Risk-adapted breast imaging in population breast cancer screening: A UK National Screening Committee Evidence Summary

External review against programme appraisal criteria for the UK National Screening Committee

Version: 1

NIHR Evidence Synthesis Group: Aberdeen Belfast Evidence Collaboration (ABEC)

Date: February 2025

The UK National Screening Committee secretariat is hosted by the Department of Health and Social Care

Contents

About the UK National Screening Committee (UK NSC)

The UK NSC advises ministers and the NHS in the 4 UK countries about all aspects of population screening and supports implementation of screening programmes.

Conditions are reviewed against evidence review criteria according to the UK NSC's evidence review process.

Read a complete list of UK NSC recommendations.

UK National Screening Committee, Southside, 39 Victoria Street, London, SW1H 0EU

www.gov.uk/uknsc

Blog: https://nationalscreening.blog.gov.uk/

For queries relating to this document, please contact: uknsc@dhsc.gov.uk.

Published March 2025

Plain English summary

In the UK, people aged 50-70 are offered mammograms (x-ray images of the breasts) every three years to help detect breast cancer early. This national screening programme has successfully reduced breast cancer deaths by 20-40%. However, mammograms are less effective for persons with dense breasts, who also have a higher risk of cancer. About half of screened persons have dense breasts, which can lead to missed cancers.

A 2019 review by the National Screening Committee found that there was no evidence that adding ultrasound to mammograms reduced cancers found between screenings, lowered mortality, or reduced NHS costs in participants with dense breasts who had negative screening results.

Recently, other screening methods have been considered. Clinical guidelines in the US and Europe now recommend that people be informed about their breast density and, in some cases, receive additional screening like Magnetic Resonance Imaging (MRI).

We have gathered information from clinical studies published in the medical literature on the performance of advanced imaging methods compared to traditional mammography for people with dense breasts. Our goal was to determine which imaging method -— such as MRI, handheld ultrasound (HHUS), automated breast ultrasound (ABUS), 3D mammography also known as digital breast tomosynthesis (DBT), or contrast-enhanced mammography (CEM) — could be added to standard breast mammography to improve cancer detection in people with dense breasts.

We have identified 36 studies. Key findings include:

- DBT, ABUS, and HHUS detected 1-3 extra cancers per 1,000 exams.
- MRI detected about 19 extra cancers per 1,000 exams - a much higher rate.
- In a single study (BRAID), CEM performed similarly to MRI.
- The definition of recall rates (how often patients are called back for further testing) varied between studies, making direct comparisons difficult, however, the recall rates were generally below the recommended 10% threshold, making them acceptable given the additional cancers detected.

- There were small increases in biopsy rates (proportions of people who undergo a biopsy after an abnormal screening result) for ABUS and MRI compared with standard mammography.
- Two large studies showed that adding MRI reduced cancers found between screenings compared with standard mammography.

Overall, MRI and CEM appear to be superior to DBT, ABUS, and HHUS in identifying cancers missed by standard mammography. There are, however, some considerations regarding MRI studies, such as delays between mammograms and MRIs and the fact that some studies focused only on people with extremely dense breasts. Despite these limitations, current evidence suggests MRI, and possibly CEM, could improve early cancer detection in people with dense breasts who are at higher risk of cancer.

Executive summary

Background

The UK National Screening Committee (UK NSC) plays a vital role in early cancer detection, inviting women aged 50-70 years for mammography every three years. This national programme has led to a significant 20-40% reduction in breast cancer mortality risk. However, underdiagnosis remains a critical concern, particularly for women with dense breast tissue, a factor that not only increases cancer risk but also diminishes the effectiveness of mammograms. Nearly half of all women in screening programmes are classified as having heterogeneously or extremely dense breasts - categorised as BI-RADS (Breast Imaging Reporting and Data System) grades C and D, respectively - placing them at a heightened risk of missed breast cancer diagnoses. Given this substantial gap in detection, the question of whether additional imaging methods could improve outcomes has been a subject of ongoing debate.

In 2019, the UK NSC commissioned an evidence review to evaluate the potential benefits of supplemental ultrasound screening for women with dense breasts following a negative mammogram. This review concluded that existing evidence did not demonstrate that ultrasound could reduce interval cancers and mortality, nor did it prove to be a cost-effective option for the NHS, leading to its rejection as an additional screening modality.

Meanwhile, the landscape of breast imaging is rapidly evolving. In 2022, the European Society of Breast Imaging (EUSOBI) issued guidelines recommending that women be informed of their breast density and offered supplemental magnetic resonance imaging (MRI) screening for those with extremely dense breasts. In the US, the FDA has mandated that, from September 2024, mammogram reports must disclose breast density information and advice should be provided to patients that additional screening could aid cancer detection. In the UK, the recently completed multicentre BRAID (Breast Screening — Risk Adapted Imaging for Density) study investigated whether abbreviated MRI, contrast-enhanced mammography (CEM), and automated breast ultrasound (ABUS) could enhance cancer detection in women with dense breasts.

Focus of the review

The aim of this report is to examine the existing evidence related to the use supplemental imaging modalities compared with standard mammography in individuals with breast density

who underwent screening for breast cancer (UK NSC Criterion 4: There should be a simple, safe, precise and validated screening test). Comprehensive search strategies were developed by an information scientist and the following databases were searched from 2014 onwards: MEDLINE, Embase, Cochrane Database of Systematic Reviews, CENTRAL, Scopus, and Web of Science. Only articles published in full were eligible for inclusion.

Findings of the review

Thirty-six studies were included in this review, of which 18 investigated DBT as a supplementary modality, 6 investigated ABUS, 9 investigated HHUS, 6 investigated MRI, and 3 investigated 'other' modalities, including a single study assessing CEM.

DBT, ABUS and HHUS detected more cancers than mammography alone. The additional cancers detected per 1,000 screenings were on average 1.69 (95% CI 0.81, 2.58), 2.3 (95% CI 1.28, 3.33), and 2.57 (95% CI 0.99, 1.44) for DBT, ABUS, and HHUS, respectively. In contrast, MRI demonstrated a greater mean difference, detecting an additional 18.92 (95% CI 15.41, 22.43) cancers per 1,000 screenings compared to mammography. This estimate was based on five randomised studies where women with a negative mammogram underwent supplementary MRI. Sensitivity analysis, excluding small studies (<500 participants) did not substantially alter the findings, with MRI detecting 17.23 (95% CI 14.38, 20.08) cancers per 1,000 screenings across three large studies. For CEM, only one study investigated its effectiveness. The cancer detection rate (CDR) was comparable to MRI, suggesting its potential as a supplementary screening modality. The differential combined recall rate varied across supplementary modalities but was broadly similar for ABUS and HHUS and did not differ substantially for MRI. Interestingly, despite the considerably higher CDR for MRI, the recall rate did not increase as expected. Both the pooled recall rate estimate and the estimates of the four individual studies included in the MRI meta-analysis remained within the acceptable thresholds (<10%) set by the National Health System Breast Screening Programme. Notably, variability in the definition of 'recall rate' across studies limited direct comparisons, although definitions were broadly consistent across MRI studies. Regarding biopsy rates, there was no difference between DBT plus mammography and mammography-only. Even though there was considerable heterogeneity in the data, we observed small increases in biopsy rates for ABUS and MRI in the supplementary screening group compared to mammography. CEM biopsy rates were comparable to those of MRI.

Data on interval cancer rates (cancers detected between screening rounds) were often not reported or inconsistently reported across included studies. Two studies reported lower interval cancers in the HHUS supplementary group. Two studies reported interval cancer data for MRI. The DENSE study reported that in the incidence screening round, there were 2.5 (95% CI 1.6 to 3.8) interval cancers detected per 1,000 screenings in the MRI invitation group and 5.0 (95% CI 4.3 to 5.5.8) in the mammography-only group. Among women who actually underwent MRI, the interval cancer rate was lower, with 0.8 interval cancers detected per 1,000 screenings. The German Kuhl study reported that no interval cancers were detected in over 2 years of follow-up in patients who had supplemental MRI.

False positives were not widely reported across studies and imaging modalities. For MRI, reports from the DENSE study showed a notable reduction in false positive rates between screening rounds.

Of the 36 included studies, 13 reported tumour characteristics with supplementary imaging generally out-performing mammography alone in detecting invasive cancers. Supplementary MRI detected 4, 32, 14 and 64 invasive cancers across four studies involving women with a negative mammogram. The reported median tumour sizes were 1.0 cm (IQR 0.8-1.5 cm) and 0.7 cm (IQR 0.6-1.0 cm) in 2 studies. For CEM, the UK BRAID study found that 32 of 39 detected cancers were invasive, and the median size of detected cancers was 1.1 cm (IQR 0.7-1.5 cm). CEM was reported to detect three times as many invasive cancers as ABUS and have a performance comparable to MRI.

There was little evidence of a formal attempt to measure the time required to conduct or interpret supplementary imaging. Among the 36 included studies, only nine studies reported some data on this aspect.

Concerning the methodological quality of included studies, while most studies were deemed to have a low risk of bias in terms of patient selection, five studies were classified as high risk due to opportunistic recruitment. Applicability concerns were noted in 14 studies, mainly relating to the participants' age range. Two MRI studies focused exclusively on women with extremely dense breasts, while the scope of this systematic review encompasses both heterogeneous and extreme breast density. In one MRI study, the comparator group included individuals with both negative ultrasounds and negative mammography, raising some applicability concerns. The

main quality concerns stemmed from flow and design, including missing or insufficient details on follow-up, incomplete data reporting, and delays between index tests and comparator tests.

Limitations

In our review, we included both retrospective and randomised studies, which may elevate the risk of selection bias due to confounding factors. To address this, we excluded studies featuring 'selected' populations with characteristics linked to higher cancer risk, ensuring the inclusion of women at average cancer risk, aside from dense breasts. This approach limited the number of studies, along with the review's statistical power, but allowed for better representation of the general breast screening population. Additionally, since the review was focused on 'dense' breasts without distinguishing between heterogeneous and extremely dense types, we could not determine if supplementary modalities performed better for the densest breasts. As we included studies including women with negative mammography as well as studies of women assigned to a mammography-only arm, we reported differential data rather than absolute values. To capture the latest evidence in this rapidly evolving field, we conducted searches up to November 2024, including unpublished reports, which allowed us to incorporate a confidential version of the BRAID study.

Considerations and uncertainties

In terms of cancer detection, supplementary MRI and CEM demonstrated superiority over DBT, ABUS and HHUS in detecting cancers that were missed by mammography alone. Therefore, the following considerations focus mainly on these two modalities.

This systematic review did not specifically differentiate between heterogeneously and extremely dense breasts. The meta-analysis of the supplementary MRI showed a high CDR and relatively low recall rates; however, it is important to note that the DENSE study exclusively examined women with extremely dense breasts, raising concerns about its broader applicability. In contrast, the BRAID study included women with both types of density, showing similar detection rates to the prevalence screening group of the DENSE study. The EUSOBI recommends MRI screening every 2-4 years for women with extremely dense breasts, citing robust evidence from the DENSE study. Given that breast density manifests as a continuum, the BRAID study suggests that the cancer-detecting benefits of supplementary MRI extend beyond those with the highest breast density.

The BRAID study used an abbreviated MRI protocol, whereas the other three studies employed a full MRI protocol. The feasibility of widespread MRI use as a supplementary screening tool may be limited by cost, availability, and the expertise required for conventional MRI. An abbreviated protocol could mitigate these limitations by increasing throughput and reducing costs. A retrospective study of 356 women supports the accuracy of abbreviated MRI, showing comparable performance to full-protocol MRI.

In studies on supplementary MRI, the primary risk of bias relates to flow and timing. Our review identifies a potential risk of bias in the BRAID, Bakker, and Veenhuizen studies due to the time between index tests (negative mammography and supplementary MRI), with median intervals ranging from 8 to 20 weeks. The German Kaiser study, in its preliminary report, excluded participants without an index test within two months to rule out interval carcinoma being interpreted in favour of MRI. Although logistical issues can affect MRI access and availability, similar intervals were observed in the BRAID study for other modalities, such as ABUS and CEM. An applicability concern was also raised regarding the German study by Kuhl et al. as the comparator arm included participants who had both a negative mammogram and, in 65% of cases, a negative ultrasound. This could have led to an underestimation of MRI's performance as some cancers missed by mammography may have been detected by ultrasound. However, the exact numbers of such cases were not given.

Overall conclusions

Standard screening mammography misses cancers in women with dense breasts. Our analysis of randomised evidence demonstrates that supplementary MRI is superior to ABUS, HHUS and DBT in detecting cancers missed by mammography. The UK-based BRAID study provides additional evidence to support the findings of the Netherlands-based DENSE study. It is important to note that detection rates, recall and biopsy rates may decline in the incident screening rounds, as observed in the second round of the DENSE study. The interval between negative mammography and supplementary MRI could introduce a potential risk of bias. CEM was shown to have a similar performance to MRI, though this finding is based only on the BRAID study. Overall, these findings highlight supplementary screening modalities that have the potential to increase the number of cancers detected in women with dense breasts, contributing to a more refined approach to population-level screening.

Introduction and approach

Background

Breast cancer is the most common type of cancer among participants in the UK, accounting for 15% of all new cancer cases. Each year around 55,900 are diagnosed, more than 150 daily.[1] Whilst breast cancer can occur at any age, it most commonly affects postmenopausal people over 50 years of age.

The UK national screening programme invites individuals aged 50-70 years for mammograms every three years, significantly reducing breast cancer mortality risk by 20-40%.[2-4] However, underdiagnosis remains a concern, particularly for individuals with dense breast tissue, which not only increases cancer risk but also makes mammograms less effective.[5]

Breast density is determined by the proportion of fibroglandular tissue visible on a mammogram. It is classified into four categories (A,B,C,and D) according to the American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) system[6] with categories C and D indicating heterogenous to extremely dense breasts. Nearly half of people in screening programmes fall into these two groups,[7] making them more prone to missed diagnoses. In clinical practice, breast density has traditionally been assessed through a subjective manual evaluation, where radiologists visually inspect mammograms to categorise breast density.

Ethnicity also plays a role. For example, women in Japan have denser breasts. Unlike Western countries, where screening has been associated with decreasing mortality, Japan has not observed a similar reduction in mortality rates,[8] possibly due to underdiagnosis with standard mammography.[9]

Rationale for the review

Criterion 4 — There should be a simple, safe, precise and validated screening test

Question — *What is the effect of an additional imaging modality to supplement standard mammography compared with standard mammography alone for identifying breast cancer in people with dense breasts?*

In 2019, the UK National Screening Committee (NSC) commissioned an evidence review to evaluate the benefits of additional ultrasound screening for people with dense breasts after a negative mammogram.[10]

The review found that increased breast density was associated with reduced mammography sensitivity and a higher risk of interval cancers. However, it also highlighted challenges in validating the breast density measurement methods and reported high false positive rates with ultrasound. Notably, no evidence supported ultrasound in reducing interval cancers and mortality, nor demonstrated its cost-effectiveness, leading to its rejection as a supplemental screening modality.[10]

The field of breast imaging is rapidly advancing. In 2022, the European Society of Breast Imaging (EUSOBI) guidelines recommended informing women about their breast density and offering supplemental Magnetic Resonance Imaging (MRI) screening for those with extremely dense breasts.[11] In the US, the FDA has mandated that, from September 2024, mammogram reports must disclose breast density and inform patients that additional screening may aid cancer detection.[12]

In the UK, the multicentre Breast Screening - Risk Adapted Imaging for Density (BRAID) study, which investigates whether abbreviated MRI, Contrast Enhanced Mammography (CEM), and automated breast ultrasound (ABUS) improve cancer detection in women with dense breasts, has recently been completed. Its findings will contribute to the growing evidence base.[13]

To inform future policy, the UK NSC has commissioned a suite of evidence reviews to assess supplemental imaging modalities to detect breast cancer in individuals with dense breasts and the role of these modalities within the national breast screening programme.

Objectives

Specifically, these reviews address the following objectives:

- Objective 1: To determine the agreement (concordance) between automated and manual measurement of mammographic breast density
- Objective 2: To determine the effect of an additional imaging modality to supplement standard mammography compared with standard mammography alone for identifying breast cancer in individuals with dense breasts.
- Objective 3: To review evidence on existing economic models assessing the costs and consequences of enhanced mammographic screening for individuals with breast density.

This document addresses Objective 2 and complies with NSC criterion 4 for a population screening programme, which requires a simple, safe precise and validated screening test.

Methods

General

This systematic review was commissioned by the UK NSC and was conducted in accordance with the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions[14] and in adherence with the Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines.[15] The methods were pre-specified in a protocol and registered with the PROSPERO International Prospective Register of Systematic Reviews, available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42024550250.

Patient and public involvement (PPI)

Two PPI partners were part of the study Advisory Group, which also included academic and clinical experts. One PPI partner has lived experience of undergoing mammography for routine breast screening and the other has lived experience of breast cancer. PPI partners participated in regular Advisory Group meetings, where they contributed to discussions and made recommendations at each stage of the project.

Language and inclusivity statement

Most people who use the UK's breast screening programme identify as women, though not all do. While using exclusively gender-neutral language can enhance inclusivity, it may also reduce clarity. None of the studies included in our review reported data on non-binary participants. We have, therefore, chosen to use both 'women' and gender-neutral language where appropriate. We recognise this is a compromise; however, when we refer to 'women', we ask the readers to interpret this as including all individuals who use the breast screening service, not only those who identify as women.

Role of the funding source

The NIHR Aberdeen-Belfast Evidence Collaboration (ABEC) was funded by the NIHR Evidence Synthesis Programme to conduct this review (project no. NIHR164221). The funder of the study and the UK NSC contributed to the conceptualisation of the research question and study design, but had no role in data collection, data analysis, data interpretation, or writing of the report.

Eligibility for inclusion in the review

Search strategy

Comprehensive literature search strategies were developed by an Information Specialist (PM) to identify relevant published peer-reviewed articles. Major electronic databases were searched, including MEDLINE, Embase, Scopus, The Cochrane Database of Systematic Reviews, Web of Science, and CENTRAL. No language restrictions were applied to the searches, but results were limited to articles published within the last 10 years to capture the most relevant developments in imaging screening techniques. Abstracts of conference proceedings were screened for late-breaking studies; however, those without subsequent full-text publications were not included. Searches were conducted up to end July 2024, with an additional search extending to early November 2024 to ensure the inclusion of potential new studies. The searches focused on imaging modalities to detect breast cancer in individuals with breast density and on manual and automated measurements of breast density. The search strategies included both relevant database index terms and text words. The reference lists of articles selected for full-text appraisal were screened for additional sources of evidence. Ongoing trials were identified by searching major clinical trial registries. All references were exported to Endnote for recording and deduplication. Details of the MEDLINE search are reported in Appendix 1.

Study selection

One reviewer (SND) screened all citations identified by the search strategies using EPPI-reviewer.[16] A second reviewer (CR) independently screened a random sample of citations (20%) to ensure consistency by comparing their results. Potentially relevant articles were retrieved in full, and while the protocol stated that each paper was to be assessed independently by two reviewers, due to time constraints, a 20% sample was assessed by a second reviewer (CR). Any disagreements were resolved through discussion between reviewers or consultation with the wider research team. Multiple publications of the same study were linked and considered together. The number of excluded studies was noted, and the main reasons for exclusion were documented. The study selection process for Objective 2 is depicted through a PRISMA flow diagram (Appendix 2, Figure 1).

The following information were recorded from the included studies: characteristics of publication: (first author, year of publication, recruitment dates, geographical location (urban/rural), name/number of centres, language, screening setting, objectives, inclusion and exclusion criteria); characteristics of participants; number and experience of the health professional involved in the measurement of breast density; frequency of screening, and characteristics of relevant imaging modalities (MRI, CEM, HHUS, ABUS, DBT). Data were extracted by one reviewer using a bespoke data extraction form and checked by a second reviewer (EMcC). Any disagreements were resolved through discussion. Extracted data were recorded using Microsoft Excel®.

A priori, the following approach was planned for prioritising studies for extraction:

- Studies reporting predefined outcomes for supplementary imaging plus mammography versus mammography alone. The comparator for mammography had to be standard 2D mammography and not 3D or synthetic mammography.
- In line with the inclusion criteria, only studies based on breast cancer screening programmes were considered eligible for inclusion. Studies involving selected populations (e.g. people with prior breast cancer, people with genetic risk for breast cancer, and people with strong family history of breast cancer) were excluded.
- Only data relevant to people with dense breasts (heterogeneously dense and extremely dense) were deemed relevant. Studies that recruited a mixed population were included only if they reported data for people with dense breasts separately.
- Studies involving people with negative DM results were deemed suitable for inclusion. In such cases, only data relevant to the supplementary imaging modality were extracted.
- Some studies reported outcomes for people who received either supplementary imaging plus mammography, or mammography alone while others included only people who underwent supplementary imaging. If studies provided separate data for the supplementary imaging plus mammography and mammography alone, those were extracted.
- Studies were examined for potential duplicate reporting. If concerns arose, study authors were contacted. Preliminary reports of subsequently published data were excluded from the final included set.

- Meta-analysis was performed only if at least three studies reported relevant outcome measures.

- Studies with a small sample size (<500 participants) were not excluded but flagged. Sensitivity analyses were planned for those eligible for meta-analysis.

- Attempts were made to contact authors for clarification where study details were unclear.

Table 1 Inclusion criteria for the key question

Key question	Inclusion criteria					
	Population	**Target condition**	**Intervention**	**Comparator**	**Outcome**	**Study type**
What is the effect of an additional imaging modality to supplement standard mammography compared with standard mammography alone for identifying breast cancer in people with dense breasts?	Individuals between 40 and 70 years of age undergoing screening who have been stratified by breast density categories using either visual or automated methods	Breast cancer in people with breast density	Supplemental imaging modalities for detection of breast cancer in people with breast density. These included: Magnetic resonance imaging (full MRI/abbreviated MRI) Contrast-enhanced mammography (CEM) Ultrasound (hand-held HHUS/automated ABUS) Digital breast tomography (DBT)	Mammography (standard 2D mammography) We excluded articles that report direct comparisons of the diagnostic performance of mammography versus another imaging modality or articles that assessed the diagnostic performance of single imaging modalities.	• Cancer detection rate • Interval cancer rate • Recall rate • Positive predictive values • False positive rate • Sensitivity • Specificity • Cancer stage and nodal involvement at detection • Time needed for the additional imaging modality to be performed	Studies published in English in the last 10 years that assessed the performance of supplemental imaging modalities for the detection of breast cancer in people with dense breasts

Data extraction and risk of bias

Two reviewers (SND and EMC) conducted data extraction using a prespecified data extraction form that was developed with input from the Advisory Group. This form was designed in accordance with PRO-EDI[17] initiative guidelines to ensure consideration of equality, diversity and inclusion in participant characteristics within evidence syntheses. The Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was used to evaluate the risk of bias in individual studies.[18] QUADAS-2 consists of four domains: patient selection, index test, reference standard, and flow and timing. For comparative accuracy studies, we used the QUADAS-C version of the tool to assess the risk of bias. Assessments were initially conducted by one reviewer (SND) and subsequently discussed with a second reviewer (GV), who also performed accuracy checks. To develop the risk of bias guidance table (Appendix 3, Table 6), input was sought from the review team and clinical experts.

Data analysis

The findings of included studies were summarised narratively and, when appropriate, through meta-analyses. For studies that compared the performance of supplemental imaging screening to standard mammography for dense breasts, we followed the methods recommended by the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy and presented the results for each screening modality (e.g. HHSU, ABUS, MRI, CEM, DBT) separately. When possible, we presented pooled estimates with 95% CIs for efficacy screening outcomes (e.g., detection rate, recall rate, biopsy rate). Results of random-effects meta-analyses are presented in summary tables and displayed graphically using forest plots. The width of the CIs was used as a measure of the precision and reliability of the effect estimate. The I^2 statistic was used to describe the percentage of total variation across included studies due to heterogeneity rather than chance.[19] We used the following thresholds for the interpretation of I^2: <30% will indicate low heterogeneity, 30–60% moderate heterogeneity and >60% substantial heterogeneity.[17]

Description of the evidence

Database searches yielded 1,740 reports, which were screened at the title and abstract stage. Of those, 178 were selected for full-text screening. The main reasons for exclusion during full-text assessment included: incorrect population (n=34), ineligible intervention (n=29), absence of relevant comparator [mammography] (n=24), incorrect publication type (n=15), no breast density stratification (n=14), inappropriate setting (n=11), duplicates (n=6), irrelevant outcomes (n=4) and incorrect study design (n=1). A total of 39 reports were judged to be relevant to address the research question, and an additional two articles were retrieved from other sources; one accessed in its draft version (pre-publication stage).[20] One was retrieved during the Review 1 search phase and was deemed applicable for Review 2.[21] Of these 41 potentially relevant reports, 5 were subsequently excluded as they reported duplicate data, bringing the final inclusion to 36 reports of original data. As there were multiple reports of individual trials (reporting distinct population subsets), the 36 included studies represented findings from 32 unique trials. The 36 included studies examined a total of 42 supplementary imaging modalities, as some studies assessed more than one imaging modality. Comprehensive results tables, covering demographic details, study characteristics, study design, and screening information for each supplementary screening modality, are provided in Appendix 3 (Tables 1.A to 5.B). A PRISMA flow diagram is presented in Appendix 2, Figure 1, alongside lists of both included and excluded studies. To explore the impact of small study sizes, we agreed with the study's Advisory Group to conduct a sensitivity analysis excluding studies with fewer than 500 participants. This represented a deliberate deviation from the original research protocol.

Characteristics of the included studies

Of the 36 identified studies investigating the role of supplementary imaging for the detection of breast cancer, 10 were published in the last two years (2023/4).[20, 24-32] Geographically, 16 studies focused on studies conducted in Europe,[20-22, 28, 30, 32-42] 10 on studies conducted in North America,[39-48] and 10 on studies conducted in Asia.[23-27, 29, 31, 53-55] Across these 36 studies, a total of 42 supplementary modalities were evaluated. The most frequently investigated modality was digital breast tomosynthesis (DBT), assessed in 18 studies, followed by handheld ultrasound (HHUS) in 9, automated breast ultrasound (ABUS) in 6, and magnetic resonance imaging (MRI) in 6. In addition, one study each evaluated contrast-enhanced mammography (CEM),[20] whole breast sonography (WBS, a combination of ABUS/HHUS),[49] and combined WBS/DBT[49] (Appendix 3, Tables 1.A to 5.B). As data extraction focused exclusively on people with dense

breasts, the mean age of these participants was not consistently reported across studies and some studies failed to provide age-related data. Few studies reported on the ethnicity of participants.

Quality assessment

The overall results of the risk of bias assessment for the 36 included studies are summarised in Figure 1 and detailed results for individual studies are provided in Table 2. In the text below, we describe the main risk of bias assessment issues according to the QUADAS-2 domains.

Patient selection

In most studies, patient selection was judged to be at low risk of bias. However, one study (3%) had an unclear risk due to insufficient reporting,[29] while 5 studies (14%) were deemed at high risk due to opportunistic patient selection.[25, 29, 45, 52, 53] Examples of this include people opting for supplementary screening, access being limited to those with health insurance, or availability through workplace programmes.

There was no concern regarding case-control studies since these were excluded, as per our guidance. Similarly, our eligibility criteria restricted inclusion to studies conducted within the breast cancer screening setting. Studies focusing on high-risk populations, such as people with genetic mutations or prior history of breast cancer were excluded, except for those specifically assessing people with dense breasts. Inappropriate exclusions were rare but present in 5 studies (14%).[25, 29, 45, 52, 53] In these cases, participants may have been excluded for reasons such as lack of insurance coverage for supplementary imaging, lack of workplace access to imaging, or personal choice to decline additional imaging after being informed of potential radiation exposure.

In terms of whether the study population aligned with the population of interest for this systematic review, 14 (39%) studies were deemed to have applicability concerns. The primary reason was that the age profile of most participants did not match the UK breast screening age range. Two studies from the Dutch DENSE trial on MRI were flagged for applicability concerns as they included participants with extremely dense breasts, whereas the scope of this systematic review includes both heterogeneously and extremely dense breast categories.[37, 39] A further study on MRI was flagged for applicability concerns as the comparator arm was not a true mammography-only group.[22] In this study, approximately 65% of participants had both a negative ultra-

sound and a negative mammography. Individuals who had a negative ultrasound were not eligible for inclusion in the MRI imaging study; however, the exact number was not specified. As a result, the comparator arm had an advantage compared with a mammography-only arm, likely leading to an underestimation of MRI's cancer detection performance.

We considered studies to be at low risk of bias if they used paired or randomised designs or provided evidence of an absence of key confounders imbalance; 17 (47.2%) studies [20-21, 23-24, 26, 29, 30, 32, 34-41, 54] met these criteria. However, 6 studies (17%)[22, 25, 28, 45, 52-53] were deemed at high risk of bias because of concerns over their non-randomised design, evidence of confounders imbalance with no statistical adjustment, or evidence that participants were assigned to the index groups based on non-random factors. In these high-risk studies, willingness to participate,[25, 28, 52] workplace or private screening availability,[53] and willingness to accept a higher radiation dose[45] were identified as key selection influences. Lastly, 13 studies (38%)[27, 31, 33, 42-44, 46-51, 55] were deemed to have an unclear risk of bias. These studies lacked a randomised or paired design and provided little or no detail on how participants were assigned to the index test groups, making it difficult to assess potential selection biases.

Index test

In some studies, supplementary imaging was conducted with knowledge of the comparator test, while in others, blinding was applied. Nevertheless, this was not considered to be a risk of bias for this systematic review, as supplementary imaging is intended to provide additive benefits to the comparator test by offering different perspectives. Instead, we regarded reciprocal masking as a potential source of heterogeneity warranting further investigation. Regarding pre-specific thresholds, most studies used a common index test threshold, such as BI-RADS or an equivalent internationally recognised standard definition, which was assessed as low risk of bias. However, one study conducted in Spain was deemed high risk in terms of applicability because the supplemental imaging consisted of a combination of DBT and synthesised mammography.[28] Although this study met the general inclusion criteria, it was excluded from the meta-analysis for this reason. For comparative questions, all studies were assessed as low risk of bias.

Reference standard

Differential verification and incorporation biases are intrinsic in breast cancer screening studies. Typically, index test-positive participants undergo further tests, often biopsy, while those with negative index test results are verified through follow-up mammography or combined tests, after one or more years. Notably, statistical simulation studies suggest that the risk of biased estimates due to incorporation and differential verification biases is generally low across most

breast cancer screening programmes.[56] We expected low-risk studies to employ and clearly describe multiple case-finding methods beyond standard imaging follow-up. These could include linkage to cancer registries, electronic medical records, death certificates, or follow-up surveys. When studies did not specify such additional verification methods, we considered them to have an unclear risk of bias. Furthermore, if studies did not confirm repeat imaging after 1 or more years, we classified them at high risk of bias. Reassuringly, 24 studies (67%) adopted at least one of these methods, ensuring a low risk of bias. [20-21, 24, 26-28, 31-34, 36-44, 46, 47, 52, 54, 55]

However, a study from Japan raised some concerns, as it mentioned referral to a specialist institute but lacked sufficient details about the reference standard for positive or negative imaging cases.[25] Ten studies (28%) were deemed to have a high risk of bias because of either the absence or inadequacy of follow-up details for participants with negative mammography results. [23, 29, 30, 35, 45, 48-51, 53]

Flow and timing

Most studies showed either an unclear (17, 47%)[21, 25, 29, 30, 34, 35, 38, 40, 43-47, 49, 51, 53, 55] or high (8, 22%)[20, 22, 23, 27, 46, 48] risk of bias concerning the flow and timing domain of the QUADAS-2 tool. This was mainly due to insufficient details or missing data, particularly among participants who tested negative on the index screening and were subsequently verified through follow-up. In several cases where the risk of bias was high, participants with inadequate follow-up were either excluded from the analysis,[22, 24, 27, 31] lost to follow-up,[26, 50, 52] or there were no details given.[23]

For comparative analyses, we expected the supplementary imaging to be performed within a reasonable timeframe from digital mammography to ensure consistency in the assessment of their screening performance. However, in some studies (n=11, 31%), the interval between imaging modalities was not reported, resulting in an assessment of unclear risk of bias. One study reported that the MRI supplementary imaging was scheduled "*not to exceed two months*"[28] while another study reported that some imaging occurred on the same day but did not specify the timing of the remaining tests.[31] Moreover, some studies assessing MRI as a supplementary modality reported several months between screening tests and, therefore, were judged to have a high risk of bias.[20, 37, 39] This was due to the possibility that cancers could develop or previous undetected cancers could become visible within that timeframe. In addition, 12 studies (33%)

were judged at unclear risk of bias because of potential discrepancies in missing data between single and combined imaging tests. [22-23, 34, 43, 45-48, 50-51, 53, 55]

Figure 1 Risk of Bias and Applicability Assessments using QUADAS-2 and QUADAS-C

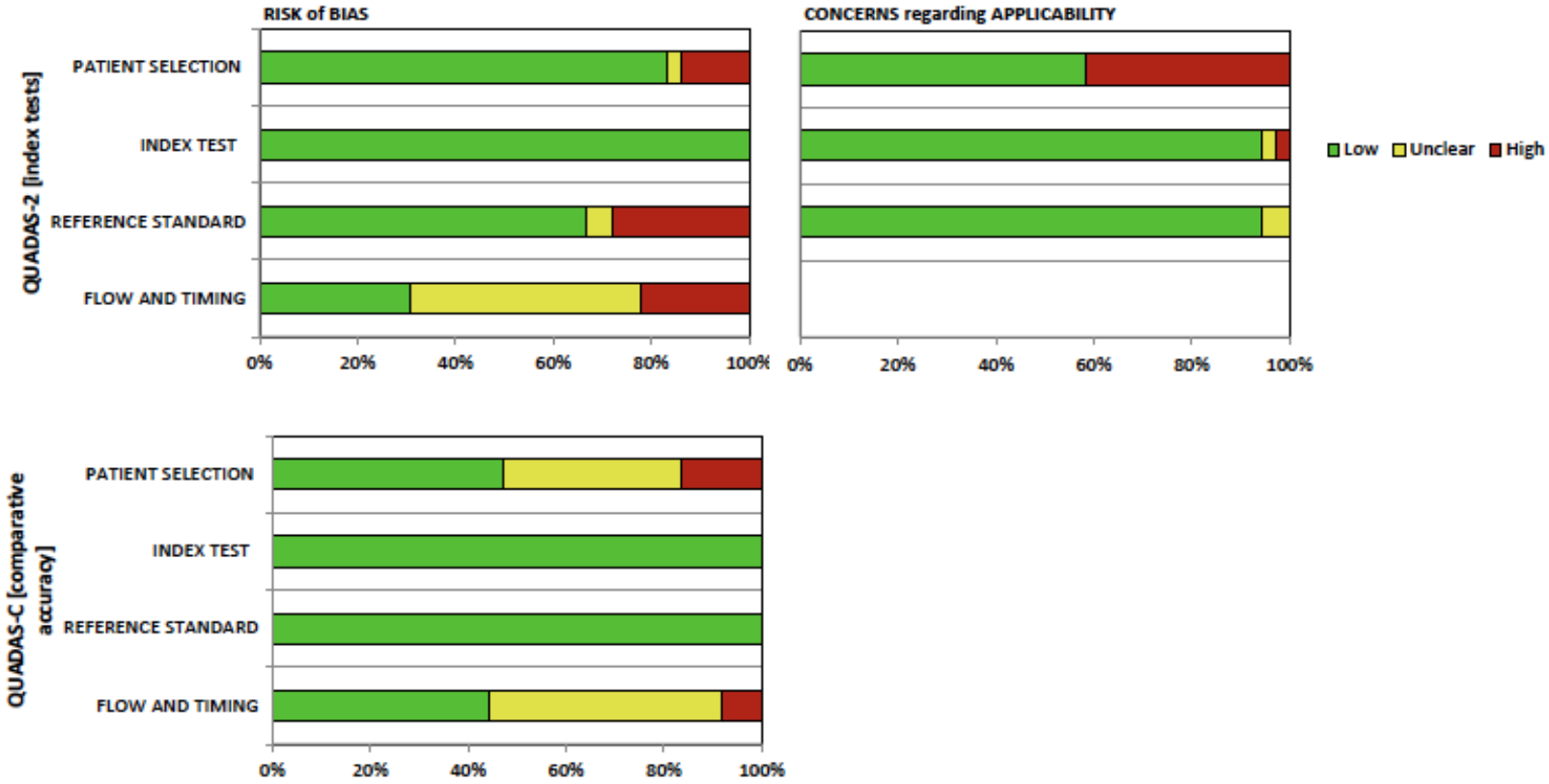


Table 2 Risk of Bias Judgements for Individual Studies

Study	Risk of bias (QUADAS-2)				Applicability concerns (QUADAS-2)			Risk of bias (QUADAS-C)			
	P	I	R	FT	P	I	R	P	I	R	FT
Pulida-Carmona, 2024[28]	✗	✓	✓	✓	✗	✗	✓	✗	✓	✓	✓
Gilbert, 2024[20]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Ha, 2024[31]	✓	✓	✓	✗	✓	✓	✓	?	✓	✓	?
Lee, 2024[27]	✓	✓	✓	✗	✓	✓	✓	?	✓	✓	✓
Nakamura, 2024[25]	✗	✓	?	?	✗	✓	✓	✗	✓	✓	✓
Kaiser, 2024[32]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	?
Olinder, 2023[30]	✓	✓	✗	?	✓	✓	✓	✓	✓	✓	✓
Kwon, 2023[24]	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	?
Ren, 2023[26]	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
Rani, 2023[29]	?	✓	✗	?	✓	✓	✓	✓	✓	✓	✓
Pattacini, 2022[36]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Pang, 2022[51]	✓	✓	✗	?	✓	?	✓	?	✓	✓	?
Gatta, 2021[33]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	?
Harada-Shoji,2021[55]	✓	✓	✓	?	✗	✓	✓	?	✓	✓	?
Durand, 2021[46]	✓	✓	✓	?	✓	✓	✓	?	✓	✓	?
Ban, 2021[53]	✗	✓	✗	?	✗	✓	?	✗	✓	✓	?
Veenhuizen, 2021[39]	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗
Bakker, 2019[37]	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗
Johnson, 2019[41]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Stepenak, 2019[52]	✗	✓	✓	✗	✗	✓	✓	✗	✓	✓	✓

Osteras, 2019[21]	✓	✓	✓	?	✓	✓	✓	✓	✓	✓	✓
Buchberger, 2018[34]	✓	✓	✓	?	✓	✓	✓	✓	✓	✓	?
Upadhyay, 2018[38]	✓	✓	✓	?	✗	✓	✓	✓	✓	✓	✓
Kuhl, 2017[22]	✓	✓	?	✗	✗	✓	✓	✗	✓	✓	?
Chen, 2017[23]	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	?
Freer, 2017[45]	✗	✓	✗	?	✗	✓	?	✗	✓	✓	?
Rose, 2017[48]	✓	✓	✗	?	✗	✓	✓	?	✓	✓	?
McDonald, 2016[43]	✓	✓	✓	?	✓	✓	✓	?	✓	✓	?
Conant, 2016[47]	✓	✓	✓	?	✓	✓	✓	?	✓	✓	?
Bernardi, 2016[35]	✓	✓	✗	?	✓	✓	✓	✓	✓	✓	✓
Starikov, 2016[49]	✓	✓	✗	?	✓	✓	✓	?	✓	✓	✓
Tagliafico, 2016[40]	✓	✓	✓	?	✓	✓	✓	✓	✓	✓	✓
Wilczek, 2016[42]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	✓
McDonald, 2015[44]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	?
Brem, 2015[50]	✓	✓	✗	✗	✗	✓	✓	?	✓	✓	?
Chang, 2015[54]	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓

P = patient selection; I = index test; R = reference standard; FT = flow and timing
✓ indicates low risk of bias; ✗ indicates high risk of bias; ? indicates unclear risk of bias

Review findings

Meta-analyses were performed for four imaging modalities, DBT, ABUS, HHUS and MRI versus standard digital mammography (see Table 3 below). Sensitivity analyses were conducted to determine the impact of removing studies with very small sample sizes (<500). Detailed results tables are presented in Appendix 3, Tables 1.A to 5.B. An MRI-specific results summary is presented in Table 6.

Table 3 Combined pooled estimates of the screening performance measures of four supplementary imaging modalities compared to digital mammography, with and without sensitivity analysis

Parameter		DBT	ABUS	HHUS	MRI
CDR	Number of studies	13	6	8	6[ϕ]
	CDR for mammography only, per 1,000 exams	5.43 (3.75, 7.11)	4.52 (3.03, 6.02)	2.68 (1.77, 3.59)	-[#]
	Additional CDR, per 1,000 exams	1.69 (0.81, 2.58)	2.3 (1.28, 3.33)	2.57 (0.99, 4.14)	18.92 (15.41, 22.43)
	*Sensitivity analysis^*	-	-	-	*17.23 (14.38, 20.08)*
Recall rate	Number of studies	11	5	5	3
	Recall rate for mammography only, %	12.56 (8.96,16.15)	5.97 (-0.32,12.27)	4.25 (1.37,7.12)	-[#]
	Difference in recall rate, %	-2.19 (-4.03, -0.34)	5.49 (0.85, 10.13)	6.65 (1.02, 12.27)	7.98 (4.62, 11.34)
	Sensitivity analysis[§]	*-1.94 (-3.83, -0.05)*	-	-	*9.55 (8.85,10.24)*
Biopsy rate	Number of studies	3	5	2	3
	Biopsy rate for mammography only, %	2.07 (1.57,2.57)	1.83 (0.30,3.36)	-	-
	Difference in biopsy rate, %	0.32 (0.15, 1.32)	1.62 (0.54, 2.70)	n/a	n/a

CDR, cancer detection rate; DBT, digital breast tomosynthesis; ABUS, automated breast ultrasound; HHUS, hand-held ultrasound; MRI, magnetic resonance imaging. Data in parentheses are 95% confidence intervals
^ Kaiser[32] and Chen[23] removed for MRI sensitivity analysis; § Upadhyay[38] removed for DBT sensitivity analysis; # These values are zero because the comparator arm in the MRI studies is 'negative mammography'; [ϕ] Chen reports both abbreviated and full protocol MRI, so there are 6 comparisons in 5 studies represented in the MRI meta-analysis

Cancer detection rate (CDR)

Pooled estimates indicate that all four supplementary modalities detect more cancers compared to DM only (Figure 2, A-D). MRI demonstrates the most significant difference in CDR, with a pooled effect size of 18.92 (95% CI 15.41 to 22.43) across four studies, indicating that MRI detected nearly 19 additional cases per 1,000 patients compared to mammography alone. Notably, the MRI meta-analysis shows the lowest heterogeneity and provides the most precise estimate of the CDR increase. A sensitivity analysis excluding two studies with small sample sizes (Kaiser et al.,2024; Chen et al, 2017; <500 participants), resulted in a slightly lower, but consistent, CDR difference of 17.23 (95% CI 14.38 to 20.08), indicating that MRI detected over 17 additional cancers per 1,000 participants, across three large studies.

Findings from the DENSE studies illustrate the impact of screening rounds on CDR. The study by Veenhuizen et al.[39] was not included in the meta-analysis as it reports data for the incident screening round. This study reported a CDR of 5.8 (95% CI 3.8 to 9) per 1,000 examinations, which is considerably lower than the Bakker study,[37] which analysed the prevalent screening round (CDR 16.5, 95% CI 12.92 to 20.12). The BRAID study by Gilbert et al. (2024, confidential communication),[20] and the German study by Kuhl et al.[22] more closely align with the prevalent round of the DENSE trial in both context and outcome.[37] However, both DENSE studies raised some concerns regarding patient selection applicability, as they exclusively included participants with extremely dense breasts, potentially limiting generalisability.[37,39] The study by Kuhl et al. provided data for both prevalent and incident MRI screening rounds; however, only the prevalent round included data for individuals with dense breasts. Consequently, only the prevalent round data were included in this review.[22] This study raised some concerns about applicability, as 65% of participants with negative mammography results also had negative ultrasound imaging. The number of participants excluded from the MRI screening study due to cancer detection on ultrasound - but not on mammography - was not specified. Consequently, the reported MRI detection rate of 20.28 additional cancers per 1,000 screenings is likely to be an underestimation of its true performance.

The BRAID study utilised an abbreviated MRI protocol,[20] while the DENSE and Kuhl studies employed a full MRI protocol.[22, 37, 39] The Chinese study by Chen et al. assessed both abbreviated and full-protocol MRI.[23] While we were not able to conduct separate meta-analyses,

Figure 2.D indicates that the studies using an abbreviated MRI protocol showed cancer detection rates comparable to those using a full MRI protocol.

The other three supplementary modalities - DBT, ABUS, and HHUS - show smaller, yet statistically significant increases in CDR, with narrower confidence intervals compared to MRI. The pooled differences in CDR are 1.69 (95% CI 0.81 to 2.58), 2.30 (95% CI 1.28 to 3.33), and 2.57 (95% CI 0.99 to 4.14) for DBT, ABUS, and HHUS, respectively. Both ABUS and HHUS show similar increases in CDR compared to mammography alone. It is worth noting that while DBT and HHUS analyses show considerable heterogeneity, ABUS demonstrates comparatively lower heterogeneity. The UK BRAID study (see Figure 2 B) reported the highest differential CDR among the six ABUS studies included in the meta-analysis (4.2, 95% CI 1.15 to 7.25).[20]

CEM was not included in the meta-analysis, as only one eligible study - the BRAID trial - investigated this supplementary modality. In this study, the CDR for CEM following a negative mammogram was 19.2 (95% CI: 13.7–26.1) per 1,000 screenings, indicating that CEM identified, on average, 19 additional cancers per 1,000 women with a negative mammography result. This detection rate was comparable to that reported for MRI in the same study.[20]

Interval cancers

Only 2 DBT studies provided data on interval cancers. One study from Spain combined DBT with synthesised mammography in the supplemental arm and reported interval cancer rates (%) of 0.95 (95% CI 0.55 to 1.35) in the supplementary group and 3.17 (95% CI 2.64 to 3.62) in the mammography-only group.[28] The Italian RETomo Trial reported interval cancer rates (%) of 0.23% in the supplementary group and 0.25% in the mammography-only group.[36]

None of the included ABUS studies assessed interval cancers. However, 2 HHUS studies provided relevant data. One study from Japan examined individuals aged 40-49 years and found a lower interval cancer rate (%) in the HHUS/mammography group (0.5, 95% CI 0.1 to 1.1) compared to the mammography-only group (1.8, 95% CI 0.7 to 2.9).[55] Similarly, 1 study from China studied a relatively younger cohort of individuals and reported “no interval cancers” in the HHUS/negative mammography group.[54]

Two European studies reported on interval cancers for MRI. In the DENSE study,[37] participants who’d had negative mammography were assigned to either MRI supplementary imaging, or to mammography only. There were 4 interval cancers detected in participants in the

supplementary MRI arm, and 16 were detected among those who were invited to MRI but did not participate. Participants in the mammography-only group had 161 interval cancers. These numbers were equivalent to 2.5 (95% CI 1.6 to 3.8) interval cancers per 1,000 exams in the MRI invitation group and 5.0 (95% CI 4.3 to 5.5.8) interval cancers per 1,000 in the mammography-only group. Among those who underwent MRI, the interval cancer rate per 1,000 exams was 0.8. The study by Kuhl et al. reported that no interval cancers were detected during a follow-up period of at least two years.[22]

Recall rate

Across all 4 supplementary modalities, we observed substantial heterogeneity (I^2 >90%) leading to wide confidence intervals around the pooled effect sizes. The pooled differences in recall rates range from a 2.19% reduction to a 6.67% increase. Among studies comparing DBT with DM alone, most report a reduction in recall rate. This is reflected in the pooled difference of -2.19 (95% CI -4.03 to 0.34). The sensitivity analysis excluding the Upadhyay 2018[38] study is consistent with a pooled difference of -1.94 (95% CI -3.83 to -0.05). For the other 3 modalities, the pooled differences represent increases in the percentage recall rate of 5.49% (95% CI 0.85 to 10.13), 6.65% (95% CI 1.02 to 12.27), and 7.98% (95% CI 4.62 to 11.34) from ABUS, HHUS, and MRI, respectively. The MRI sensitivity analysis excluding one study[32] of small sample size, results in a larger increase in recall rate (9.55%, 95% CI 8.85 to 10.24).

There are notable differences among studies in their definition of recall rate, as illustrated by the individual studies' results in the meta-analysis forest plots (Figure 3). In the DBT meta-analysis (Figure 3, A), which includes 11 studies, the recall rate was defined as BI-RADS ‘0’ in three studies,[48, 49, 52] BI-RADS 0,4 or 5 in two studies[44, 51], BI-RADS 0,3,4, or 5 in another two and ‘*a need for additional evaluation*’ in two others,[41, 49] while 2 studies did not report their definition.[30,46]

For ABUS (Figure 3, B), among the 5 studies included in the meta-analysis, one defined recall rate as BI-RADS ‘0’,[50] another as BI-RADS 3,4, or 5,[20] and three other studies as ‘*abnormal or suspicious findings*’.[24, 33, 42] The highest recall differential rate was observed in a US study (Brem et al),[50] despite their reported definition of recall as BI-RADS ‘0’ (inconclusive), which seems inconsistent given their high recall rate. The authors attributed this increase to improved sensitivity with the combined ABUS/mammography approach. However, compared to the other

ABUS studies, the CDR reported by Brem et al. was not significantly higher.[50] It is plausible that the younger age of the study population (>25 years) might partially explain the high recall rate.

For HHUS (Figure 3, C), the recall rate was described variably across studies, including as BI-RADS 0, 4 or 5,[34] an abnormality score >9,[27] BI-RADS 3, 4 or 5 (also referred to as Abnormal Interpretation Rate),[27] or 'additional testing required'.[55] One study did not provide a definition.[21] While most HHUS studies reported a modest differential recall rate, the study by Ha (2024) had an exceptionally high difference in recall rate between the HHUS/mammography arm and the mammography-alone arm.[31] However, it was not accompanied by a correspondingly large CDR increase (Figure 2, D).

For MRI, the definition of recall was relatively consistent across all four studies reporting this parameter.[20, 32, 37, 39] Three studies define recall as BI-RADS 3,4, or 5,[20, 37, 39] while 1 study uses the definition 'positive MRI findings'.[32] The second round of the DENSE study[39] was excluded from the meta-analysis as it was not deemed comparable to the other studies, each of which represents a prevalent screening context (Figure 3, D). Nevertheless, this study reported a lower recall rate for the second round of MRI screening (3.2%, 95% CI 2.61 to 3.79) compared to the pooled estimate of 7.98% (95% CI 4.62 to 11.34) for the MRI meta-analysis, mirroring the lower CDR observed in subsequent screening rounds.[39] Consistent with the greater differential observed for CDR, MRI/mammography demonstrated a higher recall rate than mammography alone in both the BRAID[20] study and the DENSE study (first round).[37]

The BRAID study reported that the recall rate for CEM as a supplementary imaging modality was 9.7% (95% CI 8.4 to 11), matching the recall rate for MRI in the same study.[20]

Table 4 Sensitivity, specificity, PPV1 and PPV3, per supplementary modality; presented as % (95% CI) where reported

	Author, year	Arm	Sensitivity, % (95% CI)	Specificity, % (95% CI)^	PPV1 for recall, % (95% CI)	PPV3 for biopsy, % (95% CI)
DBT	Olinder, 2023[30]	DBT plus DM	Grade C 78.4 (67.7, 86.2) Grade D 90.9 (72.2, 97.5)	Grade C 96.1 (95.6, 96.6) Grade D 94.5 (93.1, 95.7)	Grade C 23.6 (18.7, 29.3) Grade D 23.3 (15.6, 33.2)	Grade C 40.6 (32.9, 48.8) Grade D 37.7 (25.9, 51.2)
		DM only	Grade C 66.2 (54.9, 76) Grade D 50 (30.7, 69.3)	Grade C 97.4 (96.9, 97.8) Grade D 96.3 (95.1, 97.2)	Grade C 27.8 (21.7, 34.9) Grade D 19.6 (11.3, 31.8)	Grade C 41.9 (33.3, 50.9) Grade D 32.4 (19.1, 49.2)
	Pattacini, 20226	DBT plus DM			18.4	
		DM only			10.8	
	Pang, 2021[51]	DBT plus DM			Grade C 7.7 Grade D 5.2	
		DM only			Grade C 5.2 Grade D 5.3	
	Rose, 2017[48]	DBT plus DM			2.5	17.2
		DM only			1.6	13.4
	Conant, 2016[47]	DBT plus DM			6	
		DM only			3.7	
	Tagliofico, 2016[40]	DBT plus negative DM				37 (21.3, 55.4)
ABUS	Kwon, 2023[24]	ABUS plus DM	63.6 (40.9, 81.8)		6.6 (5.6, 7.4)	27.8 (23.4, 31)
		DM only	94.6 (93.6, 95.5)		10.8 (8.8, 12.2)	33.3 (26.7, 37.8)
	Ren, 2023[26]	ABUS plus DM	Age 45-54y: 96.55 (80.37, 99.82) Age 55-64y: 100 (69.87, 100)	Age 45-54y: 97.11 (96.72, 97.45) Age 55-64y: 98.01 (97.26, 98.56)		Age 45-54y: 16.87 (11.68, 23.63) Age 55-64y: 34.29 (19.69, 52.27)

		DM only	Age 45-54y: 72.41 (52.51, 86.55) Age 55-64y: 75 (42.84, 93.31)	Age 45-54y: 98.9 (98.65, 99.11) Age 55-64y: 99.08 (98.52, 99.44)		Age 45-54y: 24.71 (16.27, 35.47) Age 55-64y: 40.91 (21.48, 63.32)
	Gatta, 2021[33]	ABUS plus DM	93.5 (79.2, 98.2)	87 (71, 94.8)	24.8 (13.7, 43.2)	41 (26.4, 59.2)
		DM only	58.8 (30.9, 78.3)	94 (73, 98)	68 (41.3, 82.2)	58.8 (30.9, 78.3)
	Wilczek, 2016[42]	ABUS plus DM	100	98.4 (97.8, 98.9)	28.9 (14.3, 42.3)	47.8 (27, 66.7)
		DM only	63.6 (33.3, 90.9)	99 (98.5, 99.4)	30.4 (12.3, 49.1)	63.6 (33.3, 90)
	Brem, 2015[50]	ABUS plus DM	100		2.6 (2.1, 3.1)	9.8 (8.1, 11.7)
		DM only	73.2		3.6 (3.28, 4.4)	14 (11.2, 16.8)
HHUS	Ha, 2024[31]	HHUS plus DM	97 (84.7, 99.5)	77.6 (76.5, 78.6)	2.5 (1.7, 3.4)	
		DM only	57.6 (40.8, 72.8)	94.3 (93.6, 94.8)	5.5 (3.6, 8.5)	
	Lee, 2024[27]	HHUS plus DM	100 (73.5, 100)	89.1 (87.3, 90.7)		
		DM only	66.7 (34.8, 90.1)	96.2 (95, 97.2)	13.8 (9, 20.6)	53.3 (35.3, 70.6)
	Nakamura, 2024[25]	HHUS plus DM			10.5	
		DM only			9.5	
	Ren, 2023[26]	HHUS plus DM	Age 45-54y: 93.1 (75.78, 98.8) Age 55-64y: 100 (69.87, 100)	Age 45-54y: 97.34 (96.97, 97.67) Age 55-64y: 97.91 (97.15, 98.48)		Age 45-54y: 15.34 (10.52, 21.72) Age 55-64y: 30 (17.09, 46.71)
		DM only	Age 45-54y: 72.41 (52.51, 86.55) Age 55-64y: 75 (42.84, 93.31)	Age 45-54y: 98.9 (98.65, 99.11) Age 55-64y: 99.08 (98.52, 99.44)		Age 45-54y: 24.71 (16.27, 35.47) Age 55-64y: 40.91 (21.48, 63.32)
	Rani, 2023[29]	HHUS plus DM	100	87.5		
		DM only	25	100		
	Harada-Shoji, 2021[55]	HHUS plus DM	93.2 (85.7, 100)	85.4 (84.5, 86.3)		
		DM only	70.6 (55.3, 85.9)	91.7 (91, 92.4)		
	Buchberger, 2018[34]	HHUS plus DM	81.3 (72, 88.5)		10.5 (8.4, 12.9)	37.7 (31.1, 44.7)
		DM only	61.5 (51, 71.2)		13.3 (10.3, 16.8)	52.7 (43, 62.2)

	Tagliafico, 2016[40]	HHUS plus negative DM				48 (34.1, 63.9)
MRI	Kaiser, 2024[32]	Fp-MRI plus negative DM			50.0 (15.7, 84.3)	
	Veenhuizen, 2021[39]	Fp-MRI plus negative DM			18.2 (12.1, 26.4)	23.8 (16, 33.9)
	Bakker, 2019[37]	Fp-MRI plus negative DM	95.2 (88.1, 98.7)	92	17.4 (14.25, 21.2)	26.3 (21.7, 31.6)
	Chen, 2017[23]	Fp- MRI plus negative DM	100	94.6	41	
		Ab-MRI plus negative DM	93.8	88.3	27.7	

DBT, digital breast tomosynthesis; ABUS, automated breast ultrasound; HHUS, hand-held ultrasound; DM, digital mammography; MRI, magnetic resonance imaging; Fp-MRI, full protocol MRI; Ab-MRI, abbreviated MRI; CI confidence interval; PPV1, positive predictive value (number of cancers diagnosed per number of positive screens; PPV3, positive predictive value (biopsy proven predictive value); ^ Grade C and D refer to Breast Imaging Reporting and Data System (BI-RADS)[6] density categories heterogeneously and extremely dense, respectively

Biopsy rate

Biopsy rate (%) was reported in 13 studies; however, as only 2 studies[34, 55] compared HHUS/mammography with mammography alone, no meta-analysis was performed for this modality. For DBT (Figure 4, A), the combined effect size of 0.32 (95% CI -0.12 to 0.76) indicates there is no significant difference in biopsy rate between DBT/mammography and mammography only. In contrast, ABUS showed a pooled effect size of 1.62 (95% CI 0.54 to 2.70), suggesting small increases in the biopsy rates for this modality (Figures 4, B). However, both meta-analyses showed substantial heterogeneity (I^2 >75%), precluding the ability to draw definitive conclusions. Among the ABUS studies, Brem et al, which reported the greatest difference in CDR, also demonstrated the largest increase in biopsy rate.[50] While the biopsy rate was reported for three MRI studies, the second round of the DENSE study[39] is not directly comparable to either the first round of the DENSE study[37] or the BRAID study,[20] as both of the latter present biopsy rates for the prevalent rounds. The biopsy rates in the BRAID[20] and first-round DENSE[37] studies were 4.93% (95% CI 3.98 to 5.88) and 6.27% (95% CI 5.58 to 6.96), respectively, while the biopsy rate in the second round of the DENSE study[39] was 2.44% (95% CI 1.93 to 2.96), illustrating that, like the CDR and recall rates, the biopsy rate decreases in subsequent screening rounds. In the BRAID study, the CEM biopsy rate (4.4, 95% CI 3.5 to 5.4) was comparable to that of the MRI.[20]

Sensitivity and specificity, PPV1 and PPV3 values

These estimates of accuracy were not widely reported, and meta-analysis was not performed. Detailed values are shown in Table 4. Sensitivity was higher in most cases in the supplementary groups compared to the mammography-only groups. One exception was in a study from South Korea in which the ABUS plus mammography group had a lower sensitivity (63.6%, 95% CI 40.9 to 81.8) than the mammography-only group (94.6%, 95% CI 93.6 to 95.5).[24] One MRI study reported a sensitivity of 95.2% (95% CI 88.1 to 98.7) for MRI following negative mammography. A second study reported 100% sensitivity for full MRI protocol and 93.8% sensitivity for abbreviated MRI protocol, both following negative mammography. Specificity was >90% for most studies with a few exceptions: a study conducted in Italy reported a specificity of 87.0% (95% CI 71.0 to 94.8) in the ABUS plus mammography arm compared to 94.0% (95% CI 73.0 to 98.0) in the mammography only arm.[33] A study from South Korea reported a specificity of 77.6% (95% CI 76.5 to 78.6) in the HHUS plus mammography arm compared to a specificity of 94.3% (95% CI 93.6 to 94.8) in the mammography-only arm.[31] Similarly, a Japanese study found a

specificity of 83.4% (95% CI 84.5 to 86.3) for HHUS plus mammography versus 97.7% (95% CI 91 to 92.4) in the mammography-only arm.[55] For supplementary MRI, one study reported 92% specificity for MRI following negative mammography,[37] and a second small study of <500 participants[23] reported specificity of 94.6% and 88.5% for Fp-MRI and Ab-MRI, respectively. PPV1 and PPV3 values are detailed in Table 4.

False positives

False positive rates were inconsistently reported across studies, with data available from only three DBT studies, two ABUS studies, two MRI studies, and one HHUS study. Due to this limited reporting, pooling of estimates of effects was deemed unsuitable. However, in general, false positives were more frequent in the supplementary imaging group than in the mammography group across all modalities.

DBT

Among the three DBT studies, two found a higher false positive rate in the DBT/mammography group compared to mammography alone. In a Swedish study, for participants with grade C breast density, the false positive rate was higher in the DBT/mammography group (3.8%, 95% CI 3.3 to 4.4) compared to mammography alone (2.6%, 95% CI 2.2 to 3.1).[40] Similarly, for those with extremely dense breasts, false positives were more frequent in the supplementary imaging group than in the mammography alone group. An Italian study reported similar findings, reinforcing the trend of increased false positives with DBT/mammography.[33] However, a Norwegian study found the opposite, a higher false positive rate in the mammography-only group compared to the supplementary DBT/mammography group.[21]

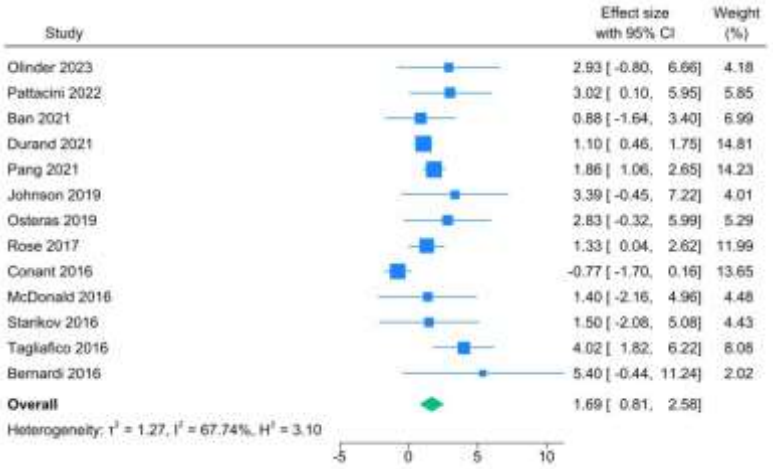
ABUS

Two studies reported false positive rates for ABUS, both indicating a higher rate in the ABUS/mammography group compared to mammography alone. A South Korean study found false positives to be significantly higher in the ABUS/mammography group (2.4 %, 95% CI 2.2 to 2.6) compared to the mammography alone (1.3%, 95% CI 1.1 to 1.4).[24] A Chinese study examined false positives across age groups, reporting that rates were higher in the supplementary imaging group, particularly among younger participants (age 45-54 years).[26]

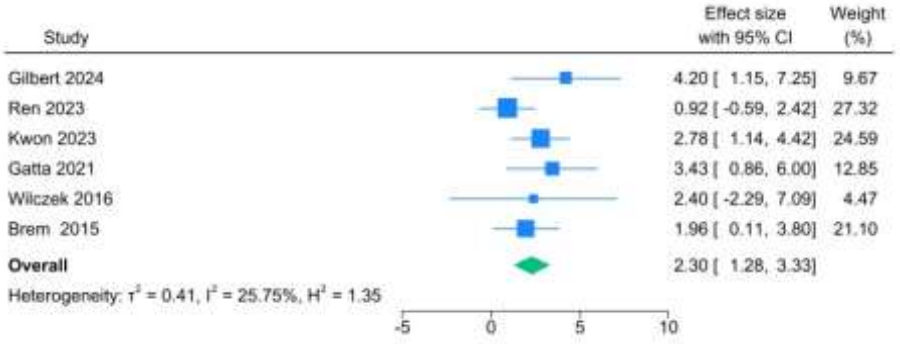
Figure 2 Forest plots showing the effect size of supplementary A) DBT B) ABUS C) HHUS and D) MRI versus DM alone for Cancer Detection Rate

^AB, abbreviated MRI; FP, full protocol MRI

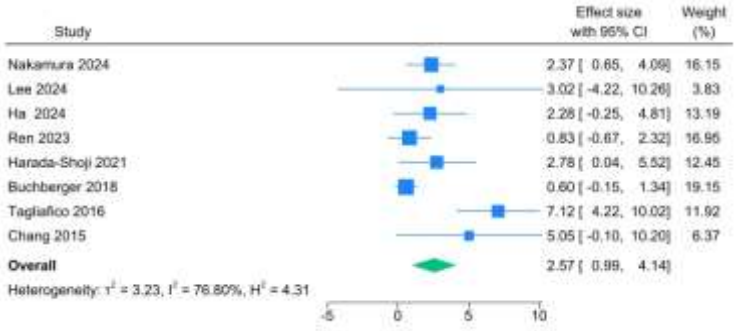
A



B



C



D^

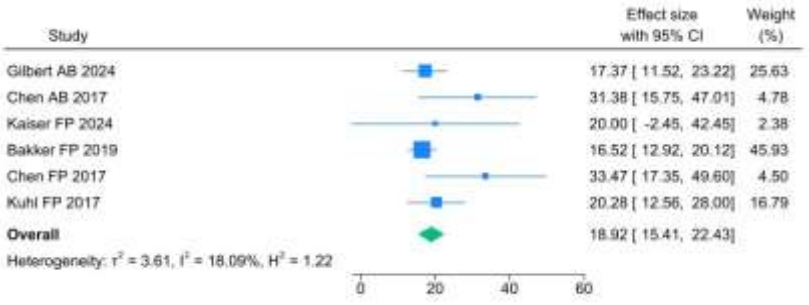


Figure 3 Forest plots showing the effect size of supplementary A) DBT B) ABUS C) HHUS and D) MRI versus DM alone for % Recall Rate

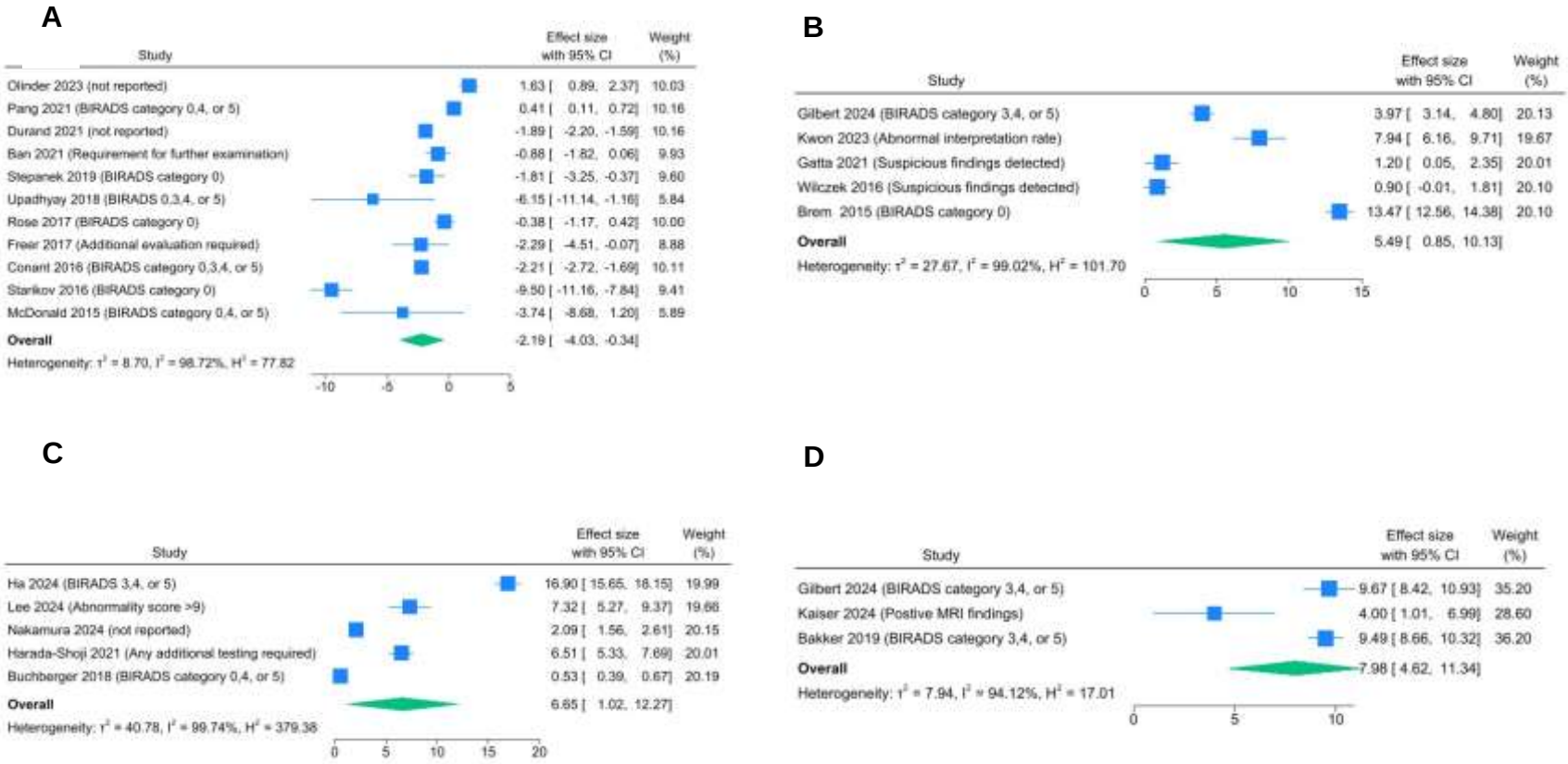
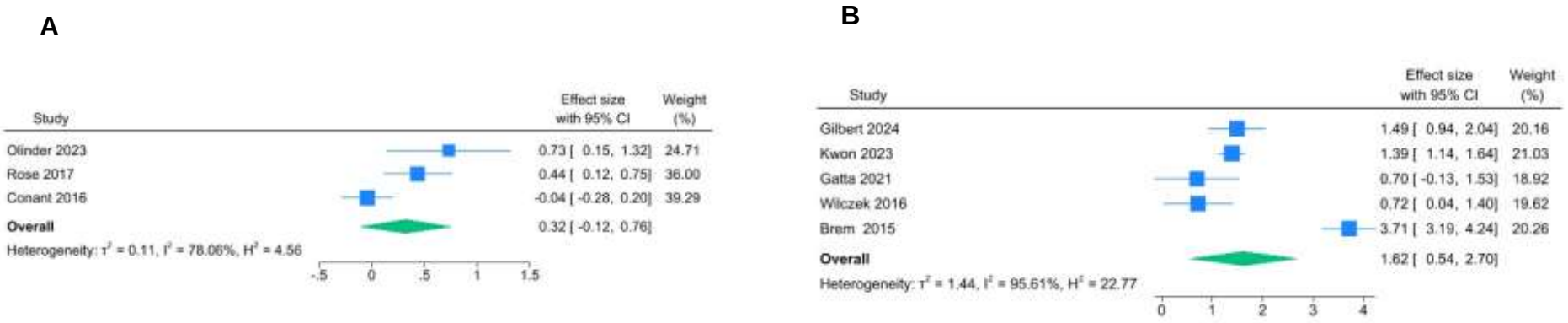


Figure 4 Forest plots showing the effect size of supplementary A) DBT B) ABUS versus DM alone for % Biopsy Rate



HHUS

The same Chinese study also reported false positive rates for HHUS and showed similar findings, false positives were more frequent in the supplementary imaging group and among younger participants.[26]

MRI

Two studies reported false positive rates in DM-negative participants undergoing supplementary MRI as part of the DENSE study in the Netherlands.[37, 39] The false positive rate was 79.8 per 1,000 screenings (95% CI 72.4 to 87.9) in the first screening round, equivalent to 8%,[37] decreasing to 26.3 per 1,000 screenings (95% CI 21.5 to 32.3) in the second round, equivalent to 2.6% (Table 6).[39]

Pathological characteristics of cancers detected

Only 13 studies described the pathological characteristics of cancers detected across various imaging modalities. The nature of the detected cancers, alongside their tumour size where reported, is summarised below for each modality and further detailed in Table 5 below.

DBT

An Italian study found that DBT plus mammography detected 40 invasive cancers and ductal carcinoma in situ (DCIS) compared to 26 invasive cancers and five DCIS detected by mammography alone.[36] Regarding interval cancers, 13 invasive cancers were diagnosed in the DBT/DM arm versus 15 invasive cancers in the mammography-only arm. A second Italian study reported that the mean (SD) invasive tumour size detected by supplementary DBT was 1.52 cm (SD 0.61) in participants who had negative mammograms.

ABUS

A UK study reported the identification of 9 invasive cancers in participants who had negative mammograms.[20] A US study found that ABUS/mammography detected 79 invasive cancers compared to 51 invasive cancers detected by mammography alone.[50] The mean tumour size of detected cancers did not differ between the supplementary imaging and mammography-only groups.

HHUS

A Japanese study reported that 28 invasive cancers were detected in the HHUS/ mammography arm compared to 18 in the mammography-only arm.[55] Among cancers detected in the HHUS group, 93% were <20mm, whereas in the mammography-only group, the corresponding proportion was 76%. A Chinese study detected five cancers of mixed histology (including invasive and high-grade DCIS) in participants who had negative mammograms.[54] All detected cancers were <1.5 cm. An Italian study found that the mean (SD) invasive tumour size detected by supplementary HHUS was 1.51 cm (0.48 cm) in participants with negative mammograms.

MRI

Four studies reported the detection of 4,[32] 14,[39] 32,[20] and 64[37] invasive cancers, respectively, in participants who had been classified as negative for breast cancer following standard screening mammography. The median (IQR) tumour size of cancers detected by supplementary MRI was 1.0 cm (0.8-1.5) and 0.7 cm (0.6-1.0) in two studies that reported these data.

CEM

In one UK study that investigated CEM as a supplementary modality, 32 of 39 cancers detected were invasive (15.7 of 1,000 exams were invasive).[20] The median (IQR) cancer size detected was 1.1 (0.7-.1.5) cm. CEM detected three times as many invasive cancers as ABUS, with a similar performance to that of MRI.

Time taken for additional supplementary imaging

The total time needed to conduct additional imaging, including acquisition and interpretation time, was not consistently reported across the included studies. Only 9 studies reported the time taken to perform and/or interpret supplementary imaging modalities. One study reported that supplementary DBT took 3.7 seconds.[39] Three studies reported that supplementary ABUS took additional i) 10 minutes before data were sent to the workstation,[33] ii) 60 seconds acquisition time, with a total exam time of 15 minutes,[50] and iii) 15 minutes per patient with radiologist interpretation taking 5-7 minutes.[42] Two studies reported that supplementary HHUS took an average additional time of i) 10 minutes[55] and ii) 15-20 minutes, respectively.[54] One study reported that supplementary full-protocol MRI took 9 minutes and 57 seconds.[32] This was mirrored by a second full protocol MRI study, which reported that the entire protocol took less

than ten minutes to perform.[22] A third MRI study reported that the mean (SD) interpretation time for the full protocol MRI was 192 (44) seconds, while the interpretation time for the abbreviated MRI protocol in the same study was 42 (18) seconds.[23]

Acceptance of MRI invitation

The DENSE study[37] reported that only 59% of invitees accepted the invitation for MRI. In the BRAID study, 188 women withdrew from the MRI arm of the study for a variety of reasons, most commonly 'moved out of the area', 'unable to attend', 'health/personal reasons', or 'contraindication to contrast', or women opted to join a different study.[20] In the study by Kaiser et al.,[32] preliminary results were presented for 200 women; however, 80 women were screened but not scheduled for MRI for various reasons, including personal/health issues, language barriers, withdrawal of consent, or unwillingness to receive contrast medium. Neither the BRAID study nor the Kaiser study reported results according to acceptance rates. The other supplementary MRI studies did not mention women declining or failing to attend the MRI screening.[22, 23]

Table 5 Reported tumour characteristics per supplemental imaging modality

Author (year)	Study arm	Cancer stage at detection, n (%)	Invasive tumour size [cm] Mean (SD) (unless otherwise stated)	Cancer histology, n (%)	Nodal involvement, positive or negative, n (%) unless otherwise stated
DBT					
Pattacini (2022)[36]	DBT/DM	NR	NR	SCREEN-DETECTED CANCERS: DCIS: 9 (0.15) Invasive: 40 (0.67) INTERVAL CANCERS: DCIS: 1 (0.02) Invasive: 13 (0.22)	NR
	DM only	NR	NR	SCREEN-DETECTED CANCERS: DCIS: 9 (0.15) Invasive: 40 (0.67) INTERVAL CANCERS: DCIS: 0 Invasive: 15 (0.25)	NR
Tagliafico (2016)[40]	DBT/DM negative	NR	1.52 (0.61)	NR	NR
ABUS					
Gilbert (2024)[20]	ABUS/DM negative	NR		DCIS: 0 Invasive: 9	
Brem (2015)[50]	ABUS/DM	Stage IA or IB: 57 Stage IIA or IIB: 14 Stage IIIA, IIIB, or IIIC: 4 Stage IV: 0 Unknown stage: 4	1.3 (0.78)^	DCIS: 33 Invasive: 79 (IDC: 59; ILC: 15; other invasive type: 5)	Positive: 4
	DM only	Stage IA or IB: 37 Stage IIA or IIB: 9	1.3 (0.79)^	DCIS: 31 Invasive: 51	Positive: 2

		Stage IIIA, IIIB, or IIIC: 2 Stage IV: 0 Unknown: 3		-(IDC: 38; ILC: 1; other invasive type: 2)	
HHUS					
Nakamura (2024)[25]	HHUS/DM	Stage 0: 3 Stage I: 13 Stage II: 4 Stage III: 0 Stage IV: 1 Unknown: 4	NR	NR	NR
	DM only	Stage 0: 2 Stage I: 11 Stage II: 5 Stage III: 0 Stage IV: 0 Unknown: 1	NR	NR	NR
Harada Shoji (2021)[55]	HHUS/DM	Stage 0 and I: 35 (85.4) Stage II or higher: 6 (14.6)	n (%) <10 mm: 11 (39.3) 11-20 mm: 15 (53.6) >20 mm: 2 (7.1)	Non-invasive: 13 (31.7) Invasive: 28 (68.3)	Positive: 5 (17.9) Negative: 23 (82.1)
	DM only	0 and I: 19 (79.2) II or higher: 5 (20.8)	n (%) <10 mm: 9 (50) 11-20 mm: 4 (22.2) >20 mm:4 (22.2)	Non-invasive: 6(25) Invasive: 18 (75)	Positive: 2 (11.1) Negative: 15 (83.3) Data missing: 1 (5.6)
Tagliofico (2016)[40]	HHUS/DM negative	NR	1.51 (0.48)	NR	NR
Chang (2015)[54]	HHUS/DM negative	NR	(5 cancers detected) 1. 1.2 2. 1.0 1. 0.5 2. 0.5 3. 0.8 (ultimately 1.5cm, 1.0cm, 0.7cm, 1.0cm and 1.0cm on pathology, respectively)	(5 cancers detected) 1. invasive mixed ductal and lobular carcinoma, grade II 2. ICD, grade 1, with high grade DCIS 3. IDC, grade II 4. DCIS involving adenomyoepitheliosis 5. DCIS, high grade	NR
MRI					

Kaiser (2024)[32]	fpMRI/DM negative	NR	NR	DCIS: 1 Invasive: 4	NR
Gilbert (2024)[20]	AbMRI/DM negative	NR	NR	DCIS: 5 Invasive: 32	NR
Veenhuizen (2021)[39]	FpMRI/DM	-Stage II to IV: 0 -High-grade (invasive cancers) 6.7%	Median (IQR) 0.7 (0.6, 1)	DCIS: 6 (30) Invasive: 14	Positive: 0
Bakker (2019)[37]	fpMRI/DM negative	NR	NR	(%) DCIS: (19) Invasive: (81)	Positive: 1.9 per 1,000 exams Negative: 14.6 per 1,000 exams
CEM					
Gilbert (2024)[20]	CEM/DM negative	NR	NR	DCIS: 7 Invasive: 32	NR

DBT, digital breast tomosynthesis; ABUS, automated breast ultrasound; HHUS, hand-held ultrasound; MRI, magnetic resonance imaging; CEM, contrast enhanced mammography; SD, standard deviation; NR, not reported; DCIS, ductal carcinoma in situ; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; IQR, inter-quartile range; ^for invasive cancers only

Table 6 Summary of MRI study findings, additional details provided in Appendix 3, Table 4.A and 4.B

Study Name Author, year (country)	Study Question Supplemental screening modality / comparator imaging	Number with dense breasts in study arm/ number with dense breasts in DM only arm ϕ	Screening programme interval (method of assessment of breast density; manual or automated)	CDR per 1,000 exams (lower CI, upper CI)	Recall rate % (lower CI, higher CI) (definition of recall rate^)	PPV1 for recall % (lower CI, upper CI)	Biopsy rate % (lower CI, upper CI)	PPV3 for biopsy % (lower CI, upper CI)	Interval cancers % (lower CI, higher CI)	False negatives % (lower CI, higher CI)	False positives % (lower CI, higher CI)
MA-DETECT Kaiser, 2024[32] (Germany)	Negative DM plus Fp-MRI	200 / 200 (participants underwent both imaging modalities)	NR (Manual)	20 (5.5, 50.4)	40 (17.4, 77.3) (per 1,000 exams) (definition: positive MRI findings)	50.0 (15.7, 84.3)	NR	NR	NR	NR	NR
BRAID Gilbert, 2024[20] (UK)	Negative DM plus Ab-MRI	2,130 / 6,306 (participants underwent both imaging modalities)	Triennial (Manual)	17.4 (12.2, 23.9)	9.7 (8.4, 11) (definition: BI-RADS 3, 4, or 5)	NR	4.9 (4, 5.9)	NR	NR	NR	NR
DENSE (2nd Round) Veenhuizen, 2021[39]	Negative DM plus Fp-MRI	3,436 / 3,436 (participants underwent both imaging modalities)	Biennial (Automated; Volpara)	5.8 (3.8, 9)	32 (26.6, 38.4) (per 1,000 exams)	18.2 (12.1, 26.4)	24.4 (19.8, 30.2)	23.8 (16, 33.9)	NR	NR	2.6 26.3 (21.5, 32.3) per

(Netherlands)					(definition: BI-RADS 3,4,5)						1,000 exams
DENSE (1st Round) Bakker, 2019[37] (Netherlands)	Negative DM plus Fp-MRI	4,783 / 4,783 (participants underwent both imaging modalities)	Biennial (Automated; Volpara)	16.5 (13.3, 20.5)	9.5 94.9 (86.9, 103.6) (per 1,000 exams) (definition: BI-RADS 3,4,5)	17.4 (14.25, 21.2)	6.3 62.7 (56.2, 70) (per 1,000 exams)	26.3 (21.7, 31.6)	n=4 (n=161 in mammography only arm)	NR	8.0 79.8 (72.4, 87.9) per 1,000 exams
Kuhl, 2017[22] (Germany)	Negative DM plus Fp-MRI	1,282 / 1,282 (participants underwent both modalities)	NR (Manual)	n=26 (of 1,282 exams)	NR	NR	NR	NR	n=0	NR	NR
Chen, 2017,[23] (China)	Negative DM plus FP-MRI	478 / 478 (participants underwent MM, FP-MRI and Ab-MRI)	NR (Manual)	n=16 (of 478 exams)	NR	41	NR	NR	NR	NR	NR
	Negative DM plus Ab-MRI	478 / 478 (participants underwent MM, FP-MRI and Ab-MRI	NR (Manual)	n=15 per 478 exams	NR	27.7	NR	NR	NR	NR	NR

^Recall rate definition as reported by study authors

ϕ Number of people with dense breasts in the MRI plus DM study arm, and in the DM only arm (according to BIRADS categories D and C corresponding to heterogeneously and extremely dense breasts, respectively)

MRI, magnetic resonance imaging; Ab-MRI, abbreviated protocol MRI; Fp-MRI full protocol MRI; DENSE, Dense Tissue and Early Breast Neoplasm Screening; BRAID, Breast Screening; Risk Adaptive Imaging for Density; DM digital mammography; PPV1, positive predictive value (number of cancers diagnosed per number of positive screens); PPV3, positive predictive value (biopsy proven predictive value); CI, confidence interval; NR, not reported; CDR, cancer detection rate; BIRADS breast imaging reporting and data system

Discussion of findings

Standard screening mammography misses cancers in people with dense breasts. When considering supplementary modalities, DBT detects the fewest additional cancers in this group, identifying an additional 1-2 cancers per 1,000 screenings. Both ultrasound techniques, ABUS and HHUS, perform similarly, detecting 2-3 additional cancers per 1,000 screenings among participants with dense breasts. MRI is superior to both DBT and ultrasound in the detection of missed cancers, identifying almost 19 additional cancers per 1,000 screenings compared to mammography alone. Taking a conservative view by excluding studies with small sample sizes (<500 participants), MRI identified more than 17 additional cancers per 1,000 screenings. In the UK BRAID randomised study, which focused on participants with dense breasts of otherwise average risk, CEM identified 19 additional cancers per 1,000 screenings and was found to be comparable to MRI in detecting cancers missed by standard mammography. Although CEM has often been studied in participants at elevated risk of breast cancer, the BRAID study, which included participants with 79% heterogeneous and 21% extremely dense breasts in the CEM group, demonstrates promising results for CEM as a supplemental screening modality for participants with dense breasts and average cancer risk.[20]

Whilst our review focused on people with dense breasts and average breast cancer risk, a systematic review published two years ago examined data for participants of both average and intermediate risk who had negative mammography results.[57] That review, which was published in 2023, included 22 studies up to March 2020. The authors concluded that MRI was the most effective supplemental imaging modality for participants with dense breasts at average or intermediate risk for breast cancer. They also highlighted the need for more research to assess the relative effectiveness of other modalities and to determine MRI advantages in terms of mortality rate and cost-effectiveness. In line with the fast-moving nature of this topic, our review included 10 studies published in 2023 and 2024, with 18 studies published since 2020.

Whilst CDR is a key outcome measure, breast screening programmes routinely refer to '*referral to assessment rate*' (RAR), which is equivalent to 'recall rate' and useful to minimise harms like anxiety and interventions associated with false-positive findings.[59] In the UK, a recall rate of <7% for prevalent screenings and <5% for incident screenings are considered to be achievable targets. Performance thresholds of <10% for prevalent screening and <7% for incident screening are deemed to be acceptable levels.[55] Although elevated recall rates are undesirable, higher recall was associated with lower interval cancer rates in the National Health System Breast

Screening Programme (NHSBSP), with 80-84 additional recalls estimated to be required to avoid one interval cancer.[59] Therefore, a minimal threshold for recall may be needed to maximise value for individuals undergoing breast cancer screening.

Our meta-analyses found that recall rates varied across supplementary modalities but were broadly similar for ABUS, HHUS and MRI. This similarity was noted despite the substantially higher CDR per 1000 screenings observed in the MRI meta-analysis, which would typically be expected to translate into a higher recall rate. Moreover, the recall rate was higher in the BRAID[20], Kuhl,[22] and Bakker[37] studies, which focused on prevalent (first-round) screening, compared to the Veenhuizen study[39] which focused on incident screening. As prevalent screenings tend to detect more cancers than subsequent incident screenings, it is plausible that both recall rates and CDRs would decrease over time within the context of a population screening programme. Notably, the pooled recall rate for the MRI meta-analysis as well as that of the three individual studies included in the meta-analysis were below 10%, which is considered an acceptable level.[59] The NHSBSP sets positive predictive value thresholds for referral to assessment (PPV1) as follows:[60]

Prevalent screening = acceptable level >8%, achievable level >12%,

Incident screening = acceptable level >24% and achievable level >33%.

PPV1 represents the proportion of people who are diagnosed with breast cancer among those who had a positive result on screening mammography. According to the NHSBSP criteria, the results of the prevalent screening round of the DENSE study (17.4%, 95% CI 14.25 to 21.2)[37] met the desired level while those of the incident screening round did not (18.2%, 95% CI 12.1 to 26.4).[39] The Kaiser study conducted in Germany reported a favourable PPV1 of 50% (95% CI 15.7, 83.3), although this finding was based on a preliminary report from a larger ongoing trial.[32] It is worth noting that the BRAID study did not report PPV1.[20]

Stark differences in how individual reports define ‘recall’ have likely contributed to the substantial heterogeneity and wide confidence intervals observed in the meta-analysis estimates for each supplementary screening modality. The inclusion of BI-RADS 0, which indicates an ‘inconclusive’ result, may inflate recall rates. However, several studies[48-50, 52] exclusively included BI-RADS 0 in their definition of recall, further complicating direct comparisons. Given that recall rates serve as a key metric for evaluating the success of screening programmes, inconsistencies in definitions across institutions and countries undermine their reliability as a comparative

benchmark. However, within the MRI group, recall definitions were more consistently applied, enabling greater confidence in the pooled estimates for this imaging modality.

Interval rates were not widely reported across the included studies, but two MRI studies reported this parameter.[22, 37] One of these studies reported an interval cancer rate of 2.5 per 1,000 exams in the supplementary MRI group compared to 5.0 per 1,000 in the mammography-only group.[37] Notably, only 59% of participants invited for MRI screening accepted the invitation. Among those who declined, the interval cancer rate was 4.9 per 1,000 screenings (16 cancers) whereas it was significantly lower - 0.8 per 1,000 exams (4 cancers) - among those who underwent supplementary MRI. The second MRI study reported that no interval cancers were detected in more than two years of follow-up in individuals who had supplemental MRI after negative mammography and ultrasound.[22] These findings highlight the potential of supplemental MRI screening for individuals with extremely dense breasts, demonstrating a significantly lower incidence of interval cancers compared to mammography alone.

Sensitivity and specificity values were inconsistently reported across included studies, preventing a meaningful meta-analysis. Across the studies that reported these measures of accuracy, sensitivity was lower in the mammography-only groups, compared to supplementary modalities, while specificity was generally high for all groups.

For our primary outcome, CDR, MRI and CEM demonstrated superior performance compared to DBT, ABUS and HHUS, detecting a greater number of cancers that were missed by mammography alone in people with dense breasts. Given their higher screening yield, the rest of this discussion will focus specifically on MRI and CEM.

Acceptance of MRI invitation

The willingness of women to undergo MRI as a supplementary imaging modality for breast cancer screening is an important consideration that could impact the feasibility of implementing supplementary MRI screening. The DENSE, BRAID and Kaiser studies[37, 39, 20, 32] report that some women did not accept the invitation to attend MRI screening, although the number was far greater in the DENSE study (only 59% accepted the invitation). The reasons for declination were not always due to personal preference or unwillingness/inability to undergo MRI but also included moving out of the area, travel issues, or health issues.

Heterogeneously versus extremely dense breasts

In conducting this review, we were not asked to differentiate between heterogenous and extremely dense breasts. Although the supplementary MRI group meta-analysis showed a high CDR and a relatively low recall rate, the two DENSE studies[37, 39] included *only* participants with extremely dense breasts, introducing an applicability concern. The BRAID study included individuals with both heterogeneously and extremely dense breasts (16.5% extremely dense in the MRI group).[20] Following the findings of the DENSE studies, the EUSOBI, acknowledged that individuals with both heterogeneously and extremely dense breasts are underserved by mammography or DBT alone. Consequently, they recommended that people be informed of their breast density, and that those with extremely dense breasts should be offered MRI screening every 2-4 years.[10] Their guidance cited that robust evidence from the DENSE studies for extremely dense breasts, but noted the lack of such evidence for heterogeneously dense breasts. The UK-based BRAID study results were shared with the authors of this report in draft format via personal communication, but have not yet been published.[20] BRAID is the second randomised, multicentre study evaluating supplementary MRI, corroborating the findings of the DENSE studies[37] by demonstrating comparable detection and recall rates in a population with mixed breast density. Since breast density is a continuum, the BRAID study suggests that the cancer-detecting benefits of supplementary MRI screening extend beyond those with the highest level of breast density.[20] A preliminary report from the German MA-DETECT study was included in our MRI meta-analysis, although it was removed during sensitivity analysis due to an *a priori* decision to conduct sensitivity analysis on studies with fewer than 500 participants. The MA-DETECT RCT includes participants with both heterogeneous and extremely dense breasts and aims to determine the CDR and recall rate among individuals attending the national German screening programme. While the preliminary findings were based on a small cohort of 200 participants, the study identified five cancers in cases where standard mammography had yielded negative results.[32]

MRI protocols

Of the six MRI studies, four[32, 37, 39] used a full MRI protocol, the UK BRAID study[20] used an abbreviated MRI protocol, while participants in a small study from China received either a full MRI protocol or an abbreviated MRI protocol, following negative mammography. Looking at the effect estimates of the individual studies included in our meta-analysis, we did not find that the cancer detection rates differed meaningfully between those that adopted an abbreviated MRI

protocol and those that adopted a full MRI protocol. The feasibility of implementing MRI in clinical practice as a supplementary imaging modality at the population screening level may be constrained by the limited availability of MRI machines, higher costs, and potential shortage of expertise in image acquisition and interpretation. Adopting an abbreviated MRI protocol could improve accessibility by increasing throughput and reducing examination costs while still maintaining comparable diagnostic accuracy, as current evidence suggests.[10] A retrospective study involving 356 individuals with dense breasts and negative screening mammograms found no difference in cancer detection rates between abbreviated and full diagnostic MRI protocols.[61]

Study quality

In studies evaluating supplementary MRI, aside from applicability concerns relating to the recruitment of participants with extremely dense breasts and the nature of the comparator, the most important risk of bias relates to flow and timing. The median interval between index tests is 10 (IQR 8-14) weeks for the Bakker study,[37] 8 (IQR 3-13) weeks for the Veenhuizen study,[39] and 143 (IQR 98-183) days [around 20 (IQR 14-26) weeks] for the BRAID study.[20] The Kaiser study did not report the median interval but excluded those who underwent the index test more than two months apart from the comparator test to prevent interval carcinoma from being misinterpreted in favour of breast MRI.[32] The Chen study did not report the interval between MRI and mammography.[23] The Kuhl study reported a median interval of 5 days (range of 1-28 days) between MRI and comparator tests.[22] While delays between negative mammography and subsequent MRI may be attributed to logistical challenges such as MRI access, availability, cost, and expertise, similar intervals were observed for other imaging modalities in the BRAID study. Specifically, the reported median interval was 16 weeks (IQR 11, 21) for ABUS and 19 weeks (IQR 13, 25) for CEM.[20]

Need for additional evidence

CEM was found to have a similar screening performance to MRI, but only in a single UK study, which leaves some uncertainty as to the generalisability of this finding. Additional evidence would reduce the uncertainty about the ability of supplementary CEM to detect additional cancers compared to mammography alone in participants with dense breasts undergoing screening. In particular, additional evidence regarding the potential effectiveness of CEM compared to MRI would be welcome. There are several studies in process, including the CMIST[62] (primary outcome data complete in early 2025) and C-MERIT[63] studies (primary outcome data complete at the end of 2026), both of which are investigating supplementary CEM in participants with

dense breasts in a screening context, although neither compares its effectiveness directly with MRI.

Limitations and strengths

Our analysis included both retrospective studies and studies with randomised or paired designs, which could have introduced confounding factors associated with selection bias. However, to mitigate this risk, we excluded any studies that appeared to include a 'pre-selected' population, such as participants with characteristics linked to a higher cancer risk (e.g., family history, previous breast cancer, genetic risk). Our strict eligibility criteria ensured the inclusion of individuals at average cancer risk (apart from dense breasts) while excluding studies involving individuals at intermediate or high risk. While this approach reduced the number of studies retrieved - thereby affecting the statistical power of our review - it provided a more accurate representation of a typical breast screening population. Moreover, our commissioned review question referred broadly to 'dense' breasts rather than distinguishing between heterogeneously and extremely dense categories. As a result, we did not specifically assess whether supplementary imaging modalities performed better in the densest breast subgroups. Furthermore, we included studies examining both participants with negative DM findings and those assigned to mammography alone. Consequently, our combined estimates reflect differences in outcomes rather than absolute values. Finally, given the fast-evolving nature of this research area, we conducted additional searches up to November 2024, including unpublished reports. This allowed us to include the most up-to-date evidence, including a draft version of the UK BRAID study, which has been recently completed but was still unpublished at the time of this report's preparation.

Discussion and Review Summary

Our findings confirm that standard mammography alone is insufficient in detecting breast cancer in some individuals with dense breasts. While adjunctive imaging modalities such as ABUS, HHUS, and, to a lesser extent, DBT offer incremental improvement, their ability to identify missed cancers remains limited. Our meta-analysis combined five randomised studies (representing six comparisons) and demonstrated that supplementary MRI considerably enhances cancer detection rates.[20, 32, 37, 39]

The UK-based BRAID study[20] provides additional robust evidence, reinforcing the findings of the Netherlands-based DENSE studies,[37, 39] and the German study by Kuhl et al. Two smaller studies on supplementary MRI also provided similar findings.[32, 23] It is worth pointing out that detection, recall, and biopsy rates may decrease in the incident screening rounds, as indicated by the second phase of the DENSE study.[39] The DENSE study also showed that significantly fewer interval cancers were subsequently detected in the supplemental MRI group compared to the mammography-only group,[37] with the Kuhl study reporting no interval cancers with at least 2 years of follow-up in those who had supplemental MRI screening.[22] The interval between negative mammography and supplementary MRI in some studies could introduce a potential risk of bias in these results. CEM has shown a screening performance comparable to MRI, albeit in a recent single study.[20] Overall, these findings highlight supplementary screening modalities that have the potential to improve cancer detection rates for individuals with dense breasts, supporting a more refined and personalised approach to population-level breast screening.

Appendix 1 — Search strategy

Search Terms

Search terms included combinations of free text and subject headings (Medical Subject Headings [MeSH] for MEDLINE, and Emtree terms for Embase), grouped into the following categories. Results were imported into Endnote and duplicates removed.

Table 1 Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase

Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions
1. Mass Screening/ or Early Detection of Cancer/
2. Breast/ or breast?.tw.
3. 1 and 2
4. (breast adj3 (screen* or imag*)).tw,kf.
5. 3 or 4
6. breast density/
7. ((breast? or mammog*) adj5 dens*).tw,kf.
8. 6 or 7
9. *Mammography/ or mammogra*.tw,kf.
10. *Magnetic Resonance Imaging/ or ("magnetic resonance imaging" or MRI).tw,kf.
11. *Ultrasonography, Mammary/ or (sonogra* or ultrasound* or ultrasonogra* or echomammogra* or ABUS or HHUS).tw,kf.
12. (("contrast-enhanced" adj3 mammogra*) or CEM).tw,kf.
13. (tomosynthesis or "3D mammogra*" or "3-D mammogra*" or "digital breast tomogra*" or DBT).tw,kf.
14. ((supplement* or enhance* or adjunct* or addit* or "risk-adapted" or "risk adapted") adj5 (screen* or imag*)).tw,kf.
15. or/10-14
16. 5 and 8 and 9 and 15
17. (case reports or comment or editorial or letter or news).pt.
18. 16 not 17
19. limit 18 to yr="2014 -Current"

Appendix 2 — Included and excluded studies

PRISMA flowchart

Figure 1 summarises the volume of publications included and excluded at each stage of the review.

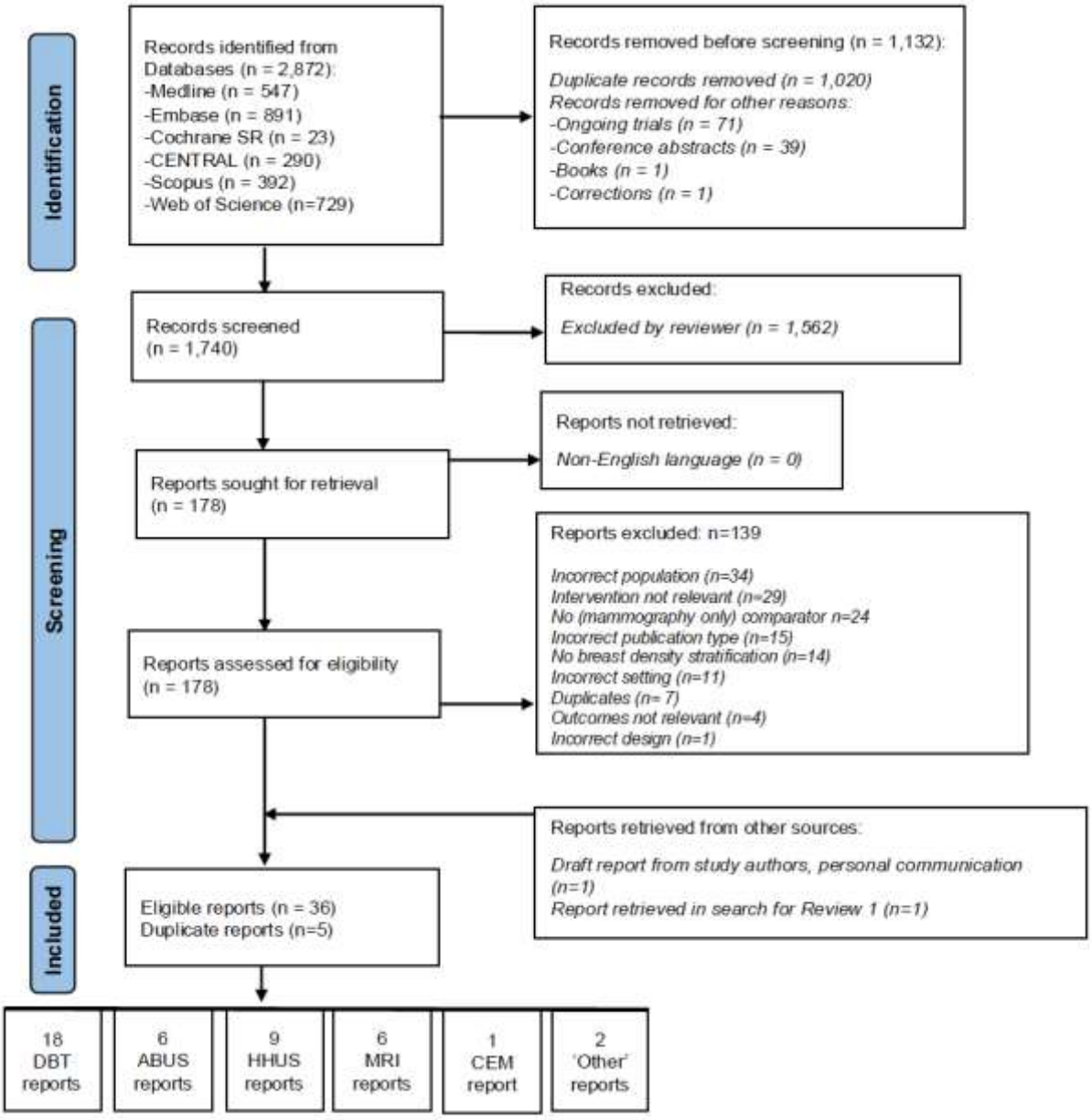


Figure 1 PRISMA flowchart showing summary of publications included and excluded at each stage of the review [14]

Publications included after review of full text articles

Of the 178 publications included after the review of titles and abstracts, 34 were eligible for inclusion. An additional one report was retrieved in draft format, and one report was retrieved during the Review 1 and deemed eligible for Review 2, making 36 included studies in total. A further five studies were eligible for inclusion but contained duplicate reporting and were therefore not included in the final total. Details of the included reports are listed below:

Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. N Engl J Med. 2019;**381**(22):2091-102.

Ban K, Tsunoda H, Togashi S, et al. Breast cancer screening using digital breast tomosynthesis compared to digital mammography alone for Japanese women. Breast Cancer. 2021;**28**(2):459-64.

Bernardi D, Macaskill P, Pellegrini M, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. Lancet Oncol. 2016;**17**(8):1105-13.

Brem RF, Tabar L, Duffy SW, et al. Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue: the SomoInsight Study. Radiology. 2015;**274**(3):663-73.

Buchberger W, Geiger-Gritsch S, Knapp R, Gautsch K, Oberaigner W. Combined screening with mammography and ultrasound in a population-based screening program. Eur J Radiol. 2018;**101**:24-9.

Chang JM, Koo HR, Moon WK. Radiologist-performed hand-held ultrasound screening at average risk of breast cancer: results from a single health screening center. Acta Radiol. 2015;**56**(6):652-8.

Chen, Shuang-Qing et al. Application of Abbreviated Protocol of Magnetic Resonance Imaging for Breast Cancer Screening in Dense Breast Tissue. Academic Radiology, Volume 24, Issue 3, 316 – 320

Conant EF, Beaber EF, Sprague BL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography compared to digital mammography alone: a cohort study within the PROSPR consortium. Breast Cancer Res Treat. 2016;**156**(1):109-16.

Durand MA, Friedewald SM, Plecha DM, et al. False-negative rates of breast cancer screening with and without digital breast tomosynthesis. Radiology. 2021;**298**(2):296-305.

Freer PE, Riegert J, Eisenmenger L, et al. Clinical implementation of synthesized mammography with digital breast tomosynthesis in a routine clinical practice. Breast Cancer Res Treat. 2017;**166**(2):501-9.

Gatta G, Cappabianca S, La Forgia D, et al. Second-generation 3D automated breast ultrasonography (Prone ABUS) for dense breast cancer screening integrated to mammography: effectiveness, performance and detection rates. J Pers Med. 2021;**11**(9).

Gilbert FJ. Breast Screening - Risk Adaptive Imaging for Density. BRAID trial results (personal communication: Brazzelli, M. 2024). 2024.

Ha SM, Jang M-J, Youn I, et al. Screening outcomes of mammography with AI in dense breasts: a comparative study with supplemental screening US. Radiology. 2024;**312**(1):e233391.

Harada-Shoji N, Suzuki A, Ishida T, et al. Evaluation of adjunctive ultrasonography for breast cancer detection among women aged 40-49 years with varying breast density undergoing screening mammography: a secondary analysis of a randomized clinical trial. JAMA Netw Open. 2021;**4**(8):e2121505.

Johnson K, Zackrisson S, Rosso A, et al. Tumor characteristics and molecular subtypes in breast cancer screening with digital breast tomosynthesis: the Malmö Breast Tomosynthesis Screening Trial. Radiology. 2019;**293**(2):273-81.

Kaiser C, Wilhelm T, Walter S, Singer S, Keller E, Baltzer PAT. Cancer detection rate of breast-MR in supplemental screening after negative mammography in women with dense breasts. Preliminary results of the MA-DETECT-Study after 200 participants. Eur J Radiol. 2024;**176**:111476.

Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental Breast MR Imaging Screening of Women with Average Risk of Breast Cancer. Radiology. 2017 May;283(2):361-370.

Kwon M-R, Choi JS, Lee MY, et al. Screening outcomes of supplemental automated breast US in Asian women with dense and nondense breasts. Radiology. 2023;**307**(4):e222435.

Lee SE, Yoon JH, Son N-H, Han K, Moon HJ. Screening in patients with dense breasts: comparison of mammography, artificial intelligence, and supplementary ultrasound. AJR Am J Roentgenol. 2024;**222**(1):e2329655.

McDonald ES, McCarthy AM, Akhtar AL, Synnestvedt MB, Schnall M, Conant EF. Baseline screening mammography: performance of full-field digital mammography versus digital breast tomosynthesis. AJR Am J Roentgenol. 2015;**205**(5):1143-8.

McDonald ES, Oustimov A, Weinstein SP, Synnestvedt MB, Schnall M, Conant EF. Effectiveness of digital breast tomosynthesis compared with digital mammography: outcomes analysis from 3 years of breast cancer screening. JAMA oncology. 2016;**2**(6):737-43.

Nakamura A, Ohnuki K, Takahashi H, et al. The effects of breast density on the benefits of mammograms with adjunctive ultrasonography in breast screening. Breast Cancer. 2024;**31**(2):228-33.

Olinder J, Johnson K, Akesson A, Fornvik D, Zackrisson S. Impact of breast density on diagnostic accuracy in digital breast tomosynthesis versus digital mammography: results from a European screening trial. Breast Cancer Res. 2023;**25**(1):116.

Osteras BH, Martinsen ACT, Gullien R, Skaane P. Digital mammography versus breast tomosynthesis: impact of breast density on diagnostic performance in population-based screening. Radiology. 2019;**293**(1):60-8.

Pang JX, Newsome J, Sun M, et al. Impact of switching from digital mammography to tomosynthesis plus digital mammography on breast cancer screening in Alberta, Canada. J Med Screen. 2022;**29**(1):38-43.

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;**303**(2):256-66.

Pulido-Carmona C, Romero-Martín S, Raya-Povedano JL, et al. Interval cancer in the Córdoba Breast Tomosynthesis Screening Trial (CBTST): comparison of digital breast tomosynthesis plus digital mammography to digital mammography alone. Eur Radiol. 2024.

Rani VR, Reddy BN. Diagnostic performance in detecting breast cancer on combined mammography and ultrasound. J Cardiovasc Dis Res. 2023;**14**(5):1649-52.

Ren W, Yan H, Zhao X, et al. Integration of handheld ultrasound or automated breast ultrasound among women with negative mammographic screening findings: a multi-center population-based study in China. Acad Radiol. 2023;**30**(Supplement 2):S114-S26.

Rose SL, Shisler JL. Tomosynthesis impact on breast cancer screening in patients younger than 50 years old. AJR Am J Roentgenol. 2018;**210**(6):1401-4.

Starikov A, Drotman M, Hentel K, Katzen J, Min RJ, Arleo EK. 2D mammography, digital breast tomosynthesis, and ultrasound: which should be used for the different breast densities in breast cancer screening? Clin Imaging. 2016;**40**(1):68-71.

Stepanek T, Constantinou N, Marshall H, et al. Changes in the utilization of the BI-RADS Category 3 assessment in recalled patients before and after the implementation of screening digital breast tomosynthesis. Acad Radiol. 2019;**26**(11):1515-25.

Tagliafico AS, Calabrese M, Mariscotti G, et al. Adjunct screening with tomosynthesis or ultrasound in women with mammography-negative dense breasts: Interim report of a prospective comparative trial. J Clin Oncol. 2016;**34**(16):1882-8.

Upadhyay N, Soneji N, Stewart V, Ralleigh G. The effect of the addition of tomosynthesis to digital mammography on reader recall rate and reader confidence in the UK prevalent screening round. Clin Radiol. 2018;**73**(8):744-9.

Veenhuizen SGA, de Lange SV, Bakker MF, et al. Supplemental breast MRI for women with extremely dense breasts: results of the second screening round of the DENSE trial. Radiology. 2021;**299**(2):278-86.

Wilczek B, Wilczek HE, Rasouliyan L, Leifland K. Adding 3D automated breast ultrasound to mammography screening in women with heterogeneously and extremely dense breasts: report from a hospital-based, high-volume, single-center breast cancer screening program. Eur J Radiol. 2016;**85**(9):1554-63.

Publications excluded after review of full text articles

Of the 176 publications included after the review of titles and abstracts, 139 were ultimately judged not to be relevant to this review. Among there, there were two sets of duplicate studies, therefore 137 excluded publications are list below, along with main reasons for exclusion.

Incorrect population

Boca I, Ciurea AI, Vesa SC, Ciortea CA, Dudea SM, Manole S. Associating Automated Breast Ultrasound (ABUS) and Digital Breast Tomosynthesis (DBT) with Full-Field Digital Mammography (FFDM) in Clinical Practice in Cases of Women with Dense Breast Tissue. Diagnostics. 2022;12(2):459.

Brunetti N, De Giorgis S, Tosto S, et al. A Prospective Comparative Evaluation of Handheld Ultrasound Examination (HHUS) or Automated Ultrasound Examination (ABVS) in Women with Dense Breast. Diagnostics. 2022;12(9):2170.

Chan YS, Hung WK, Yuen LW, Chan HYY, Chu CWW, Cheung PSY. Comparison of Characteristics of Breast Cancer Detected through Different Imaging Modalities in a Large Cohort of Hong Kong Chinese Women: Implication of Imaging Choice on Upcoming Local Screening Program. Breast Journal. 2022;2022:3882936.

Cho KR, Seo BK, Woo OH, et al. Breast cancer detection in a screening population: Comparison of digital mammography, computer-aided detection applied to digital mammography and breast ultrasound. J Breast Cancer. 2016;19(3):316-23.

Choudhary S, Axmacher J, Conners A, Geske J, Brandt K. Masses in the era of screening tomosynthesis: Is diagnostic ultrasound sufficient? Br J Radiol. 2019;92(1096):20180801.

Chudgar AV, Conant EF, Weinstein SP, et al. Assessment of disease extent on contrast-enhanced MRI in breast cancer detected at digital breast tomosynthesis versus digital mammography alone. Clin Radiol. 2017;72(7):573-9.

Elder K, Matheson J, Nickson C, et al. Contrast enhanced mammography in breast cancer surveillance. Breast Cancer Res Treat. 2023;199(2):221-30.

Fallenberg EM, Dromain C, Diekmann F, et al. Contrast-enhanced spectral mammography: Does mammography provide additional clinical benefits or can some radiation exposure be avoided? Breast Cancer Res Treat. 2014;146(2):371-81.

Fischer U, Luftner-Nagel S, Baum F, Marten-Engelke K, Wienbeck S. The Value of Quality-Assured Magnetic Resonance Imaging of the Breast for the Early Detection of Breast Cancer in Asymptomatic Women. J Comput Assist Tomogr. 2018;42(1):1-5.

Fonseca MM, Alhassan T, Nisha Y, et al. Randomized trial of surveillance with abbreviated MRI in women with a personal history of breast cancer- impact on patient anxiety and cancer detection. BMC Cancer. 2022;22(1):774.

Gao F, Wu T, Li J, et al. SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis. Comput Med Imaging Graph. 2018;70:53-62.

Gareth ED, Nisha K, Yit L, et al. MRI breast screening in high-risk women: Cancer detection and survival analysis. Breast Cancer Res Treat. 2014;145(3):663-72.

Giger ML, Inciardi MF, Edwards A, et al. Automated breast ultrasound in breast cancer screening of women with dense breasts: Reader study of mammography-negative and mammography-positive cancers. AJR Am J Roentgenol. 2016;206(6):1341-50.

Gilbert FJ, Tucker L, Gillan MG, et al. Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial). Radiology. 2015;277(3):697-706.

Kim S-Y, Cho N, Kim SY, et al. Supplemental Breast US Screening in Women with a Personal History of Breast Cancer: A Matched Cohort Study. Radiology. 2020;295(1):54-63.

Li J, Zhang H, Jiang H, et al. Diagnostic Performance of Digital Breast Tomosynthesis for Breast Suspicious Calcifications From Various Populations: A Comparison With Full-field Digital Mammography. Comput Struct Biotechnol J. 2019;17:82-9.

Mansour S, Adel L, Mokhtar O, Omar OS. Comparative study between breast tomosynthesis and classic digital mammography in the evaluation of different breast lesions. Egyptian Journal of Radiology and Nuclear Medicine. 2014;45(3):1053-61.

Maxwell AJ, Michell M, Lim YY, et al. A randomised trial of screening with digital breast tomosynthesis plus conventional digital 2D mammography versus 2D mammography alone in younger higher risk women. Eur J Radiol. 2017;94:133-9.

Moshina N, Aase HS, Danielsen AS, et al. Comparing Screening Outcomes for Digital Breast Tomosynthesis and Digital Mammography by Automated Breast Density in a Randomized Controlled Trial: results from the To-Be Trial. Radiology. 2022;303(1):E23.

Nissan N, Comstock CE, Sevilimedu V, et al. Diagnostic Accuracy of Screening Contrast-enhanced Mammography for Women with Extremely Dense Breasts at Increased Risk of Breast Cancer. Radiology. 2024;313(1):e232580.

Park JH, Lim JH, Kim S, Heo J. A Multi-label Artificial Intelligence Approach for Improving Breast Cancer Detection With Mammographic Image Analysis. In Vivo. 2024;38(6):2864-72.

Patel BK, Carnahan MB, Northfelt D, et al. Prospective Study of Supplemental Screening With Contrast-Enhanced Mammography in Women With Elevated Risk of Breast Cancer: Results of the Prevalence Round. J Clin Oncol. 2024:JCO2202819.

Powell JL, Hawley JR, Lipari AM, Yildiz VO, Erdal BS, Carkaci S. Impact of the Addition of Digital Breast Tomosynthesis (DBT) to Standard 2D Digital Screening Mammography on the Rates of Patient Recall, Cancer Detection, and Recommendations for Short-term Follow-up. Acad Radiol. 2017;24(3):302-7.

Rhodes DJ, Hruska CB, Conners AL, et al., Molecular breast imaging at reduced radiation dose for supplemental screening in mammographically dense breasts. AJR Am J Roentgenol; 2015.

Shen S, Zhou Y, Xu Y, et al. A multi-centre randomised trial comparing ultrasound vs mammography for screening breast cancer in high-risk Chinese women. Br J Cancer. 2015;112(6):998-1004.

Singla D, Chaturvedi A, Aggarwal A, Rao S, Hazarika D, Mahawar V. Comparing the diagnostic efficacy of full field digital mammography with digital breast tomosynthesis using BIRADS score in a tertiary cancer care hospital. Indian J Radiol Imaging. 2018;28(1):115-22.

Song SE, Cho N, Chang JM, Chu AJ, Yi A, Moon WK. Diagnostic performances of supplemental breast ultrasound screening in women with personal history of breast cancer. Acta Radiol. 2018;59(5):533-9.

Sorin V, Yagil Y, Yosepovich A, et al. Contrast-enhanced spectral mammography in women with intermediate breast cancer risk and dense breasts. AJR Am J Roentgenol. 2018;211(5):W267-W74.

Sudhir R, Sannapareddy K, Potlapalli A, Krishnamurthy PB, Buddha S, Koppula V. Diagnostic accuracy of contrast-enhanced digital mammography in breast cancer detection in comparison to tomosynthesis, synthetic 2D mammography and tomosynthesis combined with ultrasound in women with dense breast. Br J Radiol. 2021;94(1118):20201046.

Sung JS, Stamler S, Brooks J, et al. Breast cancers detected at screening mr imaging and mammography in patients at high risk: Method of detection reflects tumor histopathologic results. Radiology. 2016;280(3):716-22.

Upneja A, Long JB, Aminawung JA, et al. Comparative Effectiveness of Digital Breast Tomosynthesis and Mammography in Older Women. J Gen Intern Med. 2022;37(8):1870-6.

Yamashita MW, Larsen LH, Perez J, Edwards A, Papaioannou J, Jiang YL. Comparison of Mammography and Mammography with Supplemental Whole-Breast US Tomography for Cancer Detection in Patients with Dense Breasts. Radiology. 2024;311(3).

Yu X, Hu G, Zhang Z, et al. Retrospective and comparative analysis of 99mTc-Sestamibi breast specific gamma imaging versus mammography, ultrasound, and magnetic resonance imaging for the detection of breast cancer in Chinese women. BMC Cancer. 2016;16(1):450.

Zhu H, Polat D, Evans P, et al. Is There a Difference in the Diagnostic Outcomes of Calcifications Initially Identified on Synthetic Tomosynthesis Versus Full-Field Digital Mammography Screening? Eur J Radiol. 2020;133:109365.

Intervention not relevant

Alsheik NH, Dabbous F, Pohlman SK, et al. Comparison of Resource Utilization and Clinical Outcomes Following Screening with Digital Breast Tomosynthesis Versus Digital Mammography: Findings From a Learning Health System. Acad Radiol. 2019;26(5):597-605.

Brentnall AR, Cuzick J, Buist DSM, Bowles EJA. Long-Term accuracy of breast cancer risk assessment combining classic risk factors and breast density. JAMA Oncology. 2018;4(9):e180174.

Caumo F, Zorzi M, Brunelli S, et al. Digital breast tomosynthesis with synthesized two-dimensional images versus full-field digital mammography for population screening: Outcomes from the Verona screening program. Radiology. 2018;287(1):37-46.

Dibble EH, Singer TM, Jimoh N, Baird GL, Lourenco P. Dense breast ultrasound screening after digital mammography versus after digital breast tomosynthesis. AJR Am J Roentgenol. 2019;213(6):1397-402.

Giess CS, Pourjabbar S, Ip IK, Lacson R, Alper E, Khorasani R. Comparing diagnostic performance of digital breast tomosynthesis and full-field digital mammography in a hybrid screening environment. AJR Am J Roentgenol. 2017;209(4):929-34.

Ha SM, Yi A, Yim D, et al. Digital Breast Tomosynthesis Plus Ultrasound Versus Digital Mammography Plus Ultrasound for Screening Breast Cancer in Women With Dense Breasts. Korean J Radiol. 2023;24(4):274-83.

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(5):601-11.

Ho TQH, Bissell MCS, Kerlikowske K, et al. Cumulative Probability of False-Positive Results after 10 Years of Screening with Digital Breast Tomosynthesis vs Digital Mammography. JAMA Network Open. 2022;5(3):E222440.

Hofvind S, Moshina N, Holen Å, et al. Interval and Subsequent Round Breast Cancer in a Randomized Controlled Trial Comparing Digital Breast Tomosynthesis and Digital Mammography Screening. Radiology. 2021;300(1):66-76.

Jia M, Lin X, Zhou X, et al. Diagnostic performance of automated breast ultrasound and handheld ultrasound in women with dense breasts. Breast Cancer Res Treat. 2020;181(3):589-97.

Kerlikowske K, Su YR, Sprague BL, et al. Association of Screening With Digital Breast Tomosynthesis vs Digital Mammography With Risk of Interval Invasive and Advanced Breast Cancer. JAMA. 2022;327(22):2220-30.

Khanani S, Hruska C, Lazar A, Hoernig M, Hebecker A, Obuchowski N. Performance of Wide-Angle Tomosynthesis with Synthetic Mammography in Comparison to Full Field Digital Mammography. Acad Radiol. 2023;30(1):3-13.

Kim SY, Kim MJ, Moon HJ, Yoon JH, Kim EK. Application of the downgrade criteria to supplemental screening ultrasound for women with negative mammography but dense breasts. Medicine (United States). 2016;95(44):e5279.

Kim WH, Chang JM, Lee J, et al. Diagnostic performance of tomosynthesis and breast ultrasonography in women with dense breasts: a prospective comparison study. Breast Cancer Res Treat. 2017;162(1):85-94.

Lowry KP, Coley RY, Miglioretti DL, et al. Screening Performance of Digital Breast Tomosynthesis vs Digital Mammography in Community Practice by Patient Age, Screening Round, and Breast Density. JAMA network open. 2020;3(7):e2011792.

McCarthy AM, Kontos D, Synnestvedt M, et al. Screening outcomes following implementation of digital breast tomosynthesis in a general-population screening program. J Natl Cancer Inst. 2014;106(11).

Miglioretti DL, Abraham L, Lee CI, et al. Digital Breast Tomosynthesis: Radiologist Learning Curve. Radiology. 2019;291(1):34-42.

Moshina N, Aase HS, Danielsen AS, et al. Comparing Screening Outcomes for Digital Breast Tomosynthesis and Digital Mammography by Automated Breast Density in a Randomized Controlled Trial: results from the To-Be Trial. Radiology. 2020;297(3):522-31.

Plimpton SR, Milch H, Sears C, et al. External Validation of a Commercial Artificial Intelligence Algorithm on a Diverse Population for Detection of False Negative Breast Cancers. J Breast Imaging. 2024.

Sharpe RE, Venkataraman S, Phillips J, et al. Increased cancer detection rate and variations in the recall rate resulting from implementation of 3D digital breast tomosynthesis into a population-based screening program. Radiology. 2016;278(3):698-706.

Strahle DA, Pathak DR, Sierra A, Saha S, Strahle C, Devisetty K. Systematic development of an abbreviated protocol for screening breast magnetic resonance imaging. Breast Cancer Res Treat. 2017;162(2):283-95.

Verburg E, Van Gils CH, Van Der Velden BHM, et al. Deep Learning for Automated Triaging of 4581 Breast MRI Examinations from the DENSE Trial. Radiology. 2022;302(1):29-36.

Verburg E, Van Gils CH, Van Der Velden BHM, et al. Validation of Combined Deep Learning Triaging and Computer-Aided Diagnosis in 2901 Breast MRI Examinations from the Second Screening Round of the Dense Tissue and Early Breast Neoplasm Screening Trial. Invest Radiol. 2023;58(4):293-8.

Vourtsis A, Kachulis A. The performance of 3D ABUS versus HHUS in the visualisation and BI-RADS characterisation of breast lesions in a large cohort of 1,886 women. Eur Radiol. 2018;28(2):592-601.

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.

Weigel S, Heindel W, Hense HW, Decker T, Gerß J, Kerschke L. Breast Density and Breast Cancer Screening with Digital Breast Tomosynthesis: a TOSYMA Trial Subanalysis. Radiology. 2023;306(2):e221006.

Weigel S, Hense HW, Weyer-Elberich V, Gerss J, Heindel W. Breast cancer screening with digital breast tomosynthesis: is independent double reading still required? RöFo. 2023.

Weyer-Elberich V, Kerschke L, Nekolla E, et al. Radiation exposure and screening yield by digital breast tomosynthesis compared to mammography: results of the TOSYMA Trial breast density related. Eur Radiol. 2024.

Xin Y, Zhang X, Yang Y, et al. A multicenter, hospital-based and non-inferiority study for diagnostic efficacy of automated whole breast ultrasound for breast cancer in China. Sci Rep. 2021;11(1):13902.

No (mammography only) comparator

Chough DM, Berg WA, Bandos AI, et al. A prospective study of automated breast ultrasound screening of women with dense breasts in a digital breast tomosynthesis-based practice. J Breast Imaging. 2020;2(1):125-33.

Comstock CE, Gatsonis C, Newstead GM, et al. Comparison of Abbreviated Breast MRI vs Digital Breast Tomosynthesis for Breast Cancer Detection Among Women With Dense Breasts Undergoing Screening. JAMA. 2020;323(8):746-56.

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 Oncology. 2019;5(5):635-42.

Conant EF, Talley MM, Parghi CR, et al. Mammographic Screening in Routine Practice: Multisite Study of Digital Breast Tomosynthesis and Digital Mammography Screenings. Radiology. 2023;307(3):e221571.

Destounis S, Arieno A, Santacroce A. Comparison of Cancers Detected by Screening Breast Ultrasound and Digital Breast Tomosynthesis. Acad Radiol. 2022;29(3):339-47.

Dibble EH, Singer TM, Baird GL, Lourenco AP. BI-RADS 3 on dense breast screening ultrasound after digital mammography versus digital breast tomosynthesis. Clin Imaging. 2021;80:315-21.

Dou E, Ksepka M, Dodelzon K, Shingala PY, Katzen JT. Assessing the Positive Predictive Value of Architectural Distortion Detected with Digital Breast Tomosynthesis in BI-RADS 4 Cases. J Breast Imaging. 2020;2(6):552-60.

Edmonds CE, Weinstein SP, McDonald ES, et al. Abbreviated Breast MRI for Supplemental Screening in Patients With Dense Breasts: Comparison of Baseline Versus Subsequent-Round Examinations. AJR Am J Roentgenol. 2024;223(1).

Gouda W, Yasin R, Yasin MI, Omar S. Automated breast ultrasound in breast cancer screening of mammographically dense breasts: added values. Egyptian Journal of Radiology and Nuclear Medicine. 2024;55(1):86.

Houssami N, Lockie D, Giles M, Noguchi N, Marr G, Marinovich ML. Two-year follow-up of participants in the BreastScreen Victoria pilot trial of tomosynthesis versus mammography: breast density-stratified screening outcomes. Br J Radiol. 2023;96(1148):20230081.

Kerlikowske K, Zhu W, Su Y-R, et al. Supplemental magnetic resonance imaging plus mammography compared with magnetic resonance imaging or mammography by extent of breast density. J Natl Cancer Inst. 2024;116(2):249-57.

Klein Wolterink F, Ab Mumin N, Appelman L, et al. Diagnostic performance of 3D automated breast ultrasound (3D-ABUS) in a clinical screening setting-a retrospective study. Eur Radiol. 2024.

Klevos GA, Collado-Mesa F, Net JM, Yepes MM. Utility of supplemental screening with breast ultrasound in asymptomatic women with dense breast tissue who are not at high risk for breast cancer. Indian J Radiol Imaging. 2017;27(1):52-8.

Korpraphong P, Limsuwarn P, Tangcharoensathien W, Ansusingha T, Thephamongkhol K, Chuthapisith S. Improving breast cancer detection using ultrasonography in asymptomatic women with non-fatty breast density. Acta Radiol. 2014;55(8):903-8.

Marshall H, Burkard-Mandel L, Hsu J, Durieux J, Shikhman R, Plecha D. Abbreviated Breast MRI: Our Two-Year Initial Experience. J Breast Imaging. 2023;5(3):318-28.

Mesurolle B, El Khoury M, Travade A, Bagard C, Petrou A, Monghal C. Is there any added value to substitute the 2D digital MLO projection for a MLO tomosynthesis projection and its synthetic view when a 2D standard digital mammography is used in a one-stop-shop immediate reading mammography screening? Eur Radiol. 2021;31(12):9529-39.

Moffa G, Galati F, Maroncelli R, et al. Diagnostic Performance of Contrast-Enhanced Digital Mammography versus Conventional Imaging in Women with Dense Breasts. Diagnostics. 2023;13(15):2520.

Paul S, McPeek C, Lieberenz J, et al. Supplemental Automated Breast US Screening in Patients With Dense Breasts: 5-Year Experience From an Academic Medical Center. J Breast Imaging. 2024.

Tagliafico AS, Mariscotti G, Valdora F, et al. A prospective comparative trial of adjunct screening with tomosynthesis or ultrasound in women with mammography-negative dense breasts (ASTOUND-2). Eur J Cancer. 2018;104:39-46.

Tutar B, Esen Icten G, Guldogan N, et al. Comparison of automated versus hand-held breast US in supplemental screening in asymptomatic women with dense breasts: is there a difference regarding woman preference, lesion detection and lesion characterization? Arch Gynecol Obstet. 2020;301(5):1257-65.

Wu T, Warren LJ. The Added Value of Supplemental Breast Ultrasound Screening for Women With Dense Breasts: A Single Center Canadian Experience. Can Assoc Radiol J. 2022;73(1):101-6.

Yi A, Jang MJ, Yim D, Kwon BR, Shin SU, Chang JM. Addition of screening breast US to digital mammography and digital breast tomosynthesis for breast cancer screening in women at average risk. Radiology. 2021;298(3):568-75.

Yuge S, Miyake KK, Ishimori T, et al. Performance of dedicated breast PET in breast cancer screening: comparison with digital mammography plus digital breast tomosynthesis and ultrasound. Ann Nucl Med. 2023;37(9):479-93.

Zuckerman SP, Sprague BL, Weaver DL, Herschorn SD, Conant EF. Multicenter evaluation of breast cancer screening with digital breast tomosynthesis in combination with synthetic versus digital mammography. Radiology. 2020;297(3):545-53.

Incorrect publication type

Bae MS, Seo BK. Breast Cancer Screening with Digital Breast Tomosynthesis Improves Performance of Mammography Screening. Radiology. 2023;307(3).

Fujita T. Adjunctive ultrasonography for breast cancer screening. Lancet. 2016;387(10036):2380-1.

Gao Y, Reig B, Heacock L, Bennett DL, Heller SL, Moy L. Magnetic Resonance Imaging in Screening of Breast Cancer. Radiol Clin North Am. 2021;59(1):85-98.

Glechner A, Wagner G, Mitus JW, et al. Mammography in combination with breast ultrasonography versus mammography for breast cancer screening in women at average risk. Cochrane Database Syst Rev. 2023;3:CD009632.

Lee AY, Ray KM. Surrogate Clinical End Points for Breast Cancer Screening Using Digital Breast Tomosynthesis. Radiology. 2022;303(2):267-8.

Li T, Houssami N, Noguchi N, Zeng A, Marinovich ML. Differential detection by breast density for digital breast tomosynthesis versus digital mammography population screening: a systematic review and meta-analysis. Br J Cancer. 2022;127(1):116-25.

Libesman S, Zackrisson S, Hofvind S, et al. An individual participant data meta-analysis of breast cancer detection and recall rates for digital breast tomosynthesis versus digital mammography population screening. Clin Breast Cancer. 2022;22(5):e647-e54.

Lobig F, Caleyachetty A, Forrester L, et al. Performance of Supplemental Imaging Modalities for Breast Cancer in Women With Dense Breasts: Findings From an Umbrella Review and Primary Studies Analysis. Clin Breast Cancer. 2023;23(5):478-90.

Moy L, Gao Y. Supplemental MRI in Extremely Dense Breasts: sharp Reduction in False-Positive Rate in the Second Screening Round of the DENSE Trial. Radiology. 2021;299(2):287-9.

Scheel JR, Lee JM, Sprague BL, Lee CI, Lehman CD. Screening ultrasound as an adjunct to mammography in women with mammographically dense breasts. Am J Obstet Gynecol. 2015;212(1):9-17.

Shahan CL, Layne GP. Advances in Breast Imaging with Current Screening Recommendations and Controversies. Obstet Gynecol Clin North Am. 2022;49(1):1-33.

Sorin V, Sklair-Levy M. Dual-energy contrast-enhanced spectral mammography (CESM) for breast cancer screening. Quant Imaging Med Surg. 2019;9(11):1914-7.

Tagliafico AS, Houssami N. Towards consensus on managing high mammographic density in population breast screening? Breast. 2023;69:422-3.

Tuite CM, Dinome ML, Goldstein LJ. Benefits of Digital Breast Tomosynthesis beyond Baseline Screening. JAMA Network Open. 2020;3(7).

Xu Y, Pan B, Zhou YD, et al. Mammography-detected ultrasound-negative asymptomatic microcalcifications in Chinese women: Would it be safe to watch and wait? Med Hypotheses. 2018;118:9-12.

No breast density stratification

Abdi AJ, Fieselmann A, Pfaff H, Mertelmeier T, Larsen LB. Comparison of screening performance metrics and patient dose of two mammographic image acquisition modes in the Danish National Breast Cancer Screening Programme. Eur J Radiol. 2018;105:188-94.

Armaroli P, Frigerio A, Correale L, et al. A randomised controlled trial of digital breast tomosynthesis vs digital mammography as primary screening tests: screening results over subsequent episodes of the Proteus Donna study. Int J Cancer. 2022;151(10):1778-90.

Destounis S, Arieno A, Morgan R. Initial experience with combination digital breast tomosynthesis plus full field digital mammography or full field digital mammography alone in the screening environment. J Clin Imaging Sci. 2014;4(1):127838.

Dong H, Huang Y, Song F, et al. Improved Performance of Adjunctive Ultrasonography After Mammography Screening for Breast Cancer Among Chinese Females. Clin Breast Cancer. 2018;18(3):e353-e61.

Durand MA, Haas BM, Yao X, et al. Early clinical experience with digital breast tomosynthesis for screening mammography. Radiology. 2015;274(1):85-92.

Elias-Cabot E, Romero-Martin S, Raya-Povedano JL, Brehl AK, Alvarez-Benito M. Impact of real-life use of artificial intelligence as support for human reading in a population-based breast cancer screening program with mammography and tomosynthesis. Eur Radiol. 2024;34(6):3958-66.

Friedewald SM, Rafferty EA, Rose SL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA. 2014;311(24):2499-507.

Giorgi Rossi P, Mancuso P, Pattacini P, et al. Comparing accuracy of tomosynthesis plus digital mammography or synthetic 2D mammography in breast cancer screening: baseline results of the MAITA RCT consortium. Eur J Cancer. 2024;199:113553.

Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial. Lancet. 2016;387(10016):341-8.

Poplack SP, Patel AK, Salter A, Langley Blanton H, Murray D, McGuire C. The impact of adjunctive tomosynthesis on screening mammography outcomes in two widely diverse radiology practices. The Breast Journal. 2021;27(1):13-20.

Rafferty EA, Rose SL, Miller DP, et al. Effect of age on breast cancer screening using tomosynthesis in combination with digital mammography. Breast Cancer Res Treat. 2017;164(3):659-66.

Skaane P, Bandos AI, Niklason LT, et al. Digital Mammography versus Digital Mammography Plus Tomosynthesis in Breast Cancer Screening: The Oslo Tomosynthesis Screening Trial. Radiology. 2019;291(1):23-30.

Sumkin JH, Ganott MA, Chough DM, et al. Recall Rate Reduction with Tomosynthesis During Baseline Screening Examinations: An Assessment From a Prospective Trial. Acad Radiol. 2015;22(12):1477-82.

Xu H-F, Wang H, Liu Y, et al. Baseline Performance of Ultrasound-Based Strategies in Breast Cancer Screening Among Chinese Women. Acad Radiol. 2024.

Incorrect setting

Abdelattef SAA, Ibrahim SF, Abdelhamid WR, Mahmoud FM. Three-dimensional tomosynthesis versus two-dimensional mammography in detection and characterization of different breast lesions. Egyptian Journal of Radiology and Nuclear Medicine. 2021;52(1).

Ahmadinejad N, Rasoulighasemlouei S, Rostamzadeh N, Arian A, Mohajeri A, Miratashi Yazdi SN. Our experience using synthesized mammography vs full field digital mammography in population-based screening. European Journal of Radiology Open. 2023;10:100475.

Arslan ZF, Altunkeser A, Aksoy N, Korez MK, Omeroglu E. Which Modality Should Be Integrated to Increase the Diagnostic Efficiency of BI-RADS 0, 3, and 4 Lesions? Ultrasonography or Digital Breast Tomosynthesis? Iranian Journal of Radiology. 2021;18(4):e113845.

Bahl M, Mercaldo S, Vijapura CA, McCarthy AM, Lehman CD. Comparison of performance metrics with digital 2D versus tomosynthesis mammography in the diagnostic setting. Eur Radiol. 2019;29(2):477-84.

Behrens A, Fasching PA, Schwenke E, et al. Predicting mammographic density with linear ultrasound transducers. Eur J Med Res. 2023;28(1):384.

Carbonaro LA, Di Leo G, Clauser P, et al. Impact on the recall rate of digital breast tomosynthesis as an adjunct to digital mammography in the screening setting. A double reading experience and review of the literature. Eur J Radiol. 2016;85(4):808-14.

Choi JS, Han BK, Ko EY, Kim GR, Ko ES, Park KW. Comparison of synthetic and digital mammography with digital breast tomosynthesis or alone for the detection and classification of microcalcifications. Eur Radiol. 2019;29(1):319-29.

Gatta G, Di Grezia G, Donatello D, et al. Prone 3D abus vs hhus: Diagnostic accuracy and potentialities. Pharmacologyonline. 2020;1(Special Issue):23-38.

Huang JL, Lin Q. Benefit of digital breast tomosynthesis in symptomatic young women (<=30 years) diagnosed with BI-RADS category 4 or 5 on ultrasound. Clin Radiol. 2022;77(1):e55-e63.

Uematsu T, Nakashima K, Harada TL, Nasu H, Igarashi T. Comparisons between artificial intelligence computer-aided detection synthesized mammograms and digital mammograms when used alone and in combination with tomosynthesis images in a virtual screening setting. Jpn J Radiol. 2023;41(1):63-70.

Uematsu T, Nakashima K, Harada TL, Nasu H, Igarashi T. Artificial intelligence computer-aided detection enhances synthesized mammograms: comparison with original digital mammograms alone and in combination with tomosynthesis images in an experimental setting. Breast Cancer. 2023;30(1):46-55.

Duplicates

Elder K, Matheson J, Nickson C, et al. Contrast enhanced mammography in breast cancer surveillance. Breast Cancer Res Treat. 2023;199(2):221-30.

Hruska CB, Conners AL, Jones KN, et al. Diagnostic workup and costs of a single supplemental molecular breast imaging screen of mammographically dense breasts. AJR Am J Roentgenol. 2015;204(6):1345-53.

Lowry KP, Coley RY, Miglioretti DL, et al. Screening Performance of Digital Breast Tomosynthesis vs Digital Mammography in Community Practice by Patient Age, Screening Round, and Breast Density. JAMA Network Open. 2020;3(7):e2011792.

Song SE, Cho N, Chang JM, Chu AJ, Yi A, Moon WK. Diagnostic performances of supplemental breast ultrasound screening in women with personal history of breast cancer. Acta Radiol. 2018;59(5):533-9.

Outcomes not relevant

Foster VM, Trentham-Dietz A, Stout NK, et al. Supplemental breast cancer screening after negative mammography in U.S. women with dense breasts. J Natl Cancer Inst. 2024.

Lee JM, Arao RF, Sprague BL, et al. Performance of Screening Ultrasonography as an Adjunct to Screening Mammography in Women Across the Spectrum of Breast Cancer Risk. JAMA Internal Medicine. 2019;179(5):658-67.

Sprague BL, Ichikawa L, Eavey J, et al. Breast cancer risk characteristics of women undergoing whole-breast ultrasound screening versus mammography alone. Cancer. 2023;129(16):2456-68.

Wernli KJ, DeMartini WB, Ichikawa L, et al. Patterns of breast magnetic resonance imaging use in community practice. JAMA Internal Medicine. 2014;174(1):125-32.

Incorrect design

Stout NK, Miglioretti DL, Su Y-R, et al. Breast Cancer Screening Using Mammography, Digital Breast Tomosynthesis, and Magnetic Resonance Imaging by Breast Density. JAMA internal medicine. 2024.

Appendix 3 Summary and appraisal of individual studies

Data Extraction – Characteristics and results tables for each supplementary imaging modality versus mammography alone

Characteristics and results pertaining to DBT plus mammography versus mammography alone are presented below

Table 1.A Reported characteristics of studies comparing digital breast tomosynthesis (DBT) plus mammography with mammography-only in women with dense breasts

Study name Author, year (country)	Supplementary screening modality / comparator imaging (interval between tests)	Design (number of centres)	Study years	Number with dense breasts in study arm/ number with dense breasts in DM only arm ϕ	Brief population description Age§	Screening programme interval	Reference standard
CBTST Pulido-Carmona, 2024[26] (Spain)#	DBT / DM (Interval: simultaneous)	Retrospective study (1)	2015 Jan to 2016 Dec	4,207 / 8,950	People 50-69y resident in Cordoba participating in breast cancer screening Age NR	Biennial	Linkage with multiple registries and database of centre's Breast Unit
MBTST Olinder, 2023[30] (Sweden)	DBT / DM (Interval: simultaneous -acquired at one screening occasion)	Prospective screening trial (1)	2010 Jan to 2015 Feb	6,645 / 6, 645 (participants underwent both imaging modalities)	Random sample of people 40-74y selected from screening registry in Malmö, Sweden Age NR (reported for highest density quintiles only)	NR	NR
RETomo Pattacini, 2022[36] (Italy)	DBT / DM (Interval: simultaneous)	RCT (3)	2014 Mar to 2017 Aug	5,970 / 5,978	People aged 45–69y attending screening in one of three clinics Age NR	45–49 y annual 50–74 y biennial	All recruited people followed up in screening programme and Cancer Registry

							Evaluation of surgical specimen considered final diagnosis of lesion
Durand, 2021[46] (USA)	DBT / DM (Interval: NR)	Retrospective study (10)	DBT start dates between 2011 Apr and 2012 May to 2013 Jun DM start dates NR	86,571 / 95,914	Screening examinations from 10 academic and community practices Age: 40-49y: 113,950 (29.9%) 50-69y: 209,753 (55.1%) ≥70y: 45,019 (11.8%)	Annual	Linkage to site and state cancer registries
Pang, 2021[51] (Canada)	DBT / DM (Interval: simultaneous; combined DBT and DM)	Observational (2)	2015 to 2018	58,281 / 67,489	People ≥40y who underwent screening at two large volume multisite radiology groups in Alberta, Canada Age NR	Biennial	Linkage to Alberta Cancer Registry
Ban, 2021[53] (Japan)	DBT / DM (Interval: NR)	Prospective trial (1)	2017 May to 2019 Mar	1,739 / 4,226	People attending population-based screening in Japan who also had opportunistic DBT screening (workplace/private) and were >30y Age: NR	Biennial	NR
Oslo Tomosynthesis Screening Trial	DBT / DM	Prospective clinical trial (1)	2010 Nov to 2012 Dec	8,466 / 8,466 (participants underwent both imaging modalities)	People aged 50-69y attending population-based screening program, BreastScreen Norway	Biennial	2 years of follow-up to assess interval cancers

Osteras, 2019[21] (Norway)	(Interval: simultaneous [combined DBT and DM])				Age: 50–54y: 3,158 55–59y: 2,390 60–64y: 1,574 65–69y: 1,624		
Stepanek, 2019[52] (USA)	DBT / DM (Interval: simultaneous; combined DBT and DM)	Retrospective review (1)	DM: 2010 Sep to 2011 Aug DBT/DM: 2014 Jan to 2015 Jun	4,389 / 4,895	Screening programme with DM; DBT made available at an additional charge Age NR	Annual	Category 3 lesions followed for minimum 2 y or cross-referenced with the state tumour registry Negative findings BI-RADS 1/2 resumed annual screenings
MBTST Johnson, 2019[41] (Sweden)	DBT / DM (Interval: simultaneous -one screening occasion)	Prospective, population-based screening trial (NR)	2010 Jan to 2015 Feb	6,202 / 6,202 (participants underwent both imaging modalities)	Random sample of people from screening registry in the city of Malmö, Sweden Age NR	NR	Pathologic assessment of surgical specimens Linkage with Cancer Registry
Upadhyay, 2018[38] (UK)	DBT / DM (Interval: simultaneous)	Retrospective reader study (1)	2013 Jan to 2013 Dec	423 / 423 (participants underwent both imaging modalities)	People aged 46-53 were recruited from the prevalent screening round of the NHS Breast Screening Programme Age NR	Triennial	Recalls: Assessment within 3 weeks (further imaging, biopsy as appropriate)
Rose, 2017[48] (USA)	DBT / DM (Interval: Simultaneous -combined DBT/DM	Retrospective review of screening mammography audit data (31)	2015 Jan to 2015 Dec	10,360 / 21,929	Community-based screening programme for people <50y Age NR	NR	Data validated by auditing patient outcomes, imaging reports, and pathology results for biopsy

	for all or part of the analysis period)						and surgical specimens
Freer, 2017[45] (USA)	DBT / DM (Interval: simultaneous)	Retrospective study with institutional databases (NR)	2013 Oct to 2014 Dec (cohort 3 end date 2015 Dec)	521 / 8302	People who presented for screening MM to any screening facility of their academic medical centre Cohort 2: Under 40: 1.7% ; 40–49: 27.5%; 50–59: 34.2%; 60–69: 25.3%; 70–74: 6.9%; Over 74: 4.4% Cohort 3: Under 40: 1.4% ; 40–49: 26.0%; 50–59: 31.7%; 60–69: 25.5%; 70–74: 7.8%; Over 74: 7.6%	NR	NR
STORM-2 Bernardi, 2016[35] (Italy)	DBT / DM (Interval: simultaneous)	Prospective population-based screening study (1)	2013 May to 2015 May	2,592 / 2,592 (participants underwent both imaging modalities)	People >49y attending population-based screening programme Age NR	Biennial	Biopsy Recalled patients: completed assessment outcome, inclusive of work-up imaging (with or without biopsy)
Starikov, 2016[49] (USA)	DBT / MM (Interval: same day)	Retrospective observational case–control study (1)	2013 Jan to 2013 Dec	1,875 / 7,117	People presenting for screening mammography Age: Majority of patients >40y	Annual	List of biopsy proven screen-detected breast cancers
PROSPR Conant, 2016[47] (USA)	DBT / DM (Interval NR)	Retrospective analysis of prospective cohort data (3)	2011 to 2014	21,133 / 44,303	People 40-74y with no known prior breast cancer Age:	NR	Biopsy information from electronic health records and pathology databases

					40–49 years: 37,155 (26.0%) for DM, 18,668 (33.3%) for DBT 50–59 years: 51,096 (35.8%) for DM, 20,839 (36.4%) for DBT 60–74 years: 54,632 (38.2%) for DM, 16,941 (30.3%) for DBT		-≥ 1 year of imaging follow-up Cancer data from local institutional tumour registries, state registries, and one statewide surveillance system
ASTOUND Tagliofico, 2016[40] (Italy)	DBT / negative DM (Interval: simultaneous	Prospective multicentre study (5)	2012 Dec to 2015 Mar	3,231 / 3,231 (participants underwent both imaging modalities)	Asymptomatic people with negative MM aged 38 years or older with dense breasts Age: median (IQR) age 51 y (44-78) (of those invited to study; 64 declined to participate))	NR	Biopsy Recalled subjects: work-up imaging with or without core-needle biopsy
MacDonald, 2016[43] (USA)	DBT / DM (Interval NR)	Retrospective analysis (1)	2010 Sep to 2011 Aug	3,611 / 3,489	People underdoing screening mammography at the University of Pennsylvania Mean (SD) age: DM cohort: 56.9 (11.0)	Unclear, likely annual	Biopsy Pennsylvania State Cancer Registry queried to determine interval cancer rate
MacDonald, 2015[44] (USA)	DBT / DM (Interval NR)	Retrospective analysis of a natural outcomes experiment (1)	2010 Sept to 2013 Feb	632 / 405	People underdoing screening mammography at the University of Pennsylvania Mean (SD) age: DM cohort: 49.2y (10.1) DBT cohort: 48.9y (10.6)	NR	Biopsy or surgical outcomes tracked within 180 days of screening recall Min 12 months of follow-up data

							available for all patients in both cohorts Linkage to Pennsylvania State Cancer Registry

^ NR, but obtained via author correspondence
ϕ Number of people with dense breasts in the DBT plus DM study arm, and in the DM only arm (according to BIRADS categories D and C corresponding to heterogeneously and extremely dense breasts, respectively)
§ Age shown if reported specifically for people with dense breasts; presented as mean, median, proportions as reported
# Study was eligible for inclusion however as the intervention (supplemental modality) included synthesis mammography as well as digital breast tomosynthesis, study was not amenable to meta-analysis

RCT, randomised controlled trial; CBTST, Córdoba Breast Tomosynthesis Screening Trial; MBTST, Malmö Breast Tomosynthesis Screening Trial; RETomo, Reggio Emilia Tomosynthesis; PROSPR, Population-based research optimizing screening through personalized regimens; STORM, Screening with Tomosynthesis Or standard Mammograph; ASTOUND, Adjunct Screening With Tomosynthesis or Ultrasound in People With Mammography-Negative Dense Breasts; NR, not reported; DM, digital mammography; DBT, digital breast tomosynthesis; MM, mammography; BIRADS, breast imaging reporting and data system; IQR, inter-quartile range; SD, standard deviation

Table 1.B Summary of CDR, recall rate, biopsy rate, interval cancers, false negatives, false positive, sensitivity and specificity for studies comparing digital breast tomosynthesis (DBT) plus mammography with mammography-only in people with dense breasts

Study name Author, year (country)	**Study Question Supplemental screening modality /comparator imaging**	**CDR per 1,000 exams (lower CI, upper CI)**	**Recall rate % (lower CI, higher CI) (definition of recall rate^)**	**PPV1 for recall % (lower CI, upper CI)**	**Biopsy rate % (lower CI, upper CI)**	**PPV3 for biopsy % (lower CI, upper CI)**	**Interval cancers % (lower CI, higher CI)**	**False negatives % (lower CI, higher CI)**	**False positives % (lower CI, higher CI)**	**Sensitivity % (lower CI, upper CI)**	**Specificity % (lower CI, upper CI)**

CBTST Pulido-Carmona, 2024[28] (Spain)	DBT plus DM	NR	NR	NR	NR	NR	0.95 (0.55, 1.35)	NR	NR	NR	NR
	DM only	NR	NR	NR	NR	NR	3.17 (2.64, 3.62)	NR	NR	NR	NR
MBTST Olinder, 2023[30] (Sweden)	DBT plus DM	Grade C density 11.8 (9.1, 15.2) Grade D density 16.3 (10.6, 25)	Grade C density 5 (4.4, 5.7) Grade D density 7 (5.7, 8.6)	Grade C density 23.6 (18.7, 29.3) Grade D density 23.3 (15.6, 33.2)	Grade C density 2.9 (2.5, 3.4) Grade D density 4.3 (3.3, 5.6)	Grade C density 40.6 (32.9, 48.8) Grade D density 37.7 (25.9, 51.2)	NR	NR	Grade C density 3.8 (3.3, 4.4) Grade D density 5.4 (4.2, 6.8)	Grade C density 78.4 (67.7, 86.2) Grade D density 90.9 (72.2, 97.5)	Grade C density 96.1 (95.6, 96.6) Grade D density 94.5 (93.1, 95.7)
	DM only	Grade C density 10 (7.6, 13.2) Grade D density 8.9 (5, 15.9)	Grade C density 3.6 (3.1, 4.1) Grade D density 4.6 (3.5, 5.9) (definition: NR)	Grade C density 27.8 (21.7, 34.9) Grade D density 19.6 (11.3, 31.8)	Grade C density 2.4 (2, 2.8) Grade D density 2.8 (2, 3.8)	Grade C density 41.9 (33.3, 50.9) Grade D density 32.4 (19.1, 49.2)	NR	NR	Grade C density 2.6 (2.2, 3.1) Grade D density 3.7 (2.7, 4.9)	Grade C density 66.2 (54.9, 76) Grade D density 50 (30.7, 69.3)	Grade C density 97.4 (96.9, 97.8) Grade D density 96.3 (95.1, 97.2)
RETomo Pattacini, 2022[36] (Italy)	DBT plus DM	n=49	NR	18.4	NR	NR	0.23 %	NR	NR	NR	NR
	DM only	n=31	NR	10.8	NR	NR	0.25 %	NR	NR	NR	NR
Durand, 2021[46] (USA)	DBT plus DM	2.6	114.6 (per 1,000 ex-ams)	NR	NR	NR	NR	n=70	NR	NR	NR

	DM only	2.1	133.6 (per 1,000 exams) (definition: NR)	NR	NR	NR	NR	n=72	NR	NR	NR
Pang, 2021[51] (Canada)	DBT plus DM	Grade C density 6.5 Grade D density 4	Grade C density 8.5 Grade D density 7.7	Grade C density 7.7 Grade D density 5.2	NR	NR	NR	NR	NR	NR	NR
	DM only	Grade C density 4.2 Grade D density 3.7	Grade C density 8.0 Grade D density 7.0 (definition: reported as Abnormal Call Rate; refers to BI-RADS 0,4,5)	Grade C density 5.2 Grade D density 5.3	NR	NR	NR	NR	NR	NR	NR
Ban, 2021[53] (Japan)	DBT plus DM	n=4	n=46	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	n=6	n=149 (definition: reported as 'cases requiring further examination')	NR	NR	NR	NR	NR	NR	NR	NR
	DBT plus DM	n=106	NR	NR	NR	NR	NR	NR	n=995	NR	NR

Oslo Tomosyn-thesis Screening Trial Osteras, 2019[21] (Norway)	DM only	n=82	NR	NR	NR	NR	NR	NR	n=1044	NR	NR
Stepanek, 2019[52] (USA)	DBT plus DM	NR	n=602	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	NR	n=760 (definition: BIRADS 0)	NR	NR	NR	NR	NR	NR	NR	NR
MBTST Johnson, 2019[41] (Sweden)	DBT plus DM	n=85	NR	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	n=64	NR	NR	NR	NR	NR	NR	NR	NR	NR
Upadhyay, 2018[38] (UK)	DBT plus DM	NR	13.5	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	NR	19.6 (Definition: NR)	NR	NR	NR	NR	NR	NR	NR	NR
Rose, 2017[48] (USA)	DBT plus DM	3.5	132 (per 1,000 ex-ams)	2.5	20.4 (per 1,000 exams)	17.2	NR	NR	NR	NR	NR
	DM only	2.1	136 (per 1,000 ex-ams) (definition: BIRADS 0)	1.6	16 (per 1,000 exams)	13.4	NR	NR	NR	NR	NR

Freer, 2017[45] (USA)	DBT plus DM	NR	6.63	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	NR	8.92 (definition: screening for additional evaluation)	NR	NR	NR	NR	NR	NR	NR	NR
STORM-2 Bernardi, 2016[35] (Italy)	DBT plus DM	13.1 (9.1, 18.3)	NR	NR	NR	NR	NR	NR	5.01 (4.2, 5.93)~	NR	NR
	DM only	7.7 (4.7, 11.9)	NR	NR	NR	NR	NR	NR	3.95 (3.23, 4.78)~	NR	NR
Starikov, 2016[49] (USA)	DBT plus DM	5.3	10.4	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	3.8	19.9 (definition: BIRADS 0)	NR	NR	NR	NR	NR	NR	NR	NR
PROSPR Conant, 2016[47] (USA)	DBT plus DM	6.8	10.3	6	2.2	NR	NR	NR	NR	NR	NR
	DM only	4.7	12.69 (definition: BIRADS 0,3,4,5)	3.7	2.2	NR	NR	NR	NR	NR	NR
ASTOUND Tagliofico, 2016[40] (Italy)	Negative DM plus DBT	4 (1.8, 6.2)	NR	NR	NR	37 (21.3, 55.4)	NR	NR	n=53	NR	NR
MacDonald, 2016[41] (USA)	DBT plus DM	6.6	NR	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	5.2	NR	NR	NR	NR	NR	NR	NR	NR	NR

MacDonald, 2015 (USA)[44]	DBT plus DM	NR	n=109	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	NR	n=85 (definition: BIRADS 0,4,5)	NR	NR	NR	NR	NR	NR	NR	NR

± Grade C breast density refers to heterogeneously dense breasts and grade D density refers to extremely dense breasts, and may be described as such in some reports
^Recall rate definition as reported by study authors
~ reported as 'false positive recalls'

CDR, cancer detection rate; CI, confidence interval; CBTST, Córdoba Breast Tomosynthesis Screening Trial; MBTST, Malmö Breast Tomosynthesis Screening Trial; RE-Tomo, Reggio Emilia Tomosynthesis; PROSPR, Population-based research optimizing screening through personalized regimens; STORM, Screening with Tomosynthesis Or standard Mammograph; ASTOUND, Adjunct Screening With Tomosynthesis or Ultrasound in Women With Mammography-Negative Dense Breasts; NR, not reported; DM, digital mammography; DBT, digital breast tomosynthesis; MM, mammography; PPV1, positive predictive value (number of cancers diagnosed per number of positive screens); PPV3, positive predictive value (biopsy proven predictive value); BIRADS breast imaging reporting and data system

Characteristics and results of ABUS plus mammography versus mammography alone are presented below

Table 2.A Reported characteristics of studies comparing automated breast ultrasound (ABUS) plus mammography with mammography-only in people with dense breasts

Study name Author, year (country)	**Supplementary screening modality / comparator imaging (interval between tests)**	**Design (number of centres)**	**Study years**	**Number with dense breasts in study arm/ number with dense breasts in DM only arm ϕ**	**Brief population description Age§**	**Screening programme interval**	**Reference standard**

BRAID, Gilbert, 2024[20] (UK)	ABUS/ negative screening MM (Interval median days [IQR]: 111 [77-150])	RCT (10)	2019 Sept to 2024 Mar	2,141 / 6,306	MM-negative people attending UK population breast screening at 10 centres Median (IQR) age: 56 (52-61)	Triennial	Recall cases: further imaging and biopsy if lesion was confirmed Where doubt about lesion or not: repeat CEM or MRI/ short term follow up offered
Kwon, 2023[24] (South Korea)	ABUS / DM (Mean interval 4 days; range 0-90 days)	Retrospective analysis involving database search (1)	2018 Jan to 2019 Dec	2,155 / 2,155 (participants underwent both imaging modalities)	Database of consecutive asymptomatic people breast cancer screening Age profile: <50y 40% >50y 60%	NR	Histopathologic results from surgical excision, US-guided core biopsy or vacuum-assisted biopsy, stereotactic biopsy, or stability on follow-up imaging
Ren, 2023[26] (China)	ABUS/ DM (Interval NR, but people with an odd number at end of ID receive ABUS first while those with even number attend DM first – suggesting immediate)	Multicentre, population-based study (6)	2018 Feb to 2022 Aug	Total group 10,884	Asymptomatic people attending breast cancer screening in China Median (IQR) age: 51.54 (4.61)	NR	Biopsy or 1-year follow-up
Gatta, 2021[33] (Italy)	ABUS/ FFDM (Interval NR)	Prospective observational study (1)	2017 Jun to 2019 Feb	1,165 / 1,165	People with dense breasts attending health screening centre Age:	Annual (for dense breasts)	No signs of malignancy: dense breast annual checkup Suspected malignancy: Biopsy

					40-50y 68% 50-60y 25% 60-70y 7%		
Wilczek, 2016[42] (Sweden)	ABUS/ DM (Interval: immediate)	Prospective evaluation of a breast screening programme (1)	2020 Nov to 2012 Feb	1,668 / 1,668	Asymptomatic people with dense breasts who had been invited for breast cancer service screening mammography Age: 40-49y: 59.7% 50-59y: 26.7% 60-69y: 13.5%	People aged 40-49y every 18 months	Suspicious findings (DM or ABUS): recalled and subjected to mammography work–up with complementary views and HHUS ^Code 1–2: invited to the next screening after appropriate interval ^Code 3–5: biopsy
SomoInsight Brem, 2015[50] (USA)	ABUS/ DM (Interval: Immediate)	Observational, multicentre study (10)	2009 to 2019	15,318 / 15,318	People aged >25 years with dense breast tissue Mean (SD) age 53.3 (10)	Annual	Normal, benign, or probably benign: followed up for 12 months Developed breast symptoms during the follow-up period: clinically indicated evaluation

§ Age shown if reported specifically for people with dense breasts; presented as mean, median, proportions as reported
ϕ Number of people with dense breasts in the ABUS plus DM study arm, and in the DM only arm (according to BIRADS categories D and C corresponding to heterogeneously and extremely dense breasts, respectively)
^ Five step coding: (1) normal, (2) benign, (3) probably benign, (4) highly suspicious of malignancy, (5) malignant

SD, standard deviation; IQR, Inter-quartile range; RCT, randomised controlled trial; MM mammography; DM digital mammography; ABUS, automated breast ultrasound; BRAID, Breast Screening; Risk Adaptive Imaging for Density; HHUS, hand-held ultrasound; NR, not reported; CEM contrast enhanced mammography; BIRADS, breast imaging reporting and data system; MRI, magnetic resonance imaging

Table 2.B Summary of CDR, recall rate, biopsy rate, interval cancers, false negatives, false positive, sensitivity and specificity for studies comparing automated breast ultrasound (ABUS) plus mammography with mammography-only in people with dense breasts

Study name Author, year (country)	Study Question Supplementary screening modality / comparator	CDR per 1,000 exams (lower CI, upper CI)	Recall rate % (lower CI, higher CI) (definition of recall rate^)	PPV1 for recall % (lower CI, upper CI)	Biopsy rate % (lower CI, upper CI)	PPV3 for biopsy % (lower CI, upper CI)	Interval cancers % (lower CI, higher CI)	False negatives % (lower CI, higher CI)	False positives % (lower CI, higher CI)	Sensitivity % (lower CI, upper CI)	Specificity % (lower CI, upper CI)
BRAID, Gilbert, 2024[20] (UK)	Negative DM plus ABUS	4.2 (1.9, 8)	4 (3.2, 4.9) (Definition: BI-RADS 3, 4, or 5)	NR	1.5% (1, 2.1)	NR	NR	NR	NR	NR	NR
Kwon, 2023[24] (South Korea)	ABUS plus DM	9.3 (7.7, 10.3)	14 (13.4, 14.4)	6.6 (5.6, 7.4)	3.3 (3.1, 3.5)	27.8 (23.4, 31)	NR	NR	2.4 (2.2, 2.6)	63.6 (40.9, 81.8)	NR
	DM only	6.5 (5.2, 7.2)	6 (5.7, 6.3) (Definition: reported as Abnormal Interpretation Rate)	10.8 (8.8, 12.2)	1.9 (1.8, 2.1)	33.3 (26.7, 37.8)	NR	NR	1.3 (1.1, 1.4)	94.6 (93.6, 95.5)	NR
Ren, 2023[26] (China)	ABUS plus DM	Age 45-54y: 3.3 (2.2, 4.8) Age 55-64y: 6.09 (3.3, 10.9)	NR	NR	NR	Age 45-54y: 16.87 (11.68, 23.63) Age 55-64y:	NR	NR	Age 45-54y: 1.63 (1.38, 1.93) Age 55-64y: 1.17	Age 45-54y: 96.55 (80.37, 99.82) Age 55-64y: 100	Age 45-54y: 97.11 (96.72, 97.45) Age 55-64y: 98.01

						34.29 (19.69, 52.27)			(0.76, 1.78)	(69.87, 100)	(97.26, 98.56)
	DM only	Age 45-54y: 2.48 (1.6, 3.9) Age 55-64y: 4.56 (2.3, 9)	NR	NR	NR	Age 45-54y: 24.71 (16.27, 35.47) Age 55-64y: 40.91 (21.48, 63.32)	NR	NR	Age 45-54y: 0.75 (0.58, 0.96) Age 55-64y: 0.66 (0.37, 1.16)	Age 45-54y: 72.41 (52.51, 86.55) Age 55-64y: 75 (42.84, 93.31)	Age 45-54y: 98.9 (98.65, 99.11) Age 55-64y: 99.08 (98.52, 99.44)
Gatta, 2021[33] (Italy)	ABUS plus DM	6.8 (5, 8.1)	26.6 (16.2, 30) (per 1,000 exams)	24.8 (13.7, 43.2)	14 (5, 28) (per 1,000 persons)	41 (26.4, 59.2)	NR	NR	NR	93.5 (79.2, 98.2)	87 (71, 94.8)
	DM only	3.4 (1.7, 5.8)	14.5 (9, 19.8) (per 1,000 exams) (Definition: suspicious findings)	68 (41.3, 82.2)	7 (4.1, 8.2) (per 1,000 persons)	58.8 (30.9, 78.3)	NR	NR	NR	58.8 (30.9, 78.3)	94 (73, 98)
Wilczek, 2016[42] (Sweden)	ABUS plus DM	6.6 (3, 10.2)	22.8 (16.2, 30) (per 1,000 exams)	28.9 (14.3, 42.3)	13.8 (8.4, 19.8) (per 1,000 exams)	47.8 (27, 66.7)	NR	NR	NR	100	98.4 (97.8, 98.9)
	DM only	4.2 (1.2, 7.2)	13.8 (9.0, 19.8) (per 1,000 exams)	30.4 (12.3, 49.1)	6.6 (3, 10.8) (per 1,000 exams)	63.6 (33.3, 90)	NR	NR	NR	63.6 (33.3, 90.9)	99 (98.5, 99.4)

			(Definition: suspicious findings)								
SomoInsight Brem, 2015[50] (USA)	ABUS plus DM	7.3 (5.9, 8.7)	284.9 (278, 292.2) (per 1,000 exams)	2.6 (2.1, 3.1)	77 (72.9, 81) (per 1,000 exams)	9.8 (8.1, 11.7)	NR	NR	NR	100	NR
	DM only	5.4 (4.2, 6.6)	150.2 (144.1, 155.7) (per 1,000 exams (Definition: BI-RADS category 0 or combined impression of immediate management)	3.6 (3.28, 4.4)	39.8 (36.7, 43.2) (per 1,000 exams)	14 (11.2, 16.8)	NR	NR	NR	73.2	NR

Data presented as % (lower confidence interval, upper confidence interval) unless otherwise specified

CDR, cancer detection rate; DM, digital mammography; ABUS, automated breast ultrasound; BRAID, Breast Screening; Risk Adaptive Imaging for Density; NR, not reported; BIRADS breast imaging reporting and data system; PPV1, positive predictive value (number of cancers diagnosed per number of positive screens); PPV3, positive predictive value (biopsy proven predictive value); CI, confidence interval

Characteristics and results of HHUS plus mammography versus mammography alone are presented below

Table 3.A ***Reported characteristics of studies comparing hand-held ultrasound (HHUS) plus mammography with mammography-only in people with dense breasts***

Study name Author, year (country)	Supplementary screening modality / comparator imaging (interval between tests)	Design (number of centres)	Study years	Number with dense breasts in study arm/ number with dense breasts in DM only arm ϕ	Brief population description Age§	Screening programme interval	Reference standard
Ha, 2024[31] (South Korea)	HHUS / DM (Interval NR, but some on same day)	Retrospective study (1)	2017 Jan to 2018 Dec	5,707 / 5,707	Consecutive asymptomatic people (≥40 years) with dense breasts who underwent routine screening DM and supplemental US Age: Mean (SD) age 52.4 (7.9) y	Annual	Histologic examination and 1-year follow-up data Positive cases: short-interval follow-up (6 months) and additional DM or biopsy recommended
Lee, 2024[27] (South Korea)	US/ DM (Interval: within 1 month)	Retrospective study (1)	2017 Jan to 2017 Dec	1,325 / 1,325	Patients with dense breasts who underwent both screening DM and screening US Mean 53 years, median 53 years	At least biennial	Cancer diagnosis: pathological evaluation Benign diagnoses: MM/US follow-up documenting at least 24 months of imaging stability
J-Start Nakamura, 2024[25] (Japan)	HHUS / DM (Interval NR)	Retrospective analysis of breast cancer screening data (NR)	2018 to 2021	6,271 / 11,765	People in their 40s undergoing breast cancer screening Age: NR	NR	Positive cases referred to specialised institution

					for people with dense breasts		
Ren, 2023[26] (China)	HHUS / negative DM (Interval NR, but people with an odd number at end of ID receive ABUS first while those with even number attend DM first – suggesting immediately)	Multicentre, population-based study (6)	2018 Feb to 2022 Aug	10,884 / 10,884 (participants underwent both imaging modalities)	Asymptomatic people attending breast cancer screening in China with negative DM Median (IQR) age: 51.54 (4.61)	NR	Biopsy or 1-year follow-up
Rani, 2023[29] (India)	US / DM (Interval: 'mammography followed by ultrasound')	Prospective, observational study (1)	NR	125 / 125 (participants underwent both imaging modalities)	Asymptomatic people coming for breast cancer screening Age > 40y	Annual health screening - not clear if this was annual breast screening	Biopsy BIRADS 3 lesions subjected to short interval follow up and few lesions had surgical excision
J-Start Harada-Shoji, 2021[55] (Japan)	HHUS / DM (Interval NR)	Secondary analysis of a randomised clinical trial (42)	2007 Jul to 2011 Mar	5,797 / 5,593	Asymptomatic people aged 40 to 49 years who were enrolled in J-START Mean (SD) age: 44.8 (2.8)	NR	Linkage with hospital discharge records and cancer registry databases used to identify breast cancer diagnosis information

Buchberger, 2018[34] (Austria)	HHUS / DM (Interval: at same visit)	Observational data from a population-based screening program (22)	2008 Jun to 2010 May	31,918 / 31,918 (participants underwent both imaging modalities)	Participants identified through the Tyrolean breast cancer screening program, included all people aged 40–69 who lived in Tyrol and were covered by compulsory social insurance Age NR	Aged 40– 59: annual Aged 60–69: biennial	Linking the screening dataset to all breast cancer cases collected in Cancer Registry of Tyrol
ASTOUND Tagliafico, 2016[40] (Italy)	HHUS / negative DM (Interval unclear, but likely immediate)	Prospective multicentre study (5)	2012 Dec to 2015 Mar	3,231 / 3,231 (participants underwent both imaging modalities)	Asymptomatic people with negative MM aged 38 years or older with dense breasts Age: Median (IQR) 51 (44-78)	NR	Biopsy Recalled subjects: work-up imaging with or without core-needle biopsy
Chang, 2015[54] (South Korea)	HHUS / negative DM (Interval unclear: 'Scheduling simultaneously performed')	Retrospective review of a database (1)	2008 Jan to 2008 Dec	990 / 990 (participants underwent both imaging modalities)	MM-negative people attending health screening centre 62% of all participants <50y	NR	Most severe biopsy results within 1 year of screening and clinical follow-up at 1 year or both

Data presented as % (lower confidence interval, upper confidence interval) unless otherwise specified

CDR, cancer detection rate; DM, digital mammography; ABUS, automated breast ultrasound; BRAID, Breast Screening; Risk Adaptive Imaging for Density; NR, not reported; BIRADS breast imaging reporting and data system; PPV1, positive predictive value (number of cancers diagnosed per number of positive screens); PPV3, positive predictive value (biopsy proven predictive value); CI, confidence interval

Table 3.B Summary of CDR, recall rate, biopsy rate, interval cancers, false negatives, false positive, sensitivity and specificity for studies comparing hand-held breast ultrasound (HHUS) plus mammography with mammography-only in people with dense breasts

Study name Author, year (country)	Study Question **Supplementary screening modality / comparator**	**CDR per 1,000 exams (lower CI, upper CI)**	**Recall rate % (lower CI, higher CI) (definition of recall rate^)**	**PPV1 for recall % (lower CI, upper CI)**	**Biopsy rate % (lower CI, upper CI)**	**PPV3 for biopsy % (lower CI, upper CI)**	**Interval cancers % (lower CI, higher CI)**	**False negatives % (lower CI, higher CI)**	**False positives % (lower CI, higher CI)**	**Sensitivity % (lower CI, upper CI)**	**Specificity % (lower CI, upper CI)**
Ha, 2024[31] (South Korea)	HHUS plus DM	5.6 (3.9, 7.9)	22.9 (21.8, 24.0)	2.5 (1.7, 3.4)	NR	NR	NR	NR	NR	97 (84.7, 99.5)	77.6 (76.5, 78.6)
	DM only	3.3 (2.1, 5.2)	6 (5.4, 6.7) (reported as Abnormal Interpretation Rate)	5.5 (3.6, 8.5)	NR	NR	NR	NR	NR	57.6 (40.8, 72.8)	94.3 (93.6, 94.8)
Lee, 2024[27] (South Korea)	HHUS plus DM	9.1 (4.7, 15.8)	11.7 (10, 13.6)	NR	NR	NR	NR	NR	NR	100 (73.5, 100)	89.1 (87.3, 90.7)
	DM only	6 (2.6, 11.9)	4.4 (3.3, 5.6) (Definition NR)	13.8 (9, 20.6)	NR	53.3 (35.3, 70.6)	NR	NR	NR	66.7 (34.8, 90.1)	96.2 (95, 97.2)
J-Start Nakamura, 2024[25] (Japan)	HHUS plus DM	n=25	3.8	10.5	NR	NR	NR	NR	NR	NR	NR
	DM only	n=19	1.7	9.5	NR	NR	NR	NR	NR	NR	NR

			(Definition NR)								
Ren, 2023[26] (China)	HHUS plus DM	Age 45-54y: 3.18 (2.2, 4.7) Age 55-64y: 6.09 (3.3, 10.9)	NR	NR	NR	Age 45-54y: 15.34 (10.52, 21.72) Age 55-64y: 30 (17.09, 46.71)	NR	NR	Age 45-54y: 1.76 (1.5, 2.07) Age 55-64y: 1.42 (0.96, 2.07)	Age 45-54y: 93.1 (75.78, 98.8) Age 55-64y: 100 (69.87, 100)	Age 45-54y: 97.34 (96.97, 97.67) Age 55-64y: 97.91 (97.15, 98.48)
	DM only	Age 45-54y: 2.48 (1.6, 3.9) Age 55-64y: 4.56 (2.3, 9)	NR	NR	NR	Age 45-54y: 24.71 (16.27, 35.47) Age 55-64y: 40.91 (21.48, 63.32)	NR	NR	Age 45-54y: 0.75 (0.58, 0.96) Age 55-64y: 0.66 (0.37, 1.16)	Age 45-54y: 72.41 (52.51, 86.55) Age 55-64y: 75 (42.84, 93.31)	Age 45-54y: 98.9 (98.65, 99.11) Age 55-64y: 99.08 (98.52, 99.44)
Rani, 2023[29] (India)	HHUS plus DM	NR	NR	NR	NR	NR	NR	NR	NR	100	87.5
	DM only	NR	NR	NR	NR	NR	NR	NR	NR	25	100
J-Start Harada-Shoji, 2021[55] (Japan)	HHUS plus DM	7.1 (4.9, 9.2)	15.2	NR	6.2%	NR	0.5% (0.1, 1.1)	NR	NR	93.2 (85.7, 100)	85.4 (84.5, 86.3)
	DM only	4.3 (2.6, 6)	8.7 (Definition: The need for any additional diagnostic testing after screening, including	NR	2.3%	NR	1.8% (0.7, 2.9)	NR	NR	70.6 (55.3, 85.9)	91.7 (91, 92.4)

			imaging and/or biopsy)								
Buchberger, 2018[34] (Austria)	HHUS plus DM	2.4 (1.9, 3)	11 (9.9, 12.2) (per 1,000 exams)	10.5 (8.4, 12.9)	6.5 (5.6, 7.4) (per 1,000 exams)	37.7 (31.1, 44.7)	NR	NR	NR	81.3 (72, 88.5)	NR
	DM only	1.8 (1.4, 2.4)	5.7 (4.9, 6.6) (per 1,000 exams) (Definition: BI-RADS 0, 4 and 5)	13.3 (10.3, 16.8)	3.5 (2.9, 4.2) (per 1,000 exams)	52.7 (43, 62.2)	NR	NR	NR	61.5 (51, 71.2)	NR
ASTOUND Tagliafico, 2016[40] (Italy)	Negative DM plus HHUS	7.1 (4.2, 10)	NR	NR	NR	48 (34.1, 63.9)	NR	NR	n=65	NR	NR
Chang, 2015[54] (South Korea)	Negative DM plus HHUS	5.1 (1.8, 12.1)	NR	NR	NR	NR	n=0	NR	NR	NR	NR

^Recall rate definition as reported by study authors

DM digital mammography; CDR, cancer detection rate; HHUS, hand-held ultrasound; J-START, Japan Strategic Anti-cancer Randomized Trial; CI confidence interval; ASTOUND; PPV1, positive predictive value (number of cancers diagnosed per number of positive screens); PPV3, positive predictive value (biopsy proven predictive value); CI, confidence interval; ASTOUND, Adjunct Screening With Tomosynthesis or Ultrasound in Women With Mammography-Negative Dense Breasts

Characteristics and results of MRI plus mammography versus mammography alone are presented below

Table 4.A Reported characteristics of studies comparing digital breast tomosynthesis (DBT) plus mammography with mammography-only in people with dense breasts

Study Name Author, year (country)	Supplementary screening modality / comparator imaging (Interval between tests)	Design (number of centres)	Study years	Number with dense breasts in study arm/ number with dense breasts in DM only arm ϕ	Brief population description Age§	Screening programme interval	Reference standard
MA-DETECT Kaiser, 2024[32] (Germany)	Fp-MRI / negative screening DM (Interval: not to exceed 2 months)	Prospective screening study, preliminary report (1)	2021 Jun to 2023 Jun	200 / 200 (participants underwent both imaging modalities)	MM-negative people attending national German breast cancer screening programme; report for first 200 people Age NR	NR	BI-RADS 3: MRI follow-up after 6 months BI-RADS 4 and 5: Biopsy (as per EUSOBI recommendations)
BRAID Gilbert, 2024[20] (UK)	Ab-MRI / negative screening DM (Interval median days [IQR]: 143 [98-183])	RCT (10)	2019 Sept to 2024 Mar	2,130 / 6,306 (participants underwent both imaging modalities)	MM-negative people attending UK population breast screening at 10 centres Median (IQR) age: 56 (52-61)	Triennial	Recall cases: further imaging and biopsy if lesion was confirmed Ab-MRI group: Fp-MRI conducted when no lesion found on assessment Where doubt about lesion or not: repeat CEM or MRI/ short term follow up offered

DENSE (2nd Round) Veenhuizen, 2021[39] (Netherlands)	Fp-MRI / negative screening DM (Median [IQR] interval 8 [3-13] weeks)	Prospective, multicentre trial embedded within the screening programme (NR)	2011 Dec to 2016 Jan	3,436 / 3,436 (participants underwent both imaging modalities)	MM-negative people attending Dutch population based biennial screening programme for second MRI (**all extremely dense breasts**) Median (IQR) age: 54 (51-59)	Biennial	MRI screening and diagnostic work-up (biopsy and histopathological examination)
DENSE (1st Round) Bakker, 2019[37] (Netherlands)	FpMRI/ negative screening DM (Median [IQR] interval: 10 [I8-14] weeks)	Multicentre RCT (8)	2011 Dec to 2015 Nov	4,783 / 4,783 (participants underwent both imaging modalities)	MM-negative people attending Dutch National population-based biennial screening for first round MRI (**all extremely dense breasts**) Median (IQR) age: DM only: 54 (51-61) MRI: 55 (51-61)	Biennial	Linkage with the Netherlands Cancer Registry MRI invitation group: 6-month repeat screening
Kuhl, 2017[22] (Germany)	Fp-MRI / negative screening DM (Median interval: 5 days; range 0-28 days)	Prospective observational study (2)	2005 Jan to 2013 Dec	1,282 / 1,282 (participants underwent both modalities)	MM-negative women of average breast cancer risk invited to supplemental screening at one of two academic centres. Some women also negative on ultrasound. First round data presented. Age NR	NR	Negative screens followed with MM with or without US for at least another 24 months BI-RADS 3 underwent short-term follow-up and returned to regular screening Suspicious findings underwent biopsy with MRI screening High risk lesion: biopsy / surgical excision

Chen, 2017,[23] (China)	Fp-MRI / negative DM Ab-MRI / negative DM	NR (1)	2013 Jan to 2015 Mar	478 / 478 (participants underwent MM, FP-MRI and Ab-MRI)	MM-negative women with dense breasts invited to supplemental screening at a single institution in China Mean age: 49.3	NR	Biopsy and surgical pathology

§ Age shown if reported specifically for people with dense breasts; presented as mean, median, proportions as reported
ϕ Number of people with dense breasts in the MRI plus DM study arm, and in the DM only arm (according to BIRADS categories D and C corresponding to heterogeneously and extremely dense breasts, respectively)

RCT, randomised controlled trial; BIRADS, breast reporting imaging data system; MRI, magnetic resonance imaging; Ab-MRI, abbreviated MRI; Fp-MRI, full protocol MRI; CEM, contrast-enhanced mammography; DENSE, Dense Tissue and Early Breast Neoplasm Screening; BRAID, Breast Screening; Risk Adaptive Imaging for Density; IQR interquartile range; DM digital mammography; EUSOBI, European Society of Breast Imaging; NR, not reported; MM, mammography; DM, digital mammography

Table 4.B Summary of CDR, recall rate, biopsy rate, interval cancers, false negatives, false positive, sensitivity and specificity for studies comparing magnetic resonance imaging (MRI) / abbreviated magnetic resonance imaging plus mammography with mammography-only in people with dense breasts

Study Name Author, year (country)	**Study Question Supplemental screening modality / comparator imaging**	**CDR per 1,000 exams (lower CI, upper CI)**	**Recall rate % (lower CI, higher CI) (definition of recall rate^)**	**PPV1 for recall % (lower CI, upper CI)**	**Biopsy rate % (lower CI, upper CI)**	**PPV3 for biopsy % (lower CI, upper CI)**	**Interval cancers % (lower CI, higher CI)**	**False negatives % (lower CI, higher CI)**	**False positives % (lower CI, higher CI)**	**Sensitivity % (lower CI, upper CI)**	**Specificity % (lower CI, upper CI)**
MA-DETECT Kaiser, 2024[32] (Germany)	Negative DM plus Fp-MRI	20 (5.5, 50.4)	40 (17.4, 77.3) (per 1,000 exams) (definition: positive MRI findings)	50.0 (15.7, 84.3)	NR	NR	NR	NR	NR	NR	NR

BRAID Gilbert, 2024[20] (UK)	Negative DM plus Ab-MRI	17.4 (12.2, 23.9)	9.7 (8.4, 11) (definition: BI-RADS 3, 4, or 5)	NR	4.9 (4, 5.9)	NR	NR	NR	NR	NR	NR
DENSE (2nd Round) Veenhuizen, 2021[39] (Nether-lands)	Negative DM plus Fp-MRI	5.8 (3.8, 9)	32 (26.6, 38.4) (per 1,000 exams) (definition: BIRADS 3,4,5)	18.2 (12.1, 26.4)	24.4 (19.8, 30.2)	23.8 (16, 33.9)	NR	NR	26.3 (21.5, 32.3) per 1,000 ex-ams	NR	NR
DENSE (1st Round) Bakker, 2019[37] (Nether-lands)	Negative DM plus Fp-MRI	16.5 (13.3, 20.5)	9.5 94.9 (86.9, 103.6) (per 1,000 exams) (definition: BI-RADS 3,4,5)	17.4 (14.25, 21.2)	6.3 62.7 (56.2, 70) (per 1,000 exams)	26.3 (21.7, 31.6)	n=4 (n=161 in mam-mogra-phy only arm)	NR	8.0 79.8 (72.4, 87.9) per 1,000 ex-ams	95.2 (88.1, 98.7)	92
Kuhl, 2017[22] (Germany)	Negative DM plus Fp-MRI	n=26 (of 1,282 ex-ams)	NR	NR	NR	NR	n=0	NR	NR	NR	NR
Chen, 2017,[23] (China)	Negative DM plus FP-MRI	n=16 (of 478 ex-ams)	NR	41	NR	NR	NR	NR	NR	100	94.6
	Negative DM plus Ab-MRI	n=15 per 478 ex-ams	NR	27.7	NR	NR	NR	NR	NR	88.3	99.8

^Recall rate definition as reported by study authors

MRI, magnetic resonance imaging; Ab-MRI, abbreviated protocol MRI; Fp-MRI full protocol MRI; DENSE, Dense Tissue and Early Breast Neoplasm Screening; BRAID, Breast Screening; Risk Adaptive Imaging for Density; DM digital mammography; PPV1, positive predictive value (number of cancers diagnosed per number of positive screens); PPV3, positive predictive value (biopsy proven predictive value); CI, confidence interval; NR, not reported; CDR, cancer detection rate; BIRADS breast imaging reporting and data system

Characteristics and results of 'other imaging modalities' plus mammography versus mammography alone are presented below

Table 5.A A Reported characteristics of studies comparing 'other imaging modalities' plus mammography with mammography-only in people with dense

Study name Author, year (country)	Supplementary screening modality / comparator imaging (interval between tests)	Design (number of centres)	Study years (number of centres)	Number with dense breasts in study arm/ number with dense breasts in DM only arm ϕ	Brief population description Age§	Screening programme interval	Reference standard
BRAID Gilbert, 2024[20] (UK)	CEM / DM (Interval median days [IQR]: 134 [91-173])	RCT (10)	2019 Sept to 2024 Mar	2,035 / 6,303	MM-negative people attending UK population breast screening at 10 centres Median (IQR) age: 56 (52-61)	Triennial	Recall cases: further imaging and biopsy if lesion confirmed Where doubt about lesion or not: repeat CEM or MRI/ short term follow up offered
Starikov, 2016[49] (USA)	WBS / DM One third of WBS performed with ABUS machine, two thirds performed with HHUS	Retrospective observational case–control study (1)	2013 Jan to 2013 Dec	1,397 / 7,117	People presenting for screening mammography Majority of patients >40y	Annual	List of biopsy proven screen-detected breast cancers

	(Interval: same day)						
Starikov, 2016[49] (USA)	WBS plus DBT / DM One third of WBS performed with ABUS machine, two thirds performed with HHUS (Interval: same day)	Retrospective observational case–control study (1)	2013 Jan to 2013 Dec	526 / 7,117	People presenting for screening mammography Majority of patients >40y	Annual	List of biopsy proven screen-detected breast cancers

§ Age shown if reported specifically for people with dense breasts; presented as mean, median, proportions as reported
ϕ Number of people with dense breasts in the 'other imaging modality' plus DM study arm, and in the DM only arm (according to BIRADS categories D and C corresponding to heterogeneously and extremely dense breasts, respectively)

WBS, whole breast sonography; DM, digital mammography; RCT, randomised controlled trial; CEM contrast-enhanced mammography; MM, mammography; IQR, interquartile range; MRI, magnetic resonance imaging; HHUS, hand-held ultrasound; ABUS, automated breast ultrasound; DBT digital breast tomosynthesis; BRAID, Breast Screening; Risk Adaptive Imaging for Density; BIRADS, breast imaging reporting and data system

Table 5.B Summary of CDR, recall rate, biopsy rate, interval cancers, false negatives, false positive, sensitivity and specificity for studies comparing magnetic resonance imaging (MRI) / abbreviated magnetic resonance imaging plus mammography with mammography-only in people with dense breasts

Study name Author, year (country)	**Study Question Supplemental screening modality /**	**CDR per 1,000 exams (lower**	**Recall rate % (lower CI, higher CI)**	**PPV1 for recall % (lower**	**Biopsy rate % (lower CI, upper CI)**	**PPV3 for biopsy % (lower**	**Interval cancers % (lower CI,**	**False negatives % (lower CI, higher CI)**	**False positives % (lower CI,**	**Sensitivity % (lower CI, upper CI)**	**Specificity % (lower CI, upper CI)**

	comparator imaging	CI, upper CI)	(definition of recall rate^)	CI, upper CI)		CI, upper CI)	higher CI)		higher CI)		
BRAID Gilbert, 2024[20] (UK)	CEM plus DM negative	19.2 (13.7, 26.1)	9.7 (8.4, 11) (definition BIDRADS 3,4,5)	NR	4.4 (3.5, 5.4)	NR	NR	NR	NR	NR	NR
Starikov, 2016[49] (USA)	WBS plus DM	7.2	20.8 (definition: BIRADS 0)	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	3.8	19.9	NR	NR	NR	NR	NR	NR	NR	NR
Starikov, 2016[49] (USA)	DBT plus WBS plus DM	7.6	23.4 (definition: BIRADS 0)	NR	NR	NR	NR	NR	NR	NR	NR
	DM only	3.8	19.9	NR	NR	NR	NR	NR	NR	NR	NR

^Recall rate definition as reported by study authors

CDR, cancer detection rate; WBS, whole breast sonography; DM, digital mammography; CEM contrast-enhanced mammography; PPV1, positive predictive value (number of cancers diagnosed per number of positive screens); PPV3, positive predictive value (biopsy proven predictive value); CI, confidence interval; DBT, digital breast tomosynthesis; NR, not reported; BRAID, Breast Screening; Risk Adaptive Imaging for Density

Appraisal for quality and risk of bias

Risk of Bias Guidance and summaries are reported below.

Table 6 Guidance for Risk of Bias Assessment (QUADAS-2 and QUADAS-C)

Signalling questions	Guidelines for Rating	Notes
Patient Selection: Risk of Bias		
Q1-1. Was a consecutive or random sample of patients enrolled? (Yes/No/Unclear)	1. When "consecutive" or "random" sampling was explicitly described: YES 2. When "convenience sampling" or "volunteer enrolment" was described: NO 3. All other descriptions: UNCLEAR	
Q1-2. Was a case-control (two-gate) design avoided? (Yes/No/Unclear)	Not applicable, as case-control studies excluded as per eligibility criteria	
Q1-3. Did the study avoid inappropriate exclusions? (Yes/No/Unclear)	1. All eligible participants were included who would normally attend screening programmes: YES 2. Participants were excluded based on other reasons: NO 3. Lack of detailed information about participant (or image) selection after enrollment: UNCLEAR	*Acceptable exclusions* were those inherent to the screening programme, such as breast implants, symptoms, high risk (e.g. high-risk gene), family history, age. Studies that include *only people with dense breasts* in the analysis are also acceptable
Overall rating: Could the selection of patients (or input images) have introduced bias? (Low/High/Unclear)	1. All "Yes": LOW 2. One or two "Unclear": UNCLEAR 3. One "No" or more than three "Unclear": HIGH	
Patient Selection: Concerns Regarding Applicability		
Q1-4. Is there concern that the included patients do not match the review question? (Low/High/Unclear)	Were the participant characteristics comparable to those encountered in the screening setting where the index test was intended to be used? Yes: LOW No: HIGH Not sure: UNCLEAR	Participants must be from a population screening cohort and not a selected group (aside from a dense breast cohort), and should not be selected or volunteer for supplementary imaging. We did not expect the age profile to match the UK profile exactly, but studies with participants who were outside the UK range were deemed to be high risk of bias, along with those with a high proportion outside the UK age range. While data from people with dense breasts only were included, studies of participants with extremely dense breasts only were considered to have

		high risk of bias in terms of applicability.
Patient Selection: Comparative Accuracy		
C1-1. Was the risk of bias for each index test judged “low” for this domain? (Yes/No)		Applicable to "within-study" or "between-study" comparative studies
C1-2. Was a fully paired or randomized design used? (Yes/No/Unclear)	1. Either 1) fully paired or randomised or 2) not randomised but evidence of absence of imbalance of key confounders; YES 2. Not enough details reported: UNCLEAR 3. Evidence of imbalanced confounders and no statistical adjustment: NO	i.e. "within-study" comparative studies in which all participants received both index tests and the reference test; or, participants were randomised to receiving either index test and the reference test
C1-3. Was the allocation sequence random? (Yes/No/Unclear/Not applicable)		Only applicable to randomised designs
C1-4. Was the allocation sequence concealed until patients were enrolled and assigned to index tests? (Yes/No/Unclear/Not applicable)		Only applicable to randomised designs
Overall rating: Could the selection of patients have introduced bias in the comparison? (Low/High/Unclear)	1. All "Yes": LOW 2. One or two "Unclear": UNCLEAR 3. One "No" or more than three "Unclear": HIGH	
Index Tests: Risk of Bias		
Q2-1. Were the index test results interpreted without knowledge of the results of the reference standard? (Yes/No/Unclear)		All studies deemed to be of low risk as image readers will not know results of future verification tests.
Q2-2. If there was a threshold, was it pre-specified?	1. Yes: LOW 2. No description or mention: UNCLEAR 3. Mentions that various cut-offs were explored; NO	We expect that all studies on breast cancer screening use an accepted threshold (e.g. BI-RADS category or equivalent) for judgement
Overall rating: Could the conduct or interpretation of the index test have introduced bias? (Low/High/Unclear)	1. All "Yes”: LOW 2. One or two "Unclear": UNCLEAR 3. One "No" or more than three "Unclear": HIGH	
Index Test: Concerns regarding Applicability		
The index tests are clearly MM plus additional modality, and MM only (Low/High/Unclear)		Not applicable as studies that are not index tests plus MM versus MM alone were excluded as per study eligibility criteria
Index test: Comparative Accuracy		
Is there concern that the index test, its conduct, or inter-		As above, strict eligibility criteria meant that only studies comparing index tests plus MM versus MM alone were included

pretation differ from the review question? (Low/High/Unclear)		
Reference Standard: Risk of Bias		
Q3-1. Is the reference standard likely to correctly classify the target condition? (Yes/No/Unclear)	1. Reference standard is well-described including standard biopsy for positive cases; surveillance, linkage with cancer registries/databases, clinical evaluation for negative cases as appropriate for negative cases: YES 2. Additional capture method of index test negatives is not reported: UNCLEAR 3. NO: unclear whether repeat imaging after 1 or more years	We expect that breast screening programmes will have reference standard as routine; however, this should be described in the study. We will not rate RoB as HIGH for this item since bias with repeat imaging as reference standard (whether single or combined) is likely to be small.
Q3-2. Were the reference standard results interpreted without knowledge of the results of the index test? (Yes/No/Unclear)		Not applicable as due to screening nature of the study, results will be known (e.g. positive MM result will follow cancer pathway)
Overall rating: Could the reference standard, its conduct, or its interpretation have introduced bias? (Low/High/Unclear)	1. All "Yes": LOW 2. One or two "Unclear": UNCLEAR 3. One "No" or more than three "Unclear": HIGH	
Reference Standard: Concerns Regarding Applicability		
q3-3. Is there concern that the target condition as defined by the reference standard does not match the review question? (Low/High/Unclear)		Where no information given on reference standard: Unclear
Reference Standard: Comparative Accuracy		
c3-1. Was the risk of bias for each index test judged “low” for this domain? (Yes/No)	All "Yes": LOW All others: NO	
c3-2. Did the reference standard avoid incorporating any of the index tests? (Yes/No/Unclear)	Not applicable since all studies will use repeated imaging as a reference standard in people not receiving biopsy,	
Flow and Timing: Risk of Bias		

Q4-1. Was there an appropriate interval between index test(s) and reference standard? (Yes/No/Unclear)	1. Description of time frame depending on participant and setting :YES 2. No information: UNCLEAR 3. More than 1 month between index test and biopsy for test positives; or more than the standard adopted in each national context for test negatives	We did not define an ideal interval for conducting reference standard (e.g. biopsy or imaging), accepting that this may be specific to the institution/country In some cases, studies reported the nature of the reference standard but did not describe the interval between index tests and reference standard; or the interval for one aspect of reference standard was described but not others. We expected that the interval would be reported, and lack of information was deemed unclear
Q4-2. Did all patients receive a reference standard? (Yes/No/Unclear)	1. All study participants received a reference standard: YES 2. Not all study participants received a reference standard, and this was related to DM results: NO 3. Lack of detailed information: Unclear	If some participants did not receive appropriate reference standard (dropouts, loss to follow-up etc), this was deemed to be high risk of bias If there was insufficient detail this was deemed to be unclear
Q4-3. Were all patients included in the analysis? (Yes/No/Unclear)	1. All included patients were analysed: YES 2. There were missing data: NO 3. Uncertainty or lack of detailed information: UNCLEAR	
Overall rating: Could the patient flow have introduced bias? (Low/High/Unclear)	All "Yes": LOW One, two, or three "Unclear": UNCLEAR At least one “No”: NO	
Flow and Timing: Comparative Analysis		
C4-1. Was the risk of bias for each index test judged “low” for this domain? (Yes/No)		
C4-2. Was there an appropriate interval between the index tests? (Yes/No/Unclear)		We expected that the comparator (MM) and supplementary imaging would be conducted within a reasonable timeframe. The reported interval varied according to the nature of the imaging (with some being done simultaneously on the same machine). We did not specify an ideal timeframe, however, where the supplementary test was conducted up to several months later, this was deemed to be a high risk of bias. Where the interval was not reported we deemed this to be an unclear risk of bias

C4-3. Was the same reference standard used for all index tests? (Yes/No/Unclear)		This was not applicable
C4-4. Are the proportions and reasons for missing data similar across index tests? (Yes/No/Unclear)		
Overall rating: Could the patient flow have introduced bias in the comparison? (Low/High/Unclear)	Both “Yes”: LOW At least one “Unclear”: UNCLEAR At least one “No”: NO	

MM, mammography; BI-RADS, Breast Imaging Reporting and Data System

Appendix 4 – UK NSC reporting checklist for evidence summaries

All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table xx.

Table 7 UK NSC reporting checklist for evidence summaries

Section	Item	Page no.
Title and summaries		
Title Sheet	Identify the review as a UK NSC Evidence summary	Title page
Plain English summary	Plain English description of the executive summary.	5
Executive summary	Structured overview of the whole report. To include: the purpose/aim of the review; background; previous recommendations; findings and gaps in the evidence; recommendations on the screening that can or cannot be made on the basis of the review	7
Introduction and Approach		

Section	Item	Page no.
Background and objectives	Background – Current policy context and rationale for the current review – for example, reference to details of previous reviews, basis for current recommendation, recommendations made, gaps identified, drivers for new reviews	7
	Objectives – What are the questions the current evidence summary intends to answer? – statement of the key questions for the current evidence summary, criteria they address, and number of studies included per question, description of the overall results of the literature search.	15
	Method – briefly outline the rapid review methods used.	16
Eligibility for inclusion in the review	State all criteria for inclusion and exclusion of studies to the review clearly(PICO, dates, language, study type, publication type, publication status etc.) To be decided a priori	17
Appraisal for quality/ risk of bias tool	Details of tool/ checklist used to assess quality, e.g. QUADAS 2, CASP, SIGN, AMSTAR.	20

Section	Item	Page no.
Search strategy and study selection		
Databases/ sources searched	Give details of all databases searched (including platform/ interface and coverage dates) and date of final search.	54
Search strategy and results	Present the full search strategy for at least one database(usually a version of Medline), including limits and search filters if used.	54
	Provide details of the total number of (results from each database searched), number of duplicates removed, and the final number of unique records to consider for inclusion.	55
Study selection	State the process for selecting studies – inclusion and exclusion criteria, number of studies screened by title/abstract and full text, number of reviewers, any cross checking carried out.	17
Study level reporting of results (for each key question)		

Section	Item	Page no.
Study level reporting, results and risk of bias assessment	For each study, produce a table that includes the full citation and a summary of the data relevant to the question (for example, study size, PICO, follow-up period, outcomes reported, statistical analyses etc.).	19
	Provide a simple summary of key measures, effect estimates and confidence intervals for each study where available.	
	For each study, present the results of any assessment of quality/risk of bias.	26 / 28 / 75
Additional analyses	Describe additional analyses (for example, sensitivity, specificity, PPV, etc.) carried out by the reviewer. [Remove if not performed]	28
Question level synthesis		
Description of the evidence	For each question, give numbers of studies screened, assessed for eligibility, and inclusion in the review, with summary reasons for exclusion	21

Section	Item	Page no.
Combining and presenting the findings	Provide a balanced discussion of the body of evidence which avoids over reliance on one study or set of studies. Consideration of four compartments should inform the reviewer's judgement on whether the criterion is "met", "not met" or "uncertain": quantity; quality; applicability and consistency.	47
Summary of findings	Provide a description of the evidence reviewed and included for each question, with reference to their eligibility for inclusion. Summarise the main findings including the quality/ risk of bias issues for each question. Have the criteria addressed been "met", "not met" or "uncertain"?	47
Review Summary		
Conclusions and implications for policy	Do findings indicate whether screening should be recommended? IS further work warranted? Are there gaps in the evidence highlighted by the review?	52
Limitations	Discuss limitations of the available evidence and of the review methodology if relevant.	50

References

1. Cancer Research UK. Cancer statistics for the UK. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk (Accessed January 2025).

2. Lauby-Secretan B, Loomis D, Straif K. Breast cancer screening: viewpoint of the IARC Working Group. N Engl J Med. 2015;**373**(15):1479.

3. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet. 2012;**380**(9855):1778-86.

4. Duffy SW, Tabár L, Yen AM, et al. Mammography screening reduces rates of advanced and fatal breast cancers: results in 549,091 women. Cancer. 2020;**126**(13):2971-9.

5. Edmonds CE, O'Brien SR, Conant EF. Mammographic breast density: current assessment methods, clinical implications, and future directions. Semin Ultrasound CT MR. 2023;**44**(1):35-45.

6. American College of Radiology. ACR BI-RADS Atlas, Breast Imaging Reporting and Data System 2013. 2013. Available from: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-Rads (Accessed January 2025).

7. Sprague BL, Kerlikowske K, Bowles EJA, et al. Trends in Clinical Breast Density Assessment from the Breast Cancer Surveillance Consortium. J Natl Cancer Inst. 2019;**111**(6):629-32.

8. Katanoda K, Ito Y, Sobue T. International comparison of trends in cancer mortality: Japan has fallen behind in screening-related cancers. Jpn J Clin Oncol. 2021;**51**(11):1680-6.

9. Uematsu T. Rethinking screening mammography in Japan: next-generation breast cancer screening through breast awareness and supplemental ultrasonography. Breast Cancer. 2024;**31**(1):24-30. doi: 10.1007/s12282-023-01506-w.

10. Patterson J, Stinton C, Alkhudairy L, et al. UK NSC Evidence Review. Additional Breast Screening with Ultrasound in Women with Dense Breasts Final 27 02 2019 2019. 2019. Available from: https://view-health-screening-recommendations.service.gov.uk/document/465/download. (Accessed January 2025)

11. Mann RM, Athanasiou A, Baltzer PAT, et al. Breast cancer screening in women with extremely dense breasts recommendations of the European Society of Breast Imaging (EUSOBI). Eur Radiol. 2022;**32**(6):4036-45.

12. Federal Register. Mammography Quality Standards. 2023. Available from: https://www.federalregister.gov/documents/2023/03/10/2023-04550/mammography-quality-standards-act (Accessed January 2025).

13. Clinical Trials.Gov. Breast Screening - Risk Adaptive Imaging for Density (BRAID) Protocol ID NCT04097366. 2019. Available from: https://www.clinicaltrials.gov/study/NCT04097366 (Accessed January 2025).

14. Higgins J, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023) 2024. Available from: www.training.cochrane.org/handbook. (Accessed December 2024)

15. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;**372**:n71.

16. Thomas, J., Graziosi, S., Brunton, J., Ghouze, Z., O'Driscoll, P., & Bond, M. & Koryakina, A. (2023) EPPI-Reviewer: advanced software for systematic reviews, maps and evidence synthesis. EPPI Centre, UCL Social Research Institute, University College London

17. Trial Forge. PRO EDI participant characteristics table 22/3/2024. 2024. Available from: https://www.trialforge.org/trial-diversity/pro-edi-improving-how-equity-diversity-and-inclusion-is-handled-in-evidence-synthesis/ (Accessed October 2024).

18. Whiting PF, Rutjes AWS, Westwood ME, et al. Quadas-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;**155**(8):529-36.

19. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;**327**(7414):557-60.

20. Gilbert FJ. Breast Screening - Risk Adaptive Imaging for Density. BRAID trial results (personal communication: Brazzelli, M. 2024). 2024.

21. Osteras BH, Martinsen ACT, Gullien R, Skaane P. Digital mammography versus breast tomosynthesis: impact of breast density on diagnostic performance in population-based screening. Radiology. 2019;**293**(1):60-8.

22. Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental Breast MR Imaging Screening of Women with Average Risk of Breast Cancer. Radiology. 2017;**283**(2):361-370. doi: 10.1148/radiol.2016161444.

23. Chen SQ, Huang M, Shen YY, Liu CL, Xu CX. Application of Abbreviated Protocol of Magnetic Resonance Imaging for Breast Cancer Screening in Dense Breast Tissue. Acad

Radiol. 2017;**24**(3):316-320. doi: 10.1016/j.acra.2016.10.003. Epub 2016 Dec 1. PMID: 27916594.

24. Kwon M-R, Choi JS, Lee MY, et al. Screening outcomes of supplemental automated breast US in Asian women with dense and nondense breasts. Radiology. 2023;**307**(4):e222435.

25. Nakamura A, Ohnuki K, Takahashi H, et al. The effects of breast density on the benefits of mammograms with adjunctive ultrasonography in breast screening. Breast Cancer. 2024;**31**(2):228-33.

26. Ren W, Yan H, Zhao X, et al. Integration of handheld ultrasound or automated breast ultrasound among women with negative mammographic screening findings: a multi-center population-based study in China. Acad Radiol. 2023;**30**(Supplement 2):S114-S26.

27. Lee SE, Yoon JH, Son N-H, Han K, Moon HJ. Screening in patients with dense breasts: comparison of mammography, artificial intelligence, and supplementary ultrasound. AJR Am J Roentgenol. 2024;**222**(1):e2329655.

28. Pulido-Carmona C, Romero-Martín S, Raya-Povedano JL, et al. Interval cancer in the Córdoba Breast Tomosynthesis Screening Trial (CBTST): comparison of digital breast tomosynthesis plus digital mammography to digital mammography alone. Eur Radiol. 2024.

29. Rani VR, Reddy BN. Diagnostic performance in detecting breast cancer on combined mammography and ultrasound. J Cardiovasc Dis Res. 2023;**14**(5):1649-52.

30. Olinder J, Johnson K, Akesson A, Fornvik D, Zackrisson S. Impact of breast density on diagnostic accuracy in digital breast tomosynthesis versus digital mammography: results from a European screening trial. Breast Cancer Res. 2023;**25**(1):116.

31. Ha SM, Jang M-J, Youn I, et al. Screening outcomes of mammography with AI in dense breasts: a comparative study with supplemental screening US. Radiology. 2024;**312**(1):e233391.

32. Kaiser C, Wilhelm T, Walter S, Singer S, Keller E, Baltzer PAT. Cancer detection rate of breast-MR in supplemental screening after negative mammography in women with dense breasts. Preliminary results of the MA-DETECT-Study after 200 participants. Eur J Radiol. 2024;**176**:111476.

33. Gatta G, Cappabianca S, La Forgia D, et al. Second-generation 3D automated breast ultrasonography (Prone ABUS) for dense breast cancer screening integrated to mammography: effectiveness, performance and detection rates. J Pers Med. 2021;**11**(9).

34. Buchberger W, Geiger-Gritsch S, Knapp R, Gautsch K, Oberaigner W. Combined screening with mammography and ultrasound in a population-based screening program. Eur J Radiol. 2018;**101**:24-9.

35. Bernardi D, Macaskill P, Pellegrini M, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. Lancet Oncol. 2016;**17**(8):1105-13.

36. 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;**303**(2):256-66.

37. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. N Engl J Med. 2019;**381**(22):2091-102.

38. Upadhyay N, Soneji N, Stewart V, Ralleigh G. The effect of the addition of tomosynthesis to digital mammography on reader recall rate and reader confidence in the UK prevalent screening round. Clin Radiol. 2018;**73**(8):744-9.

39. Veenhuizen SGA, de Lange SV, Bakker MF, et al. Supplemental breast MRI for women with extremely dense breasts: results of the second screening round of the DENSE trial. Radiology. 2021;**299**(2):278-86.

40. Tagliafico AS, Calabrese M, Mariscotti G, et al. Adjunct screening with tomosynthesis or ultrasound in women with mammography-negative dense breasts: Interim report of a prospective comparative trial. J Clin Oncol. 2016;**34**(16):1882-8.

41. Johnson K, Zackrisson S, Rosso A, et al. Tumor characteristics and molecular subtypes in breast cancer screening with digital breast tomosynthesis: the Malmö Breast Tomosynthesis Screening Trial. Radiology. 2019;**293**(2):273-81.

42. Wilczek B, Wilczek HE, Rasouliyan L, Leifland K. Adding 3D automated breast ultrasound to mammography screening in women with heterogeneously and extremely dense breasts: report from a hospital-based, high-volume, single-center breast cancer screening program. Eur J Radiol. 2016;**85**(9):1554-63.

43. McDonald ES, Oustimov A, Weinstein SP, Synnestvedt MB, Schnall M, Conant EF. Effectiveness of digital breast tomosynthesis compared with digital mammography: outcomes analysis from 3 years of breast cancer screening. JAMA oncology. 2016;**2**(6):737-43.

44. McDonald ES, McCarthy AM, Akhtar AL, Synnestvedt MB, Schnall M, Conant EF. Baseline screening mammography: performance of full-field digital mammography versus digital breast tomosynthesis. AJR Am J Roentgenol. 2015;**205**(5):1143-8.

45. Freer PE, Riegert J, Eisenmenger L, et al. Clinical implementation of synthesized mammography with digital breast tomosynthesis in a routine clinical practice. Breast Cancer Res Treat. 2017;**166**(2):501-9.

46. Durand MA, Friedewald SM, Plecha DM, et al. False-negative rates of breast cancer screening with and without digital breast tomosynthesis. Radiology. 2021;**298**(2):296-305.

47. Conant EF, Beaber EF, Sprague BL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography compared to digital mammography alone: a cohort study within the PROSPR consortium. Breast Cancer Res Treat. 2016;**156**(1):109-16.

48. Rose SL, Shisler JL. Tomosynthesis impact on breast cancer screening in patients younger than 50 years old. AJR Am J Roentgenol. 2018;**210**(6):1401-4.

49. Starikov A, Drotman M, Hentel K, Katzen J, Min RJ, Arleo EK. 2D mammography, digital breast tomosynthesis, and ultrasound: which should be used for the different breast densities in breast cancer screening? Clin Imaging. 2016;**40**(1):68-71.

50. Brem RF, Tabar L, Duffy SW, et al. Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue: the SomoInsight Study. Radiology. 2015;**274**(3):663-73.

51. Pang JX, Newsome J, Sun M, et al. Impact of switching from digital mammography to tomosynthesis plus digital mammography on breast cancer screening in Alberta, Canada. J Med Screen. 2022;**29**(1):38-43.

52. Stepanek T, Constantinou N, Marshall H, et al. Changes in the utilization of the BI-RADS Category 3 assessment in recalled patients before and after the implementation of screening digital breast tomosynthesis. Acad Radiol. 2019;**26**(11):1515-25.

53. Ban K, Tsunoda H, Togashi S, et al. Breast cancer screening using digital breast tomosynthesis compared to digital mammography alone for Japanese women. Breast Cancer. 2021;**28**(2):459-64.

54. Chang JM, Koo HR, Moon WK. Radiologist-performed hand-held ultrasound screening at average risk of breast cancer: results from a single health screening center. Acta Radiol. 2015;**56**(6):652-8.

55. Harada-Shoji N, Suzuki A, Ishida T, et al. Evaluation of adjunctive ultrasonography for breast cancer detection among women aged 40-49 years with varying breast density undergoing screening mammography: a secondary analysis of a randomized clinical trial. JAMA Netw Open. 2021;**4**(8):e2121505.

56. Alonzo TA, Brinton JT, Ringham BM, Glueck DH. Bias in estimating accuracy of a binary screening test with differential disease verification. Stat Med. 2011;**30**(15):1852-64.

57. Hussein H, Abbas E, Keshavarzi S, et al. Supplemental breast cancer screening in women with dense breasts and negative mammography: a systematic review and meta-analysis. Radiology. 2023;**306**(3):e221785.

58. Burnside ES, Vulkan D, Blanks RG, Duffy SW. Association between screening mammography recall rate and interval cancers in the UK Breast Cancer Service Screening Program: a cohort study. Radiology. 2018;**288**(1):47-54.

59. NHS England. Breast screening programme: standards. 2017. Available from: https://www.gov.uk/government/publications/breast-screening-consolidated-programme-standards (Accessed November 2024).

60. NHS England. NHS Breast screening programme screening standards valid for data collected from 1 April 2021. 2024. Available from: https://www.gov.uk/government/publications/breast-screening-consolidated-programme-standards/nhs-breast-screening-programme-screening-standards-valid-for-data-collected-from-1-april-2021#bsp-s13-referral-positive-predictive-value-of-referral (Accessed February 2025).

61. Chen SQ, Huang M, Shen YY, Liu CL, Xu CX. Abbreviated MRI protocols for detecting breast cancer in women with dense breasts. Korean J Radiol. 2017;**18**(3):470-5.

62. Clinical Trials.Gov. Comparison of Breast Cancer Screening With CESM to DBT in Women With Dense Breasts (CMIST). 2022. Available from: https://clinicaltrials.gov/study/NCT05625659 (Accessed 5 February 2025).

63. Clinical Trials.Gov. The C-MERIT Screening Cohort: Contrast-enhanced Mammography for Breast Cancer Screening and Risk Assessment in Women With Dense Breasts. 2022. Available from: https://www.clinicaltrials.gov/study/NCT05667532 (Accessed 5 February 2025).