

# Performance of Digital Breast Tomosynthesis, Synthetic Mammography, and Digital Mammography in Breast Cancer Screening: A Systematic Review and Meta-Analysis

Mostafa Alabousi, MD,<sup>1</sup> Akshay Wadera, MD,<sup>1</sup> Mohammed Kashif Al-Ghita,<sup>2</sup> Rayeh Kashef Al-Ghetaa, MPS,<sup>3</sup> Jean-Paul Salameh, MSc,<sup>4</sup> Alex Pozdnyakov, MD,<sup>5</sup> Nanxi Zha, MD,<sup>1</sup> Lucy Samoilov, MD,<sup>1</sup> Anahita Dehmoobad Sharifabadi, MD,<sup>1</sup> Behnam Sadeghirad, PharmD, MPH, PhD,<sup>6</sup> Vivianne Freitas, MD, FRCPC<sup>7</sup> Matthew DF McInnes, PhD, MD, FRCPC,<sup>8</sup> Abdullah Alabousi, MD, FRCPC<sup>9,\*</sup>

<sup>1</sup>Department of Radiology, McMaster University, Hamilton, ON, Canada; <sup>2</sup>Faculty of Biomedical Sciences, Western University, London, ON, Canada; <sup>3</sup>Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada; <sup>4</sup>Department of Radiology, University of Ottawa, Ottawa, ON, Canada; <sup>5</sup>Faculty of Medicine, McMaster University, Hamilton, ON, Canada; <sup>6</sup>Department of Health Research Methods, Evidence, and Impact (HEI), The Michael G. DeGroot Institute for Pain Research and Care, McMaster University, Hamilton, ON, Canada; <sup>7</sup>Joint Department of Medical Imaging, University of Toronto, Toronto, Ontario, Canada; <sup>8</sup>Department of Radiology and Epidemiology, University of Ottawa, Ottawa Hospital Research Institute, Clinical Epidemiology Program, Ottawa, ON, Canada; and <sup>9</sup>Department of Radiology, McMaster University, St Joseph's Healthcare Hamilton, Hamilton, ON, Canada

\*Correspondence to: Abdullah Alabousi, MD, FRCPC, Department of Radiology, McMaster University, St Joseph's Healthcare Hamilton, 50 Charlton Ave E, Hamilton, ON L8N 4A6, Canada (e-mail: abduallah.alabousi@medportal.ca).

## Abstract

**Background:** Our objective was to perform a systematic review and meta-analysis comparing the breast cancer detection rate (CDR), invasive CDR, recall rate, and positive predictive value 1 (PPV1) of digital mammography (DM) alone, combined digital breast tomosynthesis (DBT) and DM, combined DBT and synthetic 2-dimensional mammography (S2D), and DBT alone.

**Methods:** MEDLINE and Embase were searched until April 2020 to identify comparative design studies reporting on patients undergoing routine breast cancer screening. Random effects model proportional meta-analyses estimated CDR, invasive CDR, recall rate, and PPV1. Meta-regression modeling was used to compare imaging modalities. All statistical tests were 2-sided. **Results:** Forty-two studies reporting on 2 606 296 patients (13 003 breast cancer cases) were included. CDR was highest in combined DBT and DM (6.36 per 1000 screened, 95% confidence interval [CI] = 5.62 to 7.14,  $P < .001$ ), and combined DBT and S2D (7.40 per 1000 screened, 95% CI = 6.49 to 8.37,  $P < .001$ ) compared with DM alone (4.68 per 1000 screened, 95% CI = 4.28 to 5.11). Invasive CDR was highest in combined DBT and DM (4.53 per 1000 screened, 95% CI = 3.97 to 5.12,  $P = .003$ ) and combined DBT and S2D (5.68 per 1000 screened, 95% CI = 4.43 to 7.09,  $P < .001$ ) compared with DM alone (3.42 per 1000 screened, 95% CI = 3.02 to 3.83). Recall rate was lowest in combined DBT and S2D (42.3 per 1000 screened, 95% CI = 37.4 to 60.4,  $P < .001$ ). PPV1 was highest in combined DBT and DM (10.0%, 95% CI = 8.0% to 12.0%,  $P = .004$ ), and combined DBT and S2D (16.0%, 95% CI = 10.0% to 23.0%,  $P < .001$ ), whereas no difference was detected for DBT alone (7.0%, 95% CI = 6.0% to 8.0%,  $P = .75$ ) compared with DM alone (7.0%, 95% CI = 5.0% to 8.0%). **Conclusions:** Our findings provide evidence on key performance metrics for DM, DBT alone, combined DBT and DM, and combined DBT and S2D, which may inform optimal application of these modalities for breast cancer screening.

Digital breast tomosynthesis (DBT) has been attracting attention for application in breast cancer screening to improve accuracy in detection and diagnosis compared with digital mammography (DM). Some centers in North America and Europe have already transitioned to adopt DBT as the standard of care for

breast cancer screening (1-3). An increasing body of research shows that DBT alone, and combined DBT and DM are more accurate than DM alone in detecting breast cancer in average-risk women (1,4-7). The utilization of DBT has also been associated with lower recall rates and faster diagnosis due to shorter time

to biopsy (1,5-7). Multiple studies have compared the diagnostic performance of DBT and DM, including a systematic review and meta-analysis of almost 500 000 patients, which found a higher sensitivity of DBT with or without additional DM compared with DM alone (8).

Furthermore, although the evidence suggests that the diagnostic performance of DBT is better than DM, there is still a question of the need for a 2-dimensional (2D) image to allow for comparison with previous DMs as well as assessment of calcifications (1,4-7). While the radiation dose of DBT may be comparable with DM, the combination of DBT and DM as a single examination effectively doubles the radiation dose received by the patient (4,9-11). Synthetic 2D (S2D)-reconstructed images from DBT examinations have also been assessed as a potential replacement for DM as a strategy to reduce the radiation dose associated with combined DBT and DM (4,9-11). However, a lack of consensus has persisted in breast-imaging guidelines regarding the optimal utilization of DBT and S2D (3,12-15).

The assessment of the utility of breast tomosynthesis in the breast cancer-screening population, specifically with the comparison of DM vs S2D images in addition to DBT, warrants further examination. Our primary objective was to perform a systematic review and meta-analysis to assess the breast cancer detection rate (CDR) of DM alone; combined DBT and DM; combined DBT and S2DM; and DBT alone in women undergoing breast cancer screening who are unselected for risk. Our secondary objectives included assessing the invasive breast CDR, recall rate, and positive predictive value (PPV1).

## Methods

A protocol for this study was registered a priori (CRD42020180758) in the International Prospective Register of Systematic Reviews. A systematic review and meta-analysis were performed based on contemporary methodological guidance for imaging test accuracy systematic reviews (16-18). The study was reported as per the Preferred Reporting Items for Systematic reviews and Meta-Analysis Diagnostic Test Accuracy guidelines (18,19).

### Literature Search

A systematic literature search of the electronic databases MEDLINE and EMBASE was conducted to identify all relevant studies. The search was limited to English language studies published from inception of the databases to April 15, 2020. Details of the search strategy are included in the [Supplementary Methods](#) (available online).

### Eligibility Criteria

Inclusion criteria were defined as the following: the full text was available in English; women undergoing routine breast cancer screening; a comparative study design, where either similar populations underwent at least 2 imaging examinations (DM alone, combined DBT and DM, combined DBT and S2D, DBT alone) or patients were randomly assigned to the imaging tests being compared (20); the reference standard used for confirmed cases of breast cancer was histopathology; and the results reported sufficient data to calculate the CDR, that is, reporting of the total number of patients screened and the number of cancer cases detected. Exclusion criteria were defined as the following: the study investigated patients at high risk for developing

breast cancer (21); the study only assessed patients undergoing diagnostic studies, rather than screening studies; or the study employed a “single arm” design assessing only 1 imaging modality. For any duplicate studies, the study with the larger sample size was included.

### Study Selection

Results of the literature search were imported into a reference manager software (Reference Manager 11; Thomson Reuters, Toronto, ON, Canada) for title and abstract review (phase I) completed by multiple investigators (M.A., M.K.A., N.Z., J.P.S., L.S., and A.P.) independently in duplicate. A pilot screen was conducted for the first 50 studies in duplicate to improve familiarity and consistency for the remaining ones. The full texts of potentially eligible articles were retrieved and assessed for inclusion by multiple investigators (M.A., A.W., M.K.A., N.Z., J.P.S., L.S., and A.P.) independently. Any discrepancies were resolved by consensus or adjudication by a third reviewer (A.A.), if needed.

### Data Extraction

Data extraction was performed independently in duplicate on the included studies by multiple investigators (M.A., A.W., and M.K.A.). To improve familiarity and consistency among the investigators, a pilot phase was performed for the first 3 studies by all data extractors. Discrepancies were resolved by consensus.

The following data were extracted into a spreadsheet program (Microsoft Excel 2016; Microsoft, Redmond, United States) using predefined forms: first author, title of study, publication year, country of corresponding author, journal of publication, study design, patient demographics (sample size, indication for imaging, mean age of population, setting of patient sample, and reasons for patient exclusion), Breast Imaging Reporting and Data System classification, imaging protocol (scanner used, number of interpreters, independent or consensus reporting), reference standard used (histopathology and/or imaging follow-up), the total number of screening examinations and/or patients included, the number of recalled patients, the number of breast cancer cases detected, and the number of invasive breast cancer cases detected.

### Quality Assessment

Quality assessment of all included studies was conducted using the revised tool for the Quality Assessment of Diagnostic Accuracy Studies-2 (22). Multiple investigators (M.A., A.W., and M.K.A.) assessed all articles independently in duplicate for the following criteria: patient selection, index test, reference standard, flow, and timing (22). The following criteria were defined as being at high risk of bias: 1) random or consecutive patient selection was not used (patient selection); 2) radiologists reviewing the index test were not blinded to previous clinical or imaging data (index test); 3) the method by which patients are assigned to a specific imaging test may have introduced bias, for example, if the patients were able to choose (index test); 4) the reference standard, histopathology, was not offered equally to all patients who were considered to need biopsy to exclude an underlying cancer (reference standard); and 5) flow and timing for at least 2 index tests whereby they are performed more than 3 months apart (flow and timing). Discrepancies were resolved by consensus. A pilot was performed by all reviewers in

duplicate for the initial 3 studies to improve familiarity and consistency.

## Outcomes

The primary outcome was the CDR of DM alone, combined DBT and DM, combined DBT and S2D, and DBT alone. The CDR was defined as the number of cancer cases detected in a population (based on histopathology) divided by the total number of screening examinations performed, reported per 1000 examinations. Multiple secondary outcomes were included in the analysis: invasive CDR, recall rate, and PPV1. The invasive CDR was defined as the number of invasive cancer cases detected (based on histopathology) divided by the total number of screening examinations performed, reported per 1000 examinations. The recall rate was defined as the number of “recalled” screening examinations for further imaging or testing divided by the total number of screening examinations performed, reported per 1000 examinations. The PPV1 was defined as the total number of cancer cases detected (based on histopathology) divided by the total number of recalled screening examinations. The primary and secondary outcomes were compared across each of the imaging modalities included.

## Data Synthesis and Statistical Analysis

Proportional meta-analyses were performed to determine estimates of the mean CDR, invasive CDR, recall rate, and PPV1 with 95% confidence intervals (CIs) for DM alone, combined DBT and DM, combined DBT and S2D, and DBT alone using a random effects model with arcsine transformation (16,23,24). Forest plots were created using the estimated model parameters. Comparative meta-regression models were performed to compare CDR, invasive CDR, recall rate, and PPV1 for DM alone, combined DBT and DM, combined DBT and S2D, and DBT alone. Beta coefficients with 95% CIs, standard errors, and corresponding *P* values were calculated for each meta-regression analysis. For these models, DM alone was treated as the baseline comparator group. A post hoc head-to-head meta-regression comparison of combined DBT and DM vs combined DBT and S2D was performed for CDR, invasive CDR, recall rate, and PPV1. All statistical tests were 2-sided. A *P* value less than .05 was considered statistically significant. As per contemporary guidance for diagnostic accuracy systematic reviews, publication bias was not assessed (18,19). Analysis was performed using the “metaprop” and “meta” packages in STATA version 11.2 (College Station, TX) and R version 3.5.1 (Vienna, Austria) (25). Heterogeneity was assessed using the  $I^2$  value, with values greater than 50% considered at risk for substantial variability (26,27).

## Results

### Study Demographics and Risk of Bias

A study flow diagram is shown in Figure 1. An initial 1591 studies underwent title and abstract screening, of which 167 studies were retrieved for full-text review. In all, 42 articles reporting on 32 studies and 2 606 296 screened patients (13 003 breast cancer cases) met the inclusion criteria (1,4-7,28-63); 29 studies reported on DM alone (1 279 056 patients screened; 5656 cases of breast cancer), 16 studies reported on combined DBT and DM (494 145 patients screened; 2885 cases of breast cancer), 10

studies reported on combined DBT and S2D (226 126 patients screened; 1798 cases of breast cancer), and 11 studies reported on DBT alone (606 969 patients screened; 2664 cases of breast cancer). Table 1 provides a summary of the included studies.

Supplementary Table 1 (available online) provides a risk of bias and applicability summary of the included studies. All but one study were at high risk of bias. The main sources contributing to a high or unclear risk of bias related to the index test due to a lack of blinding of the radiologists to previous imaging or clinical information, or a lack of reporting of the blinding status. Other sources of bias included the following: limited reporting of sampling methods (ie, potential volunteer bias), index test assignment for patients (eg, allowing patients to choose which test they would like to be assigned to), and inclusion of cohorts of patients from different years during which testing was performed (in some cases retrospective cohort for a control group used in a prospective study).

## Data Synthesis and Pooling

Pooled estimates of the mean CDR, invasive CDR, recall rate, and PPV1 for DM alone, combined DBT and DM, combined DBT and S2D, and DBT alone are listed in Table 2. A summary of the meta-regression results for CDR, invasive CDR, recall rate, and PPV1 is shown in Table 3. Forest plots and pooled estimates of CDR and invasive CDR are illustrated in Figures 2 and 3, respectively, for DM alone, combined DBT and DM, combined DBT and S2D, and DBT alone. Forest plots and pooled estimates of the recall rate per 1000 patients screened are illustrated in Figure 4 for DM alone, combined DBT and DM, combined DBT and S2D, and DBT alone. Forest plots and pooled estimates of the PPV1 are illustrated in Figure 5 for DM alone, combined DBT and DM, combined DBT and S2D, and DBT alone.

Estimates of the mean CDR and invasive CDR for each imaging modality are shown in Table 2, and statistical comparisons of CDR and invasive CDR between each imaging modality via meta-regression are shown in Table 3. The CDR for both combined DBT and DM (6.36 per 1000 screened, 95% CI = 5.62 to 7.14,  $P < .001$ ) as well as combined DBT and S2D (7.40 per 1000 screened, 95% CI = 6.49 to 8.37,  $P < .001$ ) was higher than DM alone (4.68 per 1000 screened, 95% CI = 4.28 to 5.11). Invasive CDR for combined DBT and DM (4.53 per 1000 screened, 95% CI = 3.97 to 5.12,  $P = .003$ ) as well as combined DBT and S2D (5.68 per 1000 screened, 95% CI = 4.43 to 7.09,  $P < .001$ ) was statistically significantly higher than DM alone (3.42 per 1000 screened, 95% CI = 3.02 to 3.83). DBT alone was not statistically significantly different for CDR (5.20 per 1000 screened, 95% CI = 4.59 to 5.85,  $P = .29$ ) or invasive CDR (3.68 per 1000 screened, 95% CI = 3.06 to 4.35,  $P = .59$ ) compared with DM alone.

Estimates of the mean recall rate and PPV1 for each imaging modality are shown in Table 2, and statistical comparisons of recall rate and PPV1 between each imaging modality via meta-regression are shown in Table 3. Combined DBT and S2D had a lower recall rate (42.3 per 1000 screened, 95% CI = 37.4 to 60.4,  $P < .001$ ) compared with DM alone (78.8 per 1000 screened, 95% CI = 67.4 to 91.1). Meanwhile, no difference was identified in the recall rate of combined DBT and DM (64.6 per 1000 screened, 95% CI = 55.5 to 74.4,  $P = .09$ ) or DBT alone (82.4 per 1000 screened, 95% CI = 66.5 to 100.0,  $P = .67$ ) compared with DM alone. PPV1 was higher in both combined DBT and DM (10.0%, 95% CI = 8.0% to 12.0%,  $P = .004$ ) as well as combined DBT and S2D (16.0%, 95% CI = 10.0% to 23.0%,  $P < .001$ ) compared with DM alone (7.0%, 95% CI = 5.0% to 8.0%). No difference in PPV1 was

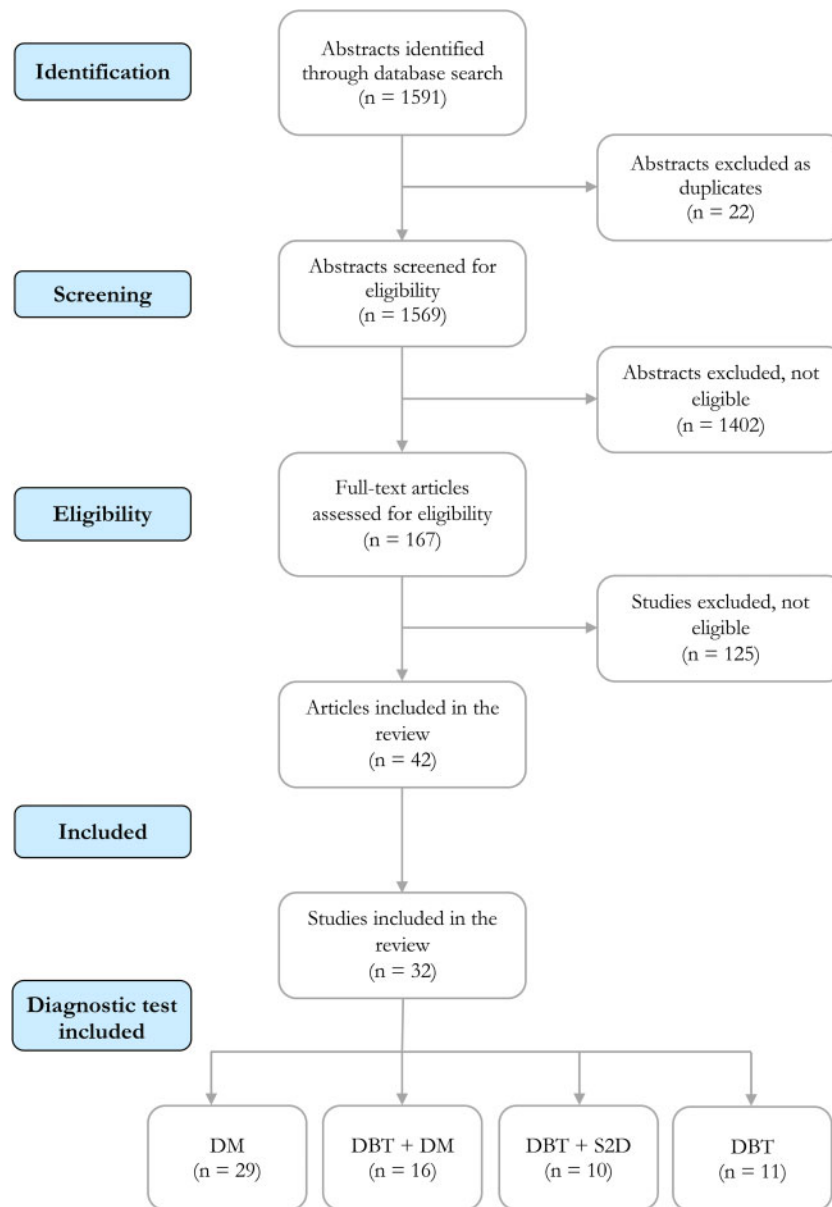


Figure 1. Study flow diagram. DBT = digital breast tomosynthesis; DM = digital mammography; S2D = synthetic 2D image.

identified between DBT alone (7.0%, 95% CI = 6.0% to 8.0%) and DM alone ( $P = .75$ ).

Based on post hoc meta-regression analysis, a head-to-head comparison of combined DBT and DM as well as combined DBT and S2D found no statistically significant difference in CDR ( $P = .44$ ) and invasive CDR ( $P = .12$ ). Combined DBT and S2D resulted in statistically significantly fewer recalled patients compared with combined DBT and DM ( $P = .006$ ). PPV1 was statistically significantly higher in combined DBT and S2D compared with combined DBT and DM ( $P = .047$ ).

Quantification of heterogeneity using the  $I^2$  statistic is shown in Table 2. All imaging modalities had a risk for substantial heterogeneity for CDR, invasive CDR, recall rate, and PPV1.  $I^2$  for CDR ranged from 84.6% to 89.4% ( $P < .001$ ) for each of the modalities, and it ranged from 73.7% to 91.3% ( $P < .001$ ) for invasive CDR.  $I^2$  for recall rate was greater than 99.0% ( $P < .001$ ) for all

imaging modalities. The  $I^2$  statistic for PPV1 ranged from 95.0% to 98.3% ( $P < .001$ ).

## Discussion

This systematic review and meta-analysis assessed the performance of DM alone, combined DBT and DM, combined DBT and S2DM, and DBT alone for breast cancer screening in a total of 2 606 296 screened patients (13 003 breast cancer cases). To date, our study includes the largest sample size for the evaluation of mammography and tomosynthesis in the setting of breast cancer screening. Our findings indicated that using combined DBT and DM or combined DBT and S2D resulted in statistically and clinically significant higher CDRs, invasive CDRs, and PPV1s compared with DM alone. Furthermore, we found that the

Table 1. Characteristics of included studies

Study	Modality	Region	Design	Center	Analysis level	BI-RADS used?	BI-RADS positive test results
Alsheik et al., 2019 (1)	DM, DBT	USA	Retrospective	Multi	Patient	Yes	≥3
Ambinder et al., 2018 (28)	DBT + S2D, DBT + DM	USA	Retrospective	Multi	Patient	No	NR
Aujero et al., 2017 (38)	DM, DBT + S2D	USA	Retrospective	Single	Lesion	Yes	0
Bahl et al., 2018 (48)	DM, DBT	USA	Retrospective	Single	Examination	Yes	≥3
Bahl et al., 2019 and 2020 (58,60)	DM, DBT	USA	Retrospective	Single	Examination	Yes	≥3
Bernardi et al., 2020 (62)	DM, DBT + S2D	Italy	Prospective	Multi	Examination	No	NR
Cochon et al., 2020 (33)	DM, DBT + DM, DBT	USA	Retrospective	Multi	Examination	Yes	0
Conant et al., 2020 (35)	DM, DBT	USA	Retrospective	Single	Examination	Yes	≥3
Dang et al., 2020 (36)	DM, DBT	USA	Retrospective	Multi	Examination	Yes	0
Destounis et al., 2014 (37)	DM, DBT + DM	USA	Retrospective	Single	Patient	Yes	≥4
Freer et al., 2017 (39)	DM, DBT + S2D, DBT + DM	USA	Retrospective	Single	Patient	Yes	NR
Friedewald et al., 2014 (6)	DM, DBT + DM	USA	Retrospective	Multi	Examination	No	NR
Fuji et al., 2019 (40)	DM, DBT + DM	USA	Retrospective	Multi	Examination	Yes	≥4
Geiss et al., 2017 (41)	DM, DBT	USA	Retrospective	Multi	Examination	Yes	≥4
Greenburg et al., 2014 (42)	DM, DBT	USA	Retrospective	Multi	Patient	Yes	≥4
Hofvind et al., 2018 (44), Hovda et al., 2020 (46)	DM, DBT + S2D	Norway	Prospective	Multi	Patient	No	NR
Hofvind et al., 2019 (43)	DM, DBT + S2D	Norway	Prospective	Single	Patient	No	NR
Lourenco et al., 2014 (47)	DM, DBT	USA	Retrospective	Single	Patient	Yes	≥4
McCarthy et al., 2014 (49)	DM, DBT	USA	Retrospective	Single	Patient	Yes	≥4
McDonald et al., 2015 <sup>a</sup> (50)	DM, DBT + DM	USA	Retrospective	Single	Patient	Yes	≥4
Oslo Tomosynthesis Trial (51,56,57)	DM, DBT + S2D, DBT + DM	Norway	Prospective	Single	Examination	No	NR
Pan et al., 2018 (52)	DM, DBT + DM	China	Retrospective	Single	Patient	Yes	≥4
Pattacini et al., 2018 (7)	DM, DBT + DM, DBT	Italy	Prospective	Single	Patient	No	NR
PROSPR Trial (5,34)	DM, DBT + DM	Italy	Retrospective	Multi	Examination	Yes	≥3
Rose et al., 2013 (53)	DM, DBT + DM	USA	Retrospective	Single	Examination	Yes	0
Rose et al., 2014 (54)	DM, DBT + DM	USA	Retrospective	Single	Patient	Yes	0
Sharpe et al., 2015 (55)	DM, DBT + DM	USA	Prospective	Single	Examination	No	NR
STORM 1 Trial (29,32,45,61)	DM, DBT + DM	Italy	Prospective	Multi	Patient	Yes	NR
STORM 2 Trial (4,63)	DM, DBT + S2D, DBT + DM	Italy	Prospective	Multi	Patient	Yes	NR
VERONA Trial (30,31)	DM, DBT + S2D	Italy	Prospective	Single	Examination	Yes	≥3
Zuckerman et al., 2016 <sup>a</sup> (59)	DBT + S2D, DBT + DM	USA	Retrospective	Single	Patient	No	NR

<sup>a</sup>Studies that share the same patient population for the DBT + DM group. BI-RADS = Breast Imaging Reporting and Data System; DBT = digital breast tomosynthesis; DM = digital mammography; S2D = synthetic 2-dimensional image.

Table 2. Summary of meta-analyses of proportions for breast cancer screening populations<sup>a</sup>

Modality	CDR, per 1000 screened (95% CI) [I <sup>2</sup> , %]	iCDR, per 1000 screened (95% CI) [I <sup>2</sup> , %]	Recall rate, per 1000 screened (95% CI) [I <sup>2</sup> , %]	PPV1, % (95% CI) [I <sup>2</sup> , %]
DM	4.68 (4.28 to 5.11) [89.4]	3.42 (3.02 to 3.83) [89.9]	78.8 (67.4 to 91.1) [99.8]	7.0 (5.0 to 8.0) [97.9]
DBT + DM	6.36 (5.62 to 7.14) [87.6]	4.53 (3.97 to 5.12) [73.7]	64.6 (55.5 to 74.4) [99.4]	10.0 (8.0 to 12.0) [96.8]
DBT + S2D	7.40 (6.49 to 8.37) [84.6]	5.68 (4.43 to 7.09) [91.3]	42.3 (37.4 to 60.4) [99.3]	16.0 (10.0 to 23.0) [98.3]
DBT	5.20 (4.59 to 5.85) [86.6]	3.68 (3.06 to 4.35) [84.9]	82.4 (66.5 to 100.0) [99.8]	7.0 (6.0 to 8.0) [95.0]

<sup>a</sup>CI = confidence interval; CDR = cancer detection rate; DBT = digital breast tomosynthesis; DM = digital mammography; iCDR = invasive cancer detection rate; PPV1 = positive predictive value; S2D = synthetic 2D image.

combination of DBT and S2D also reduced the recall rate for additional imaging and biopsies. Meanwhile, the utilization of DBT alone provided no additional benefit in breast cancer screening compared with DM alone with regards to the CDR, invasive CDR, recall, and PPV1.

We have expanded on the work of Marinovich et al. (64), which compared the incremental breast CDR and recall rate for DBT alone vs DM alone in 17 studies and 1 009 790 patients, because we included multiple additional studies, diagnostic test groups, and outcomes. Furthermore, whereas Marinovich et al. (64) reported that DBT is superior to DM in terms of its CDR and

recall rates, our findings indicated that the CDRs and recall rates improved only when DM or S2D was added to DBT compared with DM alone. We also found that the improved rates of breast cancer detection were not limited to in situ cases only, with these findings persisting for invasive breast cancer detection.

Another systematic review and meta-analysis assessed the overall diagnostic accuracy of DBT alone, combined DBT and DM, and DM alone and found that DBT, with or without DM, had a higher sensitivity than DM alone (8). However, this study was primarily limited to diagnostic populations as well as mixed diagnostic and screening populations. Its results were of limited

Table 3. Summary of meta-regression analyses<sup>a</sup>

Modality	CDR			iCDR			Recall rate			PPV		
	Beta coefficient (95% CI)	Standard error	P <sup>b</sup>	Beta coefficient (95% CI)	Standard error	P <sup>b</sup>	Beta coefficient (95% CI)	Standard error	P <sup>b</sup>	Beta coefficient (95% CI)	Standard error	P <sup>b</sup>
DBT + DM	0.30 (0.15 to 0.44)	0.07	<.001	0.30 (0.10 to 0.49)	0.10	.003	-0.19 (-0.40 to 0.03)	0.11	.09	0.51 (0.17 to 0.86)	0.18	.004
DBT + S2D	0.38 (0.22 to 0.54)	0.08	<.001	0.49 (0.28 to 0.71)	0.11	<.001	-0.51 (-0.77 to -0.26)	0.13	<.001	1.00 (0.59 to 1.41)	0.21	<.001
DBT	0.09 (-0.08 to 0.25)	0.08	.29	0.06 (-0.16 to 0.28)	0.11	.59	0.06 (-0.20 to 0.31)	0.13	.67	0.07 (-0.35 to 0.48)	0.21	.75

<sup>a</sup>DM treated as the reference group for comparison in the meta-regression models. CI = confidence interval; CDR = cancer detection rate; DBT = digital breast tomosynthesis; DM = digital mammography; iCDR = invasive cancer detection rate; PPV1 = positive predictive value; S2D = synthetic 2D image.

<sup>b</sup>The P values correspond to the 2-sided meta-regression analyses comparing each imaging modality.

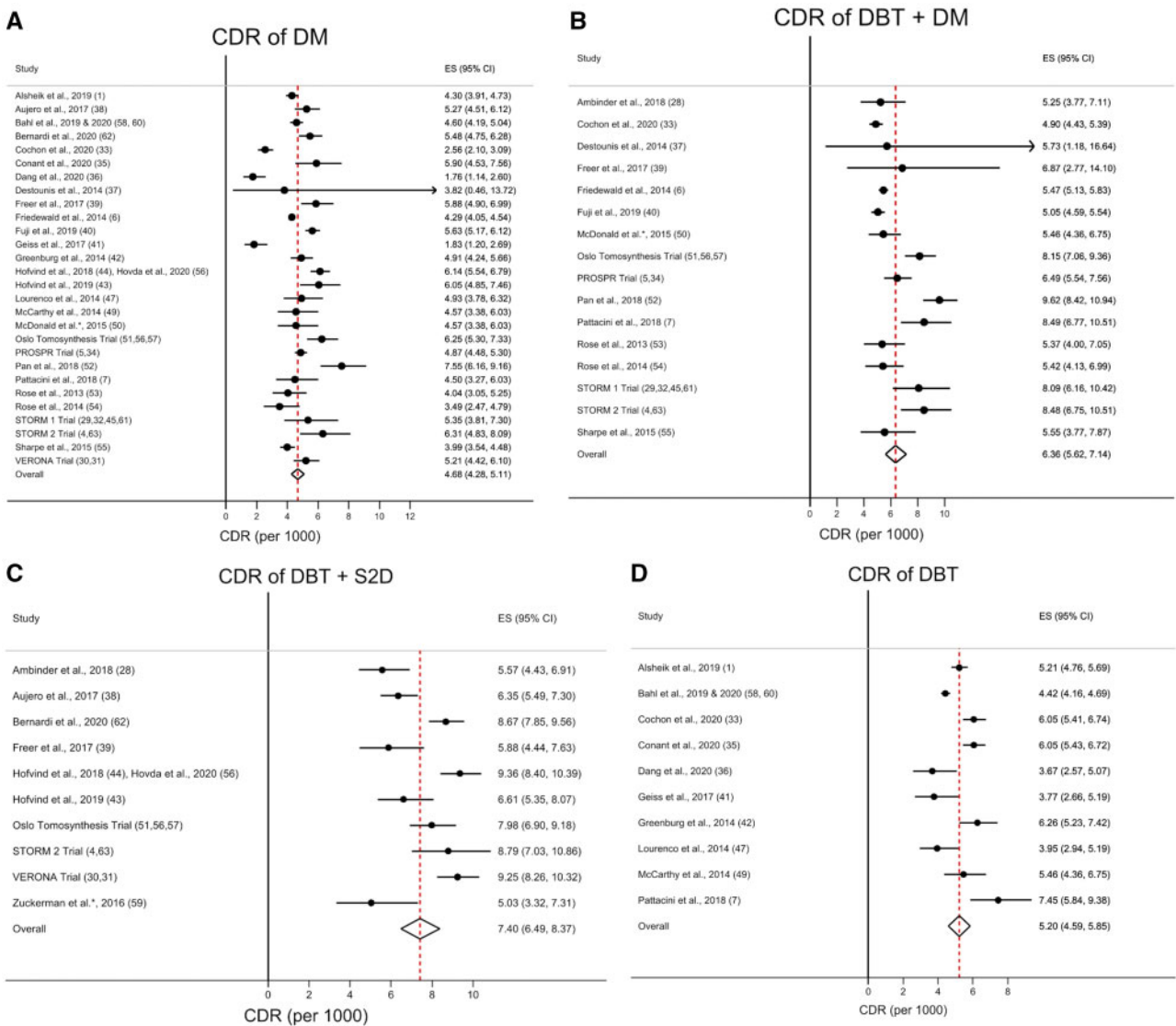
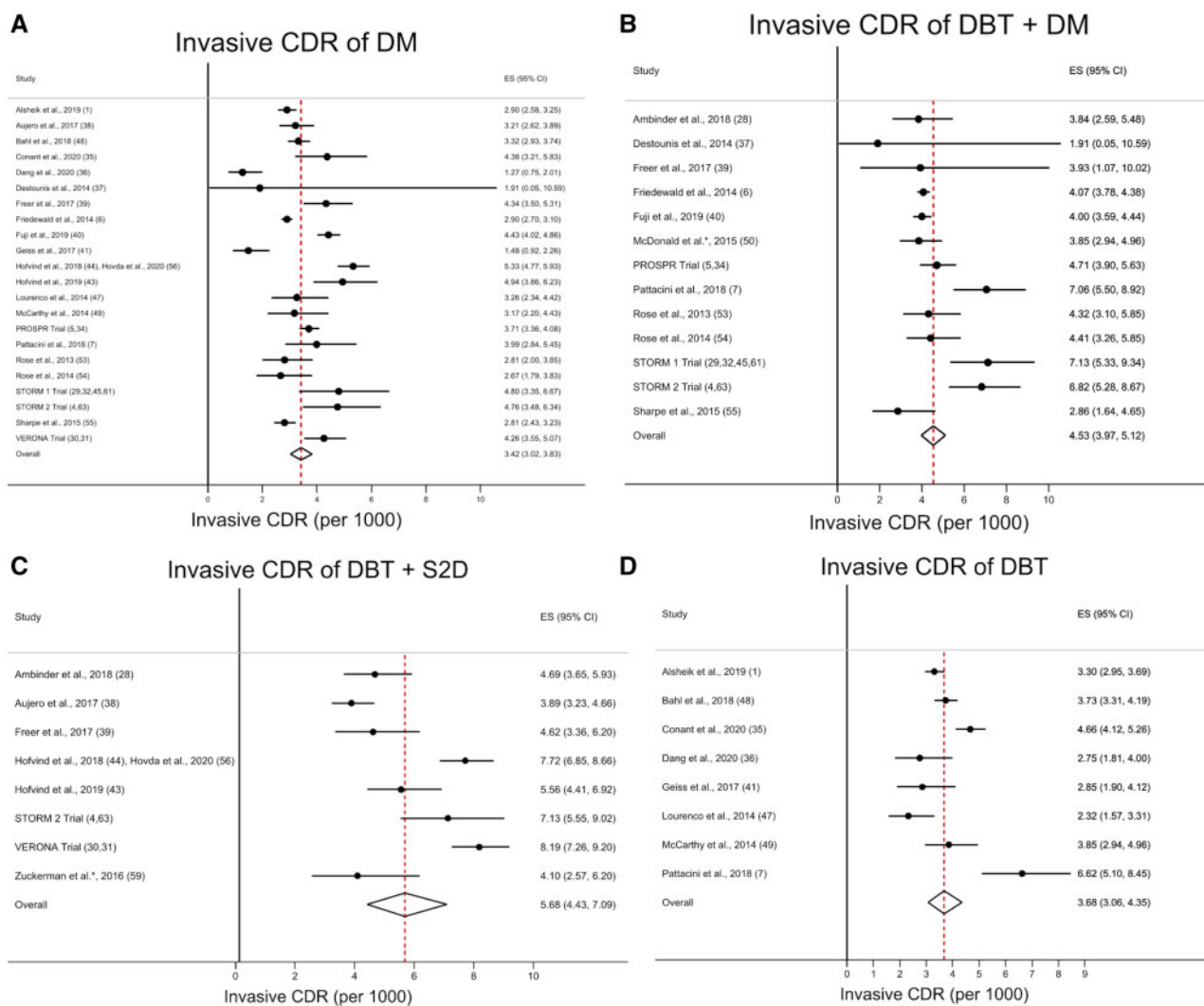


Figure 2. Forest plots and pooled estimates of cancer detection rate (CDR). Results are shown for (A) digital mammography (DM), (B) digital breast tomosynthesis (DBT) + DM, (C) DBT + synthetic 2-dimensional image (S2D), and (D) DBT. The dotted line represents the pooled summary estimate of the pooled CDR with the diamond illustrating the associated 95% confidence interval (CI). The squares represent the CDR for individual studies, with the solid lines associated with them representing the 95% confidence intervals. ES = effect size.



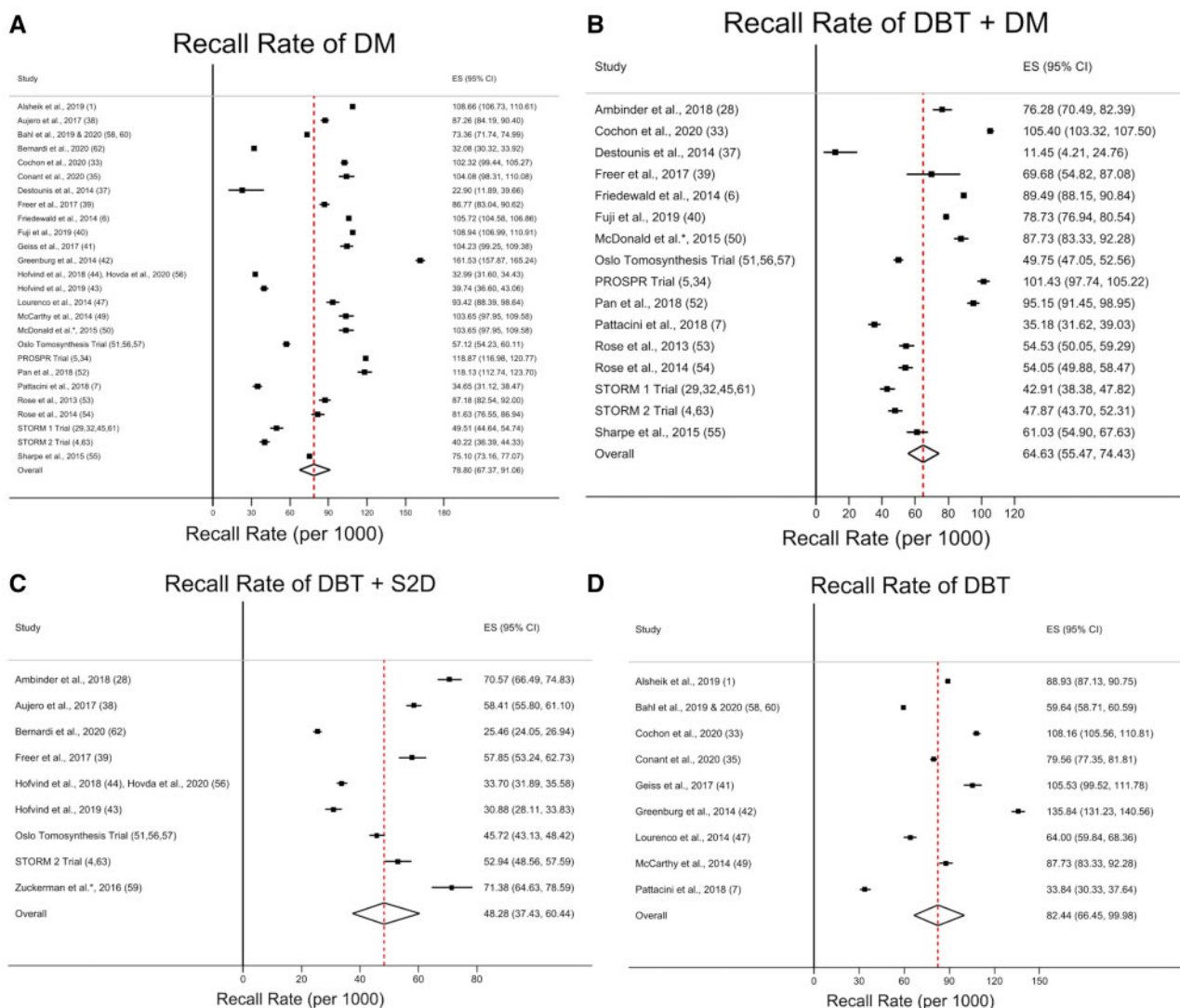
**Figure 3.** Forest plots and pooled estimates of invasive cancer detection rate (CDR). Results are shown for (A) digital mammography (DM), (B) digital breast tomosynthesis (DBT) + DM, (C) DBT + synthetic 2-dimensional image (S2D), and (D) DBT. The dotted line represents the pooled summary estimate of the pooled invasive CDR with the diamond illustrating the associated 95% confidence interval. The squares represent the invasive CDR for individual studies, with the solid lines associated with them representing the 95% confidence interval (CI). ES = effect size.

applicability in the assessment of screening populations alone, which may be more clinically applicable for population-level breast cancer screening programs. Our study has bridged this gap in knowledge, because its results are more clinically relevant and valid for screening populations alone and provided additional assessment of S2D imaging. Moreover, whereas the other study's findings indicated DBT alone is superior to DM alone in diagnostic and mixed populations, our results indicated that DBT alone does not provide an additional benefit in cancer detection or recall rate in the screening populations (8). We found that DM or S2D imaging should be performed in conjunction with DBT to improve the diagnostic performance compared with DM alone. These findings may, in part, be due to the much lower prevalence of breast cancer in screening populations compared with mixed diagnostic and screening populations, because this may specifically influence the PPV1.

The findings of our study contribute to the growing evidence that DBT combined with DM or S2D is superior to DM alone in breast cancer detection (6,8,34,64). These findings support the adoption of DBT into clinical practice for breast cancer

screening (1-3,65), with the combination of DBT with S2D preferred over the current practice of combining DBT with DM (4,31,44,66-68). This is because although both combinations had improved CDRs, invasive CDRs, and PPV1s, combined DBT and S2D would lead to a reduced overall cost and patient radiation exposure compared with combined DBT and DM. Moreover, the reduction in recall rates compared with DM alone will also reduce unnecessary additional imaging as well as invasive biopsies and their associated patient anxiety and potential complications. Of note, the overall effect on cost will depend on the type of health insurance coverage in the region. For instance, in the United States, screening mammography is fully covered with no out-of-pocket costs, but this is not the case for DBT coverage, which varies by state and breast density (69). In light of the findings of our study, reassessment of these policies may be warranted to ensure all eligible women have equal access to screening with DBT (69).

This study has several limitations. First, most of the included studies demonstrated a high risk of bias due to lack of blinding or lack of reporting of blinding status, although this



**Figure 4.** Forest plots and pooled estimates of recall rate. Results are shown for (A) digital mammography (DM), (B) digital breast tomosynthesis (DBT) + DM, (C) DBT + synthetic 2-dimensional image (S2D), and (D) DBT. The dotted line represents the pooled summary estimate of the pooled recall rate with the diamond illustrating the associated 95% confidence interval (CI). The squares represent the recall rate for individual studies, with the solid lines associated with them representing the 95% confidence interval (CI). ES = effect size.

may result in a higher external validity because these methods more closely resemble clinical practice. Moreover, there was statistically significant heterogeneity reported with the pooling of the population. We attempted to account for this with a random effects meta-analysis. Although the statistical heterogeneity may, in part, be due to between-study variability, a component may be due to the high power of this test due to the large number of included studies, resulting in the ability of this test to detect even a small amount of heterogeneity (16). Furthermore, we were not able to adequately assess the influence of breast density, specific tumor types detected, nodal status of invasive cancers detected, the median size of invasive cancers detected, and the use of direct DBT-guided biopsies because this information was not provided in all included studies, and it was beyond the scope of this systematic review. Finally, we excluded articles that did not have an English full text, which may have led us to miss relevant studies.

In summary, this systematic review and meta-analysis found additional benefit with the combined use of DBT and DM in breast cancer screening. As well, the utilization of combined DBT and S2D would provide comparable improved overall and invasive CDRs while reducing the recall rate, radiation dose, and overall cost to the health-care system. All in all, the findings support the replacement of DM by S2D combined with DBT for breast cancer–screening populations.

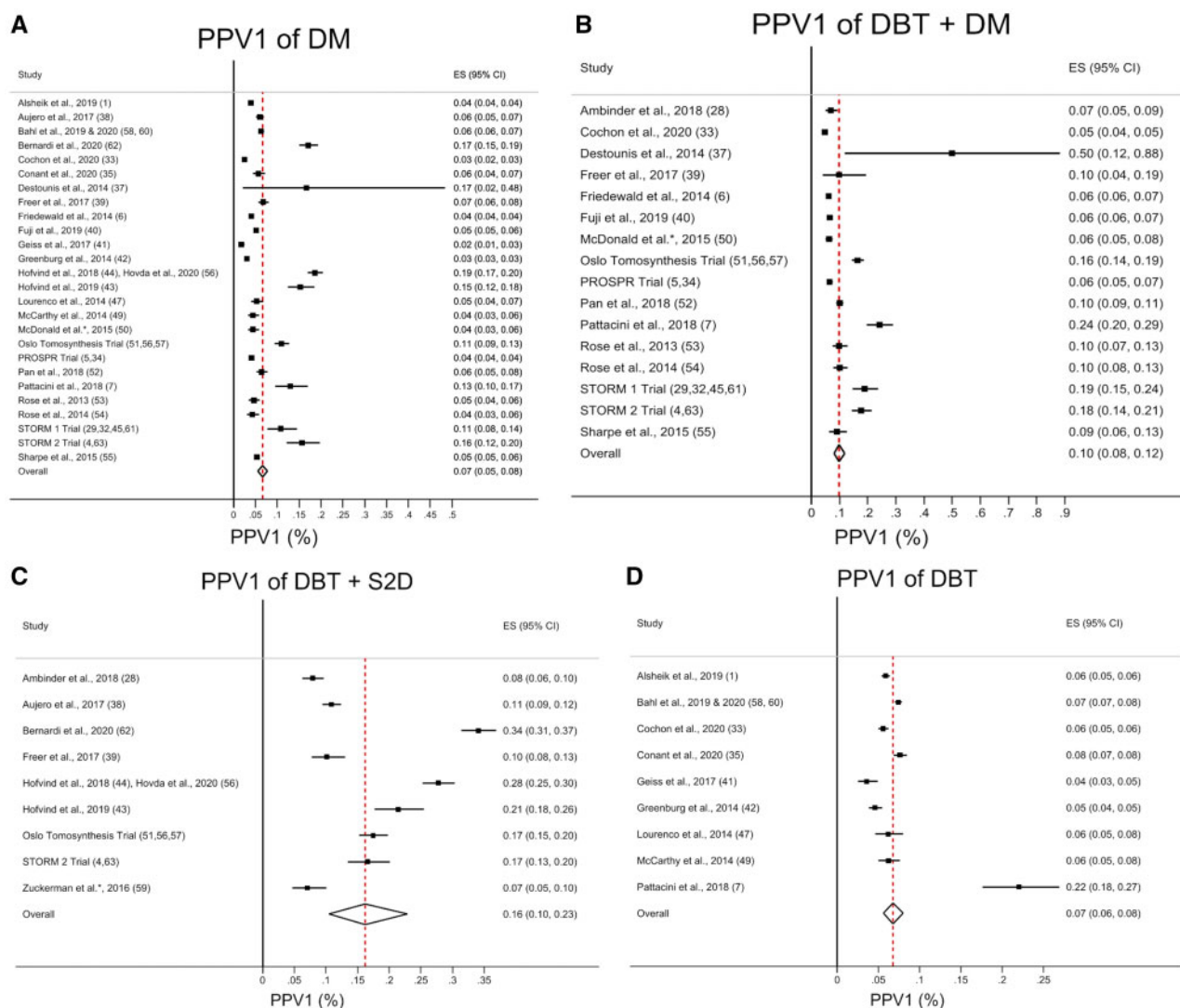
## Funding

There was no funding source for this study.

## Notes

**Role of the funder:** Not applicable.





**Figure 5.** Forest plots and pooled estimates of PPV1. Results are shown for (A) digital mammography (DM), (B) digital breast tomosynthesis (DBT) + DM, (C) DBT + synthetic 2-dimensional image (S2D), and (D) DBT. The dotted line represents the pooled summary estimate of the pooled positive predictive value (PPV1) with the diamond illustrating the associated 95% confidence interval. The squares represent the PPV1 for individual studies, with the solid lines associated with them representing the 95% confidence interval (CI). ES = effect size.

**Disclosures:** The authors have no conflicts of interest to disclose.

**Author contributions:** Mostafa Alabousi: conceptualization, methodology, formal analysis, investigation, data curation, writing—original draft, writing—reviewing and editing. Akshay Wadera: conceptualization, investigation, data curation, writing – original draft, writing – reviewing and editing. Mohammed Kashif Al-Ghita: formal analysis, investigation, data curation, writing – reviewing and editing. Rayeh Kashef Al-Ghetaa: data curation, writing – original draft, writing – reviewing and editing. Jean-Paul Salameh: methodology, formal analysis, data curation, writing – reviewing and editing. Alex Pozdnyakov: data curation, writing – reviewing and editing. Nanxi Zha: data curation, writing – reviewing and editing. Lucy Samoilov: data curation, writing – reviewing and editing. Anahita Dehmoobad Sharifabadi: data curation, writing – reviewing and editing. Behnam Sadeghirad: methodology, formal analysis, writing – reviewing and editing. Vivianne Freitas: conceptualization, methodology, writing – reviewing and editing. Matthew DF

McInnes: conceptualization, methodology, writing – reviewing and editing. Abdullah Alabousi: conceptualization, methodology, investigation, writing – reviewing and editing, supervision.

## Data Availability

The data underlying this article will be shared on reasonable request to the corresponding author.

## References

1. Alsheik NH, Dabbous F, Pohlman SK, et al. Comparison of resource utilization and clinical outcomes following screening with digital breast tomosynthesis versus digital mammography: findings from a learning health system. *Acad Radiol.* 2019;26(5):597–605.
2. Boroumand G, Teberian I, Parker L, Rao VM, Levin DC. Screening mammography and digital breast tomosynthesis: utilization updates. *Am J Roentgenol.* 2018;210(5):1092–1096.
3. Sardanelli F, Fallenberg EM, Clauser P, et al.; for the European Society of Breast Imaging (EUSOBI), with language review by Europa Donna-The European Breast Cancer Coalition. Mammography: an update of the EUSOBI

- recommendations on information for women. *Insights Imaging*. 2017;8(1):11–18.
4. Bernardi D, Macaskill P, Pellegrini M, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. *Lancet Oncol*. 2016;17(8):1105–1113.
  5. Conant EF, Beaber EF, Sprague BL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography compared to digital mammography alone: a cohort study within the PROSPR consortium. *Breast Cancer Res Treat*. 2016;156(1):109–116.
  6. Friedewald SM, Rafferty EA, Rose SL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA*. 2014;311(24):2499.
  7. Pattacini P, Nitrosi A, Giorgi Rossi P, et al.; for the RETomo Working Group. Digital mammography versus digital mammography plus tomosynthesis for breast cancer screening: the Reggio Emilia tomosynthesis randomized trial. *Radiology*. 2018;288(2):375–385.
  8. Alabousi M, Zha N, Salameh JP, et al. Digital breast tomosynthesis for breast cancer detection: a diagnostic test accuracy systematic review and meta-analysis. *Eur Radiol*. 2020;30(4):2058–2071.
  9. Gennaro G, Bernardi D, Houssami N. Radiation dose with digital breast tomosynthesis compared to digital mammography: per-view analysis. *Eur Radiol*. 2018;28(2):573–581.
  10. Svahn TM, Houssami N, Sechopoulos I, Mattsson S. Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography. *Breast*. 2015;24(2):93–99.
  11. Yaffe MJ. Reducing radiation doses for breast tomosynthesis? *Lancet Oncol*. 2016;17(8):1027–1029.
  12. Qaseem A, Lin JS, Mustafa RA, Horwitch CA, Wilt TJ; for the Clinical Guidelines Committee of the American College of Physicians. Screening for breast cancer in average-risk women: a guidance statement from the American College of Physicians. *Ann Intern Med*. 2019;170(8):547.
  13. Klarenbach S, Sims-Jones N, Lewin G, et al. Recommendations on screening for breast cancer in women aged 40–74 years who are not at increased risk for breast cancer. *CMAJ*. 2018;190(49):E1441–E1451.
  14. Oeffinger KC, Fontham ETH, Etzioni R, et al. Breast cancer screening for women at average risk. *JAMA*. 2015;314(15):1599.
  15. Siu AL, on behalf of the U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. 2016;164(4):279.
  16. Deeks JJ, Bossuyt PM, Gatsonis C. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. 2013. <https://methods.cochrane.org/sdt/handbook-dta-reviews>. Accessed December 1, 2018.
  17. McGrath TA, Alabousi M, Skidmore B, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. *Syst Rev*. 2017;6(1):194.
  18. McInnes MDF, Moher D, Thoms BD, et al. and the PRISMA-DTA Group. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies. *JAMA*. 2018;319(4):388.
  19. Moher D, Liberati A, Tetzlaff J, Altman DG. PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg*. 2010;8(5):336–341.
  20. Dehmoabad Sharifabadi A, Leeflang M, Treanor L, et al. Comparative reviews of diagnostic test accuracy in imaging research: evaluation of current practices. *Eur Radiol*. 2019;29(10):5386–5394.
  21. Zha N, Alabousi M, Abdullah P, et al. Breast cancer screening in high-risk patients during pregnancy and breastfeeding: a systematic review of the literature. *J Breast Imaging*. 2019;1(2):92–98.
  22. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529.
  23. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58(10):982–990.
  24. Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. *J Epidemiol Community Health*. 2013;67(11):974–978.
  25. Nyaga VN, Arbyn M, Aerts M. Metaprop: a Stata command to perform meta-analysis of binomial data. *Arch Public Health*. 2014;72(1):1–10.
  26. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539–1558.
  27. Ryan R; Cochrane Consumers and Communication Group. Heterogeneity and subgroup analyses in Cochrane Consumers and Communication Group reviews: planning the analysis at protocol stage. <http://cccg.cochrane.org>. Published 2016. Accessed May 1, 2020.
  28. Ambinder EB, Harvey SC, Panigrahi B, Li X, Woods RW. Synthesized mammography: the new standard of care when screening for breast cancer with digital breast tomosynthesis? *Acad Radiol*. 2018;25(8):973–976.
  29. Caumo F, Bernardi D, Ciatto S, et al. Incremental effect from integrating 3D-mammography (tomosynthesis) with 2D-mammography: increased breast cancer detection evident for screening centres in a population-based trial. *Breast*. 2014;23(1):76–80.
  30. Caumo F, Romanucci G, Hunter K, et al. Comparison of breast cancers detected in the Verona screening program following transition to digital breast tomosynthesis screening with cancers detected at digital mammography screening. *Breast Cancer Res Treat*. 2018;170(2):391–397.
  31. Caumo F, Zorzi M, Brunelli S, et al. Digital breast tomosynthesis with synthesized two-dimensional images versus full-field digital mammography for population screening: outcomes from the Verona screening program. *Radiology*. 2018;287(1):37–46.
  32. Ciatto S, Houssami N, Bernardi D, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol*. 2013;14(7):583–589.
  33. Cochran LR, Giess CS, Khorasani R. Comparing diagnostic performance of digital breast tomosynthesis and full-field digital mammography. *J Am Coll Radiol*. 2020;17(8):999–1003. doi: 10.1016/j.jacr.2020.01.010.
  34. Conant EF, Barlow WE, Herschorn SD, et al.; for the Population-based Research Optimizing Screening Through Personalized Regimen (PROSPR) Consortium. Association of digital breast tomosynthesis vs digital mammography with cancer detection and recall rates by age and breast density. *JAMA Oncol*. 2019;5(5):635–642.
  35. Conant EF, Zuckerman SP, McDonald ES, et al. Five consecutive years of screening with digital breast tomosynthesis: outcomes by screening year and round. *Radiology*. 2020;295(2):285–293.
  36. Dang PA, Wang A, Senapati GM, et al. Comparing tumor characteristics and rates of breast cancers detected by screening digital breast tomosynthesis and full-field digital mammography. *Am J Roentgenol*. 2020;214(3):701–706.
  37. Destounis S, Arieno A, Morgan R. Initial experience with combination digital breast tomosynthesis plus full field digital mammography or full field digital mammography alone in the screening environment. *J Clin Imaging Sci*. 2014;4(1):9. doi: 10.4103/2156-7514.127838.
  38. Ajuero MP, Gavenonis SC, Benjamin R, Zhang Z, Holt JS. Clinical performance of synthesized two-dimensional mammography combined with tomosynthesis in a large screening population. *Radiology*. 2017;283(1):70–76.
  39. Freer PE, Riegert J, Eisenmenger L, et al. Clinical implementation of synthesized mammography with digital breast tomosynthesis in a routine clinical practice. *Breast Cancer Res Treat*. 2017;166(2):501–509.
  40. Fujii MH, Herschorn SD, Sowden M, et al. Detection rates for benign and malignant diagnoses on breast cancer screening with digital breast tomosynthesis in a statewide mammography registry study. *Am J Roentgenol*. 2019;212(3):706–711.
  41. Giess CS, Pourjabbar S, Ip IK, Lacson R, Alper E, Khorasani R. Comparing diagnostic performance of digital breast tomosynthesis and full-field digital mammography in a hybrid screening environment. *Am J Roentgenol*. 2017;209(4):929–934.
  42. Greenberg JS, Javitt MC, Katzen J, Michael S, Holland AE. Clinical performance metrics of 3D digital breast tomosynthesis compared with 2D digital mammography for breast cancer screening in community practice. *Am J Roentgenol*. 2014;203(3):687–693.
  43. Hofvind S, Hølen ÅS, Aase HS, et al. Two-view digital breast tomosynthesis versus digital mammography in a population-based breast cancer screening programme (To-Be): a randomised, controlled trial. *Lancet Oncol*. 2019;20(6):795–805.
  44. Hofvind S, Hovda T, Hølen ÅS, et al. Digital breast tomosynthesis and synthetic 2D mammography versus digital mammography: evaluation in a population-based screening program. *Radiology*. 2018;287(3):787–794.
  45. Houssami N, Macaskill P, Bernardi D, et al. Breast screening using 2D-mammography or integrating digital breast tomosynthesis (3D-mammography) for single-reading or double-reading - Evidence to guide future screening strategies. *Eur J Cancer*. 2014;50(10):1799–1807.
  46. Hovda T, Hølen ÅS, Lång K, et al. Interval and consecutive round breast cancer after digital breast tomosynthesis and synthetic 2D mammography versus standard 2D digital mammography in BreastScreen Norway. *Radiology*. 2020;294(2):256–264.
  47. Lourenco AP, Barry-Brooks M, Baird GL, Tuttle A, Mainiero MB. Changes in recall type and patient treatment following implementation of screening digital breast tomosynthesis. *Radiology*. 2015;274(2):337–342.
  48. Bahl M, Gaffney S, McCarthy AM, Lowry KP, Dang PA, Lehman CD. Breast cancer characteristics associated with 2D digital mammography versus digital breast tomosynthesis for screening-detected and interval cancers. *Radiology*. 2018;287(1):49–57.
  49. McCarthy AM, Kontos D, Synnestvedt M, et al. Screening outcomes following implementation of digital breast tomosynthesis in a general-population screening program. *J Natl Cancer Inst*. 2014;106(11):dj316.
  50. McDonald ES, McCarthy AM, Akhtar AL, Synnestvedt MB, Schnall M, Conant EF. Baseline screening mammography: performance of full-field digital mammography versus digital breast tomosynthesis. *Am J Roentgenol*. 2015;205(5):1143–1148.
  51. Østerås BH, Martinsen ACT, Gullien R, Skaane P. Digital mammography versus breast tomosynthesis: impact of breast density on diagnostic performance in population-based screening. *Radiology*. 2019;293(1):60–68.
  52. Pan HB, Wong KF, Yao A, et al. Breast cancer screening with digital breast tomosynthesis - 4 year experience and comparison with national data. *J Chin Med Assoc*. 2018;81(1):70–80.

53. Rose SL, Tidwell AL, Bujnoch LJ, Kushwaha AC, Nordmann AS, Sexton R. Implementation of breast tomosynthesis in a routine screening practice: an observational study. *Am J Roentgenol*. 2013;200(6):1401–1408.
54. Rose SL, Tidwell AL, Ice MF, Nordmann AS, Sexton R, Song R. A reader study comparing prospective tomosynthesis interpretations with retrospective readings of the corresponding FFDM examinations. *Acad Radiol*. 2014;21(9):1204–1210.
55. Sharpe RE, Venkataraman S, Phillips J, et al. Increased cancer detection rate and variations in the recall rate resulting from implementation of 3D digital breast tomosynthesis into a population-based screening program. *Radiology*. 2016;278(3):698–706.
56. Skaane P, Bandos AI, Eben EB, et al. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. *Radiology*. 2014;271(3):655–663.
57. Skaane P, Sebuødegård S, Bandos AI, et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. *Breast Cancer Res Treat*. 2018;169(3):489–496.
58. Bahl M, Mercaldo S, Dang PA, McCarthy AM, Lowry KP, Lehman CD. Breast cancer screening with digital breast tomosynthesis: are initial benefits sustained? *Radiology*. 2020;295(3):529–539.
59. Zuckerman SP, Conant EF, Keller BM, et al. Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. *Radiology*. 2016;281(3):730–736.
60. Bahl M, Pinnamaneni N, Mercaldo S, McCarthy AM, Lehman CD. Digital 2D versus tomosynthesis screening mammography among women aged 65 and older in the United States. *Radiology*. 2019;291(3):582–590.
61. Bernardi D, Caumo F, Macaskill P, et al. Effect of integrating 3D-mammography (digital breast tomosynthesis) with 2D-mammography on radiologists' true-positive and false-positive detection in a population breast screening trial. *Eur J Cancer*. 2014;50(7):1232–1238.
62. Bernardi D, Gentilini MA, De Nisi M, et al. Effect of implementing digital breast tomosynthesis (DBT) instead of mammography on population screening outcomes including interval cancer rates: results of the Trento DBT pilot evaluation. *Breast*. 2020;50:135–140.
63. Bernardi D, Li T, Pellegrini M, et al. Effect of integrating digital breast tomosynthesis (3D-mammography) with acquired or synthetic 2D-mammography on radiologists' true-positive and false-positive detection in a population screening trial: a descriptive study. *Eur J Radiol*. 2018;106:26–31.
64. Marinovich ML, Hunter KE, Macaskill P, Houssami N. Breast cancer screening using tomosynthesis or mammography: a meta-analysis of cancer detection and recall. *J Natl Cancer Inst*. 2018;110(9):942–949.
65. Gao Y, Babb JS, Toth HK, Moy L, Heller SL. Digital breast tomosynthesis practice patterns following 2011 FDA approval. *Acad Radiol*. 2017;24(8):947–953.
66. Gilbert FJ, Tucker L, Gillan MG, et al. The TOMMY trial: a comparison of TOMosynthesis with digital MammographY in the UK NHS Breast Screening Programme – a multicentre retrospective reading study comparing the diagnostic performance of digital breast tomosynthesis and digital mammography with digital mammography alone. *Health Technol Assess (Rockv)*. 2015;19(4):1–136.
67. Alshafeiy TI, Wadih A, Nicholson BT, et al. Comparison between digital and synthetic 2D mammograms in breast density interpretation. *Am J Roentgenol*. 2017;209(1):W36–W41.
68. Choi JS, Han B-K, Ko EY, et al. Comparison between two-dimensional synthetic mammography reconstructed from digital breast tomosynthesis and full-field digital mammography for the detection of T1 breast cancer. *Eur Radiol*. 2016;26(8):2538–2546.
69. Dense Breast Info Group. Comparative Analysis of State Density Inform Efforts and Insurance Coverage States Listed in Order of Effective "Inform" Law Date. 2020. <http://densedev.epsilonium.com/img/table.laws.insurance.3.26.18.pdf>. Accessed August 27, 2020.