	12	NA	204	0.07
Dorrius, 2012	Р	No	3	25	48.7 (32– 69)	2	no FU	no FU	10-80 ^m 10-21 ^b	26	0.50
Lobrano, 2012	Р	No	3	66*	55 (20– 89)	1	36°°	NA	NA	126	0.08
Olsen, 2012	Р	No	3	51*	52 (31– 84)	1	29	12	NA	51	0.04
Bick, 2013	P, R	No	3	135	57.3 (50– 70)	1	52°°°	24	12 ≤10 16 >10	135	0.21
Oztekin, 2014	Р	No	3	868	47 (16– 82)	1	690	12	NA	868	0.07
Spick, 2015	Р	No	1	111	51 (20– 83)	1	NA	> 12	13.4 (6– 44) ^{ml} 23.8 (6– 75) ^{nml}	111	0.14
Strobel, 2015	Р	No	3	340*	53 (23– 81)	1	218	> 18	NA	275	0.15
Total										2,316	0.14

P: patient selection; R: reference standard; H: histology; b: benign; m: malignant; FU: follow-up; NA: not available;

doi:10.1371/journal.pone.0160346.t001

decreases the post-test odds for breast cancer 100-fold. Applying Bayes´ theorem, malignancy can thus be ruled out up to pretest probabilities of 60%. Under this premise, the resulting likelihood of malignancy of a negative MRI is less than 2% (Fig 5). The malignancy likelihood of a negative breast MRI falls to below 1% if pretest probabilities are set to 40% (Fig 5). In our study

^(*) patient subset fulfilling the selection criteria;

^{(**) 1 =} BI-RADS 0, 2 = BI-RADS 3, 3 = other;

^{(***) 1 =} histology and/or follow-up, 2 = histology only; m: malignant, b: benign, ml; mass lesions, nml: non-mass lesions



Table 2. Detailed indications for MRI, and extracted cross-tabulation data of MRI results against the reference standard.

Author, year	Indication for MRI	Total cases	True positives	False positives	False negatives	True negatives
Lee, 1999	Indeterminate significance of mammographic abnormality; ambiguous abnormality; lesion could not be located for biopsy; scar vs tumor at lumpectomy site or at benign biopsy site		10	28	0	60
Gökalp, 2006	BI-RADS 3		1	2	0	53
Moy, 2009	Further evaluation of inconclusive mammographic findings (asymmetry, architectural distortion, scar after benign biopsy)		6	9	0	100
Pediconi, 2009	Dense breast, problem solving, suspicious lesions		54	2	1	40
Benndorf, 2010	BI-RADS 0 and 3		14	6	1	92
El Barhoun, 2011	Suspicious imaging but no pre-MR diagnosis (30) or symptoms but mammogram/ultrasound unhelpful (11)		5	4	4	28
Yau, 2011	Problem solving (clinical, imaging, nipple discharge)		14	28	1	161
Dorrius, 2012	BI-RADS 3+4, lesion size ≥ 1cm		13	2	0	11
Lobrano, 2012	Problem solving (abnormal mammogram or ultrasound)		10	16	0	100
Olsen, 2012	Palpable masses with negative mammogram/ultrasound		2	6	0	43
Bick, 2013	Equivocal or minimal sign (single/multiple findings), lesion couldn't be sufficiently localized during biopsy attempt, benign biopsy with insufficient radiological-pathological concordance	135	28	25	0	82
Oztekin, 2014	Problem-solving: BI-RADS 0, BI-RADS 1&2 on MG/US but clinical findings, BI-RADS 3,4		63	66	0	739
Spick, 2015	Problem solving	111	15	11	0	85
Strobel, 2015	Problem solving		41	15	0	219
Total		2,316	276	220	7	1,813

population with a pooled prevalence of malignancy of 14.3%, the probability of malignancy in case of a negative MRI scan falls beneath 0.2%.

Meta-regression analysis

Using random-effects models, multivariate meta-regression analysis was performed. A minor but significant influence on sensitivity was identified for the covariates case number and cancer prevalence (P<0.05, respectively). This tendency towards higher sensitivity in larger studies and studies with a higher prevalence of malignancy was no longer statistically significant after the outlier study by El-Barhoun and Pitman [18] was removed from the meta-regression (P>0.05, see S1 File). Meta-regression on factors influencing specificity showed a minor but significantly higher specificity in studies investigating lesions either classified as BI-RADS 0 and/or 3 as opposed to other or less defined indications (P = 0.024, S1 File).

Discussion

According to our results, breast MRI demonstrates an excellent diagnostic performance in case of non-calcified equivocal breast findings detected at conventional imaging. In particular, and despite substantial heterogeneity, sensitivity and NPV were nearly 100%, and, consequently, the negative likelihood ratio was found to be very low. These findings indicate, that breast MRI can reliably exclude malignancy in most cases.



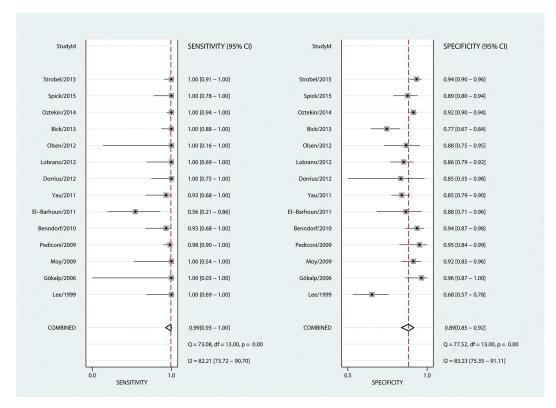


Fig 3. Forest plots of sensitivity and specificity. Sensitivity was defined as $\left(\frac{tp}{tp+fn}\right)$ and specificity as $\left(\frac{tn}{tn+fp}\right)$. tn: true negative; tp: true positive; fn: false negative; fp: false positive. All numbers have been rounded up or down to the closest second decimal.

A regularly brought up argument against breast MRI is that of a low specificity and, subsequently low PPV. In other words, it is assumed that breast MRI causes unnecessary biopsies and costly diagnostic follow-up procedures. This argument is not *per se* wrong: while pooled specificity was found to be as high as 88% in our study, we were not able to extract data on short-term follow-up examinations initiated by breast MRI in the investigated setting. Therefore, the number of short-term follow-up examinations necessary to achieve the high observed specificity remains elusive. On the other hand, as shown in this meta-analysis, PPV values were within the acceptable range of PPV for conventional imaging [46]. In addition, it has been shown that the majority of MRI-detected findings can be further identified by targeted ultrasound, providing a cost-effective and broadly available means of biopsy and follow-up in MRI-positive cases [47].

Our high diagnostic performance indices warrant some comparison with prior reports. The most recent systematic review that queried the effectiveness of breast MRI in resolving equivocal findings dates to 2010. The study comprised five studies and a total of 376 lesions including both calcified and non-calcified breast lesions [48]. The studies including calcified lesions showed inferior results as compared to those investigating non-calcified lesions. As stated above, workup of microcalcifications either by biopsy or follow-up does usually not require additional imaging modalities. The NPV for non-calcified lesions was reported as 100% in that systematic review which is in line with our findings. Among 283 malignant lesions in our study, only seven were missed by MRI. Five out of these seven false negative findings were actually visible as abnormal enhancements but were misinterpreted by the reporting radiologists



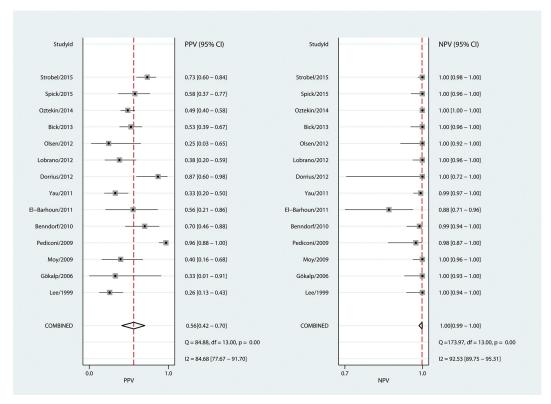


Fig 4. Forest plots of the positive and negative predictive values. The positive predictive value PPV was defined as $\left(\frac{tp}{tp+fp}\right)$ and the negative predictive value NPV as $\left(\frac{tn}{tn+fn}\right)$. tn: true negative; tp: true positive; fn: false negative; fp: false positive. All numbers have been rounded up or down to the closest second decimal.

[16,18]. One was a Morbus Paget of the mammilla and one was not visible by MRI [17,19]. A prior meta-analysis on breast MRI did report somewhat lower pooled diagnostic estimates: sensitivity of 90% and specificity of 72% [49]. This prior work is actually quoted as a reason by the EUSOMA recommendations why a negative MRI does not exclude breast cancer [50]. The lower diagnostic performance reported by the referenced work is likely due to data pooling of diagnostic accuracy studies on breast MRI without a focus on specific indications nor study designs. In particular, the prior meta-analysis included exploratory studies, such as those on artificial neural networks, and data on single diagnostic criteria such as the "peripheral washout" or "blooming sign" that refer to cancer specific diagnostic criteria [49]. We think that most would agree that such single diagnostic criteria (e.g. lesion margins) are not representative of the diagnostic potential of an imaging method as radiological diagnosis relies on the empirical or algorithmic combination of multiple imaging findings. On the other hand, the meta-analysis herein focuses on clinical reading results in a specific breast MRI indication subgroup and thus provides a more representative estimate of what is to be expected in clinical practice. This meta-analysis, therefore, provides a much need complement of the currently available meta-analytic literature on breast MRI focusing on other clinical indications such as preoperative staging [5], high-risk screening [6] and preoperative systemic therapy assessment [51,52].

As detailed in the introduction, there are various reasons to assigning conventional findings as equivocal. In its 5th and most recent edition of the BI-RADS atlas, the ACR lexicon states that breast MRI is not an appropriate follow-up measure for minimal or equivocal findings



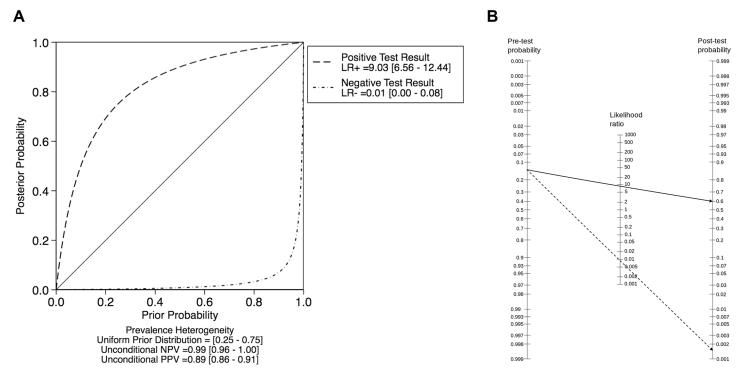


Fig 5. Pre- and post-test probabilities. A. Probability modifying plot. Note that post-test probabilities below 2% are achieved up to pre-test probabilities of 60%. **B.** Fagan's Nomogram applying pooled positive (plain line) and negative (dashed line) likelihood ratios to a pretest probability of 25% (the 95% CI upper bond of pooled prevalence of malignancy in all selected studies). Resulting posttest probabilities were 60% and 0.15% for a positive or a negative MRI respectively.

[53]. Current European EUSOMA recommendations stipulate that MRI should not be used as a problem solving tool if percutaneous biopsy can be performed while the less specific EUSOBI guidelines list problem-solving as an MRI indication [50,54]. In addition to the results of the already discussed general meta-analysis on breast MRI, the reticence to routinely use breast MRI as a problem-solving tool stems at least in part from the concern of whether conclusions derived from data obtained in one institution are pertinent to other institutions that may have applied somewhat different problem-solving indications, used different imaging equipment, or catered to different populations, to cite a few variables that can impact findings. The substantial variations in cancer prevalence identified by our analysis, ranging between 2-56%, confirms this concern and calls for a better definition of problem solving criteria. This finding is backed up by our QUADAS 2 assessment, identifying high or unclear risk of bias regarding patient selection in a number of studies. In particular, we found problem-solving indications rather generally defined in the inclusion criteria of the investigated studies while the majority did not stratify the reported results regarding the specific findings that led to breast MRI. Meta-regression unvealed a tendency towards higher sensitivity in larger studies and studies with a higher prevalence of malignancy. This could be due to higher reader experience in more experienced centers with a higher patient throughput and possibly also a better patient preselection were patients with a higher chance of malignancy were referred to MRI. On the other hand, this effect could be attributed to the clear outlier study by El-Barhoun and Pitman [18], as removal of their data eliminated statistical significance.

Regarding specificity, meta-regression found significantly higher values for studies focusing on BI-RADS 0 and 3 lesions as compared to all other indications. However, this is based only



on two out of 14 studies and although the effect is significant, the lack of data on lesion presentation in conventional imaging (e.g. architectural distortion vs solid lesion) clearly warrants further exploration. Therefore, due to the lack of indication-based subgroup data, this meta-analysis does not provide conclusions on adequate or inadequate indications for problem solving MRI. It appears reasonable to assume an only minor oncologic value of additional breast MRI in case of very low pre-test probabilities after negative conventional workup. In other words, our meta-analysis-analysis indicates that breast MRI ought not to be used to simply confirm negative conventional imaging.

This study has several limitations. First, it did not address the financial concerns that may be associated with the more frequent use of breast MRI, a subject that is best discussed within the framework of affordable healthcare for cancer patients [55]. It is however worth mentioning that recent indicators show that breast MRI costs are becoming reasonably low in some settings [56,57]. The recent increase in utilization of breast MRI has substantially expanded the data on breast MRI used as a problem-solving tool, which among other things enabled this study. Second, between-study heterogeneity was high, a finding in line with—if not in particular caused by—the highly variable prevalence of malignancy. As detailed elsewhere, PPV and NPV are affected by disease prevalence. While PPV positively correlates with prevalence, NPV negatively associates with it. Our results show a strong variation of PPV between studies (26%-96%), while NPV ranged only from 98% to 100% in thirteen out of fourteen studies. The remaining outlier had an NPV of 88%. Bivariate meta-analysis considers sensitivity and specificity as dependent variables. Consequently, a high heterogeneity on one of these parameters affects the other. This holds true for our moderately high PPV findings that do reflect the underlying heterogeneity in study populations as evidenced by their prevalence of malignancy. Conversely, the NPV of the studies included in our meta-analysis stayed robust over a wide range of prevalence values, a finding due to two reasons: high sensitivity and relatively low prevalence of malignancy. The high heterogeneity should thus not be seen as a limitation but rather a strength of our analysis: the robust and high NPV of breast MRI in the investigated setting underlines that MRI can exclude breast cancer with high accuracy despite the observed heterogeneity of problem-solving indications. Third, and as outlined above, due to a lack of indication-based subgroup data, this meta-analysis does not provide data on adequate or inadequate indications for problem solving MRI.

Conclusion

According to our results, breast MRI demonstrates an excellent diagnostic performance in case of non-calcified equivocal breast findings detected at conventional imaging. However, considering the substantial heterogeneity regarding prevalence of malignancy, problem-solving criteria need to be better defined.

Supporting Information

S1 Fig. QUADAS 2 graph illustrating the results of risk of bias and applicability assessment.

(PNG)

S2 Fig. sROC curve.

(TIFF)

S1 File. Meta-Regression analysis on factors influencing sensitivity and specificity. (DOCX)



S1 Table. PRISMA 2009 Checklist.

(DOC)

S2 Table. Technical details on MRI studies included in meta-analysis. (*) unless otherwise specified; NA: Not applicable.

(DOCX)

Acknowledgments

This article was supported by the Open Access Publishing Fund of the University of Vienna.

Author Contributions

Conceived and designed the experiments: BBB PAB.

Performed the experiments: BBB PAB.

Analyzed the data: BBB PAB.

Contributed reagents/materials/analysis tools: BBB PAB.

Wrote the paper: BBB NBB PAB.

References

- Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin. 2014; 64: 9–29.: doi: 10.3322/caac.21208 PMID: 24399786
- Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition—summary document. Ann Oncol. 2008; 19: 614–622.: doi: 10.1093/annonc/mdm481 PMID: 18024988
- Kreienberg R (coordinator). Interdisciplinary German S3 Guidelines for the Diagnosis, Treatment and Follow-up Care of Breast Cancer [Internet]. 2012. Available: www.awmf.org
- Early and locally advanced breast cancer: diagnosis and treatment | Guidance and guidelines | NICE [Internet]. [cited 21 Feb 2016]. Available: https://www.nice.org.uk/guidance/cg80
- Houssami N, Ciatto S, Macaskill P, Lord SJ, Warren RM, Dixon JM, et al. Accuracy and surgical impact of magnetic resonance imaging in breast cancer staging: systematic review and meta-analysis in detection of multifocal and multicentric cancer. J Clin Oncol Off J Am Soc Clin Oncol. 2008; 26: 3248–3258.: doi: 10.1200/JCO.2007.15.2108
- Warner E, Messersmith H, Causer P, Eisen A, Shumak R, Plewes D. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. Ann Intern Med. 2008; 148: 671–679. PMID: 18458280
- Hyun SJ, Kim E-K, Moon HJ, Yoon JH, Kim MJ. Preoperative axillary lymph node evaluation in breast cancer patients by breast magnetic resonance imaging (MRI): Can breast MRI exclude advanced nodal disease? Eur Radiol. 2016;: doi: 10.1007/s00330-016-4235-4
- Si L, Zhai R, Liu X, Yang K, Wang L, Jiang T. MRI in the differential diagnosis of primary architectural distortion detected by mammography. Diagn Interv Radiol Ank Turk. 2016; 22: 141–150.: doi: 10.5152/ dir.2016.15017
- Jalali S, Wohlin C. Systematic literature studies: database searches vs. backward snowballing. ACM Press; 2012. p. 29.: doi: 10.1145/2372251.2372257
- Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155: 529–536. PMID: 22007046
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997; 315: 629–634. PMID: 9310563
- Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003; 327: 557–560.: doi: 10.1136/bmj.327.7414.557 PMID: 12958120



- Lee CH, Smith RC, Levine JA, Troiano RN, Tocino I. Clinical usefulness of MR imaging of the breast in the evaluation of the problematic mammogram. AJR Am J Roentgenol. 1999; 173: 1323–1329.: doi: 10. 2214/ajr.173.5.10541112 PMID: 10541112
- Gökalp G, Topal U. MR imaging in probably benign lesions (BI-RADS category 3) of the breast. Eur J Radiol. 2006; 57: 436–444.: doi: 10.1016/j.ejrad.2005.10.004
- Moy L, Elias K, Patel V, Lee J, Babb JS, Toth HK, et al. Is breast MRI helpful in the evaluation of inconclusive mammographic findings? AJR Am J Roentgenol. 2009; 193: 986–993.: doi: 10.2214/AJR.08. 1229 PMID: 19770320
- 16. Pediconi F, Catalano C, Roselli A, Dominelli V, Cagioli S, Karatasiou A, et al. The challenge of imaging dense breast parenchyma: is magnetic resonance mammography the technique of choice? A comparative study with x-ray mammography and whole-breast ultrasound. Invest Radiol. 2009; 44: 412–421.: doi: 10.1097/RLI.0b013e3181a53654 PMID: 19448554
- 17. Benndorf M, Baltzer PAT, Vag T, Gajda M, Runnebaum IB, Kaiser WA. Breast MRI as an adjunct to mammography: Does it really suffer from low specificity? A retrospective analysis stratified by mammographic BI-RADS classes. Acta Radiol Stockh Swed 1987. 2010; 51: 715–721.: doi: 10.3109/02841851.2010.497164
- El-Barhoun EN, Pitman AG. Impact of breast MR in non-screening Australian clinical practice: audit data from a single-reader single-centre site. J Med Imaging Radiat Oncol. 2011; 55: 461–473.: doi: 10. 1111/j.1754-9485.2011.02302.x PMID: 22008164
- Yau EJ, Gutierrez RL, DeMartini WB, Eby PR, Peacock S, Lehman CD. The utility of breast MRI as a problem-solving tool. Breast J. 2011; 17: 273–280.: doi: 10.1111/j.1524-4741.2011.01075.x PMID: 21477168
- Dorrius MD, Pijnappel RM, van der Weide Jansen MC, Jansen L, Kappert P, Oudkerk M, et al. The added value of quantitative multi-voxel MR spectroscopy in breast magnetic resonance imaging. Eur Radiol. 2012; 22: 915–922.: doi: 10.1007/s00330-011-2322-0 PMID: 22076317
- Lobrano MB, Stolier A, L'Hoste R, Luttrell CA. Breast MRI: patterns of utilization and impact on patient management in the community hospital setting. J La State Med Soc Off Organ La State Med Soc. 2012; 164: 38–42.
- Olsen ML, Morton MJ, Stan DL, Pruthi S. Is there a role for magnetic resonance imaging in diagnosing palpable breast masses when mammogram and ultrasound are negative? J Womens Health 2002. 2012; 21: 1149–1154.: doi: 10.1089/jwh.2012.3735
- 23. Bick U, Engelken F, Diederichs G, Dzyuballa R, Ortmann M, Fallenberg EM. MRI of the breast as part of the assessment in population-based mammography screening. RöFo Fortschritte Auf Dem Geb Röntgenstrahlen Nukl. 2013; 185: 849–856.: doi: 10.1055/s-0033-1335518
- **24.** Oztekin PS, Kosar PN. Magnetic resonance imaging of the breast as a problem-solving method: to be or not to be? Breast J. 2014; 20: 622–631.: doi: 10.1111/tbj.12334 PMID: 25200378
- Spick C, Szolar DHM, Preidler KW, Tillich M, Reittner P, Baltzer PA. Breast MRI used as a problem-solving tool reliably excludes malignancy. Eur J Radiol. 2015; 84: 61–64.: doi: 10.1016/j.ejrad.2014.10.
 O05 PMID: 25454098
- 26. 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: 343–351.: doi: 10.1148/radiol.14140645 PMID: 25271857
- Nakahara H, Namba K, Fukami A, Watanabe R, Maeda Y, Furusawa H, et al. Three-dimensional MR imaging of mammographically detected suspicious microcalcifications. Breast Cancer Tokyo Jpn. 2001; 8: 116–124.
- 28. Bazzocchi M, Zuiani C, Panizza P, Del Frate C, Soldano F, Isola M, et al. Contrast-enhanced breast MRI in patients with suspicious microcalcifications on mammography: results of a multicenter trial. AJR Am J Roentgenol. 2006; 186: 1723–1732.: doi: 10.2214/AJR.04.1898 PMID: 16714666
- 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: 29–38.: doi: 10.1016/j. breast.2005.05.002
- Cilotti A, Iacconi C, Marini C, Moretti M, Mazzotta D, Traino C, et al. Contrast-enhanced MR imaging in patients with BI-RADS 3–5 microcalcifications. Radiol Med (Torino). 2007; 112: 272–286.: doi: 1007/s11547-007-0141-9
- Uematsu T, Yuen S, Kasami M, Uchida Y. Dynamic contrast-enhanced MR imaging in screening detected microcalcification lesions of the breast: is there any value? Breast Cancer Res Treat. 2007; 103: 269–281.: doi: 10.1007/s10549-006-9373-y PMID: 17063274



- **32.** Ballesio L, Maggi C, Savelli S, Angeletti M, De Felice C, Meggiorini ML, et al. Role of breast magnetic resonance imaging (MRI) in patients with unilateral nipple discharge: preliminary study. Radiol Med (Torino). 2008; 113: 249–264.: doi: 10.1007/s11547-008-0245-x
- Dorrius MD, Pijnappel RM, Sijens PE, van der Weide MCJ, Oudkerk M. The negative predictive value of breast Magnetic Resonance Imaging in noncalcified BIRADS 3 lesions. Eur J Radiol. 2012; 81: 209–213.: doi: 10.1016/j.ejrad.2010.12.046 PMID: 21251784
- Fiaschetti V, Pistolese CA, Perretta T, Cossu E, Arganini C, Salimbeni C, et al. 3–5 Bl-RADs Microcalcifications: Correlation between MRI and Histological Findings. ISRN Oncol. 2011; 2011: 643890.: doi: 10.5402/2011/643890 PMID: 22084735
- Friedlander LC, Roth SO, Gavenonis SC. Results of MR imaging screening for breast cancer in highrisk patients with lobular carcinoma in situ. Radiology. 2011; 261: 421–427.: doi: 10.1148/radiol. 11103516 PMID: 21900618
- 36. Linda A, Zuiani C, Furlan A, Lorenzon M, Londero V, Girometti R, et al. Nonsurgical management of high-risk lesions diagnosed at core needle biopsy: can malignancy be ruled out safely with breast MRI? AJR Am J Roentgenol. 2012; 198: 272–280.: doi: 10.2214/AJR.11.7040 PMID: 22268168
- Linda A, Zuiani C, Londero V, Cedolini C, Girometti R, Bazzocchi M. Magnetic resonance imaging of radial sclerosing lesions (radial scars) of the breast. Eur J Radiol. 2012; 81: 3201–3207.: doi: 10.1016/j. ejrad.2012.01.038 PMID: 22386132
- Ozcan UA, Ocak F, Altun E. Can breast MRI facilitate decision-making in equivocal mammography and ultrasonography findings? Eur J Radiol. 2012; 81 Suppl 1: S112–113.: doi: 10.1016/S0720-048X(12) 70046-3 PMID: 23083552
- Jiang Y, Lou J, Wang S, Zhao Y, Wang C, Wang D. Evaluation of the role of dynamic contrastenhanced MR imaging for patients with BI-RADS 3–4 microcalcifications. PloS One. 2014; 9: e99669.: doi: 10.1371/journal.pone.0099669 PMID: 24927476
- 40. Kikuchi M, Tanino H, Kosaka Y, Sengoku N, Yamashita K, Minatani N, et al. Usefulness of MRI of microcalcification lesions to determine the indication for stereotactic mammotome biopsy. Anticancer Res. 2014; 34: 6749–6753. PMID: <u>25368286</u>
- Lee CI, Ichikawa L, Rochelle MC, Kerlikowske K, Miglioretti DL, Sprague BL, et al. Breast MRI BI-RADS assessments and abnormal interpretation rates by clinical indication in US community practices. Acad Radiol. 2014; 21: 1370–1376.: doi: 10.1016/j.acra.2014.06.003 PMID: 25126973
- 42. Merckel LG, Verkooijen HM, Peters NHGM, Mann RM, Veldhuis WB, Storm RK, et al. The added diagnostic value of dynamic contrast-enhanced MRI at 3.0 T in nonpalpable breast lesions. PloS One. 2014; 9: e94233.: doi: 10.1371/journal.pone.0094233 PMID: 24713637
- 43. Manganaro L, D'Ambrosio I, Gigli S, Di Pastena F, Giraldi G, Tardioli S, et al. Breast MRI in patients with unilateral bloody and serous-bloody nipple discharge: a comparison with galactography. BioMed Res Int. 2015; 2015: 806368.: doi: 10.1155/2015/806368 PMID: 25685810
- Price J, Chen SW. Mammographically screen-detected asymmetric densities with architectural distortion and normal ultrasound at assessment: Value of MRI as a problem-solving tool. J Med Imaging Radiat Oncol. 2015; 59: 312–319.: doi: 10.1111/1754-9485.12281 PMID: 25707846
- 45. 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: e111217.: doi: 10.1371/journal.pone.0111217 PMID: 25365327
- 46. Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. Radiology. 2006; 239: 385–391.: doi: 10.1148/radiol.2392042127 PMID: 16569780
- Spick C, Baltzer PAT. Diagnostic utility of second-look US for breast lesions identified at MR imaging: systematic review and meta-analysis. Radiology. 2014; 273: 401–409.: doi: 10.1148/radiol.14140474
 PMID: 25119022
- Dorrius MD, Pijnappel RM, Jansen-van der Weide MC, Oudkerk M. Breast magnetic resonance imaging as a problem-solving modality in mammographic BI-RADS 3 lesions. Cancer Imaging Off Publ Int Cancer Imaging Soc. 2010; 10 Spec no A: S54–58.: doi: 10.1102/1470-7330.2010.9020
- 49. Peters NHGM, Borel Rinkes IHM, Zuithoff NPA, Mali WPTM, Moons KGM, Peeters PHM. Meta-analysis of MR imaging in the diagnosis of breast lesions. Radiology. 2008; 246: 116–124.: doi: 10.1148/radiol.2461061298 PMID: 18024435
- 50. Sardanelli F, Boetes C, Borisch B, Decker T, Federico M, Gilbert FJ, et al. Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. Eur J Cancer Oxf Engl 1990. 2010; 46: 1296–1316.: doi: 10.1016/j.ejca.2010.02.015



- Marinovich ML, Houssami N, Macaskill P, Sardanelli F, Irwig L, Mamounas EP, et al. Meta-analysis of magnetic resonance imaging in detecting residual breast cancer after neoadjuvant therapy. J Natl Cancer Inst. 2013; 105: 321–333.: doi: 10.1093/jnci/djs528 PMID: 23297042
- Marinovich ML, Macaskill P, Irwig L, Sardanelli F, Mamounas E, von Minckwitz G, et al. Agreement between MRI and pathologic breast tumor size after neoadjuvant chemotherapy, and comparison with alternative tests: individual patient data meta-analysis. BMC Cancer. 2015; 15: 662.: doi: 10.1186/ s12885-015-1664-4 PMID: 26449630
- 53. D'Orsi C, Sickles E, Mendelson E, Morris E. ACR BI-RADS[®] Atlas, Breast Imaging Reporting and Data System. American College of Radiology; 2013.
- Mann RM, Balleyguier C, Baltzer PA, Bick U, Colin C, Cornford E, et al. Breast MRI: EUSOBI recommendations for women's information. Eur Radiol. 2015;: doi: 10.1007/s00330-015-3807-z
- 55. Chalkidou K, Marquez P, Dhillon PK, Teerawattananon Y, Anothaisintawee T, Gadelha CAG, et al. Evidence-informed frameworks for cost-effective cancer care and prevention in low, middle, and high-income countries. Lancet Oncol. 2014; 15: e119–131.: doi: 10.1016/S1470-2045(13)70547-3 PMID: 24534293
- 56. Keller DL. In defense of screening for breast cancer with magnetic resonance imaging. JAMA Intern Med. 2014; 174: 1417.: doi: 10.1001/jamainternmed.2014.810
- Kuhl CK, Schrading S, Strobel K, Schild HH, Hilgers R-D, Bieling HB. 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 Off J Am Soc Clin Oncol. 2014; 32: 2304–2310.: doi: 10.1200/JCO.2013.52.5386