REVIEW

Accuracy of magnetic resonance in suspicious breast lesions: a systematic quantitative review and meta-analysis

**Lidia Rosi Medeiros · Célia Scapin Duarte ·
Daniela Dornelles Rosa · Maria Isabel Edelweiss ·
Marcia Edelweiss · Fábio Rosa Silva · Erik Paul Winnnikow ·
Patrícia Duarte Simões Pires · Maria Inês Rosa**

Abstract Dynamic contrast-enhanced breast magnetic resonance (MR) is a promising emerging technique for evaluating breast lesions. A quantitative systematic review was performed to estimate the accuracy of breast MR in the diagnosis of high-risk breast lesions and breast cancer. A comprehensive search of the Cochrane Library, MEDLINE, CANCERLIT, LILACS, and EMBASE databases was performed from January 1985 to August 2010. The medical subjects heading (MeSH) and text words for the terms "*breast neoplasm*", "*breast lesions*", "*breast cancer*" and "*magnetic resonance*" were combined with the MeSH term *diagnosis* ("sensitivity and specificity"). Studies that compared breast MR with paraffin-embedded sections parameters for the diagnosis of breast lesions (benign, high-risk borderline, and breast cancer) were included. Sixty-nine studies were analyzed, which included 9,298 women with 9,884 breast lesions. Interrater overall agreement between breast MR and paraffin section diagnosis was 79% ($\kappa = 0.55$), indicating moderate agreement. Pooled sensitivity and specificity were 90% [95% CI 88–92%] and 75% [95% CI 70–79%], respectively. The pooled likelihood positive ratio was 3.64 (95% CI 3.0–4.2) and the negative ratio was 0.12 (95% CI 0.09–0.15). For breast cancer or high-risk lesions versus benign lesions, the AUC was 0.91 for breast MR and the point Q^* was 0.84. In summary, breast MR is a useful pre-operative test for predicting the diagnosis of breast lesions.

Keywords Magnetic resonance mammography · Breast magnetic resonance · MRI · Systematic review · Meta-analysis · Breast cancer · Diagnosis · Sensitivity · Specificity · ROC

L. R. Medeiros · D. D. Rosa · M. I. Edelweiss
Postgraduate Program in Medicine, Medical Sciences at Federal University of Rio Grande do Sul, Porto Alegre, Brazil
e-mail: lidia.rosi@terra.com.br

C. S. Duarte · F. R. Silva · E. P. Winnnikow ·
P. D. Simões Pires · M. I. Rosa (✉)
Laboratory of Epidemiology and National Institute for Translational Medicine Health Sciences Unit, University of Extremo Sul Catarinense (UNESC), Rua Cruz e Souza, 510.
Bairro Pio Correa-CEP, 88811-550 Criciúma, SC, Brazil
e-mail: mir@unesc.net

C. S. Duarte · E. P. Winnnikow · P. D. Simões Pires ·
M. I. Rosa
Postgraduate Program in Health Sciences, UNESC, Criciúma, Brazil

D. D. Rosa
Oncology Unit, Hospital Fêmina, Grupo Hospitalar Conceição, Porto Alegre, Brazil

M. I. Edelweiss
Pathology Unit of Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil

M. Edelweiss
Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Introduction

Dynamic contrast-enhanced breast magnetic resonance (MR) is a promising emerging technique for evaluating breast lesions [1]. Several investigations have demonstrated the potential higher sensitivity of breast MR imaging relative to mammography or other conventional imaging techniques [1–3]. In breast MR, diagnosis is based on the dynamic contrast-enhancement, e.g. the increase in signal intensity, the speed and pattern of enhancement, and the degree of regularity of the border of the lesion [1–3].

Levels of suspicion are reported according to the Breast Imaging Reporting and Data System (BI-RADS) lexicon, with morphology rated on a scale of 0–5 (0, needs additional imaging evaluation; 1, normal enhancement; 2, benign enhancement; 3, probably benign, recommend short-term follow up; 4, suspicious; 5, highly suggestive of malignancy) [4].

Lesions of uncertain malignant potential include atypical ductal hyperplasia (ADH), atypical lobular hyperplasia (ALH), atypical columnar cell hyperplasia (ACH), lobular carcinoma in situ (LCIS), lobular neoplasia (LN), papillary lesions (PL), radial sclerosing lesions (RSL), fibroepithelial lesions, mucocele-like lesions, and columnar cell lesions [5]. According to some studies, atypical epithelial hyperplasia (ADH, ALH), LCIS, PL, multiple micropapillomas (MMP), and radial scars are considered high-risk breast lesions that predispose toward the future development of non-invasive or invasive breast cancer [6–9]. Dynamic contrast-enhanced breast MR is emerging as a powerful tool for detecting high-grade ductal carcinoma in situ. It has major capabilities for the diagnosis, detection and monitoring of malignancy. Its benefits include being non-invasive and three-dimensional, allowing visualization of the extent of disease and its angiogenic properties, visualization of lesion heterogeneity, detection of changes in angiogenic properties before morphological alterations, and the potential to predict the overall response either before the start of therapy or early during treatment.

The purpose of this systematic review and meta-analysis was to assess the accuracy of breast MR in patients with suspicious breast lesions.

Materials and methods

Search strategy

A comprehensive search of the PubMed, Cancerlit, Lilacs, Scopus, and Embase databases for relevant publications from January 1985 to August 2010 was conducted. Databases were searched using the following terms, both as text words and, as appropriate, Medical Subjects Heading (MeSH) or equivalent subject heading/thesaurus terms: *"breast neoplasms"*, *"breast lesions"*, *"breast cancer"*, *"breast tumor"* *"mammary neoplasms"*, *"mammary cancer"*, *"mammary tumor"*, and *"magnetic resonance"* were combined with the MeSH term *diagnosis* (sensitivity, specificity, false positive, false negative, predictive value, reference value, ROC, likelihood ratio, accuracy, diagnosis, measurement and analysis). This sensitive filter was created by combining three filters for the identification of diagnostic studies via the Boolean operator "OR" and "AND". The search was limited to human studies but had no language restrictions. In addition, the Cochrane Central Register of Controlled Trials was searched up to July 2010. Reference lists of all retrieved primary diagnostic studies were checked for additional relevant diagnostic studies. Additionally, we checked references of relevant reviews, meta-analyses, guidelines, and commentaries identified in PubMed and Embase. The authors who published the studies were not contacted. The complete search strategy is available on request.

Screening of abstract for eligibility

Abstracts/titles identified from the search were screened by six reviewers (M.I.R, M.I.E., F.R.S., E.P.W., M.E., and D.D.R.). Disagreements about study inclusion or exclusion were initially solved by consensus, and when this was not possible, they were arbitrarily resolved by one reviewer (L.R.M.).

Study selection

We included primary diagnostic studies with prospective or retrospective cohort and cross-sectional designs that evaluated clinically suspected breast lesions (target conditions) by breast MR (index test) according to the American College of Radiology practice guidelines [3].

Patients

Studies that included women with mammographic and sonographic abnormalities (BI-RADS 3–5) and physical examination findings were inconclusive [3]. We excluded studies with patients who (a) were known carriers of a BRCA 1 or BRCA 2 mutation, (b) had a first-degree family member known to carry a BRCA 1 or BRCA 2 mutation but did not know their own status, (c) had a personal or family history highly suggestive of BRCA 1 or BRCA 2 mutation, (d) had a 15–20% lifetime risk for breast cancer on the basis of a personal history of breast or ovarian cancer, (e) biopsy proven lobular neoplasia or ADH, and/or had (f) radiation therapy to the chest before the age of 30 years and at least 8 years previously.

Index test

The diagnostic test was performed only with high-field strength using a commercially available 0.5 or 3 T system with a dedicated surface breast coil. Breast MR was performed using a phase array coil and T1-weighting for the transverse plane and T2-weighting for the transverse and sagittal planes. The increase of signal intensity of the solid components was evaluated on dynamic contrast-enhancement of breast MR images when available. In the absence

of dynamic contrast-enhancement, the uptake of the contrast medium from solid tissues was evaluated and compared to pre- and post-contrast breast MR images. The following signs were judged as indicators of malignancy in the dynamic criteria: (a) fast wash-in, (b) plateau phenomenon, (c) wash-out phenomenon, (d) blooming sign, (e) inhomogeneous enhancement, and (f) centripetal enhancement. In the morphological criteria, the indicators of malignancy were as follows: (a) unsharp borders of lesion on T1-weighted images, (b) iso-or hypointensity on T1-weighted images, (c) adjacent vessel sign (AVS), (d) cutaneous thickening, (e) iso- or hypointensity on T2-weighted images, (f) perifocal, diffuse, unilateral, or bilateral edema, (g) connection of the lesion to the *M. pectoralis*, and (h) lymph nodes larger than 1 cm. The results of the index test were compared to the results of a reference standard.

Reference standard

The diagnostic reference test was the result of the histological analysis of standard paraffin-embedded sections after surgery. The breast MR diagnosis was considered correct if it did not differ from that derived from the paraffin section. For inclusion in this systematic review, the final histological diagnoses of the breast lesions had to include at least two of the following diagnosis as reference standards: benign, high-risk, or malignant breast lesion. We excluded studies presenting exclusively benign, high-risk lesions, or breast cancer as the reference standard. Breast lesions were grouped into high-risk lesions and breast cancer and compared with benign breast lesions. We grouped high-risk breast lesions with malignant lesions because they are considered to predispose the individual to the future development of non-invasive or invasive breast cancer [6–8]. Thus, the primary outcome analyzed was the accuracy of breast lesion diagnoses (high-risk lesions or breast cancer vs. benign lesions) by breast MR. A secondary outcome was the distribution of histological types of breast lesions according to paraffin section diagnosis (benign, high-risk, and breast cancer).

Data collection and quality assessment

We extracted data on studies, patients, and test characteristics by using a standardized form. Two reviewers (F.R.S., C.S.D.) independently abstracted data regarding the prevalence of benign, high-risk, and malignant breast lesions. They also calculated the sensitivities, specificities, positive and negative likelihood ratios, and positive and negative post-test probabilities from the primary studies of breast MR diagnoses. Studies that lacked the data needed to construct 2×2 contingency tables were excluded. The

assessment of non-English articles was performed independently (D.D.R.) following translation (when necessary). Any disagreement was resolved by consensus for studies published in all languages. Final inclusion or exclusion was made with reference to a selection criteria checklist. Disagreements about study inclusion or exclusion were initially solved by consensus, and when this was not possible, they were arbitrarily resolved by one reviewer (L.R.M.). The agreement statistics among reviewers were computed.

Methodological quality assessment of diagnostic accuracy studies was performed according to QUADAS criteria, as modified for use by the Cochrane Collaboration (items that related to reporting quality were removed) [10, 11]. The modification version consists of 11 items on study characteristics with the potential to introduce bias. Items were scored as positive (no bias), negative (potential bias), or insufficient information [10, 11]. This involved scrutinizing the studies' designs, such as the methods of data collection, the relevant features of the patient population/selection, a description of the breast MR test and the histological reference standard, and the presence of verification biases [10–14].

Data synthesis and statistical analysis

To evaluate the agreement between assessments of methodological quality, as well as between the results of breast MR and paraffin-embedded section analyses, the observed percentage of agreement and the κ coefficient for interrater reliability were calculated [13]. For each study, 2×2 contingency tables were constructed in which all biopsies were classified as benign lesions, high-risk breast lesions, or breast cancer. We calculated the true-positive rate (TPR; sensitivity), specificity, false-positive rate (FPR; 1—specificity) and likelihood ratios (LH). When 2×2 tables had 0 cells, correction of calculations were made, and when a study contained two cells with the value 0, it was excluded from the analysis [14]. Bivariate analysis was used to calculate pooled estimates of sensitivity, specificity, and likelihood ratios (LRs) along with 95% confidence intervals (CIs) for the summary estimates [15, 16]. The bivariate model preserves the 2-dimensional nature of diagnostic data by analyzing the logit transformed sensitivity and specificity of each study in a single model, and takes into account both within-study and between-study variability [15–20]. The model produces the following results: a random-effects estimate of the mean sensitivity and specificity with corresponding 95% CIs, the amount of between-study variation for sensitivity and specificity separately, and the strength and shape of the correlation between sensitivity and specificity. Pooled estimates were calculated only for studies showing sufficient clinical and statistical homogeneity (I^2 or Q test commonly used in meta-analysis). I^2 or Q test is not recommended for

assessing statistical homogeneity in diagnostic reviews because it does not take into account the association between sensitivity and specificity [21]. Statistical homogeneity was defined as overlapping 95% CIs of both sensitivity and specificity and differences in point estimates among the studies of less than 20%. When assessing heterogeneity, we always simultaneously considered sensitivity and specificity [22, 23]. A summary receiver operating characteristic curve (SROC) was generated using data from all thresholds [16, 18, 19, 24]. It can change according to the threshold and to the ROC curve used to define an abnormal examination, resulting in the expected trade-off between sensitivity and specificity. The SROC is an excellent graphical summary, but for comparison purposes, we calculated a further statistic, Q^* [16]. Q^* is the point on the SROC where sensitivity and specificity are equal. Like area under the curve, the ROC curve and the Q^* point indicate how closely a test approaches the desirable performance of 100% sensitivity and specificity [16]. The higher the Q^* value, the better the diagnostic test performance [16]. Sensitivity analyses were performed to assess the exclusion of the effect of verification bias by calculating pooled estimates of diagnostic performance, excluding studies before the year of 2000. The statistical analysis was performed with the softwares Meta-DiSc® (Clinical Biostatistics Unit, Ramón y Cajal Hospital, Madrid, Spain) (version 1.4) [25], RevMan 5.0.21 (The Nordic Cochrane Centre, Copenhagen, Denmark) [26], and STATA version 11 [27].

Results

Study identification and eligibility

The process of study selection is summarized in Fig. 1. We identified 5,054 citations from electronic searches in Pub-Med, and 2,017 additional citations from EMBASE. After initial evaluation, 355 full articles were retrieved, 69 of which were finally considered eligible for the review [8, 28–95]. A complete list of the excluded studies is available from the authors. Sixty-nine primary studies, including 9,298 women with 9,884 breast lesions, met the criteria for inclusion and were analyzed (Table 1) [8, 28–95]. Interrater overall agreement for study eligibility and methodological quality was 88% ($\kappa = 0.67$), indicating good agreement. Disagreement between reviewers related to inclusion or exclusion criteria occurred during analysis of the 69 studies. This disagreement was resolved by consensus.

Descriptions of studies

Details of the participants, text standard, and index test are summarized in Table 1. The mean age of participants was
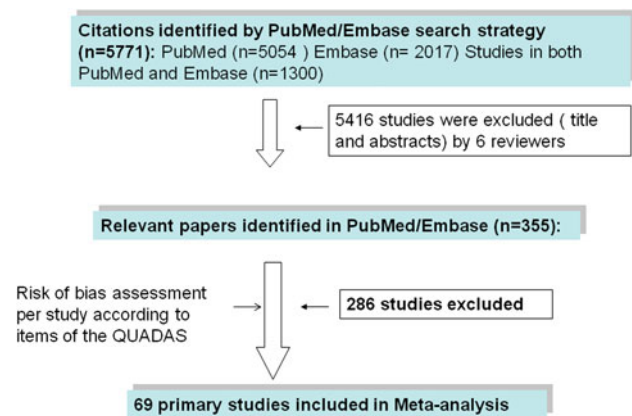

Fig. 1 Study selection process

reported in 59 studies [8, 28, 30, 32, 34–36, 38, 39, 41, 43, 45, 46, 49–72, 75–89, 91–94]. Forty-three studies were prospective and consecutive with a small population, but included sufficient experimental details, proper diagnostic tests and diagnostic reference standards [8, 28–34, 37–45, 51, 52, 55–59, 62, 63, 65, 70, 73, 75–77, 79, 82–84, 86–95]. The results of quality assessment are presented in Fig. 2 according to QUADAS [10]. On average, the reviewers disagreed on 3 of 11 items (range, 0–6). All disagreements were resolved by consensus. The most common source of bias was results that were difficult to interpret. Patient withdrawals were explained in only 10% of studies, and we were unable to determine if interpretation was done in a blinded manner in 90% of the studies included.

Breast cancer was found in 5,751 cases (58%) and 4,133 (42%) cases were benign lesions. Table 2 shows the results of contingency tables (TP, FP, FN, and TN) from each study considered in the systematic review.

Diagnostic performance and summary of results

Interrater overall agreement between breast MR and paraffin sections was 79% ($\kappa = 0.55$), indicating moderate agreement (Table 2). Pooled sensitivity and specificity was calculated according to bivariate analysis because of homogeneity (patients, breast lesions), and the results had values less than 20% (Table 3). Pooled sensitivity was 90% (95% CI 88–92%) and specificity was 75% (95% CI 70–79%). The pooled positive likelihood ratio was 3.64 (95% CI 3.0–4.2) and the negative likelihood ratio was 0.12 (95% CI 0.09–0.15). For breast cancer or high-risk lesions versus benign lesions, the area under the curve (AUC) was 0.91 for breast MR, and the point Q^* (summary point) was 0.84 (Fig. 3) (Q^* is the point on the SROC where sensitivity and specificity are equal). The AUC and point Q^* for the ROC curve were estimated by trapezoidal rule (STATA®; version 11) [27].

Table 1 Characteristics of primary diagnostic studies on diagnosing breast lesions

First author	Design and settings	N	Mean age (range)	Number total of lesions	Breast cancer prevalence (%)	MRI technique (T)
Abdolmaleki et al. [28]	Cross-sectional prospective, USA	120	49 (15–79)	120	94 (78.3)	1.5
Alamo and Fisher [29]	Cross-sectional prospective, Germany	40	NA*	40	23 (57.5)	1.5
Bagni et al. [30]	Cross-sectional prospective, Italy	45	51 (32–84)	45	38 (84.4)	1.5
Bazzocchi et al. [31]	Cross-sectional, prospective, multicenter, Italy	112	NA	112	80 (71.4)	1.0–1.5
Berge et al. [32]	Cross-sectional prospective, USA	120	48.7 (26–81)	258	177 (68.6)	1.5
Blohmer et al. [33]	Cross-sectional prospective, Germany	68	NA	68	41 (60.3)	1.5
Blumke et al. [34]	Cross-sectional prospective, Multicenter, USA	821	53 (±11.6)	821	404 (49.2)	1.5
Boetes [35]	Cross-sectional, Netherland	83	51 (24–79)	87	65 (74.7)	1.5
Boetes et al. [36]	Cross-sectional, Netherland	204	52 (23–83)	208	145 (69.7)	1.5
Boné et al. [37]	Cross-sectional prospective, Sweden	231	NA	250	144 (57.6)	1.5
Boné et al. [38]	Cross-sectional prospective Sweden	220	52 (23–83)	238	155 (65.1)	1.5
Buadu et al. [39]	Cross-sectional prospective, Japan	71	51 (30–78)	71	59 (83.1)	1.5
Cecil et al. [40]	Cross-sectional prospective, USA	38	NA	38	23 (60.5)	1.5
Choi et al. [41]	Cross-sectional prospective, Korea	89	48 (32–72)	99	65 (65.7)	1.5
Cilotti [42]	Cross-sectional prospective, Italy	55	NA	55	30 (54.5)	1.5
Daldrup et al. [43]	Cross-sectional prospective, multicenter, Europe	63	52 (25–83)	63	37 (58.7)	1.5
Daniel et al. [44]	Cross-sectional prospective, USA	52	NA	57	23 (40.3)	1.5
Dietzel et al. [45]	Cross-sectional prospective, Korea	1012	55 (16–91)	1084	648 (59.8)	1.5
Fischer et al. [46]	Cross-sectional, Germany	463	54.3 (21–89)	548	405 (73.9)	1.5
Fischer et al. [47]	Cross-sectional, Germany	793	NA	793	514 (64.8)	1.5
Fobben et al. [48]	Cross-sectional, USA	89	NA	91	21 (23.0)	1.5
Gatzemeier et al. [49]	Cross-sectional, prospective, Germany	112	58.5 (33–91)	112	91 (81.2)	1.5
Gibss et al. [50]	Cross-sectional retrospective, England	43	49 (23–74)	49	32 (65.3)	1.5
Gilles et al. [51]	Cross-sectional, prospective, France	172	53 (30–81)	172	86 (50.0)	1.5
Goto et al. [52]	Cross-sectional, Prospective, Japan	190	50 (25–80)	204	144 (70.6)	1.5
Harms et al. [53]	Cross-sectional, USA	74	56 (32–87)	74	47 (63.5)	1.5
Heiberg [54]	Cross-sectional, USA	56	48 (25–83)	83	23 (27.7)	1.5
Heinisch et al. [55]	Cross-sectional, prospective, Austria	36	48 (25–77)	40	25 (62.5)	1.0
Helbich et al. [56]	Cross-sectional, prospective, USA	66	47 (19–85)	75	26 (34.6)	1.5
Hickman et al. [57]	Cross-sectional, England	24	57 (36–79)	25	4 (16.0)	1.5
Huang et al. [58]	Cross-sectional, prospective, USA	50	50 (34–71)	50	18 (36.0)	1.5
Hulka et al. [59]	Cross-sectional, Prospective, USA	63	47 (13–70)	71	24 (33.8)	1.5
Ikeda et al. [60]	Cross-sectional, retrospective, Japan	61	48 (22–77)	64	41 (64.0)	1.5
Jacobs et al. [61]	Cross-sectional, retrospective, USA	36	45 (18–80)	36	18 (50.0)	1.5
Kelcz et al. [62]	Cross-sectional, prospective, Israel	57	52 (31–80)	68	31 (45.6)	1.5
Khouli et al. [63]	Cross-sectional, prospective, Egypt	91	50 (24–73)	101	78 (77.2)	1.5–3
Kinkel et al. [64]	Cross-sectional, retrospective, Switzerland	57	52 (29–74)	57	34 (59.6)	1.5
Kuhl et al. [65]	Cross-sectional, prospective, Germany	230	45 (21–75)	204	101 (49.5)	1.5
Liu et al. [66]	Cross-sectional, Switzerland	90	54 (19–87)	119	69 (57.9)	1.5
Lucht et al. [67]	Cross-sectional, Germany	264	48 (27–79)	264	162 (61.4)	1.5
Morris et al. [68]	Cross-sectional, retrospective, USA	101	51 (28–79)	101	40 (39.6)	1.5
Moy et al. [69]	Cross-sectional, retrospective, USA	115	54 (30–72)	115	15 (13.0)	1.5
Nakahara et al. [70]	Cross-sectional, prospective, Japan	40	49 (27–76)	40	20 (50.0)	0.5
Nakano et al. [71]	Cross-sectional, Japan	88	47 (31–72)	88	65 (73.9)	1.5
Newel et al. [72]	Cross-sectional, retrospective, Japan	116	53 (31–88)	116	88 (75.9)	1.5
Nunes et al. [73]	Cross-sectional, prospective, USA	454	NA	454	192 (42.3)	1.5
Obdeijn et al. [74]	Cross-sectional, Netherland	54	NA	54	33 (61.1)	1.5

Table 1 continued

First author	Design and settings	N	Mean age (range)	Number total of lesions	Breast cancer prevalence (%)	MRI technique (T)
Orel et al. [75]	Cross-sectional, prospective, USA	175	56 (35–88)	175	63 (36.0)	1.5
Palmedo et al. [76]	Cross-sectional, prospective, Germany	56	58 (22–81)	56	27 (48.2)	1.5
Partridge et al. [77]	Cohort prospective, USA	91	49 (22–76)	116	29 (25.0)	1.5
Pediconi et al. [78]	Cross-sectional, retrospective, Italy	98	47.6 (16–46)	98	56 (57.2)	1.5
Reinkainen et al. [79]	Cross-sectional, prospective, Finland	40	49 (23–81)	40	25 (62.5)	1.5
Sardanelli et al. [80]	Cross-sectional, retrospective	67	54 (27–77)	67	50 (74.6)	1.5
Schedel et al. [81]	Cross-sectional, Germany	90	52 (21–78)	90	67 (74.4)	1.5
Schimitz et al. [82]	Cross-sectional, prospective USA	54	51.5	56	25 (44.6)	3
Sherif et al. [83]	Cross-sectional, prospective Egypt	62	55 (23–79)	79	55 (69.6)	1.5
Siegmann et al. [84]	Cross-sectional, prospective Germany	57	52.4 (22–76)	57	37 (64.9)	1.5
Sinha et al. [85]	Cross-sectional, retrospective, USA	53	50	43	20 (46.5)	1.5
Sivarajan et al. [86]	Cross-sectional, prospective, Malaysia	41	47 (21–72)	43	14 (32.5)	1.5
Stomper et al. [87]	Cross-sectional, prospective USA	49	54 (26–82)	51	25 (49.0)	1.5
Tilling et al. [88]	Cross-sectional, prospective Germany	40	52 (39–70)	40	22 (55.0)	1.5
Trecate et al. [89]	Cross-sectional, prospective, Italy	28	33–65	28	15 (53.6)	1.5–3
Turkat et al. [90]	Cross-sectional, prospective,	35	NA	39	15 (38.5)	1.5
Vassiou et al. [91]	Cross-sectional, prospective, Greece	69	54 (39–68)	78	53 (67.9)	1.5
Wiberg et al. [92]	Cross-sectional, prospective, Sweden	93	54 (33–81)	114	82 (71.9)	1.5
Wurdinger et al. [93]	Cross-sectional, prospective, Germany	424	54 (16–86)	488	234 (47.9)	1.5
Yabuuchi et al. [94]	Cross-sectional, prospective, Japan	22	55 (33–82)	22	15 (68.1)	1.5
Yuen et al. [8]	Cross-sectional, prospective, Japan	70	48	70	57 (81.4)	1.5
Zhu et al. [95]	Cross-sectional, prospective, Japan	55	NA	52	26 (50.0)	1.5
Total		9298		10493		

All patients were submitted to breast MR with 0.5 or 3 T system (index test) and biopsy for histology (reference standard)
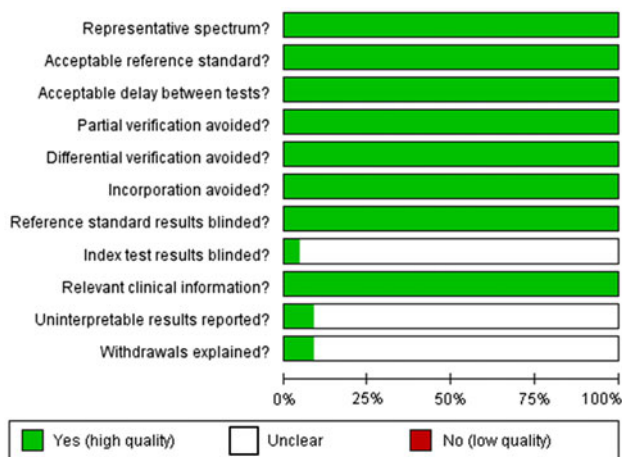
Fig. 2 Quality assessment by QUADAS [10]

Fig. 3 Summary ROC curve with summary Q* point

Sensitivity analysis

The robustness of the results was tested by repeating the analysis and excluding studies that were published before the year 2000. We found 42 studies meeting this exclusion
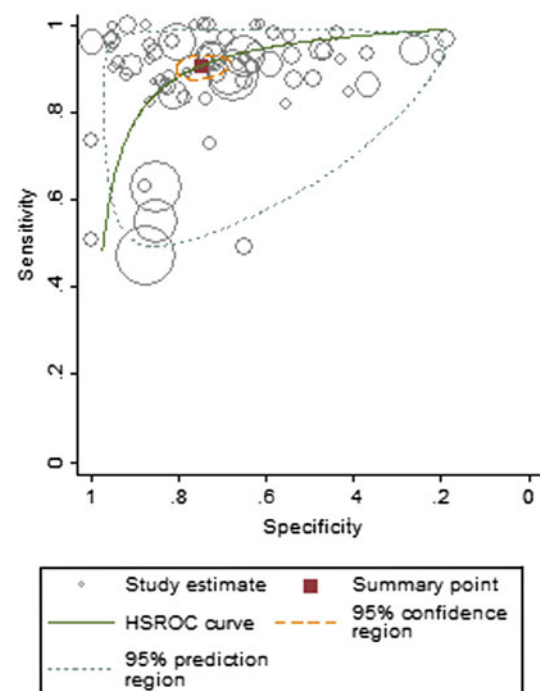
Table 2 Contingency table

Study	True positive	False positive	False negative	True negative
Abdolmaleki et al. [28]	82	12	12	14
Alamo and Fisher [29]	23	4	0	13
Bagni et al. [30]	35	4	3	3
Bazzocchi et al. [31]	75	12	5	20
Berg et al. [32]	167	60	10	21
Blohmer et al. [33]	37	10	4	17
Blumke et al. [34]	356	136	48	281
Boetes et al. [35]	62	3	3	19
Boetes et al. [36]	139	0	6	63
Bone et al. [37]	144	26	11	69
Bone et al. [38]	134	26	11	67
Buado et al. [39]	58	5	1	7
Cecil et al. [40]	19	2	4	13
Choi et al. [41]	61	5	4	29
Cilotti et al. [42]	19	11	22	25
Daldrup et al. [43]	27	7	10	19
Daniel et al. [44]	21	2	2	32
Dietzel et al. [45]	306	54	342	382
Fischer et al. [46]	375	50	30	93
Fischer et al. [47]	324	41	190	238
Fobben et al. [48]	18	12	3	58
Gatzemeier et al. [49]	88	17	3	4
Gibss et al. [50]	27	10	5	7
Gilles et al. [51]	81	45	5	41
Goto et al. [52]	124	38	20	22
Harms et al. [53]	44	17	3	10
Heiberg et al. [54]	23	16	0	44
Heinisch et al. [55]	23	4	2	11
Helbich et al. [56]	25	9	1	40
Hickman et al. [57]	4	1	0	20
Huang et al. [58]	18	12	0	20
Hulka et al. [59]	20	10	4	37
Ikeda et al. [60]	34	6	7	17
Jacobs et al. [61]	16	4	2	14
Kelcz et al. [62]	27	6	4	31
Khouli et al. [63]	68	8	10	15
Kinkel et al. [64]	33	1	1	22
Kuhl et al. [65]	92	42	9	61
Liu et al. [66]	64	13	5	37
Lucht et al. [67]	136	19	26	83
Morris et al. [68]	35	31	5	30
Moy et al. [69]	15	8	0	92
Nakahara et al. [70]	18	1	2	19
Nakano et al. [71]	63	7	2	16
Newel et al. [72]	78	8	10	20
Nunes et al. [73]	184	52	8	210
Obdeijn et al. [74]	30	7	3	14

Table 2 continued

Study	True positive	False positive	False negative	True negative
Orel et al. [75]	57	10	6	102
Palmedo et al. [76]	25	23	2	6
Partridge et al. [77]	27	40	2	47
Pediconi et al. [78]	54	2	2	40
Reinkainen et al. [79]	23	2	2	13
Sardanelli et al. [80]	44	3	6	14
Schedel et al. [81]	33	8	34	15
Schimitz et al. [82]	25	8	0	23
Sherif et al. [83]	28	0	27	24
Siegmann et al. [84]	36	9	1	11
Sinha et al. [85]	19	1	1	22
Sivarajan et al. [86]	12	5	2	24
Stomper et al. [87]	23	9	2	17
Tilling et al. [88]	18	8	4	10
Trecate et al. [89]	15	5	0	8
Turkat et al. [90]	15	3	0	21
Vassiou et al. [91]	52	14	1	11
Wiberg et al. [92]	77	17	5	15
Wurdinger et al. [93]	207	80	27	174
Yabuuchi et al. [94]	13	1	2	6
Yuen et al. [8]	42	0	15	13

criterion and performed the same bivariate calculus using the revised group of studies. Pooled sensitivity was 90% (95% CI 85–92%) and specificity was 73% (95% CI, 67–79%). The pooled positive likelihood was 3.4 (95% CI 2.7–4.2) and the negative likelihood was 0.14 (95% CI 0.10–0.19). The pooling of sensitivity, specificity, and likelihood ratios from the 42 studies published before the year 2000 did not alter the accuracy rate for breast MR diagnoses of benign breast lesions, high-risk breast lesions, or breast cancer. Therefore, the 69 studies were included in the sensitivity analysis [8, 28–95].

Discussion

The results of this meta-analysis show that breast MR is an accurate, non-invasive test for identifying breast cancer with high sensitivity and specificity (90 and 75%, respectively), both with a very restricted confidence interval. Breast MR has been shown to be the best imaging modality for determining the extent of index lesions in the breast before surgery [75]. Breast MR may reveal unsuspected multifocal, multicentric, or contralateral breast carcinoma and may result in therapy changes [46]. Therefore, breast

Table 3 Diagnostic performance of magnetic resonance mammary test

Study	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood positive (95% CI)	Likelihood negative (95% CI)
Abdolmaleki et al. [28]	0.87 (0.79–0.93)	0.54 (0.33–0.73)	1.89 (1.24–2.88)	0.24 (0.13–0.45)
Alamo and Fisher [29]	1.00 (0.85–1.00)	0.76 (0.50–0.93)	3.92 (1.76–8.74)	0.03 (0.002–0.44)
Bagni et al. [30]	0.92 (0.79–0.98)	0.43 (0.10–0.82)	1.61 (0.84–3.08)	0.18 (0.05–0.73)
Bazzocchi et al. [31]	0.94 (0.86–0.98)	0.63 (0.44–0.79)	2.50 (1.59–4.92)	0.10 (0.04–0.24)
Berg et al. [32]	0.94 (0.90–0.97)	0.26 (0.17–0.37)	1.27 (1.11–1.46)	0.22(0.11–0.44)
Blohmer et al. [33]	0.90 (0.77–0.97)	0.63 (0.42–0.81)	2.44 (1.48–4.03)	0.16 (0.06–0.41)
Blumke et al. [34]	0.88 (0.85–0.91)	0.67 (0.63–0.72)	2.70 (2.34–3.12)	0.18 (0.13–0.23)
Boetes et al. [35]	0.95 (0.87–0.99)	0.86 (0.65–0.97)	6.99 (2.44–20.05)	0.05 (0.02–0.16)
Boetes et al. [36]	0.96 (0.91–0.98)	1.00 (0.94–1.00)	122.30 (7.73–1934.6)	0.05 (0.02–0.10)
Bone et al. [37]	0.93 (0.88–0.96)	0.73 (0.63–0.81)	3.40 (2.44–4.72)	0.10 (0.06–0.18)
Bone et al. [38]	0.92 (0.87–0.96)	0.72 (0.62–0.81)	3.31 (2.38–4.60)	0.11 (0.06–0.19)
Buado et al. [39]	0.98 (0.91–1.00)	0.58 (0.28–0.85)	2.36 (1.21–4.61)	0.03 (0.004–0.22)
Cecil et al. [40]	0.83 (0.61–0.95)	0.87 (0.60–0.98)	6.20 (1.68–22.82)	0.20 (0.08–0.50)
Choi et al. [41]	0.94 (0.85–0.98)	0.85 (0.69–0.95)	6.38 (2.83–14.37)	0.07 (0.03–0.19)
Cilotti et al. [42]	0.46 (0.31–0.63)	0.69 (0.52–0.84)	1.51 (0.84–2.74)	0.77 (0.54–1.11)
Daldrup et al. [43]	0.73 (0.56–0.86)	0.73 (0.52–0.88)	2.71 (1.40–5.26)	0.37 (0.21–0.67)
Daniel et al. [44]	0.91 (0.72–0.99)	0.94 (0.80–0.99)	15.52 (4.02–59.90)	0.09 (0.03–0.35)
Dietzel et al. [45]	0.47 (0.43–0.51)	0.88 (0.84–0.91)	3.81 (2.93–4.96)	0.60 (0.56–0.65)
Fischer et al. [46]	0.93 (0.90–0.95)	0.65 (0.57–0.73)	2.65 (2.11–3.32)	0.11 (0.08–0.160)
Fischer et al. [47]	0.63 (0.59–0.67)	0.85 (0.81–0.89)	4.29 (3.21–5.64)	0.43 (0.83–0.49)
Fobben et al. [48]	0.86 (0.64–0.97)	0.83 (0.72–0.91)	5.00 (2.90–8.61)	0.17 (0.06–0.49)
Gatzemeier et al. [49]	0.97 (0.91–0.99)	0.19 (0.05–0.42)	1.20 (0.97–1.48)	0.17 (0.04–0.72)
Gibss et al. [50]	0.84 (0.67–0.95)	0.41 (0.18–0.67)	1.43 (0.94–2.19)	0.38 (0.14–1.02)
Gilles et al. [51]	0.94 (0.87–0.98)	0.48 (0.37–0.59)	1.80 (1.46–2.22)	0.12 (0.05–0.29)
Goto et al. [52]	0.86 (0.79–0.91)	0.37 (0.25–0.50)	1.36 (1.11–1.67)	0.38 (0.22–0.64)
Harms et al. [53]	0.94 (0.82–0.99)	0.37 (0.19–0.58)	1.49 (1.10–2.00)	0.17 (0.05–0.57)
Heiberg et al. [54]	1.00 (0.85–1.00)	0.73 (0.60–0.84)	3.62 (2.39–5.49)	0.03 (0.002–0.45)
Heinisch[55]	0.92 (0.74–0.99)	0.73 (0.45–0.92)	3.45 (1.48–8.05)	0.11 (0.03–0.43)
Helbich et al. [56]	0.96 (0.80–1.00)	0.82 (0.68–0.91)	5.24 (2.89–9.49)	0.05 (0.01– 0.32)
Hickman et al. [57]	1.00 (0.40–1.00)	0.95 (0.76–1.00)	13.20 (2.74–63.59)	0.11 (0.01–1.49)
Huang et al. [58]	1.00 (0.81–1.00)	0.63 (0.44–0.79)	2.37 (1.65–4.00)	0.04 (0.003–0.66)
Hulka et al. [59]	0.83 (0.63–0.95)	0.79 (0.64–0.89)	3.92 (2.20–6.98)	0.21 (0.09–0.52)
Ikeda et al. [60]	0.83 (0.68–0.93)	0.74 (0.52–0.90)	3.18 (1.58–6.41)	0.23 (0.11–0.47)
Jacobs et al. [61]	0.89 (0.65–0.99)	0.78 (0.52–0.94)	4.00 (1.66–9.64)	0.14 (0.04–0.54)
Kelcz et al. [62]	0.87 (0.70–0.96)	0.84 (0.68–0.94)	5.37 (2.55–11.31)	0.15 (0.06–0.39)
Khouli et al. [63]	0.87 (0.78–0.94)	0.65 (0.43–0.84)	2.51 (1.42–4.42)	0.20 (0.10–0.38)
Kinkel et al. [64]	0.97 (0.85–1.00)	0.96 (0.78–1.00)	22.32 (3.28–151.93)	0.03 (0.004–0.21)
Kuhl et al. [65]	0.91 (0.84–0.96)	0.59 (0.49–0.69)	2.23 (1.76–2.84)	0.15 (0.08–0.29)
Liu et al. [66]	0.93 (0.84–0.98)	0.74 (0.60–0.85)	3.57 (2.23–5.72)	0.10 (0.04–0.23)
Lucht et al. [67]	0.84 (0.77–0.89)	0.81 (0.72–0.88)	4.51 (2.99–6.80)	0.20 (0.14–0.28)
Morris et al. [68]	0.88 (0.73–0.96)	0.49 (0.36–0.62)	1.72 (1.31–2.26)	0.25 (0.11–0.60)
Moy et al. [69]	1.00 (0.78–1.00)	0.92 (0.85–0.96)	11.51 (6.01–22.04)	0.03 (0.002–0.52)
Nakahara et al. [70]	0.90 (0.68–0.99)	0.95 (0.75–1.00)	18.00 (2.65–122.28)	0.11 (0.03–0.39)
Nakano et al. [71]	0.97 (0.89–1.00)	0.70 (0.47–0.87)	3.18 (1.71–5.92)	0.04 (0.01–0.18)
Newel et al. [72]	0.89 (0.80–0.94)	0.71 (0.51–0.87)	3.10 (1.72–5.60)	0.16 (0.09–0.30)
Nunes et al. [73]	0.96 (0.92–0.98)	0.80 (0.75–0.85)	4.83 (3.78–6.17)	0.05 (0.03–0.10)
Obdeijn et al. [74]	0.91 (0.76–0.98)	0.67 (0.43–0.85)	2.73 (1.48–5.04)	0.14 (0.04–0.42)
Orel et al. [75]	0.90 (0.80–0.96)	0.91 (0.84–0.96)	10.13 (5.58–18.41)	0.11 (0.05–0.22)

Table 3 continued

Study	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood positive (95% CI)	Likelihood negative (95% CI)
Palmedo et al. [76]	0.93 (0.76–0.99)	0.21 (0.08–0.40)	1.17 (0.94–1.45)	0.36 (0.08–1.62)
Partridge et al. [77]	0.93 (0.77–0.99)	0.54 (0.43–0.65)	2.03 (1.58–2.60)	0.13 (0.03–0.49)
Pediconi et al. [78]	0.96 (0.88–1.00)	0.95 (0.84–0.99)	20.25 (5.23–78.38)	0.04 (0.01–0.15)
Reinkainen et al. [79]	0.92 (0.74–0.99)	0.87 (0.6–0.98)	6.90 (1.89–25.20)	0.09 (0.02–0.35)
Sardanelli et al. [80]	0.88 (0.76–0.95)	0.82 (0.57–0.96)	4.99 (1.78–14.00)	0.15 (0.07–0.32)
Schedel et al. [81]	0.49 (0.37–0.62)	0.65 (0.43–0.84)	1.42 (0.77–2.60)	0.78 (0.53–1.14)
Schimitz et al. [82]	1.00 (0.86–1.00)	0.74 (0.55–0.88)	3.69 (2.07–6.59)	0.03 (0.002–0.41)
Sherif et al. [83]	0.51 (0.37–0.65)	1.00 (0.86–1.00)	25.45 (1.62–400.44)	0.50 (0.38–0.66)
Siegmann et al. [84]	0.97 (0.86–1.00)	0.55 (0.32–0.77)	2.16 (1.33–3.52)	0.05 (0.01–0.35)
Sinha et al. [85]	0.95 (0,75–1.00)	0.96 (0.78–1.00)	21.85 (3.21–148.97)	0.05 (0.01–0.35)
Sivarajan et al. [86]	0.86 (0.57–0.98)	0.83 (0.64–0.94)	4.97 (2.18–11.35)	0.17 (0.05–0.63)
Stomper et al. [87]	0.92 (0.74–0.99)	0.65 (0.44–0.83)	2.66 (1.55–4.56)	0.12 (0.03–0.48)
Tilling et al. [88]	0.82 (0.60–0.95)	0.56 (0.31–0.79)	1.84 (1.06–3.20)	0.33 (0.12–0.87)
Trecate et al. [89]	1.00 (0.78–1.00)	0.62 (0.32–0.86)	2.47 (1.28–4.76)	0.05 (0.003–0.81)
Turkat et al. [90]	1.00 (0.78–1.00)	0.88 (0.68–0.97)	6.92 (2.61–18.36)	0.04 (0.002–0.56)
Vassiou et al. [91]	0.98 (0.90–1,00)	0.44 (0.24–0.65)	1.75 (1.24–2.49)	0.04 (0.01–0.31)
Wiberg et al. [92]	0.93 (0.86–0.98)	0.47 (0.29–0.65)	1.77 (1.27–2.46)	0.13 (0.05–0.33)
Wurdinger et al. [93]	0.88 (0.84–0.92)	0.68 (0.62–0.74)	2.81 (2.33–3.39)	0.17 (0.12–0.24)
Yabuuchi et al. [94]	0.87 (0.59–0.98)	0.85 (0.42–1.00)	6.07 (0.98–37.65)	0.16 (0.04–0.59)
Yuen et al. [8]	0.74 (0.63–0.84)	1.00 (0.75–1.00)	20.52 (1.34–313.43)	0.28 (0.18–0.43)
Zhu et al. [95]	0.88 (0.70–0.98)	0.92 (0.75–0.99)	11.50 (3.02–43.86)	0.13 (0.04–0.36)
Pooled	0.90 (0.88–0.93)	0.75 (0.70–0.80)	3.64 (3.0–4.42)	0.12 (0.09–0.15)

MR is a useful preoperative test for prediction of the benign or malignant nature of breast lesions.

We identified four systematic reviews evaluating breast MR in the diagnosis of breast cancer [96–99]. Only one meta-analysis has used bivariate analysis and hierarchical summary ROC approach, which included studies that analyzed patients with suspect lesions in mammography (BI-RADS 4 or 5) and studies that used an MR system with a field strength of at least 1.5 T [99]. We consider our meta-analysis of breast MR in women with suspicious lesions of the breast to be an update and extension of Peters et al.'s meta-analysis [99] that retrieved 40 studies. We identified almost all of these studies as well as 29 more. When pooling sensitivities and specificities according to a random-effects model, Peters et al. [99] found sensitivity and specificity of 90 and 72%, respectively. In comparison, our covariate analysis with bivariate approach yielded sensitivities and specificities of 90 and 75% [99]. In the SROC, we found an AUC of 0.91 for breast MR, and the point $Q*$ was 0.84 ($Q*$ is the point on the SROC where sensitivity and specificity are equal). This showed the discriminatory power of breast MR for suspicious lesions of the breast.

The potentially limited specificity of breast MR imaging has been attributed to the fact that, in addition to cancer, many benign lesions as well as presumably normal breast tissue may enhance after administration of contrast material [51]. Enhancement has been seen in many benign lesions, including fibroadenomas, proliferative, and non-proliferative fibrocystic changes and mastitis [51]. Similarly, diagnoses of high-risk breast lesions such as radial scars, atypical ductal hyperplasia, and lobular carcinoma in situ may be suggested by breast MR [5]. Therefore, breast MR may be able to assess the malignant potential of borderline lesions, thus identifying those women for whom a surgical biopsy can be avoided [5]. Our systematic review found false-positive results in 10.8% of the lesions, most of them being fibroadenomas and high-risk breast lesions. The screening of high-risk women by breast MR has high sensitivity for detecting breast cancer compared to conventional techniques. Breast MR showed 81% sensitivity, compared with mammography (40%) and ultrasonography (43%) [2]. Annual breast MR screening should be offered to patients with the following characteristics: BRCA1, BRCA2 and TP53 mutation carriers, women at 50% risk for BRCA1, BRCA2 or TP53 mutation that runs in their family, women who have had previous mantle radiotherapy before age 30, women who have been diagnosed and treated for breast cancer (5.3%) or developed synchronous or metachronous contralateral breast cancer. In BRCA

mutation carriers, the risk of contralateral disease was reported to be 29% at 10 years and 40% overall [2].

Current indications for breast MR include the following: invasive carcinoma and ductal carcinoma in situ, to determine the extent of disease and the presence of multifocality and multicentricity, evaluation of breast carcinoma prior to surgical treatment, to define the relationship of the tumor to the fascia and its extension into the pectoralis major, serratus anterior, and/or intercostals muscles, evaluation of residual disease in post-lumpectomy patients whose pathology specimens demonstrate close or positive margins for residual disease, assessment of response to neoadjuvant chemotherapy and the extent of residual disease prior to surgical treatment [2]. In patients who are potential lumpectomy candidates, determining the extent of disease is important in surgical planning, as the goal is to completely excise the tumor with clean margins with as few surgical procedures as possible [101]. Breast MR has been beneficial in improving surgical planning when assessing for extent of disease, and, therefore, may be a valuable tool in this field of medicine [100].

Strengths and weaknesses of our review

We extracted and reconstructed diagnostic data collected from cross-sectional, retrospective, and prospective studies. The methodological quality of included studies was very high, and the main problems were non-reporting of blinding status for the index test, difficult-to-interpret results, and non-explained withdrawals [10]. Bivariate analysis was used to calculate pooled estimates of sensitivity, specificity, and likelihood ratios (LRs) along with 95% confidence intervals (CIs) for the summary estimates [15, 16]. The bivariate model preserves the 2-dimensional nature of diagnostic data and takes into account both within-study and between-study variability [15–20]. We adhered to the most recent guidelines for conducting diagnostic reviews, as described in the *Cochrane Diagnostic Reviewers' Handbook* [10, 11]. We used an extensive search strategy, but by using a methodological filter, we may have missed some relevant publications. We did not use a language restriction. There were quite a few discrepancies in the phase of abstract selection and the agreement on some items of QUADAS. All articles were evaluated by six reviewers independently for methodological quality, and consensus was reached by discussing disagreements on individual scores. We defined statistical homogeneity as overlapping 95% CIs combined with less than 20% variation between point estimates. The population of interest was adults presenting with breast tumors (benign and malignant) and high-risk lesions.

Recommendations

Diagnostic tests as first-line investigations in primary care need to be valid, and should be easy to be performed, well tolerated by patients and sensitive, especially in the setting of serious diseases. Our systematic review shows that breast MR has an excellent sensitivity for suspicious breast lesions and could help doctors to make decisions about surgical intervention for diagnosis. In future research, cancer location and stage of disease should be important factors in analysis, especially as tests that are able to diagnose early stages of breast cancer become increasingly used as tools to reduce the burden of the disease.

Conflict of interest   Authors declare no conflicts of interest.

References

1. Chatterji M, Mercado CL, Moy L (2010) Optimizing 1.5 Tesla and 3-Tesla dynamic contrast-enhanced magnetic resonance imaging of the breasts. Magn Reson Imaging Clin N Am 18:207–224
2. Sardanelli F, Boetes C, Borisch B, Decker T, Federico M, Gilbert FJ, Helbich T, Heywang-Köbrunner SH, Kaiser WA, Kerin MJ, Mansel RE, Marotti L, Martincich L, Mauriac L, Meijers-Heijboer H, Orecchia R, Panizza P, Ponti A, Purushotham AD, Regitnig P, Del Turco MR, Thibault F, Wilson R (2010) Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. Eur J Cancer 46:1296–1316
3. Lee CH, Dershaw D, Kopans D, Evans P, Monsees B, Monticciolo D, Brenner RJ, Bassett L, Berg W, Feig S, Hendrick E, Mendelson E, D'Orsi C, Sickles E, Burhenne LW (2010) Breast cancer screening with imaging recommendations from the society of breast imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. J Am Coll Radiol 7:18–27
4. Liberman L, Morris EA, Joo-Young Lee M, Kaplan JB, LaTrenta LR, Menell JH, Abramson AF, Dashnaw SM, Ballon DJ, Dershaw DJ (2002) Breast lesions detected on MR imaging: features and positive predictive value. AJR 179:171–178
5. Linda A, Zuiani C, Bazzochi M, Furlan A, Londero V (2008) Borderline breast lesions diagnose at core needle biopsy: can magnetic resonance mammography rule out associated malignancy? Preliminary results bases on 79 surgically excised lesion. Breast 17:125–131
6. Simmons RM, Osborne MP (2009) The evaluation of high risk and pre-invasive breast lesions and the decision process for follow up and surgical intervention. Surg Oncol 8:55–65
7. Houssami N, Ciatto S, Bilous M, Vezzosi V, Bianchi S (2007) Borderline breast core needle histology: predictive values for malignancy in lesions of uncertain malignant potencial (B3). Br J Cancer 96:1253–1257
8. Sydnor MK, Wilson JD, Hijaz TA, Massey HD, Shaw de Paredes ES (2001) Underestimation of the presence of breast

carcinoma in papillary lesions initially diagnosed at core-needle biopsy. Radiology 242(1):58–62

9. Bilous M (2008) Management of lesions of uncertain malignant potential on breast core needle histology: vacuum-assisted excision as an alternative to surgical excision. Breast 17:543–544

10. Whiting P, Weswood ME, Rutjes A, Reittsma JB, Bossuyt P, Kleijnen J (2006) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 6:9–17

11. Reitsma JB, Rujes AW, Whiting P, Vlassov VV, Leeflang MM, Deeks JJ (2009) Assessing methodological quality, chap 9. In: Deek JJ, Bossuyt PM, Gatsonis C (eds) Cochrane handbook for systematic review of diagnostic test accuracy. Version 1.0.0. The Cochrane Collaboration 2009. htpp://srdta.cochrane.org/en/authors.html. Accessed July 19 2010

12. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC (2003) Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Ann Intern Med 138:40–45

13. Altman DG (1999) Some common problems in medical research. In: Altman DG (ed) Practical statistics for medical research, 9th edn. Chapman, London, pp 396–439

14. Schlesselman JJ, Stolley PD (1982) Csse control studies. Design, conduct, analysis. Oxford University Press, New York, pp 174–177

15. Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC et al (1994) Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med 120:667–676

16. Reitsma JB, Glas AS, Rutjes AWS, Scholten R, Bossuy PM, Zwinderman AH (2005) Bivariate analysis of sensitivity and specificity produce informative summary measures in diagnostic reviews. J Clin Epidemiol 58:982–990

17. Littenberg B, Moses LE (1993) Estimating diagnostic accuracy from multiple conflicting reports. Med Decis Making 13:313–321

18. Moses LE, Shapiro D, Littenberg B (1993) Combining independent studies of a diagnostic test into a summary ROC curve: data analytic approaches and some considerations. Stat Med 12:1293–1316

19. Deeks JJ (2001) Systematic reviews of evaluation of diagnostic and screening tests. In: Egger M, Smith GD, Altman D (eds) Systematic reviews in health care: meta-analysis in context, 2nd edn. BMJ Publishing, London, pp 248–282

20. Gatsonis C, Paliwal P (2006) Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. AJR Am J Roentgenol 187:271–281

21. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM (2008) Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. Ann Intern Med 149:889–897

22. Jellema P, van der Windt DA, Bruinvels DJ, Kneepkens CM, van der Horst HE (2010) Value of symptoms and additional diagnostic tests for colorectal cancer in primary care. BMJ 340:c1269

23. van der Windt D, Jellema P, Mulder C, Kneepkens F, van der Hors H (2010) Diagnostic testing for celiac disease among patients with abdominal symptoms. JAMA 303:1738–1746

24. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB (2000) Diagnosis and screening. In: Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB (eds) Evidence-base medicine: how to practice and teach EBM, 3rd edn. Churchill Livingstone, London, England, pp 67–93

25. AV Zamora J, Muriel A, Khan KS, Coomarasamy A (2006) Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 6:31

26. Review Manager (RevMan) [Computer program] (2008) Version 5.0. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration

27. Stata Corporation (2009) Stata Statistical Software version 11. Stata Corporation, College Station, TX

28. Abdolmaleki P, Buadu LD, Naderimansh H (2001) Feature extraction and classification of breast cancer on dynamic magnetic resonance imaging using artificial network. Cancer Lett 171:183–191

29. Alamo L, Fisher U (2001) Contrast-enhanced color Doppler ultrasound characteristics in hypervacular breast tumors: comparison with MRI. Eur Radiol 11:970–977

30. Bagni B, Francescheto A, Casolo A, De Santis M, Bagni I, Pansini F, Di Leo C (2003) Scintimammography with 99mTc-MIBI and magnetic resonance imaging in the evaluation of breast cancer. Eur J Nucl Med Mol Imaging 30:1383–1388

31. Bazzocchi M, Zuiani C, Panizza P, Del Frate C, Soldano F, Isola M, Sardanelli F, Giusepptti GM, Simonetu G, Lattazio V, Del Maschio AD (2006) Contrast-enhanced breast MRI in patients with suspicious microcalcifications on mammography: results of multicenter trial. Am J Roentgenol 186:1723–1732

32. Berge WA, Gutierrez L, Nessaiver MS, Carter WB, Bhargavan M, Lewis RS, Ioffe OB (2004) Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in pre-operative assessment of breast cancer. Radiology 233:830–849

33. Blohmer JU, Oellinger H, Schmidt C, Hufnagl P, Felix R, Lictenegger W (1999) Comparison of various imaging methods with particular evaluation of color Doppler sonography for planning surgery for breast tumors. Arch Gynecol Obstet 262:159–171

34. Bluemke DA, Gatsonis CA, Chen MH, DeAngelis GA, DeBruhl N, Harms S, Heywang-Köbrunner SH, Hylton N, Kuhl CK, Lehman C, Pisano ED, Causer P, Schnitt SJ, Smazal SF, Stelling CB, Weatherall PT, Schnall MD (2004) Magnetic resonance imaging of the breast prior to biopsy. JAMA 292:2735–2742

35. Boetes C, Barentsz JO, Mus RD, van der Sluis RF, van Erning L, Hendriks JH, Holland (1994) MR characterization of suspicious breast lesions with a gadolinium-enhanced turboFLASH subtraction technique. Radiology 193:777–781

36. Boetes C, Strijk SP, Barentsz JO, van der Sluis RF, Ruijs JH (1997) False-negative MR imaging of malignant breast tumors. Eur Radiol 7:1231–1234

37. Boné B, Aspelin P, Bronge L, Isberg B, Perbeck L, Veress B (1996) Sensitivity and specificity of MR mammography with histopathological correlation in 250 breasts. Acta Radiol 37:208–213

38. Boné B, Péntek Z, Perbeck L, Veress B (1997) Diagnostic accuracy of mammography and contrast-enhanced MR imaging in 238 histologically verified breast lesions. Acta Radiol 38:489–496

39. Buadu LD, Murakami J, Murakami S, Hashiguchi N, Sakai S, Masuda K, Toyoshima S, Kuroki S, Ohno S (1996) Breast lesions: correlations of contrast medium enhancement patterns on MR images with histopatologic findings and tumor angiogenesis. Radiology 200:639–649

40. Cecil KM, Schanall MD, Siegelman ES, Lenkinski RE (2001) The evaluation of human breast lesions with magnetic resonance imaging and proton magnetic resonance spectroscopy. Breast Cancer Res Treat 68:45–54

41. Choi N, Han B, Choe TH, Kim HS (2005) Three-phase dynamic breast magnetic resonance imaging with two-way subtraction. J Comput Assist Tomogr 29:834–841

42. Cilotti A, Iacconi C, Marini C, Moreti M, Mazzotta D, Traino C, Naccarato AG, Piagneri V, Giaconi C, Bevilaqua G, Bartolozzi C (2007) Contrast-enhanced MR imaging in patients with BI-RADS 3–5 microcalcifications. Radiol Med 112:272–286

43. Daldrup H, Rydland J, Helbich TH, Bjornerud A, Turetschek K, Kvistad KA, Kaindl E, Link TM, Staudacher K, Shames D, Brasch RC, Haraldseth O, Rummeny EJ (2003) Quantification

of breast tumor microvascular permeability with feruglose-enhanced MR imaging: initial Phase II multicenter trial. Radiology 229:885–892

44. Daniel BL, Yen YF, Glover HG, Ikeda DM, Birdwell RL, Sawyer-Glover AM, Black JW, Black JW, Plevritis SK, Jeffrey SS, Herfkens RJ (1998) Breast disease:dynamic spiral MR imaging. Radiology 209:499–509

45. Dietzel M, Baltzer PA, Vag T, Herzog A, Gajda M, Camara O, Kaiser WA (2010) The adjacent vessel sign on breast MRI: new data and subgroup analysis for 1084 histologically verified cases. Korean J Radiol 11:178–186

46. Fischer U, Kopka L, Grabbe E (1999) Breast carcinoma:effect of preoperative contrast-enhanced MR imaging on the therapeutic approach. Radiology 213:881–888

47. Fisher DR, Baltzer P, Malich A, Wurdinger S, Freesmeyer MG, Marx C, Kaiser WA (2004) Is the "blooming sign" a promising additional tool to determine malignancy in MR mammography? Eur Radiol 14:394–401

48. Fobben ES, Rubin CZ, Kalisher L, Dembner AG, Seltzer MH, Santoro EJ (1995) Breast MR imaging with commercially available techniques: radiologic-pathologic correlation. Radiology 196:143–152

49. Gatzemeier W, Liersch T, Stylianou A, Buttler A, Becker H, Fisher U (1999) Präoperative MR-mammographie beim mammacarcinom. Sicht Der Chirurg 70:1460–1468

50. Gibbs P, Liney GP, Lowry M, Kneedhaw PJ, Turnbull LW (2004) Differentiation of benign and malignant sub-1 cm breast lesions using dynamic contrast enhanced MRI. Breast 13:115–121

51. Gilles R, Meunier M, Lucidarme O, Zafrani B, Guinebretière JM, Tardivon AA, Le Gal M, Vanel D, Neuenschwander S, Cl Arriagada R (1996) Clustered breast microcalcifications: evaluation by dynamic contrast-enhanced subtraction. MRI J Comput Assit Tomogr 20:9–14

52. Goto M, Ito H, Akazawa K, Kubota T, Kizu O, Yamada K, Nishimura T (2007) Diagnosis of breast tumors by contrast-enhanced MR imaging: comparison between the diagnostic performance of dynamic enhancement patterns and morphologic features. J Magn Reson Imaging 25:104–112

53. Harms SE, Flamig DP, Hesley KL, Meiches MD, Jensen RA, Evans WP, Savino DA, Wells RV (1993) MR imaging of the breast with rotating delivery of excitation off resonance clinical experience with pathologic correlation. Radiology 187:493–501

54. Heiberg EV, Perman WH, Herrmann VM, Janney CG (1996) Dyamic sequential 3D gadolinium-enhanced MRI of the whole breast. Magnet Reson Imaging 14:337–348

55. Heinisch M, Gallowitsch HJ, Mikosch P, Kresnik E, Kumnig G, Gomez I, Lind P, Umschaden HW, Gasser J, Forsthuber EP (2003) Comparison of FDG-PET and dynamic contast-enhanced MRI in the evaluation of suggestive breast lesions. Breast 12:17–22

56. Helbich TH, Becherer A, Trattnig S, Leitha T, Kelkar P, Seifert M, Gnant M, Staudenhererz A, Rudas M, Wolf G, Mostbeck GH (1997) Differentiation of benign and malignant breast lesions: MR imaging versus Tc-99 m sestamibi scintimammography. Radiology 202:421–429

57. Hickmann PF, Moore NR, Shepstone BJ (1994) The indeterminate breast mass: assessment using contrast enhanced magnetic resonance imaging. Br J Radiol 67:14–20

58. Huang W, Fisher PR, Dulaimy K, Tudorica LA, O'Hea B, Button TM (2004) Detection of breast malignancy: diagnostic MR protocol for improved specificity. Radiology 232(2):585–591

59. Hulka CA, Edmister WB, Smith BL, Tan L, Sgroi DC, Camphell T, Kopans D (1997) Dynamic echo-planar imaging of the breast: experience in diagnosing breast carcinoma and correlation with tumor angiogenesis. Radiology 205:837–842

60. Ikeda O, Yamashita Y, Morishita S, Kido T, Kitajima M, Okamura K, Fukuda S, Takahashi (1999) Characterization of breast masses by dynamic enhancer. MR Imaging Acta Radiol 40:585–592

61. Jacobs MA, Barker PB, Bluemke DA, Maranto C, Arnold C, Herskovits EH, Bhujwalla Z (2003) Benign and malignant breast lesions: diagnosis with multiparametric MR imaging. Radiology 229:225–232

62. Kelcz F, Furman-Haran E, Grobgeld D, Degani H (2002) Clinical testing of high-spatial resolution parametric contrast-enhanced MR imaging of the breast. Am J Roentgenol 179:1485–1492

63. Khouli RE, Macura KJ, Jacobs MA, Khalil TH, Kamel IR, Kamel IR, Dwyer A, Bluemke DA (2009) Dynamic contrast-enhanced MRI oh the breast: quantitative method for kinetic curve type assessment. Am J Roentgenol 193:W295–W300

64. Kinkel K, Helbich TH, Esserman LJ, Barclay J, Schwerin EH, Sickles EA, Hylton NM (2000) Dynamic high-apatial-resolution MR imaging of suspicious breast lesion: diagnostic criteria and interobserver variability. Am J Roentgenol 175:35–43

65. Kuhl CK, Mielcareck P, Klaschik P, Leutner C, Wardelmann E, Gieseke J, Schild HH (1999) Dynamic breast MR imaging are signal intensity time course data useful for differential diagnosis of enhancing lesions? Radiology 211:101–110

66. Liu PF, Debatin JF, Caduff RF, Kacl G, Garzoli E, Krestin GP (1998) Improved diagnostic accuracy in dynamic contrast enhanced MRI of the breast by combined quantitative and qualitative analysis. Br J Radiol 71:501–509

67. Lucht RE, Knopp MV, Brix G (2001) Classification of signal-time curves from dynamic MR mammography by neural networks. Magn Reson Imaging 19:51–57

68. Morris EA, Liberman L, Dershaw DD, Kaplan JB, LaTrenta LR, Abramson AF, Ballon DJ (2002) Preoperative MR imaging-guided needle localization of breast lesion. Am J Roentgenol 178:1211–1220

69. Moy L, Elias K, Patel V, Lee J, Babb JS, Toth HK, Mercado CL (2009) Is breast MRI helpful in the evaluation of inconclusive mammographic findings? Am J Roentgenol 193:986–993

70. Nakahara H, Namba K, Fukami A, Watanabe R, Maeda Y, Furusawa H, Matsu T, Akiyama F, Nakagawa H, Ifuku H, Nakahara M (2001) Three-dimensional MR imaging of mammogaphically detected suspicious microcalcifications. Breast Cancer 8:116–124

71. Nakano S, Yoshida M, Fujii K, Yorozuya K, Mouri Y, Kousaka, Fukutomi T, Kimura J, Ishiguchi T, Ohno K, Mizumoto T, Harao M (2009) Fusion of MRI and sonography image for breast cancer evaluation using real-time virtual sonography with magnetic navigation: first experience. Jpn J Clin Oncol 39:552–559

72. Newel D, Nie K, Chen JH, Hsu CC, Yu HJ, Nalcioglu O, Su MY (2010) Selection of diagnostic features on breast MRI to differentiate between malignant and benign lesions using computer-aided diagnosis: differences in lesions presenting as mass and non-mass-like enhancement. Eur Radiol 20:771–781

73. Nunes LW, Schnall MD, Orel SG (2001) Update of breast MR imaging architectural interpretation model Radiology 219:484–494

74. Obdeijn IM, Kuijpers TA, van Dijk P, Wiggers T, Oudkerk M (1996) MR lesion detection in a breast cancer population. J Magn Reson Imaging 6:849–854

75. Orel SG, Schnall M, Powell C, Hochman M, Solin LJ, Fowble BL, Torosian Rosato EF (1995) Staging of suspected breast cancer: effect of MR imaging and MR-guide biopsy. Radiology 196:115–122

76. Palmedo H, Grünwald F, Bender H, Schomburg A, Mallmann P, Krebs D, Biersack HJ (1996) Scintimammography with

technetium-99 m methoxyisobutylisonitrile: comparison with mammography and magnetic resonance imaging. Eur J Nucl Med 23:940–946

77. Partridge SC, Mullins CD, Kurland BF, Allain MD, DeMartini WB, Eby PR, Lehman CD (2010) Apparent diffusion coefficient values for discriminating benign and malignant breast MRI lesions: effects of lesion type and size. Am J Roentgenol 194:1664–1673

78. Pediconi F, Catalano C, Roselli A, Dominelli V, Cagioli S, Karatasiou A, Pronio AM, Kirchin AM (2009) The challenge of imaging dense breast parenchyma. Invest Radiol 44:412–421

79. Reinikainen H, Pääkkö E, Suramo I, Päivänsalo M, Jauhiainen J, Rissanen T (2002) Dynamics of contrast enhancement in MR imaging and Power Doppler ultrasonography of solid breast lesions. Acta Radiol 43:492–500

80. Sardanelli F, Lozzeli A, Fausto A, Carriero A, Kirchin MA (2005) Gadobenate dimeglumine-enhanced MR imaging breast vascular maps: association between invasive cancer and ipsilateral increased vascularity. Radiology 235:791–797

81. Schedel H, Oellinger H, Kohlschein P, Siewert C, Hadjuana J, Blohmer JU, Kissner T, Felix R (2002) Magnetic resonance female breast imaging—evaluation of the changes in signal intensity over time pre- and post-administration of 0, 2 mmol/kg Gd-DTPA. Zentralbi Gynakol 124:104–110

82. Schmitz AC, Peters NG, Gallardo FAM, van Diest PJ, Stapper G, van Hillegersberg R, Mali WP, van den Bosch MA (2008) Contrast-enhanced 3.0-T breast MRI for characterization of breast lesions: increased specificity by using vascular maps. Eur Radiol 18:355–364

83. Sherif H, Mahfouz AE, Oellinger H, Hadijuana J, Blohmer JU, Taupitz, Felix R, Hamm B (1997) Peripheral washout sign on contrast-enhanced MR images of the Breast. Radiology 205:209–213

84. Siegmann KC, Xydeas T, Sinkus R, Kraemer B, Vogel U, Claussen CD (2010) Diagnostic value of MR elastography in addition to contrast-enhanced MR imaging of the breast—initial clinical results. Eur Radiol 20:318–325

85. Sinha S, Lucas-Quesada FA, DeBruhi N, Sayre J, Farria D, Gorczyca DP, Bassett LW (1997) Multifeature analysis of Gd-enhanced MR images of Breast Lesions. J Magn Reson Imaging 7:1016–1026

86. Sivarajan U, Jayapragasam KJ, Aziz A, Rahmat K, Bux SI (2009) Dynamic contrast enhancement magnetic resonance imaging evaluation of breast lesions: a morphological and quantitative analysis. J HK Coll Radiol 12:43–52

87. Stomper PC, Herman S, Klippenstein DL, Winston JS, Edge SB, Arredondo MA, Mazurchuk RV, Blumenson LE (1995) Suspect breast lesions: findings at dynamic gadolinium-enhanced MR imaging correlated with mammographic and pathologic features. Radiology 197:387–395

88. Tiling R, Khalkhali I, Sommer H, Linke R, Moser R, Willemsen F, Pfluger T, Tatsch Hahn K (1998) Limited value of scintimammography and contrast-enhanced MRI in the evaluation of microcalcification detecded by mammography. Nucl Med Commun 19:55–62

89. Trecate G, Tess JD, Vergnaghi D, Bergonzi S, de Simone T, Mariani G, Musumeci R (2002) Breast microcalcifications studied with 3d contrast-enhanced high field magnectic resonance imaging: more accuracy in the diagnosis of breast cancer. Tumori 88:224–233

90. Turkat TJ, Klein BD, Polan R, Richman R. Dynamic MR (1994) Mammography: a technique for potentially reducing the biopsy rate for benign breast disease. J Magn Reson Imaging 4:563–568

91. Vassiou K, Kanavou T, Vlychou M, Poultsidi A, Athanasiou E, Arvanitis DL, Fezoulidis IV (2009) Characterization of breast lesions with CE-MR multimodal morphological and kinetic analysis: comparison with conventional mammography and high-resolution ultrasound. Eur J Radiol 70:69–76

92. Wiberg MK, Aspelin P, Perbeck L, Boné B (2002) Value of MR imaging in clinical evaluation of breast lesion. Acta Radiol 43:275–281

93. Wurdinger S, Kamprath S, Eschrich D, Schneider A, Kaiser WA (2001) False-negative findings of malignant breast lesions on preoperative magnetic resonance mammography. Breast 10:131–139

94. Yabuuchi H, Matsuo Y, Kamitani T, Setoguchi T, Okafuji T, Soeda H, Sakai S, Hatakenaka M, Kubo M, Tokunaga E, Yamamoto H, Honda H (2010) Non-mass-like enhancement on contrast-enhanced breast MR imaging: lesion characterization using combination of dynamic contrast-enhanced and diffusion-weighted MR images. Eur J Radiol 75:e126–e132

95. Zhu J, Kurihara Y, Kanemaki Y, Ogata H, Fukuda M, Nakajima Y, Maeda I (2007) Diagnostic accuracy of high-resolution MRI using a microscopy coil for patients with presumed DCIS following mammography screening. J Magn Reson Imaging 25:96–103

96. Houssami N, Ciatto S, Macaskill P, Lord S, Warren RM, Dixon JM, Irwig L (2008) 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 26:3248–3258

97. Hrung JM, Sonnad SS, Schwartz S, Langlotz CP (1999) Accuracy of MR imaging in the work-up of suspicious breast lesions: a diagnostic meta-analysis. Acad Radiol 6:387–397

98. Warner E, Messersmith H, Causer P, Eisen A, Shumak R, Plewes D (2008) Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. Ann Intern Med 148:671–679

99. Peters NH, Rinkes IH, Zuithorff N, Mali W, Moons K, Peeters P (2008) Meta-analysis of MR imaging in the diagnosis of breast lesions. Radiology 246:116–124

100. Yeh ED (2010) Breast magnetic resonance imaging: current clinical indications. Magn Reson Imaging Clin N Am 18:155–169

101. Goscin C, Berman CG, Clark RA (2001). Magnectic resonance imaging of the breast. Cancer Control 8:399–405