Check for updates

OPEN

Accuracy and Effectiveness of Mammography versus Mammography and Tomosynthesis for Population-Based Breast Cancer Screening: A Systematic Review and Meta-Analysis

Rodrigo Rosa Giampietro[1], Marcos Vinicius Gama Cabral[1], Silvana Andrea Molina Lima[2,3], Silke Anna Theresa Weber[3,4] & Vania dos Santos Nunes-Nogueira [1,3] ✉

We proposed to compare the accuracy and effectiveness of digital breast tomosynthesis (DBT), plus digital or synthetic mammography, with digital mammography alone in women attending population-based breast cancer screenings. We performed a systematic review and included controlled studies comparing DBT with digital mammography for breast cancer screening. Search strategies were applied to the MEDLINE, Embase, LILACS, and CENTRAL databases. With moderate quality of evidence, in 1,000 screens, DBT plus digital mammography increased the overall and invasive breast cancer rates by 3 and 2 (RR 1.36, 95% CI 1.18 to 1.58 and RR 1.51, 95% CI 1.27 to 1.79, respectively). DBT plus synthetic mammography increased both overall and invasive breast cancer rates by 2 (RR 1.38, 95% CI 1.24 to 1.54 and RR 1.37, 95% CI 1.22 to 1.55, respectively). DBT did not improve recall, false positive and false negative rates. However due to heterogeneity the quality of evidence was low. For women attending population-based breast cancer screenings, DBT increases rates of overall and invasive breast cancer. There is no evidence with high or moderate quality showing that DBT compared with digital mammography decreases recall rates, as well as false positive and false negative rates.

Breast cancer is one of the most frequently diagnosed cancers among women, and population-based breast cancer screenings with mammography have been one of the worldwide health strategies to reduce breast cancer mortality[1].

Technological advances in image acquisition provided the transition from film screen to digital mammography. In more recent years, as an advancement from mammography, digital breast tomosynthesis (DBT) has been introduced into screening practices which has the potential to overcome limitations of digital mammography[2]. Most diagnostic centres perform DBT with digital mammography. However some software can synthesize digital mammography images (synthetic mammography) from data acquired during DBT, thus reducing the radiation dose[3].

Several studies have shown that adding DBT to digital mammography significantly increases the detection of breast cancer[4–6]. However, results from previous studies regarding recall rates are inconsistent; some studies have shown reduction in false recalls[7,8], while others have shown that the proportion of women recalled for further assessment has increased[9,10].

[1]Department of Internal Medicine, São Paulo State University/UNESP, Medical School, Botucatu, Sao Paulo, Brazil. [2]Department of Nursing, São Paulo State University/UNESP, Medical School, Botucatu, Sao Paulo, Brazil. [3]Health Technology Assessment Nucleus, Botucatu Medical School Clinical Hospital, Sao Paulo, Brazil. [4]Ophthalmology, Otorhinolaryngology and Head & Neck Surgery Department, São Paulo State University/UNESP, Medical School, Botucatu, Sao Paulo, Brazil. ✉e-mail: vania.nunes-nogueira@unesp.br

Although there are several systematic reviews on this topic in the literature, none have included DBT with synthetic mammography in their analysis[11–23]. In addition, at least three clinical trials evaluating the effectiveness of DBT on breast cancer screening have been reported since these reviews were published[24–26].

Thus, we proposed to evaluate the accuracy and effectiveness of DBT (with digital mammography or synthetic mammography) compared to digital mammography alone in women with a standard risk for developing this neoplasia, who attended population-based breast cancer screenings.

Methods

A systematic review was conducted according to the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy[27] and was reported on according to the PRISMA- Diagnostic Test Accuracy Studies (DTAs) Statements[28,29]. Our protocol was registered in the International Prospective Registry of Systematic Reviews, under the ID, CRD42017070890.

Eligibility criteria. *Type of studies.* We included randomized (RCT) and quasi-randomized controlled trials (quasi-RCT), cohort studies, and diagnostic test accuracy studies (cross-sectional studies involving patients who received mammography and DBT, and in which screen-reading was performed in two sequential phases, mammography only versus mammography integrated with DBT). The included studies followed the PICO protocol described below:

Patients (P). We included studies involving women, over 45 years of age and with no breast cancer related symptoms, from among a population with a standard risk of developing breast cancer, who attended population-based breast cancer screenings.

Index test (I). We considered DBT, either with digital mammography or synthetic mammography, as the index test.

Comparison (C). We considered digital mammography alone as the comparison test.

Types of outcome measures (O) of the included studies. Primary outcomes were overall and breast cancer mortalities, overall invasive and non-invasive breast cancer detection rates, proportion of women recalled for additional examinations (recall rate), adverse events, and irradiation dose per examination.

Secondary outcomes were the true positive, false positive, false negative, and true negative rates. If such data were available, the accuracy of each index test was calculated (sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, and diagnostic odds ratios).

Reference test. As a reference test to confirm the positive cases of breast cancer, we considered the results of histological tests conducted after surgery or by biopsy. To confirm the negative cases, we considered the absence of breast cancer detected via examinations during a follow-up period.

Exclusion criteria. We excluded studies in which participants consisted of women with established risk factors for breast cancer, and studies in which most participants were already diagnosed with some type breast disease or were called for additional examinations.

Further, we excluded studies in which the index and comparator tests were performed at different times.

For studies that met the eligibility criteria but also included women who were under 45 years old, an e-mail was sent to the corresponding author requesting the outcome data for patients over 45 years old. Studies that did not provide this information were included if most of the sample comprised of women aged according to our eligibility criteria.

Search methods for identification of studies. Four general and adaptive search strategies were created for the electronic databases: Embase (1980-01/March/2020), PubMed (1966-01/March/2020), LILACS (1982-01/March/2020), and CENTRAL (Cochrane Collaboration Controlled Trials Registry-01/March/2020) (Supplementary File). The mesh terms—breast cancer and DBT— were used to construct each search strategy; there were no language or year restrictions (Supplementary File).

Additionally, we surveyed the Trip Medical Database, SCOPUS, Web of Science, and CINAHL. Furthermore, we searched thesis banks for unpublished studies and ClinicalTrials.gov for ongoing studies.

We used the Endnote software to download references, remove duplicates, and facilitate the selection process.

Selection of studies. Two reviewers independently (VSNN and RRG) selected titles and abstracts from the ones identified via the bibliographic research. Potentially eligible studies were selected for a full reading and, subsequently, evaluated for conformance to the proposed PICO. In case of disagreements during the selection process, we arrived at a consensus via discussions. The reasons behind each excluded study were justified.

Data extraction and management. Both reviewers applied a data extraction form to the studies to compute the corresponding participant-related information.

Risk of bias and applicability. We evaluated the risk of bias corresponding to the included studies via the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool[30].

Unit of analysis. The unit of analysis was the aggregated data extracted from the journal publications.
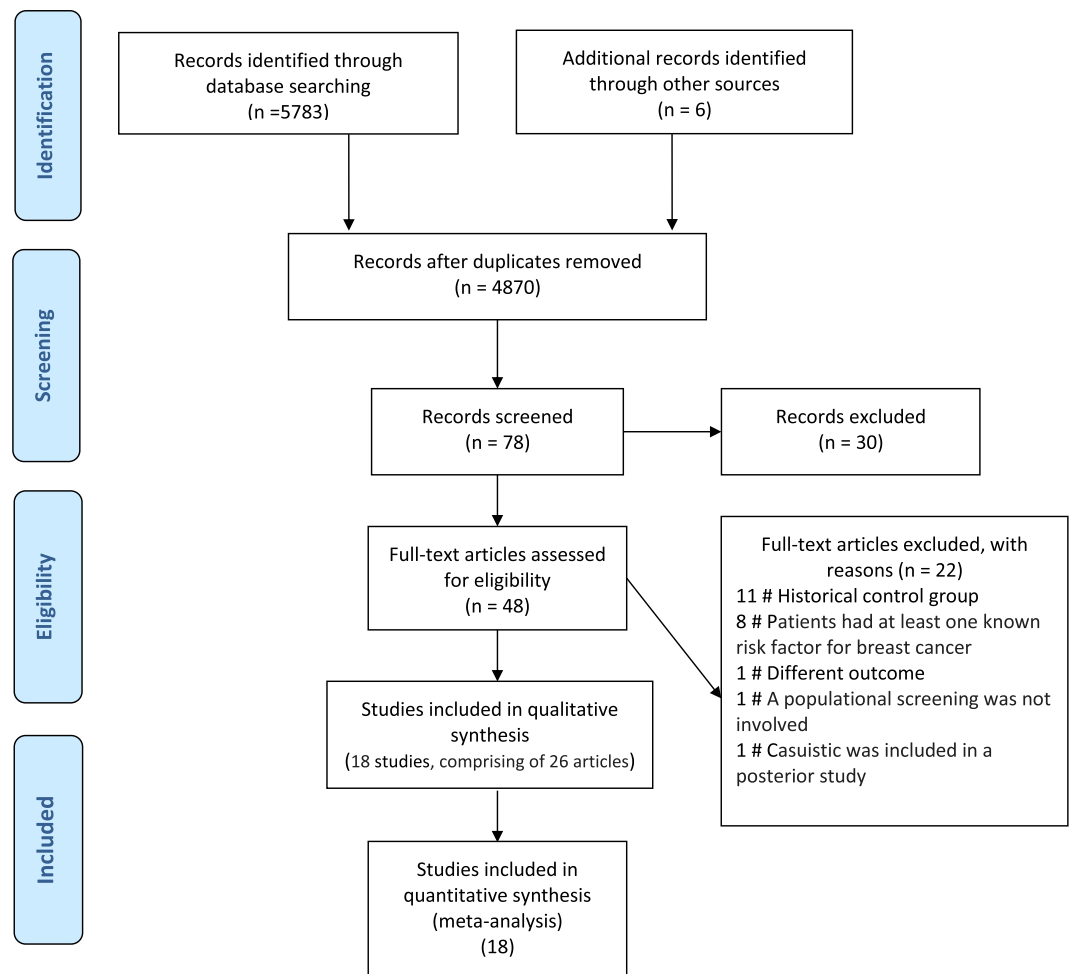
Figure 1. Flow of selection of articles for the systematic review.

Assessment of heterogeneity.   Inconsistencies among the study results were ascertained by visually inspecting a forest plot and with the Higgins or I^2 statistic, in which an $I^2 > 50\%$ indicated a moderate probability of heterogeneity.

Synthesis of results (Meta-analysis).   Similar outcomes, measured in at least two trials, were plotted in the meta-analysis using Review Manager 5.3 (Review Manager. [RevMan], version 5.3, Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). For dichotomic outcomes, the relative risk (RR) was calculated with a 95% confidence interval (CI) as an effect size of the effectiveness of the index test. We selected the random effects model for the meta-analysis, and the studies were evaluated separately according to their designs.

Grading the quality of evidence.   For each outcome, a tabulated summary of the findings was produced in order to report the effectiveness of the index test. The certainty of the evidence was measured using the GRADE approach (Grading of Recommendations Assessment, Development, and Evaluation Working Group)[31].

Ethical standards.   As no primary data collection was undertaken, no formal ethical assessment is required in our institution.

Results
Study selection.   The search strategies yielded 5,783 references, and after removing duplicates, 4,870 studies remained. We selected 48 studies that had a high probability of meeting our inclusion criteria for a complete reading (Fig. 1).

After completely reading these references, 18 studies (comprising 26 articles, since some had more than one published article) met our eligibility criteria and therefore were included in this review[4–8,24–26,32–49].

A total of 22 studies were excluded for the following reasons: 11 used a historic cohort as control;[50–60] eight studies included patients that had at least one known risk factor for breast cancer or were invited to perform additional tests due to alterations in previous examinations;[61–68] in one study the outcome evaluated was interpretation time of screening examinations[69], in one study a populational screening was not involved[70], and in one study the patients included were also included in a posterior study[71].

Study characteristics.   Only two trials explicitly include asymptomatic women at a standard population risk for breast cancer[7,35]. In the other studies, the eligibility criterion was women who attended population screening programs. Therefore, it is inferred that most participants were asymptomatic and at a population standard risk.

Fourteen studies evaluated DBT in combination with digital mammography versus digital mammography alone. These studies included: one RCT[24], five accuracy studies[7,34,35,40,44], and eight retrospective cohort studies[4–6,8,36–38,46]. Six studies, three accuracy studies[35,43,48], one prospective cohort study[47], one RCT[25] and one quasi-RCT[26] evaluated the effectiveness of DBT with synthetic mammography versus digital mammography alone.

In all included studies, the radiologists were experienced in breast imaging and had received trainings on DBT. Three studies had an eligibility criterion of including women older than 40 years, two of women older than 45 years, and the remaining included women older than 50.

Table X of supplementary data presents descriptive data of all the included studies.

Risk of bias and applicability.   Figure 2 shows the risk of bias corresponding to the included studies. Most retrospective cohort studies were assessed as having a high probability of bias in patient selection (the DBT group had more risk factors for breast cancer). The studies involving patients under 45 years old were deemed to have an uncertain risk regarding the applicability of the patient selection. All studies were evaluated as having a high risk of bias in the reference test, since the pathologists who evaluated the biopsies and pathological results had prior knowledge of the screening tests. Follow-ups were also evaluated as having a high risk of bias, since patients who were not recalled missed the reference test.

Meta-analysis DBT plus digital mammography versus digital mammography alone.   *Breast Cancer Detection Rate (Fig. 3 to Fig. 7, Supplementary File).*   Based on RCT and accuracy studies, with a moderate quality of evidence and in 1,000 screened women, DBT plus digital mammography increased the overall breast cancer rates by 3 (RR 1.36, 95% CI 1.18 to 1.58, Table 1), and the rate of invasive breast cancer detection was increased by 2 (RR 1.51, 95% CI 1.27 to 1.79, Table 1). Regarding the rate of ductal carcinoma *in situ*, there was no clear evidence to support a difference between the two interventions (RR 1.26 95% CI 0.86 to 1.83).

Based on retrospective cohort studies, the rates of overall, invasive and ductal breast cancer are very similar to those of RCT and accuracy studies. However, the certainty of the evidence was lower due to the fact that women in the DBT group had more risk factors for breast cancer than those in the digital mammography alone group.

Recall Rate (Fig. 8 and Fig. 9, Supplementary File).   RCT and accuracy studies with DBT plus digital mammography did not reveal differences in recall rates compared to those with digital mammography alone (RR 1.13, 95% CI 0.96 to 1.32, Table 1). However, due to serious inconsistencies (DBT increased, decreased, and did not change the recall rates among different studies) the certainty of evidence was low. Due to very serious inconsistencies among retrospective cohort studies there was no clear effect of DBT on this outcome.

Positive Predictive Value (PPV) and False Positive Recalls Rate (Fig. 10 and Fig. 11, Supplementary File).   Based on RCT and accuracy studies, the effects of DBT plus digital mammography on false positive recalls and in the PPV for breast cancer were not different between the groups, however the quality of evidence was low due to imprecision and inconsistencies in the meta-analyses. The same occurred with the retrospective cohort studies.

DBT plus synthetic mammography versus digital mammography alone.   *Breast Cancer Detection Rate (Fig. 12 to Fig. 14, Supplementary File).*   With a moderate quality of evidence, and with 1,000 women screened, DBT plus synthetic mammography increased the overall and invasive breast cancer rates by 2 (RR 1.38, 95% CI 1.24 to 1.54 and RR 1.37, 95% CI 1.22 to 1.55, respectively, Table 1). The ductal breast cancer rates were marginally higher for DBT, but this difference was not statistically significant (RR 1.41, 95% CI 0.94 to 2.11).

Recall Rate (Fig. 15, Supplementary File).   DBT plus synthetic mammography results in no differences in recall rates (RR 1.08, 95% CI 0.92 to 1.26, Table 1). However, due to serious inconsistencies (recall rates increased, decreased, and did not change among the studies) the quality of evidence was low.

Positive Predictive Value and False Positive Recalls (Fig. 16 and Fig. 17, Supplementary File).   The effects of DBT plus synthetic mammography on false positive recalls for breast cancer were not different between the groups. However, the quality of evidence was low due to imprecision and inconsistencies in the meta-analyses,

Conversely, regarding patients recalled for additional assessment, DBT plus synthetic mammography resulted in little increase in the positive predictive value for breast cancer (RR 1.26, 95% CI 1.09 to 1.46), but due to serious imprecisions the quality of evidence was low.

Additional analysis.   The assessment of the accuracy of DBT (with digital or synthetic mammography versus digital mammography alone) could not be verified from the 2×2 contingency table data because it was impossible to confirm true and false negatives in all studies included.

None of the included studies evaluated overall or breast cancer mortalities or adverse events associated with DBT plus digital or synthetic mammography.

Regarding false negative rates, STORM (DBT plus digital mammography) was the only study that evaluated this outcome. In this accuracy study, the authors estimated the interval cancer rate at two-year follow-up and compared this result with a concurrent group of women who had attended the same screening services and received only digital mammography. The interval breast cancer rate in the STORM trial was not statistically different from that estimated amongst women screened with digital mammography (9/7292 screens versus

Figure 2. Risk of bias assessment according to QUADAS 2.

40/25,058 screens, respectively, RR 0,77, 95% CI 0.38 to 1.59), however the quality of evidence was low (wide confidence interval)[32].

Only two studies presented the radiation dose per examination, Pattaccini et al. and Hofvind et al. (To-Be study), interventions which used DBT plus digital mammography and DBT plus synthetic mammography, respectively[24,25]. In the first study the median radiation dose per examination was 6.40 mGy (IQR, 5.68.–7.36 mGy) for DBT plus digital mammography and 4.84 mGy (IQR, 4.24–5.72 mGy) for digital mammography alone. In the second study, the mean radiation dose per examination was 2.96 mGy for DBT with synthetic mammography and 2.95 mGy for digital mammography alone. The remaining controlled studies only stated that the radiation dose levels of DBT plus digital mammography were approximately twice of those of digital mammography alone.

Ongoing studies. There are two important clinical trials which currently in the recruitment phase. The first one is the Tomosynthesis Mammographic Imaging Screening Trial (TMIST)[72]. In this study, which is taking place in the United States, women aged 45 to 75 and attending a populational-based breast screening will be

	DBT plus either conventional digital mammography or synthetic mammography compared to conventional digital mammography alone in women attending population-based breast cancer screenings					
	Patient: Women attending population-based breast cancer screenings Setting: Population-based screening programs Intervention: DBT plus either conventional digital mammography or synthetic mammography Comparison: Conventional digital mammography alone					
Outcomes	**Anticipated absolute effects* (95% CI)**		**Relative effect (95% CI)**	**№ of participants (studies)**	**Certainty of the evidence (GRADE)**	**Comments**
	Risk with DM alone	**Risk with DBT + DM or SM**				
Breast cancer detection rate from RCT and DAT studies - **DBT + DM**	7 per 1.000	**10 per 1.000**(8 to 11)	**RR 1.36**(1.18 to 1.58)	58265(1 RCT, 5 DTAs)	⊕⊕⊕◯MODERATE [a]	**DBT plus DM** likely increases breast cancer detection rate slightly.
Invasive breast cancer detection rate from RCT and DAT studies- **DBT + DM**	5 per 1.000	**7 per 1.000**(6 to 8)	**RR 1.51**(1.27 to 1.79)	56650(1 RCT, 4 DTAs)	⊕⊕⊕◯MODERATE [a]	**DBT plus DM** likely increases invasive breast cancer detection rate.
Recall Rate from RCT and DAT studies- **DBT + DM**	34 per 1.000	**38 per 1.000**(32 to 45)	**RR 1.13**(0.96 to 1.32)	58265(1 RCT, 5 DTAs)	⊕⊕◯◯LOW [a, b, c]	**DBT + DM** may increase/ have little to no effect on recall Rate DBT + DM but the evidence is very uncertain.
Breast Cancer Detection Rate - **DBT + SM**	6 per 1.000	**8 per 1.000**(7 to 9)	**RR 1.38**(1.24 to 1.54)	175572(2 RCT, 3 DTAs, 1 PC)	⊕⊕⊕◯MODERATE [a]	**DBT + SM** likely increases breast cancer detection rate.
Recall Rate - **DBT + SM**	33 per 1.000	**35 per 1.000**(30 to 41)	**RR 1.08**(0.92 to 1.26)	175572(2 RCT, 3 DTAs, 1 PC)	⊕⊕◯◯LOW [a, d]	**DBT + SM** may result in no difference in recall rate.
Invasive breast cancer detection rate - **DBT + SM**	5 per 1.000	**7 per 1.000**(6 to 8)	**RR 1.37**(1.22 to 1.55)	163604(2 RCT, 2 DTAs, 1 PC)	⊕⊕⊕◯MODERATE [a]	**DBT + SM** likely increases the rate of invasive breast cancer.
*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).**DBT**: Digital breast tomosynthesis**; DM:** Digital Mammography**; SM:** Synthetic mammography**, RCT:** Randomized Clinical Trial; **DAT:** Diagnostic test accuracy study; **PC:** Prospective cohort; **CI:** Confidence interval; **RR:** Risk ratio						
GRADE Working Group grades of evidenceHigh certainty: We are very confident that the true effect lies close to that of the estimate of the effect**Moderate certainty:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different**Low certainty:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect**Very low certainty:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect						

Table 1. Summary of findings from the main comparisons. Explanations a. The pathologists who evaluated the biopsies and pathological results had prior knowledge of the screening tests. Follow-ups were also evaluated as a high risk of bias, since patients who were not recalled missed the reference test b. Wide confidence interval c. In three studies DBT plus DM increased the recall rates, in two studies there was no difference between the groups d. In three studies DBT plus SM did not show difference between the groups, in two studies it increased the recall rate, and in one study DBT plus SM decreased the recall rate.

randomized to DBT or digital mammography. The researchers plan to enrol nearly 165,000 patients, and the primary outcome is the proportion of women diagnosed with advanced breast cancer at any time during a period of 4.5 years from randomization, including the period of active screening and a period of follow up after the last screen (time frame: 4.5 years after registration).

Another important ongoing clinical trial in this topic is called Digital Breast Tomosynthesis plus Synthesised Images versus Standard Full-Field Digital Mammography in Population-Based Screening (TOSYMA), which is being carried out in Germany. The authors aim to include 80,000 women aged 50 to 69 years who are attending their routine mammography screening programme[73]. The primary endpoints are the detection rate of invasive breast cancers during screening examinations and the cumulative incidence of interval cancers in the two years after a negative examination.

Discussion

In order to present the best available evidence to help clinicians with decision making, we conducted a systematic review. The aim of this review was to compare the effect of DBT with digital mammography in over 45 year-old women attending a routine screening mammogram programme. Eighteen studies were included in this review. Our results show, with a moderate quality of evidence, that implementing DBT plus digital or synthetic mammography in population-based breast cancer screening increases overall breast cancer detection rates, as well as invasive breast cancer detection rates.

Although some studies have shown lower recall and false positives with DBT[6–8], this was not confirmed in the present review. Our analyses did not find evidence for differences in recall rates between DBT (with digital or synthetic mammography) and digital mammography alone. However, due to the high heterogeneity between the results of the included studies, the quality of evidence was low.

In the context of breast cancer screening, a false negative finding can have devastating implications for the woman concerned, since a delay in cancer diagnosis can lead to an unfavourable evolution. In this review we did not find evidence of lower rates of false negatives with DBT. Conant *et al.*, who compared the results of DBT with a historic cohort of digital mammography, evaluated the proportion of negative examinations in which cancer

was diagnosed within 1 year. Results showed that the false-negative rates were slightly lower for DBT, but this difference was not statistically significant[74]. With the same study design Bahl *et al*. showed that the rate of interval cancers was similar with DM and DBT[75].

We found 13 systematic reviews on this topic published in the literature[11–23]. Only one of these reviews had its protocol registered[15], and all of their eligibility criteria were different from ours. Most reviews indicated that DBT with digital mammography was more effective, as it resulted in greater overall breast cancer detection and fewer false positives. However, none of them evaluated the quality of the evidence according to GRADE or included RCTs and prospective cohort studies in their analyses[24,47].

Our systematic review had some limitations, the main one being related to the fact that none of the included studies evaluated the effects of DBT on improving breast cancer-related mortality, morbidity and quality of life. In a population-based cancer screening, besides the early cancer diagnosis, we sought to analyse the damage inflicted by these programs, including overdiagnosis and overtreatment at a very early stage of the disease. The Cochrane review estimated that for every 2,000 women invited to a mammography screening over a period of 10 years, one would have a long life, ten healthy women would suffer from overdiagnosis and overtreatment, and 200 women would suffer psychological damage due to false positive results[76]. Further, it has been estimated that breast cancer does not become symptomatic or health threatening in the lifetime of 1% to 10% of women with a positive diagnosis[77]. Additionally, it is estimated that overtreatment causes lifelong chronic pain in half of overdiagnosed women[76].

Conclusion

Implications for clinical practice.   Use of DBT with digital or synthetic mammography for women attending population-based breast cancer screenings increases the rates of overall and invasive breast cancer detection. There is no evidence, with high or moderate quality, showing that DBT, compared with digital mammography, decreases recall rates, as well as false positive and false negative rates.

Implications for future research.   Longitudinal studies are necessary to evaluate the effects of DBT on improving important patient outcomes (i.e. mortality, morbidity, test procedure complications, resource utilization, and quality of life).

Data availability

All data generated or analysed during this systematic review are included in this published article (and its Supplementary File).

References

1. Coldman, A. *et al*. Pan-Canadian study of mammography screening and mortality from breast cancer. *Journal of the National Cancer Institute* **106**, https://doi.org/10.1093/jnci/dju261 (2014).
2. Michell, M. J. *et al*. A comparison of the accuracy of film-screen mammography, full-field digital mammography, and digital breast tomosynthesis. *Clinical Radiology* **67**, 976–981, https://doi.org/10.1016/j.crad.2012.03.009 (2012).
3. Alakhras, M. *et al*. Digital tomosynthesis: A new future for breast imaging? *Clinical Radiology* **68**, e225–e236, https://doi.org/10.1016/j.crad.2013.01.007 (2013).
4. Conant, E. F. *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 research and treatment* **156**, 109–116, https://doi.org/10.1007/s10549-016-3695-1 (2016).
5. Giess, C. S. *et al*. Comparing Diagnostic Performance of Digital Breast Tomosynthesis and Full-Field Digital Mammography in a Hybrid Screening Environment. *AJR. American journal of roentgenology*, 1–6, https://doi.org/10.2214/ajr.17.17983 (2017).
6. Alsheik, N. H. *et al*. Comparison of Resource Utilization and Clinical Outcomes Following Screening with Digital Breast Tomosynthesis Versus Digital Mammography: Findings From a Learning Health System. *Academic radiology* **26**, 597–605, https://doi.org/10.1016/j.acra.2018.05.026 (2019).
7. Ciatto, S. *et al*. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): A prospective comparison study. *Lancet oncology* **14**, 583–589 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/664/CN-00916664/frame.html (2013).
8. Cohen, E. O., Tso, H. H., Phalak, K. A., Mayo, R. C. & Leung, J. W. T. Screening Mammography Findings From One Standard Projection Only in the Era of Full-Field Digital Mammography and Digital Breast Tomosynthesis. *AJR. American journal of roentgenology*, 1–7, https://doi.org/10.2214/ajr.17.19023 (2018).
9. Bernardi, D. *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. *The Lancet. Oncology* **17**, 1105–1113, https://doi.org/10.1016/S1470-2045(16)30101-2 (2016).
10. Skaane, P. *et al*. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. *European radiology* **23**, 2061–2071, https://doi.org/10.1007/s00330-013-2820-3 (2013).
11. Blue, C., Blue Shield, A., Kaiser Foundation Health, P. & Southern California Permanente Medical, G. Use of digital breast tomosynthesis with mammography for breast cancer screening or diagnosis. *Technology Evaluation Center Assessment Program. Executive summary* **28**, 1–6 (2014).
12. Coop, P., Cowling, C. & Lawson, C. Tomosynthesis as a screening tool for breast cancer: A systematic review. *Radiography* **22**, e190–e195, https://doi.org/10.1016/j.radi.2016.03.002 (2016).
13. Svahn, T. M. & Houssami, N. Digital breast tomosynthesis in one or two views as a replacement or adjunct technique to full-field digital mammography. *Radiation protection dosimetry* **165**, 314–320, https://doi.org/10.1093/rpd/ncv078 (2015).
14. Svahn, T. M., 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* **24**, 93–99, https://doi.org/10.1016/j.breast.2014.12.002 (2015).
15. Hodgson, R. *et al*. Systematic review of 3D mammography for breast cancer screening. *Breast (Edinburgh, Scotland)* **27**, 52–61, https://doi.org/10.1016/j.breast.2016.01.002 (2016).

16. Houssami, N. *et al*. Digital breast tomosynthesis (3D-mammography) screening: A pictorial review of screen-detected cancers and false recalls attributed to tomosynthesis in prospective screening trials. *Breast (Edinburgh, Scotland)* **26**, 119–134, https://doi.org/10.1016/j.breast.2016.01.007 (2016).

17. Lei, J., Yang, P., Zhang, L., Wang, Y. & Yang, K. Diagnostic accuracy of digital breast tomosynthesis versus digital mammography for benign and malignant lesions in breasts: A meta-analysis. *European radiology* **24**, 595–602, https://doi.org/10.1007/s00330-013-3012-x (2014).

18. Garcia-Leon, F. J., Llanos-Mendez, A. & Isabel-Gomez, R. Digital tomosynthesis in breast cancer: A systematic review. *Radiologia* **57**, 333–343, https://doi.org/10.1016/j.rx.2014.06.006 (2015).

19. Pozz, A., Corte, A. D., Lakis, M. A. & Jeong, H. Digital Breast Tomosynthesis in Addition to Conventional 2DMammography Reduces Recall Rates and is CostEffective. *Asian Pac J Cancer Prev* **17**, 3521–3526 (2016).

20. Yun, S. J., Ryu, C. W., Rhee, S. J., Ryu, J. K. & Oh, J. Y. Benefit of adding digital breast tomosynthesis to digital mammography for breast cancer screening focused on cancer characteristics: a meta-analysis. *Breast Cancer Res Treat* **164**, 557–569, https://doi.org/10.1007/s10549-017-4298-1 (2017).

21. Skaane, P. Breast cancer screening with digital breast tomosynthesis. *Breast Cancer* **24**, 32–41, https://doi.org/10.1007/s12282-016-0699-y (2017).

22. Alabousi, M. *et al*. Digital breast tomosynthesis for breast cancer detection: a diagnostic test accuracy systematic review and meta-analysis. *European radiology* **30**, 2058–2071, https://doi.org/10.1007/s00330-019-06549-2 (2020).

23. Marinovich, M. L., Hunter, K. E., Macaskill, P. & Houssami, N. Breast Cancer Screening Using Tomosynthesis or Mammography: A Meta-analysis of Cancer Detection and Recall. *Journal of the National Cancer Institute* **110**, 942–949, https://doi.org/10.1093/jnci/djy121 (2018).

24. Pattacini, P. *et al*. Digital Mammography versus Digital Mammography Plus Tomosynthesis for Breast Cancer Screening: The Reggio Emilia Tomosynthesis Randomized Trial. *Radiology* **288**, 375–385, https://doi.org/10.1148/radiol.2018172119 (2018).

25. Hofvind, S. *et al*. Two-view digital breast tomosynthesis versus digital mammography in a population-based breast cancer screening programme (To-Be): a randomised, controlled trial. *The Lancet. Oncology* **20**, 795–805, https://doi.org/10.1016/S1470-2045(19)30161-5 (2019).

26. Houssami, N. *et al*. Pilot trial of digital breast tomosynthesis (3D mammography) for population-based screening in BreastScreen Victoria. *The Medical journal of Australia* **211**, 357–362, https://doi.org/10.5694/mja2.50320 (2019).

27. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. https://methods.cochrane.org/sdt/handbook-dta-reviews (2020).

28. McInnes, M. D. F. *et al*. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA* **319**, 388–396, https://doi.org/10.1001/jama.2017.19163 (2018).

29. Shamseer, L. *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ (Clinical research ed.)* **350**, g7647, https://doi.org/10.1136/bmj.g7647 (2015).

30. Whiting, P. F. *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* **155**, 529–536, https://doi.org/10.7326/0003-4819-155-8-201110180-00009 (2011).

31. Brozek, J. L. *et al*. Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 2 of 3. The GRADE approach to grading quality of evidence about diagnostic tests and strategies. *Allergy* **64**, 1109–1116, https://doi.org/10.1111/j.1398-9995.2009.02083.x (2009).

32. Houssami, N. *et al*. Interval breast cancers in the 'screening with tomosynthesis or standard mammography' (STORM) population-based trial. *Breast (Edinburgh, Scotland)* **38**, 150–153, https://doi.org/10.1016/j.breast.2018.01.002 (2018).

33. Houssami, N. *et al*. Breast cancer detection using single-reading of breast tomosynthesis (3D-mammography) compared to double-reading of 2D-mammography: Evidence from a population-based trial. *Cancer epidemiology* **47**, 94–99, https://doi.org/10.1016/j.canep.2017.01.008 (2017).

34. Skaane, P. *et al*. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology* **267**, 47–56, https://doi.org/10.1148/radiol.12121373 (2013).

35. Bernardi, D. *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. *The Lancet. Oncology* **17**, 1105–1113, https://doi.org/10.1016/s1470-2045(16)30101-2 (2016).

36. Destounis, S., Arieno, A. & Morgan, R. Initial experience with combination digital brea st tomosynthesis plus full field digital mammography or full field digital mammography alone in the screening environment. *Journal of Clinical Imaging Science* **4**, https://doi.org/10.4103/2156-7514.127838 (2014).

37. Durand, M. A. *et al*. Early clinical experience with digital breast tomosynthesis for screening mammography. *Radiology* **274**, 85–92, https://doi.org/10.1148/radiol.14131319 (2015).

38. Greenberg, J. S., Javitt, M. C., Katzen, J., Michael, S. & Holland, A. E. Clinical performance metrics of 3D digital breast tomosynthesis compared with 2D digital mammography for breast cancer screening in community practice. *AJR. American journal of roentgenology* **203**, 687–693, https://doi.org/10.2214/ajr.14.12642 (2014).

39. Houssami, N. *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. *European Journal of Cancer* **50**, 1799–1807, https://doi.org/10.1016/j.ejca.2014.03.017 (2014).

40. Lång, K. *et al*. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study. *European radiology* **26**, 184–190, http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/059/CN-01171059/frame.html (2016).

41. Lang, K., Nergarden, M., Andersson, I., Rosso, A. & Zackrisson, S. False positives in breast cancer screening with one-view breast tomosynthesis: An analysis of findings leading to recall, work-up and biopsy rates in the Malmo Breast Tomosynthesis Screening Trial. *European radiology* **26**, 3899–3907, https://doi.org/10.1007/s00330-016-4265-y (2016).

42. Skaane, P. *et al*. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a populationbased screening program. *Radiology* **267**, 47–56, http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/174/CN-00908174/frame.html (2013).

43. Skaane, P. *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* **271**, 655–663, https://doi.org/10.1148/radiol.13131391 (2014).

44. Zervoudis, S. *et al*. Tomosynthesis improves breast cancer detection: our experience. *European journal of gynaecological oncology* **35**, 666–669 (2014).

45. Skaane, P. *et al*. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. *European Radiology* **23**, 2061–2071, https://doi.org/10.1007/s00330-013-2820-3 (2013).

46. Powell, J. L. *et al*. Impact of the Addition of Digital Breast Tomosynthesis (DBT) to Standard 2D Digital Screening Mammography on the Rates of Patient Recall, Cancer Detection, and Recommendations for Short-term Follow-up. *Academic radiology* **24**, 302–307, https://doi.org/10.1016/j.acra.2016.10.001 (2017).

47. Hofvind, S. *et al*. Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program. *Radiology* **287**, 787–794, https://doi.org/10.1148/radiol.2018171361 (2018).

48. Romero Martin, S. *et al.* Prospective study aiming to compare 2D mammography and tomosynthesis + synthesized mammography in terms of cancer detection and recall. From double reading of 2D mammography to single reading of tomosynthesis. *Eur Radiol* **28**, 2484–2491, https://doi.org/10.1007/s00330-017-5219-8 (2018).

49. Aase, H. 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. *European radiology* **29**, 1175–1186, https://doi.org/10.1007/s00330-018-5690-x (2019).

50. Lourenco, A. P., Barry-Brooks, M., Baird, G. L., Tuttle, A. & Mainiero, M. B. Changes in recall type and patient treatment following implementation of screening digital breast tomosynthesis. *Radiology* **274**, 337–342, https://doi.org/10.1148/radiol.14140317 (2015).

51. McCarthy, A. M. *et al.* Screening outcomes following implementation of digital breast tomosynthesis in a general-population screening program. *Journal of the National Cancer Institute* **106**, https://doi.org/10.1093/jnci/dju316 (2014).

52. Aujero, M. P., Gavenonis, S. C., Benjamin, R., Zhang, Z. & Holt, J. S. Clinical Performance of Synthesized Two-dimensional Mammography Combined with Tomosynthesis in a Large Screening Population. *Radiology* **283**, 70–76, https://doi.org/10.1148/radiol.2017162674 (2017).

53. Friedewald, S. M. *et al.* Breast cancer screening using tomosynthesis in combination with digital mammography. *Jama* **311**, 2499–2507, https://doi.org/10.1001/jama.2014.6095 (2014).

54. Rose, S. L. *et al.* Implementation of breast tomosynthesis in a routine screening practice: An observational study. *American Journal of Roentgenology* **200**, 1401–1408, https://doi.org/10.2214/AJR.12.9672 (2013).

55. Rose, S. L. *et al.* A reader study comparing prospective tomosynthesis interpretations with retrospective readings of the corresponding FFDM examinations. *Academic radiology* **21**, 1204–1210, https://doi.org/10.1016/j.acra.2014.04.008 (2014).

56. Roth, R. G., Maidment, A. D., Weinstein, S. P., Roth, S. O. & Conant, E. F. Digital breast tomosynthesis: lessons learned from early clinical implementation. *Radiographics: a review publication of the Radiological Society of North America, Inc* **34**, E89–102, https://doi.org/10.1148/rg.344130087 (2014).

57. McDonald, E. S. *et al.* Baseline Screening Mammography: Performance of Full-Field Digital Mammography Versus Digital Breast Tomosynthesis. *AJR. American journal of roentgenology* **205**, 1143–1148, https://doi.org/10.2214/ajr.15.14406 (2015).

58. McDonald, E. S. *et al.* Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening. *JAMA oncology* **2**, 737–743, https://doi.org/10.1001/jamaoncol.2015.5536 (2016).

59. Sharpe, R. E. *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* **278**, 698–706, https://doi.org/10.1148/radiol.2015142036 (2016).

60. Zuckerman, S. P. *et al.* Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. *Radiology* **281**, 730–736, https://doi.org/10.1148/radiol.2016160366 (2016).

61. Gilbert, F. J. *et al.* Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial). *Radiology* **277**, 697–706, https://doi.org/10.1148/radiol.2015142566 (2015).

62. Gur, D. *et al.* Digital breast tomosynthesis: observer performance study. *AJR. American journal of roentgenology* **193**, 586–591, https://doi.org/10.2214/ajr.08.2031 (2009).

63. Gur, D. *et al.* Dose Reduction in Digital Breast Tomosynthesis (DBT) Screening using Synthetically Reconstructed Projection Images. An Observer Performance Study. *Academic Radiology* **19**, 166–171, https://doi.org/10.1016/j.acra.2011.10.003 (2012).

64. Kontos, D. *et al.* Analysis of parenchymal texture with digital breast tomosynthesis: Comparison with digital mammography and implications for cancer risk assessment. *Radiology* **261**, 80–91, https://doi.org/10.1148/radiol.11100966 (2011).

65. Rafferty, E. A. *et al.* Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: Results of a multicenter, multireader trial. *Radiology* **266**, 104–113, https://doi.org/10.1148/radiol.12120674 (2013).

66. Bernardi, D. *et al.* Application of breast tomosynthesis in screening: Incremental effect on mammography acquisition and reading time. *British Journal of Radiology* **85**, e1174–e1178, https://doi.org/10.1259/bjr.19385909 (2012).

67. Bernardi, D. *et al.* Prospective study of breast tomosynthesis as a triage to assessment in screening. *Breast cancer research and treatment* **133**, 267–271, https://doi.org/10.1007/s10549-012-1959-y (2012).

68. Sumkin, J. H. *et al.* Recall Rate Reduction with Tomosynthesis During Baseline Screening Examinations: An Assessment From a Prospective Trial. *Academic radiology* **22**, 1477–1482, https://doi.org/10.1016/j.acra.2015.08.015 (2015).

69. Dang, P. A., Freer, P. E., Humphrey, K. L., Halpern, E. F. & Rafferty, E. A. Addition of tomosynthesis to conventional digital mammography: Effect on image interpretation time of screening examinations. *Radiology* **270**, 49–56, https://doi.org/10.1148/radiol.13130765 (2014).

70. El Bakry, R. A. R. Breast tomosynthesis: A diagnostic addition to screening digital mammography. *Egyptian Journal of Radiology and Nuclear Medicine*, https://doi.org/10.1016/j.ejrnm.2017.12.004 (2018).

71. Haas, B. M. *et al.* Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. *Radiology* **269**, 694–700, https://doi.org/10.1148/radiol.13130307 (2013).

72. Institute, N. N. C. *TMIST (Tomosynthesis Mammographic Imaging Screening Trial)* 2020).

73. Weigel, S. *et al.* Digital breast tomosynthesis plus synthesised images versus standard full-field digital mammography in population-based screening (TOSYMA): protocol of a randomised controlled trial. *BMJ Open* **8**, e020475, https://doi.org/10.1136/bmjopen-2017-020475 (2018).

74. Conant, E. F. *et al.* Five Consecutive Years of Screening with Digital Breast Tomosynthesis: Outcomes by Screening Year and Round. *Radiology*, 191751, https://doi.org/10.1148/radiol.2020191751 (2020).

75. Bahl, M. *et al.* Breast Cancer Characteristics Associated with 2D Digital Mammography versus Digital Breast Tomosynthesis for Screening-detected and Interval Cancers. *Radiology* **287**, 49–57, https://doi.org/10.1148/radiol.2017171148 (2018).

76. Gotzsche, P. C. & Jorgensen, K. J. Screening for breast cancer with mammography. *The Cochrane database of systematic reviews*, CD001877, https://doi.org/10.1002/14651858.CD001877.pub5 (2013).

77. Nelson, H. D. *et al.* Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation. *Ann Intern Med* **164**, 244–255, https://doi.org/10.7326/M15-0969 (2016).

Acknowledgements

Author contributions

All authors developed the systematic review; the manuscript was drafted by V.S.N.N. V.S.N.N. developed the search strategies. R.R.G., M.V.G.C., V.S.N.N. independently screened eligible studies, extracted data from included studies, and assessed the risk of bias. S.A.M.L. and S.A.T.W. elaborated the standard extraction form. V.S.N.N. supervised all phases of this review and refereed any disagreement to avoid errors. All authors participated in the data synthesis and quality of evidence. All authors critically revised the manuscript and approved its final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-64802-x.

Correspondence and requests for materials should be addressed to V.d.S.-N.

Reprints and permissions information is available at www.nature.com/reprints.

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