diagnostics

MDPI

Systematic Review

Contrast-Enhanced Mammography for Preoperative Breast Cancer Staging: A Systematic Review of Additional Cancer Detection and Meta-Analysis of Surgical Management Changes

Romuald Ferré [1,*,†], Thad Benefield [2] and Cherie M. Kuzmiak [1,†]

[1] Division of Breast Imaging, Department of Radiology, UNC School of Medicine, Physicians' Office Building, Room 118, 170 Manning Drive, Chapel Hill, NC 27599, USA
[2] Carolina Mammography Registry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
* Correspondence: romuald@ad.unc.edu; Tel.: +1-984-974-8123
† These authors contributed equally to this work.

Abstract

Background/Objectives: This systematic review and meta-analysis evaluates additional cancer detection and observed changes in surgical management associated with preoperative contrast-enhanced mammography (CEM). **Methods:** PubMed/MEDLINE was searched through October 2026. Eligible reports evaluated preoperative CEM in adults with breast cancer. R.F. and C.M.K. assessed eligibility and extracted data; QUADAS-2 assessed diagnostic/staging studies. Only patient-level management-change outcomes were quantitatively pooled, using logit-scale random-effects models with restricted maximum likelihood and Hartung–Knapp 95% confidence intervals (CIs). Additional cancer detection and other outcomes were synthesized descriptively and narratively. Certainty was discussed qualitatively without formal GRADE ratings. PROSPERO: CRD420261551170. No external funding was received. **Results:** Thirteen studies were included. Four observational cohorts reported breast-surgery or localization changes in 149/713 patients; the pooled proportion was 21.8% (95% CI, 14.9–30.7%; $I^2 = 52.5\%$). Excluding a mixed neoadjuvant-treatment cohort yielded 22.5% (95% CI, 11.0–40.5%). An exploratory analysis of broader treatment changes yielded 22.1% (95% CI, 18.0–27.0%) across six cohorts and 892 patients. Additional-cancer detection varied across four lesion-level comparisons with magnetic resonance imaging (MRI); no pooled paired comparative effect was estimated. Single-database coverage, only four primary cohorts, imprecision, and selection and verification concerns limited confidence. **Conclusions:** Preoperative CEM was associated with observed changes in breast surgery or localization in selected cohorts. The pooled estimate of 21.8% represents management-change frequency and does not establish clinical benefit or diagnostic equivalence to MRI. Prospective comparative evidence is needed to determine whether these changes improve patient outcomes.

Keywords: breast cancer; contrast-enhanced mammography; preoperative staging; additional malignancy; surgical management; systematic review; meta-analysis

Check for updates

Academic Editor: Takouhie Catherine Maldjian

Received: 1 November 2026
Revised: 22 November 2026
Accepted: 22 November 2026
Published: 26 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Accurate preoperative assessment of breast cancer extent informs the choice and scope of breast-conserving surgery, localization, mastectomy, and contralateral evaluation. The clinical objective is to identify disease that changes appropriate treatment while limiting unnecessary biopsy and surgical escalation.

CEM combines low-energy mammographic morphology with dual-energy recombined images after intravenous iodinated contrast administration [1]. Previous evidence syntheses have evaluated its diagnostic performance [2], and a recent narrative review considered its role in staging and surgical planning [3]. However, index-cancer detection, additional cancer detection, and management change address different clinical questions. Their denominators and reference standards must remain explicit.

Breast magnetic resonance imaging (MRI) provides sensitive assessment of disease extent, but its use requires consideration of cost, availability, examination time, and patient tolerance [1,4]. Additional findings can require further imaging or biopsy before treatment is decided. A retrospective single-center study reported a mean biopsy-to-surgery interval of 55.8 days with preoperative MRI versus 42.8 days without it [5], whereas another institutional series found no significant difference in treatment timing [6]. These findings indicate that the effect on scheduling depends on the clinical setting. Evidence concerning surgical outcomes also varies across study designs and populations [7,8]. Evaluation of CEM should therefore consider both additional disease detection and the complete pathway to definitive treatment.

The staging literature includes prospective and retrospective cohorts, paired comparisons with MRI, and multireader studies. Some reports use lesions as the unit of analysis, whereas others assess breasts, patients, or changes in a proposed operation. Studies also differ in baseline imaging, the threshold for biopsy, and whether imaging findings were verified before treatment.

This systematic review and meta-analysis examines the contribution of CEM to preoperative staging through additional cancer detection, associated changes in management, endpoint-specific diagnostic performance, and the distribution of additional malignant findings. Patient-level changes in breast surgery or localization were the focus of quantitative pooling; other outcomes were interpreted using their source definitions, denominators, and verification methods. MRI evidence provides context for this interpretation.

2. Materials and Methods

2.1. Protocol and Reporting

This systematic review and meta-analysis used PRISMA 2020 reporting guidance [9]. The review is registered in PROSPERO (CRD420261551170). The review focused on adults with breast cancer undergoing preoperative CEM, with additional malignant disease detection and changes in breast surgical management as the principal outcomes. Other staging and treatment endpoints were considered separately according to their source definitions. This distinction guided outcome-specific meta-analysis and descriptive synthesis, because detection of the known index cancer, detection of otherwise occult disease, and a change in treatment represent different clinical questions. Diagnostic concordance and economic outcomes were considered secondary outcomes and were summarized separately.

2.2. Information Sources and Search Strategy

PubMed (including MEDLINE; NCBI PubMed) was selected as the sole bibliographic search platform for its coverage of biomedical literature relevant to the clinical question of preoperative breast imaging. The search covered literature through October 2026 and combined CEM terminology with staging, additional-disease and management concepts, without requiring an MRI comparator. Additional bibliographic databases were not considered in the original review design. The complete Boolean query and active filter settings are not reproduced in this report.

Author-supplied reports were reconciled by DOI. All 13 included reports also matched the supplied PubMed export.

2.3. Eligibility Criteria

Eligible reports evaluated preoperative CEM in adults with breast cancer and reported additional malignancy, multifocal or multicentric disease, management change, or diagnostic-concordance and economic outcomes. Single-arm and comparative studies were considered. Reviews, protocols, case reports and studies without an extractable preoperative cancer subgroup were excluded from the original-study evidence set. Conference abstracts without a full report were not included in the main synthesis and were considered separately for clinical context. Diagnostic-concordance and economic outcomes were evaluated separately from primary additional-disease and management outcomes. Final eligibility depended on the complete report and the stated population, setting and outcome criteria, rather than its preliminary outcome-category label.

Additional malignancy comprised malignant ipsilateral secondary foci or contralateral cancers beyond the index malignancy, according to the source study definition. Breast-level recognition of multifocal or multicentric disease was distinguished from counts of individual additional lesions. Confirmation could comprise pathology or an explicitly defined follow-up reference standard; limitations of verification were retained.

Management outcomes comprised changes in breast surgery or localization and broader treatment changes, including axillary management or neoadjuvant therapy. Study definitions were retained. Potentially overlapping reports were not combined. Participant denominators were summed only within each outcome-specific analysis, rather than across the entire review.

2.4. Study Selection and Data Extraction

R.F. and C.M.K. were responsible for study selection, with disagreements resolved by mutual consensus. Inter-reviewer agreement statistics are not reported. Following full-text review, studies that did not meet the inclusion criteria were excluded. Final eligibility decisions were made by the authors, who confirmed the 13-study evidence set summarized in Table 1.

R.F. and C.M.K. extracted and checked the study data. Extraction covered study design, recruitment setting, patient and lesion counts, CEM indication, comparator imaging, reference standard, additional cancer detection, and management changes. Exact numerators and denominators were retained and checked against the reported endpoint; missing information was not treated as zero. Lesion-level findings, breast-level disease classifications, and patient-level management changes were kept separate. Differences in extraction or endpoint interpretation were resolved by mutual consensus. Thad Benefield (T.B.) provided biostatistical input and validation of the quantitative summaries.

Preliminary title/abstract classifications guided assessment but did not determine final inclusion. A secondary-outcome label or insufficient information in an abstract was not itself a final exclusion decision. Final inclusion was based on the population, setting and outcome criteria in Section 2.3. Figure 1 summarizes the study selection accounting.

2.5. Risk of Bias

The diagnostic and staging components of 12 of the 13 studies summarized in Table 1 were assessed using QUADAS-2 [10]. The concordance and cost analysis by Colarieti [4] was summarized separately and was not assessed with QUADAS-2; no formal economic appraisal was undertaken. Risk of bias was judged as low, high, or unclear in four domains: patient selection, index test, reference standard, and flow and timing. Applicability concerns were considered separately for patient selection, the index test, and the reference standard. Judgments were based on the reported sampling method, exclusions, interpretation procedures, diagnostic thresholds, verification of additional findings, and completeness of patient

and lesion follow-up. Unreported safeguards were rated unclear rather than assumed to be present. No numerical quality score was calculated.

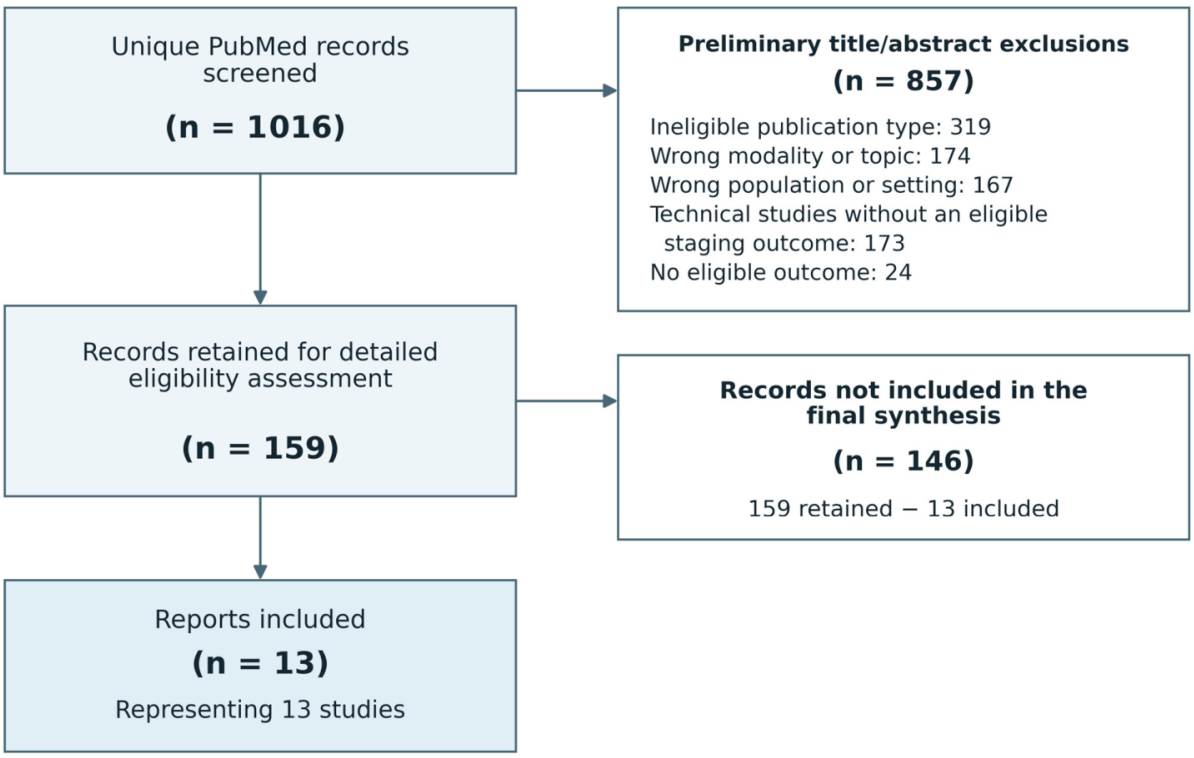

Figure 1. Study selection flow for the 1016-record PubMed set. The exclusion reason categories apply to the 857 preliminary title/abstract exclusions. The aggregate of 146 records not included in the final synthesis is calculated as 159 retained records minus 13 included reports; it is not a reason-specific full-text exclusion tally. Each included report represents one study.

Table 1. Characteristics of the reports summarized in this review.

Study	Design and Sample	Comparator or Population	Interpretive Issue
Ali-Mucheru [11]	Retrospective; 101 patients	Conventional imaging; MRI subset	14 patients received neoadjuvant therapy; mixed cohort
Lee-Felker [12]	Retrospective; 52 women	MRI; 46 referred for MRI-suspicious lesions	MRI-selected population
Kim [13]	Prospective; 84 patients	MRI	Separate index, ipsilateral, contralateral, and management endpoints
Sumkin [14]	Prospective; 99 women	MRI and MBI	Additional-lesion analysis excluded neoadjuvant cases
Cheung [15]	Retrospective; 46 women, 51 cancerous breasts	MRI	Multicentricity is a breast-level endpoint
Bicchierai [16]	Retrospective; 326 patients	DM, DBT, ultrasound	Staging accuracy defined using management and pathology
Girometti [17]	Retrospective multireader; 78 women	DM + DBT; MRI unsuitable	Four readers; 100 cancers including 22 additional lesions
Taylor [18]	Prospective; 59 women	MRI	Additional findings not seen on conventional imaging
Terzoni [19]	Prospective; 70 women	MRI	Discordant-lesion analysis; incomplete verification of MRI-only findings
MacCallum [20]	Retrospective; 202 patients	No MRI comparator; screen-detected cancer staging	Additional findings, patient yield, and surgery changes reported separately
Colarieti [4]	Retrospective; 280 patients	MRI	Diagnostic concordance and modeled costs; secondary outcomes summarized separately
Åhsberg [21]	Prospective feasibility; 47 patients	DM and ultrasound	Primary-treatment endpoint includes neoadjuvant treatment
Montrognon [22]	Retrospective; 132 patients, 134 lesions	CEM plus ultrasound and clinical staging	Multimodal and CEM-only effects are distinct

DM = digital mammography; DBT = digital breast tomosynthesis; MBI = molecular breast imaging.

The appraisal focused on the additional-disease and staging endpoints reported in this review. Differential verification, post-imaging exclusions, and limited follow-up were considered explicitly. Knowledge of a biopsy-proven index cancer was distinguished from knowledge of the reference-standard findings for additional disease. QUADAS-2 does not assess the causal effect of a change in management; those proportions were interpreted in light of the observational designs and whether the reported effect was attributable to CEM alone. Domain judgments are provided in Table 2 and Figure 2. R.F. and C.M.K. were responsible for the final judgments, with disagreements resolved by mutual consensus.

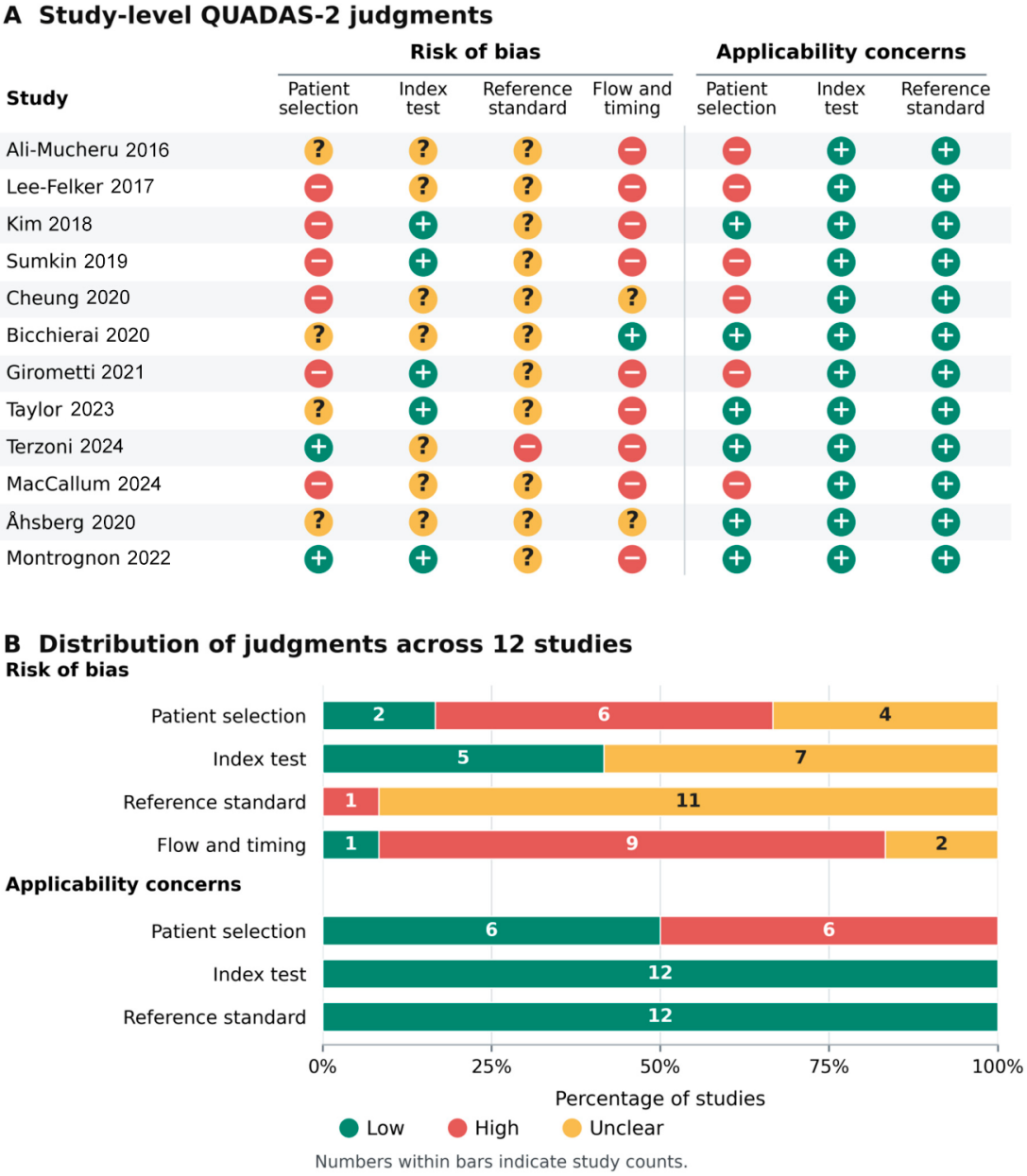

Figure 2. QUADAS-2 risk-of-bias and applicability assessment. (**A**) Study-level judgments for the 12 diagnostic or staging studies. Green (+), red (−), and yellow (?) indicate low, high, and unclear risk or concern. (**B**) Distribution of judgments within each domain; numbers in bars indicate study counts. The concordance and economic analysis was not assessed with QUADAS-2. These diagnostic and staging judgments do not establish the causal benefit of management changes [11–22].

Table 2. QUADAS-2 risk of bias and applicability concerns in 12 diagnostic or staging studies.

Study	Risk of Bias Patient Selection	Risk of Bias Index Test	Risk of Bias Reference Standard	Risk of Bias Flow and Timing	Applicability Patient Selection	Applicability Index Test	Applicability Reference Standard
Ali-Mucheru [11]	Unclear	Unclear	Unclear	High	High	Low	Low
Lee-Felker [12]	High	Unclear	Unclear	High	High	Low	Low
Kim [13]	High	Low	Unclear	High	Low	Low	Low
Sumkin [14]	High	Low	Unclear	High	High	Low	Low
Cheung [15]	High	Unclear	Unclear	Unclear	High	Low	Low
Bicchierai [16]	Unclear	Unclear	Unclear	Low	Low	Low	Low
Girometti [17]	High	Low	Unclear	High	High	Low	Low
Taylor [18]	Unclear	Low	Unclear	High	Low	Low	Low
Terzoni [19]	Low	Unclear	High	High	Low	Low	Low
MacCallum [20]	High	Unclear	Unclear	High	High	Low	Low
Åhsberg [21]	Unclear	Unclear	Unclear	Unclear	Low	Low	Low
Montrognon [22]	Low	Low	Unclear	High	Low	Low	Low

Low, high, and unclear refer to domain judgments, not numerical scores. Applicability is assessed against the preoperative CEM question. Ratings concern diagnostic and staging components; they do not establish the causal benefit of management changes. Bicchierai is assessed for its separately defined correct-staging endpoint [16]. Colarieti's concordance and economic analysis was not assessed with QUADAS-2.

2.6. Outcomes and Statistical Analysis

Additional cancer detection, diagnostic performance, and disease distribution were synthesized descriptively using the source definitions and units of analysis. Detection among confirmed additional malignant lesions was distinguished from additional cancer yield among all staged patients. Breast-level, lesion-level, patient-level, and multireader results were not combined. CEM and MRI results were presented side by side where comparable reference sets were available; no pooled paired effect was estimated because joint classifications and patient-clustering information were insufficient.

Only patient-level management-change outcomes were quantitatively pooled. The primary analysis included changes in breast surgery or localization in Bicchierai [16], MacCallum [20], Kim [13], and Ali-Mucheru [11]. Broader primary-treatment changes in Åhsberg [21] and multimodal breast, axillary, and systemic-treatment changes in Montrognon [22] were reserved for an exploratory expanded analysis. Montrognon's separately reported 24 CEM-only changes were not pooled because their exact denominator was unresolved.

Proportions were pooled on the logit scale using inverse-variance random-effects models, with restricted maximum likelihood estimation of between-study variance and Hartung–Knapp 95% confidence intervals [23,24]. Heterogeneity was summarized using Cochran Q, I^2, and τ^2. Post hoc analyses excluded the mixed neoadjuvant-treatment cohort, omitted each primary cohort in turn, and added the two broader treatment endpoints. Individual patient-level proportions used Wilson 95% confidence intervals. Mathematical details and analysis implementation are provided in the Supplementary Methods.

No formal reporting-bias test was performed because each synthesis included fewer than 10 studies. Confidence in the evidence was appraised qualitatively for additional-disease detection, management-change frequency and patient-important outcomes, considering risk of bias, consistency, precision, applicability and review-process limitations. Diagnostic findings were interpreted according to their source units of analysis (lesion, breast or patient), and the descriptive management-change proportion was not treated as a measure of clinical benefit. No formal GRADE ratings were assigned. Qualitative judgments are discussed by outcome in Section 4.3.

2.7. Ethics

This review used aggregate data from published reports and did not involve new participant recruitment or access to identifiable patient data.

3. Results

3.1. Study Selection and Characteristics

Of 1016 unique PubMed records, 857 were excluded at preliminary title/abstract screening and 159 were retained for detailed eligibility assessment. A total of 13 reports representing 13 studies were included in the final synthesis (Table 1). The remaining 146 records were not included in the final synthesis, representing the difference between the retained and included totals (159 − 13). Figure 1 shows the selection flow and the reason-specific categories for the 857 initial exclusions.

The evidence comprised 12 diagnostic or staging reports and 1 concordance and economic analysis (Table 1). Populations included MRI-selected referrals, patients unable to undergo MRI, screen-detected cancers, and a mixed-treatment cohort [11,12,17,20]. Colarieti evaluated concordance and modeled costs in 280 patients; these outcomes remained separate from additional cancer detection and management change [4].

3.2. Risk of Bias

Among 12 QUADAS-2 assessments, patient selection risk was high in 6, unclear in 4, and low in 2; flow-and-timing risk was high in 9, unclear in 2, and low in 1 (Figure 2; Table 2). Index-test risk was low in 5 and unclear in 7; reference-standard risk was unclear in 11 and high in 1. Patient applicability concerns were high in 6 and low in 6; index-test and reference-standard applicability concerns were low in all 12.

3.3. Additional Cancer Detection

Four studies reported CEM versus MRI detection among reference-standard additional malignant lesions: Lee-Felker, 11/11 (100.0%) versus 10/11 (90.9%); Kim, 31/37 (83.8%) for each modality; Sumkin, 14/23 (60.9%) versus 13/23 (56.5%); and Taylor, 1/6 (16.7%) versus 6/6 (100.0%) (Table 3; Figure 3) [12–14,18]. These lesion-level fractions use the distinct populations and endpoint definitions detailed in Table 3.

Table 3. Study-level CEM–MRI comparisons for additional-disease endpoints.

Study	Reference Set	CEM Detected	MRI Detected	Endpoint and Qualification
Lee-Felker [12]	11	11/11 (100.0%)	10/11 (90.9%)	Secondary malignant lesions in an MRI-selected referral series
Kim [13]	37	31/37 (83.8%)	31/37 (83.8%)	Additional ipsilateral and contralateral malignant lesions
Sumkin [14]	23	14/23 (60.9%)	13/23 (56.5%)	Pathology-verified nonindex malignant lesions; no neoadjuvant chemotherapy
Taylor [18]	6	1/6 (16.7%)	6/6 (100.0%)	Additional malignant lesions not detected on conventional imaging
Terzoni [19]	4	0/4 (0.0%)	4/4 (100.0%)	Selected MRI-only discordant cancers; not all additional malignancies
Cheung [15]	7	7/7 (100.0%)	7/7 (100.0%)	Breasts with pathology-confirmed multicentric disease; breast-level endpoint

Counts follow the extraction table; percentages are recalculated to one decimal place. Terzoni concerns selected MRI-only discordant cancers and cannot estimate overall relative detection. Cheung concerns breasts with multicentric disease rather than individual additional lesions. These differently defined endpoints are not combined.

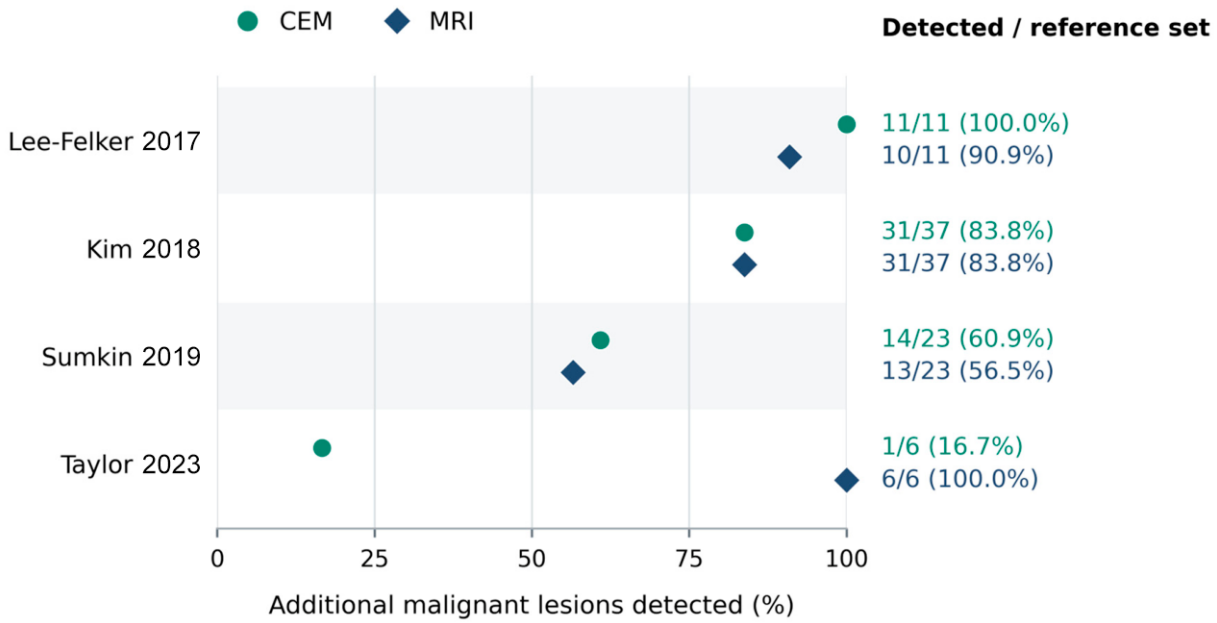

Figure 3. Study-level CEM and MRI detection of additional malignant lesions in four cohorts. Labels give detected cancers divided by the reference-standard additional malignant-lesion set. These are detection fractions, not patient-level additional cancer yield. Terzoni's MRI-only discordant subset and Cheung's breast-level endpoint are reported separately in Table 3. No pooled comparison is shown [12–14,18].

Girometti's four readers detected 19/22, 17/22, 19/22, and 19/22 additional cancers (77.3–86.4%) in the same cohort [17]. Reader observations remained separate.

Terzoni reported 102 CEM-visible areas among 114 MRI-visible enhancement areas and 4 malignant MRI-only lesions in three patients [19]. The selected discordant-cancer subset (CEM, 0/4; MRI, 4/4) is shown separately in Table 3 and does not represent all additional malignancies.

3.4. Changes in Surgical and Treatment Management

The four primary cohorts reported breast-surgery or localization changes in 149/713 patients: Bicchierai, 60/326; MacCallum, 43/202; Kim, 26/84; and Ali-Mucheru, 20/101 (Figure 4) [11,13,16,20]. The pooled proportion was 21.8% (95% CI, 14.9–30.7%), with $I^2 = 52.5\%$, $\tau^2 = 0.0415$, and $Q = 6.32$ (3 df; $p = 0.097$). Precision was limited by the four-study evidence base. Excluding Ali-Mucheru's cohort, which included 14 neoadjuvant-treatment patients, yielded 22.5% (95% CI, 11.0–40.5%) across 612 patients ($I^2 = 67.7\%$). Leave-one-out estimates ranged from 19.6% to 23.5%, with wide CIs.

Kim attributed 9/26 management changes to false-positive findings [13]. MacCallum reported additional malignant findings in 27/43 patients with changes and false-positive or nonmalignant findings in 16/43 [20]. Some nonmalignant findings were atypical proliferative lesions requiring excision, so this category does not automatically indicate inappropriate surgery.

Åhsberg reported broader primary-treatment changes, including neoadjuvant treatment, in 10/47 patients [21]. Montrognon reported multimodal breast, axillary, or systemic-treatment changes in 33/132 patients (25.0%), corresponding to 34/134 lesions (25.4%) [22]. Adding these cohorts yielded an exploratory estimate of 22.1% (95% CI, 18.0–27.0%) across 192/892 patients in six cohorts ($I^2 = 31.6\%$). This broader estimate cannot be attributed solely to CEM-induced breast-surgery changes. Montrognon's 24 CEM-only changes (18.5%, as reported) were not pooled because the denominator remained unresolved.

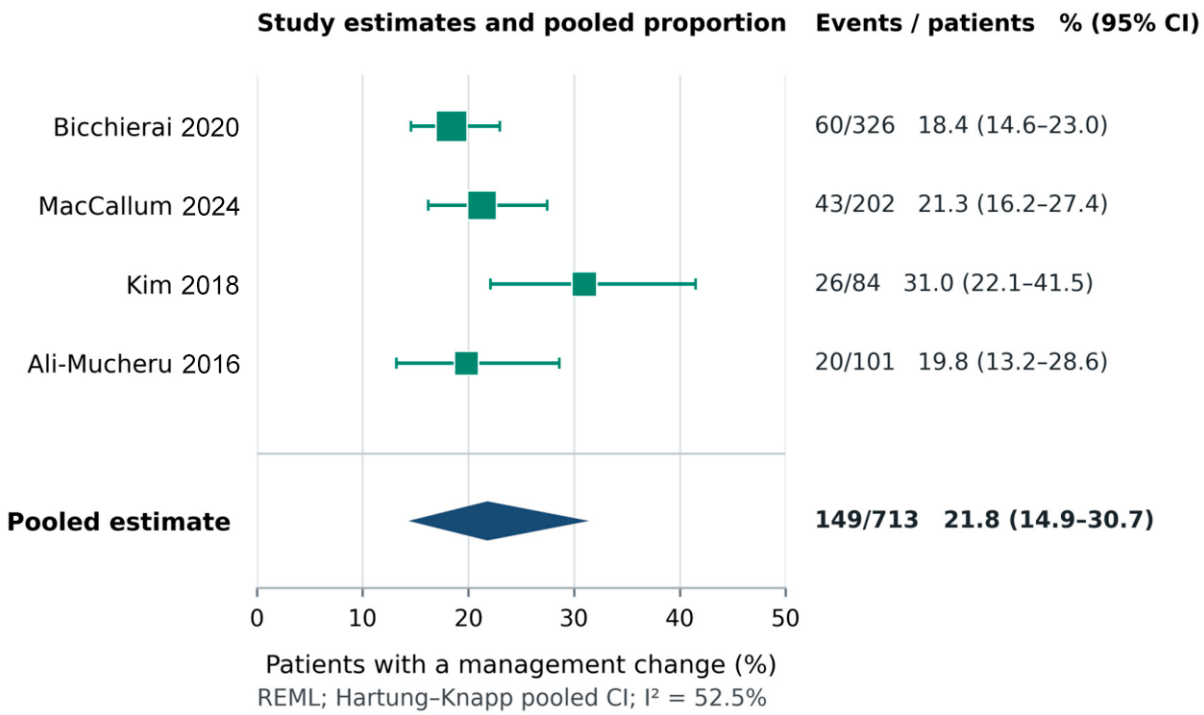

Figure 4. Meta-analysis of patient-level changes in breast surgery or localization. Squares represent study proportions, with areas proportional to random-effects weights; horizontal lines show Wilson 95% CIs. The diamond represents the pooled logit-scale REML estimate with its Hartung–Knapp 95% CI: 21.8% (95% CI, 14.9–30.7%) (I^2 = 52.5%). Åhsberg's primary-treatment and Montrognon's multimodal staging endpoints contribute only to the expanded analysis. The pooled result measures changes in management and does not establish clinical benefit [11,13,16,20].

3.5. Diagnostic Performance by Endpoint

Index-cancer detection was 101/103 (98.1%) in Ali-Mucheru after excluding two previously removed lesions [11], 67/68 (98.5%) in Taylor [18], and 78/84 (92.9%) in Kim [13]. Taylor detected 1/6 additional cancers (16.7%); Kim detected 26/31 ipsilateral secondary cancers (83.9%) and 5/6 contralateral cancers (83.3%). Cheung visualized 51/51 principal cancers across 51 cancerous breasts in 46 women [15].

CEM positive predictive value (PPV) for additional findings was 14/27 (51.9%) in Sumkin, 1/13 (7.7%) in Taylor, and 34/74 (45.9%) in MacCallum [14,18,20]. Kim's ipsilateral and contralateral PPVs were 26/36 (72.2%) and 5/13 (38.5%), respectively [13] (Figure 5). Breast-level multicentric and management-defined PPVs remained separate.

Bicchierai's correct-staging and management endpoint yielded sensitivity of 55/59 (93.2%), specificity of 261/267 (97.8%), PPV of 55/61 (90.2%), negative predictive value (NPV) of 261/265 (98.5%), and accuracy of 316/326 (96.9%) [16]. Positive results required a change in the planned operation verified against definitive histology.

Girometti's overall cancer detection ranged from 93% to 96% across four readers, including 95–96% among less experienced readers [17]. Additional-lesion detection ranged from 77.3% to 86.4%.

3.6. Multifocal and Multicentric Disease

Cheung reported multifocal disease in 9 breasts on CEM, 7 on MRI, and 7 at pathology; corresponding multicentric counts were 11, 11, and 7 [15]. Both modalities detected all 7 pathology-confirmed multicentric cases (7/7); CEM PPV was 7/11 breasts (63.6%).

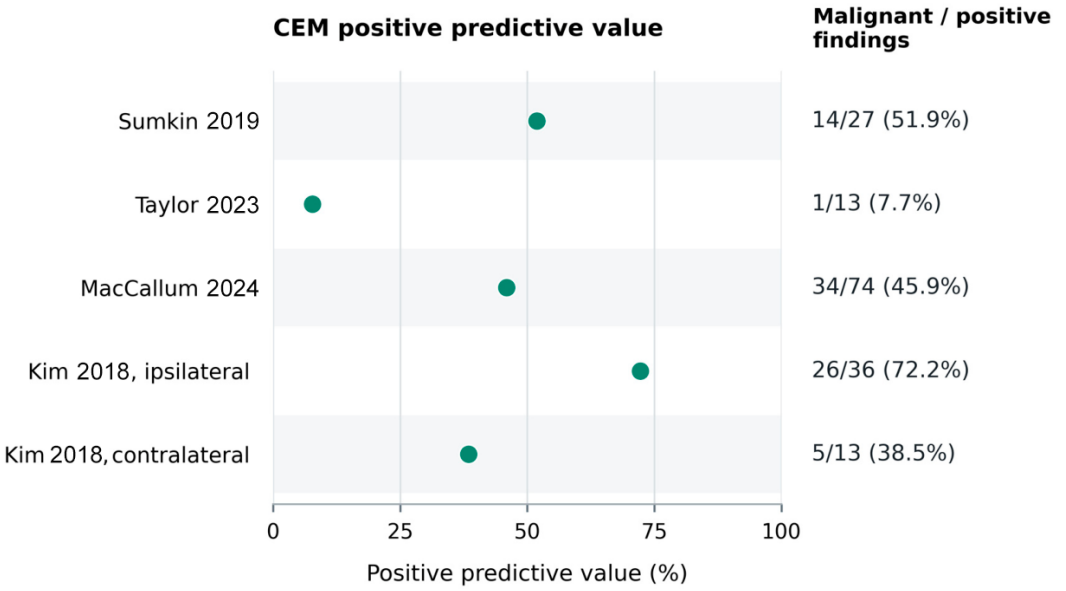

Figure 5. CEM positive predictive value for additional findings, with ipsilateral and contralateral Kim results separated. Labels show malignant findings divided by positive findings for the source endpoint. Differences in selection, indication, and reference standards limit cross-study comparisons. Multicentric breast-level PPV and correct-management PPV are not combined with these results [13,14,18,20].

3.7. Laterality and Patient Yield

Kim recorded 26 ipsilateral and five contralateral additional cancers with each modality; MacCallum recorded 28 ipsilateral and 6 contralateral CEM-detected malignant lesions [13,20]. Ali-Mucheru's management-change subset included 14 same-quadrant, 3 separate-quadrant, and 3 contralateral findings [11]. Bicchierai's nine contralateral observations represented surgical procedures, not verified malignant-lesion counts [16].

MacCallum's patient-level additional cancer yield was 29/202 (14.4%; 95% CI, 10.2–19.9%), comprising 34 malignant lesions [20]. This patient yield is distinct from lesion-level detection and PPV.

4. Discussion

Across 13 studies, CEM was associated with additional cancer detection and changes in treatment planning. Four observational cohorts yielded a pooled frequency of breast-surgery or localization changes of 21.8% (95% CI, 14.9–30.7%). This estimate describes how often plans changed in selected cohorts, without establishing that those changes improved care. Although exclusion of the mixed neoadjuvant-treatment cohort produced a similar point estimate, its wide interval and the shared risk-of-bias concerns limit reassurance from numerical stability. This analysis complements reviews of diagnostic accuracy and tumor extent [25–27]; it does not demonstrate improved surgical or oncologic outcomes.

4.1. Additional Cancer Detection and Comparison with MRI

Performance for the known index cancer should not be extrapolated to additional disease. Taylor reported CEM detection of 67/68 known malignant lesions but only 1/6 additional malignant lesions [18]. Lee-Felker reported detection of all 11 secondary cancers, but 46/52 women had been referred because MRI showed additional suspicious lesions [12]. These small, selected cohorts address different clinical populations. Girometti's multireader findings in women unable to undergo MRI also have a distinct interpretation [17].

The paired-modality studies showed variable additional-cancer detection [12–14,18]. Equal marginal totals do not establish that CEM and MRI detected the same lesions. Joint

classifications and information on clustering within patients were insufficient for a pooled paired comparison. Verification further limits interpretation: Terzoni classified eight MRI-only findings without an ultrasound correlate as false positive without biopsy [19], whereas other cohorts used pathology or imaging follow-up [14,18]. A negative second-look ultrasound does not establish benignity.

Published MRI evidence provides context rather than a comparator arm. Patient-level additional disease yield [28] cannot be compared directly with detection among a selected set of confirmed additional cancers. MRI syntheses also differ in their findings for mastectomy, re-excision and reoperation [7,8], emphasizing the importance of study design and treatment pathways. Neither these indirect comparisons nor the present descriptive CEM–MRI results establish diagnostic equivalence.

4.2. Management Changes and Clinical Implications

Management changes require interpretation alongside the initial plan, the reason for change and the final pathology. In Kim, 9/26 changes were attributed to false-positive findings; in MacCallum, 16/43 were associated with nonmalignant or atypical findings [13,20]. Some atypical lesions may warrant excision, so neither every change nor every nonmalignant finding is automatically beneficial or unnecessary. The pooled proportion describes observed changes and does not estimate a causal benefit of CEM. Broader breast, axillary and systemic-treatment endpoints retain their separate exploratory interpretation.

Preliminary randomized evidence is relevant but remains separate from the full-report synthesis. The 2025 PROCEM conference abstract reported treatment changes in 17/202 CEM participants (8.4%) and no statistically significant differences in mastectomy or re-operation [29]. A nonsignificant result does not establish equivalence, and complete trial reporting is needed to assess these findings alongside the observational cohorts.

CEM may be useful when MRI is unsuitable or unavailable, provided that additional findings can be evaluated through an appropriate biopsy pathway [1,3]. Imaging findings that would substantially alter treatment should be verified when feasible. Local assessment should include breast coverage, contrast suitability, radiation exposure, biopsy access and time to definitive treatment. The cost analysis in diagnostically concordant cases [4] does not establish equivalent additional-cancer detection or surgical outcomes.

Patient preferences should inform choice among clinically appropriate examinations. Preference studies favor CEM on average when comparable diagnostic performance is assumed, while screening and surveillance studies show that acceptance, discomfort and anxiety vary by setting and modality [30–33]. These findings do not establish diagnostic equivalence, improved adherence or oncologic benefit. Compression tolerance, iodinated contrast risk and the availability of follow-up procedures remain relevant to individual decisions [1,3,34].

4.3. Certainty of Evidence and Limitations

Overall confidence in the evidence is limited and differs by question. For additional cancer detection and comparison with MRI, confidence is reduced by small and selected cohorts, inconsistent definitions and units of analysis, incomplete verification and unavailable paired classifications. Across all 12 diagnostic/staging studies assessed with QUADAS-2, patient selection risk was high in 6/12, and flow-and-timing risk was high in 9/12; reference-standard risk was unclear in 11/12. These figures describe all 12 diagnostic/staging studies; the four-cohort management-change subset has a distinct risk profile. The evidence does not support a reliable conclusion about comparative detection or interchangeability.

For management-change frequency, only four observational cohorts contributed. Within these four cohorts, patient selection risk was high in two and unclear in two; flow-

and-timing risk was high in three and low in one, and reference-standard risk was unclear in all four (Table 2). QUADAS-2 addresses their diagnostic/staging components, not the causal effects of management changes. Differences in baseline plans, recruitment, verification, false-positive-driven changes and neoadjuvant-treatment inclusion limit applicability, while moderate heterogeneity and wide intervals limit precision. Sensitivity analyses assess numerical stability but cannot remove these biases. The pooled 21.8% should therefore be interpreted as an observed frequency in the included settings.

For patient-important benefits, including fewer reoperations, appropriate surgical extent and oncologic outcomes, this synthesis does not establish an effect. Such benefits require comparative outcomes and cannot be inferred from the frequency of changes in management. Evidence about patient experience and modeled costs is indirect to these outcomes. These qualitative judgments are not formal GRADE ratings; the absence of formal outcome-level GRADE ratings remains a limitation.

Review-level limitations further reduce confidence. Restricting the search to PubMed/MEDLINE may have missed eligible reports indexed elsewhere; additional databases were not considered. The complete Boolean query and active filter settings are not reproduced, limiting exact replication of the search. Figure 1 presents the aggregate selection flow and initial-screening categories; separate retrieval-stage counts and reason-specific full-text exclusions are not presented. No formal reporting-bias test was undertaken because each synthesis contained fewer than 10 studies. The economic analysis was not formally appraised.

4.4. Future Research

Prospective comparative studies should use prespecified definitions and report additional cancers per staged patient separately from lesion-level detection. Paired lesion classifications, patient-level clustering and verification against biopsy and surgical pathology are needed. Studies should document the initial plan, imaging-prompted proposal, verification before surgery, final operation, benign biopsies, margins, re-excision, reoperation and time to treatment. Patient-reported experience should complement these outcomes. This would clarify when CEM changes care appropriately and whether those changes improve outcomes.

5. Conclusions

Preoperative CEM was associated with additional cancer detection and changes in treatment planning. Across four observational cohorts, the pooled frequency of changes in breast surgery or localization was 21.8% (95% CI, 14.9–30.7%). This estimate describes observed management changes in selected populations; it does not establish that CEM improves surgical or oncologic outcomes. Limited search coverage, selection and verification concerns, heterogeneous endpoints and imprecision limit confidence, and the available comparisons do not establish diagnostic equivalence to MRI. Prospective comparative studies with consistent outcome definitions and verification against biopsy and surgical pathology are needed to determine whether imaging-prompted changes improve patient outcomes.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/diagnostics16233997/s1, Table S1: PRISMA 2020 reporting checklist; Supplementary Methods: Detailed statistical methods; Statistical References [23,24].

Author Contributions: R.F. and C.M.K. contributed equally to conceptualization, methodology, investigation, data curation, visualization, writing of the original draft, manuscript review and editing, supervision, and project administration. T.B. contributed biostatistical methodology, formal analysis, and validation of the quantitative summaries. The authors share responsibility for the work and approval of the final manuscript. C.M.K. completed the final review and approval of the extracted data, risk-of-bias judgments, and patient-level confidence intervals for the 13-study synthesis. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This systematic review used published aggregate data and did not involve new participant recruitment or access to identifiable patient information.

Informed Consent Statement: Not applicable.

Data Availability Statement: Search methods are described in Section 2.2 and study selection accounting in Section 3.1 and Figure 1. Study-level findings are reported in Tables 1–3, Results Sections 3.3–3.7 and Figures 3–5. The PRISMA 2020 checklist is provided in Supplementary Table S1 and detailed statistical methods in the Supplementary Methods.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (GPT-6.2, OpenAI, San Francisco, CA, USA) to assist with language editing and improve readability. The authors subsequently reviewed and edited the content and take full responsibility for the published work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jochelson, M.S.; Lobbes, M.B.I. Contrast-enhanced mammography: State of the art. *Radiology* **2021**, *299*, 36–48. [CrossRef] [PubMed]
2. Cozzi, A.; Magni, V.; Zanardo, M.; Schiaffino, S.; Sardanelli, F. Contrast-enhanced mammography: A systematic review and meta-analysis of diagnostic performance. *Radiology* **2022**, *302*, 568–581. [CrossRef] [PubMed]
3. Bellini, C.; Bicchierai, G.; de Benedetto, D.; di Naro, F.; Maiello, C.; Vidali, S.; Tonelli, P.; Vanzi, E.; Boeri, C.; Miele, V.; et al. How to optimise breast cancer staging with contrast-enhanced mammography: Current evidence and clinical implications. *Insights Imaging* **2026**, *17*, 178. [CrossRef] [PubMed]
4. Colarieti, A.; Bonetti, A.; Gianfrate, F.; Attanasio, S.; Rampi, A.M.; Gambaro, A.C.L.; Carriero, A. Contrast-enhanced mammography versus breast MRI: A multidimensional cost-effectiveness analysis. *La Radiol. Medica* **2026**, *131*, 21–30. [CrossRef] [PubMed]
5. Gauthier, I.D.; Seely, J.M.; Cordeiro, E.; Peddle, S. The impact of preoperative breast MRI on timing of surgical management in newly diagnosed breast cancer. *Can. Assoc. Radiol. J.* **2024**, *75*, 377–385. [CrossRef] [PubMed]
6. Pogatchnik, B.; Kuehn-Hajder, J.; Nelson, M.; Emory, T. No effect of pretreatment breast MRI on the timing of surgical treatment of newly diagnosed breast cancer. *J. Am. Coll. Radiol.* **2017**, *14*, 1310–1315. [CrossRef] [PubMed]
7. Houssami, N.; Turner, R.M.; Morrow, M. Meta-analysis of pre-operative magnetic resonance imaging (MRI) and surgical treatment for breast cancer. *Breast Cancer Res. Treat.* **2017**, *165*, 273–283. [CrossRef] [PubMed]
8. Eisen, A.; Fletcher, G.G.; Fienberg, S.; George, R.; Holloway, C.; Kulkarni, S.; Seely, J.M.; Muradali, D. Breast magnetic resonance imaging for preoperative evaluation of breast cancer: A systematic review and meta-analysis. *Can. Assoc. Radiol. J.* **2024**, *75*, 118–135. [CrossRef] [PubMed]
9. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
10. Whiting, P.F.; Rutjes, A.W.S.; Westwood, M.E.; Mallett, S.; Deeks, J.J.; Reitsma, J.B.; Leeflang, M.M.G.; Sterne, J.A.C.; Bossuyt, P.M.M.; QUADAS-2 Group. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann. Intern. Med.* **2011**, *155*, 529–536. [CrossRef] [PubMed]
11. Ali-Mucheru, M.; Pockaj, B.; Patel, B.; Pizzitola, V.; Wasif, N.; Stucky, C.-C.; Gray, R. Contrast-enhanced digital mammography in the surgical management of breast cancer. *Ann. Surg. Oncol.* **2016**, *23*, 649–655. [CrossRef] [PubMed]
12. Lee-Felker, S.A.; Tekchandani, L.; Thomas, M.; Gupta, E.; Andrews-Tang, D.; Roth, A.; Sayre, J.; Rahbar, G. Newly diagnosed breast cancer: Comparison of contrast-enhanced spectral mammography and breast MR imaging in the evaluation of extent of disease. *Radiology* **2017**, *285*, 389–400. [CrossRef] [PubMed]
13. Kim, E.Y.; Youn, I.; Lee, K.H.; Yun, J.-S.; Park, Y.L.; Park, C.H.; Moon, J.; Choi, S.H.; Choi, Y.J.; Ham, S.-Y.; et al. Diagnostic value of contrast-enhanced digital mammography versus contrast-enhanced magnetic resonance imaging for the preoperative evaluation of breast cancer. *J. Breast Cancer* **2018**, *21*, 453–462. [CrossRef] [PubMed]
14. Sumkin, J.H.; Berg, W.A.; Carter, G.J.; Bandos, A.I.; Chough, D.M.; Ganott, M.A.; Hakim, C.M.; Kelly, A.E.; Zuley, M.L.; Houshmand, G.; et al. Diagnostic performance of MRI, molecular breast imaging, and contrast-enhanced mammography in women with newly diagnosed breast cancer. *Radiology* **2019**, *293*, 531–540. [CrossRef] [PubMed]
15. Cheung, Y.C.; Juan, Y.H.; Lo, Y.F.; Lin, Y.-C.; Yeh, C.-H.; Ueng, S.-H. Preoperative assessment of contrast-enhanced spectral mammography of diagnosed breast cancers after sonographic biopsy: Correlation to contrast-enhanced magnetic resonance imaging and 5-year postoperative follow-up. *Medicine* **2020**, *99*, e19024. [CrossRef] [PubMed]

16. Bicchierai, G.; Tonelli, P.; Piacenti, A.; De Benedetto, D.; Boeri, C.; Vanzi, E.; Bianchi, S.; Cirone, D.; Gill, M.K.; Miele, V.; et al. Evaluation of contrast-enhanced digital mammography in the preoperative staging of breast cancer: Large-scale single-center experience. *Breast J.* **2020**, *26*, 1276–1283. [CrossRef] [PubMed]
17. Girometti, R.; Linda, A.; Conte, P.; Lorenzon, M.; De Serio, I.; Jerman, K.; Londero, V.; Zuiani, C. Multireader comparison of contrast-enhanced mammography versus the combination of digital mammography and digital breast tomosynthesis in the preoperative assessment of breast cancer. *La Radiol. Medica* **2021**, *126*, 1407–1414. [CrossRef] [PubMed]
18. Taylor, D.B.; Burrows, S.; Saunders, C.M.; Parizel, P.M.; Ives, A. Contrast-enhanced mammography versus MRI for breast cancer staging: Detection of additional malignant lesions not seen on conventional imaging. *Eur. Radiol. Exp.* **2023**, *7*, 8. [CrossRef] [PubMed]
19. Terzoni, A.; Basile, P.; Gambaro, A.C.; Attanasio, S.; Rampi, A.M.; Brambilla, M.; Carriero, A. Locoregional staging of breast cancer: Contrast-enhanced mammography versus breast magnetic resonance imaging. *La Radiol. Medica* **2024**, *129*, 558–565. [CrossRef] [PubMed]
20. MacCallum, C.; Elder, K.; Nickson, C.; Ruecker, K.; Park, A.; Mann, G.B.; Rose, A.K. Contrast-enhanced mammography in local staging of screen-detected breast cancer. *Ann. Surg. Oncol.* **2024**, *31*, 6820–6830. [CrossRef] [PubMed]
21. Åhsberg, K.; Gardfjell, A.; Nimeus, E.; Rasmussen, R.; Behmer, C.; Zackrisson, S.; Ryden, L. Added value of contrast-enhanced mammography in staging of malignant breast lesions: A feasibility study. *World J. Surg. Oncol.* **2020**, *18*, 100. [CrossRef] [PubMed]
22. Montrognon, F.; Clatot, F.; Berghian, A.; Douvrin, F.; Quieffin, F.; Defta, D.; Buquet, A.; Ferret, M.; Lequesne, J.; Leheurteur, M.; et al. Impact of preoperative staging with contrast-enhanced mammography for localized breast cancer management. *Br. J. Radiol.* **2022**, *95*, 20210765. [CrossRef] [PubMed]
23. IntHout, J.; Ioannidis, J.P.A.; Borm, G.F. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med. Res. Methodol.* **2014**, *14*, 25. [CrossRef] [PubMed]
24. Veroniki, A.A.; Jackson, D.; Viechtbauer, W.; Bender, R.; Bowden, J.; Knapp, G.; Kuss, O.; Higgins, J.P.; Langan, D.; Salanti, G. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res. Synth. Methods* **2016**, *7*, 55–79. [CrossRef] [PubMed]
25. Pötsch, N.; Vatteroni, G.; Clauser, P.; Helbich, T.H.; Baltzer, P.A.T. Contrast-enhanced mammography versus contrast-enhanced breast MRI: A systematic review and meta-analysis. *Radiology* **2022**, *305*, 94–103. [CrossRef] [PubMed]
26. Shahraki, Z.; Ghaffari, M.; Nakhaie Moghadam, M.; Parooie, F.; Salarzaei, M. Preoperative evaluation of breast cancer: Contrast-enhanced mammography versus contrast-enhanced magnetic resonance imaging: A systematic review and meta-analysis. *Breast Dis.* **2022**, *41*, 303–315. [CrossRef] [PubMed]
27. Ermans, S.J.E.; Maas, R.; Hill, T.M.; Usiskin, S.I.; Alic, L.; Douek, M. Comparison of contrast-enhanced mammography and breast MRI in tumour extent assessment: A systematic review and meta-analysis. *Eur. J. Surg. Oncol.* **2026**, *52*, 111410. [CrossRef] [PubMed]
28. Plana, M.N.; Carreira, C.; Muriel, A.; Chiva, M.; Abraira, V.; Emparanza, J.I.; Bonfill, X.; Zamora, J. Magnetic resonance imaging in the preoperative assessment of patients with primary breast cancer: Systematic review of diagnostic accuracy and meta-analysis. *Eur. Radiol.* **2012**, *22*, 26–38. [CrossRef] [PubMed]
29. Åhsberg, K.; Gardfjell, A.; Niméus, E.; Rydén, L.; Zackrisson, S. PROCEM; a prospective randomized trial on the added value of preoperative contrast-enhanced mammography in staging of malignant breast lesions [conference abstract]. *Br. J. Surg.* **2025**, *112*, znaf149.025. [CrossRef]
30. Ferre, R.; Benefield, T.; Kuzmiak, C.M. Patient Preference for Contrast-Enhanced Mammography Compared to Breast MRI. *Acad. Radiol.* **2026**, *33*, 4976–4983. [CrossRef] [PubMed]
31. Hobbs, M.M.; Taylor, D.B.; Buzynski, S.; Peake, R.E. Contrast-enhanced spectral mammography (CESM) and contrast enhanced MRI (CEMRI): Patient preferences and tolerance. *J. Med. Imaging Radiat. Oncol.* **2015**, *59*, 300–305. [CrossRef] [PubMed]
32. Vatteroni, G.; Levi, R.; Filippini, M.; Costa, M.; Trimboli, R.M.; Bernardi, D. Patient acceptance and experience with contrast-enhanced mammography in women with a personal history of breast cancer: Findings from the COMBO trial. *Insights Imaging* **2026**, *17*, 262. [CrossRef] [PubMed]
33. Allajbeu, I.; Charnley, K.R.; Yang, Y.; Field-Rayner, J.; Morris, K.; Payne, N.R.; Kilburn-Toppin, F.; Manavaki, R.; Gilbert, F.J. Acceptance, experience, and feedback for supplemental screening in dense breasts among women participating in the BRAID trial. *Insights Imaging* **2026**, *17*, 14. [CrossRef] [PubMed]
34. Gilbert, F.J.; Payne, N.R.; Allajbeu, I.; Yit, L.; Vinnicombe, S.; Lyburn, I.; Sharma, N.; Teh, W.; James, J.; Seth, A.; et al. Comparison of supplemental breast cancer imaging techniques—Interim results from the BRAID randomised controlled trial. *Lancet* **2025**, *405*, 1935–1944. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.