Published in final edited form as:

Breast Cancer Res Treat. 2020 June ; 181(2): 255–268. doi:10.1007/s10549-020-05637-y.

Surveillance for second breast cancer events in women with a personal history of breast cancer using breast MRI: a systematic review and meta-analysis

Cameron B. Haas^{1,2}, Larissa Nekhlyudov³, Janie M. Lee⁴, Sara H. Javid⁴, Mary Bush², Dianne Johnson², Timothy Gleason⁵, Cary Kaufman⁶, Jennifer Specht⁴, Sean Stitham⁷, Karen J. Wernli²

¹Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA

²Kaiser Permanente Washington Health Research Institute, 1730 Minor Ave. Ste 1600, Seattle, WA 98101, USA

³Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA

⁴Department of Radiology, School of Medicine, University of Washington, Seattle, WA, USA

⁵Swedish Medical Center, Seattle, WA, USA

⁶Bellingham Regional Breast Center, Bellingham, WA, USA

⁷Kaiser Permanente Washington, Bellevue, WA, USA

Abstract

Purpose—Women with personal history of breast cancer (PHBC) are currently recommended to receive annual mammography for surveillance of breast cancer recurrence or new primary. However, given issues in accuracy with mammography, there is a need for evolving evidence-based surveillance recommendations with supplemental imaging. In this systematic review, we compiled and compared existing studies that describe the test performance of surveillance breast MRI among women with PHBC.

Methods—We searched PubMed and EMBASE using MeSH terms for studies (2000–2019) that described the diagnostic characteristics of breast MRI in women with PHBC. Search results were reviewed and included based on PICOTS criteria; quality of included articles was assessed using QUADAS-2. Meta-analysis of single proportions was conducted for diagnostic characteristics of breast MRI, including tests of heterogeneity.

Cameron B. Haas, Cameron.b.haas@kp.org.

Conflict of interest Authors CBH, LN, JML, SHJ, MB, DJ, CK, JS, SS, and KJW declare that they have no conflicts of interest. Dr Gleason owns stock in Radia Inc.

Compliance with ethical standards

Research involving human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10549-020-05637-y>) contains supplementary material, which is available to authorized users.

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

Results—Our review included 11 articles in which unique cohorts were studied, comprised of a total of 8338 women with PHBC and 12,335 breast MRI done for the purpose of surveillance. We predict intervals (PI) for cancer detection rate per 1000 examinations (PI 9–15; $\hat{P} = 10\%$), recall rate (PI 5–31%; $\hat{P} = 97\%$), sensitivity (PI 58–95%; $\hat{P} = 47\%$), specificity (PI 76–97%; $\hat{P} = 97\%$), and PPV3 (PI 16–40%; $\hat{P} = 44\%$).

Conclusions—Studies addressing performance of breast MRI are variable and limited in population-based studies. The summary of evidence to date is insufficient to recommend for or against use of breast MRI for surveillance among women with PHBC.

Keywords

Breast imaging; Diagnostic performance; Magnetic resonance imaging; Systematic review; Meta-analysis

Introduction

Nearly 3.5 million US women are living with personal history of breast cancer (PHBC) [1, 2] and nearly 90% of women diagnosed with breast cancer will survive at least 5 years [3]. Surveillance imaging after treatment for breast cancer aims to detect second breast cancer events, either a recurrent cancer or a second primary breast cancer tumor [4]. Detection of second breast cancer events in asymptomatic women is associated with earlier stage tumors of smaller size without node metastases and ultimately lower breast cancer mortality [5]. In observational studies in women with PHBC, surveillance mammography is associated with a 17–28% absolute reduction in breast cancer mortality [4–8]. Nonetheless, mammography has been shown to be less accurate in women with PHBC than women without PHBC [9]. Using data from the Breast Cancer Surveillance Consortium (BCSC), in a geographically and racially representative sample of the United States population, Houssami et al. estimated that the overall sensitivity of mammography in women with PHBC was 65.4% (95% confidence interval [CI] 61.5–69.0%) compared with 76.5% (95% CI 71.7–80.7%) in women without PHBC [9]. Recent results from Wernli et al. indicate that mammography improved sensitivity to 70.4% (95% CI 65.8–74.8); however, this still is not as high as performance seen in women without PHBC [10].

Given the limitations of surveillance mammography in detecting all second breast cancer events, breast MRI has emerged as a potential imaging tool as an adjunct to mammography. A prior systematic review of studies published from 1993 to 2006 assessed the performance characteristics of breast MRI for the detection of recurrence in 494 women (sample size ranged from 11 to 140 women) across 10 studies [11]. The authors concluded that, while a negative MRI is informative, a positive MRI is less useful because it requires further follow-up, and ultimately there was insufficient evidence at that time to recommend MRI in routine surveillance for breast cancer recurrence. The authors also noted several limitations of the studies, including clinically heterogeneous patients (e.g., differing in risk factors and presentation, surgical treatment, histology) and inclusion of breast MRI imaging examinations prior to the American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS) guideline for standardized MRI interpretation, as well as unreported false-positives among the included studies [12]. That study did not perform any

statistical tests for heterogeneity, and provided estimates for only sensitivity and specificity [11].

The BI-RADS Atlas, which provides standardized breast imaging terminology for describing imaging findings, was instrumental in improving consistent clinical assessments and classifications for reporting results of mammography, ultrasound and MRI of the breast [12]. In 2013, screening MRI performance benchmarks were added to the BI-RADS based on data from high-risk screening MRI trials in predominantly academic medical centers [13]. These benchmarks are target performance characteristics for breast cancer detection, including recommendations for measures of cancer detection rate, sensitivity, specificity, positive predictive values, and percent invasive cancers. Our objective was to systematically review the existing evidence for breast MRI performance for surveillance among women with PHBC for second cancer events and to assess whether the findings of our review met the clinical benchmarks.

Methods

Scope of review

Our review focused on the performance of supplemental breast MRI for women with PHBC (and without other high-risk indications for breast MRI), as the American Cancer Society guidelines reported in 2007 that there was insufficient evidence at that time to recommend for or against MRI surveillance for these women [13]. We conducted a systematic review of the existing literature with reported test performance measures of breast MRI in women with PHBC. We developed specific study guide criteria for eligible patients, interventions, comparators, outcomes, timing and setting (PICOTS) that would allow for comparability of study populations (Box 1) [14].

Our systematic review was registered with PROSPERO on February 29, 2016 (CRD42016035823) [15]. The research is not human subjects and does not require IRB approval.

Data sources

We searched PubMed and EMBASE for English-language publications from January 1, 2000, to June 6, 2019, using pre-specified MeSH (Medical Subject Headings) terms (eBox). Reference lists from included studies were also reviewed for potential study articles. The results of the search were imported to Covidence [16], an open access online tool developed specifically for performing systematic reviews.

After identifying potential articles, we reviewed all abstracts. Exclusion criteria were as follows: non-English, lacking full-text article, with insufficient detail, and publication outside of 2000–2019. Studies with insufficient detail provided in the abstract continued to full-text review. We further excluded non-English articles, and studies without full-text articles (Box 1).

Study selection

Three authors (CBH, LN, KJW) were involved in three phases of study selection. LN and CBH independently made inclusion/exclusion judgments based on the title and abstracts of all included articles. Excluded articles were justified by the reviewer according to pre-specified criteria (Box 1), with the option to add reasons if not listed. Any disagreement was resolved by KJW as an arbitrator. Studies that continued onto full-text screening were evaluated in the same manner. The primary reason for which a study was not included in the final review was due to “discrepant patient population” (eFigure). Specifically, these studies focused on women considered high risk of breast cancer based upon genetic mutation or family history. Full-text articles that met eligibility criteria and full-text review were included for data abstraction and quality assessment.

Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) [17] guided the quality assessment based on four domains (i.e., patient selection, index test, reference standard, and flow and timing). We scored either low or high risk of bias, and likewise evaluated the applicability of the study in answering our question of interest [14]. Studies with a high risk of bias and applicability for the same domain were excluded ($n = 3$).

Data extraction

The goal of our study was to estimate the performance of supplemental breast MRI. Key data abstracted from study findings (either as reported or calculated based on data provided in text) included the following: cancer rate, cancer detection rate, percent invasive cancer, sensitivity, specificity, recall, positive predictive value-recall (PPV1), and positive predictive value-biopsy (PPV3) (see Box 2). Among the studies included, there was an inconsistency in reporting prevalent (those found at first-time screening examination) and incident breast cancers (those found at subsequent screening examination performed at or close to the recommended screening interval) in relation to the second cancer diagnosis. We recorded whether prevalent or incident cancer type was used in each study, and subsequently used data to determine the cancer rate/cancer detection rate per patient and per examination.

In order for all false negatives and true positives to be identified, a minimum of 1-year follow-up was required according to BI-RADS definitions [12]. We assumed that this criterion was met when not specifically stated, as the majority of studies did not state.

Statistical analysis

Studies with available data for specific diagnostic test characteristics were included in the respective meta-analyses. Using meta-analysis for single proportions from the meta (Schwarzer, 2007) package in R (R Core Team, 2014), we calculated summary estimates using fixed and random effects models, as well as a prediction interval using the Hartung-Knapp and Paule-Mandel method for the random effects model. The random effects model accounts for the fact that study estimates do not stem from one single population and is expected to be more accurate with greater heterogeneity than the fixed effects estimates. To assess the heterogeneity of study estimates, we provided both the Cochran Q test statistic, which tests whether individual study estimates are farther away from the common effect than what would be expected by chance (summarized by the p value), and the I^2 (also known

as the Higgins ²), which describes the percentage of variation across studies that is due to heterogeneity rather than chance [18].

Results

Search results

Online library searches yielded 1210 unique citations (efigure). Based on abstract review 1156 were excluded according to criteria defined in the picots, and 54 full-text articles were reviewed. After review of full articles, 11 study articles from 10 non-overlapping study populations were included.

Description of studies

Most of studies included were conducted only in the United States ($n = 8/11$) and 8 were in academic medical settings. The American College of Radiology and Imaging Network (ACRIN) 6666 trial was a multi-center primarily based in the USA and included institutions in Buenos Aires, Argentina and Toronto, Canada [19]. Two studies, Cho et al. and Gweon et al., were conducted at institutions in the Republic of Korea.

Data was collected from 1999 [20] through 2016 [21]. The age range of women included was broad, ranging from 20 to 87 years of age, mean and median ages ranged from 43 to 54 years, and 44 to 57 years, respectively, at the time of MRI examination (Table 1).

The most common study design was a retrospective cohort design, with use of institutional electronic medical records to identify all screening breast MRIs during the time period of interest, and subsequent selection of patients with PHBC [21, 23–26, 28]. Gweon et al. specifically included only MRIs with a previously negative mammogram [25], and Weinstock et al. required a mammogram to have been done within 6 months of the MRI [28]. The ACRIN 6666 trial was primarily designed to compare breast ultrasound to mammography for detection of breast cancer in women at high risk. As a secondary endpoint, the investigators offered breast MRI to participants, and reported outcomes among women in the study with PHBC [29]. One study used a breast imaging registry of examinations done in clinical practice across several data collection sites [10].

Test characteristics

Cancer rate (per 1000 patients)—Cancer rates ranged from 14 to 118 per 1000 patients. Two studies, Weinstock et al. and Brennan et al., had the highest cancer rates estimated at 52 per 1000 women with PHBC (95% CI 28–88) and 118 per 1000 women, respectively (95% CI 70–182). The remaining studies from which cancer rate could be ascertained reported between 14 (95% CI 9–30) [24] and 29 (95% CI 11–62) [21]. Meta-analysis results predict that the true cancer rate in this population lies between 8 and 84 second breast cancer events per 1000 women. Across studies, tests for heterogeneity showed a substantial amount of difference between studies, with a Cochran Q test that indicates statistically significant heterogeneity. After removing Brennan et al. from the meta-analysis because of the extreme value reported (a product of longer follow-up and more tests performed per woman than

other studies), the p value was 0.09 and the predicted interval was 16.5–33.3 per 1000 women (data not shown) (Fig. 1).

Cancer detection rate (per 1000 examinations)—Cancer detection rates of second cancer events by breast MRI ranged from 8 to 20 per 1000 examinations. Meta-analysis results showed little heterogeneity between results, with an I^2 value of only 10% and a nonsignificant Cochran's Q test statistic. Fixed and random effects models were similar, with an estimated prediction interval between 9 and 15 cases detected by breast MRI per 1000 examinations (BI-RADS Benchmark: 20–30 per 1000 examinations) (Fig. 2).

Abnormal interpretation rate—Recall rates ranged from 6 to 31% among 10,743 examinations. Brennan et al. reported the highest percentage of examinations interpreted as positive, 31% (95% CI 23–39). The remaining studies ranged from 6 to 19%. There was significant evidence of heterogeneity between these included studies, and a wide prediction interval (5–32% of examinations indicating biopsy). When Brennan et al. was removed from the analysis, there was still significant heterogeneity, although the prediction interval was more refined to between 6 and 24% (data not shown). While AIR was reported in nearly all of the included studies, this measure was removed from the current edition of BI-RADS benchmarks (Fig. 3).

Sensitivity—Sensitivity for cancer detection ranged from a low of 61% (95% CI 19–99) to as high as 100% (95% CI 74–100). The BI-RADS Benchmark suggests a sensitivity of at least 80%. With a point estimate for the random effect model of 84%, our predicted interval based on the meta-analysis results shows the true value could lie somewhere between 58 and 95%. While the Cochran's Q Test did not indicate significant heterogeneity, the I^2 value estimated that there was 47% heterogeneity in these study estimates (Fig. 4).

Specificity—Specificity ranged from 81 to 95%. Confidence intervals were much tighter, given the much larger sample size of women who tested negative as compared to positive MRI results. Berg et al. population had the lowest specificity measure with 82% (95% CI 76–86) while Weinstock et al. reported the highest value with 95% (95% CI 93–97). Meta-analysis of the studies indicated significant heterogeneity, with a point estimate of 91% based on the random effect model, with the prediction interval ranging from 76 to 97% (Fig. 5).

PPV₁ (abnormal findings at surveillance)—Based on the 1034 examinations with BI-RADS 0, 3, 4, or 5 (considered positive findings), PPV₁ ranged from 6% to as high as 19%, with the largest study at 9%. Results from the meta-analysis based on included studies estimate between 5 and 19% of breast MRI's that are categorized as BI-RADS 0, 3, 4, or 5 were determined to be malignant. However, both tests of heterogeneity indicate significant differences between studies ($I^2 = 61\%$, $p < 0.01$). It should be noted that PPV₁ is based on AIR and no longer included as a BI-RADS benchmark in the current edition, but was commonly included in these studies (Fig. 6).

PPV₂ (biopsies recommended)—The positive predictive value of biopsies recommended from MRI was only reported in a single study. According to the results of

Sippo et al. then PPV2 of MRI in this population exceeded the BI-RADS benchmark of 15% (Fig. 7).

PPV₃ (biopsy performed)—PPV3 for malignancy ranged from 18 to 44% among positive MRI examinations requiring biopsy. BI-RADS Benchmark for PPV3 is between 20 and 50%, and the predicted interval based on the results of the meta-analysis fell between 16 and 40%. We observe significant heterogeneity between studies ($I^2 = 44\%$, $p = 0.04$). The point estimate based on the random effects model was 26%, with a confidence interval that remained above the lower limit of the BI-RADS Benchmark of 20% (95% CI 20–32%) (Fig. 8).

Percent invasive cancer—Only five studies reported values for percent node negative of invasive cancers on surveillance MRI, ranging from 75 to 100%. There was no evidence of heterogeneity for this characteristic between studies, likely due to the small sample sizes. However, the model point estimate was 89% (95% CI 80–94%), which is above the BI-RADS Benchmark of 80% (Fig. 9).

Percent of detected cancers ipsilateral—Among cancers detected by MRI, approximately half were in the same breast as the original cancer according the meta-analysis model estimates. Across studies this estimate was consistent, with the lowest estimate found in the smallest sample size. Tests of heterogeneity did not indicate significant differences in this finding among the included studies (Fig. 10).

Conclusions

Our systematic review of surveillance breast MRI performance among women with PHBC found wide variation in the test performance reported from several small, single institution studies, mostly from academic medical settings, and primarily set in the United States. To date, only one large community-based study has been conducted (in the USA). While we aimed to aggregate samples and test performance in a meta-analysis, we found evidence of significant heterogeneity across study populations and several measures [30–33]. However, despite the limitations, our study provides the first meta-analysis and statistical tests of heterogeneity for the diagnostic characteristics of breast MRI for surveillance among women with PHBC and offers insight into expected performance in this growing population. Comparisons with the BI-RADS benchmarks are important to consider when determining whether breast MRI in this population is meeting clinical performance expectations.

Few of these studies explicitly compared the performance of breast MRI to mammography, the primary imaging modality used for surveillance in women with PHBC. Berg et al. and Weinstein et al. showed an improvement in sensitivity with MRI at the cost of increasing false-positive rates among women who had received both imaging modalities [28, 29]. Consequently, Berg et al. also reported lower PPV1 and PPV3, and higher biopsy rates than with mammography alone [29]. In contrast, Wernli et al. compared the performance of mammography alone among 33,938 women with PHBC to the performance of MRI in 2506 women with PHBC who received mammography with adjunct MRI from a community-based setting and found the opposite, that sensitivity was greater for mammography only

(70.3%) versus mammography with breast MRI (61.4%) [10]. However, Wernli et al. also found that adjunct breast MRI resulted in 2.3-fold increased biopsy rates and no differences in interval cancer rate [10]. Use of adjunct MRI to mammography might detect more second breast cancers, which is dependent on additional biopsies. It remains unclear as to whether additional cancer detected improve women's health outcomes, beyond mammography alone.

Various factors may be responsible for the heterogeneity of results across studies and institutions. First, the diagnostic accuracy of mammography and MRI can vary based upon patient age, breast tissue density, frequency of screening, and radiologist experience. While some studies, such as the ACRIN 6666 population [29], set a minimum age at 25 years, most studies did not have such pre-specified requirements. Cho et al. focused on women under age 50 and thus was the only study which included a maximum age [22]. Studies had different reporting criteria as to whether a minimum of 1 year since treatment was set for the examination to be considered a surveillance examination. Among examinations in the Breast Cancer Surveillance Consortium, a national imaging registry, Wernli et al. also presents the years since primary breast cancer diagnosis of the included MRI and the years since prior surveillance examination, showing that the majority received their MRI within 3 years since diagnosis, and nearly half (48%) of examinations were within a year since the prior surveillance examination [10]. Additionally, the differences in prevalence of women with and without genetic risks due to either mutation carriers or family history dictate screening modality recommendations [34, 35]. Sippo et al. reported the prevalence of multiple risk factors among those with PHBC after excluding those with BRCA mutation or a history of chest radiation, and while the majority (88%) did not have other risk factors, 12% reported having a family history or high-risk lesion [27].

However, most studies did not account for these patient characteristics either through adjustment, stratification, or even by descriptive analysis, except in the BCSC cohort [10]. Further, while studies did exclude patients with known risk factors for breast cancer, there may have been unreported exposures associated with an increased risk such as chest irradiation. The two most recent studies with the largest sample sizes, Wernli et al. and Sippo et al. provide estimates of the prevalence of these various risk factors. From the BCSC, Wernli et al. present data on the mode of detection, AJCC stage, histology, and breast cancer treatment among the women included in the MRI group [10]. Surgical treatment of previous breast cancer is likely to be an important determinant of subsequent surveillance MRI diagnostic accuracy when comparing women treated with breast conserving therapy versus mastectomy. In a study by Elmore et al., 55% of women received breast conserving therapy and remaining women had received mastectomy [23], while Gweon et al. was conducted among women with breast conserving therapy [25]. Similarly, only a few studies reported the number who had received radiation or chemotherapy as part of the previous breast cancer treatment. Among women without a second cancer detected at MRI, 76% had received radiation, and in Wernli et al. 48% of patients with PHBC had received chemotherapy for their previous cancer. Given most studies did not include details regarding surgical treatment, this is a limitation that future studies should address.

We were limited by the information contained in the many of the articles, preventing the inclusion for specific diagnostic characteristics. Brennan et al. reported results based on total

patients, rather than results per examination which would have allowed for most of the characteristics to be calculated [20]. Further, the extreme outliers for some measures from the Brennan analysis suggest that symptomatic women or those perceived to have a second cancer might have received breast MRI deemed a surveillance examination, and skewing results towards higher cancer burden in the study sample. Studies were conducted predominantly in academic settings, except in the BCSC population-based setting (Table 1). Such settings are known to have improved detection from radiologists given the specialization in skill, and the unique patient population that differs from community-based settings. Most US women are not screened in an academic setting, presenting important differences between those included in these studies and the general population of women with PHBC. Women younger than fifty and screened at the end of the observation period for Gweon et al. could be overlapping with samples included in the Cho et al. study, which included Seoul National University as one of the multi-center sites [22]. However, because the majority of the time Gweon et al. did not overlap with Cho et al. and the exclusion of women older than 50 years was not made, we chose not to exclude the study from our review [25].

Despite the limitations, our review was comprehensive and provided an updated analysis of existing studies evaluating the test characteristics of breast MRI. We focused on women with PHBC without known germline BRCA mutation and thus excluded studies conducted exclusively in this high-risk population [7, 11, 36–41]. Adjustment of patient characteristics and/or subgroup analyses based on breast cancer risk factors, original cancer stage, molecular subtype, and treatment may more precisely measure the diagnostic characteristics of breast MRI and identify women who might benefit from breast MRI [42–47].

Variation in the reporting of results from breast MRI have changed over the course of the 17-year span (1999–2016) for the included studies and the release of the BI-RADS manual which included breast MRI in 2003. While contrast-enhanced breast MRI is routinely utilized in the US for screening women at high risk of developing breast cancer [48–50], a high proportion of women screened with breast MRI have a PHBC. The American Cancer Society, American Society of Clinical Oncology, the Society of Breast Imaging, and the National Comprehensive Cancer Network have not recommended breast MRI for routine screening of women with PHBC [34, 51–53], because of insufficient evidence on its benefits and harms [54–56]. However, in 2018, the American College of Radiology recommended women with PHBC diagnosed < 50 years or with dense breast tissue should receive annual breast MRI [57]. Women with PHBC who undergo surveillance MRI have been shown to experience higher biopsy rates and significantly lower cancer yield compared to mammography alone [10, 58]. The previous systematic review on this topic by Quinn et al. stated the limited clinical value of a positive MRI result in this study population due to the need to additional follow-up [11].

In summary, the performance of breast MRI in PHBC is widely variable, lacking evaluation within population-based settings, and outside of academic settings. While our estimates predict between 9 and 15 cancers could be detected per 1000 surveillance breast MRI, and a PPV3 between 16 and 40%, these may need to be interpreted with caution. Since no single study provides sufficient generalizability of test performance of breast MRI in this

population, our meta-analyses provide better summaries in relation to recommended BI-RADS benchmarks. The summary of evidence to date still is insufficient to recommend for or against use of breast MRI for surveillance among women with PHBC. Importantly, the significant heterogeneity shown by our statistical tests indicates the need for higher-quality studies to determine precise estimates of performance, and improved reporting of patient characteristics to support future meta-analyses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding This work was supported through a Patient-Centered Outcomes Research Institute (PCORI) Project Program Award (CE-1304-6656-Comparative effectiveness of surveillance imaging modalities in breast cancer survivors). The design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication is solely the responsibility of the authors and does not necessarily represent the official views of the Patient-Centered Outcomes Research Institute. This publication was supported by grant number T32 CA094880 from the National Institute of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NCI, NIH.

Abbreviations

PHBC	Personal history of breast cancer
BCSC	Breast Cancer Surveillance Consortium
MRI	Magnetic resonance imaging

References

1. American Cancer Society (2016) Cancer treatment & survivorship: facts & figures 2016–2017. American Cancer Society, Atlanta, p 44
2. Howlader N et al. (2016) SEER cancer statistics review, 1975–2013. National Cancer Institute, Bethesda
3. Howlader N et al. (2012) SEER cancer statistics review, 1975–2009 (Vintage 2009 Publications). 2012 http://seer.cancer.gov/csr/1975_2009_pops09/index.html Accessed 16 Jan 2013
4. Lash TL et al. (2007) Mammography surveillance and mortality in older breast cancer survivors. *J Clin Oncol* 25(21):3001–3006 [PubMed: 17548838]
5. Houssami N et al. (2009) Early detection of second breast cancers improves prognosis in breast cancer survivors. *Ann Oncol* 20(9):1505–1510 [PubMed: 19297316]
6. Lu WL et al. (2009) Impact on survival of early detection of isolated breast recurrences after the primary treatment for breast cancer: a meta-analysis. *Breast Cancer Res Treat* 114(3):403–412 [PubMed: 18421576]
7. Lu W et al. (2009) The value of surveillance mammography of the contralateral breast in patients with a history of breast cancer. *Eur J Cancer* 45(17):3000–3007 [PubMed: 19744851]
8. Sardanelli F et al. (2016) Position paper on screening for breast cancer by the European Society of Breast Imaging (EUSOBI) and 30 national breast radiology bodies from Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Israel, Lithuania, Moldova, The Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Switzerland and Turkey. *Eur Radiol*

9. Houssami N et al. (2011) Accuracy and outcomes of screening mammography in women with a personal history of early-stage breast cancer. *JAMA* 305(8):790–799 [PubMed: 21343578]
10. Wernli KJ et al. (2019) Surveillance breast MRI and mammography: comparison in women with a personal history of breast cancer. *Radiology* 292(2):311–318 [PubMed: 31161975]
11. Quinn EM, Coveney AP, Redmond HP (2012) Use of magnetic resonance imaging in detection of breast cancer recurrence: a systematic review. *Ann Surg Oncol* 19(9):3035–3041 [PubMed: 22476755]
12. American College of Radiology (2013) American College of Radiology (ACR) Breast Imaging Reporting and Data System Atlas (BI-RADS® Atlas). American College of Radiology (ACR), Reston
13. American Cancer Society (2015) American Cancer Society recommendations for early breast cancer detection in women without breast symptoms. 2015 October 20 <http://www.cancer.org/cancer/breastcancer/moreinformation/breastcancerearlydetection/breast-cancer-early-detection-ac-s-recs> Accessed 1 December 2015
14. Liu Z et al. (2013) A step-by-step guide to the systematic review and meta-analysis of diagnostic and prognostic test accuracy evaluations. *Br J Cancer* 108(11):2299–2303 [PubMed: 23695015]
15. Moher D et al. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 6(7):e1000097 [PubMed: 19621072]
16. Innovation VH. Covidence systematic review software
17. Whiting PF et al. (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 155(8):529–536 [PubMed: 22007046]
18. Higgins JP, Thompson SG, Spiegelhalter DJ (2009) A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc* 172(1):137–159
19. Berg WA et al. (2008) Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA* 299(18):2151–2163 [PubMed: 18477782]
20. Brennan S et al. (2010) Breast MRI screening of women with a personal history of breast cancer. *AJR Am J Roentgenol* 195(2):510–516 [PubMed: 20651211]
21. Schacht DV et al. (2014) Importance of a personal history of breast cancer as a risk factor for the development of subsequent breast cancer: results from screening breast MRI. *AJR Am J Roentgenol* 202(2):289–292 [PubMed: 24450667]
22. Cho N et al. (2017) Breast cancer screening with mammography plus ultrasonography or magnetic resonance imaging in women 50 years or younger at diagnosis and treated with breast conservation therapy. *JAMA Oncol* 3(11):1495–1502 [PubMed: 28655029]
23. Elmore L, Margenthaler JA (2010) Breast MRI surveillance in women with prior curative-intent therapy for breast cancer. *J Surg Res* 163(1):58–62 [PubMed: 20605594]
24. Giess CS et al. (2015) Screening breast MRI in patients previously treated for breast cancer: diagnostic yield for cancer and abnormal interpretation rate. *Acad Radiol* 22(11):1331–1337 [PubMed: 26142951]
25. Gweon HM et al. (2014) Breast MR imaging screening in women with a history of breast conservation therapy. *Radiology* 272(2):366–373 [PubMed: 24635678]
26. Lehman CD et al. (2016) Screening MRI in women with a personal history of breast cancer. *J Natl Cancer Inst* 108(3):349
27. Sippo DA et al. (2019) Performance of screening breast MRI across women with different elevated breast cancer risk indications. *Radiology* 292(1):51–59 [PubMed: 31063080]
28. Weinstock C et al. (2015) Breast magnetic resonance imaging (MRI) surveillance in breast cancer survivors. *Springerplus* 4:459 [PubMed: 26322264]
29. Berg WA et al. (2012) Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA* 307(13):1394–1404 [PubMed: 22474203]
30. Leflang MM (2014) Systematic reviews and meta-analyses of diagnostic test accuracy. *Clin Microbiol Infect* 20(2):105–113 [PubMed: 24274632]
31. Rutter CM, Gatsonis CA (2001) A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 20(19):2865–2884 [PubMed: 11568945]

32. Stroup DF et al. (2000) Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA* 283(15):2008–2012 [PubMed: 10789670]
33. Zwinderman AH, Bossuyt PM (2008) We should not pool diagnostic likelihood ratios in systematic reviews. *Stat Med* 27(5):687–697 [PubMed: 17611957]
34. Lee CH et al. (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(1):18–27 [PubMed: 20129267]
35. Ahern CH et al. (2014) Cost-effectiveness of alternative strategies for integrating MRI into breast cancer screening for women at high risk. *Br J Cancer* 111(8):1542–1551 [PubMed: 25137022]
36. Kramer S et al. (1998) Magnetic resonance imaging in the diagnosis of local recurrences in breast cancer. *Anticancer Res* 18(3C):2159–2161 [PubMed: 9703776]
37. Preda L et al. (2006) Magnetic resonance mammography in the evaluation of recurrence at the prior lumpectomy site after conservative surgery and radiotherapy. *Breast Cancer Res* 8(5):R53 [PubMed: 16959028]
38. Coulthard A, Beveridge CJ, Potterton AJ (1999) MRI in routine breast cancer follow-up: correlation with clinical outcome. *Clin Radiol* 54(7):459–461 [PubMed: 10437698]
39. Arazi-Kleinman T et al. (2013) JOURNAL CLUB: is screening MRI indicated for women with a personal history of breast cancer? Analysis based on biopsy results. *AJR Am J Roentgenol* 201(4):919–927 [PubMed: 24059385]
40. Rieber A et al. (1997) Value of MR mammography in the detection and exclusion of recurrent breast carcinoma. *J Comput Assist Tomogr* 21(5):780–784 [PubMed: 9294575]
41. Viehweg P et al. (1998) Retrospective analysis for evaluation of the value of contrast-enhanced MRI in patients treated with breast conservative therapy. *MAGMA* 7(3):141–152 [PubMed: 10050940]
42. Cuzick J et al. (2010) Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. *Lancet Oncol* 11(12):1135–1141 [PubMed: 21087898]
43. Dubsy PC et al. (2012) Tamoxifen and anastrozole as a sequencing strategy: a randomized controlled trial in postmenopausal patients with endocrine-responsive early breast cancer from the Austrian Breast and Colorectal Cancer Study Group. *J Clin Oncol* 30(7):722–728 [PubMed: 22271481]
44. Early Breast Cancer Trialists' Collaborative G et al. (2011) Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. *Lancet* 378(9804):1707–1716 [PubMed: 22019144]
45. Poggi MM et al. (2003) Eighteen-year results in the treatment of early breast carcinoma with mastectomy versus breast conservation therapy: the National Cancer Institute Randomized Trial. *Cancer* 98(4):697–702 [PubMed: 12910512]
46. Punglia RS, Hassett MJ (2010) Using lifetime risk estimates to recommend magnetic resonance imaging screening for breast cancer survivors. *J Clin Oncol* 28(27):4108–4110 [PubMed: 20697089]
47. Veronesi U et al. (2002) Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *N Engl J Med* 347(16):1227–1232 [PubMed: 12393819]
48. Bassett LW et al. (2008) National trends and practices in breast MRI. *AJR Am J Roentgenol* 191(2):332–339 [PubMed: 18647898]
49. Stout NK, Nekhlyudov L (2011) Early uptake of breast magnetic resonance imaging in a community-based medical practice, 2000–2004. *J Womens Health (Larchmt)* 20(4):631–634 [PubMed: 21413899]
50. Wernli KJ et al. (2014) Patterns of breast magnetic resonance imaging use in community practice. *JAMA Intern Med* 174(1):125–132 [PubMed: 24247555]

51. Khatcheressian JL et al. (2006) American Society of Clinical Oncology 2006 update of the breast cancer follow-up and management guidelines in the adjuvant setting. *J Clin Oncol* 24(31):5091–5097 [PubMed: 17033037]
52. Saslow D et al. (2007) American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin* 57(2):75–89 [PubMed: 17392385]
53. Mainiero MB et al. (2013) ACR appropriateness criteria breast cancer screening. *J Am Coll Radiol* 10(1):11–14 [PubMed: 23290667]
54. Khatcheressian JL et al. (2013) Breast cancer follow-up and management after primary treatment: American Society of Clinical Oncology clinical practice guideline update. *J Clin Oncol* 31(7):961–965 [PubMed: 23129741]
55. Morrow M, Waters J, Morris E (2011) MRI for breast cancer screening, diagnosis, and treatment. *Lancet* 378(9805):1804–1811 [PubMed: 22098853]
56. Sickles EA (2010) The use of breast imaging to screen women at high risk for cancer. *Radiol Clin North Am* 48(5):859–878 [PubMed: 20868890]
57. Monticciolo DL et al. (2018) Breast cancer screening in women at higher-than-average risk: recommendations from the ACR. *J Am Coll Radiol* 15(3):408–414 [PubMed: 29371086]
58. Buist DSM et al. (2018) Breast biopsy intensity and findings following breast cancer screening in women with and without a personal history of breast cancer. *JAMA Intern Med* 178(4):458–468 [PubMed: 29435556]

Box 1**Patients, interventions, comparators, outcomes, timing and setting (PICOTS)****Patients***Inclusion criteria*

Women with a personal history of breast cancer and had undergone curative-intent treatment

Exclusion criteria

Pregnant, lactating, or known metastatic disease

Non-human research participants

Male study participants

Interventions*Inclusion criteria*

Magnetic resonance imaging

Mammography

CBE

No screening

Ultrasound

Exclusion criteria

Screening modalities that do not include MRI as a study group

MRI for any purpose other than screening (i.e., diagnostic)

Comparators*Inclusion criteria*

MRI vs. Mammography (plus Ultrasound)

MRI vs. no screening

Ipsilateral vs. contralateral detection

MRI at various intervals

Women with a personal history of breast cancer vs. women without

Women with a personal history of breast cancer and a family history of breast cancer vs. no family history of breast cancer

Outcomes*Inclusion criteria*

Cancer detection rate (proportion of women with apposite screen result and a positive reference standard)

Cancer detection rate for ipsilateral and contralateral breast cancer

PPV1 (malignancy rate among cases that test positive on screening) or false-positive rate

PPV3 (rate of malignancy among cases with positive results on screening who underwent biopsy)

Sensitivity (TP/[TP + FP])

Specificity (TN/[TN + FN])

Number of screens needed to detect on cancer

Recall rate (proportion of women with a positive screen result)

Emotional impact (e.g., anxiety, depression) of positive results (true- and false-positive findings)

Exclusion criteria

Outcomes not listed

Economic outcomes only

Timing*Inclusion criteria*

Articles published after 2000

Studies of any duration

Settings

Inclusion criteria

Any setting where MRI breast cancer screening is provided

Study Design*Inclusion criteria*

Controlled studies (RCTs, cohort studies, case-control studies)

Modeling/simulation studies that meet previously stated inclusion/exclusion criteria

Observational studies (prospective and retrospective cohort studies, case-control studies, or cross-sectional studies)

Exclusion criteria

Systematic reviews and/or meta-analyses

Non-research studies (e.g., editorials, letters)

Other

English language

Peer-reviewed articles

Exclusion criteria

Non-English language

Abstracts only

Box 2

Definition of surveillance breast MRI examination performance measures. [26]

Measure	Definition	What does this mean?
True positive	Positive surveillance examination followed by diagnosis of ductal carcinoma in situ (DCIS) or invasive breast cancer within a follow-up period	A radiologist finding requires additional work-up; woman diagnosed with a second breast cancer during follow-up
False-positive	Positive surveillance examination with no breast cancer diagnosed by end of follow-up	A radiologist finding requires additional work-up; woman <i>not</i> diagnosed with a second breast cancer during follow-up
False negative	Negative surveillance examination followed by a diagnosis of DCIS or invasive breast cancer by the end of a follow-up period	No radiologist findings require additional work-up, but woman diagnosed with a second breast cancer event during follow-up
True negative	A negative surveillance examination with no diagnosis of breast cancer by the end of a follow-up period	No radiologist findings require additional work-up; woman is <i>not</i> diagnosed with a second breast cancer event during follow-up
Cancer rate	The proportion of women with breast cancer among the entire population receiving surveillance examinations	In the entire population, the percentage of breast cancer both detected and undetected by surveillance
Cancer detection rate	The proportion of women with breast cancer detected by screening among the entire population receiving surveillance examinations	In the entire population, percentage of breast cancer events detected by surveillance imaging
Sensitivity	The proportion of women with a second breast cancer event who have a positive surveillance examination	Percentage of second breast cancer events during follow-up detected by the surveillance examination
Specificity	The proportion of women without a second breast cancer event who have a negative surveillance examination	Percentage of examinations with no second breast cancer event during follow-up that were also negative on the surveillance examination
Positive predictive value (PPV1)	The probability that a woman has a second breast cancer event given that her surveillance examination is positive	Likelihood of a true second breast cancer event in a woman who screens positive with a breast MRI
Positive predictive value (PPV2)	The probability that a woman has a second breast cancer event among women recommended to receive a breast biopsy	Likelihood of a true second breast cancer event in a woman who screens positive with breast MRI and is recommended to receive a breast biopsy
Positive predictive value (PPV3)	The probability that a woman has a second breast cancer event among women who received a biopsy	Likelihood of a true second breast cancer event in a woman who screens positive with breast MRI and receives a breast biopsy

Fig. 1.
Cancer rate (per 1000 women)

Fig. 2.
Cancer detection rate (per 1000 examinations)

Fig. 3.
Abnormal interpretation rate (AIR, or Recall rate)

Fig. 4.
Sensitivity

Fig. 5.
Specificity

Fig. 6.
PPV1

Fig. 7.
PPV2

Fig. 8.
PPV3

Fig. 9.
Percent node negative invasive cancers

Fig. 10.
Percent of detected cancers ipsilateral

Table 1

Included studies for systematic review and meta-analyses

Author (Publication year)	Location	Type	Years	Study population	
				Women	Mean age
Berg (2012) [19]	ACRIN 6666 (USA, Canada, Argentina)	Multi-center academic	April 2004–February 2006	275	51
Brennan [20]	Memorial Sloan-Kettering Cancer Center (USA)	Academic	January 1, 1999–December 31, 2000	144	49
Cho (2017) [22]	Multi-center (Republic of Korea)	Academic	December 1, 2010–January 31, 2016	752	43
Elmore (2010) [23]	Washington University/Barnes Jewish Hospital (USA)	Academic	January 1, 2005–December 31, 2011	141	51
Giess (2015) [24]	Brigham and Woman's Hospital, Dana-Farber Institute (USA)	Academic	January 1, 2009–December 31, 2011	691	52
Gweon (2014) [25]	Seoul National University College of Medicine (Republic of Korea)	Academic	January 2009–March 2012	607	48 (median)
Lehman (2016) [26]	Clinical Oncology Data Integration project (USA)	Community	July 2004–November 2011	915	N/A
Schacht (2014) [21]	University of Chicago, Department of Radiology (USA)	Academic	2004–2012	208	52
Sippo (2019) [27]	Massachusetts General Hospital (USA)	Community	January 2011–December 2014	2835	54
Weinstock (2015) [28]	University of Maryland Medical Center (USA)	Academic	January 1, 2005–December 31, 2011	249	46
Wernli (2019) [10]	Breast Cancer Surveillance Consortium (USA)	Multi-center community	2005–2012	1521	51