BSC6.07	Digital Breast Tomosynthesis		
Original Policy Date:	June 28, 2013	**Effective Date:**	February 1, 2019
Section:	6.0 Radiology	**Page:**	Page 1 of 33

Policy Statement

Digital mammography and digital breast tomosynthesis may be considered **medically necessary** when used for breast cancer *screening* purposes (i.e., as a preventive service).

Repeat digital mammography and repeat digital breast tomosynthesis may be considered **medically necessary** when used for breast cancer *diagnostic* purposes, provided that written radiologic interpretation of a prior digital mammogram or a prior digital breast tomosynthesis study documents the specific, medically necessary requirement for further imaging.

Policy Guidelines

The following CPT codes are specific for digital breast tomosynthesis:

- **77061**: Diagnostic digital breast tomosynthesis; unilateral
- **77062**: Diagnostic digital breast tomosynthesis; bilateral
- **77063**: Screening digital breast tomosynthesis, bilateral (List separately in addition to code for primary procedure)

The CPT Code 77063 is an add-on code, which can be reported with 77067.

Medicare established an add-on HCPCS G code specific to diagnostic breast tomosynthesis:

- **G0279***: Diagnostic digital breast tomosynthesis, unilateral or bilateral (list separately in addition to 77065 or 77066)

***Note**: The Centers for Medicare and Medicaid Services (CMS) established G-codes for certain radiation oncology services in place of the new 2015 CPT codes.

CPT codes 77061, 77062, and 77063 cannot be reported with the 3D rendering codes 76376 and 76377.

Description

Digital breast tomosynthesis (DBT) uses modified digital mammography equipment to obtain additional radiographic data that are used to reconstruct cross-sectional “slices” of breast tissue which are then assembled into a three-dimensional image of the breast. Conventional digital mammography (DM), or " a traditional mammogram,” gives only a two-dimensional image of the breast. Tomosynthesis may improve the accuracy of digital mammography by reducing problems caused by overlapping tissue. Tomosynthesis typically involves additional imaging time and radiation exposure, although recent improvements may change this.

Related Policies

- Magnetic Resonance Imaging for Detection and Diagnosis of Breast Cancer
- Positron Emission Mammography (PEM)
- Scintimammography and Gamma Imaging of the Breast and Axilla

Benefit Application

Benefit determinations should be based in all cases on the applicable contract language. To the extent there are any conflicts between these guidelines and the contract language, the contract language will control. Please refer to the member's contract benefits in effect at the

time of service to determine coverage or non-coverage of these services as it applies to an individual member.

Some state or federal mandates [e.g., Federal Employee Program (FEP)] prohibits plans from denying Food and Drug Administration (FDA)-approved technologies as investigational. In these instances, plans may have to consider the coverage eligibility of FDA-approved technologies on the basis of medical necessity alone.

Regulatory Status

Table 1 provides a summary of digital breast tomosynthesis (DBT) systems approved by the U.S. Food and Drug Administration (FDA) through the premarket approval process. FDA product code: OTE. The tomosynthesis portion of the mammography unit is considered a separate mammographic module, and in order for a facility to use this module, the facility must apply to the FDA for certification that extends to the tomosynthesis module. The U.S. Mammography Quality Standards Act requires interpreting physicians, radiologic technologists, and medical physicists to complete 8 hours of DBT training, and mandates a detailed mammography equipment evaluation before use.

Table 1. FDA-Approved DBT Systems

Device	Manufacturer	Date Approved	PMA	Indications
Selenia Dimensions 3D System	Hologic	Feb 2011	P080003	• Used to acquire 2D and 3D mammograms for screening and diagnosis of breast cancer. Screening mammogram may consist of 2D or 2D and 3D image set.
		May 2013	P080003/S001	• A hardware and software upgrade to the FFDM conventional mammography system. A 2D image can be generated from 3D image set.
		May 2017	P080003/S005	• Approval for the added indication of screening for women with dense breasts using 3D plus 2D imaging, where the 2D image can be either synthesized 2D or FFDM image vs FFDM alone
SenoClaire DBT System	GE Healthcare	Aug 2014	P130020	• A hardware and software upgrade to FFDM conventional mammography system. Same clinical applications as traditional mammography for screening mammography. A screening examination will consist of: a 2D image set consisting of a craniocaudal view and of a mediolateral oblique view, or a 2D craniocaudal view and 3D mediolateral oblique image set.
Senographe Pristina 3D		Mar 2017	P130020/S002	• Approval for multiple projection views to produce 3D digital mammography images for screening and diagnosing breast cancer. Senographe uses similar DBT technology as SenoClaire and consists of software and hardware upgrade to reconstruct tomosynthesis images.
Mammomat Inspiration with Tomosynthesis Option	Siemens	Apr 2015	P140011	• A software upgrade to FFDM conventional mammography system. It produces multiple low-dose x-ray images used to create cross-sectional views. Indication is for a 2D image set or a 2D

Device	Manufacturer	Date Approved	PMA	Indications
		Jan 2016 Mar 2017	P140011/S002 P140011/S003	and 3D image set screening and diagnosing breast cancer. • Software update resolving an error that may occur during tomosynthesis reconstruction with breast thickness >90 mm • A software upgrade, indicated for use with the EMPIRE reconstruction algorithm for acquisition of 2D and 3D digital mammography images, to be used in screening and diagnosis of breast cancer.
Aspire Cristalle Digital Breast Tomosynthesis Option	Fujifilm Medical Systems USA	Jan 2017	P160031	Approved for screening and diagnosing breast cancer consisting of images acquired in (1) FFDM mode only or (2) FFDM image set and DBT image set acquired in the ST (standard) mode. FFDM image set and DBT image set must be acquired with normal dose setting and may be acquired in 1 compression (Tomo Set mode) or separate compressions (FFDM and DBT modes).
PowerLook® Tomo Detection Software	iCAD	Mar 2017	P160009	Approved for software device intended for radiologists while reading GE SenoClaire breast tomosynthesis exams. It detects up to 5 soft tissue densities (masses, architectural distortions, asymmetries) in the 3D tomosynthesis images and then blends with the standard 2D image. These images may be confirmed or dismissed by the radiologist in the DBT images.

DBT: digital breast tomosynthesis; FDA: Food and Drug Administration; FFDM: full field digital mammography; PMA: premarket approval; 2D: 2-dimensional; 3D: 3-dimensional.

Rationale

Background

Conventional Mammography

Conventional mammography produces 2-dimensional (2D) digital images of the breast. Overlapping tissue on a 2D image can mask suspicious lesions or make benign tissue appear suspicious, particularly in women with dense breast tissue. As a result, women may be recalled for additional mammographic spot views. Inaccurate results may lead to unnecessary biopsies and emotional stress, or to a potential delay in diagnosis. Spot views often are used to evaluate microcalcifications, opacities, or architectural distortions; to distinguish masses from overlapping tissue, and to view possible findings close to the chest wall or in the retroareolar area behind the nipple.[1] The National Cancer Institute has reported that approximately 20% of cancers are missed at mammography screening.[2] Average recall rates are approximately 10%, with an average cancer detection rate of 4.7 per 1000 screening mammography examinations.[3] The U.S. Mammography Quality Standards Act audit guidelines anticipate 2 to 10 cancers detected per 1000 screening mammograms.[4] Interval cancers, which are detected between screenings, tend to have poorer prognoses.[5]

Digital Breast Tomosynthesis

Digital breast tomosynthesis (DBT) was developed to improve the accuracy of mammography by capturing a group of tomograms of the breast, further clarifying areas of overlapping tissue. Developers proposed that its use would result in increased sensitivity and specificity, as well as fewer recalls due to inconclusive results.[6] DBT produces multiple low-dose images per view along

an arc over the breast. During breast tomosynthesis, the compressed breast remains stationary while the x-ray tube moves approximately 1 for each image in a 15 to 50 arc, acquiring 11 to 49 images.[7] These images are projected as cross-sectional "slices" of the breast, with each slice typically 1-mm thick. Adding breast tomosynthesis takes about ten seconds per view. In a study in a research setting, Gur et al (2009) reported a mean time (standard deviation) for interpretation of results was 1.22 (1.15) minutes for digital mammography and 2.39 (1.65) minutes for combined digital mammography and breast tomosynthesis.[8]

With conventional 2D mammography, breast compression helps decrease tissue overlap and improve visibility. By reducing problems with overlapping tissue, compression with breast tomosynthesis may be reduced by up to 50%. This change could result in improved patient satisfaction.[7]

A machine equipped with breast tomosynthesis can perform 2D digital mammography, DBT, or a combination of both 2D mammography and DBT during a single compression. Radiation exposure from tomosynthesis is roughly equivalent to mammography. Therefore, adding tomosynthesis to mammography doubles the radiation dose, although it still is below the maximum allowable dose established in the Mammography Quality Standards Act.

Studies typically compare 1-view (i.e., mediolateral oblique view), or more commonly, 2-view (mediolateral oblique plus craniocaudal view) breast tomosynthesis either alone or combined with standard 2D mammography, against standard 2D mammography alone. A 2014 Blue Cross Blue Shield Association Technology Evaluation Center (TEC) Assessment focused on 2-view tomosynthesis.[9] The U.S. Food and Drug Administration (FDA), which reviewed this new modality in 2011, recommended that 2-view breast tomosynthesis is preferable to 1-view tomosynthesis (both used in combination with full field digital mammography).[10]

The FDA (2013) approved new tomosynthesis software that permits the creation of 2D images (called C-View) from images obtained during tomosynthesis.[11] As a result, the performance of separate 2D mammography may become unnecessary, thereby lowering radiation dose. In other words, it is possible that only the tomosynthesis procedure will be needed, with the ability to create both conventional 2D and DBT images. It is too early to gauge how conventional 2D mammography plus tomosynthesis compares with C-View plus tomosynthesis.

Literature Review

This review was informed by a TEC Assessment (2014).[9]

Evidence reviews assess whether a medical test is clinically useful. A useful test provides information to make a clinical management decision that improves the net health outcome. That is, the balance of benefits and harms is better when the test is used to manage the condition than when another test or no test is used to manage the condition.

The first step in assessing a medical test is to formulate the clinical context and purpose of the test. The test must be technically reliable, clinically valid, and clinically useful for that purpose. Evidence reviews assess the evidence on whether a test is clinically valid and clinically useful. Technical reliability is outside the scope of these reviews, and credible information on technical reliability is available from other sources.

Digital Breast Tomosynthesis for Screening

Clinical Context and Text Purpose

The purpose of 3-dimensional (3D) DBT in patients who are being screened for breast cancer is to inform a decision whether to recall women for further diagnostic testing.

The question addressed in this portion of the review is whether there is sufficient evidence that 3D DBT, used to screen for breast cancer, improves the net health outcome compared with standard techniques. Specifically, is 3D DBT as an adjunct to 2-dimensional (2D) mammography

or 3D DBT plus synthesized 2D mammography superior to mammography alone, and is 3D DBT instead of mammography at least as beneficial as mammography? For both interventions, are differences in accuracy likely to improve health outcomes via earlier diagnosis and treatment?

The following PICOTS were used to select literature to inform this review.

Patients
The relevant population of interest are asymptomatic individuals being screened for breast cancer.

Interventions
The intervention of interest is 3D DBT screening as an adjunct to 2D mammography and 3D DBT plus synthesized 2D mammography. DBT devices approved in the United States are summarized in Table 1.

Comparators
The primary comparator of interest is mammography alone.

Outcomes
The reference standard is histopathology or at least one year follow-up for women with negative findings.

The health outcomes of interest are:

- Overall and breast cancer-specific survival
- Quality of life
- Recall rates, which may lead to unnecessary follow-up testing and possibly unnecessary biopsies and treatment
- Cancer risk from radiation exposure

For breast cancer, the most important health outcome is an overall survival from the disease. DBT, as any breast screening test, may not directly improve breast cancer-specific survival; however, higher sensitivity of breast DBT could lead to earlier cancer detection, which may, in turn, lead to improved health outcomes if earlier treatment is more effective. Although there is indirect evidence that earlier detection improves health outcomes, possible overdetection also needs to be taken into account. Overdetection would subject women to testing and treatment that does not improve health outcomes. When screening leads to diagnosis at an early stage, it may also affect the quality of life by permitting the use of less invasive or otherwise less difficult to tolerate treatments for breast cancer. If using breast DBT reduces the false-positive rate, it would reduce recalls for a diagnostic workup or for biopsy. Fewer unnecessary recalls would, in turn, have a positive impact on patient's quality of life by avoiding the anxiety and additional imaging associated with recalls. Finally, adding breast DBT to traditional mammography doubles the radiation dose, even though the combined dose remains below the limit set in the Mammography Quality Standards Act of 1992. The increased dose might be offset in part by fewer diagnostic tests if the recall rate falls. If synthesized mammography permits the use of tomosynthesis to create both 2D and 3D images, then the dose would be roughly equivalent to a single mammogram and increased radiation exposure would no longer be an issue.

Timing
At least one year of follow-up is needed to detect interval cancers that were false-negatives at initial screening.

Setting
DBT would be performed in an outpatient imaging setting.

Study Selection Criteria
For the evaluation of clinical validity of DBT, studies that met the following eligibility criteria were preferred:

- Prospective studies (preferably in a U.S. setting)
- Comparing DBT plus mammography with mammography alone
- Including asymptomatic individuals being screened for breast cancer
- Including performance characteristics such as screening sensitivity and specificity (i.e., follow-up of negative findings and interval cancers for at least one year)
- Several studies did not meet the preferred selection criteria, in particular, most lacked data on follow-up of negative findings and interval cancers. The prospective studies without sufficient follow-up of negative findings are summarized briefly in tabular form following the discussion of studies with follow-up of negative findings.

Technically Reliable
Assessment of technical reliability focuses on specific tests and operators and requires review of unpublished and often proprietary information. Review of specific tests, operators, and unpublished data are outside the scope of this evidence review and alternative sources exist. This evidence review focuses on the clinical validity and clinical utility.

3D DBT as an Adjunct to 2D Mammography
Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Prospective Studies with Long-Term Follow-Up of Negative Findings
Characteristics of prospective studies with follow-up of negative findings are shown in Table 2. The table includes two publications of previously reported prospective studies that have provided additional data including follow-up for interval cancers.

Houssami et al (2018) reported on the results from STORM (Screening with Tomosynthesis OR standard Mammography) study, which assessed interval breast cancers, based on ascertainment at 2-year follow-up from screening examinations.[12] STORM examined comparative cancer detection for traditional mammography with or without DBT in a general population of 7292 asymptomatic Italian women being screened for breast cancer. In the initial screening of STORM, women were recalled if either of two independent readers recorded a positive result at either mammography alone or mammography plus DBT. Previous reports of STORM have summarized initial findings of one round of screening and partial follow-up of the cohort (summarized in the following section). The 2018 report focused on screening measures requiring completed ascertainment of interval cancers, i.e., interval cancer rates and screening sensitivity, including two years of follow-up. Interval cancers were identified using a combination of checking local hospital and pathology databases; and checking with the local cancer registry for cancer notifications. Interval cancer rates for concurrent Italian cohorts screened with 2D-mammography alone were provided for descriptive purposes. The study was not powered for formal comparisons of mammography alone to mammography plus DBT.

Similarly, Skaane et al (2018) reported performance indicators and characteristics of screen-detected and interval cancers from 24301 women in the Oslo Tomosynthesis Screening Trial (OTST) administered by the Norwegian Cancer Registry.[13] OTST was designed to compare four different reading modes for mammography with or without DBT. The results reported herein include the double-reading mammography plus DBT and double-reading mammography alone arms. Decisions regarding recalls from the initial screens were made by consensus conference review of images that were rated by any reader as any score other than negative or definitely benign. Previous reports from OTST included initial results of one round of screening without the follow-up of negative results (summarized in the following section). The 2018 publication reported a comparison of mammography plus DBT in women from OTST who had 2 years of follow-up with

2 previous mammography screening rounds in Oslo using data from the Norwegian Cancer Registry.

Table 2. Characteristics of Prospective Studies with Long-Term Follow-Up of Negatives

Study	Study Population	Reference Standard	Threshold for Positive Index Test	Timing of Reference and Index Tests	Blinding of Assessors	Comment
STORM[12]	Asymptomatic women ≥48 y attending biennial screening in Italy, 2011 to 2012	Pathology; 2-y follow-up of negatives	Double-reading by radiologists experienced in mammography	Within 24 mo of screening episode	Yes	STORM was not designed to compare interval cancer data
OTST[13]	Women ages 50-69 y, invited biennially for screening in Norway, 2010 to 2012	Pathology; 2-y follow-up of negatives	Consensus decision of multiple radiologists	Within 24 mo of screening episode	Yes	Comparison group was not concurrent

OTST: Oslo Tomosynthesis Screening Trial; STORM: Screening with Tomosynthesis OR standard Mammography.

Results of prospective studies meeting with sufficient follow-up are shown in Table 3. Nine interval cancers were detected in STORM; three were diagnosed within one year of screening and the remaining six were diagnosed between one and two years after screening. STORM reported an interval breast cancer rate in mammography plus DBT screening participants that were numerically lower (and screening sensitivity numerically higher) than the rate in 2D-screened women although confidence intervals overlapped. These findings should be interpreted with caution given that STORM was not designed to compare interval cancer data and there were a small number of interval cases. Specificity was not reported in the publication; however, based on the information provided and the data on mammography plus DBT test results in the previous publications, it appears that the specificity was 96.6% (95% confidence interval CI., 96.2% to 97.0%) in the STORM participants.

Interval cancer rates were similar in women who received mammography alone and DBT plus mammography in the report including OTST participants. OTST also reported numerically but not statistically higher sensitivity while also reporting statistically higher specificity of mammography plus DBT compared with mammography alone. Most of the additional DBT-detected cancers in OTST were reported to be small node-negative invasive cancers of molecular subtypes known to have a good prognosis.

Table 3. Results of Prospective Studies with Long-Term Follow-Up of Negatives

Study	N	Interval Cancer Rate (95% CI)	Clinical Validity (95% CI), %			
			Sensitivity	Specificity	PPV	NPV
STORM[12]						
Mammo-only concurrent cohort	25,058	1.61/1000 negative screens (1.15 to 2.18)	77.3 (70.4 to 83.2)	NR	NR	NR
STORM participants[a]	7292	1.24/1000 negative screens (0.57 to 2.36)	85.5 (75.0 to 92.8)	NR	NR	NR
OTST[13]						

Study	N	Interval Cancer Rate (95% CI)	Clinical Validity (95% CI), %			
Mammo-only non-current cohorts	59,877	2.0/1000 screens	76.2	96.4	NR	NR
OTST participants	24,301	2.1/1000 screens	80.8	97.5	NR	NR
Difference		0.1 (-0.5 to 0.8)	4.6 (-1.4 to 10.5	1.2 (0.91 to 1.40)		

CI: confidence interval; mammo: mammography; NR: not reported; OTST: Oslo Tomosynthesis Screening Trial; STORM: Screening with Tomosynthesis OR standard Mammography.
[a]STORM participants were screened with both mammography and digital breast tomosynthesis. Woman were recalled if either of 2 independent readers recorded a positive result at either mammography alone or mammography plus digital breast tomosynthesis

The purpose of the gaps tables (see Tables 4 and 5) is to display notable gaps identified in each study. This information is synthesized as a summary of the body of evidence and provides the conclusions on the sufficiency of the evidence supporting the position statement. Neither STORM nor OTST was conducted in a U.S. setting and screening practices differ in European countries. While both studies included a prospective cohort of women receiving DBT plus mammography, the comparison group in the OTST study for the purposes of the 2018 publication was a cohort previously screened with mammography alone (i.e., not concurrent) and few details were provided on selection of the women included in that cohort.

Table 4. Relevance Gaps of Prospective Studies with Long-Term Follow-Up of Negatives

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
STORM[12]	4. Italian setting; screening practices differ from those in the U.S.			3. Only screening sensitivity is reported in 2018 paper	
OTST[13]	4. Norwegian setting; screening practices differ from those in the U.S.	3. Uses consensus of multiple readers unlike single-reader relevant to U.S. clinical setting			

The evidence gaps stated in this table are those notable in the current review; this is not a comprehensive gaps assessment. OTST: Oslo Tomosynthesis Screening Trial; STORM: Screening with Tomosynthesis OR standard Mammography.
[a]Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b]Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c]Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d]Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e]Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true positives, true negatives, false positives, false negatives cannot be determined).

Table 5. Study Design and Conduct Gaps of Prospective Studies with Long-Term Follow-Up of Negatives

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
STORM[12]						2. Comparisons not provided because study not powered to make comparisons for interval cancers
OTST[13]	2. Unclear if cohort from cancer registry was consecutive or randomly selected		2. Compared with previous rounds of mammography			

The evidence gaps stated in this table are those notable in the current review; this is not a comprehensive gaps assessment. OTST: Oslo Tomosynthesis Screening Trial; STORM: Screening with Tomosynthesis OR standard Mammography.
[a]Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b]Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c]Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d]Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e]Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f]Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison to other tests not reported.

Prospective Studies without Long-Term Follow-Up of Negative Results

Other prospective studies assessing the diagnostic accuracy of DBT for screening are summarized in Table 6. The table is subdivided by the characteristics of study designs. Select studies are summarized briefly following the table. In general, these studies do not have follow-up sufficient to capture interval cancers and therefore traditional measures of sensitivity and specificity are not provided.

Table 6. Prospective Studies of DBT for Breast Cancer Screening without Long-Term Follow-Up of Negatives

Study	No. Cancers/ No. Patients	Recalls/1000 Screens (95% CI)	PPV for Recalls (95% CI), %	Cancers Detected/1000 Screens (95% CI)	PPV for Biopsies (95% CI), %
Randomized Controlled Trials					
Pattacini et al (2018)[14]					
Mammo	44/9,783	35	13	4.5	NR
Mammo plus DBT	83/9,777	35	24	8.6	
p value			<0.001		
Maxwell et al (2017)[15]	11/1,227				
Mammo		28	NR	9.0	NR
Mammo plus DBT		27		10.6	
p value					
Prospective Observational Studies					
Patients served as their own controls					

Study	No. Cancers/ No. Patients	Recalls/1000 Screens (95% CI)	PPV for Recalls (95% CI), %	Cancers Detected/1000 Screens (95% CI)	PPV for Biopsies (95% CI), %
MBTST (2016)[16] (exploratory results)	68/7,500				
Mammo		26 (23 to 30)	24	6.3 (4.6 to 8.3)	NR
Mammo plus DBT		38 (33 to 42)	24	8.9 (6.9 to 11.3)	
p value		<0.001		<0.001	
Sumkin et al (2015)[17]	6/1,074	b			
Mammo		384	NR	4.7	NR
Mammo plus DBT		274		4.7	
Skaane (2013) OTST[18]	121/12,621				
Mammo		NR	28.5	6.1	NR
Mammo plus DBT			29.1	8.0	
p value				0.001	
STORM[19,20a]	59/7,292				
Mammo		42	11	5.3 (3.8 to 7.3)	NR
Mammo plus DBT		36	19	8.1 (6.2 to 10.4)	
p value				<0.001	
		Noncancer cases[d]			
Rafferty et al (2013)[21c]	51/997				
Study 1 (range)					
Mammo		551 (223-798)[e]	43	NR	NR
Mammo plus DBT		167 (76-284)[e]	56		
Study 2 (range)					
Mammo		488 (282-691)[e]	47	NR	NR
Mammo plus DBT		301 (198-413)[e]	50		
Includes s2D		False-Positive			
Mammo		Recall, %			
Bernardi et al (2016; STORM-2)[22]	90/9,672				
Mammo		3.42 (3.07 to 3.80)		6.3 (4.8 to 8.1)	NR
Mammo plus DBT		3.97 (3.59 to 4.38)		8.5 (6.7 to 10.5)	NR
s2D mammo plus DBT		4.45 (4.05 to 4.89)		8.8 (7.0 to 10.8)	NR

CI: confidence interval; DBT: digital breast tomosynthesis; DM: digital mammography; Mammo: mammography; MBTST: Malmö Breast Tomosynthesis Screening Trial; NR: not reported; OTST: Oslo Tomosynthesis Screening Trial; PPV: positive predictive value; s2D: synthesized 2D mammography; STORM: Screening with Tomosynthesis OR standard Mammography.
[a]Data from Ciatto et al (2013) and Houssami et al (2014).
[b]U.S. population; high-risk preferentially included.
[c]Twenty-seven women with no follow-up not included in results.
[d]U.S. population; sample enriched with women referred for biopsy (22%).
[e]Range across 12 radiologist in study 1 and 15 radiologists in study 2.

Randomized Controlled Trials

Two RCTs have compared screening with mammography alone with mammography plus DBT. Pattacini et al (2018) reported on the preliminary results from the Reggio Emilia Tomosynthesis trial, which compare mammography plus DBT with mammography alone in women in Italy ages 45 to 74 who had previously been screened with mammography.[14] The trial is designed to enroll 40000 women and compare interval cancers with cumulative incidence of advanced cancer and had 4.5 years of follow-up. The 2018 publication focuses on the preliminary results for the baseline screen of 19560 women recruited from 2014 to 2016, including cancers diagnosed within 9 months from recruitment and, as such, cannot yet provide data on interval cancers and confirmation of negative findings. Results are shown in Table 6.

Maxwell et al (2017) reported on the results of a trial of asymptomatic women from 2 centers in the U.K. ages 40 to 49 years who had previously undergone mammography for an increased risk of breast cancer.[15] Participants were randomized in a crossover design to screening with 2D

mammography followed by 2D mammography plus DBT a year later, or vice versa. The trial was designed to compare recall rates. Results are shown in Table 6. The crossover design limits the utility of collecting long-term results.

In summary, recall rates did not differ for mammography alone vs mammography plus DBT in either RCT. Maxwell et al (2017) also reported no statistically significant difference in cancer detection rate. However, preliminary results from Reggio Emilia Tomosynthesis trial would suggest an almost 90% increase in detection rate for mammography plus DBT compared with mammography (relative risk RR., 1.89; 95% CI, 1.31 to 2.72) and an increase in the PPV for recalls from 13.0% to 24.1%. The gain in cancer detection was observed for all classes of cancers except for very large or late cancers. There were more instances of ductal carcinoma in situ (DCIS) with mammography plus DBT (+1 per 1000), benign lesions (+1 per 1000), and invasive cancers (+3 per 1000). There was also an increase in the risk of surgery for mammography plus DBT (RR=1.90; 95% CI, 1.35, 2.68; risk difference, 5 per 1000; 95% CI, 2 to 7).

Prospective Observational Studies

Lång et al (2016) reported exploratory results from the first half of the Malmö Breast Tomosynthesis Screening Trial, comparing 1-view (mediolateral oblique) DBT (a lower radiation dose than digital mammography DM.) with 2-view DM.[16] The Malmö Breast Tomosynthesis Screening Trial is a 1-arm, single institution, prospective study. Randomly selected women in Sweden (age range, 40-74 years) were offered 1-view DBT and 2-view DM. A sample size of 15000 was specified to detect an improvement in cancer detection sensitivity from 63% to 88% (power, 80%); 7500 were included in the exploratory analysis. In Sweden, breast cancer screening is offered to women between ages 40 and 55 every 18 months and every 24 months after that to age 74. Six experienced readers interpreted images (mean experience, 26 years; range, 8-41 years). Blinded double-reading was carried out for DBT and DM with rule-based arbitration of disagreements women in this exploratory analysis were followed at least one year for the development of cancer ascertained through the South Swedish Cancer Registry. Of 10547 women invited, 71.1% participated with 20% undergoing their first screening test. Results are shown in Table 6. DCIS detection rates were similar between both modalities. Following arbitration, the recall rate was lower for DM (2.6%; 95% CI, 2.3% to 3.0%) than for DBT (3.8%; 95% CI, 3.3% to 4.2%; p<0.001).

The results of analysis of a cohort from a large trial, the OTST comparing 4 different reading modes, was published by Skaane et al (2013) in Norway.[18,23] The Skaane et al (2013) analysis was a preplanned interim analysis of two arms in a larger 4-arm trial; findings of the other two arms are not relevant to this topic. The sample included 12621 women with 121 cancers detected during routine screening.[24] Results are shown in Table 6. After adjusting for reader differences, the ratio of cancer detection rates for mammography plus DBT vs mammography alone was 1.27 (98.5% CI, 1.06 to 1.53; p=0.001). The trialists did not ascertain any increase in detecting DCIS by adding breast tomosynthesis (i.e., additional cancers detected were mostly invasive). In Norway, as in much of Europe, women are screened every other year, and two readers independently interpret the images, which differs from usual practice in the United States. After adjusting for differences across readers, the ratio of false-positive rates for mammography plus DBT vs mammography alone was 0.85 (98.5% CI, 0.76 to 0.96; p<0.001).

The STORM study examined comparative cancer detection for traditional mammography with or without DBT in a general population of 7292 asymptomatic Italian women being screened for breast cancer.[19,20] The reference standard was pathology results for women undergoing biopsies; women with negative results on both mammography and DBT were not followed so neither sensitivity nor specificity could be calculated. Results are shown in Table 6. Mammography plus DBT revealed all 59 cancers; 20 (34%) were missed by traditional mammography (p<0.001). In the original report, incremental cancer detection by using both modalities was 2.7 cancers per 1000 screens (95% CI, 1.7 to 4.2). There were 395 false-positive results: 181 were false-positive using either mammography or both imaging modalities together; an additional 141 occurred using mammography only, and 73 occurred using mammography

and DBT combined (p<0.001). In preplanned analyses, combined results of mammography and DBT yielded more cancers in both age groups (<60 vs >=60 years) and breast density categories (1 least dense. and 2 vs 3 and 4 most dense.). In a follow-up report including available data on interval cancers diagnosed in the first year of follow-up (note, screening was repeated at two years), six additional interval cancers had been diagnosed. The cancer detection rates including the 6 additional cancers were 4.8 (95% CI, 3.3 to 6.7) vs 7.5 (95% CI, 5.7 to 9.8) for mammography vs mammography plus DBT, for an incremental cancer detection rate of 2.7 (95% CI, 1.6 to 4.2; p<0.01).

Retrospective Studies
Several retrospective studies have also been performed, many of which included several thousand patients and 3 of which included more than 100000 patients. Many of the retrospective studies have included mixed populations or unclear indications for screening and inadequate reference standards such as historical controls and are therefore not discussed in detail. Results are summarized briefly in Appendix Table 1. Retrospective studies have, in general, suggested increases in the rates of cancer detection and decreases in recall and false-positive rates.

Characteristics of Detected Cancers
Yun et al (2017) published a meta-analysis assessing the characteristics of cancers detected with DM alone vs DM plus DBT during routine breast cancer screening.[25] Eleven studies were included in the meta-analysis, four prospective and seven retrospective observational studies, all of which are described in Table 2 (above). Reviewers evaluated study quality using the Quality Assessment of Diagnostic Accuracy Studies tool and found an overall satisfactory risk of bias, but all studies had a high-risk of bias concerning the reference standard as well as flow and timing because patients who were not recalled did not have a reference standard test (i.e., did not have biopsy-confirmed negative findings).

In a pooled analysis, the overall cancer detection rate was significantly higher with DM plus DBT than with DM alone (RR=1.29; 95% CI, 1.16 to 1.43; I2=0%). Moreover, the detection of invasive cancer was significantly higher in the DM plus DBT group compared with DM alone group (RR=1.33; 95% CI, 1.17 to 1.51; I2=7%). The rate of carcinoma in situ detection did not differ significantly between the DM plus DBT group and the DM alone group (RR=1.20; 95% CI, 0.94 to 1.52; I2=29%). Fewer studies reported on cancer detection by T and/or N stage. In a pooled analysis of 5 studies, there was a significantly higher rate of detecting T1 cancers with DM plus DBT than with DM alone (RR=1.39; 95% CI, 1.14 to 1.70; I2=0%), but no significant difference for detecting stage T2 or larger cancer (RR=1.39; 95% CI, 0.90 to 2.16; I2=0%). Similarly, there was a significantly higher rate of detection of stage N0 cancers with DM plus DBT than with DM alone (RR=1.45; 95% CI, 1.21 to 1.74; I2=0%) and no significant difference in the detection of stage N1 or higher cancers (RR=1.34; 95% CI, 0.92 to 1.99; I2=0%). The numbers of more advanced cancers were relatively small, and the pooled analyses of T2 or higher and N1 or higher cancers might have been underpowered. The findings of this meta-analysis were limited by the potential biases of the included studies (e.g., many were retrospective and studies had insufficient confirmatory data on negative imaging results).

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

There is no direct evidence from trials comparing health outcomes in patients screened for breast cancer using DBT and mammography.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

A chain of evidence should demonstrate that DBT used as an adjunct to screening improves screening performance compared with standard mammography alone. Available studies have reported that adding DBT to mammography may increase cancer detection and reduce unnecessary recalls. Even if adding breast tomosynthesis simply maintained the same sensitivity as mammography, a decline in the false-positive rate would reduce the substantial number of unnecessary diagnostic workups in the United States.

Two prospective studies (STORM, OTST) with two-year follow-up for interval cancers have been published although neither was conducted in the United States. OTST had prospective data on the mammography plus DBT cohort but compared outcomes with previously screened cohorts from a cancer registry. Neither study was powered to compare interval cancer rates. STORM reported an interval breast cancer rate in mammography plus DBT screening participants that were numerically lower (and screening sensitivity numerically higher) than the rate in 2D-screened women although CIs overlapped. OTST also reported numerically but not statistically higher sensitivity. However, OTST did report statistically significantly higher specificity of mammography plus DBT compared with DBT alone.

- Two RCTs without sufficient follow-up to detect interval cancers have reported no difference in recall rates between DBT plus mammography and mammography alone. However, 1 RCT reported approximately a 90% increase in detection rate for DBT plus mammography compared with mammography with more instances of DCIS with mammography plus DBT (+1 per 1000), benign lesions (+1 per 1000), and invasive cancers (+3 per 1000) and an increase in the PPV for recalls from 13.0% to 24.1%. This RCT is ongoing and is designed to compare interval cancers and cumulative incidence of advanced cancer with 4.5 years of follow-up at completion.
- While the incremental radiation per individual is not large, the aggregate impact of that increased radiation dose over a large group can raise greater concern. Although any elevated dose related to DBT may be offset by fewer diagnostic images required for women who are recalled for further evaluation, it needs to be considered. Synthesized mammography may resolve this issue (discussed in the following section).
- There has been widespread debate over the value of mammography that hinges in large part on beliefs about whether there is substantial overdetection of breast cancer during screening. An argument in favor of tomosynthesis is that the probability of overdetection is lower because most of the additional cancers detected are invasive. On the other hand, mammography is included with tomosynthesis in part because of concern that readers of tomosynthesis images may miss microcalcifications, some of which are malignant.

In summary, estimates of sensitivity and specificity of DBT plus mammography from studies with adequate follow-up of negative results are available from two studies. The sensitivity of DBT plus mammography is likely to be at least as high as mammography alone. One study with limitations reported the specificity of DBT plus mammography was significantly higher than mammography alone. An increase in specificity (corresponding to a decrease in the false-positives) would reduce unnecessary diagnostic workups and their consequences. Two RCTs with short follow-up reported similar recall rates for DBT plus mammography and mammography alone but one of the RCTs reported a significant increase in cancer detection rate, including invasive cancer and DCIS.

Subsection Summary: Screening with 3D DBT as an Adjunct to 2D Mammography
There is also a lack of direct evidence on the clinical utility of 3D DBT from screening trials comparing health outcomes in patients screened for breast cancer with 3D DBT vs 2D mammography. Current evidence would suggest that use of mammography plus breast tomosynthesis may modestly increase the number of cancers detected, with a potential decrease in the number of women who undergo unnecessary recalls or biopsies. A 2017 meta-analysis including a pooled analysis of 11 screening studies found a significantly higher rate of invasive cancer detection with 3D DBT plus 2D DM than with 2D DM alone. Preliminary data from an RCT also found higher rates of invasive cancer with 3D DBT plus DM.

3D DBT Plus Synthesized 2D Mammography
Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

No prospective studies with sufficient follow-up for interval cancers and negative findings were identified.

One systematic review of 3d DBT plus synthesized 2D (s2D) vs DM plus DBT for breast cancer screening has been published. Characteristics are shown in Table 7. Houssami et al (2018) included studies that evaluated s2D plus DBT compared with DM plus DBT for population screening and provided quantitative data on screening detection measures (cancer detection and recall measures).[26] Five studies were identified.[22,27-30] The studies included in the Houssami et al (2018) systematic review, with the exception of Skaane et al (2014)[30] all included a comparison of DM and DM plus DBT in addition to the sDM plus DBT arm and as such were included in Table 6 and Appendix Table 1.

Table 7. Characteristics of Systematic Reviews of DBT Plus s2D Mammography

Study	Dates	Studies	Participants	N (Range)	Design	Duration
Houssami et al (2018)[26]	Through Aug 2017	5	Received s2D or DM with DBT for population breast cancer screening	NR	Any design eligible (included 2 prospective, 3 retrospective)	NR

DBT: digital breast tomosynthesis; DM: digital mammography; s2D: synthesized 2-dimensional; NR: not reported.

Results of the systematic review are shown in Table 8. Meta-analyses were not conducted; instead, qualitative summaries were provided. Cancer detection rates appear similar between DM plus DBT (range, 5.45 to 8.5 per 1000 screens) and s2D plus DBT (range, 5.03-8.8 per 1000 screens). The recall rates appear heterogeneous across included studies. The mean glandular dose for s2D plus DBT was 55% to 58% of DM plus DBT. The systematic review did not include a risk of bias or quality assessment. However, all of the included studies had limitations similar to the studies in the previous setting, i.e., lack of follow-up for interval cancers or confirmation of negative results.

Table 8. Results of Systematic Reviews of DBT Plus s2D Mammography

Study	Breast Cancer Detect Rate (per 1000 screens)	Recall, %	Mean Glandular Dose, mGy
Houssami et al (2018)[26]			
Range of N	NR (5 studies)	NR (5 studies)	NR (3 studies)
Range of effect sizes			
DM	5.3 to 6.3/1,000	3.42 to 8.7[a]	1.36 to 3.77
DM plus DBT	5.45 to 8.5/1,000	3.97 to 8.8[a]	1.87 to 4.88
s2D plus DBT	5.03 to 8.8/1,000	4.3 to 7.1[a]	3.22 to 7.97

DBT: digital breast tomosynthesis; DM: digital mammography; NR: not reported; s2D: synthesized 2-dimensional.
[a]Two studies reported recall and 3 studies reported false-positive recall.

The Skanne et al (2014) study from the systematic review and other studies published following the systematic review are briefly summarized in Table 9.[30] None has sufficient follow-up to evaluate interval cancers.

Table 9. Other Studies of DBT plus sDM for Breast Cancer Screening

Study	No. Cancers/ No. Patients	Recalls/1000 Screens (95% CI)	PPV for Recalls (95% CI), %	Cancers Detected/1000 Screens (95% CI)	PPV for Biopsies (95% CI), %
Prospective Observational Studies					
Romero Martin et al (2018)[31]	98/16,067				
DM (double)		50	9.4	4.7	39.4
sDM plus DBT		29	18.0	5.4	46.0
p value		<0.001	<0.001	0.043	0.189
Caumo et al (2018)[32]					
DM	78/14,423	4.2	12.9	9.3	NR
sDM plus DBT	155/16,666	4.0	23.3	5.4	
p value		0.32	<0.001	<0.001	
Retrospective Observational Studies					
Ambinder et al (2018)[33]					
DBT plus DM	41/7,813	76	6.9	5.3	29.2
sDM plus DBT	82/14,722	71	8.0	5.6	36.7
p value		0.04	0.33	0.75	0.16
Skaane et al (2014)[30]					
Period 1					
DBT plus DM		28	28.5	8.0	
S2D plus DBT		25	30.3	7.4	
p value			0.61		
Period 2					
DBT plus DM		24	32.1	7.8	
S2D plus DBT		22	34.9	7.7	
p value			0.47		

CI: confidence interval; DBT: digital breast tomosynthesis; DM: digital mammography; PPV: positive predictive value; sDM: synthesized digital mammography; s2D: synthesized 2-dimensional.

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

There is no direct evidence from trials comparing health outcomes in patients screened for breast cancer using DBT and mammography.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Given that the utility of breast cancer screening with mammography has been established, a chain of evidence should demonstrate that screening performance of DBT plus synthesized 2D is equivalent to that of standard mammography alone. Available studies have reported that replacing mammography with DBT plus synthesized 2D might increase cancer detection and reduce recall rates. However, performance characteristics are uncertain due to the limitations described above in the section on the clinical utility of DBT plus acquired mammography, and thus it is not possible to construct a chain of evidence.

Subsection Summary: Screening with 3D DBT Plus Synthesized 2D Mammography
Two prospective and three retrospective studies have assessed 3D DBT plus synthesized 2D mammography, which has lower radiation exposure than 3D DBT plus DM. Two studies found higher detection rates with 3D DBT plus synthesized 2D compared with DM, one found similar detection rates with 3D DBT plus synthesized 2D compared with DM, and two found similar detection rates with 3D DBT plus synthesized 2D compared with 3D DBT plus DM. When comparing the recall rate of 3D DBT plus synthesized 2D with DM alone, one prospective study found a higher recall rate in the former and one prospective study found similar rates, while the retrospective studies had mixed findings. However, the potential for overdiagnosis cannot be ascertained because of the study designs, and interval cancer rates are not yet available. The nonrandomized designs lack long-term follow-up to assess false-negative results.

There is a lack of direct evidence on the clinical utility of DBT from screening trials comparing health outcomes in patients screened for breast cancer with DBT vs mammography. Due to limitations in the studies on diagnostic accuracy, it is not possible to construct a chain of evidence.

3D DBT for Diagnosis
Clinical Context and Test Purpose
The purpose of 3D DBT in patients who have screen-detected abnormalities suspicious for breast cancer is to inform a decision whether to biopsy.

The question addressed in this portion of the evidence review is whether there is sufficient evidence that DBT used to detect breast cancer in patients with abnormal findings on breast imaging or clinical exam improves the net health outcome compared with standard techniques. Specifically, is 3D DBT at least as accurate as standard methods for diagnosing breast cancer and is this degree of increased accuracy likely to improve health outcomes via earlier diagnosis, better patient management decisions, and more appropriate treatment?

The following PICOTS were used to select literature to inform this review.

Patients
The relevant population of interest are individuals with abnormal findings on breast imaging or a clinical examination.

Interventions
The intervention of interest is 3D DBT as an adjunct to 3D mammography for diagnosis.

Comparators
The comparators of interest are standard diagnostic methods. Diagnosis includes both physical examination and imaging. Diagnostic imaging may include diagnostic mammography and ultrasonography. Magnetic resonance imaging for diagnosis of breast cancer is discussed in Blue Shield of California Medical Policy: Magnetic Resonance Imaging for Detection and Diagnosis of Breast Cancer.

Outcomes
The beneficial outcomes of a true-negative test result are avoidance of invasive procedures (e.g., biopsy or mastectomy). The beneficial outcomes of a true-positive test result are reductions in overall mortality and breast cancer-specific mortality.

The harmful outcomes of a false-negative test result are a delay in treatment and a potential increase in mortality. The harmful outcomes of false-positive test results are unnecessary invasive procedures.

Timing
DBT for diagnosis would be performed after a positive breast cancer screening examination.

Setting
The test would be performed in an outpatient imaging setting.

Study Selection Criteria
For the evaluation of clinical validity of DBT, studies that met the following eligibility criteria were selected:

- Prospective studies (preferably in a U.S. setting)
- Comparing DBT plus mammography with diagnostic evaluation alone
- Appropriate reference standard (histopathology)
- Including performance characteristics (e.g., sensitivity, specificity)

Technically Reliable
Assessment of technical reliability focuses on specific tests and operators and requires review of unpublished and often proprietary information. Review of specific tests, operators, and unpublished data are outside the scope of this evidence review and alternative sources exist. This evidence review focuses on the clinical validity and clinical utility.

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Prospective Studies
As per the selection criteria, the characteristics of prospective studies are described in Table 10. The reference standard used for all included studies was histopathology. These prospective studies were conducted in Europe and Asia. Heywang-Kobrunner et al (2017)[34] and Thibault et al (2013)[35] used single-view DBT while Seo et al (2016)[36] used double-view.

Table 10. Characteristics of Prospective Studies of DBT Diagnostic Performance

Study	Study Population	Reference Standard	Threshold for Positive Index Test	Timing of Reference and Index Tests	Blinding of Assessors
Heywang-Kobrunner et al (2017)[34]	Germany: Ages 50-69 y with a screen-detected abnormality; percent with calcifications NR	Histopathology and 2-y follow-up of negatives and registry matching	Reading by experienced radiologists, rating of BIRADS 0, 3, 4, or 5; Single-view	NR	No
Seo et al (2016)[36]	Korea: Signs and symptoms of suspicious findings on screening mammography or ultrasonography;	Histopathology and 2-y follow-up of negatives	Reading by experienced radiologists; rating of BIRADS 4 or 5; Double-view	NR	Yes

Study	Study Population	Reference Standard	Threshold for Positive Index Test	Timing of Reference and Index Tests	Blinding of Assessors
	10% with calcifications				
Thibault et al (2013)[35]	France: Ages ≥40 y with screening recalls with unresolved mammographic or ultrasound workup or with breast symptoms; 31% with calcifications	Histopathology or minimum 2-y follow-up	Reading by experienced radiologist, rating of BIRADS 4 or 5; Single-view	NR	Yes
Teertstra et al (2010)[37]	Netherlands: Abnormal screening mammogram, with clinical symptoms, or referred from other hospitals for a second opinion; percent with calcifications NR	Histopathology with 1.5-2 y follow-up of negatives	Reading by experienced radiologist, rating of BIRADS 0, 3, 4, or 5	NR	Yes

BIRADS: Breast Imaging Reporting and Data System; DBT: digital breast tomosynthesis; NR: not reported.

Results of the studies meeting selection criteria are shown in Table 11. Precision estimates for performance characteristics such as sensitivity and specificity were only provided in Teertstra et al (2010)[37] in which the diagnostic performance of DBT was very similar to DM and Seo et al (2016).[36]

Table 11. Results of Prospective Studies of DBT Diagnostic Performance

Study	Initial N	Final N	Excluded Samples	Prevalence of Condition, %	Clinical Validity (95% Confidence Interval)			
					Sensitivity	Specificity	PPV	NPV
Heywang-Kobrunner et al (2017)[34]	NR	311	Unclear	18				
DM					91	42	25	96
DBT					96	57	32	97
DM plus DBT					96	54	31	99
Seo et al (2016)[36]	219	203	Surgical clip in breast or history of vacuum-assisted breast biopsy	63				
DM					73	61	NR	NR
DBT					78	63	NR	NR
DM plus DBT					80	64	NR	NR
Thibault et al (2013)[35]	156	131	Incomplete mammo-graphic data for review	42				
DM					73	53	53	74
DM plus US					81	48	53	78
DBT					66	64	57	72
DM plus DBT					68	64	58	73
DM+US+DBT					81	52	55	79

Study	Initial N	Final N	Excluded Samples	Prevalence of Condition, %	Clinical Validity (95% Confidence Interval)			
Teertstra et al (2010)[37]	513	513	0	37				
DM					93 (87 to 96)	86 (84 to 88)	48 (41 to 54)	99 (98 to 99)
DBT					93 (87 to 96)	84 (92 to 87)	45 (38 to 52)	99 (98 to 99)

DBT: digital breast tomosynthesis; DM: digital mammography; NPV: negative predictive value; NR: not reported; PPV: positive predictive value; US: ultrasonography.

The studies included in the tables above were prospective, consecutively enrolled participants, and used an appropriate reference standard. Notable gaps identified in each study are shown in Tables 12 and 13. Only one study compared DBT with DM plus ultrasonography and one study provided precision estimates for performance characteristics such as sensitivity and specificity.

Table 12. Relevance Gaps of Prospective Studies of DBT Diagnostic Performance

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
Heywang-Kobrunner et al (2017)[34]		1. BIRADS 0 and 3 included as positive	3. Ultrasonography not included		
Seo et al (2016)[36]			3. Ultrasonography not included	3. PPV and NPV not reported	
Thibault et al (2013)[35]					
Teertstra et al (2010)[37]		3. Intervention was DBT without DM	3. Ultrasonography not included		

The evidence gaps stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
BIRADS: Breast Imaging Reporting and Data System; DBT: digital breast tomosynthesis; DM: digital mammography; NPV: negative predictive value; PPV: positive predictive value.
[a]Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b]Intervention key: 1. Classification thresholds not defined or not standard; 2. Version used unclear; 3. Not intervention of interest.
[c]Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d]Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e]Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true positives, true negatives, false positives, false negatives cannot be determined).

Table 13. Study Design and Conduct Gaps of Prospective Studies of DBT Diagnostic Performance

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Heywang-Kobrunner et al (2017)[34]		1. No blinding	1. Timing of imaging tests and reference standard not described		1. No description of whether there were inadequate images	1. CIs not reported

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Seo et al (2016)[36]			1. Timing of imaging tests and reference standard not described			1. CIs not reported
Thibault et al (2013)[35]			1. Timing of imaging tests and reference standard not described		2. 16% of breasts had incomplete mammographic data	1. CIs not reported
Teertstra et al (2010)[37]			1. Timing of imaging tests and reference standard not described			

The evidence gaps stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
DBT: digital breast tomosynthesis; CI: confidence interval.
[a]Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b]Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c]Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d]Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e]Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f]Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison to other tests not reported.

Systematic Reviews

Lei et al (2014) conducted a meta-analysis of 7 studies (total N=2014 patients; total N=2666 lesions) that compared DBT with DM in patients who had breast lesions graded as category 2 or higher using the Breast Imaging Reporting and Data System (BI-RADS).[38] All studies were rated high quality by reviewers using the Quality Assessment of Diagnostic Accuracy Studies-2 tool. However, only two studies were prospective. As shown in Table 14, compared with histologic diagnosis, the performance of both imaging modalities was approximately similar; PPVs were low (57% for breast tomosynthesis vs 50% for DM), and NPV were high. Statistical heterogeneity among these analyses was considerable (I2 90%). Studies used both 1-view (n=4) and 2-view (n=3) breast tomosynthesis. Pooled sensitivity and specificity for only 1-view breast tomosynthesis studies were 81% and 77%, respectively; for 2-view studies, pooled sensitivity and specificity were 97% and 79% respectively.[39]

Table 14. Side-by-Side Comparison of DBT and DM Diagnostic Performance with Histologic Diagnosis: Pooled Results

	Pooled Estimates (95% CI), %	
Outcomes	**DBT**	**DM**
Sensitivity, %	90 (87 to 92)	89 (86 to 91)
Specificity, %	79 (77 to 81)	72 (70 to 74)
Positive predictive value, %[a]	57 (53 to 61)	50 (46 to 53)
Negative predictive value, %[a]	96 (95 to 97)	95 (94 to 97)
Diagnostic odds ratio[b]	26.04 (8.70 to 77.95)	16.24 (5.61 to 47.04)
LR+	3.50 (2.31 to 5.30)	2.83 (1.77 to 4.52)
LR–	0.15 (0.06 to 0.36)	0.18 (0.09 to 0.38)
Summary AUROC	0.867	0.856

Adapted from Lei et al (2014).[38] AUROC: area under the receiver operating characteristic curve; DBT: digital breast tomosynthesis; DM: digital mammography; LR+: positive likelihood ratio (ratio of the

probability of positivity in cases to the probability of positivity in controls = sensitivity/1 - specificity.); LR-: negative likelihood ratio (ratio of the probability of a negative result in cases to the probability of a negative result in controls = 1 - sensitivity./specificity).
[a]Calculated by BCBSA.
[b]Calculated as the ratio of the odds of positivity in cases to the odds of positivity in controls = LR+./LR-., where LR is the likelihood ratio.

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

There is no direct evidence from trials comparing health outcomes in patients using DBT with another technique (e.g., mammography, ultrasonography) for diagnosing breast cancer.

Chain of Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

A chain of evidence should establish that DBT incrementally improves diagnosis compared with standard management and the additional diagnostic information could be used to change management decisions so that the net health outcome is improved. However, performance characteristics are uncertain due to the limitations described below, and thus it is not possible to construct a chain of evidence.

For women with suspicious lesions (e.g., BI-RADS category 4), a consistently high NPV for DBT would be needed before DBT would likely be used to avoid biopsy. For women with lesions that have a lower BI-RADS category (e.g., BI-RADS 3 probably benign finding.), a high PPV for DBT might result in a change in management from continued surveillance to biopsy. The BI-RADS classification system supports the classification of imaging findings into categories that can be meaningfully linked to recommendations for further clinical management. For example, BI-RADS 3 may be recommended for shorter interval follow-up to assess for stability. If DBT were proposed for diagnostic use in this setting, the chain of evidence would need to clarify assumptions about how DBT results would be used to change management and how those changes would affect health outcomes. The chain cannot be established due to lack of certainty about performance characteristics and intended use population.

The mixed patient populations of the validation studies reflects the lack of clarity about who might benefit from this mode of imaging. The intended use population should be defined based on clinical characteristics such as BI-RADS category, calcifications, breast density, asymmetry in densities or distortions, irregular margins, and prior biopsy or treatment.

Mixed patient populations make it difficult to draw conclusions from the studies on the diagnostic performance of DBT. Also, some concerns have been raised about the classification of microcalcification clusters with DBT alone.

Prospective studies, preferably in the U.S. setting, with an appropriate reference standard and comparison to relevant diagnostic evaluation, are needed to establish performance characteristics.

Section Summary: 3D DBT for Diagnosis
Mixed patient populations make it difficult to draw conclusions from the available studies on the diagnostic performance of 3D DBT. Few prospective studies have addressed whether the addition of 3D DBT improves diagnosis over mammography alone or mammography plus ultrasonography. Also, some concerns have been raised about the classification of microcalcification clusters with 3D DBT alone. There is no direct evidence on the clinical utility of 3D DBT from trials comparing health outcomes in patients diagnosed with breast cancer with 3D DBT vs mammography. Due to limitations in the studies on diagnostic accuracy, it is not possible to construct a chain of evidence.

Summary of Evidence
For individuals who are asymptomatic and at average risk of breast cancer who receive 3-dimensional (3D) DBT as an adjunct to 2-dimensional (2D) mammography for screening, the evidence includes results from randomized controlled trials, prospective observational studies, and retrospective observational studies. The relevant outcomes are overall survival, disease-specific survival, and test validity. There is a lack of direct evidence on the clinical utility of DBT from trials comparing health outcomes in patients screened using DBT and mammography. The available studies have provided limited data on interval cancers and follow-up of negative findings; however, available evidence would suggest that adding breast tomosynthesis to mammography may increase sensitivity) and specificity of screening, potentially reducing the number of women who are recalled unnecessarily. Many studies had methodologic limitations, including inadequate follow-up of women with negative screening results, use of historical controls, and were based on screening practices in Europe that differ from those in the United States. Preliminary results from the RETomo randomized controlled trial would suggest an almost 90% increase in detection rate for mammography plus DBT compared with mammography alone with more instances of ductal carcinoma in situ with mammography plus DBT (+1 per 1000), benign lesions (+1 per 1000), and invasive cancers (+3 per 1000). Although limitations are present among studies within the literature and the long term outcomes are lacking, it may be logical to assume DBT technology is at least non-inferior to DM; therefore DBT may be considered medically necessary when used for screening purposes.

For individuals who are asymptomatic and at average-risk of breast cancer who receive 3D DBT with synthesized 2D mammography for screening, the evidence includes several nonrandomized comparative studies. The relevant outcomes are overall survival, disease-specific survival, and test validity. Two studies found higher detection rates with 3D DBT plus synthesized 2D mammography than with digital mammography, one study found similar detection rates with 3DDBT plus synthesized 2D mammography compared with digital mammography, and two found similar detection rates between 3D DBT plus synthesized 2D mammography and DBT plus digital mammography. When comparing the recall rates of 3D DBT plus synthesized 2D mammography with digital mammography alone, a prospective study found a higher recall rate in the former, a prospective study found similar rates, while retrospective studies had mixed findings. However, the potential for overdiagnosis (i.e., diagnosis of cancer that would not cause symptoms during a patient's lifetime) cannot be ascertained because of the study designs, and interval cancer rates are not yet available. The nonrandomized designs lack long-term follow-up to assess false-negative results. Due to limitations in the studies on diagnostic accuracy, it is not possible to construct a chain of evidence. Although limitations are present among studies within the literature and the long term outcomes are lacking, it may be logical to assume DBT technology is at least non-inferior to DM; therefore DBT may be considered medically necessary when used for screening purposes.

For individuals who have abnormal findings on breast imaging or clinical exam who receive 3D DBT as an adjunct to 2Dmammography for diagnosis, the evidence includes multiple observational studies and a meta-analysis. The relevant outcomes are test validity and treatment-related morbidity. There is a lack of direct evidence on the clinical utility of DBT from diagnostic trials comparing health outcomes in patients diagnosed with breast cancer using DBT

vs mammography. Mixed patient populations make it difficult to draw conclusions from the available studies on the diagnostic performance of DBT. Few prospective studies have addressed whether DBT improves diagnosis when added to mammography or mammography plus ultrasonography. Also, some concerns have been raised about the classification of microcalcification clusters with DBT alone. Due to limitations in the studies on diagnostic accuracy, it is not possible to construct a chain of evidence. Although the literature supporting DBT for diagnostic purposes is lacking in long term outcomes and comes with certain limitations, as with the technology used for screening purposes, it may be logical to assume DBT technology is at least non-inferior to DM; therefore DBT may be considered medically necessary when used for diagnostic purposes.

Supplemental Information
Practice Guidelines and Position Statements

American College of Radiology
The ACR(2014) statement on breast tomosynthesis included the following[40]:

> "…breast tomosynthesis has shown to be an advance over digital mammography, with higher cancer detection rates and fewer patient recalls for additional testing…. Better sensitivity will likely translate into more lives saved. Lower recall rates result in fewer patients who may experience short-term anxiety awaiting test results. This is important evidence that tomosynthesis will have a positive impact on patient care…."

While the ACR has encouraged the additional study of breast tomosynthesis, focusing on long-term clinical outcomes and better definition of subgroups, it concluded that "To be clear: tomosynthesis is no longer investigational. Tomosynthesis has been shown to improve key screening parameters compared to digital mammography."

The ACR's Appropriate Criteria for breast cancer screening, last reviewed in 2017, gave digital breast tomosynthesis (DBT) a rating of "usually appropriate" for use with women at high-risk, intermediate-risk, as well as average-risk for breast cancer.[41]

The ACR's Appropriate Criteria for palpable breast masses, last reviewed in 2016, gave DBT the following ratings[42]:

- "usually appropriate" for
  - women 40 years of age or older, initial evaluation
  - short interval follow-up for women 40 years of age or older, mammography findings probably benign, next examination to perform
  - women younger than 30 years of age, U.S. findings suspicious for malignancy. Next examination to perform
  - women 30 to 39 years of age, initial evaluation
- "usually not appropriate" for
  - short interval follow-up for women 40 years of age or older, mammography findings suspicious for malignancy, next examination to perform
  - short interval follow-up for women 40 years of age or older, mammography findings benign (like lipoma) at site of palpable mass. Next examination to perform
  - women 40 years of age or older, mammography findings negative. Next examination to perform
  - women younger than 30 years of age, initial evaluation
  - women younger than 30 years of age, U.S. findings probably benign. Next examination to perform
  - women younger than 30 years of age, U.S. findings benign (like simple cyst). Next examination to perform
  - women younger than 30 years of age, U.S. findings negative. Next examination to perform

American College of Obstetricians and Gynecologists
In a 2017 Practice Bulletin on breast cancer screening, the American College of Obstetricians and Gynecologists did not discuss tomosynthesis.[43]

A 2015 committee opinion on the management of women with dense breasts identified by mammography stated: "The American College of Obstetricians and Gynecologists does not recommend routine use of alternative or adjunctive tests to screening mammography in women with dense breasts who are asymptomatic and have no additional risk factors."[44] Breast tomosynthesis or thermography were not cited in the document as alternative tests.

American Academy of Family Physicians
The American Academy of Family Physicians (2016) issued a clinical preventive service recommendation on breast cancer.[45] The recommendation stated that there was insufficient evidence for an assessment of the benefits and harms of DBT as a primary screening method for breast cancer. The recommendation also stated that there was insufficient evidence for an assessment of benefits and harms of DBT as adjunctive screening for breast cancer in women identified as having dense breast tissue on an otherwise negative screening mammogram.

National Comprehensive Cancer Network
Current National Comprehensive Cancer Network guidelines (v.2.2018) state: "Multiple studies show tomosynthesis can decrease call back rates and appears to improve cancer detection. Of note, most studies used double the dose of radiation. The radiation dose can be minimized by synthetic 2-D reconstruction."[46]

The National Comprehensive Cancer Network also suggests that tomosynthesis be considered whenever an annual screening mammogram is recommended.

International Agency for Research on Cancer
In 2014, the benefits and harms of different methods of breast cancer screening were assessed by a panel of experts from 16 different countries, convened by the International Agency for Research on Cancer.[47] Table 15 summarizes the panel's conclusions on the available evidence for the use of tomosynthesis with mammography.

Table 15. Recommendations on Use of Tomosynthesis With Mammography

Method	Strength of Evidence[a]
Mammography with tomosynthesis vs mammography alone	
Reduces breast cancer mortality	Inadequate
Increases the detection rate of in situ and invasive cancers	Sufficient
Preferentially increases the detection of invasive cancers	Limited
Reduces the rate of interval cancer	Inadequate
Reduces the proportion of false-positive screening outcomes	Limited

Adapted from Lauby-Secretan et al (2015).[47]
[a] Rating system detailed at http://handbooks.iarc.fr/workingprocedures/index.php.

U.S. Preventive Services Task Force Recommendations
The USPSTF (2016) updated its recommendations on breast cancer screening.[48] The USPSTF recommended biennial screening mammography in women ages 50 to 74 years (grade B recommendation) and that the decision to start screening mammography before age 50 should be individualized (grade C recommendation).

For all women, the USPSTF stated: "...the current evidence is insufficient to assess the benefits and harms of digital breast tomosynthesis (DBT) as a primary screening method for breast cancer" (grade I recommendation). For women with dense breasts, the USPSTF stated "...the current evidence is insufficient to assess the balance of benefits and harms of adjunctive screening for breast cancer using... DBT, or other methods in women identified to have dense breasts on an otherwise negative screening mammogram" (grade I recommendation).

Medicare National Coverage

There is no national coverage determination. In the absence of a national coverage determination, coverage decisions are left to the discretion of local Medicare carriers.

Ongoing and Unpublished Clinical Trials

Some currently unpublished trials that might influence this review are listed in Table 16.

Table 16. Summary of Key Trials

[a] Denotes industry-sponsored or cosponsored trial.

NCT No.	Trial Name	Planned Enrollment	Completion Date
Ongoing			
NCT02698202	Randomized Controlled Trial to Evaluate the Efficacy of Digital Breast Tomosynthesis in Reggio Emilia Breast Cancer Screening Program in the 45-74 Age Group (RETomo)	40,000	Dec 2018
NCT01091545[a]	Malmö Breast Tomosynthesis Screening Trial	15,000	Dec 2019
NCT02590315	Tomosynthesis Versus Digital Mammography in a Population-based Screening Program (ProteusDonna)	92,000	Dec 2019
NCT02835625	The Tomosynthesis Trial in Bergen (TOBE)	29,453	Jan 2022
NCT03377036	Prospective Randomized Comparison of Digital Breast Tomosynthesis Plus Synthesized Images Versus Standard Full-field Digital Mammography in Population-based Screening (TOSYMA)	80,000	Jan 2023
NCT03233191	Tomosynthesis Mammographic Imaging Screening Trial (TMIST)	164,946	Aug 2030

NCT: national clinical trial.

Appendix

Appendix Table 1. Retrospective Studies of DBT for Breast Cancer Screening

Study	No. Cancers/ No. Patients	Recalls/1000 Screens (95% CI)	PPV for Recalls (95% CI), %	Cancers Detected/1000 Screens (95% CI)	PPV for Biopsies (95% CI), %
Multi-reader study; patients apparently served as their own controls					
Good et al (2008)[49];Gur et al (2009, 2011)[8,50]	35/125				
DM alone		Reported separately for cancer and noncancer cases	NR	NR	NR
DBT alone					
DM before DBT					
DM plus DBT					
Patients served as their own controls					
Upadhyay et al (2018)[51]					
Mammo alone	NR	17.4 (15 to 20)			
Mammo plus DBT		11.4 (9.5 to 13.8)			
p value		<0.001			
Patient or provider choice to receive tomosynthesis and did not serve as their own controls					
Rose et al (2018)[52]	126/59,921				
Mammo alone		117	1.6	1.9	13.8
Mammo plus DBT		109	2.3	2.7	16.1
p value		0.003	0.04	0.06	0.55
Cohen et al (2018)[53]					

Study	No. Cancers/ No. Patients	Recalls/1000 Screens (95% CI)	PPV for Recalls (95% CI), %	Cancers Detected/1000 Screens (95% CI)	PPV for Biopsies (95% CI), %
Mammo alone	28/1592		1.8		32.9
Mammo plus DBT	14/354		4.0		37.8
p value			0.01		0.68
Rafferty et al (2017)[54]					
Mammo alone	1,207/278,906	106	Varies by age	4.3	Varies by age
Mammo plus DBT	950/173,414	92		5.5	
p value		0.001	0.01 all ages	0.001[a]	<0.05 ages <60
Geiss et al (2017)[55]					
Mammo	26/14,180	103	1.8	1.8	NR
DBT	37/9,817	107	3.6	3.8	
p value		0.26	0.006	0.005	
Powell et al (2017)[56]					
Mammo alone	54/10,477	160	3	5.2	25.1
Mammo plus DBT	18/2304	140	5.6	7.8	29.5
p value	0.127	0.017	0.032	NR	0.689
Destounis et al (2014)[57]	5/1,048				
Mammo alone		114	16.7	3.8	NR
Mammo plus DBT		42	50.0	5.7	
p value		<0.001		NR	
Durand et al (2015)[58b]	105/17,955				
Mammo alone		123 (117 to 130)	NR	5.7	NR
Mammo plus DBT		78 (73 to 84)		5.9	
p value		<0.001		NS	
Greenberg et al (2014)[59c]	321/59,617				
Mammo alone		155 (138 to 175)	3.0 (2.6 to 3.4)	4.9 (4.2 to 5.7)	21.5 (18.9 to 24.5)
Mammo plus DBT		134 (119 to 152)	4.5 (3.8 to 5.4)	6.2 (5.2 to 7.5)	22.7 (19.5 to 26.6)
p value		<0.001	<0.006	0.041	NS
Haas et al (2013)[60d]	71/13,158				
DM		120 (113 to 128)	NR	5.2	NR
DM plus DBT		84 (77 to 91)		5.7	
p value		<0.01		0.70	
Pre-/post-implementation of tomosynthesis					
McDonald et al (2016)[61]	NR/23,958				
DM		104	4.4	4.6	NR
DBT (year 1)		88	6.2	5.5	
DBT (year 2)		90	6.5	5.8	
DBT (year 3)		92	6.7	6.1	
p value (vs DM years 1-3)		<0.001, <0.001, <0.003	0.06, 0.03, 0.02	NS, NS, NS	NS, NS, NS

Study	No. Cancers/ No. Patients	Recalls/1000 Screens (95% CI)	PPV for Recalls (95% CI), %	Cancers Detected/1000 Screens (95% CI)	PPV for Biopsies (95% CI), %
Conant et al (2016)[62e]	NR/198,881				
DM		104	4.1	5.9	NR
DM plus DBT		87	6.4	4.4	
p value		<0.001	<0.001[f]	0.003	
Sharpe et al (2016)[63]	311/85,852				
DM		75	NR	3.5	NR
DBT		61		5.4	
p value		<0.001		<0.002	
Lourenco et al (2015)[64]	113/25,498				
DM		93 (88 to 99)	5.8	5.4	30.2
DM plus DBT		64 (60 to 68)	7.2	4.6	23.8
p value		<0.001	NS	NS	
Friedewald et al (2014)[65c]	2,157/454,850				
Mammo alone		107 (89 to 124)	4.3	4.2 (3.8 to 4.7)	24.2
Mammo plus DBT		91 (73 to 108)	6.4	5.4 (4.9 to 6.0)	29.2
p value		<0.001	<0.001	<0.001	<0.001
McCarthy et al (2014)[66]	134/26,299				
DM		104 (98 to 109)	4.4 (3.2 to 5.6)	4.6 (3.3 to 5.8)	24.7 (18.6 to 30.9)
DM plus DBT		88 (83 to 92)	6.2 (4.9 to 7.5)	5.5 (4.3 to 6.6)	25.4 (20.6 to 30.2)
p value		<0.001	0.047	NS	NS
Rose et al (2013)[67d]	107/23,357				
DM		87	4.7	4.0	NR
DM plus DBT		55	10.1	5.4	
p value		<0.001		NS	
s2D Mammo					
Freer et al (2017)[29]					
DM	126/21,435		NR	5.9	30.4
DM plus DBT	7/1,019	70		6.9	38.9
s2D plus DBT	56/9,525	58		5.9	37.8
p value		<0.001[g], 0.25[h]		0.66[g], 0.9[h]	0.3[g], 0.98[h]
Aujero et al (2017)[28]					
DM	169/32,076	87	6.0	5.3	22.2
DM plus DBT	194/30,561	58	10.9	6.4	28.5
s2D plus DBT	98/16,173	43	14.3	6.1	40.8
p value		<0.001[i], <0.001[j]	<0.001[i], 0.02[i]	0.08[i], 0.71[i]	0.01[i], 0.001[j]
Zuckerman et al (2016)[27]					
DM plus DBT	NR/15,571	88	6.2	5.45	27.0
s2D plus DBT	NR/5,366	71	7.1	5.0	38.6
p value		<0.001	0.548	0.723	0.053

BIRADS: Breast Imaging Reporting and Data System; CI: confidence interval; DBT: digital breast tomosynthesis; DM: digital mammography; Mammo: mammography; NR: not reported; PPV: positive predictive value; s2D: synthesized 2D mammography.
[a]Except in age group 70+ years in which p=0.082.
[b]Patient samples overlap in the Durand et al (2015)[58] and Haas et al (2013)[60] studies.
[c]Adjusted estimates reported in this table.

[d]Recalls at BIRADS=0.
[e]Data overlap with in McDonald et al (2016).[61]
[f]PPV defined as some cancers diagnosed per number of positive screens.
[g]s2D plus DBT vs DM.
[h]s2D plus DBT vs DM plus DBT.
[i]DM vs DM plus DBT.
[j]DM plus DBT vs s2D plus DBT.

References

1. Tagliafico A, Astengo D, Cavagnetto F, et al. One-to-one comparison between digital spot compression view and digital breast tomosynthesis. Eur Radiol. Mar 2012;22(3):539-544. PMID 21987214
2. National Cancer Institute (NCI). Factsheet: Mammograms. 2016; http://www.cancer.gov/cancertopics/factsheet/detection/mammograms. Accessed July 26, 2017.
3. Rosenberg RD, Yankaskas BC, Abraham LA, et al. Performance benchmarks for screening mammography. Radiology. Oct 2006;241(1):55-66. PMID 16990671
4. Brandt KR, Craig DA, Hoskins TL, et al. Can digital breast tomosynthesis replace conventional diagnostic mammography views for screening recalls without calcifications? A comparison study in a simulated clinical setting. AJR Am J Roentgenol. Feb 2013;200(2):291-298. PMID 23345348
5. Shen Y, Yang Y, Inoue LY, et al. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. J Natl Cancer Inst. Aug 17 2005;97(16):1195-1203. PMID 16106024
6. Smith A. Fundamentals of breast tomosynthesis WP-00007. Bedford, MA: Hologic; 2008.
7. Alakhras M, Bourne R, Rickard M, et al. Digital tomosynthesis: a new future for breast imaging? Clin Radiol. May 2013;68(5):e225-236. PMID 23465326
8. Gur D, Abrams GS, Chough DM, et al. Digital breast tomosynthesis: observer performance study. AJR Am J Roentgenol. Aug 2009;193(2):586-591. PMID 19620460
9. Blue Cross and Blue Shield Association Technology Evaluation Center (TEC). Use of digital breast tomosynthesis with mammography for breast cancer screening or diagnosis. TEC Assessments. 2014;Volume 28:Tab 6.
10. Food and Drug Administration (FDA). Summary of Safety and Effectiveness Data (SSED): Selenia Dimensions 3D System. 2011; https://www.accessdata.fda.gov/cdrh_docs/pdf8/P080003b.pdf. Accessed September 6, 2018.
11. Hologic Inc. Press release: Hologic Receives FDA Approval for a New Low-dose 3D Mammography (Breast Tomosynthesis) Solution for Breast Cancer Screening. 2013 May 21; https://investors.hologic.com/2013-05-21-Hologic-Receives-FDA-Approval-for-a-New-Low-dose-3D-Mammography-Breast-Tomosynthesis-Solution-for-Breast-Cancer-Screening. Accessed September 6, 2018.
12. Houssami N, Bernardi D, Caumo F, et al. Interval breast cancers in the 'screening with tomosynthesis or standard mammography' (STORM) population-based trial. Breast. Apr 2018;38:150-153. PMID 29328943
13. Skaane P, Sebuodegard S, Bandos AI, et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. Breast Cancer Res Treat. Jun 2018;169(3):489-496. PMID 29429017
14. Pattacini P, Nitrosi A, Rossi PG, et al. Digital mammography versus digital mammography plus tomosynthesis for breast cancer screening: The Reggio Emilia Tomosynthesis Randomized Trial. Radiology. Jun 5 2018:172119. PMID 29869961
15. 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. Sep 2017;94:133-139. PMID 28716454
16. Lång K, Andersson I, Rosso A, et al. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmo Breast

Tomosynthesis Screening Trial, a population-based study. Eur Radiol. Jan 2016;26(1):184-190. PMID 25929946
17. 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. Dec 2015;22(12):1477-1482. PMID 26391857
18. Skaane P, Bandos AI, Gullien R, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. Radiology. Apr 2013;267(1):47-56. PMID 23297332
19. Ciatto S, Houssami N, Bernardi D, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol. Jun 2013;14(7):583-589. PMID 23623721
20. Houssami N, Macaskill P, Bernardi D, et al. Breast screening using 2D-mammography or integrating digital breast tomosynthesis (3D-mammography) for single-reading or double-reading - Evidence to guide future screening strategies. Eur J Cancer. Jul 2014;50(10):1799-1807. PMID 24746887
21. Rafferty EA, Park JM, Philpotts LE, et al. Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. Radiology. Jan 2013;266(1):104-113. PMID 23169790
22. 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. Aug 2016;17(8):1105-1113. PMID 27345635
23. Skaane P, Bandos AI, Gullien R, et al. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. Eur Radiol. Aug 2013;23(8):2061-2071. PMID 23553585
24. Skaane P. Response. Radiology. Jun 2013;267(3):969. PMID 23844424
25. Yun SJ, Ryu CW. Benefit of adding digital breast tomosynthesis to digital mammography for breast cancer screening focused on cancer characteristics: a meta-analysis. Aug 2017;164(3):557-569. PMID 28516226
26. Houssami N. Evidence on synthesized two-dimensional mammography versus digital mammography when using tomosynthesis (three-dimensional mammography) for population breast cancer screening. Clin Breast Cancer. Aug 2018;18(4):255-260 e251. PMID 29066138
27. Zuckerman SP, Conant EF, Keller BM, et al. Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. Radiology. Dec 2016;281(3):730-736. PMID 27467468
28. Aujero MP, Gavenonis SC, Benjamin R, et al. Clinical performance of synthesized two-dimensional mammography combined with tomosynthesis in a large screening population. Radiology. Apr 2017;283(1):70-76. PMID 28221096
29. 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. Nov 2017;166(2):501-509. PMID 28780702
30. Skaane P, Bandos AI, Eben EB, et al. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. Radiology. Jun 2014;271(3):655-663. PMID 24484063
31. Romero Martin S, Raya Povedano JL, Cara Garcia M, et al. Prospective study aiming to compare 2D mammography and tomosynthesis + synthesized mammography in terms of cancer detection and recall. From double reading of 2D mammography to single reading of tomosynthesis. Eur Radiol. Jun 2018;28(6):2484-2491. PMID 29294150
32. 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. Apr 2018;287(1):37-46. PMID 29237146

33. Ambinder EB, Harvey SC, Panigrahi B, et al. Synthesized mammography: the new standard of care when screening for breast cancer with digital breast tomosynthesis. Acad Radiol. Aug 2018;25(8):973-976. PMID 29395801
34. Heywang-Kobrunner S, Jaensch A, Hacker A, et al. Value of digital breast tomosynthesis versus additional views for the assessment of screen-detected abnormalities - a first analysis. Breast Care (Basel). May 2017;12(2):92-97. PMID 28559765
35. Thibault F, Dromain C, Breucq C, et al. Digital breast tomosynthesis versus mammography and breast ultrasound: a multireader performance study. Eur Radiol. Sep 2013;23(9):2441-2449. PMID 23673573
36. Seo M, Chang JM, Kim SA, et al. Addition of digital breast tomosynthesis to full-field digital mammography in the diagnostic setting: additional value and cancer detectability. J Breast Cancer. Dec 2016;19(4):438-446. PMID 28053633
37. Teertstra HJ, Loo CE, van den Bosch MA, et al. Breast tomosynthesis in clinical practice: initial results. Eur Radiol. Jan 2010;20(1):16-24. PMID 19657655
38. Lei J, Yang P, Zhang L, et al. Diagnostic accuracy of digital breast tomosynthesis versus digital mammography for benign and malignant lesions in breasts: a meta-analysis. Eur Radiol. Mar 2014;24(3):595-602. PMID 24121712
39. Lei J, Yang P, Zhang L, et al. Reply to Letter to the Editor re: Diagnostic accuracy of digital breast tomosynthesis versus digital mammography for benign and malignant lesions in breasts: a meta-analysis. Eur Radiol. Apr 2014;24(4):928-929. PMID 24442445
40. American College of Radiology (ACR). Position statements: ACR statement on breast tomosynthesis. 2014; https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Breast-Tomosynthesis. Accessed September 6, 2018.
41. American College of Radiology (ACR). ACR Appropriateness Criteria®: breast cancer screening. 2017; https://acsearch.acr.org/docs/70910/Narrative/. Accessed August 2, 2018.
42. American College of Radiology (ACR). ACR Appropriateness Criteria®: Palpable Breast Masses. 2016; https://acsearch.acr.org/docs/69495/Narrative/. Accessed August 2, 2018.
43. American College of Obstetricians and Gynecologists. Practice bulletin no. 179: Breast cancer screening. 2017; https://www.acog.org/-/media/Practice-Bulletins/Committee-on-Practice-Bulletins----Gynecology/Public/pb179.pdf?dmc=1&ts=20180204T1824108792. Accessed August 2, 2018.
44. Committee opinion no. 625: management of women with dense breasts diagnosed by mammography. Obstet Gynecol. Mar 2015;125(3):750-751. PMID 25730253
45. American Academy of Family Physicians (AAFP). Clinical Preventive Service Recommendation: Breast Cancer. 2016; https://www.aafp.org/patient-care/clinical-recommendations/all/breast-cancer.html. Accessed September 6, 2018.
46. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: breast cancer screening and diagnosis. Version 2.2018. https://www.nccn.org/professionals/physician_gls/pdf/breast-screening.pdf. Accessed August 2, 2018.
47. Lauby-Secretan B, Scoccianti C, Loomis D, et al. Breast-cancer screening--viewpoint of the IARC Working Group. N Engl J Med. Jun 11 2015;372(24):2353-2358. PMID 26039523
48. U.S. Preventive Services Task Force (USPSTF). Final Recommendation Statement. Breast Cancer: Screening. 2016; https://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/breast-cancer-screening1. Accessed September 6, 2018.
49. Good WF, Abrams GS, Catullo VJ, et al. Digital breast tomosynthesis: a pilot observer study. AJR Am J Roentgenol. Apr 2008;190(4):865-869. PMID 18356430
50. Gur D, Bandos AI, Rockette HE, et al. Localized detection and classification of abnormalities on FFDM and tomosynthesis examinations rated under an FROC paradigm. AJR Am J Roentgenol. Mar 2011;196(3):737-741. PMID 21343521
51. Upadhyay N, Soneji N, Stewart V, et al. 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. Aug 2018;73(8):744-749. PMID 29724433

52. Rose SL, Shisler JL. Tomosynthesis impact on breast cancer screening in patients younger than 50 years old. AJR Am J Roentgenol. Jun 2018;210(6):1401-1404. PMID 29629810
53. Cohen EO, Tso HH, Phalak KA, et al. Screening mammography findings from one standard projection only in the era of full-field digital mammography and digital breast tomosynthesis. AJR Am J Roentgenol. Aug 2018;211(2):445-451. PMID 29792742
54. 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. Aug 2017;164(3):659-666. PMID 28523569
55. Giess CS, Pourjabbar S, Ip IK, et al. Comparing diagnostic performance of digital breast tomosynthesis and full-field digital mammography in a hybrid screening environment. AJR Am J Roentgenol. Jun 22 2017:1-6. PMID 28639832
56. Powell JL, Hawley JR, Lipari AM, et al. 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. Mar 2017;24(3):302-307. PMID 27919540
57. 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. Feb 25 2014;4:9. PMID 24744966
58. Durand MA, Haas BM, Yao X, et al. Early clinical experience with digital breast tomosynthesis for screening mammography. Radiology. Jan 2015;274(1):85-92. PMID 25188431
59. Greenberg JS, Javitt MC, Katzen J, et al. Clinical performance metrics of 3D digital breast tomosynthesis compared with 2D digital mammography for breast cancer screening in community practice. AJR Am J Roentgenol. Sep 2014;203(3):687-693. PMID 24918774
60. Haas BM, Kalra V, Geisel J, et al. Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. Radiology. Dec 2013;269(3):694-700. PMID 23901124
61. McDonald ES, Oustimov A, Weinstein SP, et al. Effectiveness of digital breast tomosynthesis compared with digital mammography: outcomes analysis from 3 years of breast cancer screening. JAMA Oncol. Jun 1 2016;2(6):737-743. PMID 26893205
62. 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. Feb 2016;156(1):109-116. PMID 26931450
63. Sharpe RE, Jr., 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. Mar 2016;278(3):698-706. PMID 26458206
64. Lourenco AP, Barry-Brooks M, Baird GL, et al. Changes in recall type and patient treatment following implementation of screening digital breast tomosynthesis. Radiology. Feb 2015;274(2):337-342. PMID 25247407
65. Friedewald SM, Rafferty EA, Rose SL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA. Jun 25 2014;311(24):2499-2507. PMID 25058084
66. 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. Nov 2014;106(11). PMID 25313245
67. Rose SL, Tidwell AL, Bujnoch LJ, et al. Implementation of breast tomosynthesis in a routine screening practice: an observational study. AJR Am J Roentgenol. Jun 2013;200(6):1401-1408. PMID 23701081
68. Blue Cross Blue Shield Association. Medical Policy Reference Manual, No. 6.01.53 (December 2018).

Documentation for Clinical Review

Please provide the following documentation (if/when requested):

- History and physical and/or consultation notes
- Previous breast imaging studies (i.e., mammogram, ultrasound, digital breast tomosynthesis)

Post Service

- Breast imaging report

Coding

This Policy relates only to the services or supplies described herein. Benefits may vary according to product design; therefore, contract language should be reviewed before applying the terms of the Policy. Inclusion or exclusion of codes does not constitute or imply member coverage or provider reimbursement.

MN/NMN

The following services may be considered medically necessary when policy criteria are met. Services may be considered not medically necessary when policy criteria are not met.

Type	Code	Description
CPT®	77061	Diagnostic digital breast tomosynthesis; unilateral
	77062	Diagnostic digital breast tomosynthesis; bilateral
	77063	Screening digital breast tomosynthesis, bilateral (List separately in addition to code for primary procedure)
HCPCS	G0279	Diagnostic digital breast tomosynthesis, unilateral or bilateral (list separately in addition to 77065 or 77066)
ICD-10 Procedure	BH00ZZZ	Plain Radiography of Right Breast
	BH01ZZZ	Plain Radiography of Left Breast
	BH02ZZZ	Plain Radiography of Bilateral Breasts

Policy History

This section provides a chronological history of the activities, updates and changes that have occurred with this Medical Policy.

Effective Date	Action	Reason
06/28/2013	BCBSA Medical Policy adoption	Medical Policy Committee
09/30/2014	Policy revision without position change	Medical Policy Committee
01/01/2015	Coding update	Administrative Review
08/31/2015	Coding update	Administrative Review
03/01/2016	Policy revision without position change	Medical Policy Committee
08/01/2016	Coding clarification to Policy Guidelines	Administrative Review
12/01/2016	Policy revision with position change	Medical Policy Committee
11/01/2017	Policy revision without position change	Medical Policy Committee
02/01/2018	Coding update	Administrative Review
11/01/2018	Policy revision without position change	Medical Policy Committee
02/01/2019	Policy revision without position change	Medical Policy Committee

Definitions of Decision Determinations

Medically Necessary: A treatment, procedure, or drug is medically necessary only when it has been established as safe and effective for the particular symptoms or diagnosis, is not investigational or experimental, is not being provided primarily for the convenience of the patient or the provider, and is provided at the most appropriate level to treat the condition.

Investigational/Experimental: A treatment, procedure, or drug is investigational when it has not been recognized as safe and effective for use in treating the particular condition in accordance with generally accepted professional medical standards. This includes services where approval by the federal or state governmental is required prior to use, but has not yet been granted.

Split Evaluation: Blue Shield of California/Blue Shield of California Life & Health Insurance Company (Blue Shield) policy review can result in a split evaluation, where a treatment, procedure, or drug will be considered to be investigational for certain indications or conditions, but will be deemed safe and effective for other indications or conditions, and therefore potentially medically necessary in those instances.

Prior Authorization Requirements (as applicable to your plan)

Within five days before the actual date of service, the provider must confirm with Blue Shield that the member's health plan coverage is still in effect. Blue Shield reserves the right to revoke an authorization prior to services being rendered based on cancellation of the member's eligibility. Final determination of benefits will be made after review of the claim for limitations or exclusions.

Questions regarding the applicability of this policy should be directed to the Prior Authorization Department. Please call (800) 541-6652 or visit the provider portal at www.blueshieldca.com/provider.

Disclaimer: This medical policy is a guide in evaluating the medical necessity of a particular service or treatment. Blue Shield of California may consider published peer-reviewed scientific literature, national guidelines, and local standards of practice in developing its medical policy. Federal and state law, as well as contract language, including definitions and specific contract provisions/exclusions, take precedence over medical policy and must be considered first in determining covered services. Member contracts may differ in their benefits. Blue Shield reserves the right to review and update policies as appropriate.