

# Breast cancer screening using synthesized two-dimensional mammography: A systematic review and meta-analysis

Baoqi Zeng<sup>a,\*,1</sup>, Kai Yu<sup>a,1</sup>, Le Gao<sup>b</sup>, Xueyang Zeng<sup>c</sup>, Qingxin Zhou<sup>c</sup>

<sup>a</sup> Department of Science and Education, Peking University Binhai Hospital, Tianjin, China

<sup>b</sup> Department of Pharmacology and Pharmacy, The University of Hong Kong, Hong Kong, China

<sup>c</sup> Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Centre, Beijing, China

## ARTICLE INFO

### Article history:

Received 18 June 2021

Received in revised form

19 July 2021

Accepted 20 July 2021

Available online 22 July 2021

### Keywords:

Breast cancer

Cancer screening

Digital breast tomosynthesis

Meta-analysis

## ABSTRACT

**Purpose:** We conducted a systematic review and meta-analysis to compare the screening performance of synthesized mammography (SM) plus digital breast tomosynthesis (DBT) with digital mammography (DM) plus DBT or DM alone.

**Methods:** Medline, Embase, Web of Science, and the Cochrane Library databases were searched from January 2010 to January 2021. Eligible population-based studies on breast cancer screening comparing SM/DBT with DM/DBT or DM in asymptomatic women were included. A random-effect model was used in this meta-analysis. Data were summarized as risk differences (RDs), with 95 % confidence intervals (CIs).

**Results:** Thirteen studies involving 1,370,670 participants were included. Compared with DM/DBT, screening using SM/DBT had similar breast cancer detection rate (CDR) (RD = -0.1/1000 screens, 95 % CI = -0.4 to 0.2,  $p = 0.557$ ,  $I^2 = 0$  %), but lower recall rate (RD = -0.56 %, 95 % CI = -1.03 to -0.08,  $p = 0.022$ ,  $I^2 = 90$  %) and lower biopsy rate (RD = -0.33 %, 95 % CI = -0.56 to -0.10,  $p = 0.005$ ,  $I^2 = 78$  %). Compared with DM, SM/DBT improved CDR (RD = 2.0/1000 screens, 95 % CI = 1.4 to 2.6,  $p < 0.001$ ,  $I^2 = 63$  %) and reduced recall rate (RD = -0.95 %, 95 % CI = -1.91 to -0.002,  $p = 0.049$ ,  $I^2 = 99$  %). However, SM/DBT and DM had similar interval cancer rate (ICR) (RD = 0.1/1000 screens, 95 % CI = -0.6 to 0.8,  $p = 0.836$ ,  $I^2 = 71$  %) and biopsy rate (RD = -0.05 %, 95 % CI = -0.35 to 0.24,  $p = 0.727$ ,  $I^2 = 93$  %).

**Conclusions:** Screening using SM/DBT has similar breast cancer detection but reduces recall and biopsy when compared with DM/DBT. SM/DBT improves CDR when compared with DM, but they have little difference in ICR. SM/DBT could replace DM/DBT in breast cancer screening to reduce radiation dose.

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Contents

1. Introduction .....	271
2. Materials and methods .....	271
2.1. Data sources and searches .....	271
2.2. Selection of studies .....	271
2.3. Data extraction .....	271
2.4. Statistical analysis .....	271
3. Results .....	272
3.1. Literature search and study characteristics .....	272
3.2. Risk of bias .....	272
3.3. Cancer detection .....	272
3.4. Recall .....	272
3.5. Interval cancer rate .....	272

\* Corresponding author.

E-mail address: [zengbaoqi@126.com](mailto:zengbaoqi@126.com) (B. Zeng).

<sup>1</sup> Baoqi Zeng and Kai Yu contributed equally to this work.

3.6.	Biopsy .....	273
3.7.	PPV for recall .....	273
3.8.	PPV for biopsy .....	274
3.9.	Sensitivity and subgroup analysis .....	276
4.	Discussion .....	276
5.	Conclusion .....	277
	Ethical approval .....	277
	Declaration of competing interest .....	277
	Acknowledgements .....	277
	Supplementary data .....	277
	References .....	277

## 1. Introduction

Digital breast tomosynthesis (DBT) has increasingly been used for breast cancer screening and diagnosis in the past decade [1,2]. In Europe and the USA, some institutions have adopted DBT as the standard method for breast cancer screening [3,4]. Breast cancer screening using DBT in combination with digital mammography (DM) has been shown to increase cancer detection rates (CDRs) compared with DM alone, but the recall rates vary between studies [5–9]. At the same time, multiple diagnostic studies showed that the addition of DBT to DM improved the diagnostic performance for breast cancer [10,11].

The primary limitation of using DM plus DBT compared with DM alone for screening is the increase in the radiation dose to the breast being imaged, which approximately doubles [12]. Therefore, concerns about radiation make DM/DBT less suitable for population screening programs. To resolve the concern, two-dimensional synthetic mammography (SM) was developed, providing reconstruction images similar to those from DM using the data from the tomosynthesis acquisitions [8,13]. SM was approved by the U.S. Food and Drug Administration in May 2013 as an alternative to digital mammography in DBT screening [14]. However, some researchers had concerns about the image quality of SM, especially for calcification detection and characterization [11,15].

Several reviews on SM screening have been published to date, but most without any quantitative analyses [12,15–18]. At the same time, some important population-based studies on breast cancer screening have been published in the past several years [19–22]. Recently, a comprehensive meta-analysis of four imaging modalities (SM/DBT, DM/DBT, DBT, and DM) has been published, which mainly used proportional meta-analysis rather than head-to-head comparison [23]. Another review of DBT performed analysis of CDR and recall rate for SM/DBT versus DM, which was simple and incomplete [24]. Therefore, a comprehensive head-to-head meta-analysis is necessary to prove the clinical performance of SM/DBT comparing with DM/DBT or DM.

The objective of this systematic review and meta-analysis was to assess CDR, interval cancer rate (ICR), recall rate, biopsy rate, and positive predictive value (PPV) of SM/DBT compared with DM/DBT or DM using head-to-head comparison in population breast cancer screening.

## 2. Materials and methods

### 2.1. Data sources and searches

We searched Medline, Embase, Cochrane Library, and Web of Science from January 2010 to January 2021 to align with the clinical application of SM. Keywords including “synthesized mammography”, “synthetic mammography”, “tomosynthesis”, “DBT”, “3D

mammography”, “breast cancer”, and “screening” were used to search, the detailed search strategy was shown in the Supplementary material. Additionally, we identified references by searching the reference lists of included studies and relevant reviews.

### 2.2. Selection of studies

Eligible population breast cancer screening studies were required to have compared SM/DBT with DM/DBT or DM. Studies using either a paired design (comparing different screening methods used within the same group of participants) or unpaired design (comparing different screening methods used in different groups of participants) were eligible for inclusion. The primary outcomes were CDR and recall rate. ICR, invasive CDR, and ductal carcinoma in situ (DCIS), positive predictive value of recall (PPV-1) and biopsy (PPV-2) were considered as secondary outcomes. We excluded studies enrolling symptomatic or high-risk women or conducted in non-screening settings.

### 2.3. Data extraction

Two investigators (BZ and XZ) reviewed titles and abstracts independently and identified eligible studies that met pre-specified inclusion criteria. Potentially eligible studies were subsequently scrutinized in duplicate, and the consensus was achieved through consulting with a third investigator (KY) for disagreement. Study type, study design (paired or unpaired), examination modality, region, DBT vendor, number of participants, and population characteristics, including age and dense breast proportion were extracted from eligible studies. The risk of bias of included studies was assessed using the modified Revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) checklist [25] for screening setting.

### 2.4. Statistical analysis

Two pairs of comparisons (SM/DBT vs. DM/DBT and SM/DBT vs. DM) were analyzed separately. Estimates and 95 % confidence intervals (CIs) of CDRs and recall rates were calculated for each study, and summary estimates were computed using the “metaprop” command with random effects in Stata. Outcomes were pooled as risk difference (RD) with 95 % CIs. RD was used as main analysis for primary outcomes, and sensitivity analysis was performed using RR. A head-to-head comparison was performed using DerSimonian and Laird random-effects meta-analysis [26]. Subgroup analysis was performed by study design (paired or unpaired). Statistical heterogeneity between the studies was assessed with the  $\chi^2$  test and the  $I^2$  statistics.  $I^2$  values of 25 %, 50 %, and 75 % have been suggested to be indicators of low, moderate, and high

heterogeneity, respectively [27]. We performed sensitivity analyses stratifying by region (US or other), reading practice (double or single), study type (prospective or retrospective), or screening interval (biennial or annual) to explore potential sources of heterogeneity for primary outcomes. Tests of subgroup differences were based on the  $\chi^2$  test. For unpaired study designs, 95% CIs of RDs were calculated based on differences in two independent proportions. For paired study designs, “mcci” in Stata was used to take account of the pairing of results within an individual when computing the 95 % CIs of the difference in proportions. Because cross-classification data for biopsy and PPVs in paired studies couldn’t been obtained, we calculated the 95 % CIs of RDs in the same way as unpaired studies. Estimates and 95 % CIs were then input into Stata for meta-analysis. All the analyses were performed with STATA 14.

### 3. Results

#### 3.1. Literature search and study characteristics

The systematic literature search identified 1889 publications, after exclusion of duplicate and irrelevant publications, 63 reports were evaluated in full-text for eligibility (Supplementary Fig. 1). Finally, 21 articles reporting on 13 studies were included in the present meta-analysis, including 1,370,670 participants undergoing SM/DBT, DM/DBT, or DM. Seven studies compared SM/DBT with DM/DBT, and eleven compared SM/DBT with DM. Eight studies were prospective trials and conducted in Europe or Australia, and the remaining five studies were retrospective and conducted in the US. Only one of 13 studies was randomized controlled trial (RCT) [19]. Three studies were paired designs, and the remaining ten studies used unpaired designs. Characteristics of individual studies are summarized in Table 1. Cross-classification of cell counts for paired studies are available in Supplementary Table 1 and Supplementary Table 2.

#### 3.2. Risk of bias

The risk of bias was rated as low for included studies after assessing by the modified QUADAS-2 checklist. The detailed risk of bias assessment is available in Supplementary Fig. 2 and Supplementary Fig. 3.

#### 3.3. Cancer detection

Study-specific data for CDR was described in Table 2. Seven

studies were included in the analysis of CDR comparing SM/DBT with DM/DBT. There was no difference in CDR between SM/DBT and DM/DBT (RD = -0.1/1000 screens, 95 % CI = -0.4 to 0.2,  $p = 0.557$ ,  $I^2 = 0 %$ ) (Fig. 1). Subgroup analysis showed that there was no significant difference between two groups in both paired studies and unpaired studies ( $p_{interaction} = 0.722$ ). Screening using SM/DBT had similar invasive CDR and DCIS detection rate comparing with DM/DBT (Supplementary material Fig. 4 and Supplementary material Fig. 5).

Eleven studies reported CDR comparing SM/DBT with DM. CDR was estimated to be significantly higher when using SM/DBT compared to DM alone (RD = 2.0/1000 screens, 95 % CI = 1.4 to 2.6,  $p < 0.001$ ,  $I^2 = 63 %$ ) (Fig. 1). Subgroup analysis showed that the CDR was higher in SM/DBT group comparing with DM alone in three paired studies (RD = 1.7/1000 screens, 95 % CI = 1.1 to 2.2,  $p < 0.001$ ,  $I^2 = 28 %$ ) and eight unpaired studies (RD = 2.1/1000 screens, 95 % CI = 1.1 to 3.0,  $p < 0.001$ ,  $I^2 = 69 %$ ) ( $p_{interaction} = 0.140$ ). Screening using SM/DBT had higher invasive CDR (RD = 1.9/1000 screens, 95 % CI = 1.2 to 2.7,  $p < 0.001$ ,  $I^2 = 60 %$ ) but similar DCIS detection rate (RD = 0.2/1000 screens, 95 % CI = -0.1 to 0.5,  $p = 0.207$ ,  $I^2 = 48 %$ ) comparing with DM (Supplementary material Fig. 6 and Supplementary material Fig. 7).

#### 3.4. Recall

The recall rate for each study was listed in Table 2. Seven studies were included in the analysis of recall rate comparing SM/DBT with DM/DBT. The recall rate was lower in SM/DBT group than DM/DBT group (RD = -0.56 %, 95 % CI = -1.03 to -0.08,  $p = 0.022$ ,  $I^2 = 90 %$ ) (Fig. 2). Subgroup analysis showed that the recall rate was lower in SM/DBT group in five unpaired studies (RD = -0.87 %, 95 % CI = -1.40 to -0.35,  $p = 0.001$ ,  $I^2 = 82 %$ ), but similar in two paired studies (RD = 0.09 %, 95 % CI = -0.67 to 0.85,  $p = 0.817$ ,  $I^2 = 90 %$ ) comparing with DM/DBT ( $p_{interaction} < 0.001$ ).

Ten studies reported recall rate comparing SM/DBT with DM alone. The recall rate was lower in SM/DBT group than DM group (RD = -0.95 %, 95 % CI = -1.91 to -0.002,  $p = 0.049$ ,  $I^2 = 99 %$ ) (Fig. 2). Subgroup analysis showed that the recall rate was lower in SM/DBT group in eight unpaired studies (RD = -1.34 %, 95 % CI = -2.46 to -0.23,  $p = 0.018$ ,  $I^2 = 99 %$ ) but similar in two paired studies (RD: 0.57 %, 95 % CI = -0.078 to 1.92,  $p = 0.408$ ,  $I^2 = 96 %$ ) comparing with DM alone ( $p_{interaction} < 0.001$ ).

#### 3.5. Interval cancer rate

Only two studies reported ICR of SM/DBT, both were in

**Table 1**  
Study characteristics and patient demographics.

Study	Region	Study type	Age range	Participants	DBT vendors	Study design	Reading	Dense breast	Comparison
STORM-2 [28,29]	Italy	Prospective	≥49	9672	Hologic	Paired	Double	26.8 %	DM/DBT; DM
Martin (2018) [30]	Spain	Prospective	50–69	16,067	Hologic	Paired	Single	26.2 %	DM
OTST [20,31–33]	Norway	Prospective	50–69	24,301	Hologic	Paired	Single	NR	DM/DBT; DM
VSP [21,34]	Italy	Prospective	50–69	31,089	Hologic	Unpaired	Double	16.8 %	DM
TSP [35]	Italy	Prospective	≥50	83,779	Hologic	Unpaired	Double	NR	DM
To-Be trial [19,36]	Norway	RCT	50–69	28,749	GE	Unpaired	Double	NR	DM
OVVV [22,37]	Norway	Prospective	50–69	98,927	Mixed <sup>a</sup>	Unpaired	Double	NR	DM
Aujero (2017) [38]	US	Retrospective	NR	78,810	Hologic	Unpaired	Single	57.1 %	DM/DBT; DM
Freer (2017) [39]	US	Retrospective	19–100	31,979	Hologic	Unpaired	Single	40.1 %	DM/DBT; DM
Ambinder (2018) [40]	US	Retrospective	NR	22,535	Hologic	Unpaired	Single	NR	DM/DBT
Houssami (2019) [41]	Australia	Prospective	40–93	10,146	Hologic	Unpaired	Double	NR	DM
Zuckerman [14,42]	US	Retrospective	NR	151,363	Hologic	Unpaired	Single	31.4 %	DM/DBT
Cohen (2020) [43]	US	Retrospective	NR	783,253	Hologic	Unpaired	Single	NR	DM/DBT; DM

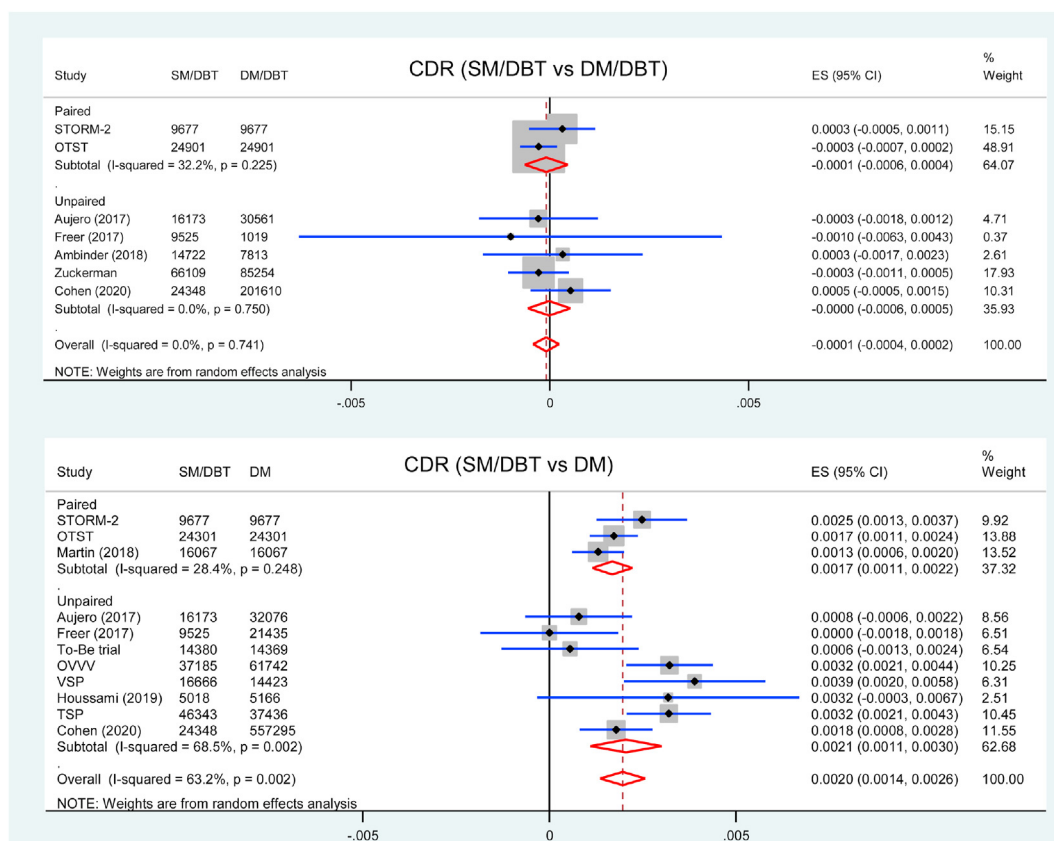
DBT digital breast tomosynthesis, DM digital mammography, RCT randomized controlled trial, NR not reported, TSP Trento screening program, VSP Verona Screening Program, OIT Oslo Tomosynthesis Screening Trial, OVVV Oslo-Vestfold-Vestre Viken.

<sup>a</sup> Including Hologic, GE and Siemens.

**Table 2**  
CDR and recall rate for each study.

Study	CDR/1000 (95 % CI)			Recall rate/100 (95 % CI)		
	SM/DBT	DM/DBT	DM	SM/DBT	DM/DBT	DM
<i>Paired studies</i>						
STORM-2 [28,29]	8.8 (7.0–10.8)	8.5 (6.7–10.5)	6.3 (4.8–8.1)	5.3 (4.9–5.8)	4.8 (4.4–5.2)	4.0 (3.6–4.4)
Martin (2018) [30]	5.4 (4.3–6.7)	NR	4.1 (3.2–5.2)	3.0 (2.7–3.3)	NR	3.1 (2.8–3.4)
OTST [20,31–33]	8.0 (6.9–9.2)	7.8 (6.8–9.0)	6.3 (5.3–7.3)	2.3 (2.1–2.5)	2.6 (2.4–2.8)	NR
<b>Summary estimate</b>	7.3 (5.4–9.4)	8.0 (7.1–9.0)	5.5 (4.1–7.1)	3.4 (2.1–5.1)	3.1 (3.0–3.3)	3.4 (3.2–3.7)
<i>Unpaired studies</i>						
VSP [21,34]	9.3 (7.9–10.9)	NR	5.4 (4.3–6.7)	4.0 (3.7–4.3)	NR	4.2 (3.9–4.6)
To-Be trial [19,36]	6.6 (5.3–8.1)	NR	6.1 (4.9–7.0)	3.1 (2.8–3.4)	NR	4.0 (3.7–4.3)
OVVV [22,37]	9.4 (8.4–10.4)	NR	6.1 (5.5–6.8)	3.4 (3.2–3.6)	NR	3.3 (3.2–3.4)
Aujero (2017) [38]	6.1 (4.9–7.4)	6.3 (5.5–7.3)	5.5 (4.7–6.4)	4.3 (3.9–4.6)	5.8 (5.6–6.1)	9.2 (8.8–9.5)
Freer (2017) [39]	5.9 (4.4–7.6)	6.9 (2.8–14.1)	5.9 (4.9–7.0)	5.8 (5.3–6.3)	7.0 (5.5–8.7)	8.7 (8.3–9.1)
Ambinder (2018) [40]	5.6 (4.4–6.9)	5.2 (3.8–7.1)	NR	7.1 (6.7–7.5)	7.6 (7.1–8.2)	NR
Houssami (2019) [41]	9.8 (7.2–12.9)	NR	6.6 (4.6–9.2)	4.2 (3.7–4.8)	NR	3.0 (2.6–3.5)
Zuckerman [14,42]	5.6 (5.0–6.2)	5.9 (5.4–6.4)	NR	7.0 (6.8–7.2)	7.9 (7.7–8.1)	NR
Cohen (2020) [43]	5.8 (4.9–6.9)	5.3 (5.0–5.6)	4.0 (3.9–4.2)	7.3 (6.9–7.6)	7.5 (7.4–7.7)	10.0 (9.9–10.1)
TSP [35]	8.7 (7.9–9.6)	NR	5.5 (4.9–7.0)	2.6 (2.4–2.7)	NR	3.2 (3.0–3.4)
<b>Summary estimate</b>	7.1 (6.0–8.3)	5.6 (5.1–6.1)	5.5 (4.6–6.5)	4.7 (3.5–6.1)	7.2 (6.5–7.9)	5.4 (3.2–8.2)
<i>All studies</i>	7.2 (6.3–8.1)	6.3 (5.5–7.2)	5.5 (4.8–6.4)	4.4 (3.4–5.6)	6.0 (4.7–7.5)	5.0 (3.0–7.4)

CDR cancer detection rate, SM synthesized mammography, DBT digital breast tomosynthesis, DM digital mammography, NR not reported, CI confidence interval.



**Fig. 1.** Forest plot for the CDR of SM/DBT compared with DM/DBT or DM.

comparison with DM alone. There was no significant difference in ICR between SM/DBT and DM alone (RD = 0.1/1000 screens, 95 % CI = -0.6 to 0.8, p = 0.836, I<sup>2</sup> = 71 %) (Fig. 3).

### 3.6. Biopsy

Five studies reported biopsy rates comparing SM/DBT with DM/DBT, all of them were unpaired studies. Screening using SM/DBT had lower biopsy rate compared with DM/DBT (RD = -0.33 %, 95 %

CI = -0.56 to -0.10, p = 0.005, I<sup>2</sup> = 78 %) (Fig. 4). Seven studies were included in the analysis of biopsy rate comparing SM/DBT with DM, and the biopsy rate was similar between two groups (RD = -0.05 %, 95 % CI = -0.35 to 0.24, p = 0.727, I<sup>2</sup> = 93 %) (Fig. 4).

### 3.7. PPV for recall

PPV-1 was reported in seven studies comparing SM/DBT with DM/DBT, two of them were paired study. Screening using SM/DBT

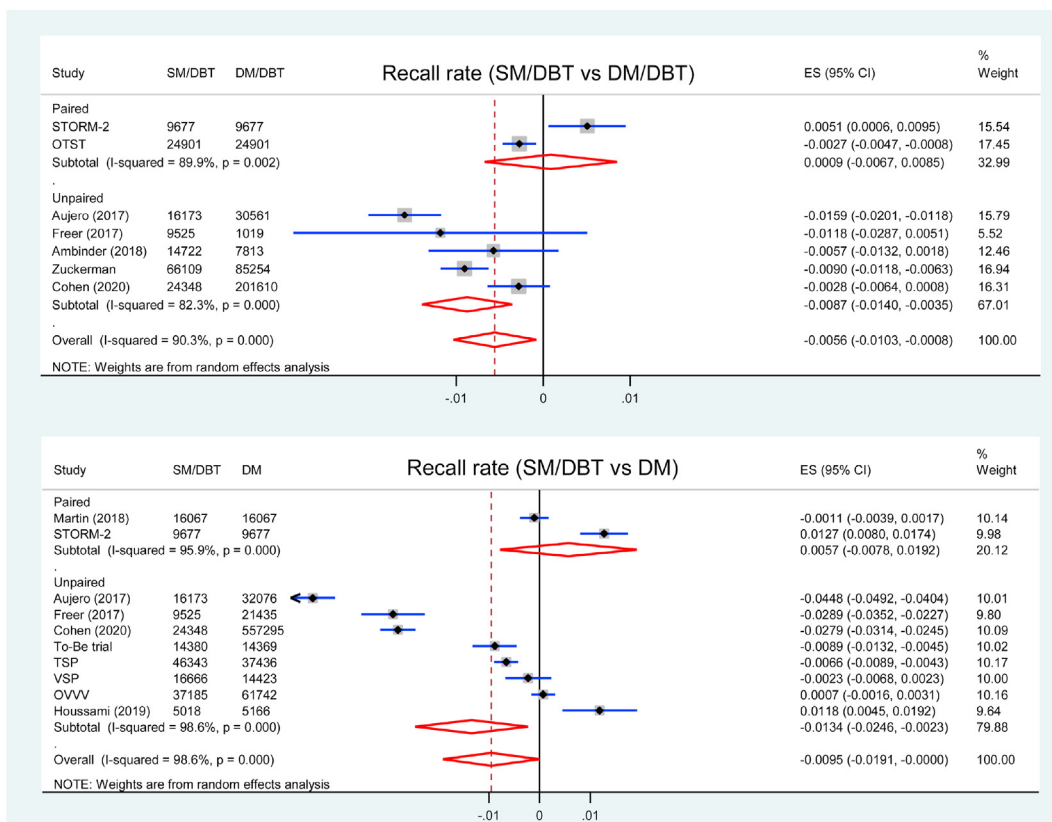


Fig. 2. Forest plot for the recall rate of SM/DBT compared with DM/DBT or DM.

had higher PPV-1 compared with DM/DBT (RD = 0.86 %, 95 % CI = 0.11 to 1.61, p = 0.026, I<sup>2</sup> = 0 %) (Fig. 5). Subgroup analysis showed that PPV-1 was higher in SM/DBT group in five unpaired studies (RD = 0.88 %, 95 % CI = 0.11 to 1.65, p = 0.025, I<sup>2</sup> = 0 %) but similar in two paired studies (RD = 0.29 %, 95 % CI = -3.27 to 4.30, p = 0.886, I<sup>2</sup> = 0 %) comparing with DM/DBT (p<sub>interaction</sub> = 0.778).

Ten studies were included in the analysis of PPV-1 comparing SM/DBT with DM, and PPV for recall was higher in SM/DBT group (RD = 6.85 %, 95 % CI = 4.09 to 9.60, p < 0.001, I<sup>2</sup> = 83 %) (Fig. 5). Subgroup analysis showed that PPV-1 was higher in SM/DBT group in eight unpaired studies (RD = 7.73 %, 95 % CI = 4.56 to 10.91, p < 0.001, I<sup>2</sup> = 86 %) but similar in two paired studies (RD = 2.98 %, 95 % CI = 0.83 to 6.80, p = 0.126, I<sup>2</sup> = 10 %) comparing with DM alone (p<sub>interaction</sub> = 0.125).

### 3.8. PPV for biopsy

Five studies reported PPV-2 for biopsy comparing SM/DBT with DM/DBT, all of them were unpaired studies. Screening using SM/DBT had higher PPV-2 compared with DM/DBT (RD = 6.44 %, 95 % CI = 2.87 to 10.01, p < 0.001, I<sup>2</sup> = 0 %) (Fig. 6).

Seven studies were included in the analysis of PPV-2 comparing SM/DBT with DM, and PPV-2 was higher in SM/DBT groups (RD = 11.29 %, 95 % CI = 7.04 to 15.54, p < 0.001, I<sup>2</sup> = 27 %) (Fig. 6). Subgroup analysis showed that PPV-2 was higher in SM/DBT group in six unpaired studies (RD = 12.20 %, 95 % CI = 8.09 to 16.31, p < 0.001, I<sup>2</sup> = 19 %) but similar in one paired study (RD = 1.74 %, 95 % CI = -12.68 to 16.15, p = 0.813) comparing with DM alone (p<sub>interaction</sub> = 0.157).

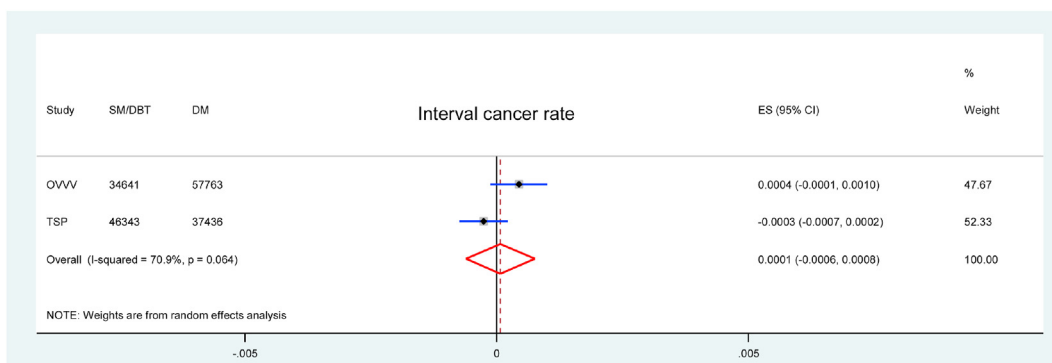


Fig. 3. Forest plot for the ICR of SM/DBT compared with DM.



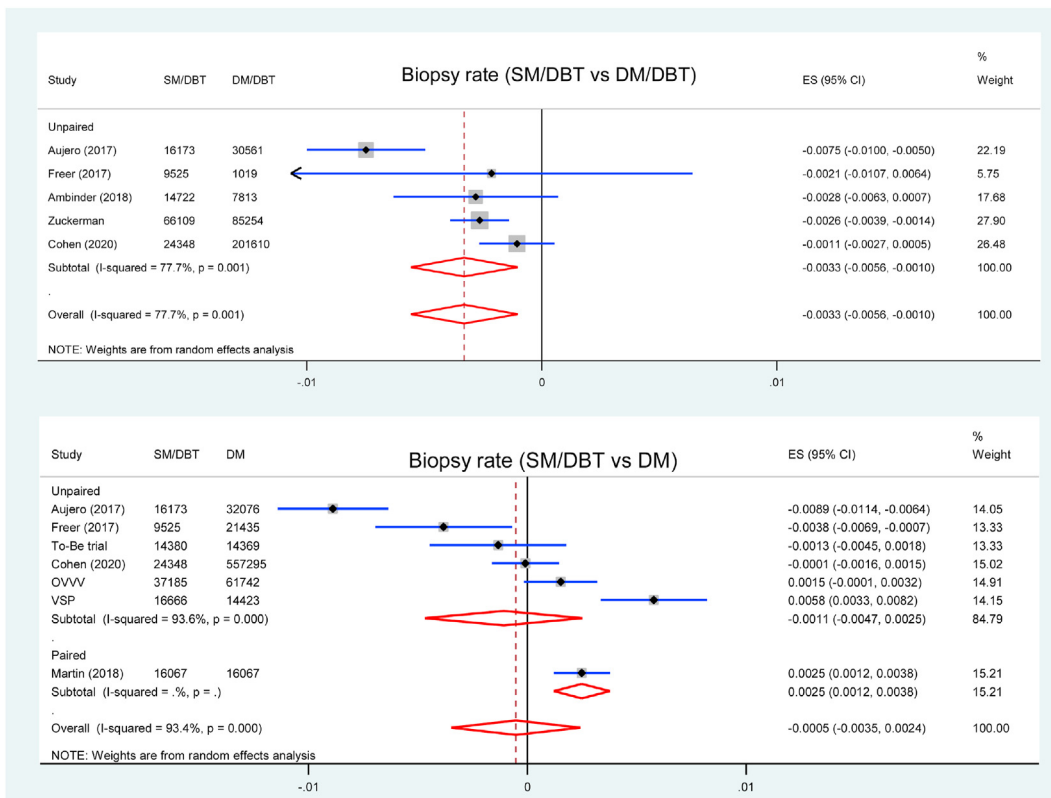


Fig. 4. Forest plot for the biopsy rate of SM/DBT compared with DM/DBT or DM.

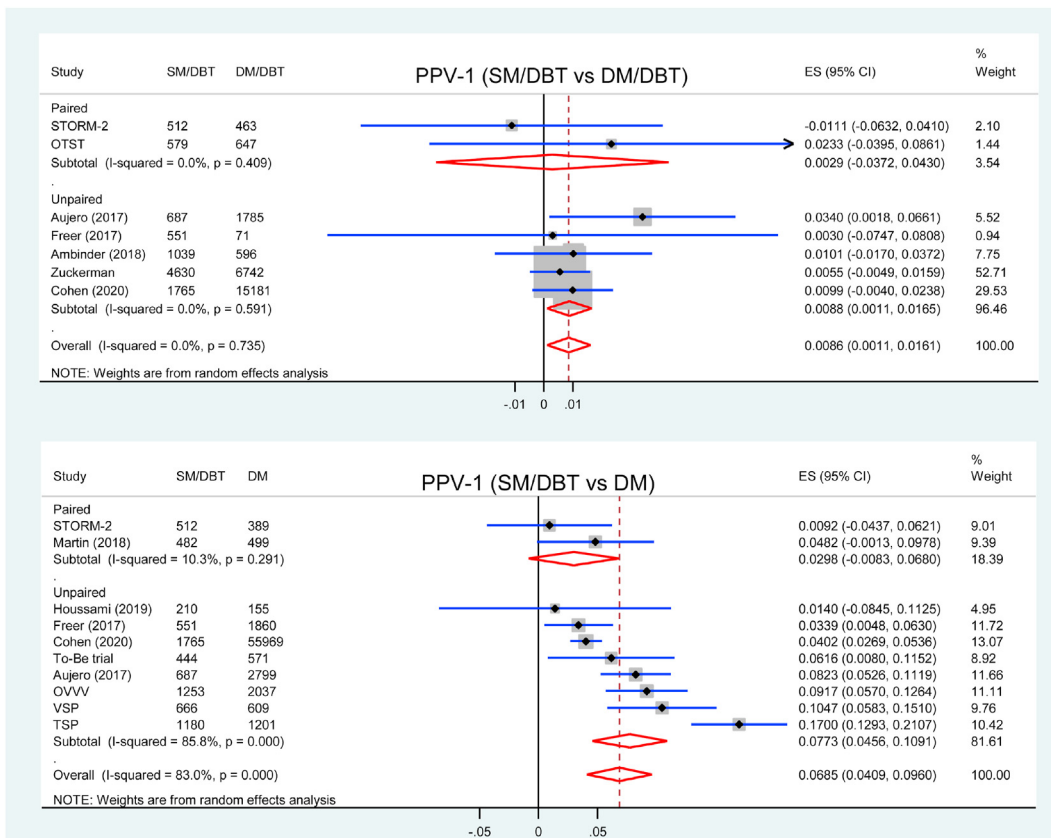


Fig. 5. Forest plot for PPV-1 of SM/DBT compared with DM/DBT or DM.

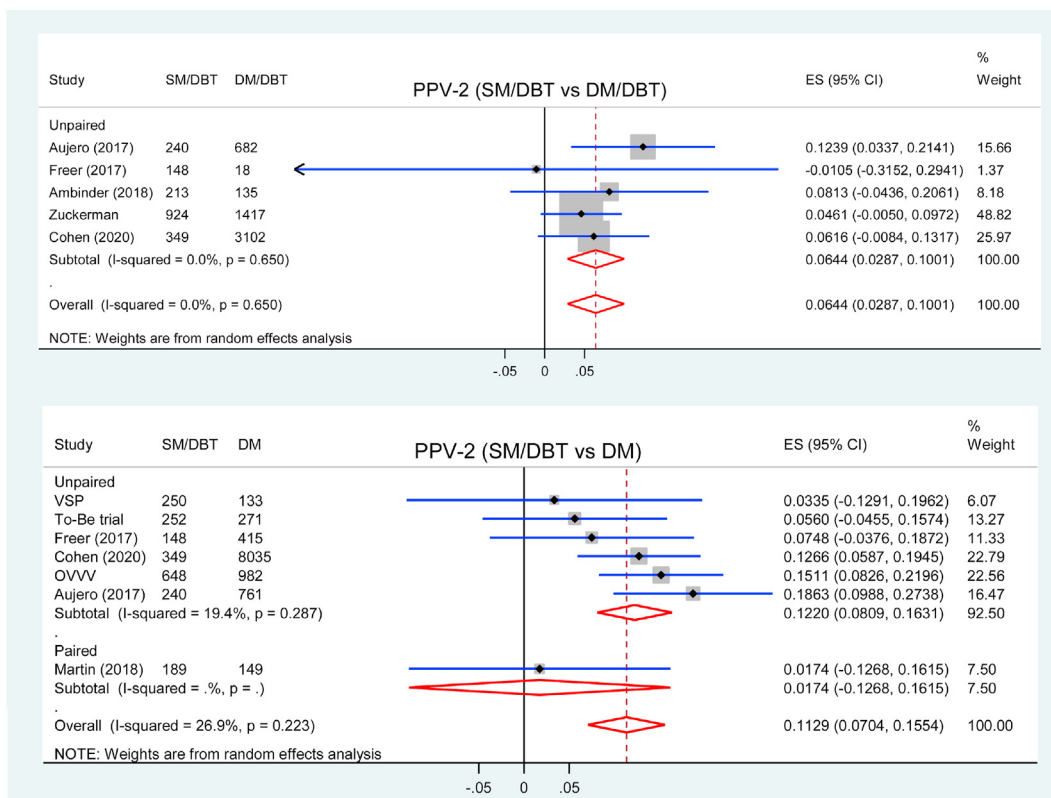


Fig. 6. Forest plot for PPV-2 of SM/DBT compared with DM/DBT or DM.

### 3.9. Sensitivity and subgroup analysis

We performed a post hoc meta-analysis using RRs for primary outcomes, the results were in consistency with RDs which used as main analysis in this study. The primary results stratifying by region/screening interval/study type and reading practice were shown in Table 3. There was a significant interaction between CDR and reading practice (double or single) when comparing SM/DBT with DM ( $p < 0.001$ ), the RD of CDR for double reading was twice that of single reading. All interaction p values were significant ( $p < 0.001$ ) for recall rate with region/screening interval/study type or reading practice when comparing SM/DBT with DM/DBT or DM.

### 4. Discussion

This study provided an overview of clinical performance metrics of SM/DBT compared with DM/DBT or DM for breast cancer screening. The SM imaging was created to maintain the benefits of combined DM and DBT imaging while decreasing patient dose [44]. Our meta-analysis indicated that screening with SM/DBT had similar benefits for detecting all cancer, invasive cancer, and DCIS compared with DM/DBT. Moreover, SM/DBT was associated with a lower biopsy rate and lower recall rate compared with DM/DBT. When comparing with DM, screening with SM/DBT improved CDR (incremental CDR: 2.0 cancers/1000 screens) and invasive cancer detection rate (incremental CDR: 1.9 cancers/1000 screens).

Table 3  
Sensitivity and subgroup analysis for CDR and recall rate.

Subgroup	CDR/1000 (95 % CI)					Recall rate/100 (95 % CI)				
	N	RD	p	I <sup>2</sup>	Pinteraction	N	RD	p	I <sup>2</sup>	Pinteraction
<i>SM/DBT vs DM/DBT</i>										
US/annual/retrospective	5	0.0 (-0.6 to 0.5)	0.947	0	0.722	5	-0.87 (-1.40 to -0.35)	0.001	82 %	<0.001
Other/biennial/prospective	2	-0.1 (-0.6 to 0.4)	0.741	32 %		2	0.09 (-0.67 to 0.85)	0.817	90 %	
Double reading	1	0.3 (-0.5 to 1.1)	0.466	—	0.299	1	0.51 (0.06–0.95)	0.025	—	<0.001
Single reading	6	-0.2 (-0.5 to 0.2)	0.345	0		6	-0.75 (-1.20 to -0.29)	0.001	88 %	
<i>SM/DBT vs DM</i>										
US/annual/retrospective	3	1.1 (0.1–2.1)	0.037	40 %	0.058	3	-3.39 (-4.53 to -2.26)	<0.01	95 %	<0.001
Other/biennial/prospective	8	2.3 (1.6–3.0)	<0.001	65 %		7	0.05 (-0.43 to 0.53)	0.839	93 %	
Double reading	6	2.8 (2.0–3.6)	<0.001	40 %	<0.001	6	-0.09 (-0.51 to 0.68)	0.773	94 %	<0.001
Single reading	5	1.4 (1.0–1.9)	<0.001	12 %		4	-2.56 (-4.60 to -0.52)	0.014	99 %	

CDR cancer detection rate, SM synthesized mammography, DBT digital breast tomosynthesis, DM digital mammography, CI confidence interval, RD risk difference, N number of study.

Furthermore, SM/DBT was associated with higher PPV-1 and PPV-2 compared with DM/DBT or DM. Our study is the first comprehensive head-to-head meta-analysis for clinical performance of SM/DBT in breast cancer screening.

All the paired trials included are European, while the unpaired studies including European, Australian, and American. Of the paired studies in our analysis, one double-reading study used an “either-positive” rule without arbitration for discordant reads [28], which increased false-positive recall. There are moderate or high heterogeneity in some analysis, such as recall rate and biopsy rate. Many factors like recall rules, equipment vendors, study designs, screening interval, reading practice, reader experience, dense breast proportion, and participant’s characteristics (e.g., age and race) may contribute to the heterogeneity. Five retrospective studies were undertaken in the US screening context, which is predominantly annual screening. The eight prospective studies were conducted in Europe or Australia, which were biennial screening. Only five biennial screening studies reported the proportion of incident (repeat) screening, most of the screens represented incident screening (over 80 %) [19,21,28,35,41]. We have performed sensitivity analysis stratifying by region/screening interval/study type and reading practice for primary outcomes to explore the potential heterogeneity, the results showed that they could account for partial heterogeneity. An interaction between CDR and reading practice was seen when comparing SM/DBT with DM, and there were significant interactions between recall rate and region/screening interval/study type or reading practice in two pairs of comparisons.

One RCT included in our meta-analysis showed that SM/DBT had a lower recall rate but similar CDR comparing with DM alone [19]. The DBT imaging equipment used in the RCT was from GE Healthcare, while most other studies used equipment from Hologic. Hologic, GE Healthcare, and Siemens are three vendors with approval for SM in the US and Europe. Due to that the SM image is created from the DBT dataset, vendor differences in DBT acquisition result in vendor-specific DBT images and subsequently varying synthesized results [15]. Most studies used Hologic Selenia Dimensions with improved version of the reconstructed synthesized image processing software (C-View) [28,30,33,38–41]. In addition, the radiologists in the RCT had insufficient experience in screen reading of DBT, so the radiologists may have not yet achieved optimal screen reading capabilities with DBT at the startup of the study [19].

Some studies have shown comparable diagnostic performance for the detection of cancer between SM/DBT and DM/DBT [11,45,46]. This suggests that SM may be used as an acceptable replacement for DM. Importantly, replacing DM with SM in DBT screening reduced the dose approximately by half to a level that was roughly comparable to that of DM alone, making DBT more widely available clinically [47]. Other strengths of SM include shorter acquisition time compared with a combined DM/DBT screening examination, and increased conspicuity of calcifications, spiculated margins, and architectural distortion [17]. Simultaneously, clinical radiologists should be aware that SM images may be lower resolution and increased noise compared with DM, and have the difficulty in assessing for motion artifact [16,17]. As the SM technology is rapidly evolving and upgrading, concerns over image quality would be alleviated [44].

Only two studies (OVVV and TSP) reported the ICR of SM/DBT, but there was no significant difference between SM/DBT and DM in pooled ICR. A recent meta-analysis including five prospective studies showed little difference between DBT/DM and DM in pooled ICR [48]. Although breast cancer screening using SM/DBT or DM/DBT improved the CDRs compared with DM alone, it could be inferred that the additional cancer detection possibly represent

over-diagnosis.

This systematic review and meta-analysis has several limitations. First, all but one included studies are nonrandomized. Although imbalance between cohorts is not a concern in paired trials, potential selection bias is possible in unpaired studies. Second, heterogeneity is moderate or high in some analyses, especially in recall rate. Third, long-term outcomes such as breast cancer mortality are not available due to the limitation of early application of the technology. Fourth, network meta-analysis for different imaging modalities is feasible and more preferably, but need more complex work, further research may consider this method.

## 5. Conclusion

In breast cancer screening, SM/DBT has similar detection for breast cancer, but higher PPV-1 and PPV-2 when compared with DM/DBT. The use of SM/DBT improves CDR, PPV-1 and PPV-2 compared with DM alone, but they has little different in ICR. SM/DBT could replace DM/DBT in breast cancer screening to reduce radiation dose.

## Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors. Hence, it does not need any human consent.

## Declaration of competing interest

The authors declare that they have no conflicts of interest.

## Acknowledgements

Fundings: No sources of funding are declared for this manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.breast.2021.07.016>.

## References

- [1] Hardesty LA, Kreidler SM, Glueck DH. Digital breast tomosynthesis utilization in the United States: a survey of physician members of the society of breast imaging. *J Am Coll Radiol* 2016;13(115):R67–73.
- [2] Richman IB, Hoag JR, Xu X, Forman HP, Hooley R, Busch SH, et al. Adoption of digital breast tomosynthesis in clinical practice. *JAMA Intern Med* 2019;179(9):1292–5.
- [3] Boroumand G, Teberian I, Parker L, Rao VM, Levin DC. Screening mammography and digital breast tomosynthesis: utilization updates. *Am J Roentgenol* 2018;210(5):1092–6.
- [4] Sardanelli F, Fallenberg EM, Clauser P, Trimboli RM, Camps-Herrero J, Helbich TH, et al. Mammography: an update of the EUSOBI recommendations on information for women. *Insights imaging* 2017;8(1):11–8.
- [5] Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. *J Am Med Assoc* 2014;311(24):2499–507.
- [6] Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol* 2013;14(7):583–9.
- [7] Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fantò C, Ostilio L, 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):1105–13.
- [8] Skaane P, Bandos AI, Eben EB, Jepsen IN, Krager M, Haakenaasen U, 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):655–63.



- [9] Marinovich ML, Hunter KE, Macaskill P, Houssami N. Breast cancer screening using tomosynthesis or mammography: a meta-analysis of cancer detection and recall. *JNCI-J Natl Cancer I* 2018;110(9):942–9.
- [10] Rafferty EA, Park JM, Philpotts LE, Poplack SP, Sumkin JH, Halpern EF, et al. Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. *Radiology* 2013;266(1):104–13.
- [11] Gilbert FJ, Tucker L, Gillan MGC, Willsher P, Cooke J, Duncan KA, et al. Accuracy of digital breast tomosynthesis for depicting breast cancer subgroups in a UK retrospective reading study (tommy trial)1. *Radiology* 2015;277(3):697–706.
- [12] Houssami N. Evidence on synthesized two-dimensional mammography versus digital mammography when using tomosynthesis (Three-dimensional mammography) for population breast cancer screening. *Clin Breast Canc* 2018;18(4):255–60.
- [13] Gur D, Zuley ML, Anello MI, Rathfon GY, Chough DM, Ganott MA, et al. Dose reduction in digital breast tomosynthesis (DBT) screening using synthetically reconstructed projection images: an observer performance study. *Acad Radiol* 2012;19(2):166–71.
- [14] Zuckerman SP, Conant EF, Keller BM, Maidment AD, Barufaldi B, Weinstein SP, et al. Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. *Radiology* 2016;281(3):730–6.
- [15] Durand MA. Synthesized mammography: clinical evidence, appearance, and implementation. *Diagnostics* 2018;8(2):15.
- [16] Freer PE, Winkler N. Synthesized digital mammography imaging. *Radiol Clin* 2017;55(3):503–12.
- [17] Ratanaprasatporn L, Chikarmane SA, Giess CS. Strengths and weaknesses of synthetic mammography in screening. *Radiographics* 2017;37(7):1913–27.
- [18] Michell MJ, Batohi B. Role of tomosynthesis in breast imaging going forward. *Clin Radiol* 2018;73(4):358–71.
- [19] Hofvind S, Holen AS, Aase HS, Houssami N, Sebuodegard S, Moger TA, 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.
- [20] Skaane P, Bandos AI, Niklason LT, Sebuodegard S, Osteras BH, Gullien R, et al. Digital mammography versus digital mammography plus tomosynthesis in breast cancer screening: the Oslo tomosynthesis screening trial. *Radiology* 2019;291(1):23–30.
- [21] Caumo F, Zorzi M, Brunelli S, Romanucci G, Rella R, Cugola L, 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.
- [22] Hofvind S, Hovda T, Holen AS, Lee CI, Albertsen J, Bjørndal H, et al. Digital breast tomosynthesis and synthetic 2D mammography versus digital mammography: evaluation in a population-based screening program. *Radiology* 2018;287(3):787–94.
- [23] Alabousi M, Wadera A, Kashif Al-Ghita M, Al-Ghetaa RK, Salameh JP, Pozdnyakov A, et al. Performance of digital breast tomosynthesis, synthetic mammography and digital mammography in breast cancer screening: a systematic review and meta-analysis. *JNCI-J Natl Cancer I* 2021;113(6):680–90.
- [24] Giampietro RR, Cabral MVG, Lima SAM, Weber SAT, Dos Santos Nunes-Nogueira V. Accuracy and effectiveness of mammography versus mammography and tomosynthesis for population-based breast cancer screening: a systematic review and meta-analysis. *Sci Rep* 2020;10:7991.
- [25] Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155(8):529–36.
- [26] DerSimonian R, Laird N. Meta-analysis in clinical trials. *Contr Clin Trials* 1986;7(3):177–88.
- [27] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557–60.
- [28] Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fanto C, Ostilio L, 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):1105–13.
- [29] Bernardi D, Li T, Pellegrini M, Macaskill P, Valentini M, Fanto C, 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.
- [30] Martin SR, Povedano JLR, Garcia MC, Romero ALS, Garriguet MP, Benito MA. Prospective study aiming to compare 2D mammography and tomosynthesis plus synthesized mammography in terms of cancer detection and recall. From double reading of 2D mammography to single reading of tomosynthesis. *Eur Radiol* 2018;28(6):2484–91.
- [31] Osteras 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):60–8.
- [32] Skaane P, Sebuodegard S, Bandos AI, Gur D, Osteras BH, Gullien R, et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. *Breast Canc Res Treat* 2018;169(3):489–96.
- [33] Skaane P, Bandos AI, Eben EB, Jepsen IN, Krager M, Haakenaasen U, 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):655–63.
- [34] Caumo F, Romanucci G, Hunter K, Zorzi M, Brunelli S, Macaskill P, 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 Canc Res Treat* 2018;170(2):391–7.
- [35] Bernardi D, Gentilini MA, De Nisi M, Pellegrini M, Fanto C, Valentini 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;50:135–40.
- [36] Aase HS, Holen AS, Pedersen K, Houssami N, Haldorsen IS, Sebuodegard S, et al. A randomized controlled trial of digital breast tomosynthesis versus digital mammography in population-based screening in Bergen: interim analysis of performance indicators from the To-Be trial. *Eur Radiol* 2019;29(3):1175–86.
- [37] Hovda T, Holen AS, Lång K, Albertsen JL, Bjørndal H, Brandal SHB, 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):256–64.
- [38] Aujero 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):70–6.
- [39] Freer PE, Riegert J, Eisenmenger L, Ose D, Winkler N, Stein MA, et al. Clinical implementation of synthesized mammography with digital breast tomosynthesis in a routine clinical practice. *Breast Canc Res Treat* 2017;166(2):501–9.
- [40] 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):973–6.
- [41] Houssami N, Lockie D, Clemson M, Pridmore V, Taylor D, Marr G, et al. Pilot trial of digital breast tomosynthesis (3D mammography) for population-based screening in BreastScreen Victoria. *MJA (Med J Aust)* 2019;211(8):357–62.
- [42] Zuckerman SP, Sprague BL, Weaver DL, Herschorn SD, Conant EF. Multicenter evaluation of breast cancer screening with digital breast tomosynthesis in combination with synthetic versus digital mammography. *Radiology* 2020;297(3):545–53.
- [43] Cohen EO, Weaver OO, Tso HH, Gerlach KE, Leung JWT. Breast cancer screening via digital mammography, synthetic mammography, and tomosynthesis. *Am J Prev Med* 2020;58(3):470–2.
- [44] Zuckerman SP, Sprague BL, Weaver DL, Herschorn SD, Conant EF. Survey results regarding uptake and impact of synthetic digital mammography with tomosynthesis in the screening setting. *J Am Coll Radiol* 2020;17(1 Pt A):31–7.
- [45] Zuley ML, Guo B, Catullo VJ, Chough DM, Kelly AE, Lu AH, et al. Comparison of two-dimensional synthesized mammograms versus original digital mammograms alone and in combination with tomosynthesis images. *Radiology* 2014;271(3):664–71.
- [46] Abdullah P, Alabousi M, Ramadan S, Zawawi I, Zawawi M, Bhogadi Y, et al. Synthetic 2D mammography versus standard 2D digital mammography: a diagnostic test accuracy systematic review and meta-analysis. *Am J Roentgenol* 2020;217(2):314–25.
- [47] 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):93–9.
- [48] Houssami N, Zackrisson S, Blazek K, Hunter K, Bernardi D, Lång K, et al. Meta-analysis of prospective studies evaluating breast cancer detection and interval cancer rates for digital breast tomosynthesis versus mammography population screening. *Eur J Canc* 2021;148:14–23.