JAMA | US Preventive Services Task Force | EVIDENCE REPORT

# Screening for Breast Cancer
## Evidence Report and Systematic Review for the US Preventive Services Task Force

Jillian T. Henderson, PhD, MPH; Elizabeth M. Webber, MS; Meghan S. Weyrich, MPH; Marykate Miller, MS; Joy Melnikow, MD, MPH

**IMPORTANCE** Breast cancer is a leading cause of cancer mortality for US women. Trials have established that screening mammography can reduce mortality risk, but optimal screening ages, intervals, and modalities for population screening guidelines remain unclear.

**OBJECTIVE** To review studies comparing different breast cancer screening strategies for the US Preventive Services Task Force.

**DATA SOURCES** MEDLINE, Cochrane Library through August 22, 2022; literature surveillance through March 2024.

**STUDY SELECTION** English-language publications; randomized clinical trials and nonrandomized studies comparing screening strategies; expanded criteria for screening harms.

**DATA EXTRACTION AND SYNTHESIS** Two reviewers independently assessed study eligibility and quality; data extracted from fair- and good-quality studies.

**MAIN OUTCOMES AND MEASURES** Mortality, morbidity, progression to advanced cancer, interval cancers, screening harms.

**RESULTS** Seven randomized clinical trials and 13 nonrandomized studies were included; 2 nonrandomized studies reported mortality outcomes. A nonrandomized trial emulation study estimated no mortality difference for screening beyond age 74 years (adjusted hazard ratio, 1.00 [95% CI, 0.83 to 1.19]). Advanced cancer detection did not differ following annual or biennial screening intervals in a nonrandomized study. Three trials compared digital breast tomosynthesis (DBT) mammography screening with digital mammography alone. With DBT, more invasive cancers were detected at the first screening round than with digital mammography, but there were no statistically significant differences in interval cancers (pooled relative risk, 0.87 [95% CI, 0.64-1.17]; 3 studies [n = 130 196]; $I^2$ = 0%). Risk of advanced cancer (stage II or higher) at the subsequent screening round was not statistically significant for DBT vs digital mammography in the individual trials. Limited evidence from trials and nonrandomized studies suggested lower recall rates with DBT. An RCT randomizing individuals with dense breasts to invitations for supplemental screening with magnetic resonance imaging reported reduced interval cancer risk (relative risk, 0.47 [95% CI, 0.29-0.77]) and additional false-positive recalls and biopsy results with the intervention; no longer-term advanced breast cancer incidence or morbidity and mortality outcomes were available. One RCT and 1 nonrandomized study of supplemental ultrasound screening reported additional false-positives and no differences in interval cancers.

**CONCLUSIONS AND RELEVANCE** Evidence comparing the effectiveness of different breast cancer screening strategies is inconclusive because key studies have not yet been completed and few studies have reported the stage shift or mortality outcomes necessary to assess relative benefits.

Editorial
Multimedia
Related articles and JAMA Patient Page
Supplemental content
CME at jamacmelookup.com
Related articles at jamaoncology.com jamanetworkopen.com

**Author Affiliations:** Kaiser Permanente Evidence-based Practice Center, Center for Health Research, Portland, Oregon (Henderson, Webber); University of California Davis Center for Healthcare Policy and Research, Sacramento (Weyrich, Miller, Melnikow).

**Corresponding Author:** Jillian T. Henderson, PhD, MPH, Kaiser Permanente Evidence-based Practice Center, Center for Health Research, Kaiser Permanente Northwest, 3800 N Interstate Ave, Portland, OR 97227 (Jillian.T.Henderson@kpchr.org).

*JAMA*. doi:10.1001/jama.2023.25844
Published online April 30, 2024.

Breast cancer is the second leading cause of cancer mortality for US women, despite a steady overall decline in breast-cancer mortality rates over the past 20 years.[1] The average age-adjusted rate for the years 2016-2020 was 19.6 per 100 000, with an estimated 43 170 deaths in 2023.[1,2] The majority of cases occur between the ages of 55 and 74 years,[1] and incidence is highest among women ages 70 to 74 (468.2 per 100 000).[3] Non-Hispanic White women have the highest breast cancer incidence,[4] but mortality is 40% higher for non-Hispanic Black women (27.6 per 100 000) compared with White women (19.7 per 100 000); non-Hispanic Black women experience lower 5-year survival regardless of the cancer subtype or stage at the time of detection.[1,5-7]

Previous reviews of breast cancer screening effectiveness established the benefits and harms of mammography based primarily on large, long-term trials.[8,9] In 2016, the US Preventive Services Task Force (USPSTF) recommended screening for breast cancer in women starting at age 50 years every 2 years continuing through age 74 years (B recommendation) and that screening from ages 40 to 49 years should be based on clinical discussions of patient preferences and individual breast cancer risk (C recommendation).[10] This comparative effectiveness systematic review of breast cancer screening strategies was conducted concurrently with a separate decision modeling study.[11] Both informed the USPSTF updated breast cancer screening recommendations.[12]

## Methods

### Scope of Review

This review addressed 3 key questions (KQs) on the comparative effectiveness and harms of different screening strategies (**Figure 1**). Methodological details including study selection, a list of excluded studies, detailed study-level results for all outcomes and for specific subpopulations, and contextual observations are available in the full evidence report.[14]

### Data Sources and Searches

Studies included in the 2016 USPSTF reviews[8,9,15,16] were evaluated for inclusion with eligibility criteria for the current review. In addition, database searches for relevant studies published between January 2014 and August 22, 2022, were conducted in MEDLINE, the Cochrane Central Register of Controlled Clinical Trials, and the Cochrane Database of Systematic Reviews (eMethods in the Supplement). Reference lists of other systematic reviews were searched to identify additional relevant studies. ClinicalTrials.gov was searched for relevant ongoing trials. Ongoing surveillance to identify newly published studies was conducted through March 2024 to identify major studies published in the interim. Two new nonrandomized studies were identified[17,18] and are not further discussed, as they would not change interpretation of the review findings or conclusions.

### Study Selection

Two independent reviewers screened titles, abstracts, and relevant full-text articles to ensure consistency with a priori inclusion and exclusion criteria (eTable 1 in the Supplement). We included English-language studies of asymptomatic screening populations not at high risk for breast cancer. The eligible population for this review is adult females (sex assigned at birth). For consistency with the underlying evidence, the term "women" is used throughout this report; however, cancer registries and studies of breast cancer generally infer gender based on physiology and medical history rather than measuring self-reported gender. Included studies compared mammography screening modalities (mammography with or without digital breast tomosynthesis [DBT]), different screening strategies with respect to interval, age to start, age to stop, or supplemental screening strategies using ultrasound or magnetic resonance imaging (MRI) with mammography.

For KQ1, randomized clinical trials (RCTs) or nonrandomized studies of interventions with contemporaneous comparison groups that reported breast cancer morbidity, mortality, all-cause mortality, or quality of life were included. For KQ2, the primary outcome of interest was progression to advanced breast cancer, defined for this review as stage IIB or higher, which encompasses tumors with local lymph node involvement or distant metastases.[19] Study-defined advanced breast cancer outcomes were used when this outcome was not reported (eg, stage II or higher). Invasive breast cancer detection outcomes from multiple screening rounds can indicate whether a screening modality or strategy reduces the risk of advanced cancer by detecting early cancers that would otherwise have progressed (stage shift), thereby potentially reducing breast cancer morbidity and mortality.[20-23]

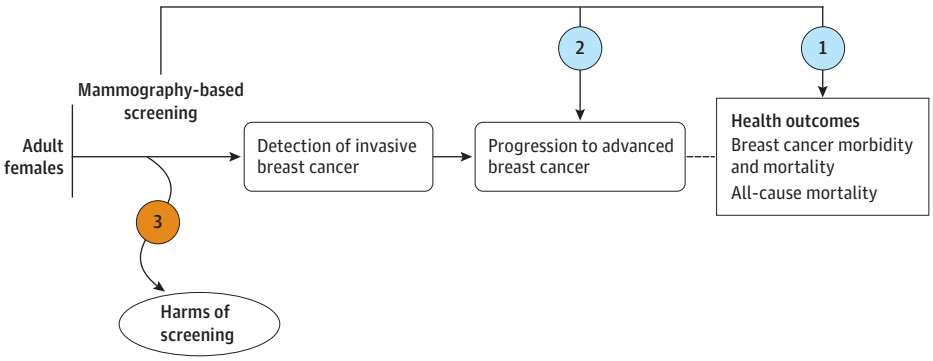
For KQ3, RCTs and nonrandomized studies of interventions reporting adverse events, including psychological harms, radiation exposure, and interval invasive cancers (incident or missed due to false-negative screening) were included, regardless of the number of screening rounds reported. False-positive recall, false-positive biopsy recommendation, and false-positive biopsy rates (individuals who underwent a biopsy for a benign lesion) were obtained from included RCTs and from nonrandomized studies reporting cumulative rates of these potential harms of screening.

### Data Extraction and Quality Assessment

Two reviewers evaluated all articles that met inclusion criteria using prespecified quality criteria (eTable 2 in the Supplement). Discordant quality ratings were resolved through discussion and input from a third reviewer. Risk-of-bias assessment was conducted using the USPSTF-specific criteria for randomized trials[13] and an adapted tool from the Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I).[24] Studies determined to be at high risk of bias were excluded. One reviewer extracted key elements of included studies into standardized evidence tables in DistillerSR (Evidence Partners) and a second reviewer checked the data for accuracy. Limited evidence on sub-KQs is available in the full report.[14] When available, reported relative risks were provided in the tables, but we calculated and reported crude effect estimates and confidence intervals when studies did not provide them. For KQ2 intermediate detection outcomes, the definition of advanced cancer reported in the studies was used for synthesis; commonly this was stage II or later. Comparisons of prognostic characteristics or markers (eg, grade, tumor size, nodal involvement, receptor status) were included for comparisons as data allowed.

All quantitative analyses were conducted in Stata version 16 (StataCorp). The presence of statistical heterogeneity was assessed among pooled studies using the $I^2$ statistic. Where effects were sufficiently consistent and clinical and statistical heterogeneity low, random-effects meta-analyses were conducted using the

Figure 1. Analytic Framework: Screening for Breast Cancer




Key questions

1. What is the comparative effectiveness of different mammography-based breast cancer screening strategies (eg, by modality, interval, initiation and stopping age, use of supplemental imaging, or personalization based on risk factors) on breast cancer morbidity and mortality?
2. What is the comparative effectiveness of different mammography-based breast cancer screening strategies (eg, by modality, interval, initiation and stopping age, use of supplemental imaging, or personalization based on risk factors) on the incidence of and progression to advanced breast cancer?
3. What are the comparative harms of different mammography-based breast cancer screening strategies (modality, interval, initiation age, use of supplemental imaging, or personalization based on risk factors)?

Evidence reviews for the USPSTF use an analytic framework to visually display the key questions that the review will address in order to allow the USPSTF to evaluate the effectiveness and safety of a preventative service. The questions are depicted by linkages that relate interventions and outcomes. A dashed line indicates a health outcome that immediately follows an intermediate outcome. For additional details see the US Preventive Services Task Force Procedure Manual.[13]

restricted maximum likelihood; all tests were 2-sided, with $P < .05$ indicating statistical significance.

Aggregate strength of evidence (ie, high, moderate, or low) was assessed for each KQ and comparison using the approach described in the Methods Guide for the Effectiveness and Comparative Effectiveness Reviews,[25] based on consistency, precision, publication bias, and study quality.

## Results

Investigators reviewed 10 378 unique citations and 419 full-text articles for all KQs (**Figure 2**). Twenty studies reported in 45 publications were included.[26-45] A full list of included studies by KQ is located in eTable 3 in the Supplement.

### Health Benefits of Screening

**Key Question 1.** What is the comparative effectiveness of different mammography-based breast cancer screening strategies (eg, by modality, interval, initiation and stopping age, use of supplemental imaging, or personalization based on risk factors) on breast cancer morbidity and mortality?

Two nonrandomized studies reported on the association of different screening programs with breast cancer morbidity and mortality. One study was designed to compare different ages to stop screening[30] and another compared annual and triennial screening intervals.[41]

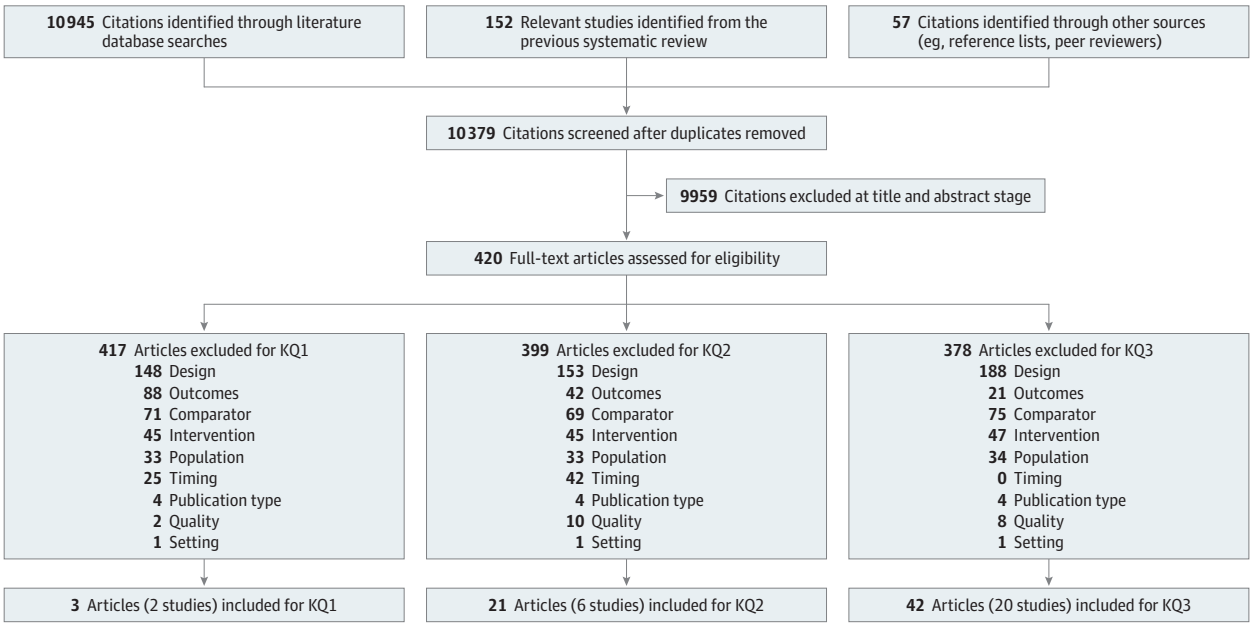
A fair-quality observational study (n = 1 058 013) on age to stop screening used an emulated trial methodology to analyze a random sample of US Medicare A and B claims data for enrollees aged 70 to 84 years (1999 to 2008), eligible for breast cancer screening, and with at least a 10-year estimated life expectancy. The study estimated the effect of stopping screening at ages 70, 75, and 80 years compared with continued annual screening.[30,46] Continuation of screening between the ages of 70 and 74 years was associated with reduced mortality risk based on survival analysis (hazard ratio, 0.78 [95% CI, 0.63 to 0.95]), but the absolute difference in the risk of death for the age group was small and the confidence interval included null (1.0 fewer deaths per 1000 screened [95% CI, −2.3 to 0.1]). These results indicate a difference in the cumulative incidence curves that approached a difference in the mortality risk for the age group. Conversely, continued screening vs no screening from ages 75 to 84 years did not result in statistically significant differences in the absolute risk of breast cancer mortality (0.07 fewer deaths per 1000 [95% CI, −0.93 to 1.3]) or the cumulative mortality incidence (hazard ratio, 1.00 [95% CI, 0.83 to 1.19]).

A fair-quality nonrandomized clinical study (n = 14 765) conducted in Finland during the years 1985 to 1995 assigned participants aged 40 to 49 years to annual or triennial screening invitations by alternating birth year.[41] The study reported no difference in breast cancer mortality: 20.3 deaths per 100 000 person-years with annual screening invitations and 17.9 deaths per 100 000 person-years with triennial screening invitations (relative risk [RR], 1.14 [95% CI, 0.59-1.27]).

### Prevention of Cancer Progression (Intermediate Outcome)

**Key Question 2.** What is the comparative effectiveness of different mammography-based breast cancer screening strategies (eg, by modality, interval, initiation and stopping age, use of supplemental imaging, or personalization based on risk factors) on the incidence of and progression to advanced breast cancer?

Figure 2. Literature Search Flow Diagram: Screening for Breast Cancer




Reasons for exclusion: Design: Study did not use an included design. Outcomes: Study did not have relevant outcomes or had incomplete outcomes. Comparator: Study used an excluded comparator. Intervention: Study used an excluded intervention/screening approach. Population: Study was not conducted in an average-risk population. Timing: Study only reported first (prevalence) round screening follow-up. Publication type: Study was published in non–English-language or only available in an abstract. Quality: Study did not meet criteria for fair or good quality. Setting: Study was not conducted in a setting relevant to US practice. KQ indicates key question.

No eligible studies of age to start or stop screening, supplemental screening, or personalized screening were included, because no RCTs or nonrandomized studies reported more than a single round of screening comparing screening strategies. For screening interval, 1 RCT[26] and 1 nonrandomized study,[41] and for comparisons of different screening modalities (DBT vs digital mammography) 3 RCTs[27,33,42] and 2 nonrandomized studies,[34,44] met eligibility criteria.

### Screening Interval

Two fair-quality studies addressed the effect of screening interval on the characteristics of detected cancers. A fair-quality United Kingdom Co-ordinating Committee on Cancer Research (UKCCCR) RCT comparing screening intervals was conducted as part of the UK National Breast Screening Program. The study randomized participants aged 50 to 62 years to annual (n = 37 530) or triennial (n = 38 492) breast cancer screening during the years 1989 to 1996.[26] After 3 years of screening (1 incidence screen in the triennial screening group), a similar number of cancers (screen-detected and interval) had been diagnosed in the annual and triennial screening groups (6.26 and 5.40 per 1000 screened, respectively; RR, 1.16 [95% CI, 0.96 to 1.40]). No statistically significant differences were found in the cancer characteristics (tumor size, nodal status, histological grade) between groups over the course of the study.

A fair-quality nonrandomized study using Breast Cancer Surveillance Consortium (BCSC) registry data (1996 to 2012)[39] found the relative risk of being diagnosed with a breast cancer with less favorable prognostic characteristics (stage IIB or higher, tumor size >15 mm, or node-positive) was not statistically different for women screened biennially compared with those screened annually for any age category (40-49, 50-59, 60-69, 70-85 years).

### Mammography Modality

Three fair-quality RCTs[27,33,42] reported cancer detection over 2 rounds of screening, comparing the effects of screening with DBT and digital mammography on the presence of advanced cancer at subsequent screening rounds (**Table 1**). Participants were randomized to the DBT intervention group or the digital mammography control group at a first round of screening, followed in 2 trials by a second round of screening with digital mammography for all second-round participants (Proteus Donna,[27] RETomo[42]) and in 1 trial with DBT for all second-round participants (To-Be[33]). The trials used an identical screening modality for both study groups at the second round because using the same instrument is a stronger design for detection of stage shift.

The RCTs reported increased detection of invasive cancer with DBT at the first round of screening (pooled RR, 1.41 [95% CI, 1.20 to 1.64]; 3 RCTs [n = 129 492]; $I^2$ = 7.6%) and no statistical difference in invasive cancer at the subsequent screening (pooled RR, 0.87 [95% CI, 0.73 to 1.05]; 3 RCTs [n = 105 064]; $I^2$ = 0%) (eFigure 1 in the Supplement).[27,33,42] There was no statistically significant difference in the incidence of advanced cancers at the subsequent screening round (progression of cancers not found at prior screening that would indicate stage shift) in the individual trials (**Figure 3**). Results were inconsistent and thus not pooled for the advanced cancer, larger tumor (>20 mm), and node-positive cancer outcomes. The results for histologic grade 3 cancer at the second screening were consistent (pooled RR, 0.97 [95% CI, 0.61-1.55]; 3 RCTs [n = 105 244];

Table 1. Study Characteristics of Included Trials and Nonrandomized Studies of Digital Breast Tomosynthesis-Based Screening Strategies[a]

| Source (quality) | Country | No. screened (round 1) | Brief population description | Study years | Screening intervention | Screening control |
|---|---|---|---|---|---|---|
| **Randomized clinical trials** | | | | | | |
| Proteus Donna<br>Armaroli et al,[27] 2022<br>Fair | Italy | 73 866 | Women aged 46 to 68 y attending a population-based screening program | 2004 to 2017 | Round 1: DBT/DM<br>Round 2: DM | DM |
| TOSYMA<br>Heindel et al,[31] 2022<br>Good | Germany | 99 634 | Women aged 50-69 y attending a population-based screening program | 2018 to 2020 | DBT/sDM | DM |
| RETomo<br>Pattacini et al,[42] 2022<br>Good | Italy | 26 877 | Women aged 45 to 69 y attending screening in 1 of 3 clinics equipped with DBT who had already participated in at least 1 round of the Reggio Emilia screening program | 2014 to 2017 | Round 1: DBT/DM<br>Round 2: DM | DM |
| To-Be<br>Hofvind et al,[33] 2021<br>Good | Norway | 28 749 | Women aged 50-69 y attending a population-based screening program | 2016 to 2020 | DBT/sDM | Round 1: DM<br>Round 2: DBT/sDM |
| **Nonrandomized studies** | | | | | | |
| BCSC 2023<br>Sprague et al,[44] 2023<br>Fair | US | 504 863 | Women aged 40 to 79 y with no personal history of breast cancer or mastectomy who had a previous mammogram within the past 30 mo | 2011 to 2020 | DBT | DM |
| BCSC 2022<br>Ho et al,[32] 2022<br>Fair | US | 903 495 | Women aged 40 to 79 y | 2005 to 2018 | DBT | DM |
| BCSC 2022<br>Kerlikowske et al,[36] 2022<br>Fair | US | 504 427 | Women aged 40 to 79 y with no history of breast cancer or mastectomy who had a screening mammogram and/or DBT | 2011 to 2018 | DBT | DM |
| MBTST<br>Johnson et al,[35] 2021<br>Fair | Sweden | 40 107 | Women enrolled in a breast cancer screening trial and population-based matched controls | 2010 to 2015 | DBT/DM | DM |
| Richman et al,[43] 2021<br>Fair | US | 4 580 698 | Women aged 40-64 y with at least 1 screening mammogram between January 1, 2015, and December 31, 2017 | 2015 to 2017 | DBT/DM | DM |
| OVVV<br>Hovda et al,[34] 2020<br>Fair | Norway | 92 404 | Women aged 50 to 69 y participating in population-based screening program | 2014 to 2017 | Round 1: DBT/DM<br>Round 2: DM | DM |
| PROSPR<br>Conant et al,[28] 2016<br>Fair | US | 103 401 | Women aged 40 to 74 y attending screening at academic medical centers participating in surveillance consortium | 2011 to 2014 | DBT/DM | DM |

Abbreviations: BCSC, Breast Cancer Surveillance Consortium; DBT, digital breast tomosynthesis; DM, digital mammography; MBTST, Malmö Breast Tomosynthesis Screening Trial; OVVV, Oslo-Vestfold-Vestre Viken; PROSPR, Population-based Research Optimizing Screening Through Personalized Regimens; RETomo, Reggio Emilia Tomosynthesis; sDM, synthetic mammography; To-Be, Tomosynthesis Trial in Bergen; TOSYMA, Tomosynthesis plus Synthesized Mammography.

[a] DBT-based screening strategies involve use DBT in addition to DM, which can be either a separate 2D digital mammography (DM) scan or a 2D image constructed from the DBT scan (sDM). Studies did not consistently specify what type of 2D image was received.

$I^2$ = 0%) (Figure 3). Due to the small number of cases, it was not possible to assess differences in the detection of cancers lacking hormone or growth factor receptors (ie, triple-negative cancers) that have the worst prognosis among breast cancer subtypes.

Two fair-quality nonrandomized studies of interventions (NRSIs), including a US study using BCSC data, compared breast cancer detection outcomes from screening over multiple rounds (≥2) with either DBT-based mammography or digital mammography alone.[34,44] The findings were generally consistent with the trial results for cancer detection and stage shift.

## Harms of Screening

**Key Question 3.** What are the comparative harms of different mammography-based breast cancer screening strategies (modality, interval, initiation age, use of supplemental imaging, or personalization based on risk factors)?

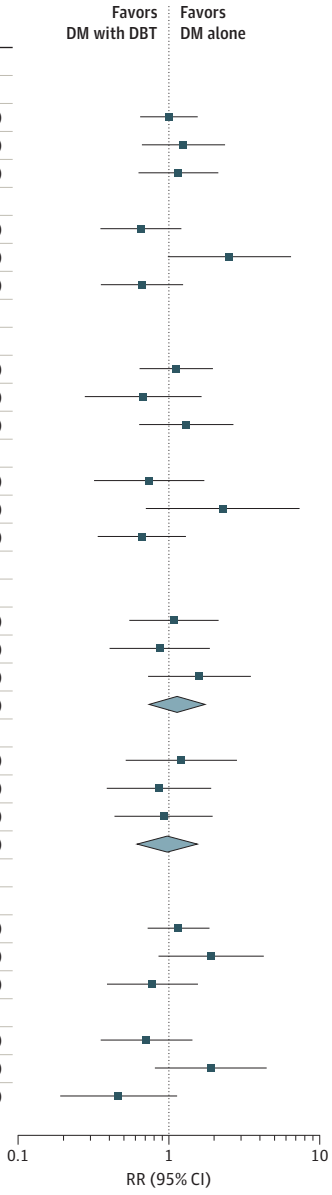
No eligible studies of age to start screening or personalized screening were identified. For age to stop screening, 1 fair-quality nonrandomized study met eligibility criteria.[30] For comparisons of potential harms associated with different screening intervals, a fair-quality RCT[26] and 2 fair-quality nonrandomized studies[39,41] were included. For comparisons of different screening modalities (DBT vs digital mammography), 4 RCTs (3 good- and 1 fair-quality)[27,31,33,42] and 7 fair-quality nonrandomized studies were included.[28,32,34-36,43,44]

### Age to Stop Screening

In the NRSI using an emulated trial methodology to evaluate the age to stop screening,[30] the 8-year cumulative proportion of participants with a breast cancer diagnosis was higher among those who continued annual screening from ages 70 to 84 years (5.5%) compared with those who discontinued screening (3.9%) at age

Figure 3. Proportion of Screen-Detected Invasive Cancers Diagnosed (Advanced Stage [II or Higher], Tumor Size >20 mm, Tumor Grade 3, Node-Positive Cancer) With Digital Breast Tomosynthesis vs Digital Mammography

| Source | Participants, No. | | Rate per 1000 | | RR (95% CI) |
|---|---|---|---|---|---|
| | DM with DBT | DM alone | DM with DBT | DM alone | |
| **Advanced stage (stage II+)** | | | | | |
| First round | | | | | |
| Proteus Donna (Armaroli et al,[27] 2022) | 30844 | 43022 | 1.23 | 1.23 | 1.00 (0.64-1.56) |
| RETomo (Pattacini et al,[42] 2022) | 13356 | 13521 | 1.57 | 1.26 | 1.25 (0.66-2.37) |
| To-Be (Hofvind et al,[33] 2021) | 14380 | 14369 | 1.53 | 1.32 | 1.16 (0.63-2.14) |
| Second round | | | | | |
| Proteus Donna (Armaroli et al,[27] 2022) | 23760 | 33534 | 0.72 | 1.10 | 0.65 (0.35-1.21) |
| RETomo (Pattacini et al,[42] 2022) | 12733 | 12911 | 1.18 | 0.46 | 2.53 (0.98-6.53) |
| To-Be (Hofvind et al,[33] 2021) | 11201 | 11105 | 1.43 | 2.16 | 0.66 (0.35-1.24) |
| **Tumor size >20 mm** | | | | | |
| First round | | | | | |
| Proteus Donna (Armaroli et al,[27] 2022) | 30844 | 43022 | 0.81 | 0.72 | 1.12 (0.64-1.96) |
| RETomo (Pattacini et al,[42] 2022) | 13356 | 13521 | 0.60 | 0.89 | 0.67 (0.28-1.65) |
| To-Be (Hofvind et al,[33] 2021) | 14380 | 14369 | 1.18 | 0.90 | 1.31 (0.63-2.69) |
| Second round | | | | | |
| Proteus Donna (Armaroli et al,[27] 2022) | 23760 | 33534 | 0.42 | 0.57 | 0.74 (0.32-1.72) |
| RETomo (Pattacini et al,[42] 2022) | 12733 | 12911 | 0.71 | 0.31 | 2.28 (0.70-7.41) |
| To-Be (Hofvind et al,[33] 2021) | 11201 | 11105 | 1.25 | 1.89 | 0.66 (0.34-1.30) |
| **Tumor grade 3[a]** | | | | | |
| First round | | | | | |
| Proteus Donna (Armaroli et al,[27] 2022) | 30844 | 43022 | 0.55 | 0.51 | 1.08 (0.54-2.14) |
| RETomo (Pattacini et al,[42] 2022) | 13356 | 13521 | 0.90 | 1.04 | 0.87 (0.40-1.88) |
| To-Be (Hofvind et al,[33] 2021) | 14380 | 14369 | 1.11 | 0.70 | 1.60 (0.73-3.52) |
| Heterogeneity: $\tau^2 = 0.00$, $I^2 = 0.00\%$, $H^2 = 1.00$ | | | | | 1.13 (0.74-1.74) |
| Second round | | | | | |
| Proteus Donna (Armaroli et al,[27] 2022) | 23760 | 33534 | 0.51 | 0.42 | 1.21 (0.51-2.85) |
| RETomo (Pattacini et al,[42] 2022) | 12733 | 12911 | 0.86 | 1.01 | 0.86 (0.38-1.91) |
| To-Be (Hofvind et al,[33] 2021) | 11201 | 11105 | 1.16 | 1.26 | 0.92 (0.43-1.96) |
| Heterogeneity: $\tau^2 = 0.00$, $I^2 = 0.00\%$, $H^2 = 1.00$ | | | | | 0.97 (0.61-1.55) |
| **Node-positive cancer** | | | | | |
| First round | | | | | |
| Proteus Donna (Armaroli et al,[27] 2022) | 30844 | 43022 | 1.13 | 0.98 | 1.16 (0.72-1.86) |
| RETomo (Pattacini et al,[42] 2022) | 13356 | 13521 | 1.27 | 0.67 | 1.91 (0.85-4.29) |
| To-Be (Hofvind et al,[33] 2021) | 14380 | 14369 | 0.97 | 1.25 | 0.78 (0.39-1.56) |
| Second round | | | | | |
| Proteus Donna (Armaroli et al,[27] 2022) | 23760 | 33534 | 0.59 | 0.83 | 0.71 (0.35-1.44) |
| RETomo (Pattacini et al,[42] 2022) | 12733 | 12911 | 1.18 | 0.62 | 1.90 (0.81-4.48) |
| To-Be (Hofvind et al,[33] 2021) | 11201 | 11105 | 0.62 | 1.35 | 0.46 (0.19-1.13) |




DBT indicates digital breast tomosynthesis; DM, digital mammography; and RR, relative risk.

[a]From random-effects restricted maximum likelihood model.

70 years. Because fewer cancers were diagnosed among those who discontinued screening, there was a lower risk of undergoing cancer treatment and experiencing related morbidity. Notably, for participants aged 75 to 84 years, screening (and treatment) were not associated with lower breast cancer mortality (see KQ1 results).

### Screening Interval

The UKCCCR trial included for KQ2[26] reported fewer interval cancers (false-negative and incident cancers) diagnosed in the annual invitation group compared with triennial screening (1.84 vs 2.70 per 1000 women screened, respectively; RR, 0.68 [95% CI, 0.50 to 0.92]). The nonrandomized clinical trial conducted in Finland included for KQ1[41] also reported interval cancers diagnosed with annual vs triennial screening and found no statistical difference in incidence ($P = .22$, data not reported). Data from 2 studies from the BCSC registry reported higher probabilities of false-positive recalls and biopsy recommendations with annual screening compared with biennial screening and no statistical difference in interval cancers in adjusted analyses.[32,39,44]

### Mammography Modality

Four RCTs (3 good-quality, 1 fair-quality)[27,31,33,42] and 7 fair-quality nonrandomized studies[28,32,34-36,43,44] reported outcomes related to potential screening harms associated with DBT-based screening compared with digital mammography–only screening, including interval cancer rates, round-specific and cumulative false-positive recalls and biopsies, and radiation exposure. Meta-analysis of 3 large trials did not show a statistically significant difference in rates of interval cancer after screening with DBT compared with digital mammography (pooled RR, 0.87 [95% CI, 0.64 to 1.17]; 3 RCTs [n = 130 196]; $I^2$ = 0%) (eFigure 2 in the Supplement).[27,33,42]

Data on interval cancers were also obtained from 7 nonrandomized studies.[28,32,34-36,43,44] The most recent BCSC analysis, reporting interval cancer rates across multiple screening rounds with either DBT or digital mammography, did not identify statistically significant differences in invasive or advanced interval cancers.[44]

The effects of DBT screening on false-positive recall and false-positive biopsy rates varied across studies[27,33,42] and by screening round, with small or no statistical differences between study groups, not consistently favoring DBT-based mammography or digital mammography.

Evidence from 2 nonrandomized BCSC studies provided false-positive results across several screening rounds.[32,44] In 1 study, rates of false-positive recall and false-positive biopsy rates were lower with DBT in initial screening rounds, but differences were attenuated and not statistically significant compared with digital mammography only after additional rounds of screening (**Table 2**).[44] The other study reported no statistical difference in 10-year cumulative false-positive biopsy recommendation rates between biennial DBT and digital mammography screening, but false-positive recall was slightly lower with DBT (eFigures 3 and 4 in the Supplement); no differences by modality were identified for individuals with extremely dense breasts in stratified analyses (eFigure 5 in the Supplement).[32]

Four RCTs[27,31,33,42] and 1 NRSI[35] reported the mean, median, or relative radiation dose received in each study group at a single screening round. The 3 studies using DBT/digital mammography screening reported radiation exposure approximately 2 times higher in the intervention group compared with the digital mammography–only group.[27,35,42] Differences between study groups in radiation exposure were smaller in studies using DBT with synthetic digital mammography.[33,47]

### Supplemental Screening

The Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial, a good-quality RCT conducted in the Netherlands, randomized (1:4) participants aged 50 to 75 years with extremely dense breasts and negative mammography findings (2011-2015) (n = 40 373) to an invitation or no invitation for supplemental MRI screening.[45] (The RCT was not included for KQ2 because second round results in the control group were unavailable). Fifty-nine percent of those randomized to the invitation underwent an MRI examination (n = 4783). In intention-to-treat analysis, 2.2 per 1000 experienced interval breast cancer diagnoses in the supplemental screening invitation group, compared with 4.7 per 1000 screened in the digital mammography control group (RR, 0.47 [95% CI, 0.29 to 0.77]). Adverse events related to the supplemental MRI screening reported in the trial included 5 classified as serious adverse events (2 vasovagal reactions and 3 allergic reactions to the contrast agent) and 2 reports of extravasation (leaking) of the contrast agents and 1 shoulder subluxation. Twenty-seven participants (0.6% of the MRI group) reported a serious adverse event within 30 days of the MRI. Those who underwent supplemental MRI screening also experienced additional recalls (94.9 per 1000 screened), false-positive recalls (80.0 per 1000 screened), and false-positive biopsies (62.7 per 1000 screened).

A fair-quality nonrandomized study used claims data from commercially insured women (MarketScan database) aged 40 to 64 years who had received at least 1 bilateral screening breast MRI (n = 9208) or mammogram (n = 9208) between January 2017 and June 2018.[29] Following propensity score matching, those undergoing screening with MRI were more likely to have additional health care cascade events such as office visits and follow-up tests unrelated to breast conditions (adjusted difference between groups, 19.6 per 100 screened [95% CI, 8.6 to 30.7]) in the subsequent 6 months.

A fair-quality RCT, the Japan Strategic Anti-cancer Randomized Trial, randomly assigned asymptomatic women aged 40 to 49 years (2007-2011) to breast cancer screening with mammography plus handheld ultrasound (digital mammography/ultrasound) (n = 36 859) or mammography only (digital mammography) (n = 36 139).[40] The relative risk of invasive interval cancer was not statistically significantly different for digital mammography/ultrasound vs digital mammography only (RR, 0.58 [95% CI, 0.31 to 1.08]). This result differs from the statistically significant population-average effect reported in the study ($P$ = .03), which included interval ductal carcinoma in situ (proportion difference, −0.05% [95% CI, −0.09 to 0]). Those undergoing ultrasound in addition to digital mammography experienced 48.0 per 1000 additional false-positive recall results compared with those assigned to digital mammography screening only.

A fair-quality nonrandomized study using data from 2 BCSC registry sites compared screening outcomes for participants receiving ultrasonography on the same day as a screening mammogram (digital mammography/ultrasound) (n = 3386, contributing 6081 screens) compared with those that received only a mammogram (digital mammography) (n = 15 176, contributing 30 062 screens).[37] However, 31% of participants had a first-degree family history of breast cancer or previous breast biopsy. There was no statistical difference in interval cancer risk (adjusted RR, 0.67 [95% CI, 0.33 to 1.37]), and rates of false-positive biopsy were twice as high for the mammography/ultrasound group (adjusted RR, 2.23 [95% CI, 1.03 to 2.58]).

## Discussion

Prior screening effectiveness reviews based on large trials initiated in previous decades established a statistically significant mortality benefit for mammography screening of women aged 50 to 69 years.[8,9,15] The current review considered comparative effectiveness questions on the relative benefits and harms of different screening start and stop ages, intervals, and modalities for women at average breast cancer risk. Findings are summarized in **Table 3**.

The evidence was insufficient for addressing the age to start or end screening. No eligible studies comparing different ages to start screening were identified. Limited evidence from 1 nonrandomized study, using an emulated trial study design, suggested that

Table 2. Harms Reported in Studies Comparing Digital Breast Tomosynthesis-Based Screening Strategies and Digital Mammography[a]

| Source | Study design | Follow-up round | Modality (previous-round modality) | No./total (rate per 1000 screened) | Effect (95% CI) |
|---|---|---|---|---|---|
| **Recalled for further assessment** | | | | | |
| Proteus Donna[b] Armaroli et al,[27] 2022 | RCT | 1 | DBT/DM | 1995/30 844 (63.4) | RR: 1.24 (1.17 to 1.32) |
| | | | DM | 2191/43 022 (50.9) | |
| | | 2 | DM (DBT/DM) | 1000/23 760 (42.1) | RR: 0.97 (0.89 to 1.05) |
| | | | DM | 1456/33 534 (43.4) | |
| RETomo[c] Pattacini et al,[42] 2022 | RCT | 1 | DBT/DM | 511/13 356 (38.3) | RR: 0.99 (0.88 to 1.10) |
| | | | DM | 522/13 521 (38.6) | |
| | | 2 | DM (DBT/DM) | 464/12 733 (36.4) | RR: 0.93 (0.82 to 1.10) |
| | | | DM | 506/12 911 (39.2) | |
| To-Be[c] Hofvind et al,[33] 2021 | RCT | 1 | DBT/sDM | 444/14 380 (30.9) | RR: 0.78 (0.69 to 0.88)[d] |
| | | | DM | 571/14 369 (39.7) | |
| | | 2 | DBT/sDM | 440/11 201 (39.3) | RR: 0.99 (0.87 to 1.13)[d] |
| | | | DBT/sDM (DM) | 441/11 105 (39.7) | |
| OVVV[c] Hovda et al,[34] 2020 | NRSI | 1 | DBT/sDM | 1253/37 185 (33.7) | RR: 1.02 (0.95 to 1.09)[d] |
| | | | DM | 2037/61 742 (33.0) | |
| | | 2 | DM (DBT/sDM) | 621/26 474 (23.5) | RR: 0.76 (0.69 to 0.83)[d] |
| | | | DM | 1408/45 543 (30.9) | |
| BCSC 2023 Sprague et al,[44] 2023 | NRSI | 1 | DBT | NR (75) | Proportion difference: −33 (−46 to -21) |
| | | | DM | NR (109) | |
| | | 2 | DBT | NR (69) | Proportion difference: −18 (−29 to -7) |
| | | | DM | NR (86) | |
| | | ≥3 | DBT | NR (61) | Proportion difference: −12 (−24 to -1) |
| | | | DM | NR (73) | |
| **Percutaneous needle biopsy** | | | | | |
| RETomo[c] Pattacini et al,[42] 2022 | RCT | 1 | DBT/DM | 159/13 356 (11.9) | RR: 1.50 (1.10 to 1.90) |
| | | | DM | 110/13 521 (8.1) | |
| | | 2 | DM (DBT/DM) | 78/12 733 (6.1) | RR: 0.76 (0.57 to 1.00) |
| | | | DM | 104/12 911 (8.1) | |
| **Biopsy (undefined)** | | | | | |
| To-Be[c] Hofvind et al,[33] 2021 | RCT | 1 | DBT/sDM | 252/14 380 (17.5) | RR: 0.93 (0.78 to 1.10)[d] |
| | | | DM | 271/14 369 (18.9) | |
| | | 2 | DBT/sDM | 248/11 201 (22.1) | RR: 0.95 (0.80 to 1.13)[d] |
| | | | DBT/sDM (DM) | 258/11 105 (23.2) | |
| BCSC 2023 Sprague et al,[44] 2023 | NRSI | 1 | DBT | NR (15) | Proportion difference: −3 (−5 to -1) |
| | | | DM | NR (18) | |
| | | 2 | DBT | NR (13) | Proportion difference: 0 (−3 to 2) |
| | | | DM | NR (14) | |
| | | ≥3 | DBT | NR (12) | Proportion difference: 0 (−2 to 3) |
| | | | DM | NR (13) | |
| **Surgical referrals** | | | | | |
| Proteus Donna[b] Armaroli et al,[27] 2022 | RCT | 1 | DBT/DM | 305/30 844 (9.9) | RR: 1.54 (1.31 to 1.82)[d] |
| | | | DM | 276/43 022 (6.4) | |
| | | 2 | DM (DBT/DM) | 103/23 760 (4.3) | RR: 0.76 (0.59 to 0.97)[d] |
| | | | DM | 191/33 534 (5.7) | |
| **Surgical procedures (including open biopsy)** | | | | | |
| RETomo[c] Pattacini et al,[42] 2022 | RCT | 1 | DBT/DM | 116/13 356 (8.7) | RR: 1.70 (1.30 to 2.30) |
| | | | DM | 68/13 521 (5.0) | |
| | | 2 | DM (DBT/DM) | 68/12 733 (5.3) | RR: 0.83 (0.60 to 1.10) |
| | | | DM | 83/12 911 (6.4) | |

(continued)

Table 2. Harms Reported in Studies Comparing Digital Breast Tomosynthesis-Based Screening Strategies and Digital Mammography[a] (continued)

| Source | Study design | Follow-up round | Modality (previous-round modality) | No./total (rate per 1000 screened) | Effect (95% CI) |
|---|---|---|---|---|---|
| **False-positive recall[e]** | | | | | |
| Proteus Donna[b] Armaroli et al,[27] 2022 | RCT | 1 | DBT/DM | 1699/30 844 (55.1) | RR: 1.22 (1.14 to 1.30)[d] |
| | | | DM | 1943/43 022 (45.2) | |
| | | 2 | DM (DBT/DM) | 900/23 760 (37.9) | RR: 0.99 (0.91 to 1.08)[d] |
| | | | DM | 1286/33 534 (38.3) | |
| RETomo[c] Pattacini et al,[42] 2022 | RCT | 1 | DBT/DM | 410/13 356 (30.7) | RR: 0.90 (0.79 to 1.00)[d] |
| | | | DM | 461/13 521 (34.1) | |
| | | 2 | DM (DBT/DM) | 403/12 733 (31.7) | RR: 0.95 (0.83 to 1.09)[d] |
| | | | DM | 430/12 911 (33.3) | |
| To-Be[c] Hofvind et al,[33] 2021 | RCT | 1 | DBT/sDM | 349/14 380 (24.3) | RR: 0.72 (0.63 to 0.83)[d] |
| | | | DM | 484/14 369 (33.7) | |
| | | 2 | DBT/sDM | 349/11 201 (31.2) | RR:1.02 (0.88 to 1.18)[d] |
| | | | DBT/sDM (DM) | 340/11 105 (30.6) | |
| BCSC 2023 Sprague et al,[44] 2023 | NRSI | 1 | DBT | NR (66) | Proportion difference: −34 (−47 to −22) |
| | | | DM | NR (101) | |
| | | 2 | DBT | NR (60) | Proportion difference: −18 (−30 to −7) |
| | | | DM | NR (78) | |
| | | ≥3 | DBT | NR (55) | Proportion difference: −11 (−23 to 2) |
| | | | DM | NR (66) | |
| OVVV[c] Hovda et al,[34] 2020 | NRSI | 1 | DBT/sDM | 905/37 185 (24.3) | RR: 0.91 (0.84 to 0.98)[d] |
| | | | DM | 1658/61 742 (26.9) | |
| | | 2 | DM (DBT/sDM) | 518/26 474 (19.6) | RR: 0.77 (0.70 to 0.86)[d] |
| | | | DM | 1154/45 543 (25.3) | |
| **False-positive biopsy result[f]** | | | | | |
| To-Be[c] Hofvind et al,[33] 2021 | RCT | 1 | DBT/sDM | 157/14 380 (10.9) | RR: 0.85 (0.69 to 1.05)[d] |
| | | | DM | 184/14 369 (12.8) | |
| | | 2 | DBT/sDM | 157/11 201 (14.0) | RR: 0.99 (0.80 to 1.24)[d] |
| | | | DBT/sDM (DM) | 157/11 105 (14.1) | |
| BCSC 2023 Sprague et al,[44] 2023 | NRSI | 1 | DBT | NR (10) | Proportion difference: −3 (−5 to −2) |
| | | | DM | NR (13) | |
| | | 2 | DBT | NR (8) | Proportion difference: −2 (−4 to 0) |
| | | | DM | NR (10) | |
| | | ≥3 | DBT | NR (8) | Proportion difference: −1 (−3 to 1) |
| | | | DM | NR (8) | |

Abbreviations: BCSC, Breast Cancer Surveillance Consortium; DBT, digital breast tomosynthesis; DM, digital mammography; NR, not reported; NRSI, nonrandomized study of intervention; OVVV, Oslo-Vestfold-Vestre Viken; RCT, randomized clinical trial; RETomo, Reggio Emilia Tomosynthesis; RR, relative risk; sDM, synthetic 2-view mammography; To-Be, Tomosynthesis Trial in Bergen.

[a] DBT-based screening strategies involve use DBT in addition to DM, which can be either a separate 2D digital mammography (DM) scan or a 2D image constructed from the DBT scan (sDM). Studies did not consistently specify what type of 2D image was received.

[b] Recalled for an assessment after double reading based on positive or suspicious screening result by either radiologist (without consensus or arbitration).

[c] Recalled for an assessment (after double reading and arbitration) based on positive or suspicious screening results.

[d] Relative risk calculated from Ns.

[e] Recalled for assessment without a finding of invasive cancer or ductal carcinoma in situ.

[f] Underwent biopsy without a finding of invasive cancer or ductal carcinoma in situ.

screening beyond age 74 years may not reduce breast cancer mortality.[30]

Evidence was also insufficient for evaluating the effect of screening intervals on breast cancer morbidity and mortality. Two nonrandomized studies found no difference in breast cancer outcomes.[26,39] Moderate evidence supported longer screening intervals (eg, biennial) to reduce the cumulative risk of false-positive recall and biopsy. The observational studies of different screening intervals compared individuals who self-selected or were referred for different screening intervals, contributing to risk of bias in the results.

Results from 3 RCTs[27,33,42] and 2 nonrandomized studies[34,44] provided moderate evidence that DBT-based mammography does not reduce the risk of invasive interval cancer or advanced cancer at subsequent screening rounds. Additional rounds of screening and longer follow-up are needed to fully evaluate whether DBT

Table 3. Summary of Evidence for Breast Cancer Screening[a]

| Screening strategy | No. of studies and study design (total sample size) by screening strategy | Summary of findings by screening strategy or outcome | Consistency and precision | Other limitations | Overall strength of evidence | Applicability |
|---|---|---|---|---|---|---|
| **KQ1: Morbidity and mortality** | | | | | | |
| Age to start or stop screening | Age to start: 0<br>Age to stop: 1 NRSI (n = 1 058 013) | Age to start: NA<br>Age to stop: screening from age 70 to 74 y: 8-y risk of breast cancer mortality was 1 fewer death per 1000 women who continued screening (RD, −1.0 [95% CI, −2.3 to 0.1])<br>Adjusted HR suggested a 22% lower hazard of 8-y breast cancer mortality with continued screening (adjusted HR, 0.78 [95% CI, 0.63 to 0.95])<br>Screening beyond age 74 y: no difference in 8-y estimated risk in breast cancer mortality (RD, 0.07 [95% CI, −0.93 to 1.3]; adjusted HR, 1.00 [95% CI, 0.83 to 1.19]) with continued screening | Age to start: NA<br>Age to stop: consistency NA, imprecise | Advanced statistical methods to emulate per protocol trial; differences in estimates of effects depending on adjustments used<br>Risk of bias from unmeasured confounding and selection | Insufficient | US Medicare Parts A, B enrollees ages 70 to 84 in years 1999 to 2008 with high probability of living >10 y; population >90% non-Hispanic White |
| Screening interval | Annual vs triennial: 1 NRSI (n = 14 765)<br>Annual vs biennial: 0 | Annual vs triennial: no difference in breast cancer mortality (RR, 1.14 [95% CI, 0.59 to 1.27]) or all-cause mortality (RR, 1.20 [95% CI, 0.99 to 1.46]) at 13 y<br>Annual vs biennial: NA | Annual vs triennial: consistency NA, imprecise<br>Annual vs biennial: NA | Assignment based on birth year, limited information on baseline characteristics, potential risk of bias due to unmeasured confounding and selection | Insufficient | Invitation to annual or triennial film mammography for ages 40 to 49 y in Finnish national screening program; treatment advances since the study conducted (1985-1995)<br>No reporting of participant characteristics |
| **KQ2: Incidence and progression to advanced cancer** | | | | | | |
| Screening interval | Annual vs triennial:1 RCT (n = 76 022)<br>Annual vs biennial: 1 NRSI (n = 15 440) | Annual vs triennial: more invasive cancers screen detected over 3 y with annual screening (RR, 1.64 [95% CI, 1.28 to 2.09])<br>Total number of invasive cancers similar (RR, 1.16 [95% CI, 0.96 to 1.40]); no statistical differences by screening interval in tumor size, nodal status, grade, or prognostic index for all cancers diagnosed<br>Annual vs biennial: no difference in risk of stage IIB or higher or of less favorable prognosis cancers diagnosed after a biennial compared with annual interval for any age group | Annual vs triennial: consistency NA, imprecise<br>Annual vs biennial: consistency NA, imprecise | Annual vs triennial: birth month used to assign intervention group first 2 y of trial, which could introduce bias, no reporting of participant characteristics<br>Study never reported mortality outcome as planned<br>Annual vs biennial: risk of bias due to limited adjustment for confounding and potential unmeasured confounding and selection into study groups | Annual vs triennial: low for greater detection of invasive cancer and no difference in tumor characteristics with annual screening<br>Annual vs biennial: insufficient | Annual vs triennial:<br>People aged 50 to 62 y screened in UK screening program 1989 to 1996; changes in population health, cancer treatment, screening modalities<br>No reporting of participant characteristics<br>Annual vs biennial:<br>Conducted using BCSC data linked with US SEER and other tumor registry sources; ages 40 to 85 y; >77% population non-Hispanic White |

(continued)

Table 3. Summary of Evidence for Breast Cancer Screening[a] (continued)

| Screening strategy | No. of studies and study design (total sample size) by screening strategy | Summary of findings by screening strategy or outcome | Consistency and precision | Other limitations | Overall strength of evidence | Applicability |
|---|---|---|---|---|---|---|
| DBT vs DM | 3 RCTs (n = 130 196)<br>2 NRSIs<br>(n = 597 267) | Three RCTs reported higher invasive cancer detection in first screening round with DBT (pooled RR, 1.41 [95% CI, 1.20 to 1.64]; 3 RCTs [n = 129 492]; $I^2$ = 8%) with absolute differences in the trials ranging from 0.6 to 2.4 additional cancers per 1000 screened<br>No detection difference was seen at round 2 (pooled RR, 0.87 [95% CI, 0.73 to 1.05]; 3 RCTs [n = 105 064]; $I^2$ = 0%)<br>Two NRSIs also reported higher invasive cancer detection rates with DBT screening<br>Three RCTs and 1 NRSI reported advanced cancer detection (stage >II) after the first screening round; there were no statistical differences in the individual trials or the NRSIs<br>Three RCTs and 2 NRSIs reported tumor characteristics that inform staging (tumor diameter, histologic grade, or node status)<br>No statistically significant differences in these or other individual tumor prognostic characteristics were reported at the first or second round of screening for any of the included studies; statistical power was limited for comparisons of less common tumor types | Detection of invasive cancer: consistent, precise<br>Stage shift: consistent, imprecise | The fair-quality RCT did not describe randomization procedures, and balance in baseline characteristics could not be assessed due to limited reporting<br>Two NRSIs provided generally consistent evidence but with higher risk of bias | Moderate for increased detection with DBT<br>Low for absence of stage shift after the first screening round | All trials conducted in European countries with national organized screening programs (Italy, Sweden) that use independent dual reading and consensus procedures different from US practice<br>Some studies used DBT paired with DM and some used DBT paired with sDM<br>Prior readings were generally available<br>All studies had limited reporting of participant characteristics with no data on racial and/or ethnic characteristics |
| **KQ3: Harms of screening** | | | | | | |
| Age to start or stop screening | Age to start: 0<br>Age to stop: 1 NRSI<br>(n = 1 058 013) | Fewer cancers diagnosed in stop screening strategy; possible overdiagnosis with continued screening<br>Cancers diagnosed in stop screening strategy more likely to receive aggressive treatments (radical mastectomy and chemotherapy vs lumpectomy and radiotherapy) | Age to start: NA<br>Age to stop: consistency NA, imprecise | Advanced statistical methods to emulate per protocol trial; differences in estimates of effects depending on adjustments used<br>Risk of bias from unmeasured confounding and selection | Insufficient | US Medicare Part A, B enrollees aged 70 to 84 y in the years 1999 to 2008 with high probability of living >10 y; population >90% White non-Hispanic |
| Screening interval | Annual vs triennial:<br>1 RCT (n = 76 022)<br>1 NRSI<br>(n = 14 765)<br>Annual vs biennial:<br>3 NRSIs<br>(n = 920 954) | Annual vs triennial:<br>Interval cancers: 1 RCT (n = 76 022) estimated 1 less invasive interval cancer per 1000 in the annual screening group (1.8 vs 2.7 per 1000 screened; RR, 0.68 [95% CI, 0.50 to 0.92])<br>One NRSI (n = 14 765) using birth year to assign screening intervals found no difference in interval cancer incidence ($P$ = .22)<br>False-positives: NR<br>Annual vs biennial:<br>Interval cancers: 1 NRSI using BCSC data (n = 15 440) reported the unadjusted interval cancer proportion for people screened negative after an annual (22.2%) or biennial screening (27.2%) interval<br>False-positive recall and biopsy higher with annual compared with biennial screening<br>One study using BCSC data (n = 903 495) reported that over 10 y of DBT screening approximately 50% of those undergoing annual screening had at least 1 false-positive recall, compared with approximately 35% of those undergoing biennial screening; annual screening resulted in ≈50 additional false-positive biopsies per 1000 screened over 10 y (annual, ≈115 per 1000 vs biennial, ≈66 per 1000)<br>One NRSI (n = 2019) reported >2 times higher odds (OR, 2.2 [95% CI, 1.7 to 2.8]) of a false-positive result over a median of 8.9 y | Annual vs triennial:<br>Interval cancer: inconsistent, precise<br>False positives: NA<br>Annual vs biennial:<br>Interval cancer: consistency NA, imprecise<br>False positives: consistent, precise | Annual vs triennial: RCT did not use random allocation for first 2 y of study (birth month) and NRSI assigned interval based on birth year (odd, even), lack of information on group baseline characteristics in both studies; potential risk of bias due to unmeasured confounding and selection<br>Annual vs biennial: NRSIs used EMR, registry, and self-report data; the largest study (n = 903 945) did not provide information on participant characteristics; data on interval cancers are unadjusted for participant characteristics; risk of bias from potential selection and confounding bias, including time-varying factors<br>BCSC NRSI with cumulative false-positives did not include prevalence screens, may underestimate cumulative false-positives from start of screening | Annual vs triennial: low for a small difference in interval cancer with annual screening; insufficient for other harms<br>Annual vs biennial: insufficient for interval cancer; moderate for higher recall, biopsy, and false-positives with annual screening | Annual vs triennial:<br>Both triennial screening interval studies conducted in Europe in 1990s; RCT screened women aged 50 to 62 y; NRSI among women screened aged 40 to 49<br>No information on participant characteristics other than age<br>Annual vs biennial:<br>Screening studies conducted in US, 1 NRSI conducted in a single academic medical center that reported false-positive recall had majority Hispanic population (76%)<br>BCSC data occurred in primarily non-Hispanic White participants (78%) |

(continued)

Table 3. Summary of Evidence for Breast Cancer Screening[a] (continued)

| Screening strategy | No. of studies and study design (total sample size) by screening strategy | Summary of findings by screening strategy or outcome | Consistency and precision | Other limitations | Overall strength of evidence | Applicability |
|---|---|---|---|---|---|---|
| DBT vs DM | 4 RCTs (n = 229 830)<br>7 NRSIs<br>(n = 6 735 868[b]) | Interval cancers: 3 RCTs did not find difference in interval cancer rates (pooled RR, 0.87 [95% CI, 0.64 to 1.17]; 3 RCTs [n = 130 196]; $I^2$ = 0%)<br>Six NRSIs with varied study designs reported interval cancer rates; 4 reported results consistent with the trial evidence<br>Recall and false-positive recalls: 3 RCTs and 2 NRSIs reported on recall rates and false-positive recall rates from at least 2 rounds of screening<br>RCT evidence from 2 screening rounds and NRSI across 2 or more cumulative screening rounds suggested little or no difference in false-positive recall<br>One NRSI reporting cumulative probability of at least 1 false-positive recall over 10 y of screening estimated slightly lower false-positive recall with annual DBT (50% vs 56%) and similar rates with biennial screening (36% vs 38%); a second NRSI found lower rates of recall with DBT, especially with fewer screening rounds<br>Biopsy and false-positive biopsy: 2 RCTs reported no difference in biopsy risk across 2 rounds of screening, with 1 reporting higher risk with DBT in the first round; 1 RCT reported no difference in false-positive biopsy<br>Two NRSIs reported no difference in false-positive biopsy risk over several screening rounds, with 1 reporting the cumulative probability of at least 1 false-positive biopsy over 10 y of screening finding no difference in cumulative false-positive biopsy for DBT v DM regardless of screening interval (11%-12% annual, 7%-8% biennial)<br>Overdetection: 3 RCTs did not find differences in DCIS, screen-detected lesions that could contribute to overdetection, at round 1 (pooled RR, 1.33 [95% CI, 0.92 to 1.93]; 3 RCTs [n = 130 196]; $I^2$ = 0%) or round 2 (pooled RR, 0.75 [95% CI, 0.49 to 1.14]; 3 RCTs [n = 130 196]; $I^2$ = 0%)<br>Adverse events: 1 RCT (n = 99 634) reported the same number of adverse events for both screening tests (DBT/sDM vs DM) (n = 6), all nonserious<br>Radiation exposure: In 3 studies using DBT/DM the dose was ≈2 × mGy higher than for DM; for 2 studies using DBT/sDM the dose was similar to DM only | Interval cancer: consistent, imprecise<br>Recall and false-positive recall: inconsistent, precise<br>Biopsy and false-positive biopsy: inconsistent, imprecise<br>Overdetection/ overtreatment: consistent, imprecise<br>Adverse events: consistency NA, imprecise<br>Radiation: consistent, imprecise | NRSIs had substantial risk of bias, limited adjustment for potential confounding and selection<br>Most NRSIs included retrospective assessment screening from records; limited adjustment for all factors that may contribute to DBT vs DM screening, including time-dependent factors<br>One NRSI that age-matched trial participants with controls from general screening population lacked adjustment for any factor other than age<br>One NRSI with geographical comparator did not describe characteristics by study group, only minor statistical adjustment elevated selection and confounding risk of bias concerns | Interval cancer: moderate for no difference<br>Recall and false-positive recall: low for lower with DBT<br>Biopsy and false positive biopsy: low for no difference<br>Overdetection/ overtreatment: low for no difference<br>Adverse events: insufficient<br>Radiation: moderate for increased radiation with DBT/DM and no increased radiation with DBT/sDM | No US-based RCTs; European RCTs and NRSIs based in organized screening programs and use independent dual mammography reading, consensus<br>Limited reporting on population characteristics, including no racial and ethnic data<br>Of the 7 NRSIs, 5 were conducted in the US and 1 each in Sweden and Norway; the US-based studies included data from the BCSC, medical administrative claims, and the multisite PROSPR study<br>Participant characteristics were only reported in 3 US studies with ≈70% non-Hispanic White participants<br>The Swedish and Norwegian studies reported no baseline characteristics |

(continued)

Table 3. Summary of Evidence for Breast Cancer Screening[a] (continued)

| Screening strategy | No. of studies and study design (total sample size) by screening strategy | Summary of findings by screening strategy or outcome | Consistency and precision | Other limitations | Overall strength of evidence | Applicability |
|---|---|---|---|---|---|---|
| Supplemental screening with MRI | 1 RCT (n = 40 373)<br>1 NRSI (n = 18 416) | Interval cancer: 1 RCT reported reduced invasive interval cancer with invitation to screening for women with extremely dense breasts and negative mammogram (2.2 vs 4.7 per 1000 invited to screening; RR, 0.47 [95% CI, 0.29 to 0.77])<br>Adverse events: RCT reported 8 adverse events (5 serious) during or immediately after MRI—vasovagal reactions, allergic reactions to contrast agent, leaking of contrast agent (extravasation), shoulder subluxation<br>Downstream consequences of supplemental imaging (including incidental findings): MRI resulted in additional recall (95 per 1000 screened), false-positive recall (80 per 1000), and biopsy (63 per 1000 screened) that did not occur for the DM-only group; RCT did not report on incidental findings from MRI<br>NRSI reported no difference in new diagnoses unrelated to breast conditions<br>Events unrelated to breast diagnostic codes were higher in the MRI group (304.5 per 100) than in the mammography group (284.8 per 100), and the adjusted difference between groups (19.6 per 100, [95% CI, 8.6 to 30.7]) mostly comprised additional health care visits | Interval cancer: consistency NA, precise<br>Adverse events: consistency NA, imprecise<br>Downstream consequences: consistent, imprecise | In the trial, 59% invited to MRI screening attended; possible unmeasured differences between population invited to screening and those attending (eg, breast cancer risk, concerns about false-positives and overdiagnosis)<br>Screening outcomes in the control group at round 2 not available, limiting interpretation of results<br>NRSI was based on US insurance claims with no clinical data to determine if follow-up was causally linked to breast screening | Interval cancer: low for reduced interval cancers with invitation to MRI<br>Adverse events: insufficient<br>Downstream consequences: low for increased follow-up | RCT conducted in the Netherlands through organized biennial breast screening program<br>Limited to women with extremely dense breasts identified using Volpara (category D)<br>Study randomized people with extremely dense breasts to MRI screening invitation—provides estimates of likely response and effects of invitation to MRI<br>No data on race or ethnicity for either study population<br>In the NRSI, 50% of individuals had a family history of breast cancer or genetic susceptibility |
| Supplemental screening with ultrasound | 1 RCT (n = 72 717)<br>1 NRSI (n = 18 562) | Interval cancer: RCT of supplemental US did not find statistical difference in invasive interval cancer (0.4 [DM/ultrasound] vs 0.8 [DM] per 1000 screened; RR, 0.58 [95% CI, 0.31 to 1.08]), nor did NRSI using BCSC data (1.5 [DM/ultrasound] vs 1.9 [DM] per 1000 screened; adjusted RR, 0.67 [95% CI, 0.33 to 1.37])<br>Downstream consequences of supplemental imaging (including incidental findings): the RCT reported recall attributable to positive findings only on ultrasound resulting in an additional 50 recalls per 1000 screened, of which 48 were false-positives; incidental findings were not reported<br>The BCSC NRSI found that referral to biopsy and false-positive biopsy were twice as high for women who underwent ultrasound | Interval cancer: consistent, imprecise<br>Downstream consequences: consistent, imprecise | Interval cancers rare in young women enrolled in RCT (aged 40-49 y), limited power to detect differences.<br>Population-averaged GEE effect estimate for interval cancer reported in RCT including DCIS lesions (23% of control group interval tumors) was statistically significant; second-round results not yet published<br>NRSI used propensity score matching to adjust for potential confounding by indication for screening; unmeasured confounding may still affect results<br>Ultrasound and mammography results not reported separately; therefore, attribution of follow-up specifically to ultrasound screening not possible | Interval cancer: low for no difference<br>Downstream consequences: low for increased follow-up with ultrasound | RCT conducted in Japan; included people aged 40 to 49 y; 23% of study population prevalence screened; 58% reported to have dense breasts, distribution not reported; ultrasound and DM results interpreted independently; performance could differ if considered together<br>BCSC NRSI included population representative of US overall; age included 30 to ≥80 y; inadequate numbers for comparisons of effects by race and ethnicity (80% White, non-Hispanic); 31% had a first-degree family history of breast cancer |

Abbreviations: BCSC, Breast Cancer Surveillance Consortium; DCIS, ductal carcinoma in situ; DBT, digital breast tomosynthesis; DM, digital mammography; EMR, electronic medical record; GEE, generalized estimating equations; HR, hazard ratio; KQ, key question; MRI, magnetic resonance imaging; NA, not available; NRSI, nonrandomized study of intervention; OR, odds ratio; PROSPR, Population-based Research Optimizing Screening Through Personalized Regimens; RCT, randomized clinical trial; RD, risk difference; RR, relative risk; sDM, synthetic mammography.

[a] Comparisons with no included studies are not included in this Summary of Evidence Table. This includes DBT vs DM (KQ1), age to start/stop screening (KQ2), supplemental screening with MRI (KQ1, KQ2), supplemental screening with ultrasound (KQ1, KQ2), and personalized screening (KQ1, KQ2, KQ3).

[b] This N includes individuals who were likely included in more than 1 of the studies that analyzed screening populations obtaining care at sites involved in the US BCSC.

reduces breast cancer morbidity and mortality. Consistent with trial findings, a nonrandomized BCSC study did not find reduced risks of advanced or interval cancers with DBT.[44] Limited evidence from trials on harms of screening with DBT[27,33,42] indicated similar false-positive recall and biopsy rates. An observational BCSC study did not show differences in the 10-year cumulative false-positive biopsy rates[32]; lower false-positive recall and biopsy with DBT screening were attenuated after several screening rounds.[44] Additional research is needed to ascertain whether DBT-based screening would reduce false-positives over a lifetime of screening.

The evidence was not adequate to evaluate the benefits and harms of supplemental MRI screening for people with dense breasts. No eligible studies were identified that provide evidence on breast cancer morbidity or mortality outcomes with supplemental MRI screening compared with mammography alone among individuals with dense breasts. The DENSE trial[45] reported fewer interval cancers with 1 round of supplemental MRI screening, but results from a second screening round are not yet published. Evidence of higher advanced cancer incidence in the mammography-only group relative to the MRI group would be needed to anticipate effects on morbidity or mortality. Supplemental MRI led to additional false-positive recalls and biopsies, and uncommon but serious adverse events were observed.[45] Two recent systematic reviews of the test performance literature reported higher cancer detection with supplemental MRI screening along with substantially increased recall and biopsy rates among individuals without cancer.[48,49]

Lack of a standardized and reliable assessment tool for measuring breast density and density variation across the lifespan pose challenges for research into the optimal screening strategy for persons with dense breasts.[16] Research is also needed to evaluate personalized risk-based screening, based on breast cancer risk factors and personal screening preferences. The ongoing WISDOM trial and My Personalized Breast Screening study (expected completion in 2025) may help to address these research gaps.[50,51]

Breast cancer is an active area of research, yet few longitudinal RCTs comparing different screening strategies have been conducted following completion of the major trials that established the effectiveness of mammography for reducing breast cancer mortality for women aged 50 to 69 years. This review included 6 new randomized trials,[27,31,33,40,42,45] 4 comparing DBT with digital mammography screening[27,31,33,42] and 2 on supplemental screening compared with mammography alone.[40,45] Three of these trials are ongoing[31,40,45] and have reported preliminary results only. Observational studies were also included, but few studies were available that followed up a screening population over time to compare the health outcomes associated with different screening approaches. These studies, while potentially more representative of a screening population, have higher risk of biased results due to confounding and selection.

### Limitations

Changes in population health, imaging technologies, and available treatments may limit the applicability of previous studies. Recent trials included in this review were conducted outside of the US and enrolled mostly White European populations. No studies evaluated screening outcomes for racial or ethnic groups in the US that experience health inequities and higher rates of breast cancer mortality. Black women are at highest risk of breast cancer mortality,[52] with lower 5-year survival than all other race and ethnicity groups.[7] Breast cancer mortality risk also increases at younger ages for Black women compared with White women.[53] This review did not address additional factors beyond screening that contribute to breast cancer mortality inequities.[54] Rigorous research is essential to understand and identify improvements needed along the pathway from screening to treatment[55] and to address inequities in follow-up time after a positive screening result, time to diagnosis,[56-60] and receipt of high-quality treatment and support services.[59,61,62]

Evidence comparing outcomes for different screening intervals and ages to start and stop screening was limited or absent. Trials of personalized screening based on risk and patient preferences are in progress and may address evidence gaps related to optimal screening start ages and intervals. Research is needed to better characterize potential harms of screening, including patient perspectives on experiencing false-positive screening results. Women with false-positive screening results may be less likely to return for their next scheduled mammogram, as reported in a large US health system study.[55,63] Rigorous studies that enroll screening populations and report advanced cancer detection, morbidity, and mortality outcomes from multiple rounds of screening are needed to overcome persistent limitations in the evidence on breast cancer screening. Multiple screening rounds are essential to determine whether a screening modality or strategy reduces the risk of advanced cancer by detecting early cancers that would otherwise have progressed (stage shift), potentially reducing breast cancer morbidity and mortality.[20-23,64]

The potential benefits of risk-stratified screening strategies, including the use of supplemental screening with ultrasound or MRI, have not been fully evaluated, although some harms are evident. Longer term follow-up on existing comparative effectiveness trials, complete results from ongoing RCTs of personalized screening programs,[65,66] and rigorous new studies are needed to further strengthen the evidence and optimize breast cancer screening strategies.

## Conclusions

Evidence comparing the effectiveness of different breast cancer screening strategies is inconclusive because key studies have not yet been completed and few studies have reported the stage shift or mortality outcomes necessary to assess relative benefits.

ARTICLE INFORMATION

**Accepted for Publication:** November 23, 2023.

**Published Online:** April 30, 2024. doi:10.1001/jama.2023.25844

**Author Contributions:** Dr Henderson had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
*Concept and design:* All authors.
*Acquisition, analysis, or interpretation of data:* All authors.
*Drafting of the manuscript:* All authors.
*Critical review of the manuscript for important intellectual content:* Henderson, Weyrich, Miller.
*Statistical analysis:* Henderson.
*Administrative, technical, or material support:* Webber, Melnikow.
*Supervision:* Henderson.

**Conflict of Interest Disclosures:** None reported.

**Funding/Support:** This research was funded under contract number 75Q80120D00004, Task Order 75Q80121F32004, from the Agency for Healthcare Research and Quality (AHRQ), US Department of Health and Human Services.

**Role of the Funder/Sponsor:** Investigators worked with USPSTF members and AHRQ staff to develop

the scope, analytic framework, and key questions for this review. AHRQ had no role in study selection, quality assessment, or synthesis. AHRQ staff provided project oversight, reviewed the report to ensure that the analysis met methodological standards, and distributed the draft for peer review. Otherwise, AHRQ had no role in the conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript findings.

**Disclaimer:** The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or the US Department of Health and Human Services.

**Additional Contributions:** The authors gratefully acknowledge the following individuals for their contributions to this project: Howard Tracer, MD (AHRQ); Heidi D. Nelson, MD, MPH, MACP (Kaiser Permanente Bernard J. Tyson School of Medicine); current and former members of the USPSTF who contributed to topic deliberations; and Evidence-based Practice Center staff members Melinda Davies, MA, Jill Pope, and Leslie A. Purdue, MPH, for technical and editorial assistance at the Kaiser Permanente Center for Health Research. USPSTF members, peer reviewers, and federal partner reviewers did not receive financial compensation for their contributions.

**Additional Information:** A draft version of this evidence report underwent external peer review from 5 content and methods experts (Nehmat Houssami, MBBS, MPH, Med, PhD [University of Sydney-Australia]; Patricia Ganz, MD [UCLA]; Gerald Gartlehner, MD, MPH [Cochrane Austria]; Karla Kerlikowske, MD [UC San Francisco]; Lisa Newman, MD, MPH [New York Presbyterian/Weill Cornell Medical Center]) and 4 scientific representatives from 3 federal partner organizations (Centers for Disease Control and Prevention; Office of Research on Women's Health; National Institute on Minority Health and Health Disparities). Comments were presented to the USPSTF during its deliberation of the evidence and were considered in preparing the final evidence review.

**Editorial Disclaimer:** This evidence report is presented as a document in support of the accompanying USPSTF Recommendation Statement. It did not undergo additional peer review after submission to *JAMA*.

## REFERENCES

1. SEER Cancer Stat Facts: female breast cancer. National Cancer Institute. Published 2023. Accessed July 31, 2023. https://seer.cancer.gov/statfacts/html/breast.html

2. US Cancer Statistics Working Group. US Cancer Statistics Data Visualizations Tool, based on 2022 submission data (1999-2020). Centers for Disease Control and Prevention. Updated June 2023. Accessed July 31, 2023. https://www.cdc.gov/cancer/dataviz

3. Cancer Statistics Explorer Network. Breast cancer: SEER incidence rates by age at diagnosis, 2016-2020, by sex (based on November 2022 SEER incidence data). National Cancer Institute. Published 2022. Accessed July 31, 2023. https://seer.cancer.gov/statistics-network/explorer/application.html

4. Cancer Statistics Explorer Network. Breast cancer: SEER 5-year age-adjusted incidence rates, 2015-2019. National Cancer Institute. Published 2022. Accessed February 16, 2023. https://seer.cancer.gov/statistics-network/explorer/application.html

5. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(6):438-451. doi:10.3322/caac.21583

6. American Association for Cancer Research. AACR cancer disparities progress report 2020. Published 2020. Accessed September 23, 2020. https://cancerprogressreport.aacr.org/disparities/

7. Cancer Statistics Explorer Network. Breast cancer: SEER 5-year relative survival rates, 2013-2019, by race/ethnicity (based on November 2022 SEER incidence data). National Cancer Institute. Published 2022. Accessed February 16, 2023. https://seer.cancer.gov/statistics-network/explorer/application.html

8. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 US Preventive Services Task Force recommendation. *Ann Intern Med*. 2016;164(4):244-255. doi:10.7326/M15-0969

9. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey L. Harms of breast cancer screening: systematic review to update the 2009 US Preventive Services Task Force recommendation. *Ann Intern Med*. 2016;164(4):256-267. doi:10.7326/M15-0970

10. Siu AL; US Preventive Services Task Force. Screening for breast cancer: US Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2016;164(4):279-296. doi:10.7326/M15-2886

11. Trentham-Dietz A, Chapman CH, Jayasekera J, et al. Collaborative modeling to compare different breast cancer screening strategies: a decision analysis for the US Preventive Services Task Force. *JAMA*. Published April 30, 2024. doi:10.1001/jama.2023.24766

12. US Preventive Services Task Force. Screening for breast cancer: US Preventive Services Task Force recommendation statement. *JAMA*. Published April 30, 2024. doi:10.1001/jama.2024.5534

13. US Preventive Services Task Force. US Preventive Services Task Force Procedure Manual. Published 2023. Accessed March 18, 2024. https://www.uspreventiveservicestaskforce.org/uspstf/about-uspstf/methods-and-processes/procedure-manual

14. Henderson JT, Webber EM, Weyrich M, Miller M, Melnikow J. *Screening for Breast Cancer: A Comparative Effectiveness Review for the US Preventive Services Task Force. Evidence Synthesis No. 231*. Agency for Healthcare Research and Quality; 2024. AHRQ publication 23-05303-EF-1.

15. Nelson HD, O'Meara ES, Kerlikowske K, Balch S, Miglioretti D. Factors associated with rates of false-positive and false-negative results from digital mammography screening: an analysis of registry data. *Ann Intern Med*. 2016;164(4):226-235. doi:10.7326/M15-0971

16. Melnikow J, Fenton JJ, Whitlock EP, et al. Supplemental screening for breast cancer in women with dense breasts: a systematic review for the US Preventive Services Task Force. *Ann Intern Med*. 2016;164(4):268-278. doi:10.7326/M15-1789

17. Wang S, Sultana F, Kavanagh A, et al. Benefits and harms of breast cancer screening: cohort study of breast cancer mortality and overdiagnosis. *Cancer Med*. 2023;12(17):18120-18132. doi:10.1002/cam4.6373

18. Richman IB, Long JB, Soulos PR, Wang SY, Gross CP. Estimating breast cancer overdiagnosis after screening mammography among older women in the United States. *Ann Intern Med*. 2023;176(9):1172-1180. doi:10.7326/M23-0133

19. National Comprehensive Cancer Network. NCCN Guidelines: breast cancer screening and diagnosis. Published 2022. Accessed March 10, 2023. http://www.nccn.org/guidelines/guidelines-detail?category=2&id=1421

20. Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness. *N Engl J Med*. 2016;375(15):1438-1447. doi:10.1056/NEJMoa1600249

21. Canelo-Aybar C, Carrera L, Beltrán J, et al. Digital breast tomosynthesis compared to diagnostic mammographic projections (including magnification) among women recalled at screening mammography: a systematic review for the European Commission Initiative on Breast Cancer (ECIBC). *Cancer Med*. 2021;10(7):2191-2204. doi:10.1002/cam4.3803

22. Ryser MD, Lange J, Inoue LYT, et al. Estimation of breast cancer overdiagnosis in a U.S. breast screening cohort. *Ann Intern Med*. 2022;175(4):471-478. doi:10.7326/M21-3577

23. Dunn BK, Woloshin S, Xie H, Kramer BS. Cancer overdiagnosis: a challenge in the era of screening. *J Natl Cancer Cent*. 2022;2(4):235-242. doi:10.1016/j.jncc.2022.08.005

24. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. doi:10.1136/bmj.i4919

25. Berkman ND, Lohr KN, Ansari M, et al. AHRQ Methods for Effective Health Care: grading the strength of a body of evidence when assessing health care interventions for the effective health care program of the Agency for Healthcare Research and Quality: an update. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Agency for Healthcare Research and Quality. Published 2008. Accessed March 18, 2024. https://www.ncbi.nlm.nih.gov/books/NBK174881/

26. Breast Screening Frequency Trial Group; United Kingdom Co-ordinating Committee on Cancer Research. The frequency of breast cancer screening: results from the UKCCCR Randomised Trial. *Eur J Cancer*. 2002;38(11):1458-1464. doi:10.1016/S0959-8049(01)00397-5

27. 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):1778-1790. doi:10.1002/ijc.34161

28. 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):109-116. doi:10.1007/s10549-016-3695-1

29. Ganguli I, Keating NL, Thakore N, Lii J, Raza S, Pace LE. Downstream mammary and extramammary cascade services and spending following screening breast magnetic resonance imaging vs mammography among commercially

insured women. *JAMA Netw Open*. 2022;5(4):e227234. doi:10.1001/jamanetworkopen.2022.7234

**30**. García-Albéniz X, Hernán MA, Logan RW, Price M, Armstrong K, Hsu J. Continuation of annual screening mammography and breast cancer mortality in women older than 70 years. *Ann Intern Med*. 2020;172(6):381-389. doi:10.7326/M18-1199

**31**. Heindel W, Weigel S, Gerß J, et al; TOSYMA Screening Trial Study Group. 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):601-611. doi:10.1016/S1470-2045(22)00194-2

**32**. Ho TH, 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 Netw Open*. 2022;5(3):e222440. doi:10.1001/jamanetworkopen.2022.2440

**33**. Hofvind S, Moshina N, Holen AS, 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. doi:10.1148/radiol.2021203936

**34**. Hovda T, Holen AS, Lång K, et al. Interval and consecutive round breast cancer after digital breast tomosynthesis and synthetic 2D mammography versus standard 2d digital mammography in BreastScreen Norway. *Radiology*. 2020;294(2):256-264. doi:10.1148/radiol.2019191337

**35**. Johnson K, Lång K, Ikeda DM, Åkesson A, Andersson I, Zackrisson S. Interval breast cancer rates and tumor characteristics in the prospective population-based Malmö breast tomosynthesis screening trial. *Radiology*. 2021;299(3):559-567. doi:10.1148/radiol.2021204106

**36**. 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):2220-2230. doi:10.1001/jama.2022.7672

**37**. 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 Intern Med*. 2019;179(5):658-667. doi:10.1001/jamainternmed.2018.8372

**38**. McGuinness JE, Ueng W, Trivedi MS, et al. Factors associated with false positive results on screening mammography in a population of predominantly hispanic women. *Cancer Epidemiol Biomarkers Prev*. 2018;27(4):446-453. doi:10.1158/1055-9965.EPI-17-0009

**39**. Miglioretti DL, Zhu W, Kerlikowske K, et al; Breast Cancer Surveillance Consortium. Breast tumor prognostic characteristics and biennial vs annual mammography, age, and menopausal status. *JAMA Oncol*. 2015;1(8):1069-1077. doi:10.1001/jamaoncol.2015.3084

**40**. Ohuchi N, Suzuki A, Sobue T, et al; J-START Investigator Groups. 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):341-348. doi:10.1016/S0140-6736(15)00774-6

**41**. Parvinen I, Chiu S, Pylkkänen L, et al. Effects of annual vs triennial mammography interval on breast cancer incidence and mortality in ages 40-49 in Finland. *Br J Cancer*. 2011;105(9):1388-1391. doi:10.1038/bjc.2011.372

**42**. Pattacini P, Nitrosi A, Giorgi Rossi P, et al; RETomo Working Group. A randomized trial comparing breast cancer incidence and interval cancers after tomosynthesis plus mammography versus mammography alone. *Radiology*. 2022;303(2):256-266. doi:10.1148/radiol.211132

**43**. Richman IB, Long JB, Hoag JR, et al. Comparative effectiveness of digital breast tomosynthesis for breast cancer screening among women 40-64 years old. *J Natl Cancer Inst*. 2021;113(11):1515-1522. doi:10.1093/jnci/djab063

**44**. Sprague BL, Coley RY, Lowry KP, et al. Digital breast tomosynthesis versus digital mammography screening performance on successive screening rounds from the Breast Cancer Surveillance Consortium. *Radiology*. 2023;307(5):e223142. doi:10.1148/radiol.223142

**45**. Veenhuizen SGA, de Lange SV, Bakker MF, et al; DENSE Trial Study Group. Supplemental breast MRI for women with extremely dense breasts: results of the second screening round of the DENSE trial. *Radiology*. 2021;299(2):278-286. doi:10.1148/radiol.2021203633

**46**. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

**47**. Aase HS, Holen AS, Pedersen K, et al. A randomized controlled trial of digital breast tomosynthesis versus digital mammography in population-based screening in Bergen: interim analysis of performance indicators from the To-Be trial. *Eur Radiol*. 2019;29(3):1175-1186. doi:10.1007/s00330-018-5690-x

**48**. Mizzi D, Allely C, Zarb F, et al. Examining the effectiveness of supplementary imaging modalities for breast cancer screening in women with dense breasts: a systematic review and meta-analysis. *Eur J Radiol*. 2022;154:110416. doi:10.1016/j.ejrad.2022.110416

**49**. Yuan WH, Hsu HC, Chen YY, Wu CH. Supplemental breast cancer-screening ultrasonography in women with dense breasts: a systematic review and meta-analysis. *Br J Cancer*. 2020;123(4):673-688. doi:10.1038/s41416-020-0928-1

**50**. Shieh Y, Eklund M, Madlensky L, et al. Breast cancer screening in the precision medicine era: risk-based screening in a population-based trial. *J Natl Cancer Inst*. 2017;109(5). doi:10.1093/jnci/djw290

**51**. Roux A, Cholerton R, Sicsic J, et al. Study protocol comparing the ethical, psychological and socio-economic impact of personalised breast cancer screening to that of standard screening in the "My Personal Breast Screening" (MyPeBS) randomised clinical trial. *BMC Cancer*. 2022;22(1):507. doi:10.1186/s12885-022-09484-6

**52**. Cancer Statistics Explorer Network. Breast cancer: US mortality rates by age at death, 2016-2020, by race/ethnicity (based on us mortality data 1969-2020). National Cancer Institute. Published 2022. Accessed February 16, 2023. https://seer.cancer.gov/statistics-network/explorer/application.html

**53**. Chen T, Kharazmi E, Fallah M. Race and ethnicity-adjusted age recommendation for initiating breast cancer screening. *JAMA Netw Open*. 2023;6(4):e238893. doi:10.1001/jamanetworkopen.2023.8893

**54**. Jatoi I, Sung H, Jemal A. The emergence of the racial disparity in US breast-cancer mortality. *N Engl J Med*. 2022;386(25):2349-2352. doi:10.1056/NEJMp2200244

**55**. Mottram R, Knerr WL, Gallacher D, et al. Factors associated with attendance at screening for breast cancer: a systematic review and meta-analysis. *BMJ Open*. 2021;11(11):e046660. doi:10.1136/bmjopen-2020-046660

**56**. Fayanju OM, Ren Y, Stashko I, et al. Patient-reported causes of distress predict disparities in time to evaluation and time to treatment after breast cancer diagnosis. *Cancer*. 2021;127(5):757-768. doi:10.1002/cncr.33310

**57**. Warner ET, Tamimi RM, Hughes ME, et al. Time to diagnosis and breast cancer stage by race/ethnicity. *Breast Cancer Res Treat*. 2012;136(3):813-821. doi:10.1007/s10549-012-2304-1

**58**. Kovar A, Bronsert M, Jaiswal K, et al. The waiting game: how long are breast cancer patients waiting for definitive diagnosis? *Ann Surg Oncol*. 2020;27(10):3641-3649. doi:10.1245/s10434-020-08484-9

**59**. Elmore JG, Nakano CY, Linden HM, Reisch LM, Ayanian JZ, Larson EB. Racial inequities in the timing of breast cancer detection, diagnosis, and initiation of treatment. *Med Care*. 2005;43(2):141-148. doi:10.1097/00005650-200502000-00007

**60**. Lawson MB, Bissell MCS, Miglioretti DL, et al. Multilevel factors associated with time to biopsy after abnormal screening mammography results by race and ethnicity. *JAMA Oncol*. 2022;8(8):1115-1126. doi:10.1001/jamaoncol.2022.1990

**61**. Emerson MA, Golightly YM, Aiello AE, et al. Breast cancer treatment delays by socioeconomic and health care access latent classes in Black and White women. *Cancer*. 2020;126(22):4957-4966. doi:10.1002/cncr.33121

**62**. Selove R, Kilbourne B, Fadden MK, et al. Time from screening mammography to biopsy and from biopsy to breast cancer treatment among Black and White, women Medicare beneficiaries not participating in a health maintenance organization. *Womens Health Issues*. 2016;26(6):642-647. doi:10.1016/j.whi.2016.09.003

**63**. Dabbous FM, Dolecek TA, Berbaum ML, et al. Impact of a false-positive screening mammogram on subsequent screening behavior and stage at breast cancer diagnosis. *Cancer Epidemiol Biomarkers Prev*. 2017;26(3):397-403. doi:10.1158/1055-9965.EPI-16-0524

**64**. Feng X, Zahed H, Onwuka J, et al. Cancer stage compared with mortality as end points in randomized clinical trials of cancer screening: a systematic review and meta-analysis. *JAMA*. Published April 7, 2024. doi:10.1001/jama.2024.5814

**65**. Women Informed to Screen Depending on Measures of Risk (Wisdom Study) [NCT02620852]. Updated December 8, 2022. Accessed August 7, 2023. https://www.clinicaltrials.gov/study/NCT02620852

**66**. My Personalized Breast Screening (MyPeBS) [NCT03672331]. Updated June 15, 2023. Accessed August 7, 2023. https://clinicaltrials.gov/study/NCT03672331