Network Meta-analysis for the Diagnostic Approach to Pathologic Nipple Discharge

Mando D. Filipe,[1] Susanna I.S. Patuleia,[2] Valentijn M.T. de Jong,[3,4] Menno R. Vriens,[1] Paul J. van Diest,[2] Arjen J. Witkamp[1]

Abstract

Pathologic nipple discharge (PND) is one of the most common breast-related complaints for referral because of its supposed association with breast cancer. The aim of this network meta-analysis (NMA) was to compare the diagnostic efficacy of ultrasound, mammogram, cytology, magnetic resonance imaging (MRI), and ductoscopy in patients with PND, as well as to determine the best diagnostic strategy to assess the risk of malignancy as cause for PND. Cochrane Library, PubMed, and Embase were searched to collect relevant literature from the inception of each of the diagnostic methods until January 27, 2020. The search yielded 1472 original citations, of which 36 studies with 3764 patients were finally included for analysis. Direct and indirect comparisons were performed using an NMA approach to evaluate the combined odd ratios and to determine the surface under the cumulative ranking curves (SUCRA) of the diagnostic value of different imaging methods for the detection of breast cancer in patients with PND. Additionally, a subgroup meta-analysis comparing ductoscopy to MRI when conventional imaging was negative was also performed. According to this NMA, sensitivity for detection of malignancy in patients with PND was highest for MRI (83%), followed by ductoscopy (58%), ultrasound (50%), cytology (38%), and mammogram (22%). Specificity was highest for mammogram (93%) followed by ductoscopy (92%), cytology (90%), MRI (76%), and ultrasound (69%). Diagnostic accuracy was the highest for ductoscopy (88%), followed by cytology (82%), MRI (77%), mammogram (76%), and ultrasound (65%). Subgroup meta-analysis (comparing ductoscopy to MRI when ultrasound and mammogram were negative) showed no significant difference in sensitivity, but ductoscopy was statistically significantly better with regard to specificity and diagnostic accuracy. The results from this NMA indicate that although ultrasound and mammogram may remain low-cost useful first choices for the detection of malignancy in patients with PND, ductoscopy outperforms most imaging techniques (especially MRI) and cytology.

Clinical Breast Cancer, Vol. 20, No. 6, e723-48 © 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Breast cancer, Cytology, Diagnosis, Ductoscopy, Mammography, MRI, Ultrasound

Introduction

Pathologic nipple discharge (PND) is defined as unilateral, spontaneous, and bloody or serous discharge, usually arising from a single duct orifice of the nipple. After pain and palpable lumps, PND is the third most common breast-related complaint.[1] PND is often associated with breast cancer and accounts for 3% to 5% of surgical breast clinic referrals.[2-5] However, the most common causes of PND—duct ectasias and intraductal papillomas—are benign.[6,7]

Mammogram and breast ultrasound are important tools for the detection of breast cancer. However, in the case of PND as the only complaint, they both have limited sensitivity.[8] Magnetic resonance imaging (MRI) has shown to be a sensitive tool for the detection of malignancy, but specificity is low. Detection of small lesions and differentiating benign from malignant masses remains difficult with MRI.[9,10] Therefore, the value of MRI is limited in patients with PND, and core needle biopsy or surgical excision is still necessary when MRI reveals a suspicious lesion.[11,12] Cytology of the nipple discharge is also used to determine the risk of malignancy in patients with PND, but its clinical relevance has been contested.[5,13,14]

[1]Department of Surgical Oncology, Cancer Centre, University Medical Centre Utrecht, Utrecht, the Netherlands
[2]Departments of Pathology and Oncology, University Medical Centre Utrecht, Utrecht, the Netherlands
[3]Cochrane Netherlands, University Medical Centre Utrecht, Utrecht, the Netherlands
[4]Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, the Netherlands

Submitted: Apr 20, 2020; Revised: May 27, 2020; Accepted: May 30, 2020; Epub: Jun 8, 2020

Address for correspondence: Mando D. Filipe, MD, Department of Surgical Oncology, Cancer Centre, University Medical Centre Utrecht, PO Box 85500, 3508 GA, Utrecht, the Netherlands
E-mail contact: m.d.filipe-2@umcutrecht.nl

https://doi.org/10.1016/j.clbc.2020.05.015

Ductoscopy is a minimally invasive microendoscopic technique providing real-time visualization of the milk ducts of the breast. Ductoscopy is performed with the patient under local anesthesia at the outpatient clinic; it is currently used as a diagnostic tool in assessing women with PND.[15-22] Previous studies and a meta-analysis show that ductoscopy is a useful tool in finding intraductal lesions causing PND (benign and malignant) before or during duct excision.[23-25]

Because PND is regarded as a possible sign of breast cancer and standard radiologic imaging often fails to reveal the cause, most women with persistent PND undergo surgical procedures, such as microdochectomy or major duct excision, to exclude malignancy.[6,8,9] However, only 5% to 8% of these patients with PND turn out to actually have malignancy,[5,26,27] meaning that 90% to 95% of these surgical procedures are performed to assess something with a nonmalignant cause. Therefore, it is important to assess the different diagnostic tools currently available and then determine the usefulness of each tool in the different phases of the diagnostic process.

To this end, we carried out a systematic review of the literature and performed a network meta-analysis (NMA) to compare the value of different diagnostic tools to detect malignancy in patients with PND. Additionally, we determined the optimal diagnostic strategy for patients with PND.

Patients and Methods

This systematic review and NMA was performed according to the guidelines of the requirements of the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) checklist for NMA (Supplemental Appendix 3).[28] A systematic literature search was performed in the PubMed, Embase, and Cochrane Library databases. The search strategy was performed on all index tests (ultrasound, mammogram, MRI, cytology, and ductoscopy) and their synonyms. The full electronic search strategy can be found in Supplemental Appendix 1 in the online version. After removal of duplicates, two authors (M.F., S.P.) independently screened articles by title and abstract. The full articles were independently screened for eligibility based on predefined inclusion and exclusion criteria. Discordant judgments were discussed by the two authors until consensus was reached.

Selection of Studies

Full text was retrieved for studies that evaluated ultrasound, mammogram, MRI, cytology, and/or ductoscopy, reported original data, and were written in English. Inclusion criteria included:

1. Participants: patients with PND as main breast complaint without history of breast cancer.
2. Intervention: ultrasound, mammogram, MRI, cytology, and/or ductoscopy.
3. Comparator: if patients were diagnosed with malignancy, they must have had definitive diagnosis of malignancy by means of biopsy or histopathologic analysis after surgery.
4. Outcome: diagnostic performance of the different diagnostic methods for the detection of (pre)cancerous lesions.

Studies were excluded from systematic review for the following reasons:

1. Not possible to determine sensitivity and specificity from the studies by means of reported true-positive, true-negative, false-positive, and true-negative rates.
2. Case report, review, and conference abstracts.

Risk of Bias

The QUADAS-2 tool was used to evaluate the quality of each eligible study.[29] The entire scale constituted 4 domains for the risk of bias: patient selection, index test, reference standard, and flow and timing. Additionally, there were 3 domains for applicability concerns: patient selection, index test, and reference standard. Each domain was judged for 3 levels of bias: low risk, intermediate/unclear risk, or high risk of bias. Full assessment criteria can be found in Supplemental Appendix 2 in the online version.

Classifications

Ultrasound, mammogram and MRI were classified according to the Breast Imaging Reporting and Data System (BI-RADS) reporting system.[30] BI-RADS I-III was considered benign, and BI-RADS IV-VI was considered malignant or suspicious for malignancy. When cytologic examination indicated atypical cells, it was considered suspicious for malignancy.

Statistical Analysis

First, sensitivity, specificity, positive predicted value (PPV), negative predicted value (NPV), and diagnostic accuracy (DA; number of truly positive and truly negative results divided by the total number of patients) were calculated for each of the 5 diagnostic methods for the diagnosis of nipple discharge for each study. After this, pooled estimates of sensitivity, specificity, PPV, NPV, and DA were calculated for each of the 5 diagnostic methods using fixed-effects models. Heterogeneity among studies was quantified by I^2 analysis and tested by the Cochran chi-square tests. Second, statistical computing software and network packages were used to draw the network graphs. Each node represents a different diagnostic method in which the size of the node reflects the number of patients, and the thickness of the line connecting the nodes represents the amount of included studies. Third, traditional pairwise meta-analyses were performed to compare different diagnostic modalities. Fourth, Bayesian network meta-analyses using the Mantel-Haenszel method were performed to combine the evidence from direct and indirect comparisons. Fifth, the surface under the cumulative ranking curve (SUCRA) was used to calculate for each intervention's being a measure of comparative diagnostic performance. A higher SUCRA value means that the intervention is likely to be ranked better than the comparators.[31] Additionally, the separate indirect from direct design evidence (SIDDE) method was used to test the local consistency assumption of the NMA.[32,33] Finally, subgroup traditional pairwise meta-analysis was performed to compare ductoscopy to MRI in studies in which all participants had negative ultrasound and/or mammogram, to compare the added value of ductoscopy and MRI to conventional imaging. $P < .05$ and 95% confidence intervals of odds ratios not containing 1 were considered statistically significant.

Comparison-adjusted funnel plots for the NMA were performed to detect the small study effects on data. The Egger,

Begg-Mazumdar, and Thomson-Sharp tests were used to quantify and test for asymmetry. For the subgroup traditional meta-analysis, the Egger test was used to quantify asymmetry.[34-36] $P \geq .05$ indicated insufficient evidence for asymmetry and therefore also for no small sample bias and no publication bias.

All calculations were performed by RStudio 1.2.5001 (with R x64 3.6.1) (https://rstudio.com/). Additionally, the following statistical packages were used for all computations of the network meta-analysis (NMA) and traditional meta-analyses: meta, mada, metafhor, gemtc, mvmeta, and netmeta. Visualization of plots was done using the ggplot2 package.

Results

Selected Articles

We followed the PRISMA NMA checklist of items to include when reporting a systematic review involving a NMA (Supplemental Appendix 4 in the online version).

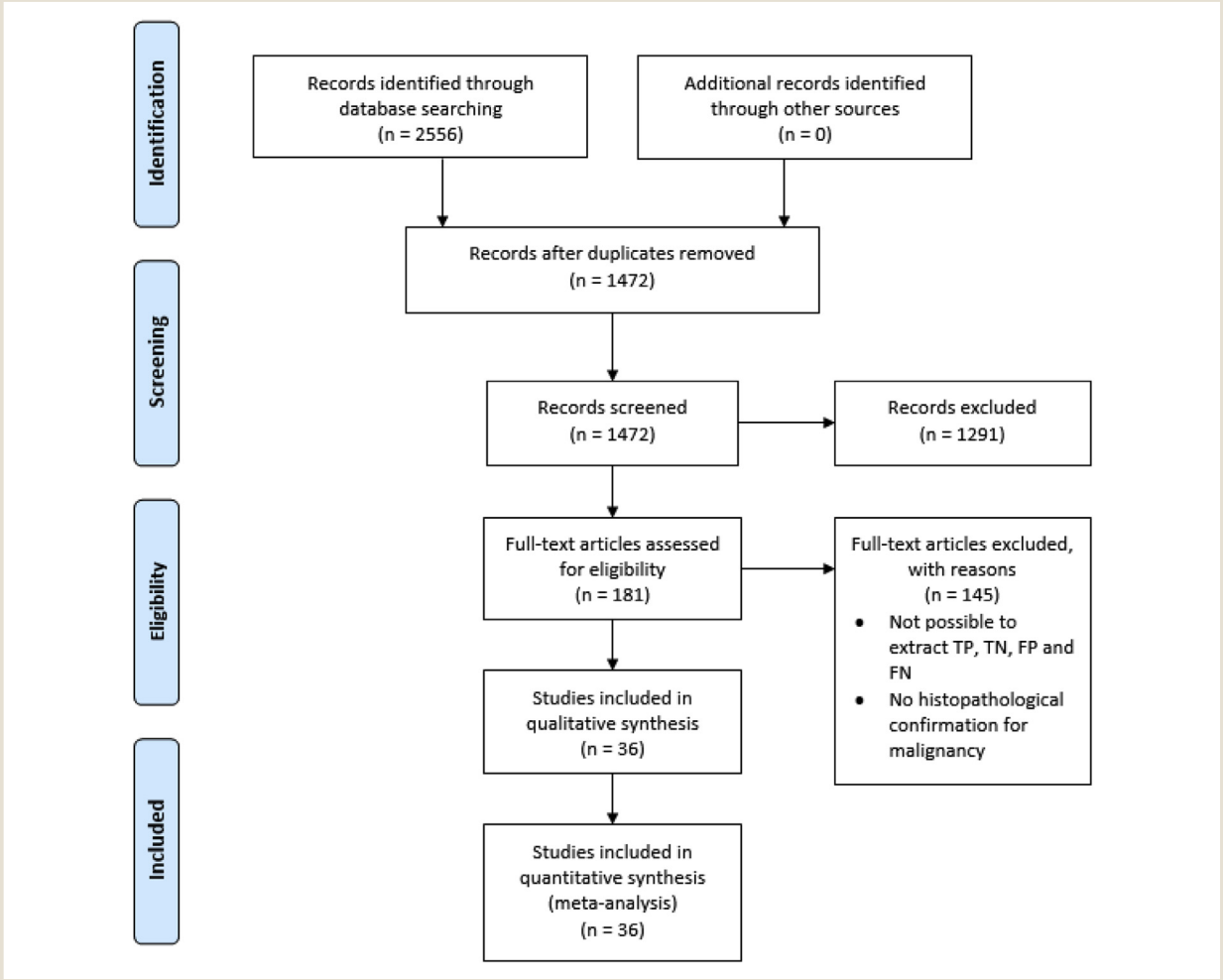
A total of 2583 citations were identified by the search and, after removing duplicates, and 181 potentially eligible articles were retrieved in full text (Figure 1). Overall, 3764 patients in 36 studies with PND underwent ultrasound, mammogram, MRI, cytology, and/or ductoscopy and were analyzed with an average of 104.6 participants per study with standard deviation of 68.7.

Figure 2 shows the network of eligible comparisons for sensitivity, specificity, NPV, PPV, and DA of the different diagnostic methods. The single-armed studies were not included in the network. Table 1 shows the studies included in the analysis and their characteristics.

Risks of Bias

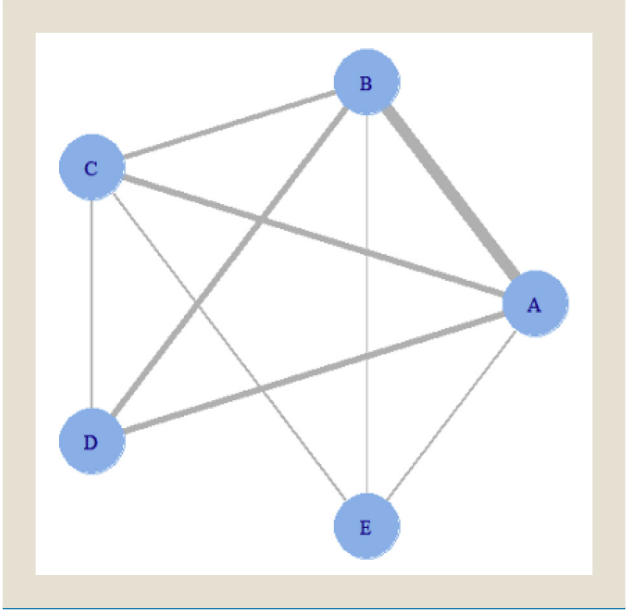
The result of the QUADAS-2 tool revealed that all the included studies were of sufficient quality. This was for both risk-of-bias domains and applicability domains (Supplemental Figure 8 in the online version). Detailed information for each enrolled study can be found in Supplemental Figure 11 in the online version.

Figure 1 Flowchart Showing Literature Search and Study Selection. A Total of 36 Relevant Studies Were Ultimately Enrolled Into Our Network Meta-analysis on Diagnostic Approach to Pathologic Nipple Discharge



Abbreviations: FN = falsely negative; FP = falsely positive; TN = truly negative; TP = truly positive.

Figure 2 Evidence Network Plot of Diagnostic Value of Imaging Methods for Diagnosis of Pathologic Nipple Discharge. Imaging Methods Included (A) Ultrasound, (B) Mammogram, (C) Magnetic Resonance Imaging, (D) Cytology, and (E) Ductoscopy



Sensitivity, Specificity, PPV, NPV, and DA

Figure 3 summarizes the different ways of assessing accuracy for the detection of breast cancer in patients with PND. In terms of pooled sensitivity for the detection of malignity in patients with PND, MRI showed highest average sensitivity (83%), followed by ductoscopy (58%), ultrasound (50%), cytology (38%), and mammogram (22%). Pooled specificity was highest for mammogram, at 93%, followed by ductoscopy (92%), cytology (90%), MRI (76%), and ultrasound (69%). PPV was highest for mammogram (46%), followed by ductoscopy (41%), MRI (40%), cytology (39%), and ultrasound (31%). Pooled NPV was highest for ductoscopy and MRI (both 96%), followed by cytology (89%), ultrasound (83%), and mammogram (80%). The highest DA was seen for ductoscopy (88%), followed by cytology (82%), MRI (77%), mammogram (76%), and ultrasound (65%). Sensitivity, specificity, PPV, NPV, and DA of individual studies can be found in Supplemental Figures 1 to 5 in the online version.

Pairwise meta-analysis from the literature (Table 2) showed mammogram to have statistically significantly lower sensitivity for the detection of malignancy in patients with PND than ultrasound, MRI, cytology, and ductoscopy. Additionally, ultrasound has a significantly lower sensitivity than MRI. Other combinations did not show statistically significant differences or were not directly measured in the currently available literature. Ultrasound showed significantly higher specificity than MRI, cytology, and mammogram. No statistical differences in specificity between ductoscopy and ultrasound were found. Mammogram has a statistically significantly higher PPV than ultrasound, MRI, and cytology. No direct comparisons in the literature were found between mammogram and ductoscopy. No other combinations showed statistical differences in PPV between the other diagnostic methods for the detection of breast cancer in patients with PND. NPV of MRI was statistically significantly higher than ultrasound and mammogram. No other comparisons between the 5 diagnostic methods showed any significant differences, although no direct comparison was available between ductoscopy and cytology.

Overall NMA

The estimates resulting from direct and indirect evidence of diagnostic methods for the detection of breast cancer in patients with PND are provided in Table 3. Sensitivity of MRI was significantly superior to ultrasound, mammogram, and cytology but did not differ significantly from ductoscopy. Ductoscopy performed significantly better than mammogram, and ultrasound performed better than mammogram. Other comparisons did not show significant differences in sensitivity. Mammogram showed to be significantly more specific than all other diagnostic methods. Additionally, ultrasound was statistically inferior to cytology but superior to MRI for specificity. Furthermore, MRI was less specific than cytology. Ductoscopy did not differ significantly from cytology, MRI, and ultrasound with regard to specificity. PPV of ultrasound was inferior to mammogram but was not statistically different from other diagnostic methods. Mammogram had a significantly higher PPV than MRI but was similar to ductoscopy. Other combinations showed no statistically significant differences. NPV of ductoscopy did not differ significantly from the other diagnostic methods. MRI had significantly higher NPV than ultrasound and mammogram but was comparable to cytologic assessment. Finally, cytology had a higher NPV than mammogram. DA of ultrasound was significantly lower than mammogram and cytology but did not differ significantly from MRI and ductoscopy. MRI had a lower DA than mammogram and cytology but did not differ significantly from ductoscopy. All other comparisons did not show significant differences (Supplemental Table 1 in the online version).

SUCRA values indicating the probability of having the best diagnostic performance are shown in Supplemental Table 2 in the online version. The highest SUCRA value for sensitivity was seen for mammogram and the lowest for ductoscopy. The highest SUCRA value for specificity was seen for MRI, followed by ductoscopy, with the lowest seen for mammogram. Ductoscopy showed the highest SUCRA value for PPV and mammogram the lowest. Ultrasound showed the highest SUCRA value for NPV and the MRI the lowest. The highest SUCRA values for DA were seen for ultrasound and the lowest for cytology. Finally, the SIDDE approach did not provide strong evidence for inconsistency with regard to sensitivity, specificity, PPV, NPV, and DA for all comparisons (all $P > .182$).

Subgroup Meta-analysis

Subgroup analysis comparing sensitivity, specificity, PPV, and NPV of MRI and ductoscopy when conventional imaging (ultrasound and mammogram) is negative can be found in Figure 4 and Table 4. Pooled sensitivity of 5 remaining studies each was 44% for ductoscopy and 76% for MRI (not significant). Specificity and DA were significantly higher for ductoscopy compared to MRI (98% vs. 84%, 95%-83%, respectively). PPV and NPV were insignificant between ductoscopy and MRI. Individual diagnostic performance of

Table 1 Baseline Characteristics of Various Diagnostic Studies on Pathologic Nipple Discharge

Study	Year	Country	Reference Standard	N	Diagnostic Method			
					D1	D2	D3	D4
Groves[37]	1996	UK	Histopathologic diagnosis	216	D			
Hou[38]	2000	Taiwan	Histopathologic diagnosis	111	D			
Orel[39]	2000	USA	Histopathologic diagnosis	15	C			
Hou[40]	2002	Taiwan	Histopathologic diagnosis	176	A	B		
Cabioglu[41]	2003	USA	Histopathologic diagnosis	142	A	B	D	
Simmons[42]	2003	USA	Histopathologic diagnosis	59	A	B	D	
Yamamoto[43]	2003	Japan	Histopathologic diagnosis	60	D			
Moncrief[44]	2005	USA	Histopathologic diagnosis	59	E			
Morrogh[45]	2007	USA	Histopathologic diagnosis	33	C			
Denewer[46]	2008	Egypt	Histopathologic diagnosis	53	E			
Bender[47]	2009	Turkey	Histopathologic diagnosis	102	E			
Kooistra[13]	2009	Netherlands	Histopathologic diagnosis	163	D			
Simpson[48]	2009	Canada	Histopathologic diagnosis	39	E			
Tekin[49]	2009	Turkey	Histopathologic diagnosis	34	D			
Vaughan[50]	2009	USA	Histopathologic diagnosis	89	E			
Dolan[14]	2010	Ireland	Histopathologic diagnosis	74	D			
Morrogh[51]	2010	USA	Histopathologic diagnosis	270	A	B	C	D
Cyr[52]	2011	USA	Histopathologic diagnosis	119	E			
Lorenzon[53]	2011	Italy	Histopathologic diagnosis	38	A	B		
Bahl[54]	2015	USA	Histopathologic diagnosis	91	C			
Bahl[8]	2015	USA	Histopathologic diagnosis	262	A	B		
Van Gelder[10]	2015	Netherlands	Histopathologic diagnosis	107	C			
Waaijer[22]	2015	Netherlands	Histopathologic diagnosis	53	E			
Zhao[55]	2015	China	Histopathologic diagnosis	153	A	D		
Park[56]	2016	South Korea	Histopathologic diagnosis	67	A			
Sanders[9]	2016	USA	Histopathologic diagnosis	85	C			
Bahl[57]	2017	USA	Histopathologic diagnosis	105	C			
Lesetedi[27]	2017	South Africa	Histopathologic diagnosis	153	A	B		
Yılmaz[58]	2017	Turkey	Histopathologic diagnosis	26	A	C	E	

Table 1 Continued

Study	Year	Country	Reference Standard	N	Diagnostic Method			
					D1	D2	D3	D4
Gui[59]	2018	UK	Histopathologic diagnosis	32	E			
Kan[60]	2018	China	Histopathologic diagnosis	95	A	B	D	
Li[61]	2018	USA	Histopathologic diagnosis	257	A	B	C	D
Baydoun[62]	2019	USA	Histopathologic diagnosis	92	A	B		
Jung[63]	2019	South Korea	Histopathologic diagnosis	46	A			
Zacharioudakis[64]	2019	UK	Histopathologic diagnosis	82	C			
Filipe[65]	2020	Netherlands	Histopathologic diagnosis	206	A	B	C	E

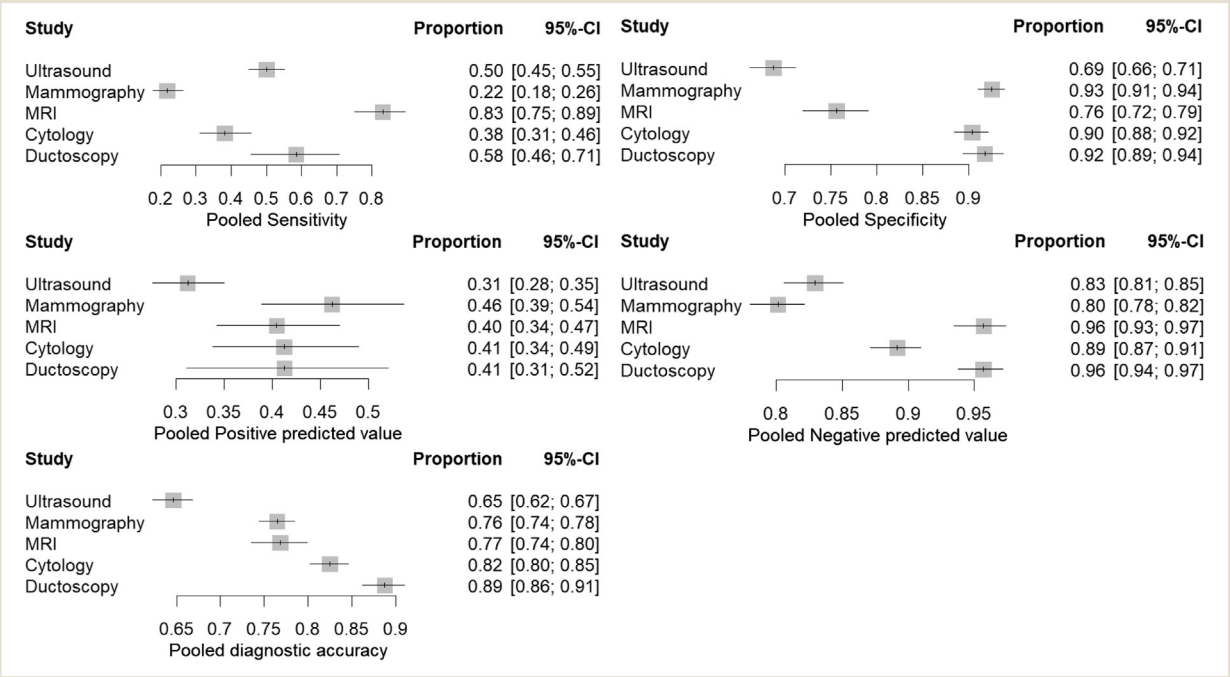
Diagnostic modalities are as follows: A = ultrasound; B = mammogram; C = magnetic resonance imaging (MRI); D = cytology; E = ductoscopy.
Abbreviations: CI = confidence interval; D1-4 = diagnostic methods.

each study can be found in Supplemental Figures 6 and 7 in the online version for, respectively, ductoscopy and MRI.

Assessment of Publication Bias

The results of assessment of publication bias showed symmetrical distributions for sensitivity, specificity, PPV, NPV, and DA. This indicates that there is no evidence to conclude small sample effects or publication bias in this NMA (Supplemental Figure 9 in the online version). Subgroup analysis also showed symmetry of the effect (Supplemental Figure 10 in the online version). For the subgroup analyses, *P* values for the Egger test for sensitivity, specificity, PPV, NPV, and DA were .0504, .7549, .8378, .2211, and .8187, respectively.

Figure 3 Usefulness of Approaches to Detection of Breast Cancer in Patients With Pathologic Nipple Discharge. Shown are Sensitivity, Specificity, Positive Predicted Value, Negative Predicted Value, and Diagnostic Accuracy of Different Approaches to Detection of Breast Cancer in Patients With Pathologic Nipple Discharge



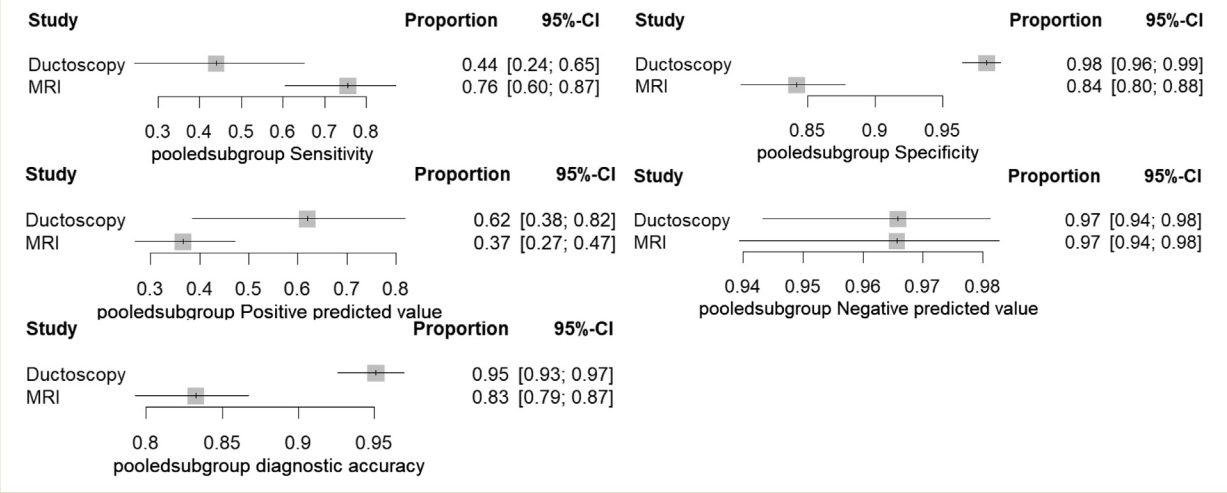
Abbreviations: CI = confidence interval; MRI = magnetic resonance imaging.

Table 2 Pairwise Meta-analysis of 5 Diagnostic Methods to Detect Breast Cancer in Patients With Pathologic Nipple Discharge

Sensitivity					Specificity				
Studies (N)	Comparison	OR	2.5% CI	97.5% CI	Studies (N)	Comparison	OR	2.5% CI	97.5% CI
10	A vs. B[b]	3.274	2.307	4.647	10	A vs. B[b]	0.185	0.141	0.243
5	A vs. C[b]	0.337	0.137	0.828	5	A vs. C[b]	1.475	1.031	2.11
5	A vs. D	0.991	0.573	1.715	5	A vs. D	0.982	0.650	1.484
2	A vs. E	0.188	0.02	1.796	2	A vs. E	3.623	0.685	19.158
4	B vs. C[b]	0.044	0.015	0.130	4	B vs. C[b]	19.532	11.903	32.051
5	B vs. D[b]	0.449	0.249	0.808	5	B vs. D[b]	2.059	1.404	3.02
1[a]	B vs. E				1	B vs. E	2.043	0.337	12.386
2	C vs. D	2.143	0.583	7.871	2	C vs. D	0.646	0.33	1.262
2	C vs. E	0.562	0.046	6.806	2	C vs. E	1.46	0.265	8.036
0	D vs. E				0	D vs. E			
Positive Predictive Value					**Negative Predictive Value**				
Studies (N)	Comparison	OR	2.5% CI	97.5% CI	Studies (N)	Comparison	OR	2.5% CI	97.5% CI
10	A vs. B[b]	0.398	0.262	0.606	10	A vs. B	1.023	0.818	1.278
5	A vs. C	0.719	0.45	1.148	5	A vs. C[b]	0.402	0.162	0.999
5	A vs. D	0.933	0.529	1.645	5	A vs. D[b]	0.320	0.203	0.505
2	A vs. E	1.406	0.206	9.619	2	A vs. E	0.326	0.085	1.252
4	B vs. C[b]	3.369	1.833	6.192	4	B vs. C[b]	0.339	0.142	0.81
5	B vs. D[b]	2.551	1.358	4.792	5	B vs. D	1.079	0.642	1.813
1[a]	B vs. E				1	B vs. E	0.303	0.064	1.422
2	C vs. D	0.893	0.416	1.918	2	C vs. D	1.867	0.515	6.765
2	C vs. E	1.688	0.249	11.416	2	C vs. E	0.678	0.100	4.586
0	D vs. E				0	D vs. E			

Diagnostic modalities are as follows: A = ultrasound; B = mammogram; C = magnetic resonance imaging (MRI); D = cytology; E = ductoscopy.
Abbreviations: CI = confidence interval; OR = odds ratio.
[a]Truly positive values were 0, so it was not possible to pairwise compare sensitivity and positive predictive value.
[b]Statistically significant.

Figure 4 Usefulness of Approaches to Detection of Breast Cancer in Patients With Pathologic Nipple Discharge and Negative Imaging Results. Shown are Pooled Sensitivity, Specificity, Positive Predicted Value, Negative Predicted Value, and Diagnostic Accuracy of Different Approaches to Detection of Breast Cancer in Subgroup of Patients With Pathologic Nipple Discharge and With Negative Mammogram And/Or Ultrasound Results



Abbreviations: CI = confidence interval; MRI = magnetic resonance imaging.

Table 3 Network Meta-analysis Comparing (in OR) 5 Diagnostic Methods to Detect Breast Cancer in Patients With Pathologic Nipple Discharge

Sensitivity				Specificity			
Comparison	**OR**	**2.5% CI**	**97.5% CI**	**Comparison**	**OR**	**2.5% CI**	**97.5% CI**
A vs. B[a]	3.9212	2.7396	5.6123	A vs. B[a]	0.2269	0.1708	0.3013
A vs. C[a]	0.2873	0.1181	0.6989	A vs. C[a]	1.8271	1.2836	2.6008
A vs. D	1.3188	0.7653	2.2728	A vs. D	0.7291	0.493	1.0783
A vs. E	0.1787	0.0197	1.6187	A vs. E	1.6403	0.4712	5.7103
B vs. C[a]	0.0733	0.0297	0.1808	B vs. C[a]	8.0542	5.3315	12.1672
B vs. D[a]	0.3363	0.1918	0.5898	B vs. D[a]	3.2138	2.1276	4.8545
B vs. E[a]	0.0456	0.0049	0.4218	B vs. E[a]	7.2304	2.0646	25.321
C vs. D[a]	4.5908	1.6963	12.4242	C vs. D[a]	0.399	0.2442	0.6521
C vs. E	0.6219	0.0645	5.9959	C vs. E	0.8977	0.2547	3.1637
D vs. E	0.1355	0.0141	1.3025	D vs. E	2.2498	0.6165	8.21
Positive Predictive Value				**Negative Predictive Value**			
Comparison	**OR**	**2.5% CI**	**97.5% CI**	**Comparison**	**OR**	**2.5% CI**	**97.5% CI**
A vs. B[a]	0.3942	0.2574	0.6038	A vs. B	0.9272	0.7309	1.1762
A vs. C	0.8313	0.5249	1.3167	A vs. C[a]	0.3108	0.1333	0.7245
A vs. D	0.8627	0.4979	1.4947	A vs. D[a]	0.5127	0.3285	0.8000
A vs. E	1.4045	0.2756	7.1581	A vs. E	0.3017	0.0830	1.0966
B vs. C[a]	2.1088	1.2226	3.6374	B vs. C[a]	0.3352	0.1470	0.7641
B vs. D[a]	2.1883	1.2002	3.9899	B vs. D[a]	0.5529	0.3500	0.8736
B vs. E	3.5629	0.6714	18.907	B vs. E	0.3254	0.0894	1.1844
C vs. D	1.0377	0.5458	1.9728	C vs. D	1.6496	0.6507	4.1822
C vs. E	1.6895	0.3317	8.6049	C vs. E	0.9708	0.2232	4.2221
D vs. E	1.6282	0.2963	8.9474	D vs. E	0.5885	0.1516	2.2854

Diagnostic modalities are as follows: A = ultrasound; B = mammogram; C = magnetic resonance imaging (MRI); D = cytology; E = ductoscopy.
Abbreviations: CI = confidence interval; OR = odds ratio.
[a]Statistically significant.

Discussion

There is no consensus on the diagnostic approach of patients with PND, especially if conventional imaging (ultrasound and mammogram) results are negative. High-quality studies on the value of ductoscopy and MRI are scarce. The first aim of this study was to compare by means of NMA the effectiveness for detection of malignancy of ultrasound, mammogram, MRI, cytology, of nipple fluid and ductoscopy in patients with PND. The second aim of this study was to compare MRI to ductoscopy when ultrasound and mammogram as conventional first-line imaging techniques have negative results.

The online search yielded 2556 hits, of which 36 studies with 3764 patients were included. The highest sensitivity was seen for MRI (83%) and ductoscopy (58%), and the highest specificity was seen for ductoscopy (92%) and mammogram (93%), whereas MRI had a relatively low specificity (76%). DA was highest for ductoscopy (88%), followed by cytology (82%) and MRI (77%). In subgroup analysis including studies in which patients had negative ultrasound and mammogram, there were no statistical differences in sensitivity, but specificity and DA were significantly higher for ductoscopy.

Ultrasound and mammogram are established, cheap breast imaging methods.[66,67] We found that the pooled average sensitivity for the detection of breast cancer in patients with PND was 50% for

Table 4 Meta-analysis Comparing (in OR) Ductoscopy With MRI for Detection of Breast Cancer in Patients With Pathologic Nipple Discharge, and Negative Mammogram and Ultrasound Results

Comparison	Measure	OR	2.5% CI	97.5% CI	*z* Score	*P*
A vs. B	Sensitivity	0.285	0.071	1.155	−1.76	.0788
A vs. B	Specificity	10.401	3.958	27.332	4.75	<.001[a]
A vs. B	PPV	2.251	0.523	9.686	1.09	.2759
A vs. B	NPV	1.247	0.485	3.21	0.46	.6468
A vs. B	DA	4.94	1.439	16.961	2.54	.0112[a]

Diagnostic modalities are as follows: A = ductoscopy; B = magnetic resonance imaging (MRI).
Abbreviations: CI = confidence interval; DA = diagnostic accuracy; NPV = negative predictive value; OR = odds ratio; PPV = positive predictive value.
[a]Statistically significant.

ultrasound but only 22% for mammogram. Pooled specificity was 69% for ultrasound but much was higher, at 93%, for mammogram. Most causes of PND (around 95%), such as papillomas and ductal ectasia, are benign,[6,7] for which ultrasound[5,8,12,63,68] and mammogram[5,12,51,60,68] indeed have a high sensitivity and specificity. Therefore, ultrasound and mammogram are likely to remain the initial approach in patients with PND.

MRI has been used more often for the detection of breast cancer in patients with PND in recent years. We found that the pooled average sensitivity of MRI is 86% and specificity is 76%. Pooled PPV of MRI was low (40%), meaning that over half of patients with a positive MRI are advised to undergo histopathologic analysis by core biopsy and/or surgery for a benign lesion (Supplemental Figure 3 in the online version).[10,11] Therefore, MRI may need to be reserved for PND cases where mammogram and ultrasound are negative.[8,10,54,64] The current NMA shows that the sensitivity of MRI then drops to 74% but specificity increases to 85% (Supplemental Figure 7). Contrast-enhanced MRI appears to be a promising approach for the detection of breast cancer in patients with PND in pilot studies.[50,69-73]

For cytology, pooled sensitivity was only 38%, although the pooled specificity was high (90%) (Supplemental Figure 4), indicating that cytology is not very useful for the detection of breast cancer in patients with PND.[13] However, biomarker analysis of nipple discharge, as by RNA assessment,[74] may be more promising.[6,69-72,75-78]

This NMA showed that ductoscopy has an average sensitivity of 58% and a high specificity of 92% for the detection of breast cancer in patients with PND. However, ductoscopy is highly suitable for detecting benign lesions causing PND.[46,47,50,59,73,79] However, without histologic sampling of the lesions found, ductoscopy images alone cannot permit reliable discrimination between benign and malignant causes; endoscopic sampling during ductoscopy is not possible; and surgery may still be warranted to exclude malignancy.[24] Ductoscopy is not a cheap technique, like MRI is, so it may be especially useful when conventional imaging is negative.[65] For this reason, we conducted a subgroup meta-analysis comparing MRI to ductoscopy in patients with PND by negative conventional imaging for the detection of breast cancer. Sensitivity for ductoscopy dropped to 44% but specificity rose to 98%, which is significantly higher than for MRI. However, the low incidence of malignancy in patients with PND leads to broad confidence intervals, meaning that the pooled relative difference in sensitivity is relatively big, but not statistically significant. Moreover, the higher incidence of malignancy in the MRI studies (around 20% vs. the 4.5% in the ductoscopy studies) may explain the relatively high sensitivity of MRI, which may therefore not be realistic. This high incidence of malignancy in MRI studies is unexpected because the reported incidence of malignancy in patients with negative echography and mammogram is around 5%.[26,27] The same applies for the PPV. Consequently, because the most common causes of PND are benign, specificity may be clinically more relevant. The present study found that ductoscopy has a statistically significantly higher specificity (and DA) than MRI, so it may be a more useful diagnostic tool in patients with PND with no signs of malignancy by conventional radiography. Additionally, intraductal biopsies are nowadays possible with the basket extraction device,[22] and new techniques surrounding ductoscopy are being developed in order to increase the sensitivity for the detection of (pre)malignant lesions, such as autofluorescent imaging.[80,81]

To our knowledge, this is the first systematic review comparing different diagnostic methods for the detection of malignancy in patients with PND. However, there are some limitations of this NMA. This study could not provide enough direct comparisons between the 5 individual imaging methods as a result of limited evidence. Nonetheless, further research is warranted comparing availability, impact to the patient, and cost-effectiveness of the different diagnostic methods. Finally, most women with PND undergo surgical procedures, such as microdochectomy or major duct excision, to exclude malignancy and treat the PND symptoms.[6,8,9] These surgical procedures are performed under general anesthesia, are expensive, and are associated with scarring, which may result in breastfeeding difficulties.[68] Additionally, heterogeneity was high within the groups and between different groups. We cannot explain the high heterogeneity because we consider the methodology and patient population of the studies to be very similar, especially in the subgroup analysis. Furthermore, we have no explanation for the high incidence of malignancy in the MRI studies.

To conclude, our findings suggest that although ultrasound and mammogram may remain low-cost useful first choices for the detection of malignancy in patients with PND, ductoscopy outperforms most imaging techniques (especially MRI) and cytology.

Disclosure

The authors have stated that they have no conflict of interest.

Supplemental Data

Supplemental tables, figures, and appendices accompanying this article can be found in the online version at https://doi.org/10.1016/j.clbc.2020.05.015.

References

1. Hussain AN, Policarpio C, Vincent MT. Evaluating nipple discharge. *Obstet Gynecol Surv* 2006; 61:278-83.
2. Dixon JM, Mansel RE. ABC of breast diseases. Symptoms assessment and guidelines for referral. *BMJ* 1994; 309:722-6.
3. Seltzer MH. Breast complaints, biopsies, and cancer correlated with age in 10,000 consecutive new surgical referrals. *Breast J* 2004; 10:111-7.
4. Santen RJ, Mansel R. Benign breast disorders. *N Engl J Med* 2005; 353:275-85.
5. King TA, Carter KM, Bolton JS, Fuhrman GM. A simple approach to nipple discharge. *Am Surg* 2000; 66:960-6.
6. Alcock C, Layer GT. Predicting occult malignancy in nipple discharge. *ANZ J Surg* 2010; 80:646-9.
7. Albrecht C, Thele F, Grunwald S, et al. Nipple discharge: role of ductoscopy in comparison with standard diagnostic tests. *Onkologie* 2013; 36:12-6.
8. Bahl M, Baker JA, Greenup RA, Ghate SV. Diagnostic value of ultrasound in female patients with nipple discharge. *AJR Am J Roentgenol* 2015; 205:203-8.
9. Sanders LM, Daigle M. The rightful role of MRI after negative conventional imaging in the management of bloody nipple discharge. *Breast J* 2016; 22:209-12.
10. Van Gelder L, Bisschops RH, Menke-Pluymers MB, Westenend PJ, Plaisier PW. Magnetic resonance imaging in patients with unilateral bloody nipple discharge; useful when conventional diagnostics are negative? *World J Surg* 2015; 39:184-6.
11. Expert Panel on Breast Imaging, Lee SJ, Trikha S, et al. ACR Appropriateness Criteria® evaluation of nipple discharge. *J Am Coll Radiol* 2017; 14:S138-53.
12. de Paula IB, Campos AM. Breast imaging in patients with nipple discharge. *Radiol Bras* 2017; 50:383-8.
13. Kooistra BW, Wauters C, van de Ven S, Strobbe L. The diagnostic value of nipple discharge cytology in 618 consecutive patients. *Eur J Surg Oncol* 2009; 35:573-7.
14. Dolan RT, Butler JS, Kell MR, Gorey TF, Stokes MA. Nipple discharge and the efficacy of duct cytology in evaluating breast cancer risk. *Surgeon* 2010; 8:252-8.

15. Yamamoto D, Shoji T, Kawanishi H, et al. A utility of ductography and fiberoptic ductoscopy for patients with nipple discharge. *Breast Cancer Res Treat* 2001; 70:103-8.
16. Yamamoto D, Ueda S, Senzaki H, et al. New diagnostic approach to intracystic lesions of the breast by fiberoptic ductoscopy. *Anticancer Res* 2001; 21:4113-6.
17. Okazaki A, Okazaki M, Asaishi K, et al. Fiberoptic ductoscopy of the breast: a new diagnostic procedure for nipple discharge. *Jpn J Clin Oncol* 1991; 21:188-93.
18. Matsunaga T, Ohta D, Misaka T, et al. Mammary ductoscopy for diagnosis and treatment of intraductal lesions of the breast. *Breast Cancer* 2001; 8:213-21.
19. Grunwald S, Heyer H, Paepke S, et al. Diagnostic value of ductoscopy in the diagnosis of nipple discharge and intraductal proliferations in comparison to standard methods. *Onkologie* 2007; 30:243-8.
20. Han Y, Li J, Han S, Jia S, Zhang Y, Zhang W. Diagnostic value of endoscopic appearance during ductoscopy in patients with pathological nipple discharge. *BMC Cancer* 2017; 17:300-3.
21. Kamali S, Bender O, Kamali GH, Aydin MT, Karatepe O, Yuney E. Diagnostic and therapeutic value of ductoscopy in nipple discharge and intraductal proliferations compared with standard methods. *Breast Cancer* 2014; 21:154-61.
22. Waaijer L, van Diest PJ, Verkooijen HM, et al. Interventional ductoscopy in patients with pathological nipple discharge. *Br J Surg* 2015; 102:1639-48.
23. Jacobs VR, Kiechle M, Plattner B, Fischer T, Paepke S. Breast ductoscopy with a 0.55-mm mini-endoscope for direct visualization of intraductal lesions. *J Minim Invasive Gynecol* 2005; 12:359-64.
24. Waaijer L, Simons JM, Borel Rinkes IH, van Diest PJ, Verkooijen HM, Witkamp AJ. Systematic review and meta-analysis of the diagnostic accuracy of ductoscopy in patients with pathological nipple discharge. *Br J Surg* 2016; 103: 632-43.
25. Makita M, Akiyama F, Gomi N, Iwase T. Mammary ductoscopy and watchful follow-up substitute microdochectomy in patients with bloody nipple discharge. *Breast Cancer* 2016; 23:242-51.
26. Galvin R, Joyce D, Downey E, Boland F, Fahey T, Hill AK. Development and validation of a clinical prediction rule to identify suspected breast cancer: a prospective cohort study. *BMC Cancer* 2014; 14:743.
27. Lesetedi C, Rayne S, Kruger D, Benn CA. Indicators of breast cancer in patients undergoing microdochectomy for a pathological nipple discharge in a middle-income country. *J Surg Res* 2017; 220:336-40.
28. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015; 162:777-84.
29. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155:529-36.
30. Rao AA, Feneis J, Lalonde C, Ojeda-Fournier H. A pictorial review of changes in the BI-RADS fifth edition. *Radiographics* 2016; 36:623-39.
31. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011; 64:163-71.
32. Efthimiou O, Debray TP, van Valkenhoef G, et al. GetReal in network meta-analysis: a review of the methodology. *Res Synth Methods* 2016; 7:236-63.
33. Efthimiou O, Rucker G, Schwarzer G, Higgins JPT, Egger M, Salanti G. Network meta-analysis of rare events using the Mantel-Haenszel method. *Stat Med* 2019; 38:2992-3012.
34. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; 315:629-34.
35. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994; 50:1088-101.
36. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med* 1999; 18:2693-708.
37. Groves AM, Carr M, Wadhera V, Lennard TWJJ. An audit of cytology in the evaluation of nipple discharge. A retrospective study of 10 years' experience. *Breast* 1996; 5:96-9.
38. Hou M, Tsai K, Lin H, Chai C, Liu C, Huang T. A simple intraductal aspiration method for cytodiagnosis in nipple discharge. *Acta Cytol* 2000; 44:1029-34.
39. Orel SG, Dougherty CS, Reynolds C, Czerniecki BJ, Siegelman ES, Schnall MD. MR imaging in patients with nipple discharge: initial experience. *Radiology* 2000; 216:248-54.
40. Hou MF, Tsai KB, Ou-Yang F, et al. Is a one-step operation for breast cancer patients presenting nipple discharge without palpable mass feasible? *Breast* 2002; 11:402-7.
41. Cabioglu N, Hunt KK, Singletary SE, et al. Surgical decision making and factors determining a diagnosis of breast carcinoma in women presenting with nipple discharge. *J Am Coll Surg* 2003; 196:354-64.
42. Simmons R, Adamovich T, Brennan M, et al. Nonsurgical evaluation of pathologic nipple discharge. *Ann Surg Oncol* 2003; 10:113-6.
43. Yamamoto D, Senzaki H, Nakagawa H, Okugawa H, Gondo H, Tanaka K. Detection of chromosomal aneusomy by fluorescence in situ hybridization for patients with nipple discharge. *Cancer* 2003; 97:690-4.
44. Moncrief RM, Nayar R, Diaz LK, Staradub VL, Morrow M, Khan SA. A comparison of ductoscopy-guided and conventional surgical excision in women with spontaneous nipple discharge. *Ann Surg* 2005; 241:575-81.
45. Morrogh M, Morris EA, Liberman L, Borgen PI, King TA. The predictive value of ductography and magnetic resonance imaging in the management of nipple discharge. *Ann Surg Oncol* 2007; 14:3369-77.
46. Denewer A, El-Etribi K, Nada N, El-Metwally M. The role and limitations of mammary ductoscope in management of pathologic nipple discharge. *Breast J* 2008; 14:442-9.
47. Bender O, Balci FL, Yuney E, Akbulut H. Scarless endoscopic papillomectomy of the breast. *Onkologie* 2009; 32:94-8.
48. Simpson JS, Connolly EM, Leong WL, et al. Mammary ductoscopy in the evaluation and treatment of pathologic nipple discharge: a Canadian experience. *Can J Surg* 2009; 52:E245-8.
49. Tekin E, Akin M, Kurukahvecioglu O, et al. The value of breast ductoscopy in radiologically negative spontaneous/persistent nipple discharge. *Breast J* 2009; 15: 329-32.
50. Vaughan A, Crowe JP, Brainard J, Dawson A, Kim J, Dietz JR. Mammary ductoscopy and ductal washings for the evaluation of patients with pathologic nipple discharge. *Breast J* 2009; 15:254-60.
51. Morrogh M, Park A, Elkin EB, King TA. Lessons learned from 416 cases of nipple discharge of the breast. *Am J Surg* 2010; 200:73-80.
52. Cyr AE, Margenthaler JA, Conway J, et al. Correlation of ductal lavage cytology with ductoscopy-directed duct excision histology in women at high risk for developing breast cancer: a prospective, single-institution trial. *Ann Surg Oncol* 2011; 18:3192-7.
53. Lorenzon M, Zuiani C, Linda A, Londero V, Girometti R, Bazzocchi M. Magnetic resonance imaging in patients with nipple discharge: should we recommend it? *Eur Radiol* 2011; 21:899-907.
54. Bahl M, Baker JA, Greenup RA, Ghate SV. Evaluation of pathologic nipple discharge: what is the added diagnostic value of MRI? *Ann Surg Oncol* 2015; 22(suppl 3):S435-41.
55. Zhao S, Gai X, Wang Y, et al. Diagnostic values of carcinoembryonic antigen, cancer antigen 15-3 and cancer antigen 125 levels in nipple discharge. *Chin J Physiol* 2015; 58:385-92.
56. Park CJ, Kim EK, Moon HJ, Yoon JH, Kim MJ. Reliability of breast ultrasound BI-RADS final assessment in mammographically negative patients with nipple discharge and radiologic predictors of malignancy. *J Breast Cancer* 2016; 19:308-15.
57. Bahl M, Gadd MA, Lehman CD. Journal Club: Diagnostic utility of MRI after negative or inconclusive mammography for the evaluation of pathologic nipple discharge. *AJR Am J Roentgenol* 2017; 209:1404-10.
58. Yilmaz R, Bender O, Yabul FC, Dursun M, Tunaci M, Acunas G. Diagnosis of nipple discharge: value of magnetic resonance imaging and ultrasonography in comparison with ductoscopy. *Balkan Med J* 2017; 34:119-26.
59. Gui G, Agusti A, Twelves D, et al. INTEND II randomized clinical trial of intraoperative duct endoscopy in pathological nipple discharge. *Br J Surg* 2018; 105:1583-90.
60. Kan WM, Chen C, Kwong A. Implications of nipple discharge in Hong Kong Chinese women. *Hong Kong Med J* 2018; 24:18-24.
61. Li GZ, Wong SM, Lester S, Nakhlis F. Evaluating the risk of underlying malignancy in patients with pathologic nipple discharge. *Breast J* 2018; 24:624-7.
62. Baydoun S, Gonzalez P, Whitman GJ, Dryden M, Xi Y, Dogan B. Is ductography still warranted in the 21st century? *Breast J* 2019; 25:654-62.
63. Jung HK, Park YM, Baek HJ, et al. Comparison between ultrasonography and galactography in detecting lesions in patients with pathologic nipple discharge. *Ultrasound Q* 2019; 35:93-8.
64. Zacharioudakis K, Kontoulis T, Vella JX, et al. Can we see what is invisible? The role of MRI in the evaluation and management of patients with pathological nipple discharge. *Breast Cancer Res Treat* 2019; 178:115-20.
65. Filipe MD, Waaijer L, van der Pol CC, van Diest P, Witkamp AJ. Interventional ductoscopy as an alternative for major duct excision or microdochectomy in women suffering pathological nipple discharge: a single centre experience. *Clin Breast Cancer* 2020; 20:e334-43.
66. Health Quality Ontario. Ultrasound as an adjunct to mammography for breast cancer screening: a health technology assessment. *Ont Health Technol Assess Ser* 2016; 16:1-71.
67. Mandrik O, Ekwunife OI, Meheus F, et al. Systematic reviews as a "lens of evidence": determinants of cost-effectiveness of breast cancer screening. *Cancer Med* 2019; 8:7846-58.
68. Sarakbi W Al, Worku D, Escobar PF, Mokbel K. Breast papillomas: current management with a focus on a new diagnostic and therapeutic modality. *Int Semin Surg Oncol* 2006; 3:1.
69. Fackler MJ, Rivers A, Teo WW, et al. Hypermethylated genes as biomarkers of cancer in women with pathologic nipple discharge. *Clin Cancer Res* 2009; 15: 3802-11.
70. Fought AJ, McGathey C, Scholtens DM, et al. Hormonal determinants of nipple aspirate fluid yield among breast cancer cases and screening controls. *Cancer Epidemiol Biomarkers Prev* 2013; 22:2277-84.
71. Shao ZM, Nguyen M. Nipple aspiration in diagnosis of breast cancer. *Semin Surg Oncol* 2001; 20:175-80.
72. Wang G, Qin Y, Zhang J, et al. Nipple discharge of CA15-3, CA125, CEA and TSGF as a new biomarker panel for breast cancer. *Int J Mol Sci* 2014; 15: 9546-65.
73. Zielinski J, Jaworski R, Irga-Jaworska N, Pikula M, Hunerbein M, Jaskiewicz J. Use of fiberoductoscopy for the management of patients with pathological nipple discharge: experience of a single center in Poland. *Breast Cancer* 2018; 25:753-8.
74. Zhang K, Zhao S, Wang Q, Yang HS, Zhu J, Ma R. Identification of microRNAs in nipple discharge as potential diagnostic biomarkers for breast cancer. *Ann Surg Oncol* 2015; 22(suppl 3):S536-44.
75. Zhao S, Mei Y, Wang J, Zhang K, Ma R. Different levels of CEA, CA153 and CA125 in milk and benign and malignant nipple discharge. *PLoS One* 2016; 11: e0157639.

76. Brunoro GVF, Carvalho PC, Barbosa VC, et al. Differential proteomic comparison of breast cancer secretome using a quantitative paired analysis workflow. *BMC Cancer* 2019.
77. Loud JT, Thiebaut AC, Abati AD, et al. Ductal lavage in women from *BRCA1/2* families: is there a future for ductal lavage in women at increased genetic risk of breast cancer? *Cancer Epidemiol Biomarkers Prev* 2009; 18:1243-51.
78. Antill YC, Mitchell G, Johnson SA, et al. Gene methylation in breast ductal fluid from *BRCA1* and *BRCA2* mutation carriers. *Cancer Epidemiol Biomarkers Prev* 2010; 19:265-74.
79. Zielinski J, Jaworski R, Irga-Jaworska N, Haponiuk I, Jaskiewicz J. The significance of ductoscopy of mammary ducts in the diagnostics of breast neoplasms. *Wideochir Inne Tech Maloinwazyjne* 2015; 10:79-86.
80. Jacobs VR, Paepke S, Schaaf H, Weber BC, Kiechle-Bahat M. Autofluorescence ductoscopy: a new imaging technique for intraductal breast endoscopy. *Clin Breast Cancer* 2007; 7:619-23.
81. Douplik A, Leong WL, Easson AM, Done S, Netchev G, Wilson BC. Feasibility study of autofluorescence mammary ductoscopy. *J Biomed Opt* 2009; 14:44036.

Supplemental Appendix 1

Search Strategy

PubMed

((("Nipple Discharge"[Mesh]) OR nipple discharge*[Title/Abstract])) AND ((("Ultrasonography"[Mesh] OR "Cytological Techniques"[Mesh] OR "Mammography"[Mesh] OR "Magnetic Resonance Imaging"[Mesh] OR "Endoscopy"[Mesh] OR "ductoscopy"[tiab] OR "ductoscope"[tiab])) OR ((Ultrasonograph*[Title/Abstract] OR echograph*[Title/Abstract] OR ultrasound[Title/Abstract] OR Cytological Technique*[Title/Abstract] OR mammograph*[Title/Abstract] OR chest X-ray*[Title/Abstract] OR Magnetic Resonance Imaging[Title/Abstract] OR MRI[Title/Abstract] OR endoscop*[Title/Abstract] OR ductoscop*[Title/Abstract] OR fiberoductoscop*[Title/Abstract] OR FDS[tiab])))

Embase

('breast discharge'/exp OR 'breast discharge*':ti,ab,kw) AND ('echography'/exp OR 'echography':ti,ab,kw OR 'cytology'/exp OR 'cytology':ti,ab,kw OR 'mammography'/exp OR 'mammography':ti,ab,kw OR 'nuclear magnetic resonance imaging'/exp OR 'nuclear magnetic resonance imaging':ti,ab,kw OR 'ductoscop*':ti,ab,kw OR 'endoscopy'/exp OR 'endoscopy':ti,ab,kw OR 'fiberoductoscop*':ti,ab,kw)

Cochrane

(nipple discharge*:ti,ab,kw AND (ductoscop*:ti,ab,kw OR Ultrasonograph*:ti,ab,kw OR echograph*:ti,ab,kw OR ultrasound:ti,ab,kw OR Cytology:ti,ab,kw OR Cytological Technique*:ti,ab,kw OR mammograph*:ti,ab,kw OR chest X-ray*:ti,ab,kw OR Magnetic Resonance Imaging:ti,ab,kw OR MRI:ti,ab,kw OR endoscop*:ti,ab,kw OR ductoscop*:ti,ab,kw OR fiberoductoscop*:ti,ab,kw OR FDS:ti,ab,kw)

Supplemental Appendix 2

QUADAS-2 Tool Signaling Questions to Assess Quality of Included Studies

A. Evaluation of bias

1. Patient selection—Could the selection of patients have introduced bias?

Risk: Low/Intermediate/High/Unclear

a. Was a consecutive or random sample of patients enrolled?
b. Was a case−control design avoided?
c. Was selection bias avoided by including patients with pathologic nipple discharge (PND) and comparable previous diagnostic workup?

2. Index test (ductoscopy/cytology/ultrasound [US]/mammogram/magnetic resonance imaging [MRI])—Could the conduct or interpretation of the index test have introduced bias?

Risk: Low/Intermediate/High/Unclear

a. Was the ductoscopy/cytology/US/mammogram/MRI outcome interpreted without knowledge of the histologic outcome?
b. Was the ductoscopic visual scoring system cytology/US/mammogram/MRI interpretation specified?

3. Reference standard (histology or follow-up in a defined selection of patients)—Could the reference standard, its conduct, or its interpretation have introduced bias?

Risk: Low/Intermediate/High/Unclear

a. Was histology used as a reference standard (especially for the malignant cases)?
b. When histology was not used as a reference standard in all patients, was follow-up performed in the other group of patients?

If there is no follow-up for one subgroup, 1 point is assigned to this subcategory. Because all studies are retrospective, this implies that there was a follow-up, namely at the moment the medical records were reviewed.

4. Flow and timing

Risk: Low/Intermediate/High/Unclear

a. Was the ductoscopy/cytology/US/mammogram/MRI performed within 1 to 3 months before histology?

If the answer is "no" or "unclear," 0 points are assigned to this subcategory.

b. Did all patients receive a reference standard?
c. Were all patients included in the analysis (even if a subgroup had mammography or cytology or MRI; it's fine if only a subgroup was analyzed, as long as it is justified/logical)?

B. Evaluation of applicability

1. Patient selection—Is there concern that the included patients do not match the review question?

Concern: Low/Intermediate/High/Unclear

a. Was the patient group studied matching with the review question?

If (i) the patient population in the selected articles presents with PND as a chief complaint, if (ii) PND has a slight different or specific definition, and/or (iii) if patients are from a non-Western country, 0 points are assigned to this subcategory.

b. Were patients included before they were referred for surgery, and not when they were already referred for surgery?

If the answer is "no" or "unclear," 0 points are assigned to this subcategory. Yes leads to 1 point in this subcategory.

2. Index test (ductoscopy/cytology/US/mammogram/MRI)—Is there concern that the index test, its conduct, or interpretation differ from the review question?

Concern: Low/Intermediate/High/Unclear

a. Was the same interpretation of (visual) findings used in every patient?

A clear explanation of the classification system used (eg, Breast Imaging Reporting and Data System [BIRADS]) or description of the

aspects evaluated, is assigned 0 points. A "positive" or "negative" classification, without explanation, is assigned 1 point to this subcategory.

b. Was ductoscopy/cytology/US/mammogram/MRI performed by an experienced operator?

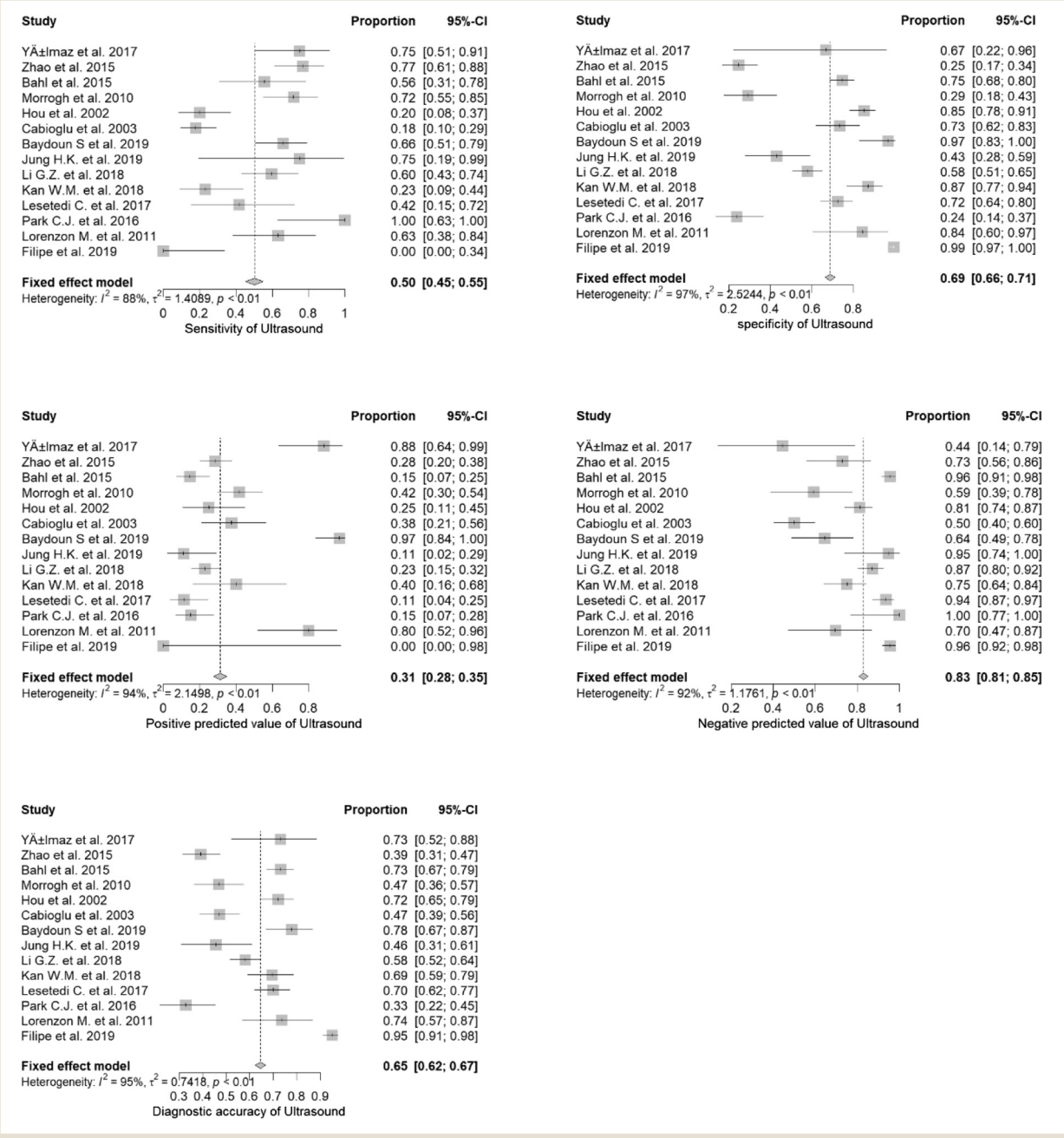
If in the article it is clearly stated that a second and/or independent and/or experienced operator revised the imagery or cytology, 0 points are assigned to this subcategory.

3. Reference standard (histology or follow-up)—Is there concern that the target condition as defined by the reference standard does not match the review question?

Concern: Low/Intermediate/High/Unclear

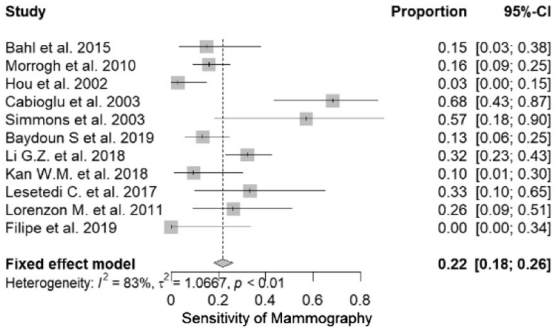
a. Was histology or follow-up in a defined selection of patients used as reference standard?

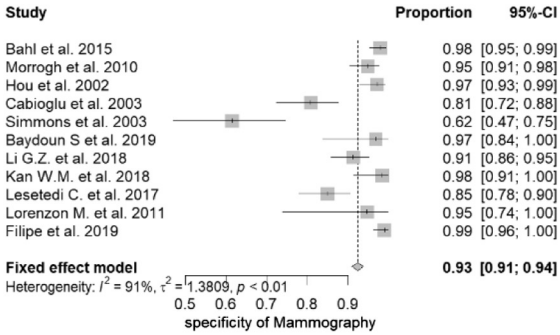
Supplemental Figure 1 **Usefulness of Ultrasound as a Modality to Detect Breast Cancer in Patients With Pathologic Nipple Discharge. Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, and Diagnostic Accuracy of Ultrasound**

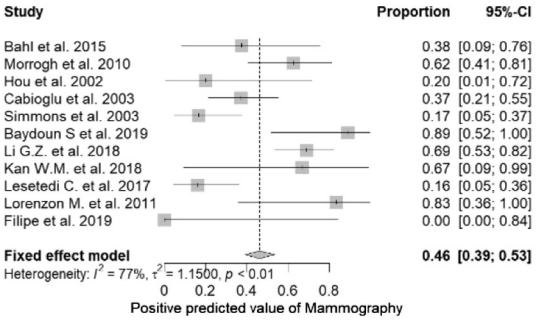


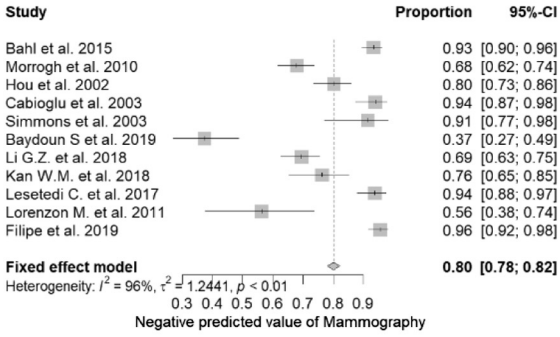
Abbreviation: CI = confidence interval.

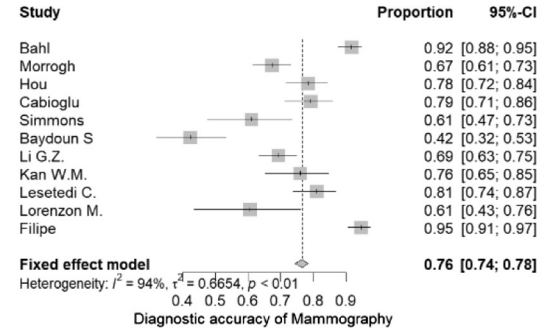
Supplemental Figure 2 **Usefulness of Mammogram as a Modality to Detect Breast Cancer in Patients With Pathologic Nipple Discharge. Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, and Diagnostic Accuracy of Mammogram**





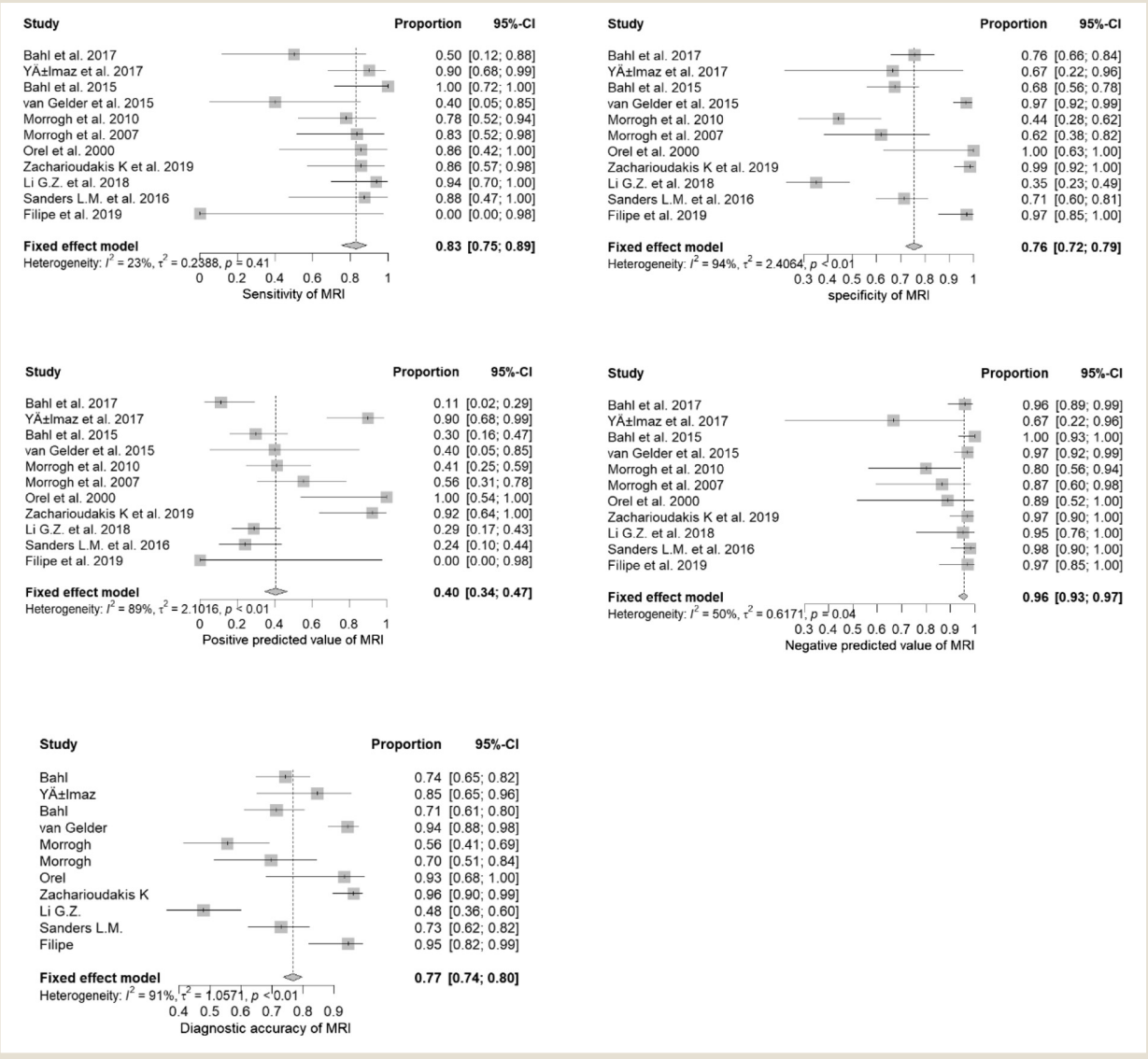






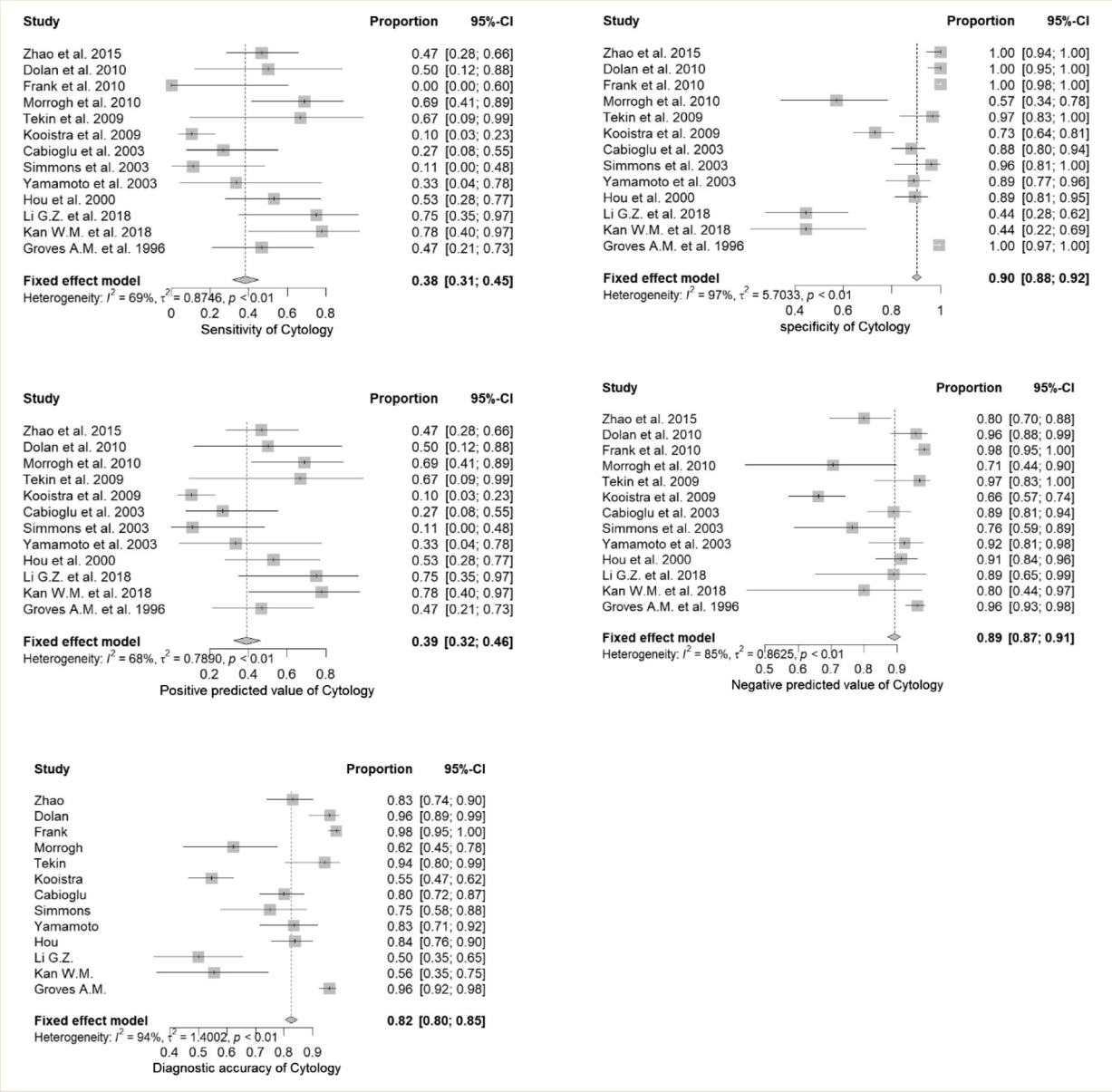
Abbreviation: CI = confidence interval.

Supplemental Figure 3 **Usefulness of MRI as a Modality to Detect Breast Cancer in Patients With Pathologic Nipple Discharge. Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, and Diagnostic Accuracy of MRI**



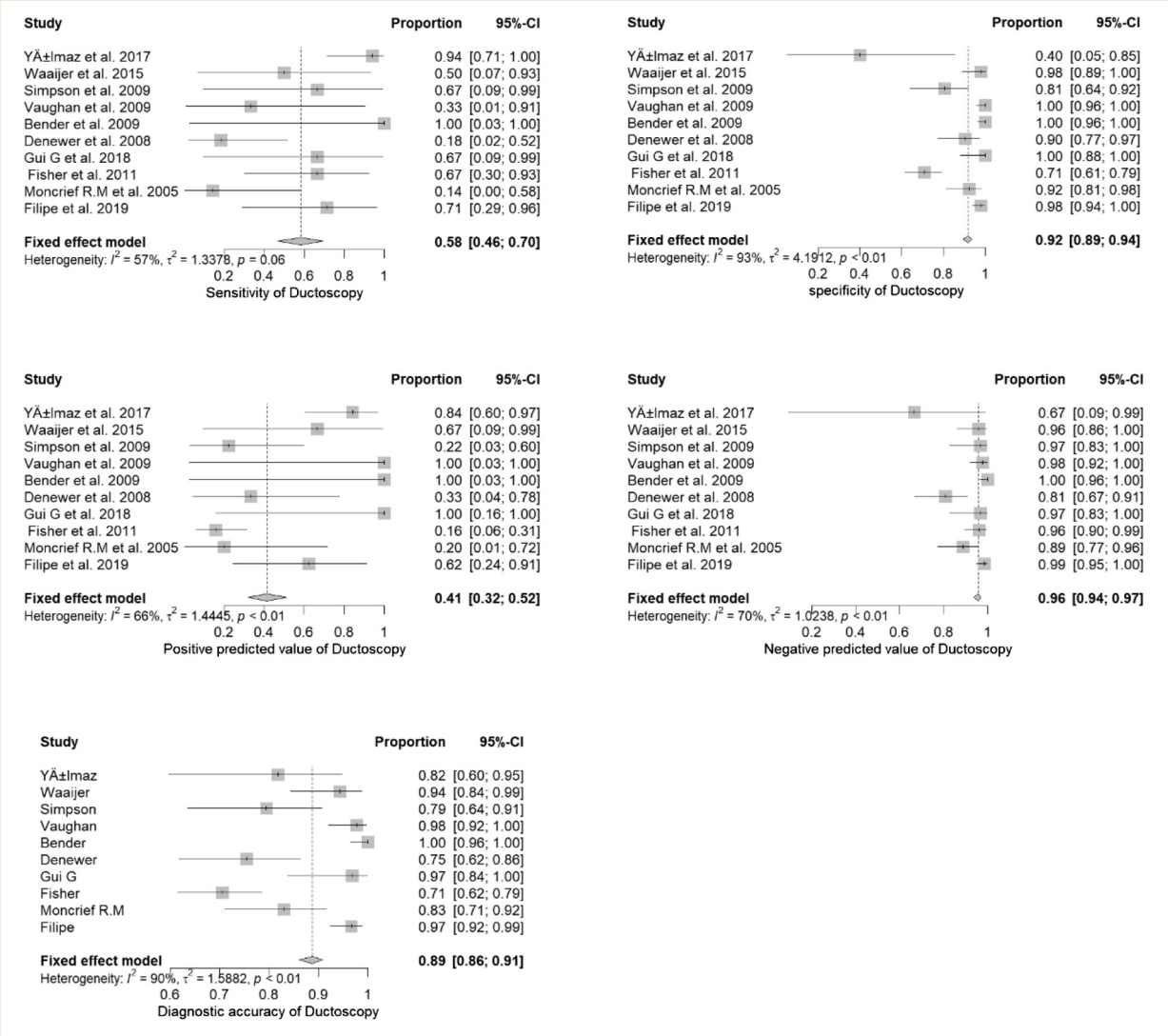
Abbreviations: CI = confidence interval; MRI = magnetic resonance imaging.

Supplemental Figure 4 **Usefulness of Cytology as a Modality to Detect Breast Cancer in Patients With Pathologic Nipple Discharge. Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, and Diagnostic Accuracy of Cytology**



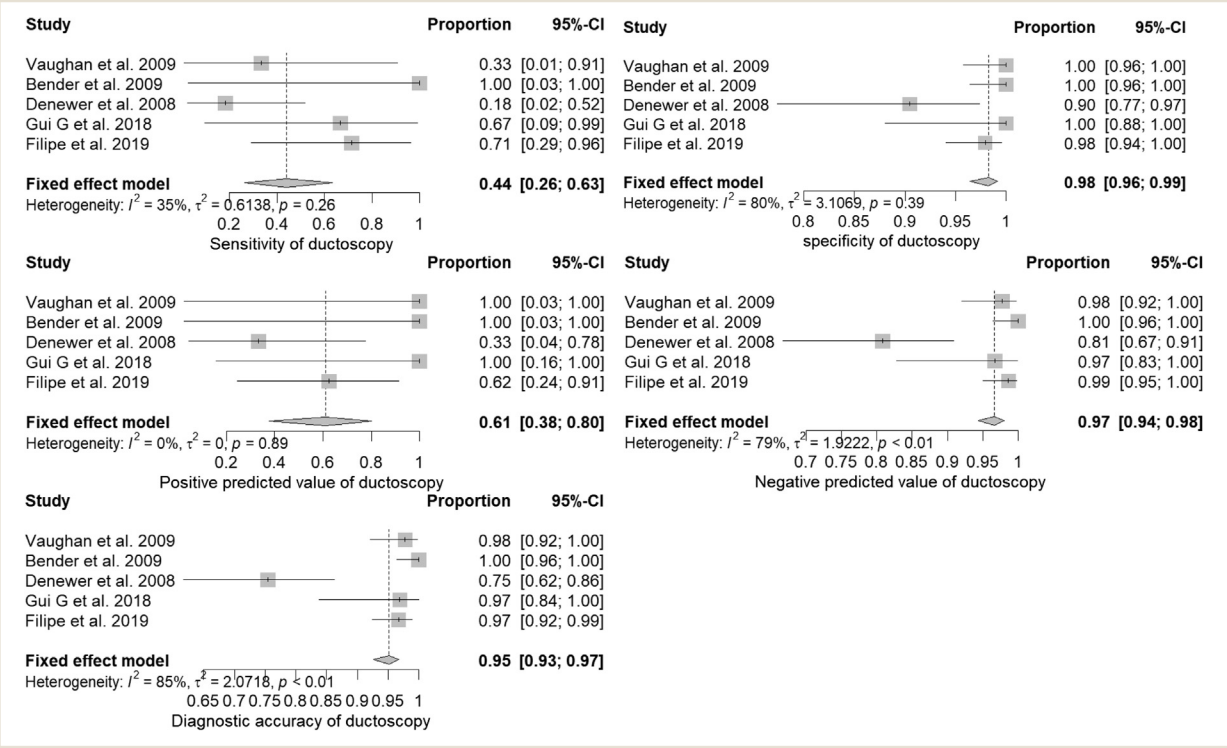
Abbreviation: CI = confidence interval.

Supplemental Figure 5 **Usefulness of Ductoscopy as a Modality to Detect Breast Cancer in Patients With Pathologic Nipple Discharge. Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, and Diagnostic Accuracy of Ductoscopy**



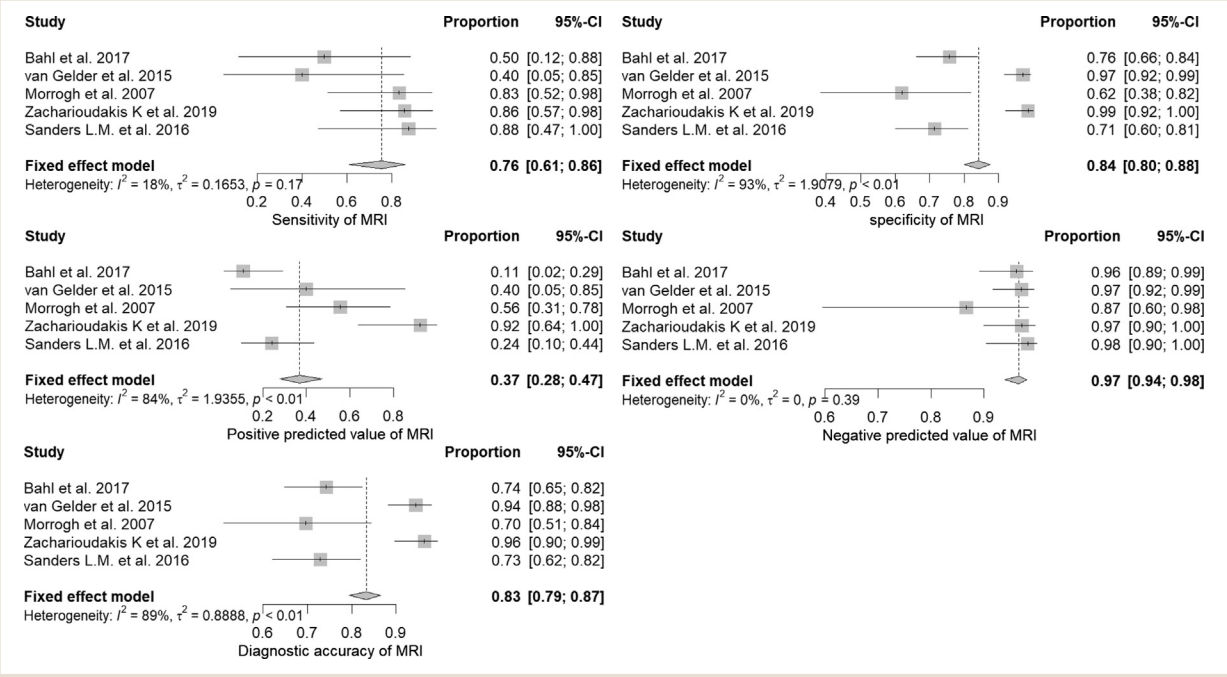
Abbreviation: CI = confidence interval.

Supplemental Figure 6 **Usefulness of Ductoscopy as a Modality to Detect Breast Cancer in Patients With Pathologic Nipple Discharge and Negative Mammogram/Ultrasound Results. Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, and Diagnostic Accuracy for Detection of Malignancy by Ductoscopy in Patients With Negative Mammogram/Ultrasound Results**



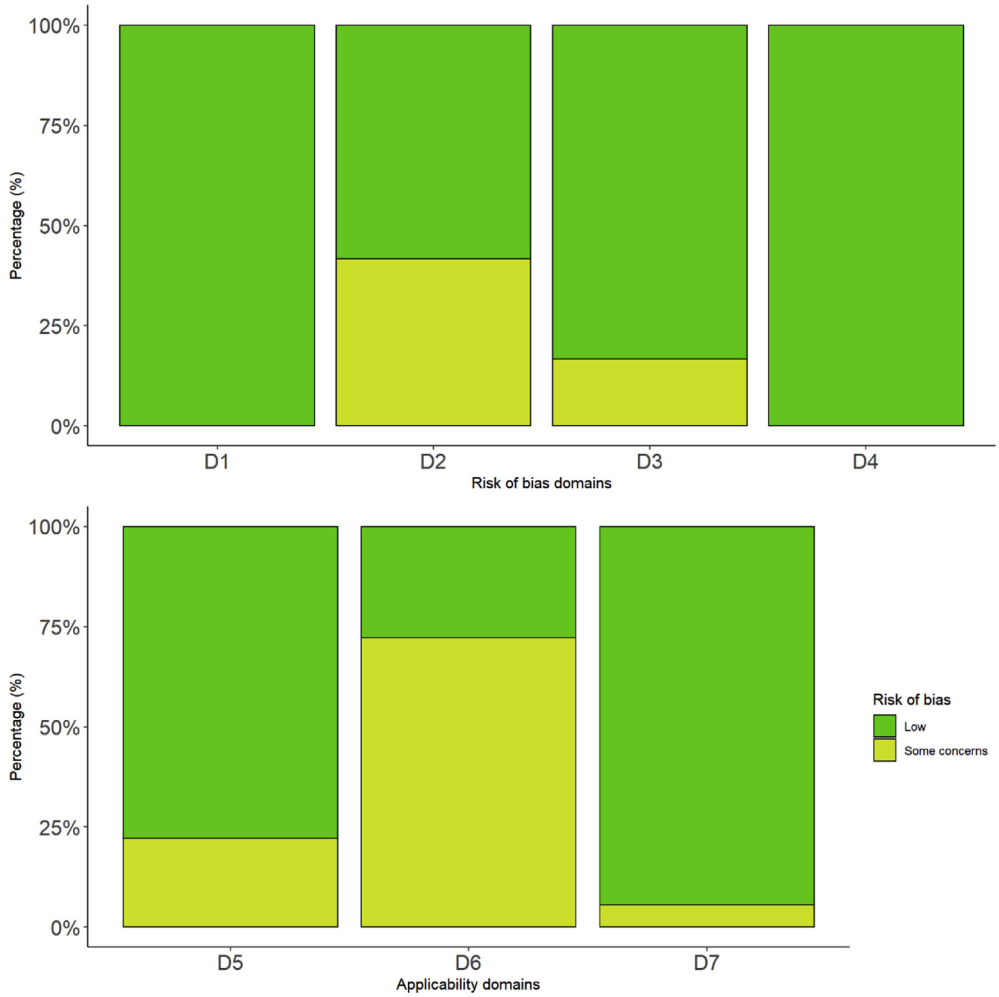
Abbreviation: CI = confidence interval.

Supplemental Figure 7 **Usefulness of Ultrasound as a Modality to Detect Breast Cancer in Patients With Pathologic Nipple Discharge. Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, and Diagnostic Accuracy for Detection Malignancy of MRI in Patients With Negative Mammogram/Ultrasound Results**



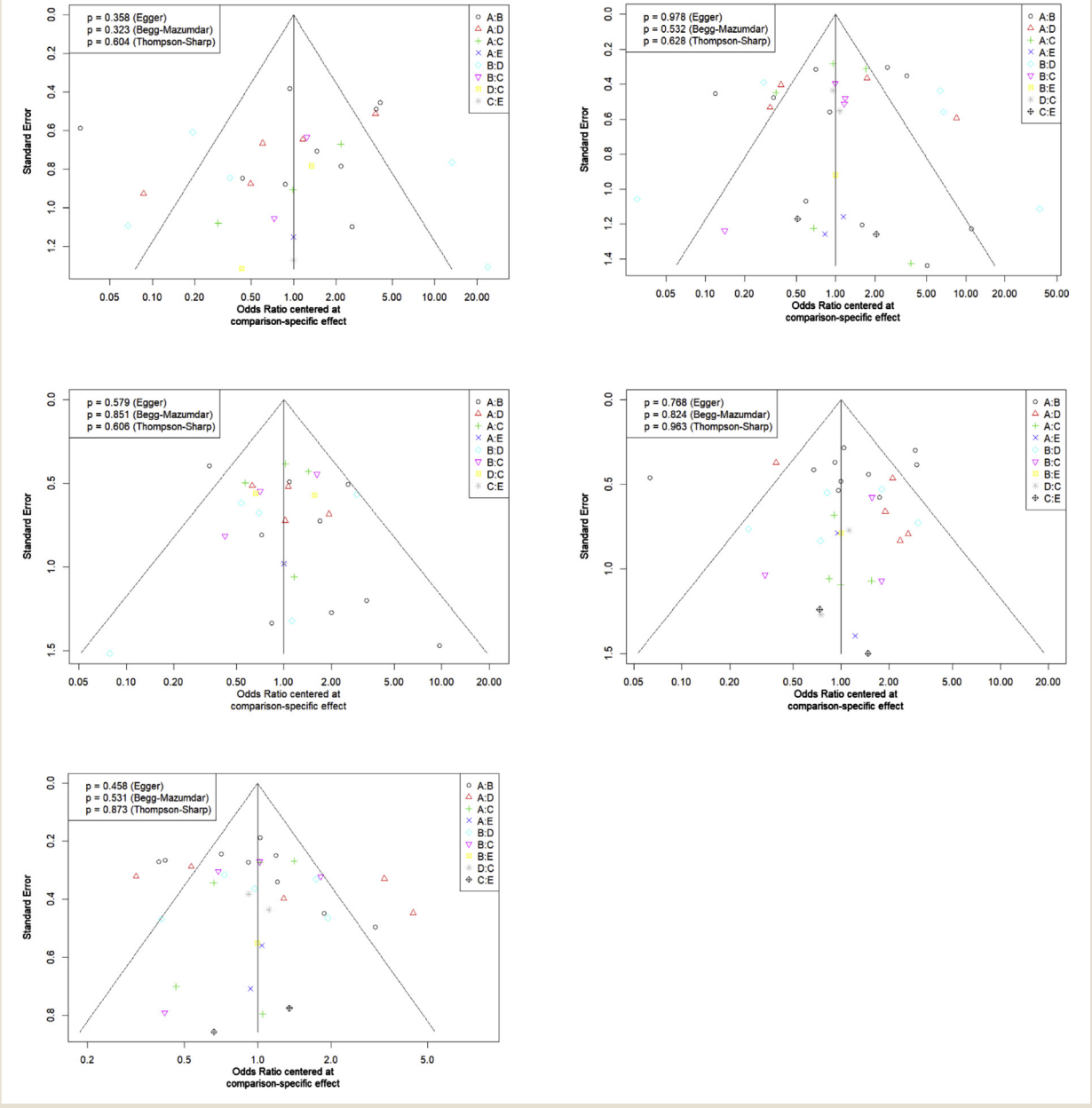
Abbreviations: CI = confidence interval; MRI = magnetic resonance imaging.

Supplemental Figure 8 Summary of Risks of Bias and Applicability Domains

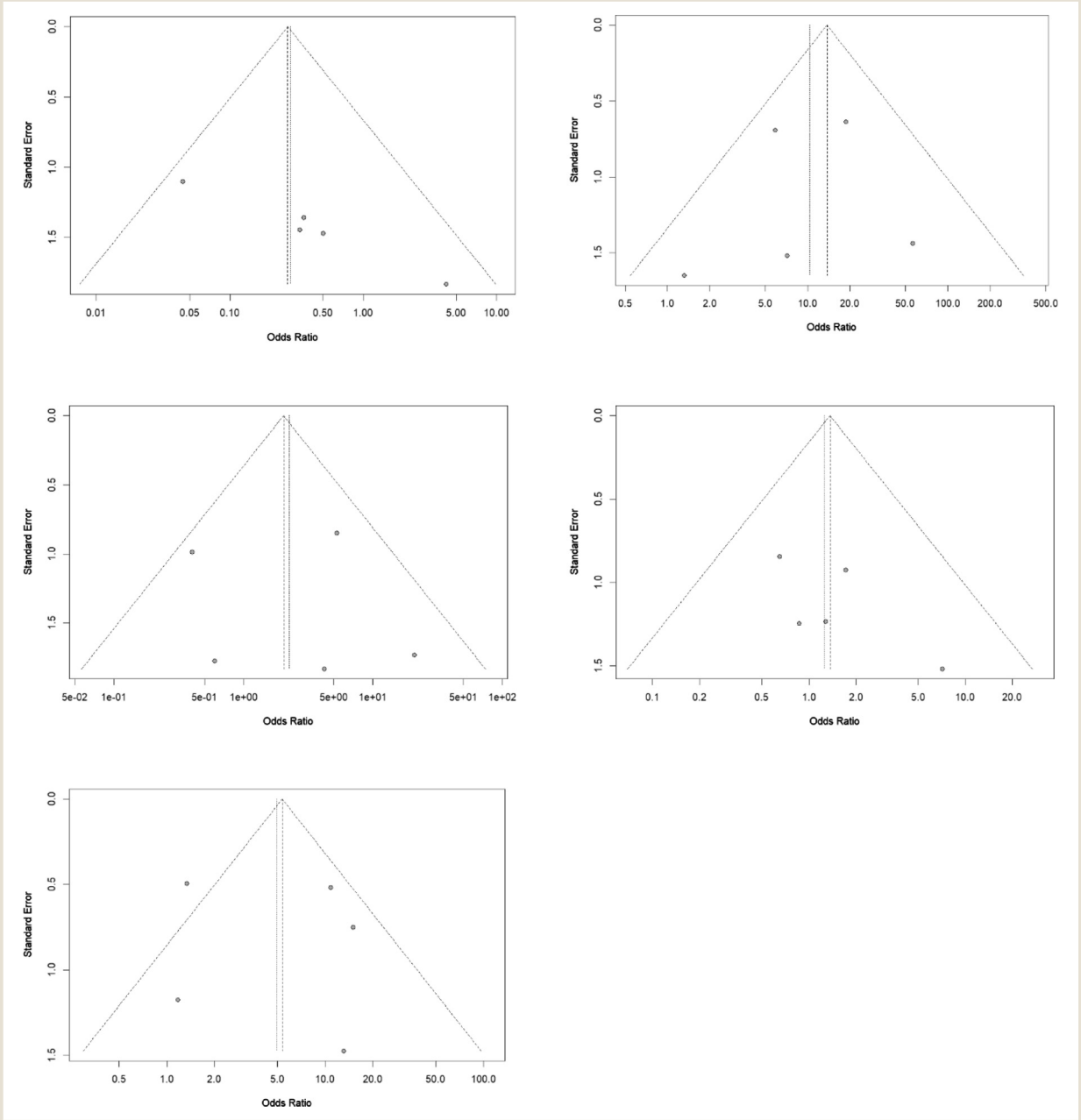


Abbreviations: D1 = patient selection; D2 = index test; D3 = reference standard; D4 = flow and timing; D5 = patient selection; D6 = index test; D7 = reference standard.

Supplemental Figure 9 **Funnel Plots of Potential Publication Bias Regarding Treatment of Pathologic Nipple Discharge. Funnel Plots Showing Potential Publication Bias for Sensitivity (Top Left), Specificity (Top Right), Positive Predictive Value (Middle Left), Negative Predictive Value (Middle Right), and Diagnostic Accuracy (Bottom Left) in a Meta-analysis of Various Approaches to Treat Pathologic Nipple Discharge. Modalities are as Follows: A = Ultrasound; B = Mammogram; C = Magnetic Resonance Imaging; D = Cytology; E = Ductoscopy**

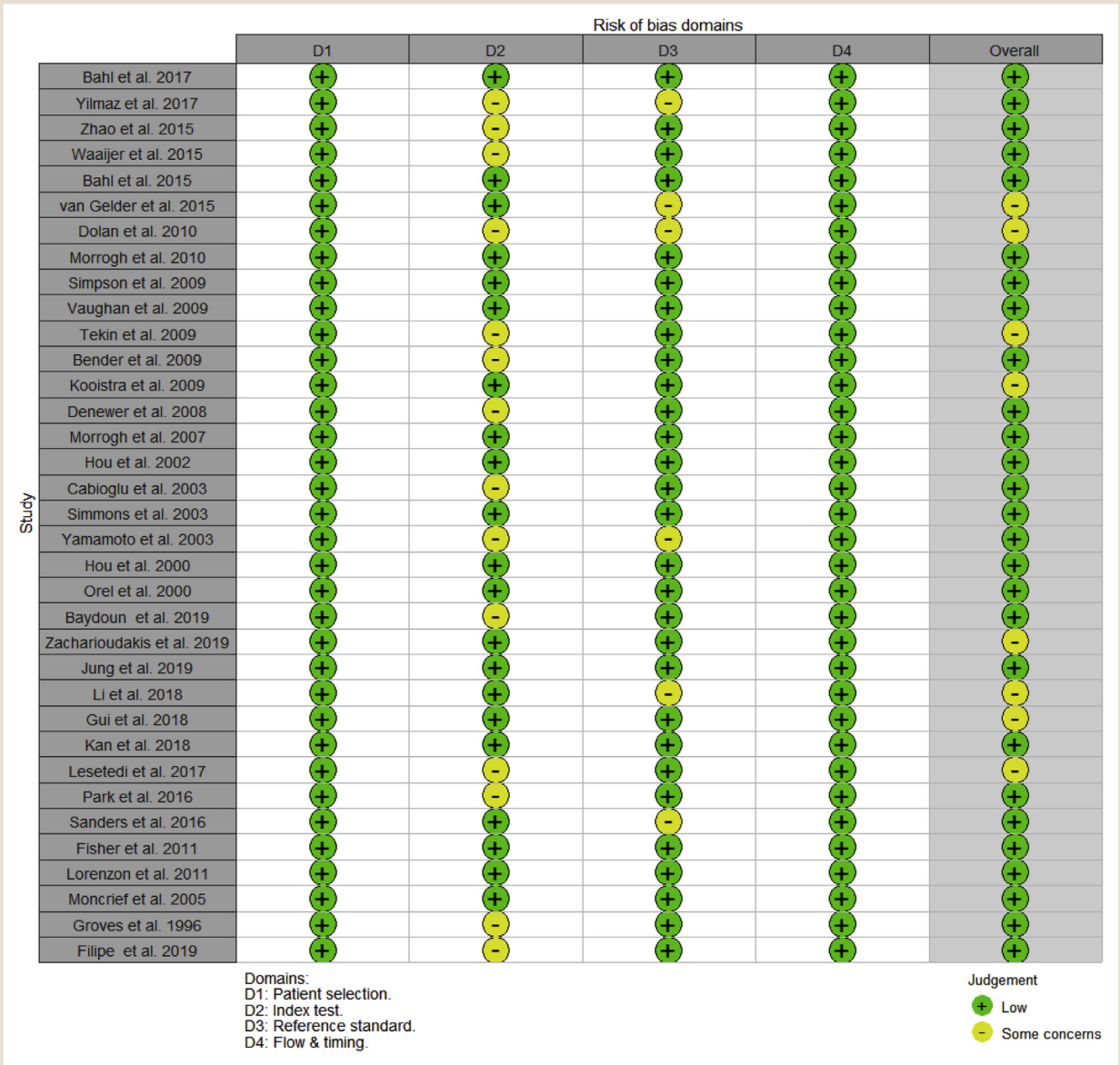


Supplemental Figure 10 **Funnel Plots of Potential Publication Bias Regarding Treatment of Pathologic Nipple Discharge. Funnel Plots Showing Potential Publication Bias for Sensitivity (Top Left), Specificity (Top Right), Positive Predictive Value (Middle Left), Negative Predictive Value (Middle Right), and Diagnostic Accuracy (Bottom Left) in a Subgroup Meta-analysis Comparing MRI to Ductoscopy in Patients With Pathologic Nipple Discharge and Negative Mammogram/Ultrasound Results**



Abbreviation: MRI = magnetic resonance imaging.

Supplemental Figure 11 Study Risk of Bias



Supplemental Table 1 Pairwise Meta-analysis and Network Meta-analysis Comparing Diagnostic Accuracy in 5 Different Diagnostic Methods

Pooled Diagnostic Accuracy

Modality	DA	2.5% CI	97.5% CI
]Ultrasound	0.6460	0.6232	0.6684
Mammography	0.7648	0.7440	0.7846
MRI	0.7684	0.7355	0.7990
Cytology	0.8248	0.8024	0.8456
Ductoscopy	0.8829	0.8560	0.9063

Pairwise Direct Meta-analysis

Comparison	No. of Studies	OR	2.5% CI	97.5% CI
A vs. B[a]	10	0.599	0.508	0.708
A vs. C	5	1.065	0.781	1.453
A vs. D[a]	5	0.418	0.314	0.555
A vs. E	2	0.645	0.273	1.523
B vs. C[a]	4	2.401	1.738	3.316
B vs. D	5	1.294	0.938	1.787
B vs. E	1	0.598	0.203	1.758
C vs. D	2	0.848	0.483	1.487
C vs. E	2	0.905	0.288	2.845
D vs. E	0			

Network Meta-analysis

Comparison	OR	2.5% CI	97.5% CI
A vs. B[a]	0.5927	0.5018	0.7000
A vs. C	1.1423	0.8475	1.5396
A vs. D[a]	0.4801	0.3638	0.6335
A vs. E	0.6044	0.2758	1.3245
B vs. C[a]	1.9274	1.4242	2.6084
B vs. D	0.8100	0.6092	1.0770
B vs. E	1.0199	0.4636	2.2435
C vs. D[a]	0.4203	0.2862	0.6172
C vs. E	0.5291	0.2347	1.1929
D vs. E	1.2590	0.5514	2.8748

Diagnostic modalties are as follows: A = ultrasound; B = mammogram; C = MRI; D = cytology; E = ductoscopy.
Abbreviations: CI = confidence interval; DA = diagnostic accuracy; MRI = magnetic resonance imaging; OR = odds ratio; PND = pathologic nipple discharge.
[a]Statistically significant.

Supplemental Table 2 Surface Under the Cumulative Ranking Curves (SUCRA) Values of 5 Diagnostic Methods to Detect Malignancy in Patients With Pathologic Nipple Discharge

Diagnostic Method	Sensitivity	Specificity	PPV	NPV	DA
Ultrasound	0.5234	0.5405	0.7066	0.9235	0.7717
Mammography	0.9992	0.0002	0.0192	0.8031	0.3517
MRI	0.1659	0.8915	0.4827	0.1674	0.9366
Cytology	0.6994	0.2917	0.5316	0.4099	0.0915
Ductoscopy	0.1121	0.7760	0.7599	0.1962	0.3485

Abbreviations: DA = diagnostic accuracy; MRI = magnetic resonance imaging; NPV = negative predictive value; PPV = positive predictive value.