

Digital breast tomosynthesis plus synthesised mammography versus digital screening mammography for the detection of invasive breast cancer (TOSYMA): a multicentre, open-label, randomised, controlled, superiority trial



Walter Heindel*, Stefanie Weigel*, Joachim Gerß, Hans-Werner Hense, Alexander Sommer, Miriam Krischke, Laura Kerschke, for the TOSYMA Screening Trial Study Group†

Summary

Background Two dimensional (2D) full-field digital mammography is the current standard of breast cancer screening. Digital breast tomosynthesis generates pseudo-three dimensional datasets of the breast from which synthesised 2D (s2D) mammograms can be reconstructed. This innovative approach reduces the likelihood of overlapping breast tissues that can conceal features of malignancy. We aimed to compare digital breast tomosynthesis plus s2D mammography with digital screening mammography for the detection of invasive breast cancer.

Methods TOSYMA was a randomised, open-label, superiority trial done at 17 screening units in two federal states of Germany. Eligible participants were women aged 50–69 years who had been invited to participate in a population-wide, quality-controlled mammography screening programme. Women were randomly assigned (1:1) to digital breast tomosynthesis plus s2D mammography or digital mammography alone using block randomisation (block size of 32), stratified by site. The primary endpoints were the detection rate of invasive breast cancer and invasive interval cancer rate at 24 months, analysed in the modified full analysis set, which included all randomly assigned participants who underwent either type of screening examination. Ten examinations, corresponding to a second study participation, were excluded. Analyses were done according to the intention-to-treat principle. Interval cancer rates will be reported in the follow-up study. Safety was assessed in the as-treated population, which included all participants who were randomly assigned. This trial is registered with ClinicalTrials.gov, NCT03377036, and is closed to accrual.

Findings Between July 5, 2018, and Dec 30, 2020, 99 689 women were randomly assigned to digital breast tomosynthesis plus s2D mammography (n=49 804) or digital mammography (n=49 830). Invasive breast cancers were detected in 354 of 49 715 women with evaluable primary endpoint data in the digital breast tomosynthesis plus s2D group (detection rate 7·1 cases per 1000 women screened) and in 240 of 49 762 women in the digital mammography group (4·8 cases per 1000 women screened; odds ratio 1·48 [95% CI 1·25–1·75]; p<0·0001). Adverse events and device deficiencies were rare (six adverse events in each group; 23 device deficiencies in the digital breast tomosynthesis plus s2D group vs five device deficiencies in the digital mammography group) and no serious adverse events were reported.

Interpretation The results from this study indicate that the detection rate for invasive breast cancer was significantly higher with digital breast tomosynthesis plus s2D mammography than digital mammography alone. Evaluation of interval cancer rates in the follow-up study will further help to investigate incremental long-term benefits of digital breast tomosynthesis screening.

Funding Deutsche Forschungsgemeinschaft (German Research Foundation).

Copyright © 2022 Elsevier Ltd. All rights reserved.

Introduction

GLOBOCAN 2020 estimates of cancer incidence have shown that breast cancer in women has surpassed lung cancer as the most commonly diagnosed cancer worldwide.¹ An estimated 2·3 million new cases of breast cancer are diagnosed annually, representing 11·7% of all cancer cases. Breast cancer is the leading cause of cancer morbidity and mortality in women worldwide, accounting for 685 000 deaths per year.

In addition to new therapeutic approaches, the International Agency for Research on Cancer have

recommended that population-based mammography screenings are offered every 2 years for women aged 50–69 years who are at average risk of breast cancer.² Mammography screening is an essential component in reducing breast cancer associated mortality by around 20%.³

One of the disadvantages of two dimensional full-field digital mammography, which is the current screening standard, is that overlapping breast tissues can conceal features of malignancy resulting in delayed diagnosis.⁴ Digital breast tomosynthesis is an imaging technology

Lancet Oncol 2022; 23: 601–11

Published Online

April 12, 2022

[https://doi.org/10.1016/S1470-2045\(22\)00194-2](https://doi.org/10.1016/S1470-2045(22)00194-2)

See [Comment](#) page 554

For the German translation of the abstract see Online for appendix 1

*Contributed equally

†Members of the TOSYMA Screening Trial Study Group are listed in appendix 2 (p 19)

Clinic for Radiology and Reference Center for Mammography Münster (Prof W Heindel MD, Prof S Weigel MD, A Sommer Dipl Ing), Institute of Biostatistics and Clinical Research (J Gerß Dr rer nat, L Kerschke Dr rer medic), Institute of Epidemiology and Social Medicine (Prof H-W Hense MD), Centre for Clinical Trials Münster (M Krischke Dr rer nat), University of Münster and University Hospital Münster, Münster, Germany

Correspondence to: Prof Walter Heindel, Clinic for Radiology and Reference Center for Mammography Münster, University of Münster and University Hospital Münster, D-48149 Münster, Germany heindel@uni-muenster.de

See Online for appendix 2

Research in context

Evidence before this study

We searched MEDLINE between July 1, 2012, and Sept 30, 2017 for English language publications that included “breast cancer screening”, “tomosyn\$”, and “synthes\$” in the title or abstract. No randomised controlled trials of breast cancer screening with digital breast tomosynthesis plus synthesised two-dimensional (s2D) mammography were identified. Two reviews comparing diagnostic studies of digital breast tomosynthesis plus digital mammography versus digital breast tomosynthesis plus s2D mammography showed by comparison that both approaches had a similar screening performance. Another study, using parallel double-readings conducted sequentially for digital mammography and digital breast tomosynthesis images acquired in the same individuals, reported significantly higher detection rates for invasive cancer with digital breast tomosynthesis plus s2D mammography. In other studies, the detection rates for invasive cancer were higher with digital breast tomosynthesis plus s2D mammography screening than digital mammography alone, whereas recall rates were similar between groups in the prospective studies and lower in the digital breast tomosynthesis plus s2D mammography group than the digital mammography group in retrospective analyses. Most studies had a substantial risk of bias due to inadequate reporting of cohort assignments and inclusion of patients from different calendar periods. At the time of TOSYMA initiation, a randomised, controlled, parallel-group screening trial comparing digital breast tomosynthesis plus

s2D mammography with digital mammography in a single centre, using one type of machine, within a population-based cancer screening programme, was ongoing in Norway, but no results had been published. Thus, robust evidence indicating the superiority of digital breast tomosynthesis plus s2D mammography versus standard digital mammography in population-based screening programmes was scarce.

Added value of this study

This randomised, controlled screening trial is—to our knowledge—the first randomised, controlled trial in a multicentre and multivendor setting to compare digital breast tomosynthesis plus s2D mammography with digital mammography alone. In this large, randomised, parallel-group setting, the detection rate of invasive breast cancers was 48% higher in groups screened with digital breast tomosynthesis plus s2D mammography than digital mammography alone.

Implications of all the available evidence

The currently available evidence, including findings from TOSYMA-1, seems to suggest that digital breast tomosynthesis plus s2D mammography is superior to standard digital mammography for the detection of invasive breast cancer. The study design with extended follow-up (TOSYMA-2) will provide further insights about the prognostic implications of increased detection of invasive breast cancer at screening by determination of interval breast cancer incidence at 24 months in both groups.

that generates pseudo-three dimensional datasets of the breast, thus potentially reducing the likelihood of overlapping tissue.^{5–8} The major advantage of supplementing digital mammography with digital breast tomosynthesis is a substantial improvement in diagnostic accuracy; however, this subsequently increases radiation dose as a result of duplicate exposures.^{7–10} This adverse effect might be avoided through the combination of digital breast tomosynthesis with synthesised two-dimensional (s2D) mammography images reconstructed without any additional radiation exposures directly from the digital breast tomosynthesis datasets (ie, digital breast tomosynthesis plus s2D mammography).¹⁰

Previous reviews of paired and unpaired studies showed that diagnostic performance was improved with digital breast tomosynthesis plus s2D mammography and concluded that new studies should preferentially investigate digital breast tomosynthesis plus s2D mammography as an advanced imaging technology.^{11–13} To date, only one randomised, controlled, screening trial has been done, reporting outcomes for screening with digital breast tomosynthesis plus s2D mammography in a Norwegian population-based breast cancer screening programme (To-Be), which was published in 2019.¹⁴ In contrast to previous studies, results from To-Be showed no significant differences in the proportion of

screen-detected breast cancers or histopathological tumour characteristics for digital breast tomosynthesis plus s2D mammography versus digital mammography alone. Thus, evidence for the superiority of the digital breast tomosynthesis screening strategy is scarce.

The European Commission has conditionally recommended the use of digital breast tomosynthesis or digital mammography for organised screening programmes. This decision takes into account the balance between desirable and undesirable effects that might favour digital breast tomosynthesis overall, but in the context of low certainty regarding available evidence and scarcity of data on the downstream impact of the interventions on mortality and morbidity.¹⁵

Here, we report the first results from the screening phase of the TOmosynthesis plus SYnthesised MAMmography Study (TOSYMA),¹⁴ a large, randomised, controlled trial embedded in the German mammography screening programme, comparing digital breast tomosynthesis plus s2D mammography with digital mammography alone (TOSYMA-1).

Methods

Study design and participants

TOSYMA is a multicentre, open-label, two-stage adaptive, parallel-group, randomised, controlled, superiority trial

embedded within the population-wide, quality-controlled mammography screening programme in Germany.¹⁶ The trial compared breast cancer screening examinations performed with digital breast tomosynthesis plus s2D mammography with full-field digital mammography.¹⁷ Participants were recruited from 17 screening units in the federal states of North Rhine-Westphalia and Lower Saxony, Germany (appendix 2 pp 6, 12).

Eligible women were residents aged 50–69 years who were eligible for participation in the German mammography screening programme. We excluded women with a breast cancer diagnosis in the 5 years before screening invitation or those who had had a mammogram in the past 12 months. Eligible women living in the catchment areas of the participating screening units received, together with the regular invitation letter, a study invitation with information about the objectives and design of TOSYMA. Breast implants or previous TOSYMA participation were specific exclusion criteria for trial participation.

Women were enrolled into the prospective study after written, informed consent was obtained. The study protocol^{16,18} (including two amendments) was approved by the ethics committee of the Medical Association of Westphalia-Lippe and the Westphalian Wilhelms University of Münster (2016-132-f-S), and by the local ethics committee at each participating study centre. An independent data monitoring committee regularly reviewed unblinded data for patient safety.

Randomisation and masking

Participants were randomly assigned (1:1) to screening with digital breast tomosynthesis plus s2D mammography (intervention group) or digital mammography (control group). Randomisation was done in blocks of 32 and stratified by site. The randomisation sequence was computer-generated by an independent statistician not involved in the data analysis and was inaccessible to the study site personnel performing the randomisation. The sequence was implemented into the certified mammography screening programme documentation software MaSc (versions 5.0 and 6.0), which automatically outputted the randomisation result. Study site personnel doing the screening examination and readers were aware of treatment allocation. Participants were not aware of allocation before completion of the screening examination. The statistician performing data analysis was masked until the unblinded interim analysis was performed.

Procedures

A trial-specific software component embedded into the MaSc software was used for participant registration and documentation. The type of mammographic devices used in the intervention and control group were the exclusive responsibility of each study centre, and generally reflected the technical state of the equipment at the time of study initiation. Details of specific technical requirements for participation in the study, trial-related training, and

qualification procedures of the study personnel are presented in appendix 2 (p 2).

Screening examinations in both groups acquired craniocaudal and mediolateral oblique views of each breast. Unlike digital mammography, for digital breast tomosynthesis, low-dose x-rays were used to acquire multiple x-ray projections across an arc; for craniocaudal and mediolateral oblique views, oblique slices of 1 mm thickness or thinner were reconstructed and shown as stacked images. Additionally, 2D images were synthesised for each view from the digital breast tomosynthesis dataset.

Independent double reading of mammograms was performed in both study groups in accordance with the European Guidelines.¹⁷ The readers received the screening examinations in their digital work list according to the acquisition order and had no option to select examinations of a specific study group before reading. In case of any suspicious abnormality, reading results were clarified with an arbitrator to decide whether women had to be recalled for further diagnostic tests. Diagnostic procedures were concordant with the regular screening programme and for both study groups included clinical examinations, ultrasound, additional mammographic views (eg, magnifications views), tomosynthesis, MRI, and invasive assessments. Specific to the study, each unit had access to digital breast tomosynthesis-guided biopsy equipment. All processes and results were documented on site. All pathologists collaborating in the study were certified breast cancer screening pathologists and had received specific study training. As in routine operation, a reference pathologist was available for consultation. The documentations of all invasive breast cancers were validated by the responsible physicians at all study sites to ensure consistence with the original data.

Adverse events and device deficiencies occurring on the day of mammography screening were assessed as safety data by the study site personnel and documented in the MaSc software. Furthermore, the glandular radiation dose was monitored for quality assurance and safety. The data monitoring committee reviewed safety data in 3-month intervals. Serious adverse events and device deficiencies potentially causing serious adverse events had to be reported to the principle investigator (WH) within 3 working days.

Outcomes

The primary endpoints were the detection rate of histologically confirmed invasive breast cancer at screening and the invasive interval cancer rate at 24-months, defined as the proportion of women who were negative at screening (ie, no ductal carcinoma in situ [DCIS] or invasive breast cancer detection) and developed an invasive breast cancer in the 24-month interval after the screening examination. Invasive interval cancer rate will be assessed in an extended prospective follow-up phase (TOSYMA-2) in collaboration with the state cancer registries and the results for this endpoint will be published elsewhere.

For the study protocol see <https://bmjopen.bmj.com/content/bmjopen/8/5/e020475.full.pdf>

Secondary endpoints were the detection rate of invasive breast cancers with tumour size pT1 (tumour ≤ 20 mm in largest dimension), the detection rate of DCIS, the recall rate for further assessment, the positive predictive value of recall (PPV1), and the cumulative incidence of invasive interval cancers at 12 months (appendix 2 p 11). The incidence of invasive interval cancers at 12 months will also be assessed after the required follow-up phase in TOSYMA-2 and will be reported elsewhere.

Statistical analysis

The trial had a two-stage adaptive design using a conditional error function with a binding futility-bound of 0.5 and a two-sided multiple significance level of 5% for both primary endpoints (appendix 2 pp 2–3). The original sample size calculation was based solely on the first primary endpoint, assuming an invasive breast cancer detection rate of 4.4 cases per 1000 women screened in the control group (based on 2009–10 evaluation data from the Reference Center for Mammography Münster for North Rhine-Westphalia) and an assumed clinically relevant increase of 33% in the intervention group. The resulting sample size at 80% power was 80 000 participants in total across both study groups.¹⁶ Considering the intensified discussion about interval cancer rates as an instrument to assess the potential for overdiagnosis,^{7,15} the data monitoring committee supported to revise the sample size calculation to achieve reasonable power for the evaluation of both primary endpoints. Therefore, a sample size amendment (invasive breast cancer detection rate of 5.3 cases per 1000 women screened in the control group; invasive interval cancer rate of 1.92 per 1000 negative examinations in the control group and 1.25 per 1000 negative examinations in the intervention group) was implemented before the planned interim analysis and without knowledge of any study data from the ongoing trial, with a new total of 120 000 participants required to achieve statistical power of 80% (appendix 2 p 2).¹⁸

Trial recruitment started on July 5, 2018.¹⁶ The interim analysis was performed as planned on Jan 31, 2020, including participants recruited until Aug 30, 2019 ($n=40\,129$ participants in the intention-to-treat population). The conditional power to confirm the one-sided (alternative) hypothesis that the detection rate for invasive breast cancer was higher with digital breast tomosynthesis plus s2D mammography than with digital mammography in the final analysis at a one-sided α of 2.5% was calculated based on a prespecified algorithm (appendix 2 pp 3–4). The result was disclosed to the data monitoring committee who informed the principal investigator that recruitment should continue as planned.

The study was affected by COVID-19 pandemic: considering the risk of infection, screening was paused between March 30 and April 30, 2020. As a consequence of this unforeseeable delay, recruitment had to be stopped through a protocol amendment on Dec 30, 2020, before

achieving recruitment of the intended sample size of 120 000 women.¹⁸ Considering the biennial invitation interval in the mammography screening programme, this termination of recruitment was necessary to prevent an increasing number of women, who had already been invited to participate in the study at the beginning of the recruitment period, from being reinvited. Since each woman could participate only once in the trial, continuation of recruitment beyond the end of 2020 would have increased the proportion of women who were attending screening for the first time in the study sample specifically, distorting the comparability with the screening population in routine practice. The amendment was approved by the data monitoring committee.

The first primary endpoint of invasive breast cancer detection rate was analysed using a one-sided Cochran-Mantel-Haenszel test with stratification by site. Reported p values are one-sided to avoid potential directional conflicts caused by two-sided p values.¹⁹ An overall p value combining the p values of both trial stages (appendix 2 p 7) was calculated as described previously,²⁰ and a p value less than 0.025 was deemed significant.

Secondary endpoints were analysed without formal significance testing. For primary and secondary endpoints, point estimates, odds ratios (ORs), and risk differences adjusted for study site, and corresponding 95% CIs are reported. Rates of invasive breast cancer detection, pT1 detection, and DCIS detection are presented per 1000 women screened. Rates of recall and PPV1 are presented as percentages. All outcomes were analysed in the modified full analysis set, which included all randomly assigned participants who underwent either type of screening examination. Ten examinations, corresponding to a second study participation, were excluded. Analyses were performed according to the intention-to-treat principle. Hereafter, we refer to the modified full analysis set analysed according to the intention-to-treat principle as the intention-to-treat population.

The detection rate of invasive breast cancer and the recall rate were additionally assessed within the predefined subgroups of first and subsequent screening round participants and by age groups. We additionally assessed the rate of performed biopsies and the positive predictive value of performed biopsies (PPV3) as performance indicators of a screening programme,¹⁷ as well as the image acquisition and reading time. These analyses were performed descriptively and were not prespecified in the study protocol (post-hoc analyses).

Safety data were analysed descriptively, in the as-treated population, which included all participants who were randomly assigned. Device deficiencies that led to a deviation from the randomised study group were evaluated on the basis of the allocated intervention.

Since the amount of missing outcome data was small ($<0.1\%$), we did a complete-case analysis for all endpoints. We did a sensitivity analysis of the first

primary endpoint in which missing outcome data were imputed according to different simulation scenarios (appendix 2 pp 4–5). A further sensitivity analysis was performed in the per protocol population, which included all participants from the full analysis set without major protocol deviations; major protocol deviations were defined as deviations from inclusion and exclusion criteria or the assigned study group, performance of the screening examination with a mammography device not part of the study, incomplete mammography examination (ie, initial breast imaging), and study withdrawal.

Analyses were performed with SAS (version 9.4) and R (version 4.0.2) statistical software. The trial is registered with ClinicalTrials.gov, NCT03377036.

Role of the funding source

The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Between July 5, 2018, and Dec 30, 2020, 99 689 participating women were randomly assigned to digital breast tomosynthesis plus s2D mammography (n=49 804) or digital mammography (n=49 830; figure). 49 762 women allocated to the digital breast tomosynthesis plus s2D mammography group and 49 796 women allocated to the digital mammography group underwent screening and were participating in the study for the first time and thus were included in the modified full analysis set; 66 women did not undergo screening and ten examinations corresponded to a second study participation and thus were excluded from the full analysis set. Allocation per trial stage is shown in appendix 2 (p 8). Mammography examinations were performed using seven different devices supplied by five vendors (appendix 2 p 13). The baseline characteristics of the study participants were similar in both groups (table 1) and consistent between groups recruited before and after the interim analysis (appendix 2 p 14).

In participants with evaluable data, invasive breast cancer was detected in 354 of 49 715 women assigned to the digital breast tomosynthesis plus s2D mammography group (detection rate 7.1 cases per 1000 women screened) and 240 of 49 762 women assigned to the digital mammography group (4.8 cases per 1000; table 2). The detection rate for invasive cancer was significantly higher in the digital breast tomosynthesis plus s2D mammography group than the digital mammography group (OR 1.48 [95% CI 1.25–1.75]; $p<0.0001$). The increase in detection rate with digital breast tomosynthesis plus s2D mammography was higher in the second stage of the trial (ie, after the interim analysis) than in the first stage (appendix 2 p 15). The results were similar in the per protocol analysis and the sensitivity analysis for missing values (appendix 2 pp 16–17).

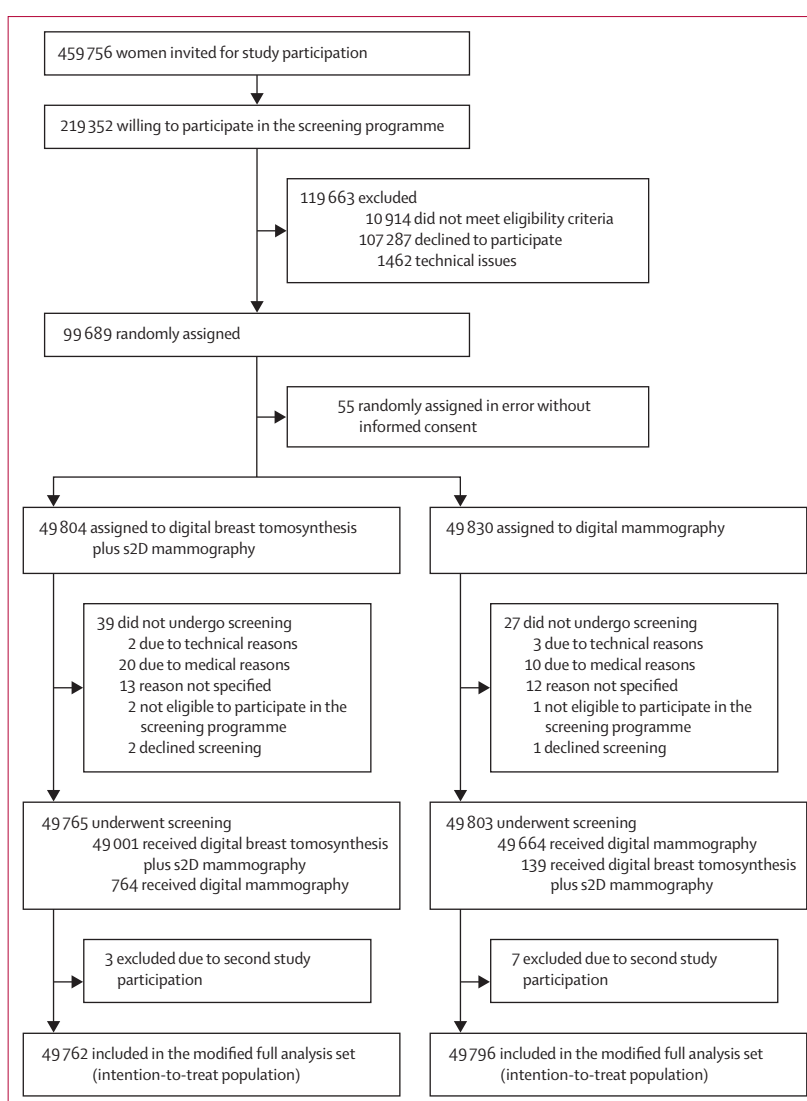


Figure: Trial profile
s2D=synthesised two dimensional.

The invasive cancer detection rate for breast cancers with tumour size pT1 was substantially higher in the digital breast tomosynthesis plus s2D mammography group than the digital mammography group (OR 1.73 [95% CI 1.41–2.13]; table 2). No marked differences were identified in the detection rate of DCIS between the groups (OR 0.94 [95% CI 0.65–1.35]). Recall rates did not differ between the study groups (OR 0.98 [0.92–1.03]), and the PPV1 was higher in the digital breast tomosynthesis plus s2D mammography group (OR 1.50 [1.28–1.77]).

Characteristics of the detected breast cancers are shown in table 3. 57 (16.1%) of 354 invasive carcinomas detected with digital breast tomosynthesis plus s2D mammography and 42 (17.8%) of 236 invasive carcinomas detected with digital mammography were treated with neoadjuvant

therapy. The frequency of node-positive invasive cancers was low in both study groups (42 [11·9%] of 354 invasive breast cancers in the digital breast tomosynthesis plus s2D mammography group vs 37 [15·4%] of 240 invasive breast cancers in the digital mammography group).

The detection rate for invasive breast cancer with digital breast tomosynthesis plus s2D mammography was substantially higher among women who had attended subsequent screening rounds and were older than 60 years than in women who participated in screening for the first time and were younger than 60 years, respectively, whereas the recall rates did not differ by screening round or age groups (tables 4, 5).

In post-hoc analyses, we found that the rate of performed biopsies was higher in the digital breast tomosynthesis plus s2D mammography group (16·3 per 1000 women

screened) than for digital mammography (12·2 per 1000 women screened; OR 1·35 [95% CI 1·21–1·50], and the PPV3 was comparable between the digital breast tomosynthesis plus s2D mammography group (412 [51·5%] of 800 biopsies) and digital mammography group (305 [50·5%] of 604 biopsies; OR 1·04 [0·83–1·29]).

The median image acquisition time was 220·0 s (IQR 178·2–285·0) for digital breast tomosynthesis plus s2D mammography and 186·0 s (154·8–225·0) for digital mammography. The median reading time was 109·0 s (71·4–172·8) for digital breast tomosynthesis plus s2D mammography and 54·0 s (33·0–91·2) for digital mammography.

Adverse events were rare (six events per group; appendix 2 p 18) and no events were classified as serious or related to the device used. Device deficiencies were also rare (23 [0·05%] of 49 179 examinations performed with digital breast tomosynthesis plus s2D mammography vs five [0·01%] of 50 455 examinations performed with digital mammography) and none was classified as a potential cause of a serious adverse event.

The median glandular dose was 1·86 mGy (IQR 1·48–2·45) for the digital breast tomosynthesis plus s2D mammography group and 1·36 mGy (1·02–1·85) for the digital mammography group, and the median thickness of the compressed breast was 59·0 mm (IQR 49·0–68·0) in both study groups.

Discussion

The large, randomised, pragmatic TOSYMA trial, conducted in a population-wide routine mammography screening programme in two federal states of Germany, showed that the detection rate of invasive breast cancers was 48% higher with tomosynthesis with synthetic mammography than standard full-field digital

	Digital breast tomosynthesis plus s2D mammography (n=49 762)	Digital mammography (n=49 796)
Age of participating women, years		
50–54	18 292 (36·8%)	18 421 (37·0%)
55–59	12 649 (25·4%)	12 573 (25·2%)
60–64	11 045 (22·2%)	10 891 (21·9%)
65–70	7776 (15·6%)	7911 (15·9%)
Screening round		
First round	8613 (17·3%)	8653 (17·4%)
Regular subsequent round	37 963 (76·3%)	37 925 (76·2%)
Irregular subsequent round	3186 (6·4%)	3218 (6·5%)
Data are n (%). s2D=synthesised two dimensional. Data on race or ethnicity were not recorded.		
Table 1: Baseline characteristics of the intention-to-treat population of the TOSYMA trial		

	Digital breast tomosynthesis plus s2D mammography (n=49 762)		Digital mammography (n=49 796)		Rate difference* (95% CI)	OR* (95% CI)
	n/N	Detection rate (per 1000 women screened)	n/N	Detection rate (per 1000 women screened)		
Primary outcome						
Invasive breast cancer detection	354/49 715	7·1	240/49 762	4·8	2·3 per 1000 (1·3 to 3·3)	1·48 (1·25 to 1·75)
Secondary outcomes						
pT1 detection†	255/49 715	5·1	148/49 758	3·0	2·2 per 1000 (1·4 to 2·9)	1·73 (1·41 to 2·13)
DCIS detection	62/49 715	1·2	66/49 762	1·3	–0·1 per 1000 (–0·5 to 0·4)	0·94 (0·65 to 1·35)
Recall	2457/49 756 (4·9%)	..	2515/49 794 (5·1%)	..	–0·1 (–0·4 to 0·2)	0·98 (0·92 to 1·03)
PPV1	416/2416 (17·2%)	..	306/2483 (12·3%)	..	5·1 (3·1 to 7·0)	1·50 (1·28 to 1·77)
n/N=women with event/total women with evaluable data. OR=odds ratio. pT1=tumour size ≤20 mm in largest dimension. DCIS=ductal carcinoma in situ. s2D=synthesised two dimensional. PPV1=positive predictive value of recall. *Adjusted for study site (ie, screening unit). †Women with invasive breast cancer for which the final pathological characterisation was obtained after neoadjuvant therapy were evaluated as having no pT1 detection.						
Table 2: Summary of outcomes in the intention-to-treat population						

mammography. In planning the sample size of TOSYMA, we considered a relative difference between study groups of at least 33% to be clinically relevant, which was consistent with the study evidence available when the study was designed. Our findings confirm our prestudy assumptions.

The results for the first primary endpoint corroborate findings from previous reports. A previous meta-analysis¹² analysed retrospective and prospective, non-randomised studies from the USA, Italy, and Norway and found higher detection rates for invasive breast cancer with digital breast tomosynthesis plus s2D mammography (median 5.68 cases per 1000 screened) than digital mammography (3.42 cases per 1000 screened). The absolute detection rate for invasive breast cancer in both groups of TOSYMA was higher (7.1 cases per 1000 women screened in the digital breast tomosynthesis plus s2D mammography group vs 4.8 cases in the digital mammography group), but they closely matched those reported by another meta-analysis.¹¹

Contrary to the findings from earlier studies,^{11–13} the To-Be study¹⁴—the only randomised controlled trial done before TOSYMA that used the innovative imaging technology of digital breast tomosynthesis plus s2D mammography done in a single centre—showed no difference in detection rate of invasive breast cancers between digital breast tomosynthesis and digital breast tomosynthesis plus s2D mammography groups. Similar to TOSYMA, To-Be was a randomised, controlled trial, done in a population-based setting in Norway, with high participation rates and a high level of data completeness because of linkage with the Cancer Registry of Norway. In contrast to the multicentre and multivendor approach used in TOSYMA, To-Be, which began in 2016, used only one first-generation device for digital breast tomosynthesis plus s2D mammography. Furthermore, the experience of readers before the start of the To-Be Trial varied such that the programme recommendation of 5000 annual screen readings was met by only half of the radiologists,¹⁴ whereas mammography readers in the TOSYMA trial had to fulfil this national screening standard at least 1 year before participating in the study.^{16,17}

The reading time in the intervention group was higher than in the control group. The median reading time for digital breast tomosynthesis including interpretation of synthetic 2D mammograms in the To-Be trial^{14,21} was 48 s (IQR 33–78), which was markedly shorter than the 109.0 s (71.4–172.8) reported in TOSYMA. Other screening studies reported median digital breast tomosynthesis reading times of 67 s and 94 s.^{22,23} Long reading time might potentially be explained by the reader instruction in TOSYMA to evaluate all reconstructed slices of 1 mm thickness or less in addition to the synthetic mammograms that were presented first to provide an overview and comparison to previous mammograms. We hypothesise that these longer reading times, the larger sample size of TOSYMA, and the multicentre setting might have

	Digital breast tomosynthesis plus s2D mammography (n=416)	Digital mammography (n=306)
Type of breast cancer		
DCIS	62 (14.9%)	66 (21.6%)
Invasive carcinoma	354 (85.1%)	240 (78.4%)
Tumour size, mm		
DCIS*	20.0 (8.0–35.0)	22.5 (6.5–37.5)
Invasive carcinoma†	12.0 (8.0–17.6)	13.0 (8.0–20.0)
Neoadjuvant treated invasive carcinoma‡	57 (16.1%)	42 (17.8%)
pT1 carcinoma§	255 (100%)	148 (100%)
1mic	3 (1.2%)	4 (2.7%)
1a	29 (11.4%)	21 (14.2%)
1b	100 (39.2%)	49 (33.1%)
1c	122 (47.8%)	74 (50.0%)
1x	1 (0.4%)	0
Nodal status of invasive carcinoma†	354 (100%)	240 (100%)
Positive	42 (11.9%)	37 (15.4%)
Negative	305 (86.2%)	195 (81.3%)
Missing	7 (2.0%)	8 (3.3%)

Data are n (%) or median (IQR). DCIS=ductal carcinoma in situ. s2D=synthesised two dimensional. *Includes one DCIS case without residuals in surgical specimen; lesion size was evaluated based on the clinical tumour size. †Includes 99 neoadjuvant-treated tumours and four tumours without final pathological characterisation (one woman died, one woman migrated, and two women did not consent to treatment); these cancers were evaluated based on the clinical tumour size and nodal status. ‡Data on neoadjuvant treatment were not available for four of 240 women who had invasive carcinoma detected with digital mammography (one woman died, one woman migrated, and two women did not consent to treatment); the relative frequency was therefore calculated based on 236 cases. §Percentages calculated based on 255 cases of pT1 carcinoma in the digital breast tomosynthesis plus s2D mammography group and 148 cases in the digital mammography group.

Table 3: Characteristics of breast cancers identified in the two screening groups in the intention-to-treat population

contributed to the findings (ie, the detection rate for invasive breast cancer in the digital breast tomosynthesis plus s2D mammography group of the TOSYMA study was higher than that of the digital mammography group). Sensitive digital breast tomosynthesis reading generally requires more time than digital mammography reading^{21–23} because thorough reading of slices is presumably a prerequisite to increase invasive cancer detection. To facilitate the work of human readers and improve cost-effectiveness, the use of artificial intelligence could help distinguishing digital breast tomosynthesis slices showing normal breast tissue from suspected breast cancer lesions. However, current evidence suggests that at present, artificial intelligence solutions for implementation into clinical practice are insufficient with regard to quality and quantity.²⁴

With the detection rate for invasive breast cancer representing an early screening surrogate parameter, results from TOSYMA point towards possible effects of digital breast tomosynthesis on long-term screening

	Digital breast tomosynthesis plus s2D mammography (n=49 762)		Digital mammography (n=49 796)		Rate difference* (95% CI)	OR* (95% CI)
	n/N	Detection rate (per 1000 women screened)	n/N	Detection rate (per 1000 women screened)		
Invasive breast cancer detection						
First round screening (N=17 266)	61/8594	7.1	52/8633	6.0	1.1 per 1000 (-1.3 to 3.5)	1.18 (0.80 to 1.75)
Regular subsequent round screening (N=75 888)	251/37 942	6.6	172/37 913	4.5	2.1 per 1000 (1.0 to 3.1)	1.46 (1.20 to 1.79)
Irregular subsequent round screening (N=6404)	42/3179	13.2	16/3216	5.0	8.3 per 1000 (3.6 to 13.0)	2.70 (1.48 to 5.15)
Recall						
First round screening (N=17 266)	900/8613 (10.4%)	..	1030/8651 (11.9%)	..	-1.5 (-2.5 to -0.6)	0.85 (0.77 to 0.94)
Regular subsequent round screening (N=75 888)	1383/37 958 (3.6%)	..	1333/37 925 (3.5%)	..	0.1 (-0.1 to 0.4)	1.04 (0.96 to 1.12)
Irregular subsequent round screening (N=6404)	174/3185 (5.5%)	..	152/3218 (4.7%)	..	0.8 (-0.3 to 1.8)	1.17 (0.93 to 1.47)

n/N=women with event/total women with evaluable data. s2D=synthesised two dimensional. OR=odds ratio. *Adjusted for study site (ie, screening unit).

Table 4: Invasive breast cancer detection rates and recall rates by screening status in the intention-to-treat population

	Digital breast tomosynthesis plus s2D mammography (n=49 762)		Digital mammography (n=49 796)		Rate difference* (95% CI)	OR* (95% CI)
	n/N	Detection rate (per 1000 women screened)	n/N	Detection rate (per 1000 women screened)		
Invasive breast cancer detection						
50-54 years (N=36 713)	94/18 173	5.1	78/18 397	4.2	0.9 per 1000 (-0.5 to 2.3)	1.22 (0.89 to 1.66)
55-59 years (N=25 222)	69/12 645	5.5	53/12 569	4.2	1.3 per 1000 (-0.5 to 3.0)	1.30 (0.90 to 1.90)
60-64 years (N=21 936)	91/11 037	8.2	49/10 890	4.5	3.7 per 1000 (1.6 to 5.9)	1.84 (1.28 to 2.66)
65-70 years (N=15 687)	100/7766	12.9	60/7906	7.6	5.3 per 1000 (2.1 to 8.4)	1.70 (1.22 to 2.39)
Recall						
50-54 years (N=36 713)	1269/18 291 (6.9%)	..	1378/18 420 (7.5%)	..	-0.5 (-1.1 to 0.0)	0.92 (0.85 to 1.00)
55-59 years (N=25 222)	466/12 649 (3.7%)	..	474/12 572 (3.8%)	..	-0.1 (-0.5 to 0.4)	0.98 (0.86 to 1.12)
60-64 years (N=21 936)	417/11 042 (3.8%)	..	367/10 891 (3.4%)	..	0.4 (-0.1 to 0.9)	1.12 (0.97 to 1.30)
65-70 years (N=15 687)	305/7774 (3.9%)	..	296/7911 (3.7%)	..	0.2 (-0.4 to 0.8)	1.05 (0.89 to 1.24)

n/N=women with event/total women with evaluable data. s2D=synthesised two dimensional. OR=odds ratio. *Adjusted for study site (ie, screening unit).

Table 5: Invasive breast cancer detection rates and recall rates by age in the intention-to-treat population

benefits. An absolute increase in the detection rate of invasive breast cancer for early tumour stages in the screening phase of TOSYMA, presumably indicating diagnostic improvements, might be expected to reduce the incidence of advanced breast cancers in screened populations and, thus, potentially exert positive effects on breast cancer mortality.^{25,26} However, increased detection of small size cancers at screening without a reduction in the incidence of invasive interval cancers among screen-negative women in the 2-year interval up to the subsequent screening examination would raise

questions regarding the screening benefit and possible overdiagnosis.²⁷⁻²⁹

A report from the prospective population-based Malmö Breast Tomosynthesis Screening Trial³⁰ found that the interval cancer rate among women screened by digital breast tomosynthesis (1.6 cases per 1000 women) was significantly lower than that among women screened with digital mammography (2.8 cases per 1000 women). An individual participant data meta-analysis found no significant differences in interval breast cancer rates between digital breast tomosynthesis and digital

mammography screening despite increased cancer detection rates.³¹ In February, 2022, the RETomo Working Group reported that the higher sensitivity of digital breast tomosynthesis plus digital mammography compared with digital mammography was not matched by a subsequent reduction in cancer incidence at the next screening examination or in number of interval cancers.³²

In the context of these heterogeneous study results, the endpoints obtained in the prospective follow-up phase of the TOSYMA-2 trial (ie, the 12-month and 24-month invasive interval cancer rates) will indicate whether the increased detection of invasive breast cancers at screening has translated into lower interval cancer incidence.¹⁸ As a result of the unforeseeable pause in recruitment due to the COVID-19 pandemic, recruitment was stopped before the revised sample size of 120 000 participants for TOSYMA-2 was achieved, resulting in a lower total sample size of approximately 100 000 participants (ie, 83% of intended sample size). This reduced sample size affected the statistical power of the primary analysis. The resulting reduction in statistical power, based on the planning assumptions used in the sample size amendment, was about 10% when considering both primary endpoints and 3% when considering only the first primary endpoint.¹⁶ This reduction was considered to be acceptable, considering that a recruitment continuation beyond the 2-year screening interval would have selectively increased the proportion of first-round participants in the study sample, affecting the external comparability with the screening population in routine practice. Follow-up data for the around 100 000 TOSYMA participants are being collected at present via population-based cancer registries; results will be available after completion of the necessary prespecified follow-up time.^{16,18}

This study showed that the detection rate of invasive breast cancers for tumours up to 20 mm in diameter was approximately 70% higher with digital breast tomosynthesis plus s2D mammography than digital mammography, consistent with the results of other studies (appendix 2 pp 9–10).^{11,12} To avoid overestimation of early invasive breast cancer detection, women with neoadjuvant therapy were evaluated as having no pT1 detection. The frequency of neoadjuvant therapy was comparable between the intervention (16·1%) and control group (17·8%). The median diameter of invasive breast cancers was comparable in the digital breast tomosynthesis plus s2D mammography and digital mammography groups (12 mm vs 13 mm), and in both groups, the incidence of pT1c breast cancer size was highest, thus indicating that no shift towards smaller invasive breast cancers was observed despite a higher detection rate. The difference in detection rate of invasive breast cancer observed in the digital breast tomosynthesis plus s2D mammography group versus digital mammography group was largest among women older than 60 years and among those with repeated screening participation. This effect has been observed before in a previous study³³ and it is probably associated

with increasing breast cancer incidence with age and a tumour spectrum that is easier to detect by digital breast tomosynthesis.²² Further exploratory evaluations of the tumour biology, tumour stages, and age at diagnosis in the context of the increased detection rate of invasive breast cancer by digital breast tomosynthesis are necessary to determine possible causes.

Considering the variation in mammography devices across the TOSYMA trial, DCIS detection rates by digital breast tomosynthesis plus s2D mammography were similar to digital mammography. This result suggests that, in a quality assured setting, modern digital breast tomosynthesis technique does not result in underdetection of DCIS compared with digital mammography. However, the evaluation of calcification-related findings remains challenging; magnification imaging with digital mammography remains necessary to further characterise calcifications with higher spatial resolution.¹⁰

We found no marked differences in the recall rates between study groups, or by age groups and screening rounds. This finding is consistent with the results of a meta-analysis that also found no evidence for decreased recall rates with digital breast tomosynthesis in European studies,¹¹ and contrasts with other studies that reported reduced recall rates with digital breast tomosynthesis.^{12,13}

TOSYMA also provides evidence that the probability of diagnosing a malignant breast lesion after recall in screenings is substantially higher with digital breast tomosynthesis plus s2D mammography than digital mammography. Notably, the PPV1 of 17·2% in the intervention group of TOSYMA was higher than reported by US studies and lower than that reported by most EU studies.^{13,29} A higher PPV1 on the basis of a higher invasive cancer detection rate with digital breast tomosynthesis plus s2D mammography than digital mammography presents a desirable screening performance.³⁴ The higher biopsy rate of the digital breast tomosynthesis group than the control group did not lower the high PPV3 of the intervention group.¹³

To our knowledge, TOSYMA is the largest screening randomised controlled trial to date to investigate the effect of digital breast tomosynthesis plus s2D mammography, involving almost 100 000 women. Detailed exploratory evaluations of the screening performance, depending on breast density and pathological tumour characteristics, will be published elsewhere. Successful randomisation procedures led to a balanced structure of the intervention and control groups. This pragmatic trial had a high degree of external validity and practical feasibility, in particular because of the inclusion of multiple centres and the use of technology from multiple vendors. Technicians, readers, and pathologists were trained before study start with the focus of performing and reading digital breast tomosynthesis examinations and about study specific aspects. Regular site visits assessed the quality of procedures and documentation. All readers were experienced and did not differ between the study groups nor between the study and

the routine screening setting. The results in the control group are consequently concordant with results reported for the national mammography screening programme.³⁵

The median of the average glandular dose in the control group was consistent with international reports³⁶ and lower than in the intervention group. The national diagnostic reference level of 2.0 mGy was not surpassed.³⁷

This study had limitations. TOSYMA analysed only one screening round, therefore differences between the study groups might be affected by a first time prevalence screening effect without subsequent digital breast tomosynthesis examinations. Furthermore, there might have been a learning curve regarding acquisition and reading of digital breast tomosynthesis examinations, potentially underestimating the detection rate of invasive breast cancer in the early study phase. TOSYMA reflects mammography performance according to technical standards of 2018.

Only about 50% of screened women consented to participation in the randomised controlled trial, which is lower than in previous studies.^{22,29} Reasons for declining participation in the randomised controlled trial were not specifically investigated. However, the composition of the study sample in terms of age and screening round did not indicate distortion by selective participation since it was closely comparable to that found in regular participants of the German mammography screening programme.³⁵

The British PROSPECTS study with a projected enrolment of 100 000 participants (NCT03733106) is ongoing, which will compare the cost-effectiveness of breast cancer screening with digital breast tomosynthesis plus digital mammography versus standard screening with digital mammography, and the interval cancer rates in both study groups to assess the study effectiveness.

The randomised pragmatic TOSYMA trial contributes robust findings to the existing evidence base, showing that the detection rate for invasive breast cancer is increased by 48% with digital breast tomosynthesis plus synthetic mammography compared with the full-field digital mammography used in routine screening. Determination of interval cancer rates in the follow-up phase of the study will enable the investigation of incremental long-term benefits of digital breast tomosynthesis screening. Thus, the findings from TOSYMA might help to close an important knowledge gap and to develop advanced strategies for an improved systematic early breast cancer detection in population-based settings.

Contributors

WH, SW, H-WH, JG, and LK were guarantors of integrity for the study. All authors conceptualised and designed the study. LK and JG had full access to and verified all data in the study and did the statistical analysis. WH and SW were responsible for one study centre and also acquired data. All authors were involved in data analysis and data interpretation. H-WH, SW, WH, LK, and AS did the literature research. All authors drafted the manuscript for important intellectual content and edited the manuscript. All authors approved the final version of the submitted manuscript.

Declaration of interests

WH has received grants from the German Research Foundation (HE 1646/5-1 and HE 1646/5-2). WH, SW, and AS run the National Reference Center Mammography Münster, a third-party-funded project at the University of Münster and the associated Reference Screening Unit. WH, SW, H-WH, and AS are involved in the educational courses of the Reference Center and receive payments for their lectures. JG has received honoraria from TESARO, QUIRIS Healthcare, Ecker+Ecker, Dr August Wolff, Roche, University Clinics Schleswig-Holstein, and RWTH Aachen University. Before starting the TOSYMA randomised controlled trial, all vendors supplied training cases without any charge; those cases were used for reader training at the Reference Center for Mammography Münster. MK and LK declare no competing interests.

Data sharing

Individual participant data will not be made available with this publication because the follow-up phase (TOSYMA-2) is ongoing. A decision about sharing individual level participant data will be made after finalisation of TOSYMA-2. The study protocol is available online.

Acknowledgments

We thank the German Research Foundation for funding this randomised controlled trial (HE 1646/5-1 and HE 1646/5-2) and all the women who participated. The authors would like to acknowledge the support of the recruiting screening units (Niedersachsen Nordwest, Hannover, Niedersachsen Nord, Niedersachsen Mitte, Niedersachsen Nordost, Duisburg, Krefeld Mönchengladbach Viersen, Wuppertal Solingen, Aachen Düren Heinsberg, Köln rechtsrheinisch, Münster-Süd Coesfeld, Bottrop Gelsenkirchen Recklinghausen, Minden Lübbecke Herford, Bielefeld Gütersloh, Hamm Unna Märkischer Kreis, Höxter Paderborn Soest, Münster-Nord Warendorf, Germany) and the central screening offices of North Rhine-Westphalia (Zentrale Stelle Mammographie Westfalen-Lippe, Münster; Zentrale Stelle Mammographie Nordrhein, Düsseldorf, Germany) and Lower Saxony (Zentrale Stelle Gesundheitsamt, Bremen, Germany). We also thank Regine Hacke and Claudia Verloh (Clinic for Radiology and Reference Center for Mammography Münster, University of Münster and University Hospital Münster, Münster, Germany) for their assistance in the study implementation and training. We also highly appreciate the support of the Data and Safety Monitoring Board represented by Elke Nekolla (Federal Office for Radiation Protection, Munich, Germany), Per Skaane (University of Oslo, Oslo, Norway), and Guido Knapp (Department of Statistics, Technical University of Dortmund, Dortmund, Germany).

References

- 1 Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; **71**: 209–49.
- 2 Lauby-Secretan B, Scoccianti C, Loomis D, et al. Breast-cancer screening—viewpoint of the IARC Working Group. *N Engl J Med* 2015; **372**: 2353–58.
- 3 Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* 2013; **108**: 2205–40.
- 4 Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005; **353**: 1773–83.
- 5 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.
- 6 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.
- 7 Houssami N, Skaane P. Overview of the evidence on digital breast tomosynthesis in breast cancer detection. *Breast* 2013; **22**: 101–08.
- 8 Zackrisson S, Lång 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.

For the study protocol see
[https://bmjopen.bmj.com/
 content/bmjopen/8/5/e020475.
 full.pdf](https://bmjopen.bmj.com/content/bmjopen/8/5/e020475.full.pdf)

- 9 Gilbert FJ, Tucker L, Young KC. Digital breast tomosynthesis (DBT): a review of the evidence for use as a screening tool. *Clin Radiol* 2016; **71**: 141–50.
- 10 Chong A, Weinstein SP, McDonald ES, Conant EF. Digital breast tomosynthesis: concepts and clinical practice. *Radiology* 2019; **292**: 1–14.
- 11 Giampietro RR, Cabral MVG, Lima SAM, Weber SAT, Dos Santos Nunes-Nogueira V. Accuracy and effectiveness of mammography versus mammography and tomosynthesis for population-based breast cancer screening: a systematic review and meta-analysis. *Sci Rep* 2020; **10**: 7991.
- 12 Alabousi M, Wadera A, Kashif Al-Ghita M, et al. Performance of digital breast tomosynthesis, synthetic mammography, and digital mammography in breast cancer screening: a systematic review and meta-analysis. *J Natl Cancer Inst* 2021; **113**: 680–90.
- 13 Heywang-Köbrunner SH, Jänsch A, Hacker A, Weinand S, Vogelmann T. Digital breast tomosynthesis (DBT) plus synthesised two-dimensional mammography (s2D) in breast cancer screening is associated with higher cancer detection and lower recalls compared to digital mammography (DM) alone: results of a systematic review and meta-analysis. *Eur Radiol* 2022; **32**: 2301–12.
- 14 Hofvind S, Holen AS, Aase HS, et al. Two-view digital breast tomosynthesis versus digital mammography in a population-based breast cancer screening programme (To-Be): a randomised, controlled trial. *Lancet Oncol* 2019; **20**: 795–805.
- 15 European Commission. Use of tomosynthesis. 2021. <https://healthcare-quality.jrc.ec.europa.eu/european-breast-cancer-guidelines/screening-tests> (accessed March 25, 2022).
- 16 Weigel S, Gerss J, Hense HW, 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* 2018; **8**: e020475.
- 17 Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis, 4th edn. Luxembourg: Office for Official Publications of the European Communities, 2006.
- 18 Heindel W, Weigel S, Kerschke L, Spieker K, Hense HW, Gerss J. The TOSYMA RCT: increasing the study sample size to assess the effect of DBT on interval cancer rates in mammography screening. Response to: Weigel S, Gerss J, Hense H, 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* 2018; **8**: e020475.
- 19 Wassmer G, Brannath W. Group sequential and confirmatory adaptive designs in clinical trials, 1st edn. Cham: Springer, 2016.
- 20 Brannath W, Posch M, Bauer P. Recursive combination tests. *J Am Stat Assoc* 2002; **97**: 236–44.
- 21 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.
- 22 Houssami N, Lockie D, Clemson M, et al. Pilot trial of digital breast tomosynthesis (3D mammography) for population-based screening in BreastScreen Victoria. *Med J Aust* 2019; **211**: 357–62.
- 23 Caumo F, Zorzi M, Brunelli S, et al. Digital breast tomosynthesis with synthesised two-dimensional images versus full-field digital mammography for population screening: outcomes from the Verona screening program. *Radiology* 2018; **287**: 37–46.
- 24 Freeman K, Geppert J, Stinton C, et al. Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy. *BMJ* 2021; **374**: n1872.
- 25 Tabár L, Yen AM, Wu WY, et al. Insights from the breast cancer screening trials: how screening affects the natural history of breast cancer and implications for evaluating service screening programs. *Breast J* 2015; **21**: 13–20.
- 26 Katalinic A, Eisemann N, Kraywinkel K, Nofz MR, Hübner J. Breast cancer incidence and mortality before and after implementation of the German mammography screening program. *Int J Cancer* 2020; **147**: 709–18.
- 27 Autier P, Boniol M, Koechlin A, Pizot C, Boniol M. Effectiveness of and overdiagnosis from mammography screening in the Netherlands: population based study. *BMJ* 2017; **359**: j5224.
- 28 Houssami N, Zackrisson S, Blazek K, et al. Meta-analysis of prospective studies evaluating breast cancer detection and interval cancer rates for digital breast tomosynthesis versus mammography population screening. *Eur J Cancer* 2021; **148**: 14–23.
- 29 Bernardi D, Gentilini MA, De Nisi M, et al. Effect of implementing digital breast tomosynthesis (DBT) instead of mammography on population screening outcomes including interval cancer rates: results of the Trento DBT pilot evaluation. *Breast* 2020; **50**: 135–40.
- 30 Johnson K, Lång K, Ikeda DM, Åkesson A, Andersson I, Zackrisson S. Interval breast cancer rates and tumor characteristics in the prospective population-based Malmö Breast Tomosynthesis Screening Trial. *Radiology* 2021; **299**: 559–67.
- 31 Houssami N, Hofvind S, Soerensen AL, et al. Interval breast cancer rates for digital breast tomosynthesis versus digital mammography population screening: an individual participant data meta-analysis. *EClinicalMedicine* 2021; **34**: 100804.
- 32 Pattacini P, Nitrosi A, Giorgi Rossi P, et al. A randomized trial comparing breast cancer incidence and interval cancers after tomosynthesis plus mammography versus mammography alone. *Radiology* 2022; published online Feb 1. <https://doi.org/10.1148/radiol.211132>.
- 33 Conant EF, Barlow WE, Herschorn SD, et al. Association of digital breast tomosynthesis vs digital mammography with cancer detection and recall rates by age and breast density. *JAMA Oncol* 2019; **5**: 635–42.
- 34 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.
- 35 Kooperationsgemeinschaft Mammographie. Jahresbericht evaluation. 2019. Deutsches Mammographie-Screening-Programm. https://www.mammoprogramm.de/download/downloads/berichte/neu_KOOPMAMMO_Jahresbericht_Eval_2019_2021 (accessed March 24, 2022).
- 36 Gennaro G, Bernardi D, Houssami N. Radiation dose with digital breast tomosynthesis compared to digital mammography: per-view analysis. *Eur Radiol* 2018; **28**: 573–81.
- 37 Bundesamt für Strahlenschutz. Bekanntmachung der aktualisierten diagnostischen Referenzwerte für diagnostische und interventionelle Röntgenanwendungen. Bundesamt für Strahlenschutz, 2016. <https://www.bfs.de/SharedDocs/Downloads/BfS/DE/fachinfo/ion/drw-roentgen.pdf> (accessed March 24, 2022).