Breast cancer screening with digital breast tomosynthesis: Is independent double reading still required?

Brustkrebs-Screening mit digitaler Brust-Tomosynthese: Ist weiterhin eine unabhängige Doppelbefundung erforderlich?

Authors
Stefanie Weigel[1] ⓘ, Hans-Werner Hense[2], Veronika Weyer-Elberich[3], Joachim Gerss[3] ⓘ, Walter Heindel[1] ⓘ

Affiliations
1 Clinic for Radiology and Reference Center for Mammography Münster, University of Münster and University Hospital Münster, Münster, Germany
2 Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany
3 Institute of Biostatistics and Clinical Research, University of Münster, Münster, Germany

Correspondence
Frau Prof. Stefanie Weigel
Clinic for Radiology and Reference Center for Mammography Münster, University of Münster Faculty of Medicine, Albert-Schweitzer-Campus 1, Building A1, 48149 Münster, Germany
Tel.: +49/2 51/8 34 56 50
Fax: +49/2 51/8 34 56 60
weigels@uni-muenster.de

ABSTRACT

Purpose The European guidelines recommend independent double reading in mammography screening programs. The prospective randomized controlled trial TOSYMA tested the superiority of digital breast tomosynthesis and synthetic mammography (DBT+SM) over digital mammography (DM) for invasive breast cancer detection. This sub-analysis compares the true-positive readings of screening-detected breast cancers resulting from independent double readings in the two trial arms.

Materials and Methods The 1:1 randomized TOSYMA trial was executed in 17 screening units between 07/2018 and 12/2020. This sub-analysis included 49,762 women in the test arm (DBT+SM) and 49,796 women in the control arm (DM). The true-positive reading results (invasive breast cancers and ductal carcinoma in situ) from 83 readers were determined and merged in a double reading result.

Results DBT+SM screening detected 416 women with breast cancer and DM screening detected 306. Double readings of DBT+SM examinations led to a single true-positive together with a single false-negative result in 26.9 % of cancer cases (112/416), and in 22.2 % of cases (68/306) in the DM examinations. The cancer detection rate with discordant reading results was 2.3 per 1,000 women screened with DBT+SM and 1.4 per 1,000 with DM. Discordant reading results occurred most often for invasive breast cancers [DBT+SM 75.9 % (85/112), DM 67.6 % (46/68)], category T1 [DBT+SM 67.9 % (76/112), DM 55.9 % (38/68)], and category 4a [DBT+SM: 67.9 % (73/112); DM: 84.6 % (55/68)].

Conclusion The higher breast cancer detection rate with DBT screening includes a relevant percentage of breast cancers that were only detected by one reader in an independent double reading. As in digital mammography, independent double reading continues to be justified in screening with digital breast tomosynthesis.

Key Points
- The percentages of discordant cancer reading results were 26.9 % and 22.2 % for DBT+SM and DM, respectively.
- The single true-positive detection rate was 2.3 ‰ for DBT+SM and 1.4 ‰ for DM.
- A relevant proportion of screening-detected cancers resulted from a single true-positive reading.

Citation Format
- Weigel S, Hense HW, Weyer-Elberich V et al. Breast cancer screening with digital breast tomosynthesis: Is independent double reading still required?. Fortschr Röntgenstr 2024; DOI 10.1055/a-2216-1109

ZUSAMMENFASSUNG

Ziel Die Europäischen Leitlinien empfehlen eine unabhängige Doppelbefundung in Mammografie-Screening-Programmen. Die prospektiv randomisiert-kontrollierte Studie TOSYMA

prüfte die Überlegenheit der digitalen Brust-Tomosynthese und synthetischen Mammografie (DBT+SM) gegenüber der digitalen Mammografie (DM) für die Detektion invasiver Mammakarzinome. In dieser Subanalyse werden die richtig-positiven Karzinombefundungen verglichen, die bei unabhängigen Doppelbefundungen in den beiden Studienarmen auftraten.

Materialien und Methoden Die 1:1 randomisierte TOSYMA-Studie wurde in 17 Screening-Einheiten von 7/2018 bis 12/2020 durchgeführt. In die Subanalyse wurden im DBT+SM-Arm 49.762 Frauen und im DM-Arm 49.796 Frauen eingeschlossen. Von 83 Befunderinnen und Befundern wurden die richtig-positiven Brustkrebsentdeckungsraten (invasive Mammakarzinome und duktales Carcinoma in situ) ermittelt und in einem Doppelbefundungsergebnis zusammengeführt.

Ergebnisse Mit DBT+SM wurde bei 416 Frauen, mit DM bei 306 Frauen Brustkrebs entdeckt. Dabei trat im DBT+SM-Arm bei 112/416 (26,9 %) Frauen eine richtig-positiv kombiniert mit einer falsch-negativen Befundung auf, im DM-Arm war dies bei 68/306 Frauen der Fall (22,2 %). Dies entsprach mit DBT+SM einer Detektionsrate von 2,3 pro 1.000 gescreenter Frauen mit diskordantem Ergebnis bzw. mit DM 1,4 pro 1.000. Unter diskordanten Diagnosen dominierten invasive Mammakarzinome [DBT+SM 75,9 % (85/112), DM 67,6 % (46/68)], die Kategorie T1 [DBT+SM 67,9 % (76/112), DM 55,9 % (38/68)] und die Befundstufe 4a [DBT+SM: 73/112 (67,6 %); DM: 55/68 (84,6 %)].

Schlussfolgerung Die insgesamt höhere Detektionsrate für Mammakarzinome im DBT-Screening umfasst bei einer unabhängigen Doppelbefundung einen relevanten Anteil von Karzinomen, der nur durch einen Befunder entdeckt wird. Wie im digitalen Mammografie-Screening behält die unabhängige Doppelbefundung ihre Rationale auch beim Screening mit digitaler Brust-Tomosynthese.

Introduction

The aim of breast cancer screening is to reduce the number of cases that progress to an advanced tumor stage through earlier diagnosis, thus enabling therapeutic benefits and reducing breast cancer-specific mortality [1]. Mammography is an evidence-based method of systematic screening which has been proven to lower the rate of breast cancer mortality [2, 3]. In Germany, a mammography screening program (MSP) based on the European guidelines has been introduced nationwide since 2005. The recommendations set out in the European guidelines include independent double reading of the mammograms, performed at different times and different locations, so as to increase the sensitivity by 5–15 % [4]; this is mandatory in the German MSP [5].

Digital breast tomosynthesis (DBT) reduces tissue overlap by moving the X-ray tube in an arc over the breast and reconstructing a pseudo 3D examination from the captured parallel layers; this results in higher breast cancer detection rates than digital mammography (DM), which is the current standard in population-based screening [6]. The randomized controlled TOSYMA study, conducted as part of the ongoing German mammography screening program, showed that the test arm with DBT plus synthetic mammography (SM) had a statistically significant higher rate of detecting invasive breast cancer compared to the control arm with DM [7, 8]. An independent, i. e., blinded, double reading of the mammograms was performed by the same qualified examiners in both study arms. Performing an independent double reading requires an investment of medical resources, especially for screening with DBT which has a higher median time per reading than screening with DM, at 109 seconds compared to 54 seconds [8]; the integration in screening therefore needs to be justified.

Due to the high structural equality of both study arms, the randomized TOSYMA study provides a valid basis for assessing the influence of an independent double reading on breast cancer detection with digital mammography compared to digital breast tomosynthesis [8].

The aim of this TOSYMA subanalysis is to compare the two study arms with regard to the proportion of discordant readings, i. e., cases in which only one of the two independent double readings led to a true-positive finding, and to characterize the breast cancers that are detected in this way.

Materials and Methods

Study Design

Phase 1 of the multicentric TOSYMA study was conducted from July 2018 to December 2020 at 17 screening units in the federal states of North Rhine-Westphalia and Lower Saxony. In this study, 99,689 women were randomized 1:1 to either the study arm (DBT +SM) or the control arm (DM). The study protocol was approved by the responsible ethics committee (2016–132-f-S) and reviewed by two other ethics committees. All of the study participants gave their written consent. The study protocol, the results of the first primary endpoint with secondary endpoints, and two subanalyses have already been published [7–10].

Study Subjects

All women aged 50 to 69 receive a written invitation every two years to participate in the German MSP. In the catchment areas of the TOSYMA study sites, in addition to the regular invitation letter, women also received a personal invitation to take part in the study, together with the study information. Women who had been diagnosed with breast cancer up to 5 years previously or who had undergone a mammography within the past 12 months were not eligible to participate in the MSP. Specific exclusion criteria for the TOSYMA study included having breast implants or having already previously participated in the study [7, 8].

Screening Examination Setup

Participation in the study was offered at 17 screening units in 21 locations (North West Lower Saxony (Wilhelmshaven), Hannover, North Lower Saxony (Stade), Central Lower Saxony (Vechta), North East Lower Saxony (Lüneburg), Duisburg, Krefeld/Mönchengladbach/Viersen, Wuppertal/Solingen (Bergisches Land/Mettmann District), Aachen-Düren-Heinsberg, Cologne Right Rhine (Bergisch Gladbach), Münster-South/Coesfeld, Bottrop, Gelsenkirchen, Recklinghausen, Minden-Lübbecke/Herford, Bielefeld/Gütersloh, Hamm/Unna/Märkischer District (Schwerte), Höxter, Paderborn, Soest (Lippstadt), and Münster North/Warendorf).

Mammography devices from five different manufacturers were used to perform the DBT+SM or DM examinations: Amulet Innovality (Fujifilm Cooperation, Tokyo, Japan; n = 10,075), Class Tomo (IMS Giotto, Sasso Marconi, Italy; n = 7,970), Lorad Selenia 3Dimensions (Hologic, Malborough, US; n = 10,955), Lorad Selenia Dimensions (Hologic, Malborough, US; n = 40,645), MAMMOMAT Inspiration (Siemens Healthineers, Erlangen, Germany; n = 6,759), MAMMOMAT Relevation (Siemens Healthineers, Erlangen, Germany; n = 12,917), Senographe Essential (GE Healthcare, Chicago, US; n = 10,237).

In both study arms, the examination included cranio-caudal and medio-lateral-oblique projections for each breast. In the test arm, stacked layers of ≤ 1 mm thickness were reconstructed to create the images for reading (DBT), in addition to the synthesized two-dimensional mammogram (SM) [7–9].

Independent Double Reading

As in the current MSP, independent double readings were performed by the same certified physicians in both study arms. The screening study involved a total of 83 experienced readers who had at least two years of previous screening experience, performing at least 5,000 screening readings per year. DBT training was provided prior to the start of the TOSYMA study. There were four to eight readers per study site. They received their list of study examinations with both study arms mixed in a random order, and it was not possible to identify the study arm in the screening software before reading.

If there were any abnormalities, the results were discussed at the consensus conference with the responsible physician of the program so as to decide whether further diagnostics were indicated. The protocol for further diagnostics after the study examination did not differ from the established protocol of the MSP; guided by the screening findings, it included, besides a clinical examination, additional mammogram projections where appropriate (e. g., magnification mammography or DBT), ultrasound, MRI, or invasive diagnostic procedures.

All of the screening data were saved in the screening documentation system MaSc (KV-IT GmbH, Dortmund, Germany) [9].

Study Data

The body of data included all of the results from the double readings; this made it possible to determine the number and proportion of concordant results (two true-positives) and discordant results (one true-positive and one false-negative finding) for the breast cancers detected in each study arm (invasive breast carcinomas and ductal carcinoma in situ (DCIS)).

A finding was considered true-positive if a subsequently diagnosed carcinoma was presented at the consensus conference due to at least one mammographic abnormality (category 4a, 4b, and 5), and false-negative if the radiological finding for this carcinoma did not result in a presentation at the consensus conference (category 1, 2) [4, 11].

Based on the DM or SM images, breast density was visually assigned to categories A (fatty), B (fibroglandular), C (heterogeneously dense), or D (extremely dense) [12]. If the two breasts differed in density, the higher category was documented [12]; in the case of discordant density categorization in the independent double reading, the highest density category was used [9]. A and B were grouped together as non-dense parenchyma, and C and D were grouped together as dense parenchyma.

The proportion of breast carcinomas detected based on concordant or discordant findings were stratified according to T categories (Tis, T1, >T1). In the case of multiple manifestations, the more advanced diagnosis was used, determined by histological size (pT), or by imaging (cT) in the case of neoadjuvant therapy. Further stratification included the histological subtype (invasive breast carcinoma of no specific type, invasive lobular breast carcinoma, other subtypes), the mammographic degree of suspicion (category 4a: suspicious abnormality, probably benign; 4b: suspicious abnormality, probably malignant; 5: high suspicion of malignancy), and the mammographic morphology (mass, microcalcification, architectural distortion, asymmetry, and density) according to the consensus conference.

Statistical Analysis

The modified full analysis set included 49,762 women from the test arm (DBT+SM) and 49,796 women from the control arm (DM) who received a screening examination after randomization. The descriptive sub-analysis included all women in whom breast cancer was detected through screening, comprising 416 women from the test arm, and 306 women from the control arm (▶ Fig. 1). Absolute and relative frequencies were calculated for the categorical variables. In addition, we calculated the detection rates for single and double true-positive breast carcinomas per 1,000 women screened.

Results

In the DBT+SM arm, breast cancer was detected in 416 out of 49,762 women (8.4 ‰). Of these, the diagnosis resulted from discordant radiology findings with only one true-positive result in 112 women (26.9 %), corresponding to a detection rate of 2.3 ‰ (112/49,762).

At 6.1 ‰ (306/49,796), the breast cancer detection rate in the DM arm was lower than in the DBT+SM arm, and the proportion of discordant findings was 22.2 %; the resulting detection rate was 1.4 ‰ (68/49,796) (▶ Table 1). Stratification according to non-
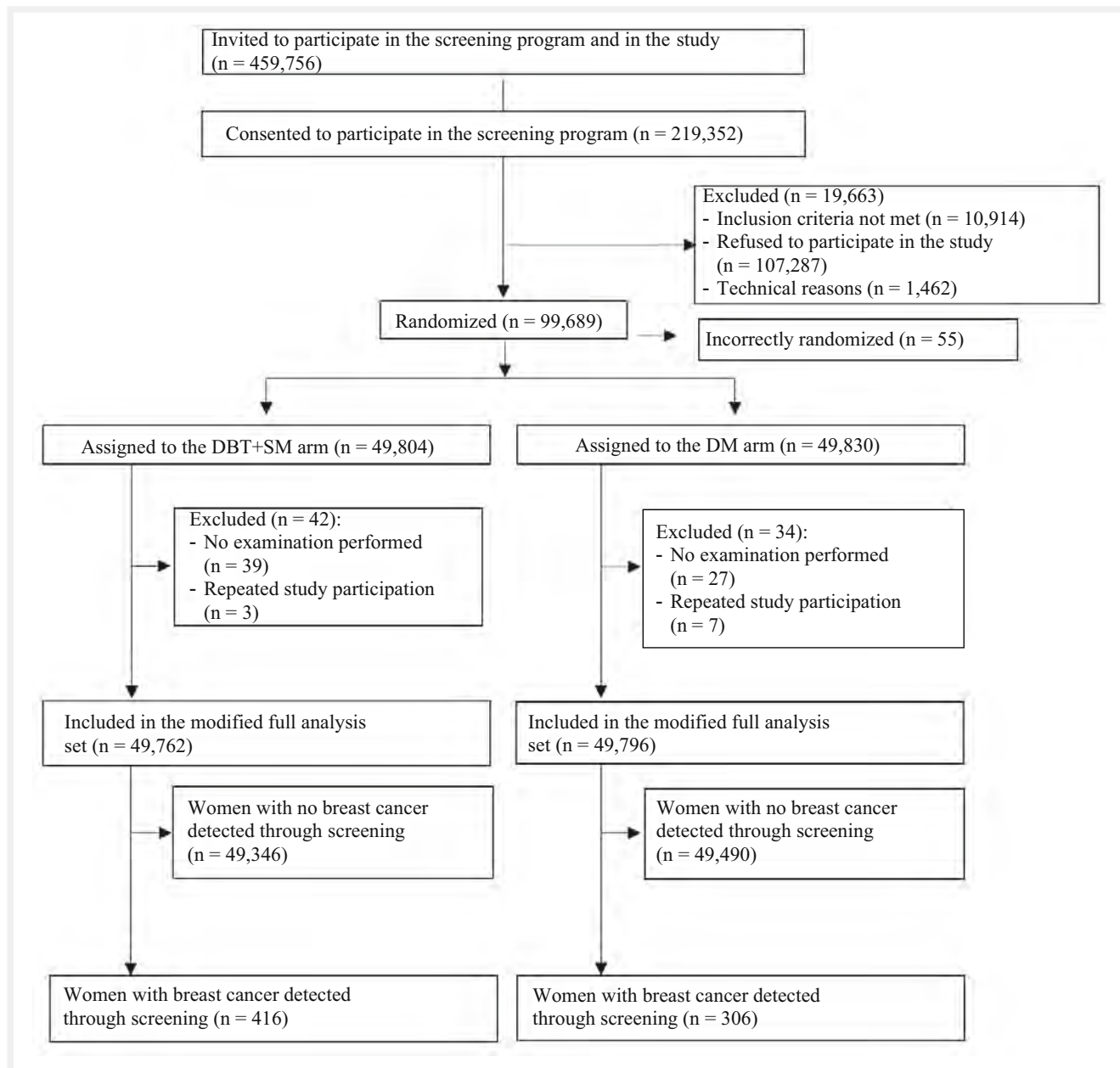
► **Fig. 1** Randomized allocation of the TOSYMA trial participants. DBT+SM = digital breast tomosynthesis plus synthetic mammography; DM = digital mammography.

dense and dense parenchyma showed comparable proportions of single true-positive breast carcinomas in both study arms (DBT+SM: 29.6% and 24.7% respectively; DM: 20.5% and 23.8% respectively).

Of the breast carcinomas in the DBT+SM arm that were only detected through a single true-positive reading, 24.1% (27/112) had DCIS, 67.9% had an invasive breast carcinoma up to 20 mm in size (67/112), and 8% (9/112) had an invasive breast carcinoma larger than 20 mm. The corresponding proportions in the DM arm were 32.4% (22/68), 55.9% (38/68), and 11.8% (8/68) respectively (► **Table 2**).

Among the invasive breast carcinomas detected through a single true-positive (discordant readings) or double true-positive

finding (concordant readings), the non special type was predominant in both study arms. In contrast, the proportion of invasive lobular carcinomas detected through a single true-positive finding was higher than the proportion detected through a double true-positive finding (DBT+SM: 23,5% (20/85) vs. 15,6% (42/269), DM: 26.1% (12/46) vs. 14.9% (29/194) (► **Table 2**).

High suspicion of malignancy (category 5) was rare in both study arms, accounting for less than 10% of carcinomas with discordant readings. Here, suspicious changes of probably benign dignity (category 4a) were predominant, accounting for 67.7% of cases (73/112) in the DBT+SM arm and 84.6% (55/68) in the DM arm (► **Table 3**).

▶ **Table 1** Number (N) and proportion (%) of single and double true-positive detected breast cancers (invasive and DCIS), based on independent double reading in the DBT+SM and DM trial arms.

Results from the independent double reading	DBT+SM n (%)	DBT+SM A+B n (%)	DBT+SM C+D n (%)	DM n (%)	DM A+B n (%)	DM C+D n (%)
Single true-positive	112 (26.9%)	56 (29.6%)	56 (24.7%)	68 (22.2%)	30 (20.5%)	38 (23.8%)
Double true-positive	304 (73.1%)	133 (70.4%)	171 (75.3%)	238 (77.8%)	116 (79.5%)	122 (76.2%)
Total (invasive breast carcinoma plus DCIS)	416 (100%)	189 (100%)	217 (100%)	306 (100%)	146 (100%)	160 (100%)

DBT+SM: Digital Breast-Tomosynthesis + Synthetic Mammography
DM: Digital Mammography
DCIS: ductal Carcinoma in situ
Visually determined breast density categories A+B (BI-RADS 5th ed. [12]): Non-dense parenchyma
Visually determined breast density categories C+D (BI-RADS 5th ed. [12]): Dense parenchyma

▶ **Table 2** Number (n) and proportion (%) of single and double true-positive detected breast cancers (invasive and DCIS), differentiated according to tumor characteristics and histological subtype, based on independent double reading in the DBT+SM and DM trial arms.

Tumor characteristics	DBT+SM Single true-positive n (%)	DM Single true-positive n (%)	DBT+SM Double true-positive n (%)	DM Double true-positive n (%)
pTis	27 + 0 (24.1%)	22 + 0 (32.4%)	35 + 0 (11.5%)	44 + 0 (18.5%)
pT1 + cT1	68 + 8 (67.9%)	34 + 4 (55.9%)	187 + 33 (72.4%)	114 + 30 (60.5%)
> pT1 + >cT1	8 + 1 (8.0%)	8 + 0 (11.8%)	34 + 15 (16.1%)	38 + 12 (21.0%)
No special type	56 (65.9%)	31 (67.4%)	210 (78.1%)	157 (80.9%)
Lobular subtype	20 (23.5%)	12 (26.1%)	42 (15.6%)	29 (14.9%)
Other subtypes	9 (10.6%)	3 (6.5%)	17 (6.3%)	8 (4.1%)

All histologies are based on the final post-operative evaluation.
pTis: Ductal carcinoma in situ
pT1: Histological tumor size up to 20 mm, > pT1: Histological tumor size greater than 20 mm
cT: In the case of histological confirmation of invasive breast cancer with indication for neoadjuvant therapy, tumor size was estimated using imaging.

Among the examinations that only resulted in a single true-positive finding, the proportion of masses and architectural distortions was higher in the DBT+SM arm than in the DM arm, while the proportion of microcalcifications was lower (▶ Table 4, ▶ Fig. 2).

The median reading time for single true-positive readings was 238.0 seconds for DBT+SM and 121.5 seconds for DM, and for single false-negative readings it was 100.0 seconds (DBT+SM) and 40.0 seconds (DM). Breast carcinomas detected through a double true-positive reading had a median reading time of 194.0 seconds in the DBT+SM arm and 99.5 seconds in the DM arm.

Discussion

The large, multicentric, randomized, controlled TOSYMA study conducted in the context of the German MSP shows that the independent double readings performed in screening with both DM (22.2%) and DBT+SM (26.9%) resulted in a relevant proportion of carcinomas being detected based on only a single true-positive reading. Comparable proportions of discordant findings have already been described in routine mammography screening programs. Of the screening-detected cancers, 23.6% were diagnosed in women who were recalled because of screenings with discordant interpretation [13], and 23% of breast carcinomas diagnosed through screening were evaluated negatively by one of the two radiologists [14]. Other reports in the literature also conclude that the double readings can help to increase the sensitivity of mammography [15–17]. Our results are consistent with results that describe a decrease in sensitivity for all density categories associated with a single reading of a mammogram compared to a double reading [18].

In the TOSYMA study, a higher total rate of breast cancer detection by DBT+SM versus DM also results in a higher rate of breast cancer detection with one true-positive and one false-neg-

▶ **Table 3** Number (n) and proportion (%) of single or double true-positive detected breast cancers, differentiated according to the degree of mammographic suspicion, based on independent double reading in the DBT+SM and DM trial arms.

Finding level at consensus conference	DBT+SM Single true-positive n (%)	DM Single true-positive n (%)	DBT+SM Double true-positive n (%)	DM Double true-positive n (%)
4a – Suspicious abnormality, probably benign	73 (67.6 %)	55 (84.6 %)	101 (33.6 %)	101 (43.7 %)
4b – Suspicion abnormality, probably malignant	26 (24.1 %)	6 (9.2 %)	83 (27.6 %)	65 (28.1 %)
5 – High suspicion of malignancy	9 (8.3 %)	4 (6.2 %)	117 (38.9 %)	65 (28.1 %)
Missing data	4	3	3	7
Total (invasive carcinomas plus DCIS)	112 (100 %)	68 (100 %)	304 (100 %)	238 (100 %)

Mammographic suspicion documented during the consensus conference, based on a single or double true-positive independent double reading of screening-detected breast cancers of both trial arms based on the BI-RADS 4th ed. [11]
DCIS: Ductal carcinoma in situ

▶ **Table 4** Number (n) and proportion (%) of single or double true-positive detected breast cancers, differentiated according to mammographic morphology, based on independent double reading in the DBT+SM and DM trial arms.

Morphology at consensus conference	DBT+SM Single true-positive n (%)	DM Single true-positive n (%)	DBT+SM Double true-positive n (%)	DM Double true-positive n (%)
Masses	29 (26.9 %)	13 (20.0 %)	115 (38.3 %)	106 (45.9 %)
Microcalcifications	26 (24.1 %)	24 (36.9 %)	50 (16.7 %)	51 (22.1 %)
Architectural distortion	23 (21.3 %)	7 (10.8 %)	29 (9.6 %)	11 (4.8 %)
Asymmetry	0 (0.0 %)	3 (4.6 %)	0 (0.0 %)	4 (1.7 %)
Density	0 (0.0 %)	5 (7.7 %)	2 (0.7 %)	9 (3.9 %)
Combination	30 (27.8 %)	13 (20.0 %)	105 (34.9 %)	50 (21.6 %)
Missing data	4	3	3	7
Total invasive carcinomas plus DCIS	112 (100 %)	68 (100 %)	304 (100 %)	238 (100 %)

Mammographic morphology documented during the consensus conference, based on a single or double true-positive independent double reading of screening-detected breast cancers of both trial arms. DCIS: Ductal carcinoma in situ

ative finding (DBT+SM arm: 2.3 ‰, DM arm: 1.4 ‰). The cancers detected by a single true-positive reading in the DBT arm include in particular invasive breast carcinomas up to 20 mm in diameter with a low degree of mammographic suspicion (category 4a). The predominant subtype here is breast carcinoma of no special type, while the mammographic morphologies vary. Screening aims to detect T1 carcinomas; however, this can be challenging, even for radiologists experienced in both mammography techniques. The time taken for the reading could have an influence on breast cancer detection, as the median reading times for the single false-negative findings are significantly lower than those of the single true-positive findings, and are slightly lower than the total median reading time for each study arm [8]. In addition, the single true-

positive breast carcinomas may have more subtle abnormalities than those with double true-positive findings, consistent with a longer median reading time for each study arm.

This study does not assess sensitivity at the level of the carcinoma lesion; instead, it is based on the radiological assessment of the screening examination. Since presentation at the consensus conference is not the same as an indication for a mandatory patient recall, but also involves, for example, requesting external mammograms or other examination results, we did not calculate a specificity parameter in relation to the individual readers. Overall, the recall rate did not differ between the two study arms (DBT+SM: 4.9 %, DM: 5.1 %), while the positive predictive value of recall
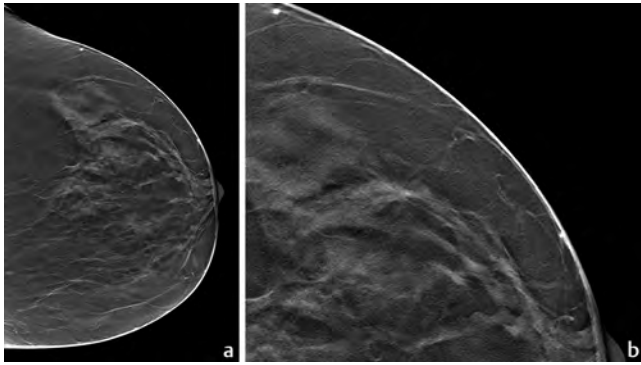
▶ **Fig. 2** Screening-detected breast cancer. **a** Single true-positive reading with depiction of an architectural distortion in digital breast tomosynthesis (cranio-caudal) of the left breast in the lateral quadrants. Histology: Invasive lobular carcinoma, pT1c (11 mm), pN0, cM0, G2. **b** Lesion-depicting magnification.

for further diagnostics (PPV1) was higher in the test arm than in the control arm (DBT+SM: 17.2 %, DM: 12.3 %) [8].

Among the single true-positive breast carcinomas, the largest difference in proportions was observed for microcalcifications, which occurred more frequently in the DM arm (DM: 36.9 %, DBT+SM: 24.1 %). This is consistent with results from 2D mammography screening, which show a significantly higher proportion of microcalcifications among breast carcinomas diagnosed based on discordant readings than those based on concordant readings [13]. Since the DCIS detection rate did not differ between the two study arms [8] and DCIS detection has a strong association with microcalcifications [19], a true-positive finding of microcalcification with DBT+SM appears to be less dependent on the independent double reading than is the case with other mammographic morphologies. In some cases, contrast enhancement of microcalcifications may lead to more obvious visualization in the test arm than in the control arm [6]. Architectural distortion accounted for the second largest difference in proportions, with a higher proportion in the DBT+SM arm (DBT+SM: 21.3 %, DM: 10.8 %). The literature describes the superiority of DBT in detecting spiculations and architectural distortions [6]. This study shows that the independent double reading has a positive influence on the frequency of those diagnosis.

Especially in the context of the longer reading times with DBT compared to DM due to the greater extent of the imaging material with reconstructed layers measuring 1 mm and a median breast compression thickness of 59 mm [8], the prospect of using artificial intelligence (AI) based systems seems promising as an alternative to performing independent double readings. Implementation of AI solutions is favored by standardized mammography setting techniques; in the future it could potentially support human reading, relieving the workload through stratified preselection. The retrospective AI evaluations conducted in the Malmö and Córdoba studies show the potential uses of this technology [20, 21]: Using DBT, the second reading was replaced by AI, resulting in the detection of 95 % of breast carcinomas that were diagnosed

through a double-reading process; this cancer detection rate was 26 % higher than for DM screening with an independent double reading – but at the expense of increasing the recall rate by 53 %. AI alone in the DBT arm had a sensitivity comparable to that of the DM arm with double readings [20]. Compared to DBT examinations with an independent double reading, AI could thus contribute to a relevant reduction in workload without loss of sensitivity [21]. Results from a randomized mammography trial evaluating AI-supported mammography reading compared to the established double reading support the assumption that a comparable breast cancer detection rate can be achieved with a much lower workload using AI [22].

The parameter we used, i. e., breast carcinomas detected through a single true-positive finding, reflects the combined performance of the readers, rather than that of individual readers. This parameter was measured in the same way for both study arms, within a randomized study that had a very low potential for bias due to selective choice of screening participants, readers, or devices. Considering that the DBT arm contained a not insignificant proportion of breast carcinomas that were detected through a single true-positive reading, this argues for the fact that a double reading is still necessary in DBT screening.

TOSYMA is the largest randomized controlled study to date investigating DBT+SM versus DM screening, comprising almost 100,000 study participants. It allows for complementary exploratory evaluations based on successful randomization. The pragmatic approach of this study has a high degree of external validity and also proves its practical feasibility, due in particular to the involvement of a high number of screening units and device technologies. Radiographers, readers, and pathologists were trained prior to the start of the study. All of the physicians were experienced, and the same physicians read examinations of both study arms, with no differences between the study examinations and routine screening.

The TOSYMA study has some limitations. It only investigated one round of screening; this means that the differences between the study arms may have been influenced by an initial prevalence-screening effect with DBT+SM. In addition, there may be a learning curve required for reading tomosynthesis images, meaning that the reading time may decrease with experience. In this sub-analysis, the "true-positive" reading refers to the level of the examination, not the level of the lesion.

Clinical Relevance

As in digital mammography screening, there is a relevant proportion of breast carcinomas that are only detected through one true-positive reading out of the two readings; this applies especially for tumors up to 20 mm in diameter or for lesions that do not give rise to high suspicion of malignancy. The mandatory independent double reading still seems necessary, even with DBT screening. In future, this could be a field for the development of artificial intelligence applications.

Funding

Conflict of Interest

S.W. Payments for lectures from the Reference Center for Mammography Münster; leadership role at the Reference Center for Mammography in Muenster and the associated Reference Screening Unit; board member of the AG Mammadiagnostik of the German Röntgen Society and German Society for Senology; before starting the TOSYMA trial, received training cases from all vendors without charge for reader training at the Reference Center for Mammography Münster. H.W.H. No relevant relationships. V.W.E. No relevant relationships. J.G. Consulting fees from Dr August Wolff, Ecker + Ecker, QUIRIS Healthcare, and TESARO; honoraria for lectures from Roche and TESARO; participation on a data and safety monitoring board or advisory board for the TOMAHAWK trial (University Medical Center Schleswig-Holstein, Campus Lübeck) and Ruxo-BEAT trial (RWTH Aachen University). W.H. Payments for lectures from the Reference Center for Mammography Münster; leadership role at the Reference Center for Mammography in Muenster and the associated Reference Screening Unit; board member/member of the European Society of Radiology, German Röntgen Society, European Society of Breast Imaging, German Society for Senology, and German Röntgen Museum; before starting the TOSYMA trial, received training cases from all vendors without charge for reader training at the Reference Center for Mammography Münster.

Acknowledgement

References

[1] Lauby-Secretan B, Scoccianti C, Loomis D et al. Breast-cancer screening – viewpoint of the IARC Working Group. N Engl J Med 2015; 372: 2353–2358

[2] Marmot MG, Altman DG, Cameron DA et al. The benefits and harms of breast cancer screening: an independent review. Br J Cancer 2013; 108: 2205–2240

[3] Waldmann A, Hübner J, Katalinic A. Trends Over Time in Breast-Cancer-Specific Mortality in Germany. Dtsch Arztebl Int 2021; 118: 538–539

[4] Perry N, Broeders M, de Wolf C et al. European guidelines for quality assurance in breast cancer screening and diagnosis, 4th edn. Luxembourg: Office for Official Publications of the European Communities; 2006

[5] Heindel W, Bock K, Hecht G et al. Systematic and quality-assured early diagnosis of sporadic breast cancer: Update on screening effects and scientific studies. Radiologe 2021; 61: 126–136

[6] Chong A, Weinstein SP, McDonald ES et al. Digital breast tomosynthesis: concepts and clinical practice. Radiology 2019; 292: 1–14

[7] 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; 14 (5): e020475. doi:10.1136/bmjopen-2017-020475

[8] Heindel W, Weigel S, Gerß J et al. 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. Lancet Oncol 2022; 23: 601–611

[9] Weigel S, Heindel W, Hense HW et al. Breast Density and Breast Cancer Screening with Digital Breast Tomosynthesis: A TOSYMA Trial Subanalysis. Radiology 2023; 306 (2): e221006. doi:10.1148/radiol.221006

[10] Weigel S, Heindel W, Decker T et al. Digital Breast Tomosynthesis versus Digital Mammography for Detection of Early-Stage Cancers Stratified by Grade: A TOSYMA Subanalysis. Radiology 2023; 309 (3): e231533. doi:10.1148/radiol.231533

[11] D'Orsi CJ, Mendelson EB, Ikeda DM et al. Breast Imaging Reporting and Data System: ACR BI-RADS – breast imaging atlas, 4th ed. Reston, VA: American College of Radiology; 2003

[12] D'Orsi CJ, Sickles EA, Mendelson EB et al. ACR BI-RADS Atlas, Breast Imaging Reporting and Data System, 5th ed. Reston, VA: American College of Radiology; 2013

[13] Hofvind S, Geller BM, Rosenberg RD et al. Screening-detected breast cancers: discordant independent double reading in a population-based screening program. Radiology 2009; 253: 652–660

[14] Martiniussen MA, Sagstad S, Larsen M et al. Screen-detected and interval breast cancer after concordant and discordant interpretations in a population based screening program using independent double reading. Eur Radiol 2022; 32: 5974–5985

[15] Thurfjell EL, Lernevall KA, Taube AA. Benefit of independent double reading in a population-based mammography screening program. Radiology 1994; 191: 241–244

[16] Ciatto S, Del Turco MR, Morrone D et al. Independent double reading of screening mammograms. J Med Screen 1995; 2: 99–101

[17] Skaane P, Bandos AI, Gullien R et al. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. Eur Radiol 2013; 23: 2061–2071

[18] Euler-Chelpin MV, Lillholm M, Napolitano G et al. Screening mammography: benefit of double reading by breast density. Breast Cancer Res Treat 2018; 171: 767–776

[19] Weigel S, Decker T, Korsching E et al. Calcifications in digital mammographic screening: improvement of early detection of invasive breast cancers? Radiology 2010; 255: 738–745

[20] Dahlblom V, Dustler M, Tingberg A et al. Breast cancer screening with digital breast tomosynthesis: comparison of different reading strategies implementing artificial intelligence. Eur Radiol 2023; 33: 3754–3765

[21] Raya-Povedano JL, Romero-Martín S, Elías-Cabot E et al. AI-based Strategies to Reduce Workload in Breast Cancer Screening with Mammography and Tomosynthesis: A Retrospective Evaluation. Radiology 2021; 300: 57–65

[22] Lång K, Josefsson V, Larsson AM et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. Lancet Oncol 2023; 24: 936–944