

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

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

Autorinnen/Autoren

Stefanie Weigel¹, Hans-Werner Hense², Veronika Weyer-Elberich³, Joachim Gerss³, Walter Heindel¹

Institute

- 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

Keywords

breast cancer, digital breast tomosynthesis, independent double reading, screening

eingereicht 25.8.2023

akzeptiert 15.11.2023

Artikel online veröffentlicht 31.1.2024

Bibliografie

Fortschr Röntgenstr 2024; 196: 834–842

DOI 10.1055/a-2216-1109

ISSN 1438-9029

© 2024, Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Korrespondenzadresse

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

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-positive 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.

Kernaussagen

- Diskordante Karzinom-Befundungen lagen mit DBT+SM bzw. DM zu 26,9% bzw. 22,2% vor.
- Die einfach-positive Detektionsrate betrug mit DBT+SM 2,3‰ und mit DM 1,4‰.
- Ein relevanter Anteil Screening-detektierter Mammakarzinome wurde durch eine einfach richtig-positive Befundung diagnostiziert.

Zitierweise

- 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; 196: 834–842

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 exami-

nations. 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: 67.6% (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.

Einleitung

Brustkrebs-Früherkennung verfolgt das Ziel, fortgeschrittene Tumorstadien durch eine Diagnosevorverlagerung zu reduzieren, somit Therapievorteile zu ermöglichen und die brustkrebspezifische Mortalität zu senken [1]. Die Mammografie ist eine evidenzbasierte Methode zur systematischen Früherkennung mit nachgewiesenem, die Sterblichkeit an Brustkrebs senkenden Effekt [2, 3]. In Deutschland wurde seit 2005 flächendeckend ein Mammografie-Screening-Programm (MSP) basierend auf den Europäischen Leitlinien implementiert. Diese empfehlen unter anderem eine räumlich und zeitlich getrennte unabhängige Doppelbefundung der Mammogramme zur Steigerung der Sensitivität um 5–15% [4], die im deutschen MSP obligat zum Einsatz kommt [5].

Die digitale Brust-Tomosynthese (DBT) mindert durch einen Schwenk der Röntgenröhre sowie Rekonstruktion von detektorparallelen Schichten Gewebsüberlagerungen; daraus resultieren höhere Brustkrebsdetektionsraten als bei der digitalen Mammographie (DM), dem derzeitigen Standard in der populationsbezogenen Früherkennung [6]. Die randomisierte, kontrollierte TOSYMA-Studie, die in das laufende deutsche Mammografie-Screening-Programm eingebettet ist, zeigte eine statistisch signifikante Überlegenheit in der Detektion invasiver Mammakarzinome bei einem Vergleich des Testarms mit DBT plus synthetischer Mammografie (SM) gegenüber dem Kontrollarm mit DM [7, 8]. In beiden Studienarmen wurde eine unabhängige, d. h. gegeneinander verblindete, Doppelbefundung der mammografischen Untersuchungen von denselben, qualifizierten Befundern durchgeführt. Die unabhängige Doppelbefundung bedeutet insbesondere im Screening mit DBT ein ärztliches Ressourceninvestment, da die mediane Befundungszeit mit 109 Sekunden höher liegt als mit 54 Sekunden unter

Einsatz der DM [8]: Ihr Einsatz muss deshalb sachlich begründet sein.

Die randomisierte TOSYMA-Studie bietet durch die hohe Strukturgleichheit beider Studienarme eine valide Grundlage, den Einfluss einer unabhängigen Doppelbefundung auf die Mammakarzinomdetektion mit digitaler Mammografie bzw. mit digitaler Brust-Tomosynthese zu vergleichen [8].

Ziel der vorliegenden TOSYMA-Subanalyse ist ein Vergleich der beiden Studienarme hinsichtlich des bei der unabhängigen Doppelbefundung aufgetretenen Anteils diskordanter, d. h. nur einfach richtig-positiver, Befunde und einer Charakterisierung der so entdeckten Mammakarzinome.

Materialien und Methoden

Studiendesign

Die Phase 1 der multizentrischen TOSYMA-Studie wurde von Juli 2018 bis Dezember 2020 in 17 Screening-Einheiten in den Bundesländern Nordrhein-Westfalen und Niedersachsen durchgeführt. Dabei wurden 99.689 Frauen mit einer 1:1-Randomisierung dem Testarm (DBT+SM) oder dem Kontrollarm (DM) zugeordnet. Das Studienprotokoll wurde von der zuständigen Ethikkommission genehmigt (2016-132-f-S) und von zwei weiteren Ethikkommissionen geprüft. Alle Studienteilnehmerinnen gaben ihr schriftliches Einverständnis. Das Studienprotokoll, die Ergebnisse des ersten primären Endpunkts mit sekundären Endpunkten sowie zwei Subanalysen wurden bereits veröffentlicht [7–10].

Studienteilnehmerinnen

Alle Frauen im Alter von 50 bis 69 Jahren werden alle zwei Jahre schriftlich zu einer Teilnahme am deutschen MSP eingeladen. In den Einzugsgebieten der TOSYMA-Studienzentren erhielten Frauen zusammen mit dem regulären Einladungsschreiben zusätzlich eine persönliche Studieneinladung mit Informationsmaterial. Nicht anspruchsberechtigt zur MSP-Teilnahme waren Frauen mit einer Brustkrebsdiagnose bis zu 5 Jahre zuvor oder mit einer Mammografie innerhalb der letzten 12 Monate. Brustimplantate oder eine wiederholte TOSYMA-Teilnahme waren spezifische Ausschlusskriterien für die TOSYMA-Studie [7, 8].

Erstellung der Screening-Untersuchung

In 17 Screening-Einheiten wurde an 21 Standorten (Niedersachsen Nordwest (Wilhelmshaven), Hannover, Niedersachsen Nord (Stade), Niedersachsen Mitte (Vechta), Niedersachsen Nordost (Lüneburg), Duisburg, Krefeld/Mönchengladbach/Viersen, Wuppertal/Solingen (Bergisches Land/Kreis Mettmann), Aachen-Düren-Heinsberg, Köln rechtsrheinisch (Bergisch Gladbach), Münster-Süd/Coesfeld, Bottrop, Gelsenkirchen, Recklinghausen, Minden-Lübbecke/Herford, Bielefeld/Gütersloh, Hamm/Unna/Märkischer Kreis (Schwerte), Höxter, Paderborn, Soest (Lippstadt), Münster-Nord/Warendorf) eine Studienteilnahme angeboten.

Fünf verschiedene Hersteller von Mammografie-Geräten wurden zur Erstellung der DBT+SM oder DM-Untersuchung eingesetzt: 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 Revelation (Siemens Healthineers, Erlangen, Germany; n = 12.917), Senographe Essential (GE Healthcare, Chicago, US; n = 10.237).

Die Untersuchungen umfassten in beiden Studienarmen die cranio-caudale und die medio-latero-oblique Projektion pro Brust. Im Testarm wurden neben den synthetisierten, zweidimensionalen Mammografien (SM) gestapelte Schichten von ≤ 1 mm Dicke zur Befundansicht rekonstruiert (DBT) [8, 9].

Unabhängige Doppelbefundung

Wie im laufenden MSP wurde in beiden Studienarmen eine unabhängige Doppelbefundung durch dieselben zertifizierten Ärztinnen und Ärzte durchgeführt. Insgesamt umfasste die Screening-Studie 83 erfahrene Befunderinnen und Befunder, die zuvor Screening-Erfahrung von mindestens zwei Jahren hatten mit mindestens 5.000 Screening-Befundungen pro Jahr. Vor Beginn der TOSYMA-Studie wurden DBT-Schulungen durchgeführt. Pro Standort gab es vier bis acht Befunderinnen bzw. Befunder. Sie erhielten ihre Liste mit Studienuntersuchungen in einer gemischten Reihenfolge beider Studienarme, ohne dass der Studienarm vor der Fallwahl in der Screening-Software zu identifizieren war.

Bei Auffälligkeiten wurden die Ergebnisse mit der/dem Programmverantwortlichen Ärztin bzw. Arzt in der Konsensuskonferenz besprochen, um zu entscheiden, ob eine weiterführende Abklärungsdiagnostik indiziert war. Die Abklärungsdiagnostik nach

Studienteilnahme unterschied sich nicht vom etablierten Ablauf des MSP und umfasste befund-orientiert neben einer klinischen Untersuchung ggf. weitere mammografische Projektionen (z. B. Vergrößerungsmammografien oder DBT), die Sonografie, MR-Untersuchungen oder invasive Abklärungsverfahren.

Alle Screening-Daten wurden im Screening-Dokumentationssystem MaSc (KV-IT GmbH, Dortmund, Deutschland) gespeichert [7–9].

Studiendaten

Die Datenbasis umfasste alle Ergebnisse der Doppelbefundungen, sodass Anzahl und Anteil der konkordanten (zweifach richtig-positiv) und diskordanten (einmal richtig-positiv und einmal falsch-negativ) Befundungen der im jeweiligen Studienarm detektierten Mammakarzinome (invasive Mammakarzinome und duktales Carcinoma in situ (DCIS)) ermittelt werden konnten.

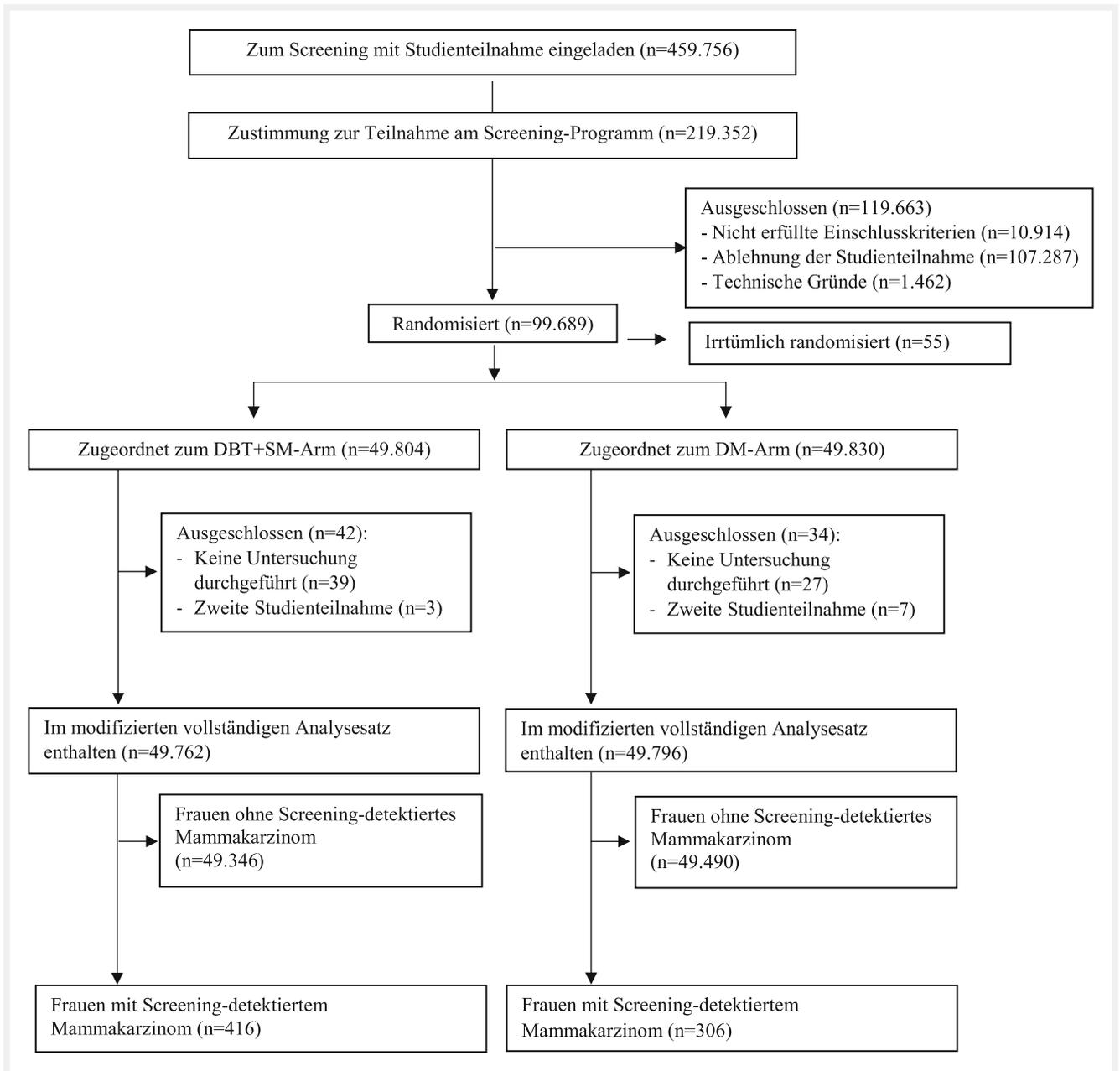
Eine Befundung wurde als richtig-positiv gewertet, wenn es bei einem nachfolgend diagnostizierten Karzinom zu einer Vorstellung in der Konsensuskonferenz aufgrund mindestens einer mammografischen Auffälligkeit kam (Befundstufe 4a, 4b, 5), und als falsch-negativ, wenn bei diesem Karzinom der radiologische Befund keine Vorstellung in der Konsensuskonferenz zur Folge hatte (Befundstufe 1, 2) [4, 11].

Die Brustdichte wurde anhand der DM- bzw. SM-Bilder visuell den Kategorien A (lipomatös), B (fibroglandulär), C (heterogen dicht), D (extrem dicht) zugeordnet [12]. Wenn sich beide Mammæ in der Dichte unterschieden, wurde die höhere Kategorie dokumentiert [12], im Fall von diskordanter Dichtekategorisierung in der unabhängigen Doppelbefundung wurde die höchste Dichtekategorie verwendet [9]. A und B wurden als nicht-dichtes Parenchym, C und D als dichtes Parenchym zusammengefasst.

Die Anteile konkordant und diskordant befundeter Mammakarzinome wurden nach T-Kategorien stratifiziert (Tis, T1, >T1). Im Falle mehrerer Manifestationen wurde die fortgeschrittenere Diagnose verwendet, bestimmt anhand der histologischen Größe (pT) bzw. im Fall einer neoadjuvanten Therapie definiert anhand der Bildgebung (cT). Weitere Stratifizierungen umfassten den histologischen Subtyp (invasives Mammakarzinom nicht spezieller Typ, invasiv lobuläres Mammakarzinom, andere Subtypen), den mammografischen Verdachtsgrad für Malignität (Befundstufe 4a: suspekter Veränderung, eher benigne; 4b: suspekter Veränderung, eher maligne; 5: hochgradig malignitätsverdächtiger Befund) sowie die mammografische Morphologie (Herd, Mikrokalk, Architekturstörung, Asymmetrie, Verdichtung) der Konsensuskonferenz.

Statistische Auswertung

In den modifizierten vollständigen Analysen wurden 49.762 Frauen aus dem Testarm (DBT+SM) und 49.796 Frauen aus dem Kontrollarm (DM), die nach der Randomisierung eine Screening-Untersuchung erhielten, aufgenommen. Die deskriptive Subanalyse umfasste alle Frauen mit einem Screening-detektierten Mammakarzinom, 416 Frauen aus dem Testarm und 306 Frauen aus dem Kontrollarm (► **Abb. 1**). Es wurden für die kategorialen Variablen absolute und relative Häufigkeiten berechnet. Zudem wurden die Detektionsraten von einfach und zweifach



► **Abb. 1** Randomisierte Zuordnung der TOSYMA Studien-Teilnehmerinnen. DBT+SM: Digitale Brust-Tomosynthese plus synthetische Mammografie; DM: Digitale Mammografie.

richtig-positiv befundenen Mammakarzinomen pro 1.000 untersuchte Frauen berechnet.

Ergebnisse

Im DBT+SM-Arm wurden 416 von 49.762 (8,4%) Frauen mit Brustkrebs detektiert. Bei 112 Frauen führte eine diskordante radiologische Befundung mit nur einem richtig-positiven Befund zur Diagnose (26,9%), was einer Detektionsrate von 2,3% (112/49.762) entsprach.

Im DM-Arm lag die Brustkrebsdetektionsrate mit 6,1% (306/49.796) niedriger als im DBT+SM-Arm, der Anteil der diskordan-

ten Befundungen betrug 22,2%, die resultierende Detektionsrate 1,4% (68/49.796) (► **Tab. 1**). Die Stratifizierung nach nicht-dichtem und dichtem Parenchym zeigte vergleichbare Anteile von einfach-positiv detektierten Mammakarzinomen in beiden Studienarmen (DBT+SM: 29,6% bzw. 24,7%; DM: 20,5% bzw. 23,8%).

Bei den nur einfach richtig-positiv detektierten Mammakarzinomen im DBT+SM-Arm lag in 24,1% (27/112) ein DCIS vor, in 67,9% ein invasives Mammakarzinom bis 20 mm (67/112) und in 8% (9/112) ein invasives Mammakarzinom größer als 20 mm. Die entsprechenden Anteile im DM-Arm betragen 32,4% (22/68), 55,9% (38/68) bzw. 11,8% (8/68) (► **Tab. 2**).

► **Tab. 1** Anzahl (N) und Anteil (%) der einfach bzw. zweifach richtig-positiv detektierten Mammakarzinome (invasiv und DCIS), basierend auf einer unabhängigen Doppelbefundung in den Studienarmen DBT+SM und DM.

Ergebnis der unabhängigen Doppelbefundung	DBT+SM n (%)	DBT+SM A+B n (%)	DBT+SM C+D n (%)	DM n (%)	DM A+B n (%)	DM C+D n (%)
einfach richtig-positiv	112 (26,9%)	56 (29,6%)	56 (24,7%)	68 (22,2%)	30 (20,5%)	38 (23,8%)
zweifach richtig-positiv	304 (73,1%)	133 (70,4%)	171 (75,3%)	238 (77,8%)	116 (79,5%)	122 (76,2%)
Summe (invasive Mammakarzinome plus DCIS)	416 (100%)	189 (100%)	217 (100%)	306 (100%)	146 (100%)	160 (100%)

DBT+SM: Digitale Brust-Tomosynthese + Synthetische Mammografie

DM: Digitale Mammografie

DCIS: duktales Carcinoma in situ

Visuell bestimmte Brustdichte-Kategorien A+B (BI-RADS 5th ed. [12]): nicht-dichtes Parenchym

Visuell bestimmte Brustdichte-Kategorien C+D (BI-RADS 5th ed. [12]): dichtes Parenchym

► **Tab. 2** Anzahl (n) und Anteil (%) der einfach bzw. zweifach richtig-positiv detektierten Mammakarzinome (invasiv und DCIS), differenziert nach Tumorcharakteristika und histologischem Subtyp, nach unabhängiger Doppelbefundung in den Studienarmen DBT+SM und DM.

Tumorcharakteristik	DBT+SM einfach richtig-positiv n (%)	DM einfach richtig-positiv n (%)	DBT+SM zweifach richtig-positiv n (%)	DM zweifach richtig-positiv 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%)
Nicht spezieller Typ	56 (65,9%)	31 (67,4%)	210 (78,1%)	157 (80,9%)
Lobulärer Subtyp	20 (23,5%)	12 (26,1%)	42 (15,6%)	29 (14,9%)
Andere Subtypen	9 (10,6%)	3 (6,5%)	17 (6,3%)	8 (4,1%)

Alle Histologien basieren auf der abschließenden post-operativen Bewertung.

pTis: Duktales Carcinoma in situ

pT1: Histologische Tumorgröße bis 20 mm, >pT1: Histologische Tumorgröße größer 20 mm

cT: Im Fall einer histologischen Sicherung eines invasiven Mammakarzinoms mit Indikation zu einer neoadjuvanten Therapie wurde die Tumorgröße anhand der Bildgebung eingeschätzt.

Unter den einfach (diskordant) und zweifach richtig-positiv (konkordant) detektierten invasiven Mammakarzinomen dominierte der nicht spezielle Typ in beiden Studienarmen. Im Vergleich zu den zweifach richtig-positiv detektierten Mammakarzinomen lag dagegen der Anteil des invasiv lobulären Subtyps unter den einfach richtig-positiven Befundungen höher (DBT+SM: 15,6% (42/269) vs. 23,5% (20/85), DM: 14,9% (29/194) vs. 26,1% (12/46) (► **Tab. 2**).

Hochgradig malignitätsverdächtige Befunde (Befundstufe 5) waren in beiden Studienarmen mit unter 10% unter den Karzinomen mit diskordanten Befunden selten. Hier überwogen suspektere Veränderungen mit eher benigner Dignität (Befundstufe 4a) mit Anteilen von 67,7% (73/112) im DBT+SM-Arm und von 84,6% (55/68) im DM-Arm (► **Tab. 3**).

Unter den nur einfach richtig-positiv befundenen Untersuchungen war der Anteil von Herden und Architekturstörungen im DBT

+SM-Arm höher als im DM-Arm, während der Mikrokalkanteil geringer war (► **Tab. 4**, ► **Abb. 2**).

Die mediane Befundungszeit bei einfach richtig-positiver Befundung lag bei 238,0 Sekunden mit DBT+SM und bei 121,5 Sekunden mit DM, bei einfach falsch-negativen Befundungen betrug sie 100,0 (DBT+SM) bzw. 40,0 Sekunden (DM). Zweifach richtig-positiv Screening-detektierte Mammakarzinome wiesen eine mediane Befundungszeit von 194,0 Sekunden im DBT+SM-Arm und von 99,5 Sekunden im DM-Arm auf.

Diskussion

Die große, multizentrische, randomisiert und kontrolliert im Rahmen des deutschen MSP umgesetzte TOSYMA-Studie zeigt, dass durch die unabhängige Doppelbefundung sowohl im Screening mit DM (22,2%) als auch im Screening mit DBT+SM (26,9%) ein

► **Tab. 3** Anzahl (n) und Anteil (%) der einfach bzw. zweifach richtig-positiv detektierten Mammakarzinome, differenziert nach dem mammografischen Verdachtsgrad, nach unabhängiger Doppelbefundung in den Studienarmen DBT+SM und DM.

Befundstufe Konsensuskonferenz	DBT+SM einfach richtig-positiv n (%)	DM einfach richtig-positiv n (%)	DBT+SM zweifach richtig-positiv n (%)	DM zweifach richtig-positiv n (%)
4a – Suspekte Veränderung eher benigne	73 (67,6 %)	55 (84,6 %)	101 (33,6 %)	101 (43,7 %)
4b – Suspekte Veränderung eher maligne	26 (24,1 %)	6 (9,2 %)	83 (27,6 %)	65 (28,1 %)
5 – Hochgradig malignitäts-verdächtiger Befund	9 (8,3 %)	4 (6,2 %)	117 (38,9 %)	65 (28,1 %)
Fehlende Angabe	4	3	3	7
Summe (invasive Karzinome plus DCIS)	112 (100 %)	68 (100 %)	304 (100 %)	238 (100 %)

Mammografischer Verdachtsgrad dokumentiert im Rahmen der Konsensuskonferenz, basierend auf einer einfach oder zweifach-positiven unabhängigen Doppelbefundung von Screening-detektierten Mammakarzinomen beider Studienarme in Anlehnung an das BI-RADS 4th ed. [11]
DCIS: Duktales Carcinoma in situ

► **Tab. 4** Anzahl (n) und Anteil (%) der einfach bzw. zweifach richtig-positiv detektierten Mammakarzinome, differenziert nach der mammografischen Morphologie, nach unabhängiger Doppelbefundung in den Studienarmen DBT+SM und DM.

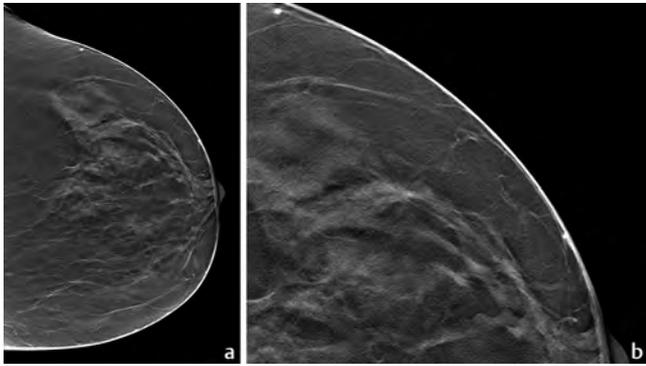
Morphologie Konsensuskonferenz	DBT+SM einfach richtig-positiv n (%)	DM einfach richtig-positiv n (%)	DBT+SM zweifach richtig-positiv n (%)	DM zweifach richtig-positiv n (%)
Herd	29 (26,9 %)	13 (20,0 %)	115 (38,3 %)	106 (45,9 %)
Mikrokalk	26 (24,1 %)	24 (36,9 %)	50 (16,7 %)	51 (22,1 %)
Architekturstörung	23 (21,3 %)	7 (10,8 %)	29 (9,6 %)	11 (4,8 %)
Asymmetrie	0 (0,0 %)	3 (4,6 %)	0 (0,0 %)	4 (1,7 %)
Verdichtung	0 (0,0 %)	5 (7,7 %)	2 (0,7 %)	9 (3,9 %)
Kombination	30 (27,8 %)	13 (20,0 %)	105 (34,9 %)	50 (21,6 %)
Fehlende Angabe	4	3	3	7
Summe invasive Karzinome plus DCIS	112 (100 %)	68 (100 %)	304 (100 %)	238 (100 %)

Mammografische Morphologie dokumentiert im Rahmen der Konsensuskonferenz, basierend auf einer einfach oder zweifach-positiven unabhängigen Doppelbefundung von Screening-detektierten Mammakarzinomen beider Studienarme. DCIS: duktales Carcinoma in situ

relevanter Anteil von Karzinomen mit nur einer einfach richtig-positiven Befundung detektiert wird. Vergleichbare Anteile diskordanter Befundungen wurden bereits für bisher in der Routine laufende Mammografie-Programme beschrieben: So wurden im norwegischen Screening 23,6% der screening-detektierten Mammakarzinome mit einer diskordanten Screeningbefundung diagnostiziert [13] bzw. 23% [14]. Die Doppelbefundung trägt also auch nach anderen Berichten in der Literatur zur Steigerung der Sensitivität der mammografischen Diagnostik bei [15–17]. Unsere Ergebnisse stehen im Einklang mit der Literatur, die eine singuläre Mammografie-Befundung im Vergleich zur Doppelbefundung

verbunden mit einer Abnahme der Sensitivität für alle Dichtekategorien beschreibt [18].

In der TOSYMA-Studie ergibt sich bei einer höheren totalen Brustkrebsdetektionsrate durch DBT+SM versus DM auch eine höhere Brustkrebsdetektionsrate mit einer einfach richtig-positiven und einer falsch-negativen Befundung (DBT+SM-Arm: 2,3%, DM-Arm 1,4%). Diese im DBT-Arm einfach richtig-positiv diagnostizierten Karzinome umfassen insbesondere invasive Mammakarzinome bis 20 mm Durchmesser mit mammografisch geringem Verdachtsgrad (Befundstufe 4a). Es führt das Mammakarzinom vom nicht speziellen Typ während die mammografischen Morphologien variieren. Die Detektion von T1-Karzinomen wird in der Früherkennung



► **Abb. 2** Screening-detektiertes Mammakarzinom. **a** Einfach richtig-positive Screeningbefundung mit dargestellter Architekturdistorsion in der digitalen Brust-Tomosynthese (cranio-caudal) der linken Brust in den lateralen Quadranten. Histologie: invasiv lobuläres Mammakarzinom, pT1c (11 mm), pN0, cM0, G2. **b** Befundrelevante Ausschnittsvergrößerung.

angestrebt; sie scheint aber auch für Erfahrene mit beiden mammografischen Techniken eine Herausforderung darzustellen. Die Befundungsdauer könnte einen Einfluss auf die Mammakarzinomdetektion haben, da die medianen Befundungszeiten der einfach falsch-negativen Befundungen deutlich niedriger liegen als die der einfach richtig-positiven Befundungen bzw. geringfügig niedriger liegen als die gesamte mediane Befundungszeit pro Studienarm [8]. Zusätzlich könnten die einfach richtig-positiven Mammakarzinome subtilere Auffälligkeiten aufweisen als die zweifach richtig-positiven bei einer pro Studienarm jeweils längeren medianen Befundungszeit.

Die Studie bewertet die Sensitivität nicht auf der Ebene der Karzinomläsion, sondern basiert auf der Bewertung der Screening-Untersuchung. Da die Vorlage in der Konsensuskonferenz nicht gleichzusetzen ist mit einer obligaten Rückrufindikation, sondern z. B. auch die Anforderung externer Mammografien oder anderer Untersuchungsergebnisse beinhaltet, wurde auf Befunderbene kein Spezifitätsparameter errechnet. Insgesamt unterschied sich die Rückrufrate zwischen den beiden Studienarmen nicht markant (DBT+SM: 4,9 %, DM: 5,1 %), der positive prädiktive Wert des Rückrufs zur Abklärungsdiagnostik (PPV1) lag im Testarm höher als im Kontrollarm (DBT+SM: 17,2 %, DM: 12,3 %) [8].

Unter den einfach richtig-positiven Mammakarzinomen war die größte Anteilsdifferenz bei Mikroverkalkungen zu verzeichnen, wobei der Anteil im DM-Arm höher war (DM: 36,9 %, DBT+SM: 24,1 %). Dies steht im Einklang mit Ergebnissen zum 2D-Mammografie-Screening, die einen signifikant höheren Anteil an Mikroverkalkungen unter diskordanten als konkordant diagnostizierten Mammakarzinomen beschreiben [13]. Da die DCIS-Detektionsrate zwischen beiden Studienarmen nicht markant unterschiedlich war [8] und die DCIS-Detektion eine starke Assoziation zu Mikroverkalkungen hat [19], erscheint die richtig-positive Befundung von Mikrokalk mittels DBT+SM weniger ausgeprägt von der unabhängigen Doppelbefundung abzuhängen als es bei anderen mammografischen Morphologien der Fall ist. Ggf. führt ein betonter Kontrast der Mikroverkalkungen zu einer markanteren Wahrnehmung im Testarm als im Kontrollarm [6]. Die zweitgrößte Anteilsdifferenz bildete die Architekturdistorsion mit einem hö-

heren Anteil im DBT+SM-Arm (DBT+SM: 21,3 %, DM: 10,8 %). Die Literatur beschreibt vielfältig die Überlegenheit der DBT in der Detektion von Spikulierungen und Architekturstörungen [6]. Diese Arbeit zeigt, dass die unabhängige Doppelbefundung einen positiven Einfluss auf dessen Diagnosehäufigkeit hat.

Insbesondere im Kontext der im Vergleich zur DM längeren Befundungszeit mit DBT aufgrund des größeren Umfangs des Bildmaterials bei 1 mm messenden rekonstruierten Schichten und einer medianen Brustkompressionsdicke von 59 mm [8], erscheint als Alternative zur unabhängigen Doppelbefundung der Einsatz von künstlicher Intelligenz (KI)-basierenden Systemen vielversprechend. Eine Realisierung ist durch standardisierte mammografische Einstelltechniken begünstigt und könnte perspektivisch die menschliche Befundung unterstützen und durch stratifizierte Vorauswahlen entlasten. So zeigen die retrospektiven KI-Bewertungen des Malmö-Trials und der Córdoba-Studie potenzielle Einsatzmöglichkeiten [20, 21]: Anhand einer DBT-Projektionsebene wurde durch KI die Zweitbefundung ersetzt und es wurden 95 % der mit Doppelbefundung diagnostizierten Mammakarzinome entdeckt; diese Karzinomdetektionsrate lag im Vergleich zu einer unabhängigen Doppelbefundung im DM-Screening um 26 % höher – allerdings auf Kosten einer Steigerung der Rückrufrate um 53 %. KI allein im DBT-Arm hatte eine vergleichbare Sensitivität zum DM-Arm mit Doppelbefundung [20]. Im Vergleich zu einer unabhängigen Doppelbefundung von DBT-Untersuchungen könnte KI somit ohne Unterlegenheit in der Sensitivität zu einer relevanten Senkung der Arbeitslast beitragen [21]. Ergebnisse einer randomisierten Studie mit Prüfung einer KI-gestützten Mammografie-Befundung im Vergleich zur etablierten Doppelbefundung stützen die Annahme einer vergleichbar erreichbaren Brustkrebsdetektionsrate bei einer wesentlich geringeren Arbeitsbelastung unter Einsatz von KI [22].

Der verwendete Parameter der einfach richtig-positiv detektierten Mammakarzinome reflektiert nicht die einzelne, sondern die kombinierte Befunderleistung. Dieser Parameter wurde in gleicher Weise für beide Arme innerhalb einer randomisierten Studie erhoben, die nur ein sehr geringes Potenzial für verzerrende Einflüsse durch eine selektive Auswahl von Screening-Teilnehmerinnen, Untersuchern oder Geräten hatte. Der weiterhin nicht geringere Anteil der einfach richtig-positiven Mammakarzinome im DBT-Arm spricht dafür, dass auf eine Doppelbefundung in einem DBT-Screening nicht verzichtet werden kann.

TOSYMA ist die bisher größte randomisiert-kontrollierte Screening-Studie zur Untersuchung von DBT+SM versus DM mit fast 100.000 Studienteilnehmerinnen. Sie ermöglicht ergänzende explorative Auswertungen auf der Basis einer erfolgreichen Randomisierung. Der pragmatische Ansatz hat ein hohes Maß an externer Validität und belegt auch die praktische Umsetzbarkeit, insbesondere aufgrund der Einbeziehung zahlreicher Screening-Einheiten und Gerätetechnologien. Radiologische Fachkräfte, Befunderinnen und Befunder und Pathologinnen und Pathologen wurden vor Studienbeginn geschult. Alle Ärzte waren erfahren und unterschieden sich weder zwischen den Studienarmen noch zwischen der Studie und dem Routine-Screening.

Die Studie hat auch Limitationen. TOSYMA analysierte nur eine Screening-Runde, daher könnten die Unterschiede zwischen den Studienarmen durch einen initialen Prävalenz-Screening-Effekt

mit DBT+SM beeinflusst sein. Darüber hinaus könnte eine Lernkurve in der Befundung von Tomosynthesen vorhanden sein, die sich mit steigender Erfahrung auf die Befundungszeit reduzierend auswirken könnte. Die Subanalyse bezieht sich in der Bewertung ‚richtig-positiv‘ auf die Untersuchungsebene, nicht auf die Läsionsebene.

Klinische Relevanz

Wie im digitalen Mammographie-Screening findet sich ein relevanter Anteil von Mammakarzinomen, der nur durch einen der beiden Befunder richtig-positiv detektiert wird, vor allem bei Tumoren bis 20 mm Durchmesser bzw. unter Läsionen ohne hochgradigen Malignitätsverdacht. Die obligate, unabhängige Doppelbefundung erscheint auch in einem DBT-Screening weiterhin erforderlich. Künftig könnte sich hier ein Feld für den Einsatz künstlicher Intelligenz entwickeln.

Förderung

Deutsche Forschungsgemeinschaft (DFG)
HE 1646/5–1, HE 1646/5–2

Interessenkonflikt

S.W. Honorare für Vorträge des Referenzzentrums Mammographie Münster; Leitungsfunktion im Referenzzentrum Mammographie Münster und der zugehörigen Referenz-Screening-Einheit; Vorstandsmitglied der AG Mammadiagnostik der Deutschen Röntgengesellschaft und der Deutschen Gesellschaft für Senologie; vor Beginn der TOSYMA-Studie wurden kostenlose Schulungsfälle von allen Herstellern für die Befunderschulung im Referenzzentrum Mammographie Münster erhalten. H.W.H. Keine relevanten Interessenkonflikte. V.W.E. Keine relevanten Interessenkonflikte. J.G. Beratungshonorare von Dr. August Wolff, Ecker + Ecker, QUIRIS Healthcare und TESARO; Honorare für Vorträge von Roche und TESARO; Mitarbeit in einem Daten- und Sicherheitsmonitoring-Gremium bzw. Beirat für die TOMAHAWK-Studie (Universitätsklinikum Schleswig-Holstein, Campus Lübeck) und Ruxo-BEAT-Studie (RWTH Aachen). W.H. Honorare für Vorträge des Referenzzentrums Mammographie Münster; Leitungsfunktion im Referenzzentrum Mammographie Münster und der zugehörigen Referenz-Screening-Einheit; Vorstandsmitglied/Mitglied der Europäischen Gesellschaft für Radiologie, der Deutschen Röntgengesellschaft, der Europäischen Gesellschaft für Senologie, der Deutschen Gesellschaft für Senologie und des Deutschen Röntgenmuseums; vor Beginn der TOSYMA-Studie wurden kostenlose Schulungsfälle von allen Herstellern für die Befunderschulung im Referenzzentrum Mammographie Münster erhalten.

Danksagung

Wir bedanken uns bei der Deutschen Forschungsgemeinschaft für die Förderung der randomisierten, kontrollierten, klinischen Studie TOSYMA (DFG HE 1646/5–1 und HE 1646/5–2) sowie bei allen teilnehmenden Frauen.

Ein besonderer Dank gilt den rekrutierenden Screening-Einheiten bzw. Studienzentren, dem interdisziplinären Studienteam der Universität Münster (Klinik für Radiologie, Institut für Epidemiologie und Sozialmedizin, Institut für Biometrie und klinische Forschung, Zentrum für Klinische Studien) sowie dem *Data and Safety Monitoring Board* repräsentiert durch:

Mary Andresen-Khanlian, Rudolf Amendt, Sonja Baier, Christina Barduna, Josiane Bedrossian- Pflingsten, Stefan Bergeest, Wilfried Berning, Markus

Blumenröther, Andrea Böhmert, Michael Bremme, Horst Bürger, Seyhan Cubukcu, Beate Dablé, Caroline Dietz, Darius Ettehad, Wilhelm Freitag, Margret Ganseforth, Reinhard Golz, Axel Gräwingholt, Bettina Grundner, Regine Hacke, Stefan Hanning, Gerold Hecht, H.-G. Heidecker, Burkhard Helmke, Ingeborg Henseleit, Michael Hirning, Reinhard Hovestadt, Axel Iffländer, Smaragdo Kapsimalakou, H.-U. Kasper, Laura Kerschke, Uwe Keske, Peter Keulers, Bernd Kirchner, Guido Knapp, Ruth Knüchel-Clarke, Ulrich Krause, Miriam Krischke, Florian Länger, Thomas Lehnen, Peter Liersch, Roland Lukaschek, Dorothea Meyer, Ulrike Meyer-Johann, Jost Müller, Elke Nekolla, Anja Peters, Annette Pfitzenmaier, Friederike Pochede-Vos, Doris Polenske, Annette Pollerberg, Jost Porrmann, Peter Prodehl, Regine Rathmann, Beate Richter-Sadocco, Annette Rudolf, Lars Rühle, Denise Rütter, Monika Salewski, Barbara Schaefer; Gudrun Schlake, Elke Schmidt-Eversheim, Filip Schmidt-Petersen, Klaus J. Schmitz, Kirsten Schulze, Mechtilde Schulze-Hagen, Henning Schubert, Regina Schwanitz-Gräwingholt, Per Skaane, Alexander Sommer, Karin Spieker, Simona Spital, Karl Struckmann, Markus Stuhmann, Renate Tewaag, Ulrike Theus, Andreas Thiel, Thilo Töllner, Mark Tsesarskiy, Gisela Urban, Claudia Verloh, Bernd Weimar, Berthold Wein, Ingo Wendt, Jürgen Wiesmann, Cornelia Wunsch, Stephanie Zobel.

Literatur

- [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