



Two-view digital breast tomosynthesis versus digital mammography in a population-based breast cancer screening programme (To-Be): a randomised, controlled trial

Solveig Hofvind, Åsne S Holen, Hildegunn S Aase, Nehmat Houssami, Sofie Sebuødegård, Tron A Moger, Ingfrid S Haldorsen, Lars A Akslen

Summary

Background Digital breast tomosynthesis is an advancement of mammography, and has the potential to overcome limitations of standard digital mammography. This study aimed to compare first-generation digital breast tomosynthesis including two-dimensional (2D) synthetic mammograms versus digital mammography in a population-based screening programme.

Methods BreastScreen Norway offers all women aged 50–69 years two-view (craniocaudal and mediolateral oblique) mammographic screening every 2 years and does independent double reading with consensus. We asked all 32 976 women who attended the programme in Bergen in 2016–17, to participate in this randomised, controlled trial with a parallel group design. A study-specific software was developed to allocate women to either digital breast tomosynthesis or digital mammography using a 1:1 simple randomisation method based on participants' unique national identity numbers. The interviewing radiographer did the randomisation by entering the number into the software. Randomisation was done after consent and was therefore concealed from both the women and the radiographer at the time of consent; the algorithm was not disclosed to radiographers during the recruitment period. All data needed for analyses were complete 12 months after the recruitment period ended. The primary outcome measure was screen-detected breast cancer, stratified by screening technique (ie, digital breast tomosynthesis and digital mammography). A log-binomial regression model was used to estimate the efficacy of digital breast tomosynthesis versus digital mammography, defined as the crude risk ratios (RRs) with 95% CIs for screen-detected breast cancer for women screened during the recruitment period. A per-protocol approach was used in the analyses. This trial is registered at ClinicalTrials.gov, number NCT02835625, and is closed to accrual.

Findings Between, Jan 14, 2016, and Dec 31, 2017, 44 266 women were invited to the screening programme in Bergen, and 32 976 (74.5%) attended. After excluding women with breast implants and women who did not consent to participate, 29 453 (89.3%) were eligible for electronic randomisation. 14 734 women were allocated to digital breast tomosynthesis and 14 719 to digital mammography. After randomisation, women with a previous breast cancer were excluded (digital breast tomosynthesis group n=314, digital mammography group n=316), women with metastases from melanoma (digital breast tomosynthesis group n=1), and women who informed the radiographer about breast symptoms after providing consent (digital breast tomosynthesis group n=39, digital mammography group n=34). After exclusions, information from 28 749 women were included in the analyses (digital breast tomosynthesis group n=14 380, digital mammography group n=14 369). The proportion of screen-detected breast cancer among the screened women did not differ between the two groups (95 [0.66%, 0.53–0.79] of 14 380 vs 87 [0.61%, 0.48–0.73] of 14 369; RR 1.09, 95% CI 0.82–1.46; p=0.56).

Interpretation This study indicated that digital breast tomosynthesis including synthetic 2D mammograms was not significantly different from standard digital mammography as a screening tool for the detection of breast cancer in a population-based screening programme. Economic analyses and follow-up studies on interval and consecutive round screen-detected breast cancers are needed to better understand the effect of digital breast tomosynthesis in population-based breast cancer screening.

Funding Cancer Registry of Norway, Department of Radiology at Haukeland University Hospital, University of Oslo, and Research Council of Norway.

Copyright © 2019 Elsevier Ltd. All rights reserved.

Introduction

Standard digital mammography in combination with digital breast tomosynthesis has been shown to increase the frequency of breast cancer detection, but its effect on recall varied in a meta-analysis¹ of non-randomised

studies.^{2–5} These randomised studies included women of different ages, with varying screening intervals, and with a variety of screening and screen reading procedures. Two study designs were generally used to compare the outcome of screening with digital breast tomosynthesis

Lancet Oncol 2019; 20: 795–805

Published Online

May 8, 2019

[http://dx.doi.org/10.1016/S1470-2045\(19\)30161-5](http://dx.doi.org/10.1016/S1470-2045(19)30161-5)

S1470-2045(19)30161-5

See [Comment](#) page 745

This online publication has been corrected. The corrected version first appeared at thelancet.com/oncology on July 1, 2019

Department of Breast Screening, Cancer Registry of Norway, Oslo, Norway (Prof S Hofvind PhD, Å S Holen MSc, S Sebuødegård MS); Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway (Prof S Hofvind); Department of Radiology (H S Aase MD, Prof I S Haldorsen MD), Department of Pathology (Prof L A Akslen MD), and Centre for Cancer Biomarkers (Prof L A Akslen), Haukeland University Hospital, Bergen, Norway; Department of Clinical Medicine, University of Bergen, Bergen, Norway (H S Aase, Prof I S Haldorsen, Prof L A Akslen); Faculty of Medicine and Health, Sydney School of Public Health, University of Sydney, Sydney, NSW, Australia (Prof N Houssami MD); and Department of Health Management and Health Economics, Institute of Health and Society, University of Oslo, Oslo, Norway (Prof T A Moger PhD)

Correspondence to: Prof Solveig Hofvind, Department of Breast Screening, Cancer Registry of Norway, 0304 Oslo, Norway solveig.hofvind@krefregisteret.no

Research in context

Evidence before this study

Two of the investigators (SH and NH) searched MEDLINE between Jan 1, 2010, and Aug 31, 2015, for all publications with “breast neoplasms” and searched “tomosyn\$” in the title, to identify studies reporting screening outcomes for population-based breast screening with digital breast tomosynthesis. No randomised, controlled trials of digital breast tomosynthesis screening were identified.

Three prospective, non-randomised trials and several retrospective studies reported screening performance measures for digital breast tomosynthesis, predominantly in combination with standard digital mammography. The results of these studies suggested that digital breast tomosynthesis and digital mammography generally improved screening outcome measures by increasing cancer detection or reducing recall, or both, compared with digital mammography alone, although effect estimates varied. Only two of these studies investigated screen reading with digital breast tomosynthesis alone, or investigated digital breast tomosynthesis including synthetic two-dimensional (2D) mammograms as one of the screen reading strategies. Both studies were prospective and reported only interim analyses. The study of digital breast tomosynthesis alone showed significantly increased proportions of cancer detection and recall, whereas the study including digital breast tomosynthesis and synthetic 2D mammograms showed equivalence in cancer detection with digital breast tomosynthesis and digital mammography.

Added value of this study

To our knowledge, this is the first randomised, controlled trial reporting outcomes for screening with digital breast tomosynthesis and synthetic 2D mammograms in a

high-throughput, population-based breast cancer screening programme. By contrast with most other studies, we did not find significant differences in the proportion of breast cancers detected or tumour characteristics for digital breast tomosynthesis and synthetic 2D mammograms versus digital mammography alone. We found a lower proportion of recall and higher positive predictive value of recalls among those screened with digital breast tomosynthesis compared with digital mammography. Our results indicate that digital breast tomosynthesis and synthetic 2D mammogram is not significantly different from standard digital mammography as a screening tool for selected breast cancer screening outcome measures in a population-based screening programme for women at average risk of breast cancer.

Implication of all the available evidence

The evidence on digital breast tomosynthesis as a screening tool points towards a higher proportion of screen-detected breast cancer and prognostically favourable tumour characteristics among women screened with digital breast tomosynthesis in combination with digital mammography or synthetic 2D mammogram, compared with standard digital mammography. We found that use of digital breast tomosynthesis in a screening setting is safe, and has the potential to reduce harms, but studies on interval and screen-detected breast cancer in consecutive screening rounds are needed to better understand the effect of digital breast tomosynthesis in a population-based screening setting. These results will be valuable in the coming policy discussions about whether to implement digital breast tomosynthesis in breast cancer screening programmes.

and digital mammography versus digital mammography alone; ie, paired studies²⁻⁵ using women as their own controls and unpaired studies⁶⁻⁹ using other geographical areas or historical data as the control group. The paired studies showed a consistent increase in the detection of breast cancer, but the design hampers future analyses of interval and consecutive breast cancer. The unpaired studies showed a less substantial increase in cancer detection, and the sample sizes varied considerably.¹ To our knowledge, results from only one randomised, controlled trial comparing digital breast tomosynthesis and digital mammography versus digital mammography have been published in an interim report.¹⁰ This study showed a 90% higher proportion of screen-detected breast cancer for women screened with digital breast tomosynthesis and digital mammography versus digital mammography alone.

Lower proportions of recall for digital breast tomosynthesis and digital mammography have been shown compared with standard mammography alone, but mostly in studies from the USA, where recalls are known

to be substantially higher than in Europe.¹ A lower frequency of recall is beneficial for women if the proportion of screen-detected breast cancer remains stable or increases; however, a higher proportion of screen-detected breast cancer is only beneficial for screened women if the tumours are progressive, as opposed to small, low-histologically graded tumours, which might represent overdiagnosis. If screen-detected breast cancers are detected at an earlier stage than in a clinical setting without screening, we expect the proportion of interval or consecutive screen-detected breast cancer, or both, to decrease.

The use of digital breast tomosynthesis and digital mammography roughly doubles the dose of radiation compared with digital mammography alone.¹¹ As a result, synthetic two-dimensional (2D) mammograms have been developed using raw data from the digital breast tomosynthesis acquisition to minimise the radiation burden on women. To our knowledge, there is insufficient evidence to draw any conclusions about the overall balance of benefits and harms of using digital breast

tomosynthesis and synthetic 2D mammograms in a population-based screening programme, and no results from randomised, controlled trials are available.^{11–13}

To fill the evidence gaps in the use of digital breast tomosynthesis in combination with synthetic 2D mammograms in breast cancer screening, we did the tomosynthesis trial in Bergen (the To-Be trial). This trial aimed to investigate performance measures and economic aspects of using digital breast tomosynthesis and synthetic 2D mammograms (hereafter referred to as digital breast tomosynthesis) versus digital mammography alone in an organised, population-based breast cancer screening programme with high attendance and complete data.¹⁴ This Article presents results of outcome measures for women screened with digital breast tomosynthesis versus digital mammography regarding the screen-detected breast cancer, consensus, recall, and distribution of histopathological tumour characteristics, in addition to the positive predictive value of recall (PPV-1) and biopsy (PPV-2), as well as time used for screen reading.

Methods

Study design and participants

The To-Be trial was a large-scale, parallel group randomised, controlled trial, in which the outcome of the new screening technique digital breast tomosynthesis was tested in an everyday screening setting. Participants were screened in Bergen, Norway, through the national screening programme, BreastScreen Norway. BreastScreen Norway offers women aged 50–69 years biennial two-view mammographic screening,¹⁴ and the Cancer Registry of Norway administers the programme. Cancer registration is mandated by law in Norway, and the Registry's databases are more than 99% complete for breast cancer.

44 266 women born between 1947 and 1966 were invited to the screening unit in Bergen during the recruitment period. All attending women were informed about the trial and received written information from an administrative assistant when they entered the screening unit. In the prescreening room, radiographers prepared women for the screening exam and asked about participation in the trial before they asked general questions, following standard procedures.

Women who agreed to participate in the trial signed a written, electronic consent form. Status of the study, a protocol synopsis, and the protocol are available online. The trial was approved by the Regional Committees for Medical and Health Research Ethics in southeastern Norway (official record number 2015/424).

Randomisation and masking

We developed a study-specific software to randomly allocate women to receive either digital breast tomosynthesis or digital mammography using a 1:1 simple randomisation method based on participants' unique 11-digit national identity numbers. This sequence included

two randomly generated control digits, one of which was used by the randomisation software to assign the screening technique. The interviewing radiographer in the prescreening room did the randomisation by entering the unique identification number of the consenting women into the software. Randomisation was done after consent was obtained and was therefore concealed from both the women and the radiographer at the time of consent. Only the primary investigator (SH) and the software developers knew the algorithm for randomisation; it was concealed to the radiographers during the study period. We evaluated the randomisation process continuously, but did not evaluate the masking of the trial group allocation. Because of the characteristics of the intervention, it was not possible to blind the intervention either in the screening or in the screen reading process.

Data on attendance in the programme and the trial were accessible for all professionals involved (both those who were working in the screening programme and those administering the trial) throughout the study. No other results were made available during the recruitment period.

Procedures

Women who consented to participate in the trial had one screening examination during the recruitment period. A standard examination consisted of two-view (craniocaudal and mediolateral oblique) mammography of each breast, either with digital breast tomosynthesis including synthetic 2D mammograms or digital mammography, using imaging equipment from GE Healthcare (Chicago, IL, USA; SenoClaire 3D Breast Tomosynthesis). Two radiographers did the screening examination as a team and all examinations took place at one of the two equally equipped examination rooms. The digital breast tomosynthesis acquisition consisted of nine exposures over an angle of 25°, reconstructed into synthetic 2D mammograms, 10 mm slabs, and 1 mm planes. Mean glandular dose per exam was 2·96 mGy for digital breast tomosynthesis and 2·95 mGy for digital mammography during the first year of the trial.¹⁵ Independent double screen reading was done on Image Diagnost International workstations, each with two 5-megapixel monitors (GE Healthcare Mammo-Workstation, version 4·70).

A standardised screen reading protocol including two-view synthetic 2D mammograms, slabs and planes of each breast for digital breast tomosynthesis or two-view digital mammography of each breast, was used for screen reading (appendix p 1) and at the consensus meetings (hereafter referred as consensus). Up to four previous screening examinations, or diagnostic images, or both, from the previous 10 years were available on the workstation for both trial groups, both for screen reading and at consensus.

We did independent double screen reading with consensus, following the programme's standard

For more on the **protocol** see <https://www.kreftregisteret.no/en/Research/Projects/to-be-studies/>

See Online for appendix

procedures.¹⁴ Each breast was assigned a score of one to five from each radiologist; a score of one indicated the screening examination was negative for abnormality; two, probably benign; three, intermediate suspicion; four, probably malignant; and five, high suspicion of malignancy. If either radiologist assigned a score of two or higher, consensus was used to determine whether to recall the woman for further assessment, hereafter referred to as recall. Consensus was done by pairs of radiologists, and a third radiologist was consulted if the pair could not agree.

A pool of eight breast radiologists did the initial screen readings and consensus. Their experience in screen reading (screen film and digital mammography) before start-up of the trial varied from zero to approximately 110 000 examinations.¹⁵ The number of digital breast tomosynthesis and digital mammography screen readings, and interpretation time in the trial were automatically recorded for each radiologist, whereas the consensus time was recorded for each meeting.

Further assessment took place 2–8 weeks after screening, at the breast centre at Haukeland University Hospital by the same pool of radiologists who did the screen reading. If indicated, additional imaging, including digital breast tomosynthesis, ultrasound, and, less frequently, contrast enhanced spectral mammo-graphy, or MRI, or both, was done. Cancer diagnoses were verified by pathologists examining histological specimens. Information about histopathological features were considered complete 12 months after further assessment took place. We considered the screening examination, with or without further assessment as one event.

Women were the unit of analysis and because all women were screened only once in the trial, the number of women and screening examinations were the same. Women with bilateral breast cancer were included once on the basis of the topography of the recalled lesion or, if recalled for lesions in both breasts, based on malignancy according to histological type (invasive before ductal carcinoma in situ), tumour diameter, and histological grade. Histopathological tumour characteristics for ductal carcinoma in situ and invasive breast cancer were determined from routine histopathology reports done by pathologists affiliated with the breast centre at Haukeland University Hospital, with no reclassification. Lobular carcinoma in situ was included in the group of benign lesions¹⁶ (digital breast tomosynthesis group $n=3$, digital mammography group $n=2$). We diagnosed one case of pleomorphic lobular carcinoma in situ in the digital breast tomosynthesis group, which we included in the lobular carcinoma in situ group.

We collected information about the screening examinations and screen reading electronically, in compliance with standard procedures at BreastScreen Norway. Additional information related to consensus and recall was manually recorded on a paper-based form,

which was designed specifically for this trial by the principal investigator (SH) and the radiologists at the breast centre. These forms were transferred electronically or by letter mail to the Cancer Registry of Norway, where the data were registered and quality assured by a dedicated research assistant before interim and final analyses were done. All study data were stored in databases at the Cancer Registry of Norway.

Outcomes

The primary outcome of this trial was to determine whether the proportion of screen-detected breast cancer was favourable for digital breast tomosynthesis versus digital mammography, as pre-specified in the protocol. Breast cancer was defined as histologically verified ductal carcinoma in situ or invasive breast cancer, or both.

Secondary outcome measures, also specified in the protocol, were percentage of recalls, positive predictive value of recall (PPV-1) and biopsy (PPV-2), histopathological tumour characteristics, and economic costs. PPV-1 was the percentage of breast cancer cases detected among the women who were recalled, and PPV-2 was the percentage of breast cancer detected among women who were recalled and had a needle biopsy.

Prognostic characteristics for invasive tumours included mean and median tumour diameter, and distribution of tumour diameter groups (<10 mm, 10 mm to <20 mm, and ≥ 20 mm), histological grade (1, 2, and 3), and lymph node involvement (negative or positive) were presented as percentages of all values. For tumour diameter, lesions treated with neoadjuvant therapy were categorised as information not available (digital breast tomosynthesis group $n=9$, digital mammography group $n=11$). Predictive biomarkers included oestrogen and progesterone receptor status (positive or negative for each), HER2 (positive or negative), and Ki67 proliferation (<30% and $\geq 30\%$). This information was used to classify the invasive tumours into subtypes (luminal A, luminal B HER2-negative, luminal B HER2-positive, HER2-positive, and triple negative).¹⁷

Other outcomes specified in the protocol were consensus, time spent on screen reading and consensus, mammographic features, radiation doses for the two techniques, and interval and breast cancers in consecutive screening rounds, for which at least 2 years of follow-up of each individual woman was needed.

We defined women's screening history as prevalent, indicating the first or incident screen in BreastScreen Norway, or subsequent, indicating previous attendance in the programme and the availability of previous screening mammograms (digital mammography) for comparison during screen reading.

Statistical analysis

The hypothesis of our trial was that the proportion of screen-detected breast cancer would be superior for digital breast tomosynthesis versus digital mammography. In a population with an estimated previous

screen-detected breast cancer proportion of 0·60%, we calculated that with 15 000 women in each group, we could observe an increase in detection from 0·60% with digital mammography to 0·88% with digital breast tomosynthesis, given 80% power and a two-sided significance threshold of 5%.

Our analyses included information from all women with a complete screening examination, who had no previous history of breast cancer or metastatic melanoma, and who did not report breast symptoms when attending for screening examination. These women represent the per-protocol population. No women revoked their consent to participate in the study. Results from interim analyses on the proportion of consensus and recall, in addition to radiation dose for digital breast tomosynthesis versus digital mammography from the first year of the recruitment period, were conducted as planned.¹⁵

We estimated crude risk ratios (RRs) with 95% CIs for screen-detected breast cancer for digital breast tomosynthesis with digital mammography as the reference, using a log-binomial regression model. Log-binomial regression models were also fitted to estimate the RR of recall for women screened with digital breast tomosynthesis versus digital mammography.

All analyses were stratified by the randomly assigned screening technique (digital breast tomosynthesis or digital mammography). Age was categorised into four groups: younger than 55 years, 55–59 years, 60–64 years, and 65 years or older. Consensus, recall, biopsy, and screen-detected breast cancer were presented as proportion per 100 women screened within the recruitment period, whereas PPV-1 and PPV-2 were presented as percentages. The distributions of histopathological tumour characteristics were reported as percentages for cases with non-missing values. Tumour diameter (mm) and time spent on initial and consensus reading (min, sec) did not follow a perfect normal distribution, thus we estimated both mean (SD) and median (IQR).

For sensitivity analyses, we did log-binomial regression models to estimate crude RR for screen-detected breast cancer for digital breast tomosynthesis versus digital mammography, stratified by screening history. Quality control and assurance was continuously done according to standard procedures as interim analyses, to ensure the women's safety according to radiation dose, recall and breast cancer detection, and for continuation of the trial. In the case of unacceptable values, the trial would have been discontinued. Results were regularly presented for the steering committee of the trial, and were only discussed with this committee. No independent review was done.

We used Stata (version 15.0) for all statistical analyses and tested differences across categories using two-sample *t* tests, χ^2 tests, one-way ANOVA or tests of proportions (*Z* test). A *p* value of less than 0·05 was considered significant.

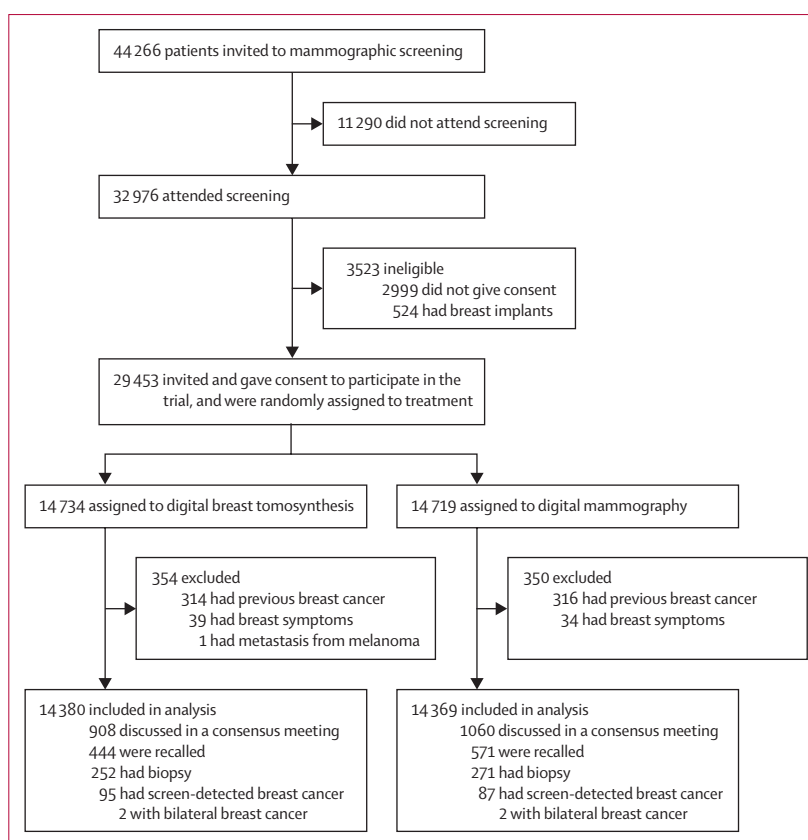


Figure 1: Trial profile

This trial is registered at ClinicalTrials.gov, number NCT02835625.

Role of the funding source

The funder of the study had no role in the study design, data collection, analyses, interpretation, or writing of the report. SH, ÅH, and SS had full access to all the data in the study, and the corresponding author had final responsibility for the decision to submit the manuscript.

Results

Among the 44 266 women invited to the screening unit in Bergen during the recruitment period, Jan 14, 2016, to Dec 31, 2017, 32 976 (74·5%) attended (figure 1). Women who did not consent to participate (2999 [9·1%] of 32 976) and women with breast implants (524 [1·6%] of 32 976) were not included in the trial, leaving 29 453 (89·3%) eligible to be randomly assigned to treatment. Of the eligible women, 14 734 were allocated to screening with digital breast tomosynthesis and 14 719 to digital mammography. Additional exclusions post-randomisation (breast cancer diagnosed before date of screening, metastatic melanoma, and reporting symptoms of breast cancer at the screening examination) resulted in a final study population of 28 749 women (digital breast tomosynthesis *n*=14 380, digital mammography *n*=14 369).

	Digital breast tomosynthesis (n=14 380)	Digital mammography (n=14 369)
Age, years		
<55	3746 (26.1%)	3813 (26.5%)
55–59	3628 (25.2%)	3668 (25.5%)
60–64	3625 (25.2%)	3492 (24.3%)
>64	3381 (23.5%)	3396 (23.6%)
Screening history		
Prevalent (first screening)	2013 (14.0%)	2053 (14.3%)
Subsequent	12 367 (86.0%)	12 316 (85.7%)

Data are n (%).

Table 1: Baseline characteristics

	Digital breast tomosynthesis (n=14 380)	Digital mammography (n=14 369)	p value*
Screen-detected breast cancer	95 (0.66%, 0.53–0.79)	87 (0.61%, 0.48–0.73)	0.56
Ductal carcinoma in situ	15 (0.10%, 0.05–0.16)	16 (0.11%, 0.06–0.17)	0.86
Invasive	80 (0.56%, 0.43–0.68)	71 (0.49%, 0.38–0.61)	0.47
Recall	444 (3.1%, 2.8–3.4)	571 (4.0%, 3.7–4.3)	<0.0001
PPV-1	95/444 (21.4%, 17.6–25.2)	87/571 (15.2%, 12.3–18.2)	0.011
PPV-2	95/252 (37.7%, 31.7–43.7)	87/271 (32.1%, 26.5–37.7)	0.18
Biopsy	252 (1.8%, 1.5–2.0)	271 (1.9%, 1.7–2.1)	0.40
Consensus	908 (6.3%, 5.9–6.7)	1060 (7.4%, 6.9–7.8)	0.0004

Data are n (%), 95% CI or n/N (%), 95% CI. PPV-1=positive predictive value of recall. PPV-2=positive predictive value of biopsy. *p value for Z test between digital breast tomosynthesis and digital mammography.

Table 2: Summary of outcomes in the To-Be trial by screening technique

during the recruitment period (figure 1). Baseline characteristics are presented in table 1, and characteristics before post-randomisation exclusions are available in the appendix (p 2).

182 women were diagnosed with screen-detected breast cancer because of mammographic findings: 31 ductal carcinoma in situ and 151 invasive cancers (table 2). Four women were diagnosed with bilateral cancer (n=2 in each group). The proportion of screen-detected breast cancer did not differ between the two groups: 95 (0.66%, 0.53–0.79) of 14 380 in the digital breast tomosynthesis group versus 87 (0.61%, 0.48–0.73) of 14 369 in the digital mammography group (RR 1.09, 95% CI 0.82–1.46; p=0.56), whereas the risk of recall was lower for digital breast tomosynthesis versus digital mammography (RR 0.78, 95% CI 0.69–0.88; p<0.0001). Summary of outcomes, without any post-randomisation exclusions, are available in the appendix (p 2).

A consensus meeting was held for 908 (6.3%) of 14 380 women in the digital breast tomosynthesis group versus 1060 (7.4%) of 14 369 in the digital mammography group (p=0.0004; table 2). 444 (3.1%) of 14 380 women in the digital breast tomosynthesis group versus 571 (4.0%) of 14 369 in the digital mammography group (p<0.0001) were recalled. The proportion of biopsies did not differ between digital breast tomosynthesis (252 [1.8%] of

14 380) and digital mammography (271 [1.9%] of 14 369; p=0.40). PPV-1 was significantly higher for digital breast tomosynthesis (95 [21.4%] of 444) compared with digital mammography (87 [15.2%] of 571; p=0.01), while PPV-2 was not (p=0.18). RR for recall and screen-detected breast cancer without any post-randomisation exclusions and results of sensitivity analyses, stratified by screening history, are available in the appendix (p 3).

Consensus, recall, and detection increased during the study period for digital breast tomosynthesis (p=0.02, p=0.01, and p=0.04, respectively; figure 2). For digital mammography, an increase was observed for recall (p=0.01), but not for consensus and detection.

The distribution of histopathological tumour characteristics did not differ significantly between the two groups for invasive breast cancer (table 3), or for ductal carcinoma in situ (table 4).

All eight radiologists read both digital breast tomosynthesis and digital mammography screens during the trial, ranging from 1079 to 5663 examinations for digital breast tomosynthesis and from 1067 to 7538 for digital mammography per radiologist (appendix p 3). Overall, mean time spent on initial screen reading was 1 min 6 sec (median 48 sec, IQR 33–78) for digital breast tomosynthesis, including interpretation of synthetic 2D mammograms and digital breast tomosynthesis planes as well as prior images (from a previous screening examination or diagnostic mammography), and 39 sec (median 23 sec, IQR 13–44) for digital mammography, including interpretation of prior images (difference in screen reader time p<0.0001). Mean time for each radiologist ranged from 39 sec to 2 min 42 sec for digital breast tomosynthesis and from 13 sec to 3 min 2 sec for digital mammography. Mean time spent on consensus was 2 min 51 sec (SD 1 min 48 sec; median 2 min 21 sec, IQR 1 min 50 sec) for digital breast tomosynthesis and 2 min 4 sec (SD 2 min 5 sec; median 1 min 42 sec, IQR 1 min 11 sec) for digital mammography (p<0.0001; appendix p 3).

Recall and screen-detected breast cancer data in the pre-trial (2008–15) and trial period (2016–17), in Bergen, and in Norway excluding Bergen, are given in the appendix (p 4).

Discussion

This large-scale trial compared results from screening with digital breast tomosynthesis including synthetic 2D mammogram, with standard digital mammography, in an organised population-based breast cancer screening programme. Breast cancer detection did not differ significantly for women screened with digital breast tomosynthesis versus digital mammography. We observed lower consensus and recall for women screened with digital breast tomosynthesis versus digital mammography, whereas the PPV for recalls was higher for digital breast tomosynthesis compared with digital mammography. The distribution of histopathological tumour

characteristics did not differ significantly between the two screening techniques.

Our finding of no significant difference in breast cancer detection for digital breast tomosynthesis versus digital mammography was inconsistent with results from the majority of both paired and unpaired studies.¹ Our intervention group used digital breast tomosynthesis and synthetic 2D mammogram, whereas most of the earlier studies used digital breast tomosynthesis and digital mammography. A similar effect of synthetic 2D mammogram and digital mammography in combination with digital breast tomosynthesis has been consistently reported;^{3,18–20} however, the quality of synthetic 2D mammograms might present differently for different digital breast tomosynthesis machines and versions of software.

Digital mammography is commonly assumed to be superior to digital breast tomosynthesis in the characterisation of microcalcifications; however, recent studies have shown that the perceptibility of microcalcifications is also adequate for digital breast tomosynthesis in combination with synthetic 2D mammogram.^{21,22} Several factors affect image quality of both digital breast tomosynthesis, including synthetic 2D mammogram, and digital mammography, such as filter and anode combinations, spatial resolution, the angular range of the x-ray tube, and radiation dose.^{21,23} It could be argued that the radiation dose measured for digital breast tomosynthesis in our study, which was lower than that reported in other studies of digital breast tomosynthesis,¹⁵ might have negatively affected the image quality; however, differences in vendor-specific technical implementations and the optimisation of mammography workstations can also affect image quality. Moreover, different requirements for training to start screen reading, reading conditions, and protocols, including the availability and use of previous mammograms, could have affected the radiologists' perception and interpretation of mammographic features, and thus contributed to the heterogeneity of results in the published literature to date.

We used two-view mammography with independent double screen reading and consensus in both trial groups and a hanging protocol with availability of screening and any diagnostic mammograms taken during the previous 10 years, in line with standard procedures. However, the radiologists' experience in digital mammography screen reading before the start of the trial varied, and the programme's recommendations of 5000 annual screen readings was met by half of the radiologists.¹⁵ Furthermore, reading volume varied between radiologists for digital breast tomosynthesis and digital mammography during the trial. Experience in screen reading and preferences of digital mammography or digital breast tomosynthesis has been suggested to affect reader sensitivity.^{24,25} Radiologists in this study were not exposed to any preliminary results during any part of the trial, which we consider a strength; and made it impossible to

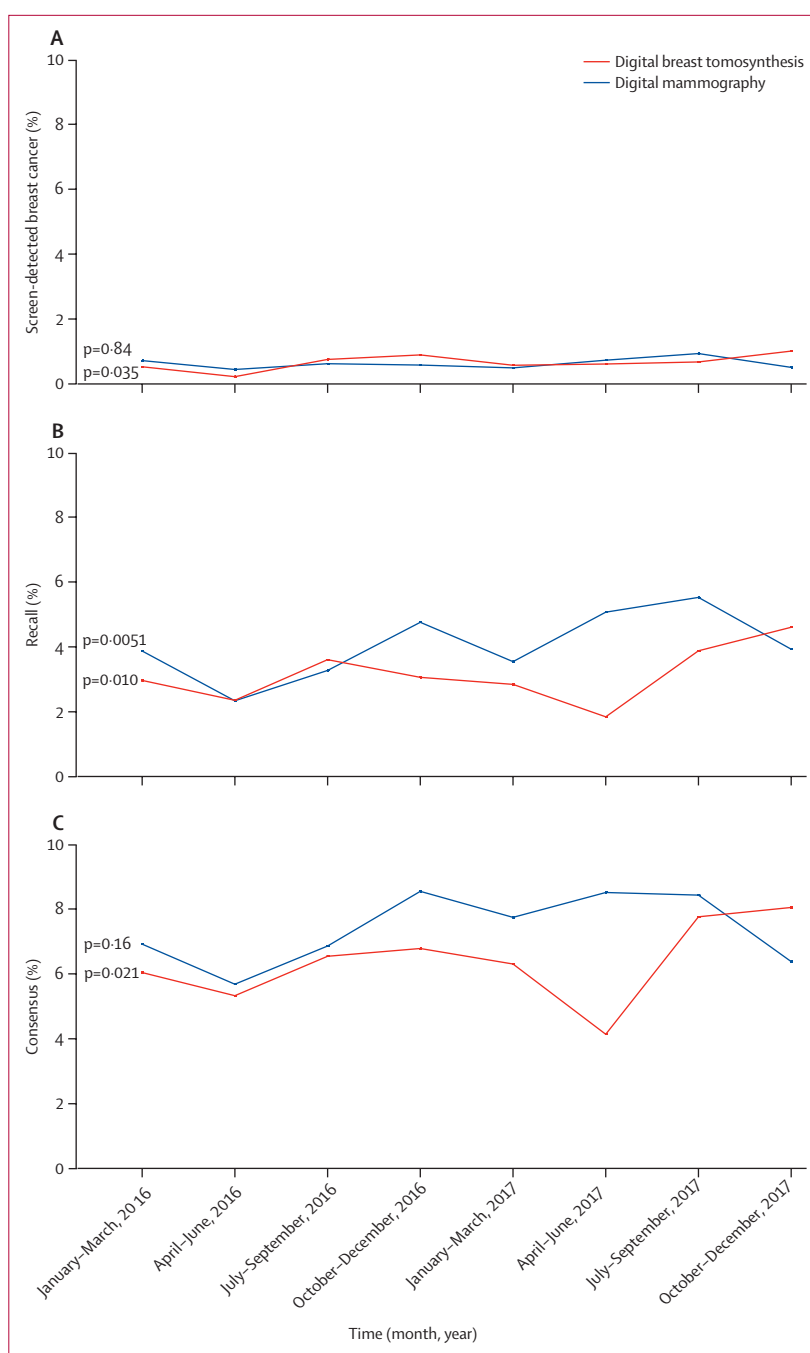


Figure 2: Screen-detected breast cancer (A), recall (B), and consensus (C) per 100 women screened. Tested for trend with one-way ANOVA.

provide individual feedback about their performance. Nonetheless, the radiologists were involved in both screen reading and recall assessments and thus not masked to the final outcome of the women's screening examination.

All participating radiologists were trained in digital breast tomosynthesis screen reading and diagnostics to some extent before the start of the trial, but they did not

	Digital breast tomosynthesis (n=80)	Digital mammography (n=71)	p value
Histological type			
Carcinoma (no special type)	62 (77.5%, 66.8–86.1)	51 (71.8%, 59.9–81.9)	0.11‡
Lobular carcinoma	6 (7.5%, 2.8–15.6)	13 (18.3%, 10.1–29.3)	..
Tubular carcinoma	2 (2.5%, 0.3–8.7)	4 (5.6%, 1.6–13.8)	..
Other carcinoma†	10 (12.5%, 6.2–21.8)	3 (4.2%, 0.9–11.9)	..
Tumour diameter			
<10 mm	10 (14.3%, 7.1–24.7)	18 (30.0%, 18.8–43.2)	0.09‡
≥10 to <20 mm	43 (61.4%, 49.0–72.8)	29 (48.3%, 35.2–61.6)	..
≥20 mm	17 (24.3%, 14.8–36.0)	13 (21.7%, 12.1–34.2)	..
Information not available	10 (12.5%, NA)	11 (15.5%, NA)	..
Lymph node positive	14 (17.7%, 10.0–27.9)	18 (25.7%, 16.0–37.6)	0.24‡
Information not available	1 (1.3%, NA)	1 (1.4%, NA)	..
Histological grade			
1	22 (28.9%, 19.1–40.5)	24 (34.8%, 23.7–47.2)	0.53‡
2	38 (50.0%, 38.3–61.7)	35 (50.7%, 38.4–63.0)	..
3	16 (21.1%, 12.5–31.9)	10 (14.5%, 7.2–25.0)	..
Information not available	4 (5.0%, NA)	2 (2.8%, NA)	..
Oestrogen receptor positive	71 (89.9%, 81.0–95.5)	69 (97.2, 90.2–99.7)	0.07‡
Information not available	1 (1.3%, NA)	0	..
Progesterone receptor positive	61 (77.2%, 66.4–85.9)	62 (87.3%, 77.3–94.0)	0.11‡
Information not available	1 (1.3%, NA)	0	..
HER2 positive	8 (10.1%, 4.5–19.0)	8 (11.3%, 5.0–21.0)	0.82‡
Information not available	1 (1.3%, NA)	0	..
Ki67 ≥30%	21 (29.6%, 19.3–41.6)	13 (20.0%, 11.1–31.8)	0.20‡
Information not available	9 (11.3%, NA)	6 (8.5%, NA)	..
Subtype			
Luminal A	44 (58.7%, 46.7–69.9)	43 (61.4%, 49.0–72.8)	0.43‡
Luminal B HER2-negative	18 (24.0%, 14.9–35.3)	18 (25.7%, 16.0–37.6)	..
Luminal B HER2-positive	5 (6.7%, 2.2–14.9)	7 (10.0%, 4.1–19.5)	..
HER2-positive	3 (4.0%, 0.8–11.2)	1 (1.4%, 0.04–7.7)	..
Triple negative	5 (6.7%, 2.2–14.9)	1 (1.4%, 0.04–7.7)	..
Information not available	5 (6.3%, NA)	1 (1.4%, NA)	..
Tumour diameter, mm‡			
Mean	70 (16.0, 8.4)	60 (14.5, 8.8)	0.33§
Median	70 (14.9, 11.0–18.0)	60 (14.0, 8.5–18.8)	..

Data are n (%; 95% CI), n (mean, SD), or n (median, IQR). NA=not applicable. Distributions are based on data with valid values, whereas the percentage of information not available is based on all cases, with and without missing values.

*p value for χ^2 test between digital breast tomosynthesis and digital mammography. §p value for t test of mean tumour diameter between digital breast tomosynthesis and digital mammography. †Including mucinous and other invasive cancers. ‡Neo-adjuvant treated women had no information available about tumour diameter.

Table 3: Histopathological tumour characteristics of invasive screen-detected breast cancer by screening technique

	Digital breast tomosynthesis (n=15)	Digital mammography (n=16)	p value
Tumour diameter			
<20 mm	3 (23.1%, 5.0–53.8)	8 (50.0%, 24.7–75.3)	0.14*
≥20 mm	10 (76.9%, 46.2–95.0)	8 (50.0%, 24.7–75.3)	..
Information not available	2 (13.3%, NA)	0	..
Grade			
1	2 (13.3%, 1.7–40.5)	1 (6.7%, 0.2–31.9)	0.82*
2	5 (33.3%, 11.8–61.6)	5 (33.3%, 11.8–61.6)	..
3	8 (53.3%, 26.6–78.7)	9 (60.0%, 32.3–83.7)	..
Information not available	0	1 (6.3%, NA)	..
Tumour diameter, mm			
Mean	13 (26.4, 15.7)	16 (28.5, 32.5)	0.83†
Median	13 (28.0, 20.0–30.0)	16 (18.5, 9.3–34.5)	..

Data are n (%; 95% CI), n (mean, SD), or n (median, IQR). Distributions are based on data with valid values, whereas the percentage of information not available is based on all cases, with and without missing values. *p value for χ^2 test between digital breast tomosynthesis and digital mammography. †p value for t test of mean tumour diameter between digital breast tomosynthesis and digital mammography.

Table 4: Histopathological tumour characteristics of screen-detected ductal carcinoma in situ by screening technique

mammography in our study, compared with 91 sec for digital breast tomosynthesis and digital mammography and 45 sec for digital mammography alone in the Oslo trial.⁴ This reduced screen reading time could represent cultural differences between the two breast centres, although our results showed substantial variation in time used for screen reading between the radiologists.

Analysis of the radiologist's sensitivity; interval and consecutive round screen-detected breast cancer; and reviews of cases dismissed at consensus, interval, and next round screen-detected breast cancers are needed to conclude whether the cancer cases were missed because of interpretation errors. Admittedly, consensus, recall, and detection increased over the 2 years recruitment period for digital breast tomosynthesis, whereas an increase was observed only for recalls for digital mammography. These findings might represent a learning effect for digital breast tomosynthesis. The increase in recall for digital mammography might represent natural variation, or study effect.

The availability of previous screening and diagnostic images at screen reading might have affected the number of cases discussed at consensus, and cases dismissed at consensus. Findings on digital breast tomosynthesis images might have been present and perceptible by radiologists on the previous digital mammography, but considered non-suspicious and interpreted as negative, or positive but dismissed on the consensus. An absence of obvious findings on the current synthetic 2D mammogram or the slabs and planes, or both, might have downgraded their interpretation or dismissed cases at consensus.²⁷ Furthermore, the availability of several

practice it in an everyday screening setting until the trial started.¹⁵ Insufficient experience in screen reading of digital breast tomosynthesis is therefore considered a limitation of our study²⁶ and it is possible that the radiologists had not yet achieved optimal screen reading capabilities with digital breast tomosynthesis at the start-up of the trial.

Despite this potential limitation, time used for screen reading was lower in the To-Be trial than reported from the Oslo Tomosynthesis Screening Trial, 66 sec for digital breast tomosynthesis and 39 sec for digital

sets of previous mammograms could have been distracting instead of elucidative to the radiologists. These factors might have affected the proportions of recall and breast cancer detection, and the histopathological tumour characteristics in the digital breast tomosynthesis group. We presume that recalling women with these findings for further assessment might have led to a higher detection of breast cancer in our study; however, we do not know whether these lesions were small tumours of low histological grade or small aggressive breast cancers, also termed killing cancers.

Our results on histopathological tumour characteristics are in line with those from the Malmö tomosynthesis trial.⁵ The distribution of tumour diameter, histological grade, and tumour cell proliferation index by Ki67 staining indicated an increase in the detection of node-negative progressive tumours, which could be considered favourable for digital breast tomosynthesis. The relatively high number of cases with no available tumour diameter might have affected the mean tumour diameter in our study and is likely to be related to new guidelines that recommend neoadjuvant treatment for some tumours 20 mm or larger in diameter.²⁸ Our histopathological findings diverge somewhat from some of the other studies, which have shown an increase in small, low histological grade breast cancers in groups screened with digital breast tomosynthesis.^{18,29}

To investigate a possible study effect, we compared recall in Bergen before the trial, 2008–15, with results from the trial. The recall before the trial did not differ from the recall for digital breast tomosynthesis, but did differ from the recall for digital mammography. The change in recall in Bergen is probably a study effect and is expected to affect both groups equally; however, our study was designed to compare results of digital breast tomosynthesis versus digital mammography regardless of this underlying effect. Recall in Norway, Bergen excluded, was lower than in Bergen irrespective of the trial. The breast cancer detection in our trial did not differ from the pretrial period in Bergen, whereas it did differ from that in Norway, Bergen excluded. This difference could be due to a study effect, including the radiologists' experience and training in tomosynthesis, first generation equipment, and other factors we have discussed. Additionally, the increase in screen-detected breast cancer in Norway could be due to a general improvement in screening such as from screen film to digital mammography, as well as random variation.

Despite several studies showing a higher proportion of screen-detected breast cancer with digital breast tomosynthesis than digital mammography, the implementation of digital breast tomosynthesis in population-based screening programmes has been anything but rapid. This slow uptake could be because of an absence of evidence from studies with sufficiently large study samples and robust study designs, as well as the problematic finding of an increased incidence of small,

low-histological-grade tumours without a corresponding decrease in the proportion of interval breast cancers. Given the results described above regarding screen-detected breast cancer for digital breast tomosynthesis, and because our trial is one of the earliest programme-embedded randomised, controlled trials of digital breast tomosynthesis screening, we present a reflection on factors that might have shaped our study findings in order to inform future breast cancer screening research.

Running a randomised, controlled trial of new technology in an everyday screening setting is a challenging task, which the trial staff managed proficiently by masking interim results, using a closed pool of radiographers and radiologists, and using standardised imaging, screen reading, and consensus procedures in both groups; our findings therefore reflect real-world screening outcomes for a randomised, controlled trial.

The completeness and data quality of the Cancer Registry of Norway provided an opportunity to link the trial data with registry data on individual screening episodes, including any diagnosis as a part of, or outside, BreastScreen Norway. The population-based randomised, controlled trial design, high participation, and a high level of data completeness because of linkage with the Cancer Registry of Norway are all strengths of the trial. However, there are some limitations to our trial. First, our assumed proportion of screen-detected breast cancer in the digital breast tomosynthesis group (based on current knowledge at study inception) was somewhat exaggerated, leading to diminished statistical power. In retrospect, a combination of a superiority and non-inferiority randomised, controlled trial might have been a better design,³⁰ given the observed detection and the final study sample size. For the observed proportion of breast cancer, we would have needed about 400 000 women in each group to detect a statistically significant difference.

Increasing the sample size in our trial would have been difficult, because an extension of study period would have resulted in the inclusion of women with digital breast tomosynthesis as a previous screening exam, instead of increasing the sample size of women first-time screened with digital breast tomosynthesis. Because we included all women who attended the screening unit in Bergen during the recruitment period, a multicentre study would be the only way to increase the study population.

Another limitation is that the trial was single centre, using equipment from one vendor. Our results might therefore have limited generalisability to other settings. Additionally, we do not currently have a sufficiently long follow-up period to report on interval breast cancers or breast cancer mortality. A longer follow-up is also needed for analyses on patient-reported outcome measures, which is in our long-term plan. Further, we expect to stratify our findings by mammographic density in later analyses.

Lastly, although we did several analyses, we found that using a statistical correction for multiple testing, such as the Bonferroni correction, would not alter our conclusions.

In summary, our population-based, breast cancer screening trial using first-generation digital breast tomosynthesis with synthetic 2D mammogram versus digital mammography alone did not identify a significant difference in the proportion of screen-detected breast cancer. The distribution of histopathological tumour characteristics did not differ between the two groups either. A lower recall and a higher positive predictive value for recall was found for women screened with digital breast tomosynthesis compared with those screened with digital mammography. Our results indicate that use of digital breast tomosynthesis in a screening setting is safe for women at average risk of breast cancer. Further studies with follow-up that analyse interval and screen-detected breast cancer in the consecutive screening rounds are needed to better understand the effect of digital breast tomosynthesis in a population-based screening setting.

Contributors

All authors are guarantors of integrity of entire study. All authors contributed to study concepts or design, or both; data acquisition; or data analysis or interpretation, or both. SH, ÅSH, NH, SS, and LAA drafted the Article or revised the Article for important intellectual content. All authors edited the Article and approved of the final submitted version. All authors agree to ensure any questions related to the work are appropriately resolved. SH, ÅSH, and NH did the literature research. SH, ÅSH, NH, SS, and LAA did the statistical analysis.

Declaration of interests

NH was supported by a grant from the National Breast Cancer Foundation Cancer Research Leadership Fellowship. SH has permanent employment as a researcher at the Cancer Registry of Norway, independent of her job as administrative leader of BreastScreen Norway. All other authors declare no competing interests.

Data sharing

Information on data sharing is in the appendix (p 4).

Acknowledgments

The study was funded by Cancer Registry of Norway, Department of Radiology at Haukeland University Hospital, the University of Oslo, and the Research Council of Norway (the To-Be trial; project number 247941/H10 and the Centres of Excellence funding scheme; project number 223250). We thank the radiographers and radiologists at the breast centre at the Department of Radiology and at the breast centre at Haukeland University Hospital, and those who screened the women and did the screen reading for their support and enthusiasm for the study. Thanks to Anders S Danielsen and Kaitlyn Tsuruda at the Cancer Registry for help during the revision of the manuscript.

References

- 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**: 942–49.
- 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**: 583–89.
- 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**: 1105–13.
- Skaane P, Bandos AI, Gullien R, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology* 2013; **267**: 47–56.
- Zackrisson S, Lang K, Rosso A, et al. One-view breast tomosynthesis versus two-view mammography in the Malmö Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. *Lancet Oncol* 2018; **19**: 1493–503.
- 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**: 109–16.
- 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. *AJR Am J Roentgenol* 2014; **203**: 687–93.
- Friedewald SM, Rafferty EA, Conant EF. Breast cancer screening with tomosynthesis and digital mammography-reply. *JAMA* 2014; **312**: 1695–96.
- 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**: pii: dju316.
- Pattacini P, Nitrosi A, Giorgi Rossi P, et al. Digital mammography versus digital mammography plus tomosynthesis for breast cancer screening: the Reggio Emilia Tomosynthesis randomized trial. *Radiology* 2018; **288**: 375–85.
- Houssami N. Evidence on synthesized two-dimensional mammography versus digital mammography when using tomosynthesis (three-dimensional mammography) for population breast cancer screening. *Clin Breast Cancer* 2017; **18**: 255–60.
- Wahab RA, Lee SJ, Zhang B, Sobel L, Mahoney MC. A comparison of full-field digital mammograms versus 2D synthesized mammograms for detection of microcalcifications on screening. *Eur J Radiol* 2018; **107**: 14–19.
- 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.
- Hofvind S, Tsuruda K, Mangerud G, et al. The Norwegian Breast Cancer Screening Program, 1996–2016: celebrating 20 years of organised mammographic screening. In: Larsen IK, ed. *Cancer in Norway 2016—cancer incidence, mortality, survival and prevalence in Norway*. Oslo: Cancer Registry of Norway, 2017.
- Aase HS, Holen AS, Pedersen K, 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**: 1175–86.
- Hortobagyi GN, Connolly JL, D'Orsi CJ, et al. Chapter 48. In: Amin MB, Edge S, Greene F, et al, eds. *AJCC Cancer Staging Manual*. Springer International Publishing, Eighth edn, 2017: p 589.
- Goldhirsch A, Winer EP, Coates AS, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol* 2013; **24**: 2206–23.
- Hofvind S, Hovda T, Holen AS, et al. Digital breast tomosynthesis and synthetic 2D mammography versus digital mammography: evaluation in a population-based screening program. *Radiology* 2018; **287**: 787–94.
- 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**: 655–63.
- 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**: 70–76.
- Mackenzie A, Kaur S, Elangovan P, Dance DR, Young KC, eds. Comparison of synthetic 2D images with planar and tomosynthesis imaging of the breast using a virtual clinical trial. *SPIE Medical Imaging* 2018; published online Mar 7. DOI:10.1117/12.2293070

- 22 Rodriguez-Ruiz A, Lardenoije S, Wanders AJT, Sechopoulos I, Mann RM. Comparison of breast cancer detection and depiction between planar and rotating synthetic mammography generated from breast tomosynthesis. *Eur Radiol* 2018; **108**: 78–83.
- 23 Rodriguez-Ruiz A, Teuwen J, Vreemann S, et al. New reconstruction algorithm for digital breast tomosynthesis: better image quality for humans and computers. *Acta Radiol* 2018; **59**: 1051–59.
- 24 Tucker L, Gilbert FJ, Astley SM, et al. Does reader performance with digital breast tomosynthesis vary according to experience with two-dimensional mammography? *Radiology* 2017; **283**: 371–80.
- 25 Alakhras MM, Brennan PC, Rickard M, Bourne R, Mello-Thoms C. Effect of radiologists' experience on breast cancer detection and localization using digital breast tomosynthesis. *Eur Radiol* 2015; **25**: 402–09.
- 26 Skaane P, Bandos A, Niklason L, et al. Digital mammography versus digital mammography plus tomosynthesis in breast cancer screening: the Oslo tomosynthesis screening trial. *Radiology* 2019; published online Feb 19. DOI:10.1148/radiol.2019182394.
- 27 Iotti V, Rossi PG, Nitrosi A, et al. Comparing two visualization protocols for tomosynthesis in screening: specificity and sensitivity of slabs versus planes plus slabs. *Eur Radiol* 2019; published online Feb 8. DOI:10.1007/s00330-018-5978-x.
- 28 Curigliano G, Burstein HJ, P Winer E, et al. De-escalating and escalating treatments for early-stage breast cancer: the St Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann Oncol* 2017; **28**: 1700–12.
- 29 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**: 391–97.
- 30 Snapinn SM. Noninferiority trials. *Curr Control Trials Cardiovasc Med* 2000; **1**: 19–21.