The New England Comparative Effectiveness Public Advisory Council

Public Meeting – December 13, 2013

The Comparative Clinical Effectiveness and Value of Supplemental Screening Tests Following Negative Mammography in Women with Dense Breast Tissue

Final Report – January 2014

Authored by:

Jeffrey A. Tice, MD
Associate Professor of Medicine
Division of General Internal Medicine
Department of Medicine
University of California San Francisco

Daniel A. Ollendorf, MPH, ARM
Chief Review Officer, Institute for Clinical and Economic Review

Janie M. Lee, MD, MS
Associate Professor of Radiology, University of Washington

Steven D. Pearson, MD, MSc
President, Institute for Clinical and Economic Review

Table of Contents

Key Abbreviations Used in This Report

ABUS:	Automated whole-breast ultrasound
ACRIN:	American College of Radiology Imaging Network
ACS:	American Cancer Society
AJCC:	American Joint Committee on Cancer
ASCO:	American Society of Clinical Oncology
BC:	Breast cancer
BCSC:	Breast Cancer Surveillance Consortium
BI-RADS:	Breast Imaging Reporting and Data System
CDR:	Cancer detection rate
CI:	Confidence interval
DBT:	Digital breast tomosynthesis
cm:	Centimeters
CEPAC:	(New England) Comparative Effectiveness Public Advisory Council
CTAF:	California Technology Assessment Forum
DARE:	Database of Abstracts of Reviews of Effects
DCIS:	Ductal carcinoma in situ
DFS:	Disease-free survival
DMIST:	Digital Mammography Imaging Screening Trial
FDA:	US Food and Drug Administration
HHUS:	Handheld ultrasound
HR:	Hazard ratio
Hz:	Hertz
IHC:	Immunohistochemistry
MRI:	Magnetic resonance imaging
MQSA:	Mammography Quality Standards Act
NCCN:	National Comprehensive Cancer Network
NPV:	Negative predictive value
NR:	Not reported
NS:	Not significant
NSABP:	National Surgical Adjuvant Breast and Bowel Project Protocol
OR:	Odds ratio
OS:	Overall survival
PPV:	Positive predictive value
RCT:	Randomized controlled trial
RR:	Relative risk
TNM:	Tumor, Node, Metastasis staging system

Introduction

To make informed healthcare decisions, patients, clinicians, and policymakers need to consider many different kinds of information. Rigorous evidence on the comparative clinical risks and benefits of alternative care options is always important; but along with this information, decision-makers must integrate other considerations. Patients and clinicians must weigh patients' values and individual clinical needs. Payers and other policymakers must integrate information about current patterns of utilization, and the impact of any new policy on access, equity, and the overall functioning of systems of care. All decision-makers, at one level or another, must also consider the costs of care, and make judgments about how to gain the best value for every healthcare dollar.

The goal of this initiative is to provide a forum in which all these different strands of evidence, information, and public and private values can be discussed together, in a public and transparent process. Funded by a consortium of state Medicaid agencies, private payers, and integrated provider groups, and backed by a diverse set of New England state policymakers, the mission of the New England Comparative Effectiveness Public Advisory Council (CEPAC) is to provide objective, independent guidance on how information on comparative effectiveness can best be used across New England to improve the quality and value of health care services. CEPAC is an independent body composed of clinicians and patient or public representatives from each New England state with skills in the interpretation and application of medical evidence in health care delivery. Representatives of state public health programs and of regional private payers are included as ex-officio members of CEPAC. The latest information on CEPAC, including guidelines for submitting public comments, is available online: cepac.icer-review.org.

The Institute for Clinical and Economic Review (ICER) is managing CEPAC and is responsible for developing evidence reviews for CEPAC consideration. ICER is an independent research organization whose mission is to lead innovation in comparative effectiveness research through methods that integrate evaluations of clinical benefit and economic value. By working collaboratively with patients, clinicians, manufacturers, insurers and other stakeholders, ICER develops tools to support patient decisions and medical policy that share the goals of empowering patients and improving the value of healthcare services. More information about ICER is available at www.icer-review.org.

The current assessment builds on a recent effort undertaken by ICER's other flagship initiative, the California Technology Assessment Forum (CTAF). CTAF's review was conducted in response to growing stakeholder interest in supplemental screening, driven in large part by recently-enacted legislation in California mandating that women with dense breast tissue on mammography be informed of this condition and its attendant risks.[1] Similar legislation has been enacted in

Connecticut (which also includes a mandate for insurance coverage of supplemental screening with ultrasound)[2] and is currently under consideration by several other New England states (see Section 4). A deliberation on this topic focused on New England was therefore deemed timely. The report that follows includes the final CTAF review as well as supplemental information for New England, including regional and national payer coverage policies, new evidence published since the CTAF review, and a population-based simulation model for New England.

This assessment will attempt to answer the key issues that patients, providers, and payers face. These include the following questions: What evidence exists to support decisions regarding the risks and benefits of supplemental screening? Are there ways to estimate the overall risk of breast cancer for women with dense breast tissue and a negative mammogram that would suggest which women are more likely, and which less, to benefit from supplemental screening? And if supplemental screening is considered, what is the potential budgetary impact of different screening options? The purpose of this assessment is to help patients, providers, and payers address these important questions and to support dialogue needed for successful action to improve the quality and value of health care for these women.

1. Background

1.1 Breast Cancer

Breast cancer is the most common form of cancer in women.[3] An American woman is estimated to have a one in eight chance of developing invasive breast cancer at some time during her life. In 2013, there will be an estimated 234,580 new cases of breast cancer in the United States and an estimated 39,620 deaths from this cancer.[3] This represents approximately 29% of all new cancer cases and 14% of all cancer deaths in women.[3] Moreover, breast cancer is the single leading cause of death for non-smoking women between the ages of 35 and 54 years, accounting for about 10% of all deaths.[4]

Mortality from breast cancer has declined by about 2.2% per year since 1990, a 28% overall decline.[5] The median values from a series of models estimated that a little more than half of the decline was due to improvements in therapy for breast cancer and that a little less than half (46%) was due to early diagnosis from mammography.[6] This remains the dominant view, but a recent analysis of 30 years of data from the United States Surveillance, Epidemiology, and End Results (SEER) data called those conclusions into question.[7] Bleyer and Welch estimated that 31% of breast cancer diagnosed with mammography represents "overdiagnosis" (i.e. , identification of cancers unlikely to cause significant morbidity or mortality) and concluded that screening mammography has had, at best, only a small effect on breast cancer mortality.[7]

1.2 Screening for Breast Cancer

The primary method used to screen for breast cancer is mammography. Nine large clinical trials established the efficacy of screening mammography by randomizing over 600,000 women and following them for ten to twenty years.[8-26] The results have been summarized in many systematic reviews and meta-analyses.[27-41] There is general consensus that, for women between the ages of 50 and 69 years, screening mammography reduces breast cancer mortality by approximately 20% to 25% after 15 years of follow-up.[37] For average-risk women between the ages of 40 to 49 years, there remains significant controversy about whether the benefits of routine mammography outweigh the harms, but most guidelines recommend either routine mammography or a discussion of the benefits and risks of mammography.[42-45]

Digital Mammography

Mammography was traditionally performed with film. It was one of the last radiographic procedures to transition from film to digital imaging because mammography requires extremely high resolution to be effective. Digital image acquisition improves the signal to noise ratio of x-ray detection over a wider contrast range than film.[46-48] Digital enhancement of the images at computer workstations may also improve the accuracy of mammographic interpretation.[47] In particular, increased contrast resolution improves the detection of low contrast lesions in radiographically dense breasts. Digital mammography has become the standard across the United States. As of July 1, 2013, 91.4% (11,705 / 12,800) of all US mammography machines accredited by the Food and Drug Administration (FDA) are full-field digital.[49]

1.3 Supplemental Screening Modalities for Breast Cancer Screening

There are many imaging approaches to screen for breast cancer in addition to mammography. Magnetic resonance imaging (MRI) has been most widely used. The American Cancer Society first recommended the use of MRI to screen women at highest risk for breast cancer in 2007, based primarily on genetic susceptibility.[50] Hand-held ultrasound has been used as a diagnostic tool to evaluate women with breast masses and has been promoted by some as a screening tool.[51] The FDA recently has approved automated whole breast ultrasound, which scans and records ultrasound images of the entire breast, for breast cancer screening.[52,53] Finally, digital breast tomosynthesis (DBT), a 3-dimensional extension of digital mammography, has been viewed as holding significant promise in breast cancer screening.[54-56] Other imaging modalities, such as contrast-enhanced mammography, thermography, diffuse optical tomography, sestamibi, positron emission mammography, dedicated breast computed tomography, electrical impedance scanning, MRI spectroscopy, and breast-specific gamma imaging are still in early investigational phases[57,58] and will not be considered further in this assessment.

All four of the advanced imaging technologies considered in this assessment generate multiple two-dimensional images representing slices of the breast. This allows the radiologist to visualize the breast in three-dimensions. This is particularly relevant in mammographically dense breasts because breast cancers may be obscured by superimposed dense tissue.

Magnetic Resonance Imaging (MRI) of the Breast

Magnetic resonance imaging uses strong magnetic fields to image the breast, rather than ionizing radiation. The system uses computational algorithms to generate detailed cross-sectional views of the breast. Mammography requires repositioning of the breast and mammography system for each

desired view. In contrast, the MRI examination is typically performed with the patient in the prone position lying on a platform placed in the MR chamber that allows the breast to extend dependently from the patient and does not require repositioning. A contrast agent, gadolinium, is injected through an intravenous catheter (IV) to improve the images of the breast.

In studies of high-risk women, MRI approximately doubles the number of breast cancers that are detected compared to film mammography or breast ultrasound.[59-66] However, several factors limit the widespread use of MRI for screening. These include an increase in false positive test results, the need for placement of an intravenous catheter to infuse contrast, the length of time required for the examination, the cost of the examination, limited availability of breast MRI facilities (with special breast-specific magnetic coils and biopsy capability), and contraindications to the use of MRI due to pacemakers and other metallic implants. In addition, mammography has been found to be more sensitive than MRI for the detection of ductal carcinoma in situ (DCIS), a noninvasive cell abnormality in the milk ducts and some invasive breast cancers, so the two are typically used together.[43,50,67]

Hand-held Ultrasonography (HHUS) of the Breast

HHUS is widely used at breast imaging centers to evaluate breast masses and to guide both cyst aspiration and percutaneous breast biopsy procedures. It is particularly useful to differentiate fluid filled cysts from solid masses (cysts are rarely cancerous). Over time, HHUS has evolved to use higher frequency sound waves to generate images of the breast with improved resolution. In addition, earlier generations of HHUS were not able to penetrate deeply into breast tissue and had a limited field of view. Advantages of ultrasound include the ability to evaluate tissue that is dense on mammography without additional ionizing radiation, which can potentially increase the risk for future cancers. It is also perceived to be more comfortable than mammography because it does not require compression of the breasts.

Ultrasound also has limitations. The primary concern with HHUS is the high number of false positive findings, which often lead to unnecessary biopsies. There are also concerns about the operator dependency and reproducibility of the examinations. Like MRI, HHUS takes time. The average length of time for breast HHUS imaging in a recent study was 19 minutes.[68] In that study and many others, a breast radiologist performed the study. At a minimum, the breast radiologist needs to be available to review static images saved by the performing technologist in real time so that additional images can be acquired if necessary.

Automated Whole Breast Ultrasonography (ABUS)

ABUS uses computer driven ultrasound transducers to scan the entire breast under the guidance of a technician. A technician compresses the woman's breasts to her chest wall and applies ultrasound gel. A breast-shaped transducer is placed on the compressed breast and automatically scans the entire breast. The entire procedure, including patient preparation, takes about 15 minutes to complete.[52] ABUS reduces the need for radiologists to perform the scan and decreases the length of time of the exam, thus addressing two of the shortcomings of HHUS. It also produces a scan that should have less operator dependence. The radiologist can review the scan independently using software that displays the images individually or sequentially in a movie mode. The primary drawbacks to ABUS are the inability to image very large breasts, the storage requirements for the data acquired during the scan, and the time required to read the scans.[52]

Digital Breast Tomosynthesis (DBT)

Digital breast tomosynthesis (DBT) uses a conventional x-ray source that sweeps along an arc around the breast to acquire multiple two-dimensional (2-D) digital images.[54,56,69] Breast compression is performed using the same device and technique as conventional mammography. The procedure to obtain each digital view is complete in less than 20 seconds. One of the advantages of DBT is that the images can be acquired immediately following the digital mammogram without needing additional compression. Like MRI, computational algorithms synthesize the resulting 2-D digital images to create tomograms (i.e., slices) allowing for a 3-D reconstruction of the breast. The tomograms can be displayed individually (similar to enhanced conventional mammograms) or in a dynamic movie mode.

There are several drawbacks to DBT. The dose of ionizing radiation for DBT is about the same as that used for a conventional mammogram. Currently, a standard digital image is also acquired, so the total dose is approximately twice that of digital mammography alone.[54,55] The technology and algorithms used for DBT are still in evolution.[54,56,69] One of the crucial areas is the development of techniques to biopsy lesions that are only seen on DBT.[70] DBT can also be used to generate a virtual 2-D digital mammogram, which could eliminate the need for performing digital mammography and thus eliminate the excess ionizing radiation. This technology is still in development and was not used in studies considered in this assessment. Finally, the reading time for DBT is about twice that required for digital mammography.[54,55]

1.4 Definitions and Statistics used in the Evaluation of Screening Tests for Breast Cancer

In the United States, the Breast Imaging Reporting and Data System (BI-RADS) of the American College of Radiology[71] sets standards for reporting of the results of breast imaging including mammography, ultrasonography, and MRI. The primary purpose of BI-RADS is to enable consistent reporting and communication regarding findings identified on breast imaging and their management recommendations. In addition widespread BI-RADS use supports quality improvement efforts in breast imaging. There are six standard BI-RADS assessment categories used for women without a known malignancy:

0 Incomplete examination, additional imaging or comparison to priors is needed;
1 Negative;
2 Benign finding(s);
3 Probably benign – short interval follow-up suggested;
4 Suspicious abnormality – biopsy should be considered; and
5 Highly suggestive of malignancy – appropriate action should be taken.

When evaluating screening tests, these results are classified into two categories: a positive test result is any of BI-RADS assessment categories 0, 4 or 5 and a negative test result is any of BI-RADS assessment categories 1, 2, or 3. A true positive is a positive imaging assessment that is followed by a diagnosis of invasive or in situ breast cancer within 12 months. A false positive is a positive imaging result that is not followed by a cancer diagnosis within 12 months. The cancer detection rate is the number of cancers detected by a positive test divided by the number of screening tests performed – for consistency and ease of comparison, we will report it as the number of breast cancers detected per 1000 screening examinations.

The most common statistics reported by scientists evaluating the diagnostic performance of a test are the sensitivity and specificity. The sensitivity is calculated among women with disease: it is defined as the number of positive tests in women with breast cancer divided by the total number of women with breast cancer and is usually reported as a percentage. In studies of breast imaging, the standard has been to follow women for one year after the screening examination and to count any cancers found during that period as interval cancers. Interval cancers are also known as false negatives because the test was negative, but cancer was likely present. True negatives are the negative test results that remain negative during follow-up. The interval cancers are added to the screen-detected cancers to give the total number of women with breast cancer for the calculation of these statistics. An important methodological point when assessing studies of diagnostic tests for breast cancer is that if the studies do not follow women with negative test results over time,

there will be no way to determine how many of the negative tests missed cancers. When there is no follow-up, there will be no false negative results and the sensitivity will always be 100%.

The specificity of these tests is calculated among women without cancer: it is defined as the number of negative tests in women without breast cancer divided by the total number of women without breast cancer over the 12 month follow-up period and is usually reported as a percentage.

Sensitivity and specificity, while helpful for comparing diagnostic tests, are not that helpful in clinical practice. What clinicians and patients want to know is how likely it is that the patient has cancer if she tests positive and how likely it is that she doesn't have cancer if she tests negative. These concepts are known as the positive predictive value (PPV) and the negative predictive value (NPV). Like sensitivity and specificity, these are usually reported as percentages. The positive predictive value is the number of true positives divided by the total number of positive tests, or the percent chance that a woman with a positive test actually has cancer. The negative predictive value is the number of true negatives divided by the total number of negative tests, or the percent chance that a woman with a negative test does not have cancer. In breast cancer screening, things are more complicated because not every woman with a positive test undergoes a biopsy. The BI-RADS audit of mammography outcomes defines three different positive predictive values. The PPV1 is the traditional definition of the number of true positives divided by the total number with a positive result on imaging, and represents the proportion of cancers identified of women recalled from screening for further diagnostic evaluation. The PPV2 is the number of true positives among those recommended for biopsy (BI-RADS 4 or 5 assessment) divided by the total number recommended for biopsy. Finally, the PPV3 is the number of true positives among all those who actually undergo biopsy divided by the total number of biopsies performed. Mammography audits that are required by law to attempt to track all positive tests allow for the calculation of the PPV but not sensitivity, because many sites do not track women over time for interval cancers.[72,73] An important methodological point here is that the predictive values are dependent on the prevalence of cancer. When a diagnostic test is evaluated in two populations, one with a high prevalence of cancer and one with a low prevalence of cancer, the PPV will be higher and the NPV will be lower in the population with the higher prevalence of cancer even though the sensitivity and specificity do not change.

Because of this complexity, two other statistics are also useful: the recall rate is the number of women recalled for additional imaging and/or biopsies divided by the total number of women screened and the biopsy rate defined as the total number of women biopsied divided by the total number of women screened. We will report these statistics per 1000 women screened to allow for comparison across studies and to allow for comparison with the cancer detection rate. Investigators have reported benchmarks for the PPV of film mammography in the United States.[74] They are based on more than two million screening mammograms in over one million women performed between 1996 and 2002. The data come from six registries in the Breast Cancer

Surveillance Consortium (BCSC), a prospective study of breast imaging across the United States. The demographics of participants in this study closely match those of the US population in terms of rural/urban mix, race, Hispanic ethnicity, education, and economic status.[74] The study sample included women ages 40-49 years (29%), 50-79 years (62%), as well as women outside this age range (9%). Approximately 6.3% of the women reported a personal history of breast cancer and 15.2% reported a family history of breast cancer.[74] The results described in this analysis did not include follow-up data so the sensitivity, specificity and negative predictive value could not be calculated. The benchmark PPV statistics come from radiologists performing at least 1000 mammograms over the study period.

Table 1: Performance benchmarks for screening mammography.

Statistic	*BCSC Value*
PPV1, %	4.8
PPV2, %	25.0
PPV3, %	32.6
Recall rate, per 1000	94
Biopsy rate, per 1000	10
Cancer detection rate, per 1000	4.7
DCIS, %	21.6
Cancers ≤ 10 mm, %	37.2
Node negative, %	79.8
Stage 0 or 1, %	75.6

PPV1 = PPV based on a positive result on initial imaging; PPV2 = PPV based on a recommendation for biopsy; PPV3 = PPV based on biopsies actually performed.

Thus, across the United States, for every 1000 mammograms performed approximately 100 women will be recalled and 10 will have a biopsy to detect about 5 cancers. One of those cancers will be DCIS (~20%), four will be lymph node negative (~80%), and 3 or 4 (~75%) will be stage 0 or 1.[74] These statistics will vary when looking at different subgroups of women or different screening technologies. For instance, younger women have more false positive mammography assessments and a lower risk for cancer, so their recall rate will be higher and the number of cancers detected will be lower. Digital mammography, which has greater sensitivity and similar specificity compared to film mammography, will have a similar recall rate, but a higher cancer detection rate.[75]

Benefits of Screening

The primary benefit of screening is a reduction in death from breast cancer. As described above, there have been nine large randomized trials evaluating the efficacy of screening mammography.[8-26] The studies found that screening mammography reduces breast cancer mortality by approximately

20% to 25% after 15 years of follow-up.[37] In absolute terms, for every 1000 women screened with mammography for 15 years, there will be 1.8 fewer deaths from breast cancer.[76] In addition to the mortality reduction, there may be other benefits, such as less need for aggressive therapies in early stage disease and decreased anxiety about breast cancer following a negative mammogram.

Harms of Screening: False Positive Results

The most common harm associated with mammography is a false positive test result. Approximately 10% of women have a false positive result at each round of mammography screening and about 50% of women will have at least one false positive result after 10 mammograms.[74,77-81] Most false positive results lead to additional imaging and not a breast biopsy. Between 7% and 19% of women have a false positive biopsy after 10 mammograms.[78,81] False positive results are associated with short-term increases in anxiety, psychological distress, and rarely, suicide.[82-88] A systematic review of 23 studies on the long-term effects of false positive mammograms found small, but significant negative impacts on health behaviors and psychological well-being. [89] False positives also usually require that a woman schedule a second appointment for additional imaging resulting in time lost with family or at work and the additional evaluation increases health care costs.

Harms of Screening: Overdiagnosis

A second important harm of screening is overdiagnosis: the diagnosis of breast cancers with mammography that, if they had been left undetected, would not have caused symptoms before the woman died of other causes.[90,91] Such patients would endure the toxicity associated with treatment of breast cancer (surgery, radiation, hormonal therapy, and chemotherapy), without receiving any benefit of reduced symptoms or longer life from treating the cancer. It is currently impossible to know whether any particular patient whose cancer is detected by mammography is or is not at risk of the cancer being "overdiagnosed," and the true magnitude of overdiagnosis for breast cancer is unclear and controversial. The most common estimates range from 10% to 30% of cancer diagnoses, although estimates range from as low as 0% to as high as 54%.[7,91-100] This is an area of active research and debate.

Harms of Screening: Radiation Exposure

Ionizing radiation, like that used in mammography, can damage DNA leading to mutations that increase the risk for the development of cancer. Evidence from those exposed to radiation from the atomic bomb explosions in Japan and from those exposed to radiation therapy as part of treatment for Hodgkin's disease demonstrates that radiation exposure increases the risk for breast cancer.[101-106] The risk is greatest for younger women and is thought to be minimal for post-menopausal

women. The radiation dose from mammography is relatively small. The dose from 20 mammograms is equivalent to about 3 years of environmental exposure to radiation; the dose from one CT scan is equivalent to about 800 mammograms.[101-106] There is no direct evidence demonstrating an increase in breast cancer due to mammography. One recent modeling study by Yaffe and colleagues estimated that among 100,000 women screened with mammography every year from ages 40 to 55 years and then every two years until age 75 (20 mammograms), the radiation would cause 86 new breast cancer diagnoses and 11 deaths from breast cancer.[106] Thus for every 1000 women screened 20 times between the ages of 40 and 75 years, the radiation from mammography will cause 0.9 additional breast cancers and 0.1 additional deaths from breast cancer.

The average dose of radiation from mammography has declined with the transition to digital mammography. In the DMIST trial, the average radiation dose was 4.7 mGy with film mammography and 3.7 mGy with digital mammography.[107] The Yaffe model[106] assumed that the dose per mammogram was 3.7 mGy based on the DMIST findings.[107] Other models using different inputs and assumptions have estimated higher rates of radiation-induced breast cancer and death from mammography.[108]

1.5 Mammographic Breast Density

As described previously, mammographic density refers to areas within the breast that absorb significant amounts of x-ray energy and show up as relatively white areas on the mammogram. These correspond to regions in the breast that are rich in epithelial and stromal tissue while the non-dense (darker gray areas) correspond to regions that are predominantly fat.

Breast Density and Masking

In the United States, the Breast Imaging Reporting and Data System (BI-RADS) of the American College of Radiology [71] classifies density in the following four breast composition categories as listed below:

1 Almost entirely fatty;
2 Scattered fibroglandular densities;
3 Heterogeneously dense; and
4 Extremely dense tissue.

The current edition of BI-RADS, published in 2003, defines these four categories more quantitatively as less than 25% dense tissue (category 1), 25 to 49% dense tissue (category 2), 50 to 74% dense

tissue (category 3), and greater than or equal to 75% dense tissue (category 4). The majority of mammograms in the United States include BI-RADS density as part of the official report.[109]

It has been known for a long time that the sensitivity of film mammography is lower in women with dense breasts than in women with fatty breasts.[110] There clearly is a masking effect due to mammographic density. In the BCSC registry, the sensitivity of film mammography decreased markedly with increasing density (see Table 2 below).[111] This study evaluated the results from 463,372 screening film mammograms performed between 1996 and 1998. Among women in the lowest density categories, the sensitivity of mammography was 88% and 82% for density categories 1 and 2 respectively, but this decreased to 69% for women with heterogeneously dense breasts and to 62% for women with extremely dense breasts.[111]

Table 2: Sensitivity of film and digital mammography by breast density.

		BI-RADS density category			
Study	***Type***	***Almost entirely fatty***	***Scattered fibroglandular densities***	***Heterogeneously dense***	***Extremely dense***
BCSC Carney 2003[111]	Film	88.2	82.1	68.9	62.2
DMIST Pisano 2005[75]	Film Digital			55* 70*	
BCSC Kerlikowske 2011[112]	Film Digital	85.7 78.3	85.1 86.6	79.3 82.1	68.1 83.6

*The DMIST study reported results for the combined high-density categories only

Breast Density and Digital Mammography

As described above, the increased contrast resolution of digital mammography improves the detection of low contrast lesions in radiographically dense breasts. Thus digital mammography should improve the sensitivity of mammography in women with dense breast tissue compared to film.

The Digital Mammography Imaging Screening Trial (DMIST) study is the largest trial directly comparing digital mammography to plain film mammography (n=42,760).[75] All women were screened with both film and digital mammography on the same visit. The mammograms were read independently by radiologists blinded to the results of the other mammogram. In DMIST, digital mammography had the same recall and biopsy rates as film mammography. Digital mammography was more sensitive than film, particularly for younger women with denser breasts (59.1% versus 27.3%, p=0. 0013).[113] Among women of all ages with either heterogeneously dense or extremely

dense breasts, digital mammography was also more sensitive than film mammography (70% versus 55%, p= 0.02, Table 2). Similarly, in women with dense breast tissue there was a trend towards greater specificity with digital mammography (91% versus 90%, p=0.09) and the overall accuracy of digital mammography, as measured by the area under the receiver operator curve, was greater than that of film mammography (0.78 versus 0.68, p=0.003).[113]

The BCSC has recently updated their earlier description of the sensitivity of mammography based on a comparison of 231,034 digital mammograms and 638,252 film mammograms performed between January 1, 2000 and December 31, 2006.[112] Similar to the prior study, the sensitivity of film mammography decreased from 86% to 68% across the four breast density categories (see Table 2 on the previous page). However, for digital mammography, the sensitivity of digital mammography remained greater than 80% for the highest density categories and did *not* appear to decrease with increasing density (see Table 2 on the previous page). As in the DMIST trial, digital mammography was significantly more sensitive than film mammography in women with dense breasts.

Table 3 below shows the cancer detection rate and specificity in addition to the sensitivity of digital mammography by BI-RADS density category in the BCSC study. Despite concerns about the test performance of mammography in dense breasts, more breast cancers are found per 1000 digital screening mammograms in the denser breast categories than in the less dense categories. This highlights the general principle that the yield of screening tests is greater as the underlying risk of the population screened goes up. Women with denser breasts are at higher risk, so the cancer detection rate is higher. These data also suggest that the masking effect of breast density is minimized when digital mammography is used.

Table 3: Cancer detection, sensitivity, and specificity of digital mammography by breast density.

		BI-RADS density category			
Study	***Type***	***Almost entirely fatty***	***Scattered fibroglandular densities***	***Heterogeneously dense***	***Extremely dense***
BCSC	Rate*	1.8	3.3	4.8	5.1
Kerlikowske 2011[112]	Sens	78	87	82	84
	Spec	95	91	87	89

*Rate = breast cancer detection rate per 1000 women screened

Table 4 on the following page summarizes important outcomes with film and digital mammograms in the three largest studies that report data on both digital and film mammography.[75,112,114] These are useful benchmarks to use when evaluating the potential yield of additional imaging compared to no additional imaging. The biopsy rate and cancer detection rate did not differ between patients screened with digital or film mammography in any of these studies, although the recall rate for

digital was higher in the BCSC (100 vs. 93 per 1000, p<.001). When cancer detection was stratified by breast density in the BCSC, no statistical differences were found between digital and film mammography. However, there was a nominal trend toward higher cancer detection in women with extremely dense breasts (5.1 vs. 3.8 per 1000, p=.17); the authors concluded that this was primarily due to better detection in women aged 40-49 with extremely dense breast tissue.

It is also worth noting in Table 4 that in Europe, the recall rate for mammography is generally about half that observed in the United States.[115-117] Thus one of the harms of mammography, recalls for false positive imaging results, is less common in Europe. It will be important to keep this in mind when evaluating how to apply the results from studies of supplemental screening performed in Europe to the United States.

Table 4: Recall rates and cancer detection using film and digital mammography in large studies of screening irrespective of density.

Study	*Type*	*Mammograms, n*	*Recall rate /1000*	*Biopsy rate /1000*	*Cancer detection /1000*	*PPV3*
DMIST	Film	42,555	86	16.0	4.1	24.4
Pisano 2005[75]	Digital	42,555	86	15.9	4.4	26.0
Vestfold	Film	324,763	42	NR	6.5	15.1
Vigeland 2008[114]	Digital	18,239	41	NR	7.7	18.5
BCSC	Film	638,252	93	10.6	3.8	24.7
Kerlikowske 2011[113]	Digital	231,034	100	11.0	3.8	25.3

PPV3 = the positive predictive value for biopsies performed

In summary, the findings from both the DMIST and BCSC studies, along with the results from other high-quality studies, highlight a critical difference between digital and film mammography in women with dense breast tissue. The studies find that digital mammography is more sensitive than film mammography in women with dense breast tissue. Therefore the masking effect of breast density observed with film mammography is substantially reduced.

Breast Density and Cancer Risk

The initial report of an association of patterns of mammographic density and breast cancer was published in 1976 by John Wolfe.[118] He described four different parenchymal patterns seen on mammography and reported that women with dysplastic pattern with sheets of dense parenchyma had a markedly increased incidence of breast cancer compared to women with normal breast parenchyma. A recent meta-analysis summarizing the literature on the BI-RADS breast density

reported a four-fold increased risk for breast cancer in women with extremely dense breasts compared to women with fatty breasts (relative risk [RR] 4.0, 95% CI 3.1 to 5.3), similar to the Wolfe patterns.[119] Risk consistently increased with increasing category of density. Using the more prevalent group of women with scattered fibroglandular density as the reference group, the risk increases linearly across the four categories (RR 0.5, 1.0 [reference group], 1.5, and 2.0).

If the lifetime risk for breast cancer in the overall population of women is about 12%, then the lifetime risk for women with dense breasts would be approximately 15%. However, lifetime risk is not helpful in deciding when to begin to screen for breast cancer or when to add additional screening – a five or ten year time frame is more clinically relevant. In one study of 629,229 women the observed five-year incidence of invasive breast cancer increased from 7.5 per 1000 women in the almost entirely fatty group to 12.4 in the scattered fibroglandular density group, 16.5 in the heterogeneously dense group, and 18.1 in the extremely dense group.[109]

Because high breast density is both a strong risk factor (relative risk of 1.5 for heterogeneously dense and 2 for extremely dense compared to scattered fibroglandular densities) and it is common (about 40% of women are in the heterogeneously dense category and 10% in the extremely dense category) it explains a greater proportion of the risk for breast cancer in the population than any risk factor other than age. For example, having a first-degree relative with breast cancer almost doubles a woman's risk for breast cancer, but only 10% to 20% of women have a positive family history. Similarly, carrying a BRCA mutation increases a woman's risk by a factor of 10 to 20, but less than 0.5% of women have a deleterious mutation.

One of the common concerns raised about the association between breast density and cancer risk is whether the elevated risk is due solely to the dense tissue masking breast cancers that are present at the time of mammography. If there were only masking, then there would be an increase in cancers detected over the next one to two years in women with dense breasts (those missed on mammography that should have been found) compared to those with fatty breasts, but this excess should not continue beyond two to three years. However, two large studies found no decrease in the strength of the association between breast density and breast cancer incidence through ten years of follow-up.[120,121] This provides strong evidence that the association of mammographic density with breast cancer represents a true association that is not an artifact arising from the masking of prevalent cancers alone.

Breast Density and Risk Assessment

Risk assessment forms the foundation of all screening and prevention programs. Screening programs using mammography for the early detection of breast cancer generally use age as the primary factor to determine eligibility for screening because age is the strongest risk factor for breast cancer. If the incidence of breast cancer is low, the harms associated with screening

outweigh the benefits through early detection and treatment of breast cancer. Conversely, for women at higher risks of breast cancer, earlier and more intensive forms of screening offer the possibility of a more favorable risk-benefit ratio. For example, the American Cancer Society (ACS) guidelines recommend annual MRI screening for women with a lifetime risk for breast cancer above 20% to 25%.[50] This risk threshold was chosen based on expert opinion.[50] The FDA indication for the use of tamoxifen to prevent breast cancer is specifically for women with a 5-year risk greater than 1.66% and, similarly, the 2013 American Society for Clinical Oncology guidelines recommend that physicians consider the use of medications to reduce the risk of breast cancer in women with a 5-year risk greater than 1.66%.[122] The 1.66% five-year risk threshold was the primary inclusion criteria for the Breast Cancer Prevention Trial, which demonstrated that tamoxifen reduces the risk of breast cancer by about 50%.[123]

It is worth noting here that using a five or ten-year time frame for estimating risk is more useful than lifetime risk when deciding when to initiate screening for breast cancer. No one would recommend that a ten year old girl with a lifetime risk for breast cancer of 25% be screened with MRI for breast cancer. Her short-term risk is too low to justify the cost and potential harms. Similarly, a woman with a 2% five year risk for breast cancer by the Gail model could have a 10% lifetime risk or a 30% lifetime risk; in either case she would be eligible for a discussion of the risks and benefits of tamoxifen to lower her risk for breast cancer.

Investigators at the National Cancer Institute developed the most commonly used model of a woman's risk for breast cancer, the Gail model or Breast Cancer Risk Assessment Tool. This model uses a woman's reproductive history and the number of first-degree relatives with breast cancer to estimate her risk for invasive breast cancer.[124,125] A web-based calculator is available for women and their physicians to use: http://www.cancer.gov/bcrisktool/. The model estimates a women's risk of developing invasive breast cancer in the next five years as well as her lifetime risk for invasive breast cancer. The Gail model remains the most widely used tool for estimating a woman's future risk for breast cancer because it was the earliest validated model and it established the entry criteria for the Breast Cancer Prevention Trial.

The limited ability of the Gail model to discriminate high risk women from low risk women [126] has encouraged investigators to develop models that incorporate additional risk factors. Because breast density is both common and a strong risk factor for breast cancer, researchers have added it to new models.[109,127,128] Investigators at the BCSC developed a model that uses BI-RADS density in combination with a woman's age, race/ethnicity, family history, and history of breast biopsies to estimate her 5-year risk for breast cancer.[109] A web-based calculator using the BCSC model is available for women and their physicians (https://tools.bcsc-scc.org/BC5yearRisk/calculator.htm). The BCSC model has better risk discrimination than the Gail model and is more accurate in non-white women.[109] Drs. Chen and Gail updated the Gail model by adding a continuous measure of

breast density to their model[128], but continuous breast density is not routinely calculated or reported with mammography at this time.

Mammography screening may be the ideal time for risk assessment because women and their physicians are thinking about breast cancer risk when mammograms are ordered and because mammographic density is the most powerful predictor of breast cancer after age. The new legislation in Connecticut and elsewhere requiring notification of women with dense breasts about the potential for dense breast tissue to mask cancers and increase overall risk makes this an opportune moment for such discussions.

2. Clinical Guidelines

2.1 Magnetic Resonance Imaging (MRI) of the Breast

The American Cancer Society (2007)

http://www.cancer.org/cancer/breastcancer/moreinformation/breastcancerearlydetection/breast-cancer-early-detection-acs-recs

The ACS recommends annual MRI screening examinations "for women with an approximately 20-25% or greater lifetime risk of breast cancer, including women with a strong family history of breast or ovarian cancer and women who were treated for Hodgkin's disease. There are several risk subgroups for which the available data are insufficient to recommend for or against screening, including women with a personal history of breast cancer, carcinoma in situ, atypical hyperplasia, and extremely dense breasts on mammography."[50]

National Comprehensive Cancer Network (NCCN)

http://www.nccn.org/professionals/physician_gls/pdf/breast.pdf

The NCCN recommends that women with a lifetime risk of breast cancer greater than 20% (using Claus, BRCAPRO, BOADICEA, or Tyrer-Cuzick models) consider screening MRI as an adjunct to mammography starting at age 30. They also recommend screening MRI for women with mutations in BRCA1, BRCA2, TP53, or PTEN and their untested first-degree relatives. In addition, they recommend annual screening MRI for those receiving radiation therapy to their chest between the ages of 10 to 30 years starting 8 to 10 years following the radiation therapy or at age 40, whichever comes first.

The NCCN guidelines also state that there is insufficient evidence to recommend for or against annual MRI screening for the following women: those with a 15% to 20% lifetime risk for breast cancer; those with a personal history of breast cancer, ductal carcinoma in situ, lobular carcinoma in situ, atypical ductal hyperplasia; or those with heterogeneously dense or extremely dense tissue on mammography.

American College of Radiology / Society of Breast Imaging

http://www.jacr.org/article/S1546-1440(09)00480-3/fulltext

Joint guidelines from the American College of Radiology and the Society of Breast imaging recommend annual screening MRI examinations starting at age 30 for BRCA mutation carriers and their untested first degree relatives, for women with greater than a 20% lifetime risk for breast

cancer on the basis of family history, women with a history of chest irradiation (usually for Hodgkin's disease), and a single screen of the contralateral breast for women with newly diagnosed breast cancer.[44] They recommend considering screening MRI for women with a lifetime risk between 15% and 20% on the basis of a personal history of breast or ovarian cancer or biopsy proven lobular neoplasia or atypical ductal hyperplasia.

The European Society of Breast Imaging

http://www.eusobi.org/html/img/pool/330_2008_863_OnlinePDF.PDF

The European Society of Breast Imaging recommends annual MRI screening examinations for women with a BRCA mutation, first degree relatives of BRCA carriers, women with radiation to their chest wall between the ages of 10 and 30 years, women with Li-Fraumeni syndrome (TP53 mutation carriers) and their untested first degree relatives, and women with Cowden syndrome (PTEN mutation carriers) and their first degree relatives).[67]

2.2 Hand-held Ultrasonography (HHUS) of the Breast

The American Cancer Society

http://www.cancer.org/cancer/breastcancer/moreinformation/breastcancerearlydetection/breast-cancer-early-detection-acs-recs

The ACS has no recommendation on HHUS for breast cancer screening.

National Comprehensive Cancer Network (NCCN)

http://www.nccn.org/professionals/physician_gls/pdf/breast.pdf

Under breast screening considerations, the NCCN guidelines state "Dense breasts limit the sensitivity of mammography. Dense breasts are associated with an increased risk for breast cancer, but there is insufficient evidence to support routine supplemental screening in women with dense breasts and no other risk factors."[129] Under the same section they also note "There are several studies supporting the use of ultrasound for breast cancer screening as an adjunct to mammography for high risk women with dense breast tissue."

American College of Radiology / Society of Breast Imaging

http://www.jacr.org/article/S1546-1440(09)00480-3/fulltext

Joint guidelines from the American College of Radiology and the Society of Breast imaging recommend considering annual screening ultrasound examinations in addition to mammography

for women eligible for MRI screening who cannot have MRI for any reason.[44] They recommend considering ultrasound in women with dense breast tissue as an adjunct to mammography.

2.3 Automated Whole Breast Ultrasonography (ABUS)

There are no guidelines currently recommending ABUS to screen for breast cancer from any major clinical society, including the American Cancer Society, the National Comprehensive Cancer Network, the American College of Radiology, and the Society of Breast Imaging.

2.4 Digital Breast Tomosynthesis (DBT)

National Comprehensive Cancer Network (NCCN)

http://www.nccn.org/professionals/physician_gls/pdf/breast.pdf

Under breast screening considerations, the NCCN guidelines state that "Early studies show promise for DBT mammography. Currently, there is insufficient evidence to recommend routine use for screening or diagnosis at this time."[129]

There are no other guidelines currently recommending DBT to screen for breast cancer.

3. Medicaid, Medicare, National and New England Private Insurance Coverage Policies

3.1 Breast Ultrasound

Medicaid

No publicly available coverage policies for breast cancer screening using ultrasound were found for Medicaid agencies in New England.

Medicare

http://www.cms.gov/medicare-coverage-database/details/lcd-details.aspx?LCDId=29871&ContrId=205&ver=25&ContrVer=1&DocType=All&bc=AgIAAAAAAAAA%3d%3d&

The national coverage determination (NCD) for breast ultrasound relates only to its use for diagnosis rather than screening. An LCD is available for New England (see link above); while the document does not specifically state whether ultrasound may be used as a screening tool, a list of sample indications suggests coverage for diagnosis or to guide treatment only (e.g., .examination of palpable masses or masses detected on mammography, radiation treatment planning).

National Private Payers

http://apps.humana.com/tad/tad_new/Search.aspx?sortfield=name&policyType=medical

Humana considers ABUS experimental and investigational for any type of breast cancer screening. No information regarding coverage for either HHUS or ABUS was publicly available from other national payers such as CIGNA, Aetna, Unicare, United Healthcare, and WellPoint/Anthem.

Regional Private Payers

No publicly available coverage policies regarding supplemental screening using either HHUS or ABUS were available from any major private payer in New England.

3.2 Breast MRI

The majority of available coverage policies for supplemental screening with breast MRI generally limit such coverage to patients with genetic risk factors (e.g., presence of BRCA mutation or first degree relatives with BRCA mutations), lifetime breast cancer risk greater than or equal to 20%, or prior radiation therapy to the chest, without specific mention of dense breast tissue as a risk factor. Where available, coverage language relevant to women with dense breast tissue is summarized below.

Medicaid

http://www.mass.gov/eohhs/docs/masshealth/guidelines/mg-breastmri.pdf

Massachusetts Medicaid (MassHealth) does not cover breast MRI for screening in asymptomatic, average risk patients.

Medicare

http://www.cms.gov/medicare-coverage-database/details/lcd-details.aspx?LCDId=29871&ContrId=205&ver=25&ContrVer=1&DocType=All&bc=AgIAAAAAAAAA&3d%3d&

The above-mentioned LCD also provides information on coverage for breast MRI. As with ultrasound, the list of sample indications relate to diagnostic or treatment-planning uses. However, one listed indication allows for use of MRI in "cases where diagnosis is inconclusive, even after standard work-up." This is of potential interest because the ICD-9 code for an inconclusive mammogram is also the recommended code for a woman with dense breast tissue.[130]

National Private Payers

WellPoint/Anthem, Unicare and Aetna cover MRI as an adjunct to mammography annually in women with dense breasts ***and*** a personal history of breast cancer. Humana covers breast MRI as an adjunct to mammography when heterogeneous or extremely dense breast tissue is identified, regardless of breast cancer history.

Regional Private Payers

http://www.bluecrossma.com/common/en_US/medical_policies/230%20MRI%20of%20the%20Breast%20prn.pdf#page=1

MRI is considered investigational by BCBS (MA) for the following indications:

a) As a screening technique in average risk patients; or
b) As a screening technique when detection of mammography is limited due to dense breast tissue, breast implants, or scarring after treatment for breast cancer.

No other major regional private payers have publicly available coverage policies regarding MRI in women with dense breast tissue.

3.3 Digital Breast Tomosynthesis (DBT)

Medicaid

No publicly available coverage policies for breast cancer screening using DBT were found for Medicaid agencies in New England.

Medicare

No NCDs or LCDs were available for coverage of DBT.

National Private Payers

DBT is considered experimental, investigational or unproven for ***any*** purpose by Aetna, CIGNA, Humana, Unicare, United Healthcare and WellPoint/Anthem.

Regional Private Payers

Connecticare provides coverage for DBT but requires prior authorization. All other major regional private payers with publicly-available policies consider DBT to be investigational and do not cover it.

4. Status of Breast Density Legislation

Detailed descriptions of the status of legislation in each New England state are presented in the sections that follow, along with summaries of status in other states as well as descriptions of two national legislative efforts. It is important to recognize that legislative status is an ever-changing landscape; accordingly, this section should be considered a "snapshot" of status *at the time of report publication.*

4.1 New England

Connecticut

The only state in New England that currently has an active law requiring provision of breast density information to women is Connecticut. In addition, according to General Statute sections 38a-503 and 38a530, insurance coverage for supplemental screening with ultrasound is *required* for women with dense breast tissue on mammography. The law was amended in 2012 to mandate coverage for breast MRI screening, but only in accordance with American Cancer Society guidelines (i.e., 20-25% or greater lifetime risk, known BRCA mutations in themselves or a first-degree relative, history of chest radiation therapy between ages 10-30, or Li-Fraumeni syndrome, Cowden syndrome, or Bannayan-Riley-Ruvalcaba syndrome in themselves or a first-degree relative) *irrespective* of whether they also have dense breast tissue. Relevant notification language from the Connecticut law can be found below:

> *"(c) On and after October 1, 2009, each mammography report provided to a patient shall include information about breast density, based on the Breast Imaging Reporting and Data System established by the American College of Radiology. Where applicable, such report shall include the following notice: "If your mammogram demonstrates that you have dense breast tissue, which could hide small abnormalities, you might benefit from supplementary screening tests, which can include a breast ultrasound screening or a breast MRI examination, or both, depending on your individual risk factors. A report of your mammography results, which contains information about your breast density, has been sent to your physician's office and you should contact your physician if you have any questions or concerns about this report."*

Maine

Maine also considered the adoption of a mandate requiring notification of women with dense breast tissue, but the bill was removed from consideration in 2012 and replaced with a law authorizing the appointment of a multi-stakeholder group to study the issue. This group consisted of government agencies, clinical societies, and patient advocacy organizations, and was tasked with recommending strategies to improve communication between physicians and patients regarding breast density and imaging options. The group recommended that a "lay letter" be provided to women with dense breast tissue, consisting of the following or similar language:

> *"According to recent literature, dense breast tissue composition may be an increased or independent risk factor for malignancy. Consider secondary screening modalities if appropriate."*

The recommendations encouraged the radiologist to communicate with the ordering physician when a mammogram shows dense breast tissue. It was also recommended that imaging facilities send a sample of their "lay letters" to providers in order to make certain that they are informed about the language used.

The final report was sent to the Joint Standing Committee on Health and Human Services in January 2013.[131] Concerns have been raised that there is no way to ensure that every woman with dense breasts is notified in the absence of a legal mandate, but to date there have been no efforts made to reintroduce the original bill or a modified version.

Massachusetts

A proposed breast density notification law in Massachusetts was favorably reviewed by the legislature's Joint Committee on Public Health in May 2013 and subsequently referred to the Joint Committee on Health Care Financing for further discussion in October 2013. The bill does not currently contain any language mandating insurance coverage. Proposed notification language can be found below:

> *"On completion of a mammogram, a mammography facility licensed by the department of public health shall provide to the patient the following notice:*
>
> *'If your mammogram demonstrates that you have dense breast tissue, which could hide abnormalities, and you have other risk factors for breast cancer that have been identified, you might benefit from supplemental screening tests that may be suggested by your ordering physician.*

Dense breast tissue, in and of itself, is a relatively common condition. Therefore, this information is not provided to cause undue concern, but rather to raise your awareness and to promote discussion with your physician regarding the presence of other risk factors, in addition to dense breast tissue.

A report of your mammography results will be sent to you and your physician. You should contact your physician if you have any questions or concerns regarding this report.'"

New Hampshire

New Hampshire introduced a breast density notification bill in 2012. The bill was deemed "inexpedient to legislate" during committee discussions and was not referred to the full legislature for debate. No information is available regarding the reason for this decision.

Rhode Island and Vermont

Vermont and Rhode Island currently do not have any legislation under discussion that is related to breast density notification or insurance coverage for supplemental screening.

4.2 Status of Legislation in Other States

Other states with mandates for breast density notification that are currently operational include Alabama, California, Maryland, New York, Texas, and Virginia. In addition, Hawaii, Nevada, North Carolina, Oregon, Pennsylvania, and Tennessee have enacted notification laws that will take effect in 2014.

In addition to Connecticut, Illinois (effective 1/2014) and Indiana have mandates in place requiring insurance coverage for supplementary screening tests. The mandates in Illinois and Connecticut require coverage of ultrasound in women with dense breasts. The breast density law in Indiana requires insurance plans to provide coverage of "appropriate medical screening, tests, or examinations of women with dense breasts," without mention of specific modalities.[132] As mentioned above, Pennsylvania will be enacting a mandate on notification, and a bill on coverage for supplemental screening tests is pending. Finally, the New Jersey Senate has passed a bill regarding breast density notification and insurance coverage for ultrasound, which is now being considered by the New Jersey Assembly.

States that are currently considering notification bills include Colorado, Delaware, Georgia, Iowa, Michigan, Ohio, South Carolina, and Washington.

4.3 Nationwide Breast Density Legislation

The federal Mammography Quality Standards Act (MQSA) was first enacted in 1994 to ensure that all mammography facilities maintain uniform quality standards. The FDA, which is responsible for enforcing MQSA standards, has acknowledged that changes in mammography technology processes have occurred over time, and that the language in the MQSA may be worth revisiting.[133] As part of this process, a recommendation was made to standardize the reporting of mammographic breast density nationwide. The FDA will hold a public meeting to debate this change in early 2014. Separately, the Breast Density and Mammography Reporting Act [134] was introduced in the U.S. House of Representatives in 2011. The act would require all mammography facilities to inform patients with mammographically-dense breast tissue about breast density, the association of density with breast cancer risk and masking, and the possible benefits of supplemental screening. The bill was referred to the House Energy and Commerce Committee (Subcommittee on Health), but was never brought to the full House for a vote. The original co-sponsor of the bill, Rep. Rosa DeLauro (D-CT), reintroduced the bill in October 2013, where it was referred back to committee.[135]

5. Previous Systematic Reviews and Technology Assessments

We were able to identify only one publicly available technology assessment focusing on the use of advanced imaging following negative mammograms (see BCBSA TEC below) and one systematic review on the use of HHUS in women at average risk for breast cancer. Because of MRI's importance, we have also summarized the major assessments of the use of MRI in women at high risk for breast cancer.

5.1 Formal Health Technology Assessments

California Technology Assessment Forum (CTAF, 2013):
http://www.ctaf.org/assessments/supplemental-cancer-screening-women-dense-breasts

In an evidence review focused on supplemental screening in women with dense breast tissue and a normal mammogram, a majority of CTAF concluded that the evidence is adequate to demonstrate that supplemental screening with any technology provides more benefit than harm in women at high overall risk of breast cancer (i.e., 5-year risk >3%). However, CTAF found the evidence to be inadequate to suggest that the benefits of supplemental screening in women at low (<1.7%) or moderate (1.7-3%) 5-year risk outweigh the harms. When asked to rank the screening modalities in terms of the preferred choice in high-risk women, CTAF ranked MRI first, followed by HHUS, ABUS, and DBT. CTAF Panel members also voted that the evidence is adequate to demonstrate that digital mammography is superior to film mammography for women with dense breast tissue, and that compared to film mammography, digital mammography greatly reduces the risk of "masking" of breast cancers.

In comparisons of value relative to the lowest-cost test available (HHUS), CTAF considered both ABUS and MRI to be of reasonable value. DBT was felt to be of low value due to limited evidence of effectiveness as a supplemental screening tool and increased radiation exposure.

Blue Cross BlueShield Association Technology Evaluation Center (BCBS TEC, 2013):
Final report is in press. Executive summary found here:
http://www.bcbs.com/blueresources/tec/vols/27/special-report-screening.html

In a technology assessment on screening asymptomatic women with dense breasts and a normal mammogram, available data suggests that digital mammography is more sensitive than film

mammography. The combination of mammography and ultrasound is more sensitive than mammography alone, but also results in more false positives and unnecessary biopsies. Evidence also suggests that MRI is more sensitive than mammography, although this evidence was generated in women with dense breasts who were also at high cancer risk from other factors. There is insufficient evidence on automated breast ultrasound and DBT in women with dense breasts.

Canadian Agency for Drugs and Technologies in Health (CADTH, 2007):
http://www.cadth.ca/media/pdf/I3010_MRI-Breast-Cancer_tr_e.pdf

Based on available evidence on screening women at high risk (no RCTs available), MRI was found to be more sensitive and cost-effective compared to mammography. The high risk category included women who were BRCA1 or BRCA2 carriers, their first degree relatives, and those with a strong family history of breast cancer. There was no mention of breast density.

Health Information and Quality Authority, Ireland (HIQA, 2013):
http://www.hiqa.ie/healthcare/health-technology-assessment/assessments/surveillance-of-women-under-50-with-increased-risk-of-breast-cancer

In women age <50 with an elevated risk of breast cancer, evidence suggests that a combination of MRI and digital mammography or MRI alone is more sensitive but less specific than mammography alone. MRI also contributes to decreased breast cancer but at an increased cost. Nevertheless, they estimated that offering annual MRI to women age 30-49 at moderate (10 year risk of 3-8%) or high (10-year risk >8%) risk would be cost-saving relative to an ad-hoc (i.e., at clinician discretion) surveillance approach or no surveillance.

New Zealand Health Technology Assessment Program (NZHTA, 2007):
http://www.otago.ac.nz/christchurch/otago014084.pdf

In women at high risk of breast cancer, ultrasound has equivalent sensitivity to mammography but produces more false positives. The high risk category includes women with a strong family history of breast cancer including women with and without known genetic mutations which predispose to breast cancer. There was no mention of breast density. As with mammography, sensitivity with ultrasound decreases as risk status increases. MRI-based surveillance is more sensitive than mammography, ultrasound, or mammography and ultrasound combined, as accuracy does not appear to diminish with increased risk status. MRI does produce more false-positives than mammography, although these decreased over time in studies evaluating whether a "learning curve" was present for MRI interpretation.

Ontario Health Technology Assessment (2010):
http://www.health.gov.on.ca/english/providers/program/mas/tech/reviews/pdf/rev_%20breast_cancer_screening_20100316.pdf

Data suggests that digital mammography is significantly more sensitive than film mammography in women with heterogeneously or extremely dense breast tissue of any age, asymptomatic women age <50, and those who are pre- or peri-menopausal. There is no evidence of differences in recall rates, however. Evidence suggests that the sensitivity of MRI is significantly higher than that of film mammography in women at high breast cancer risk due to genetic and/or familial factors, regardless of age. There is moderate evidence to suggest that the combination of mammography and MRI is significantly more sensitive than either modality alone in women at high risk of breast cancer from genetic/familial factors, although specificity is either unchanged or decreases.

5.2 Systematic Reviews of Magnetic Resonance Imaging (MRI) of the Breast for Women at High Risk

Lord 2007
Lord SJ, Lei W, Craft P, et al. A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. *Eur J Cancer.* Sep 2007;43(13):1905-1917.
http://www.sciencedirect.com/science/article/pii/S0959804907004844

Lord and colleagues published a systematic review and meta-analysis of 5 studies that evaluated the impact of MRI added to mammography in young women at high risk of breast cancer.[136] They found consistent evidence that MRI was more sensitive (93-100%) than mammography (25-59%). The addition of MRI to women with negative mammograms identified an additional 10 to 24 cancers per 1000 examinations. There was an increase in false positive results (71-74 additional false positive follow-ups per 1000 screens, a 3 to 5-fold increase) and an increase in biopsies with a benign diagnosis (7-46 additional benign biopsies per 1000 screens). There were no studies assessing whether the addition of MRI reduces patient mortality, interval, or advanced breast cancer rates. They conclude that the benefits of MRI in young, high-risk women have not been established by the evidence.

Warner 2008

Warner E, Messersmith H, Causer P, Eisen A, Shumak R, Plewes D. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. *Annals of Internal Medicine*. May 6 2008;148(9):671-679.
http://annals.org/article.aspx?articleid=740814

Warner and colleagues published a systematic review and meta-analysis of 11 studies that evaluated the impact of MRI added to mammography in women at high risk of breast cancer.[137] Using a BI-RADS assessment of 4 or 5 as the definition of a positive test, they found consistent evidence that adding MRI to mammography increased the sensitivity for breast cancer from 32% to 84% with a corresponding decrease in specificity from 98.5% to 95.2%. The PPV3 of MRI added to mammography was 25.0%. The authors conclude that annual screening with MRI and mammography is the most accurate approach to screening women with strong familial or genetic predisposition to breast cancer.

5.3 Systematic Reviews of Hand-held Ultrasonography (HHUS) of the Breast

Cochrane review

Gartlehner G, Thaler K, Chapman A, et al. Mammography in combination with breast ultrasonography versus mammography for breast cancer screening in women at average risk. *The Cochrane database of systematic reviews.* 2013;4:CD009632.
http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009632.pub2/abstract;jsessionid=1495F3C42D5F3ED564A5977F83106936.d04t03

The 2013 Cochrane review of the use of ultrasound in addition to mammography to screen average risk women for breast cancer was recently published.[138] The authors concluded, "No methodologically sound evidence is available justifying the routine use of ultrasonography as an adjunct screening tool in women at average risk for breast cancer." For women with dense breasts they concluded, "despite the increased risk for breast cancer and the limitations of mammography in women with dense breast tissue, the available evidence supporting the use of adjunct ultrasonography as a screening tool in women with dense breasts (BI-RADS 3-4) is limited and has to be interpreted cautiously."[138]

Nothacker

Nothacker M, Duda V, Hahn M, et al. Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. *BMC Cancer.* 2009;9:335.
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2760575/

Nothacker and colleagues performed a systematic review of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue.[139] They did not identify any randomized trials or prior systematic reviews. They identified six cohort studies[140-145] of fair quality, but only two of the studies reported follow-up[141,142] and both were inadequate. They estimated that the cancer detection rate with supplemental ultrasound was 3.2 per 1000 screens among women with negative mammograms and BI-RADS 2 – 4 density (scattered fibroglandular density through extremely dense). They concluded that there is limited evidence that an additional ultrasound examination after a negative mammogram is useful for the detection of breast cancer in women with mammographically dense tissue.

5.4 Systematic Reviews of Automated Whole Breast Ultrasonography (ABUS)

No systematic reviews or technology assessments were found for ABUS.

5.5 Systematic Reviews of Digital Breast Tomosynthesis (DBT)

Houssami N, Skaane P. Overview of the evidence on digital breast tomosynthesis in breast cancer detection. *Breast.* Apr 2013;22(2):101-108.
http://www.sciencedirect.com/science/article/pii/S0960977613000192

Houssame and Skaane published a systematic review that summarized the results of 14 studies of DBT. None of the studies were randomized trials, addressed breast density subgroups, or reported follow-up data on breast cancer specific mortality or distant-disease recurrence. They report that the studies are preliminary, but suggest that DBT has promise in reducing both false positives and false negatives relative to or when added to digital mammography. The authors highlight five large trials that are ongoing in population-based breast cancer screening programs internationally. They conclude that at this time there is insufficient evidence to justify the widespread use of DBT.

6. Ongoing Studies

Title/ Trial Sponsor	*Study Design*	*Comparators*	*Patient Population*	*Primary Outcomes*	*Estimated Completion Date*
Magnetic Resonance Imaging (MRI)					
Breast Cancer Screening With MRI in Women Aged 50-75 Years With Extremely Dense Breast Tissue: the DENSE Trial (Phase 4) NCT01315015	RCT	MRI (n=7,237) DM (n=28,948)	• Age: 49 - 75yrs • Density > 75% (D4) • Females only • Negative mammographic assessment (1 or 2)	Number of interval cancers between the MRI group and the control group	December 2019
Familial MRI Screening Study (FaMRIsc) NTR2789	RCT	MRI+CBE (n=1000) DM+CBE (n=1000)	• Ages 30-55 years • Lifetime risk 20% to 49% • Exclude BRCA1/2	Number and stage of screen detected cancers stratified by breast density. Secondary: false positive rate, sensitivity, PPV	January, 2015
Hand Held Ultrasound (HHUS)					
Japan Strategic Anti-cancer Randomized Trial (J-START) UMIN000000757	RCT	DM + HHUS (n=50,000) DM (n=50,000)	• Ages 40-49 years • Exclude prior breast cancer	Sensitivity and specificity, incremental cancer detection rate.	March, 2016

Title/ Trial Sponsor	*Study Design*	*Comparators*	*Patient Population*	*Primary Outcomes*	*Estimated Completion Date*
Ultrasound and Mammography for Screening Breast Cancer in Chinese Women NCT01880853	RCT	DM HHUS DM+HHUS	• Ages 30-65 years • Not pregnant or lactating • No breast implants • No metastatic disease • No symptoms	Cancer detection, sensitivity, specificity, PPV, NPV	December 2011 (no update provided)
Automated Whole Breast Ultrasound (ABUS)					
A Clinical Study to Evaluate Somo•v and Digital Mammography (XRM) Together as a Breast Cancer Screening Method, Compared to Digital Mammography Alone, in Women With Dense Breasts. (somo•InSIGHT) NCT00816530	Cohort	DM + ABUS DM	• n = 20,600 • Age > 25 • Density >50% • Females only • Not pregnant or breastfeeding • No breast surgeries or interventional breast procedures past 12 months • No signs or symptoms of breast cancer • Breast implants allowed	Sensitivity of DM and ABUS together vs DM alone	December 2012 (no update provided)

Title/ Trial Sponsor	*Study Design*	*Comparators*	*Patient Population*	*Primary Outcomes*	*Estimated Completion Date*
Earlier Breast Cancer Detection Using Automated Whole Breast Ultrasound With Mammography, Including Cost Comparisons NCT00649337	RCT	ABUS DM	• n = 4650 • Age: 35 – 90yrs • Females only • No screening mammogram in the past 10 months • No history of breast cancer for at least one year	Number of breast cancers detected one year after screening: ABUS vs blinded DM	January 2010 (no update provided)
Digital Breast Tomosynthesis (DBT)					
Comparison of Diagnostic Performance of Digital Breast Tomosynthesis (DBT) and Ultrasound (US) in Women With Dense Breasts NCT01910103	Cohort	DBT HHUS	• n = 825 • Age > 20 • Density >50% • Females only • No previous history of breast surgery or breast core biopsy performed within the prior 6 months	Performance of DBT and US in detecting breast cancer in screening and diagnostic settings	April 2014

Title/ Trial Sponsor	*Study Design*	*Comparators*	*Patient Population*	*Primary Outcomes*	*Estimated Completion Date*
Comparison of Full-Field Digital Mammography With Digital Breast Tomography for Screening Call-Back Rates NCT01236781	NonRCT	DM Combination of 2-D and 3-D DBT	• n = 500 • Age > 25 • Females only • Asymptomatic; scheduled for FFDM • Not pregnant or lactating • Breasts too large to allow for adequate positioning for the DBT examination • No breast implants • DBT of DM 11 months prior to study registration	Recall rates (1 year) between DM and DBT	June 2012 (no update provided)
Malmö Breast Tomosynthesis Screening Trial (MBTST) NCT01091545	Cohort	DM + DBT	• N = 15,000 • Ages 40-74 years • Not pregnant	Incremental cancer detection rate, sensitivity, and specificity	June 2014

Title/ Trial Sponsor	*Study Design*	*Comparators*	*Patient Population*	*Primary Outcomes*	*Estimated Completion Date*
Other					
A Multi Modality Surveillance Program for Women At High Risk for Breast Cancer NCT00989638	Cohort	DM MRI Breast biopsy	• n = 500 • Age > 18 • Females only • High risk for cancer including dense breasts, known BRCA 1 or 2 mutations, other hereditary breast cancer susceptibility genes • No active cancer at enrollment • Not pregnancy • No breast surgery within two weeks of study entry • No previous bilateral mastectomy (prophylactic or therapeutic) • No history of kidney disease or abnormal kidney tests	Early detection of small or pre-cancerous lesion(s) using a combination of screening measures including biomarkers	June 2015

7. Evidence Review (Methods & Results)

The goal of this technology assessment is to evaluate the comparative effectiveness and value of advanced imaging technologies when used to evaluate women undergoing screening digital mammography who have normal mammography results, but are noted to have heterogeneously dense or extremely dense breasts by the radiologist. The literature search revealed few studies that evaluated supplemental screening following digital mammography, so we expanded our search to include studies using film mammography. We did not assess the value of imaging as a diagnostic tool to assess women with a breast lump or an abnormal finding nor did we assess the value of imaging for high-risk women for whom digital mammography plus MRI is often used.

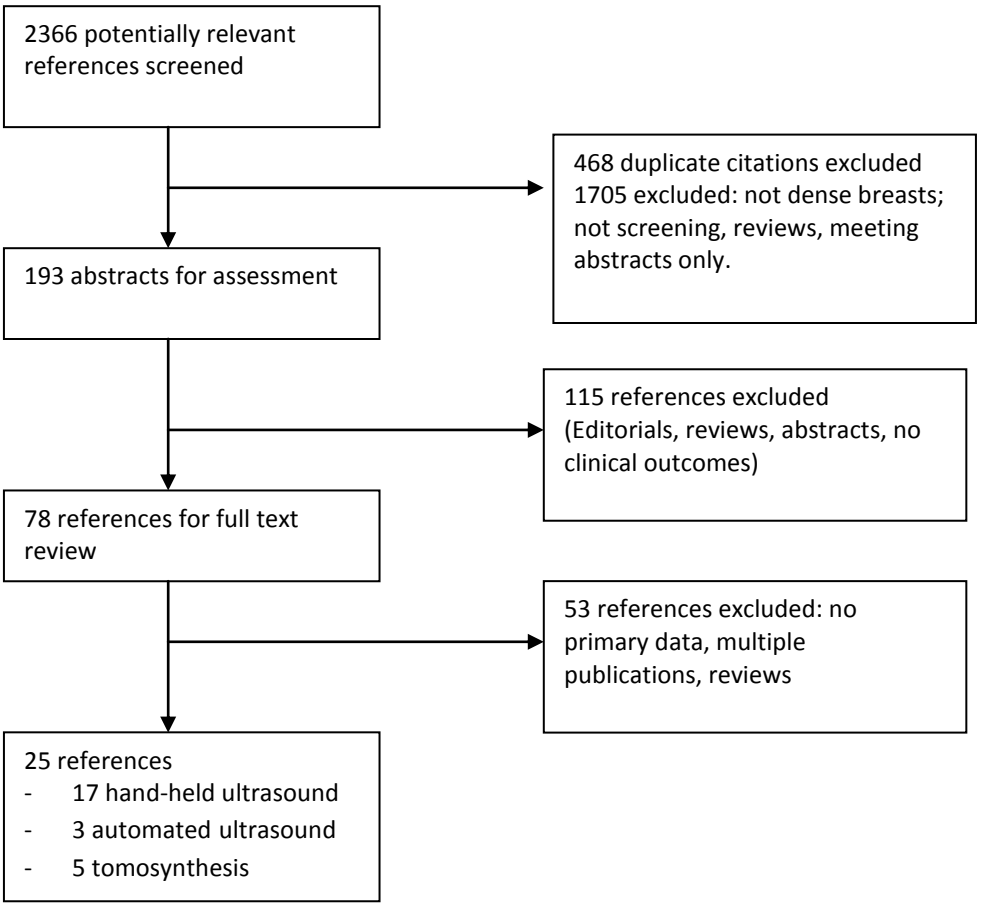
The ideal evidence would come from studies that randomize women meeting the above criteria to additional imaging or no additional imaging and follow them over time for breast cancer mortality. In the absence of such studies, this assessment focused on test characteristics of additional imaging in this population. These characteristics included the rates of true and false positive results, true and false negative results, biopsies, and cancer detection in addition to the sensitivity, specificity, positive predictive value and negative predictive value of the screening tests. In order to estimate the false negative rate and sensitivity of the test, the studies must follow the patients for 12 months to detect interval cancers. Given the paucity of such observational studies, we have included cross-sectional studies that report the cancer detection rate.

The Medline database, Embase, Cochrane clinical trials database, Cochrane reviews database and the Database of Abstracts of Reviews of Effects (DARE) were searched using the key words "ultrasound" OR "tomosynthesis" OR "magnetic resonance" AND "screen" AND "breast neoplasms" OR "breast cancer." The search was performed for the period from 1945 through June 11, 2013. Full details of the search are in the Appendix. The bibliographies of systematic reviews and key articles were manually searched for additional references. The abstracts of citations were reviewed for relevance and all potentially relevant articles were reviewed in full. We included all studies of imaging technologies used to screen for breast cancer in women with a recent (within 30 days) negative digital mammogram and high breast density. A negative mammogram was defined by BI-RADS assessments 1 or 2 (negative or benign finding) and high breast density was defined as heterogeneously dense or extremely dense using the BI-RADS density criteria. If digital mammography was not used, we included film mammography for completeness.

The search identified 2365 potentially relevant studies (see Figure 1 on the following page). After elimination of duplicate and non-relevant references, the search identified no studies evaluating MRI in this population, 17 publications describing 14 studies[140-156] evaluating hand-held ultrasound, three studies[157,158] evaluating automated whole breast ultrasound, and five publications describing

four studies[55,159-162] evaluating DBT. The primary reasons for study exclusion were (a) focus on high-risk populations defined without regard to breast density (BRCA carriers, lifetime risk > 25%, personal history of breast cancer, recent diagnosis of breast cancer); (b) use of imaging for diagnosis of a suspicious mass or abnormality on mammography only; or (c) studies were only reader reliability studies or descriptions of ongoing studies without results. Because the initial search criteria excluded many of the studies initially identified, we expanded our inclusion criteria to capture studies that defined density as scattered fibroglandular densities, heterogeneously dense or extremely dense tissue and those that only reported cancer detection rates. We also included a summary of the American College of Radiology Imaging Network (ACRIN) 6666 study[68,163] because it is widely cited and relevant to practice patterns in the United States. Finally, to give some perspective on the relative value of MRI, we summarized the larger studies and systematic reviews of MRI to screen women at very high risk due to hereditary breast cancer or its equivalent.

Figure 1. Selection of studies for inclusion in review.



The two most important outcomes in breast cancer are breast cancer specific mortality and disease-free survival. Because early stage breast cancer has such a long natural history and the majority of women do well, large randomized trials with long follow-up are needed to demonstrate the improvements in these outcomes in patients screened with additional imaging technologies. For short-term studies, the potential benefit of additional screening is best summarized by the incremental cancer detection rate. The potential harms can be assessed by evaluations of the false positive rate, the recall rate, the biopsy rate, the positive predictive values, and the radiation dose.

7.1 Screening Magnetic Resonance Imaging (MRI) of the Breast

The search did not identify any studies that evaluated the incremental benefit of MRI following negative mammography in women with dense breasts. The ACRIN 6666 study offered MRI to high-risk women who completed the third round of annual screening ultrasound and mammography in that study. The results will be described in the section on HHUS below. Several large studies have evaluated the test characteristics of MRI in conjunction with mammography and ultrasound in BRCA1 and BRCA2 mutation carriers and other women at very high risk for breast cancer. Those studies are summarized briefly below.

Magnetic Resonance imaging (MRI) for Screening High-risk Women

Magnetic resonance imaging (MRI) has been studied for breast cancer screening in women deemed to be at high risk either by personal history, family history or because they were known carriers of either a BRCA1 or BRCA2 mutation.[59-66,164-175] These women have a lifetime risk greater than 20%, rather than the 10% to 20% lifetime risk for most women with high breast density. No studies have demonstrated that MRI reduces the risk of death from breast cancer: there are no studies comparing women screened with MRI to other women screened with mammography alone and none of the studies of the test characteristics of MRI are of sufficient duration or size to evaluate patient-oriented outcomes such as the breast cancer recurrence or death from breast cancer. Table 6 on the following page summarizes the 11 larger prospective screening studies (n = 5652) that compare the use of MRI in high-risk woman to mammography with or without ultrasound. Women in these studies were primarily BRCA1 or BRCA2 mutation carriers or their first degree relatives. None of the studies followed women for more than one or two years. In addition, the majority of these studies compared MRI to film mammography only since digital mammography was not widely disseminated until after publication of the DMIST trial in 2005.

Table 6: Prospective studies comparing magnetic resonance imaging, ultrasound, and mammography to screen high-risk women for breast cancer.

Study	*Women, N*	*CDR MRI /1000*	*Bx rate /1000*	*PPV1 MRI %*	*PPV2 MRI %*	*PPV3 MRI %*	*Sensitivity*			*Specificity*		
							M %	*HHUS %*	*MRI %*	*M %*	*HHUS %*	*MRI %*
Tilanus-Linthorst 2000[65]	109	28	46	-	60	60	0*	-	100			
Podo 2002[168]	105	67	86	-	89	89	13	13	100			
Kriege 2004[61]	1909	12	29	-	57	57	40	-	71	95	-	90
Warner 2004[66]	236	30	157	-	46	46	36	33	82	99	96	81
Kuhl 2005[62]	529	36	147	-	50	50	32	40	91	97	91	97
Leach 2005[63]	649	29	-	-	-	25	40	-	77	93	-	81
Lehman 2005[166]	367	8	63	-	17	17	25	-	100	98	-	93
Lehman 2007[167]	171	23	82	-	43	43	33	17	100	91		79
Sardanelli 2007[64]	278	22	90	-	60	60	59	65	94	99	98	98
Kuhl 2010[176]	687	15	34	-	48	48	33	37	93	99	98	98
Berg 2012[163]	612	15	70	-	-	19	31	-	88	92	-	76

M: Mammography; HHUS: Ultrasound; MRI: magnetic resonance imaging

CDR = cancer detection rate

PPV1 = positive predictive value of a positive test result (BI-RADS assessment 0, 4, or 5)

PPV2 = positive predictive value of a biopsy recommended (BI-RADS assessment 4 or 5)

PPV3 = positive predictive value of biopsies actually performed

The sensitivity of MRI for breast cancer in Table 6 ranged from 77% to 100%. The sensitivity of mammography (25%-59%) and ultrasound (13%-65%) in these studies was about half that of MRI. In the largest three studies[61-63], which included 52% of the cancers in all 14 studies, the sensitivity of MRI ranged from 71% to 91% while the sensitivity of mammography ranged from 32% to 40%. However, the specificity of MRI is consistently lower than mammography. In the same three studies, the specificity of MRI ranged from 81% to 97% compared to 93% to 99% for mammography, and in each individual study the specificity of MRI was lower than that of mammography. Because breast cancer is relatively uncommon, even in these high-risk populations, the lower specificity of MRI translates into a much higher number of false positive results. One study suggested that the high false positive rate decreases after the initial MRI.[164] In that study the rate of false positive results declined from 14% initially to 8.2% on subsequent MRI's, but was still substantially higher than the 4.6% false positive rate for mammography.[164]

The cancer detection rate of MRI ranged from 8 to 36 per 1000 examinations in these studies (see Table 6 on page 44) – much higher than the 3 to 6 per 1000 examinations typically reported in studies of mammography (Tables 1, 3, and 4). This reflects in part the higher sensitivity of MRI and in part the higher incidence of breast cancer in these high-risk women. However, this higher cancer detection rate comes at a cost: the biopsy rates in the MRI studies in Table 6 range from 29 to 157 biopsies per 1000 examinations. The biopsy rates are lower in studies of screening mammography (10 to 25 per 1000 examinations, Tables 1 and 4). The PPV3 ranged from 17% to 89%, but the median was 48%, which is a very high yield per biopsy.

It is worth noting in Table 6 that the sensitivity of mammography and ultrasound were similar to each other in each of the five studies that report the sensitivity of all three screening technologies. The sensitivity of mammography and ultrasound in these studies is much lower than the sensitivity usually reported for these tests. The low sensitivity is due to the large number of cancers that are found by MRI alone – more than typically appear as interval cancers in the year following a screening examination. This suggests that many of the cancers detected by MRI would not have been diagnosed without MRI for more than one year after the examination. Early detection of cancers that would have become clinically apparent at a later date should translate into a higher cure rate and the need for less aggressive therapies, but some proportion of the cancers detected by MRI are likely to represent overdiagnosis – cancers that never would have become symptomatic in a woman's life.

The two systematic reviews described in Section 5 (Previous Systematic Reviews and Technology Assessments) both found that the addition of MRI significantly increased the sensitivity of screening for breast cancer, but increased false positive results; the effect on breast cancer mortality remained unknown because none of the studies had sufficient follow-up duration to evaluate this endpoint.[136,137] In one of the meta-analyses,[137] adding MRI to mammography increased the sensitivity from 39% to 94%, but decreased specificity from 94.7% to 77.2%. If the prevalence of breast cancer in a high-risk

population is 4.4% (the pooled prevalence across the 14 studies), then adding MRI to mammography in 1000 women would detect an additional 24 breast cancers (increased from 17 to 41) and an additional 167 women would receive false positive results (increased from 51 to 218).

Summary: Screening MRI of the Breast

There are no data evaluating MRI in a general screening population with dense breasts, nor in populations at intermediate risk (15% to 20% lifetime risk). The data from high-risk populations suggests that the addition of MRI would more than double the cancer detection rate (best estimate 2.4-fold increase) with a four-fold increase in the recall rate (best estimate 4.3-fold increase). Estimates based on these data are shown in Table 7 below. There is a high level of uncertainty around these values because of the lack of direct evidence from studies of MRI in women with dense breast tissue and because of the heterogeneity of the findings in the studies of high risk women summarized in Table 6.

Table 7: Estimated incremental yield of MRI after negative digital mammography in women with dense breast tissue.

Statistic	***Digital mammography***	***Incremental yield with MRI***	***Uncertainty***
Recall rate per 1000	128	100	High
Biopsy rate per 1000	17.8	17-36	High
CDR per 1000	4.2	8	High
PPV3	24%	22%-48%	High

These estimates suggest that MRI would find an additional 8 cancers more than those found by digital mammography, with a PPV3 between 22% and 48%. There would be approximately 100 additional recalls and between 17 and 36 additional biopsies in order to identify these cancers.

7.2 Screening Hand-held Breast Ultrasound (HHUS)

Fourteen studies of almost 70,000 women screened with HHUS met the search criteria for this assessment and are described in Table 8 beginning on page 48, with a quality assessment of these studies presented in Table 9 beginning on page 52.[140-156] In general, all participants in these studies underwent mammography first and those with negative mammograms were subsequently screened by HHUS. One study by Corsetti and colleagues is presented twice in the tables: their 2008 publication[145] had a large number of examinations; and their 2011 publication[151] included one year follow-up for a subset of the women. Results from the ACRIN 6666 trial[68,163] are also described in the

tables, although the study did not meet the inclusion criteria. However, it was the only prospective study in the United States with complete reporting of the data on the combination of mammography and HHUS with one-year follow-up after more than one round of screening.

As shown in Table 8 on the following page, the participants in these studies had a mean age usually in the 50s with a broad range (25 to 91 years). Most included asymptomatic women presenting for screening mammography who were found to have dense breasts, although the definition of high density varied somewhat. The majority of the trials were done outside of the United States. Three recent retrospective cohorts (Hooley 2012; Weigert 2012; Parris 2013)[152,154,156] described the findings in Connecticut, which was the first state to pass a law requiring breast density notification. These three studies represent the best evidence in the US population for the incremental cancer detection rate with HHUS, although they do not include any data on the interval cancer rate. The two other trials in the United States (Kaplan 2001; Kolb 2002) reported results from imaging performed in the year 2000 and earlier.[141,142] A radiologist performed the HHUS in the majority of the studies. Nine of the studies reported no follow-up on participants, two reported variable follow-up on a subset of patients, and three reported one-year follow-up. This is typical for publications of data from mammography facilities as they keep records on the follow-up of abnormal tests and cancer detection for quality assurance work, but do not routinely follow patients with normal mammography results to identify interval cancers. This means that the sensitivity, specificity, and negative predictive value reported from those studies will overestimate the true values.

Table 9 on page 52 describes elements of the study design that affect study quality. The interval between the mammogram and the HHUS examination should be relatively short. Otherwise the HHUS may find cancers that would also be visible on a mammogram at a later point in time. One study (Hooley 2012) included HHUS results from as much as 361 days after the mammogram – it is likely that mammography at that point would find additional cancers as well.[152] High quality studies performed the HHUS within one month of the mammogram. Eleven studies did not report the time interval between examinations, one study reported that there was an average of two months between the examinations, and three studies performed both examinations within the same month.

Table 8: Description of the studies.

Study	*Test*	*Years of study*	*Population*	*N*	*Follow-up*	*Age (years)*
MRI						
No Studies						
HHUS						
Maestro 1998[147] France	7.5, 10, or 13 Hz Esaote Biomedica Operator NR	1994-1995	Asymptomatic Screening "Dense" breasts	350	Variable	Mean 52
Buchberger 2000[140] Austria	5-12 MHz ATL Radiologist	1996-2000	Asymptomatic Screening D2, D3, D4	8103	None	Mean 48 Range 35-78
Kaplan 2001[141] United States	7-12 MHz GE Technician	1998-2000	Asymptomatic Screening D3, D4 subgroup	1862	Variable 72 followed for 1 year: 0 cancer	35-87
Kolb 2002[142] United States	5-12 MHz ATL 1 radiologist	1995-2000	Asymptomatic Screening D2, D3, D4	12,193 examinations in 4897 women	Variable. All participant with biopsy followed for 1 year	Mean 55
Crystal 2003[143] Israel	5-12 MHz ATL Radiologist	2000-2002	Asymptomatic Screening D2, D3, D4	1517	None	Mean 52 Range 31-84
Leconte 2003[144] Belgium	4.8-9.6 MHz Elegra, Siemens Radiologist	2000-2001	Mix 3% symptomatic, 24% breast cancer follow-up, 76% screening D1-D4	4236 3084 screening	None	NR

Study	*Test*	*Years of study*	*Population*	*N*	*Follow-up*	*Age (years)*
Brancato 2007[149] Italy	10-14 MHz Esaote Technos Radiologists	2003-2006	Asymptomatic Screening D3, D4 subgroup	5227	None	NR for subgroup
De Felice 2007[150] Italy	10-13 MHz Aloka, GE Radiologist	2000-2006	Asymptomatic Screening D3, D4 subgroup	1754	None	NR
Corsetti 2008[145] Italy	7.5-10 MHz Aloka Pro Sound Physician	2000-2007	Asymptomatic Screening D3, D4 subgroup	9157	None	Mean 52
Corsetti 2011[151] Italy	7.5-10 MHz Aloka Pro Sound Physician	2001-2006	Asymptomatic Screening D3, D4 subgroup	7224 examinations of 3356 women	One year	NR for subgroup
Hooley 2012[152] United States	12.5 – 17.5 MHz Phillips IU22 Technician	2009-2010	Asymptomatic Screening D3, D4 subgroup	935	One year	Mean 52 Range 29-89
Leong 2012[153] Singapore	7-10 MHz Toshiba PowerVision Technician	2002-2004	Asymptomatic Screening D3, D4 subgroup	141	One to two years	Mean 45 Range 30-64
Weigert 2012[154] United States	12.5 MHz Technician	2009-2010	Asymptomatic Screening D3, D4 subgroup	8647	None	NR
Girardi 2013[155] Italy	12 MHz Radiologist	2009-2010	Asymptomatic Screening D3, D4 subgroup	9960	None	Overall 51, range 33-84 NR D3/D4

Study	*Test*	*Years of study*	*Population*	*N*	*Follow-up*	*Age (years)*
Parris 2013[156] United States	12 MHz Philips HDI 5000 Technician	2009-2010	Asymptomatic Screening D3, D4 subgroup	5519	None	Mean 54
Berg 2012[163] ACRIN 6666 United States	≥ 12 MHz Radiologist	2004-2006	High-risk D3 in at least 1 quadrant of 1 breast	2659	One year for each of 3 rounds	Median 55 Range 25-91
ABUS						
Kelly 2010[157] United States	SonoCine	2003-2007	Asymptomatic Screening D3, D4 subgroup	6425 examinations for 4419 women	One year	Mean 53 Range 24 to 89
Stoblen 2011[177] Germany	SomoV, U-Systems	2008	Asymptomatic Screening Majority D2	304	None	Mean 58 Range 50 to 69
Giuliano 2013[53] United States	SomoV, U-Systems	2010-2011	Asymptomatic Screening Wolf density ≥ 50%	3418	One year	Mean 57 Range <50 to >70
DBT						
Ciatto 2013[159] STORM Italy	Selenia Dimensions, Hologic	2011-2012	Asymptomatic Screening D3, D4 subgroup	1127 D3/D4 7292 total	None	NR for D3/D4 subgroup Mean 58 Range 48-71

Study	*Test*	*Years of study*	*Population*	*N*	*Follow-up*	*Age (years)*
Rose 2013[161] United States	Selenia Dimensions, Hologic	2011-2012	Asymptomatic Screening Elected to have DBT D3, D4 subgroup	4666	None	NR for subgroup
Skaane 2013[55] Norway	Selenia Dimensions, Hologic	2010-2011	Asymptomatic Screening	12,621 total	None	NR D3/D4 subgroup Mean 59 50-69
Haas 2013[160] United States	Selenia Dimensions, Hologic	2011-2012	Asymptomatic Screening	13,158 mammography 6100 DBT	None	NR D3/D4 subgroup Mean 56 Range <40 to >70

Table 9: Quality assessment.

Study	*Interval between tests*	*Representative spectrum – consecutive patients for screening exam*	*Appropriate reference standard*	*Withdrawals*	*Design*	*Mammo*	*Quality**
MRI							
No studies							
HHUS							
Maestro 1998[147]	NR	Unclear. 19% with personal history of BC.	No, incomplete follow-up.	None reported	Unclear	Film	Poor
Buchberger 2000[140]	NR	Not reported	No, no follow-up	None reported	Unclear	Film	Poor
Kaplan 2001[141]	NR	No. Some with palpable or focal abnormal mammographic findings in other quadrants included.	Yes	6 recommended for biopsy	Prospective	Film	Fair
Kolb 2002[142]	NR	NR, but appears to be consecutive.	No, incomplete follow-up.	None reported	Unclear	Film	Poor
Crystal 2003[143]	NR	NR	No, no follow-up	None reported	Unclear	Film	Poor
Leconte 2003[144]	NR	No	No, no follow-up	None reported	Unclear	Film	Poor
Brancato 2007[149]	Within 1 month	Unclear: only 20.3% of eligible enrolled	No, no interval cancers	None reported	Unclear	Film	Poor

Study	*Interval between tests*	*Representative spectrum – consecutive patients for screening exam*	*Appropriate reference standard*	*Withdrawals*	*Design*	*Mammo*	*Quality**
De Felice 2007[150]	Same day	Yes, though no description of the participants age, family history, etc	No, no interval cancers	None reported	Prospective	Film	Poor
Corsetti 2008[145]	~50% same day ~50% within 4 weeks	Yes	No, no interval cancers	None reported	Retrospective	Film	Poor
Corsetti 2011[151]	NR	Yes	Yes	None reported	Retrospective	Film	Fair
Hooley 2012[152]	Mean 61 days Range 0-361 days	No: included BI-RADS 0 assessment on mammogram	Yes	17% did not return for one year follow-up	Retrospective	Digital	Poor
Leong 2012[153]	NR	Yes	Yes	28% of negatives with no follow-up	Prospective	Digital	Fair
Weigert 2012[154]	NR	Only 30% of eligible participated	No, not all interval cancers.	11/429 recommended for biopsy	Retrospective	NR	Poor
Girardi 2013[155]	NR	Yes	No, no interval cancers	None reported	Retrospective	Digital	Poor

Study	*Interval between tests*	*Representative spectrum – consecutive patients for screening exam*	*Appropriate reference standard*	*Withdrawals*	*Design*	*Mammo*	*Quality**
Parris 2013[156]	NR	No, abnormal mammograms included and 11% not dense	No, no interval cancers.	None reported	Retrospective	Digital	Poor
Berg 2012[163] ACRIN 6666	< 91 days	No – high-risk	Yes	<10%	Prospective	Mix film and digital	Good, but wrong population and not 100% digital
ABUS							
Kelly 2010[157]	468 women with ABUS 6 months from mammogram	No. Only 5% participation at some sites; up to 25% at others. 22% diagnostic. Data incomplete for high density subgroup.	Yes	Unclear: only 80% had mammographic follow-up > 1 year after initial imaging.	Prospective	1/3 digital, 2/3 film	Poor
Stoblen 2011[177]	Same day	Yes	No, no interval cancers	None reported	Prospective	Digital	Poor
Giuliano 2013[53]	NR	NR	Yes	None reported	Prospective	Digital	Poor

Study	*Interval between tests*	*Representative spectrum – consecutive patients for screening exam*	*Appropriate reference standard*	*Withdrawals*	*Design*	*Mammo*	*Quality**
DBT							
Ciatto 2013[159]	Same day	Yes, 95% agreed to participate.	No, no interval cancers	None reported	Prospective, consecutive	Digital	Poor
Rose 2013[161]	Same day	No, volunteer bias	No, no interval cancers	None reported	Retrospective, pre - post	Digital	Poor
Skaane 2013[55]	Same day	Yes, 70% agreed to participate	Incomplete follow-up for interval cancers	None reported	Prospective	Digital	Poor
Haas 2013[160]	Same day	Yes	No, no interval cancers	None reported	Retrospective	Digital	Poor

* **Quality rating**:

High - consecutive sample from women presenting for screening, digital mammography with a negative assessment, BI-RADS 3 or 4 density, supplemental screening test done within one month of mammogram, at least 90% follow-up of benign and negative findings at one year, and complete reporting of "positive" and "negative" results for the supplemental screening test;

Fair – same as above but film mammography and BI-RADS 2-4 density acceptable;

Poor – less than 90% follow-up or a non-consecutive sample (spectrum bias) or supplemental screening test done more than one month after the mammogram or incomplete reporting of positive and negative results for advanced imaging

The diagnostic accuracy test characteristics from these studies are summarized in Table 10 on the following page. In these studies, when one participant was diagnosed with more than one cancer or had more than one biopsy, the statistics were reported on a per participant basis rather than per cancer or biopsy. The statistics in Table 9 represent only participants who had a negative mammogram assessment and fell into one of the two high density BI-RADS categories (D3: heterogeneously dense; D4: extremely dense) except for those with separate rows for mammography and mammography plus supplemental screening.

Only four of the studies in the Tables (Hooley 2012; Leong 2012; Girardi 2012; Parris 2013) compared HHUS to digital mammography.[152,153,155,156] Ten studies compared HHUS to film mammography and one did not report the type of mammography machine used in the study. The ACRIN 6666 Trial used a mix of digital and film mammography.[68,163]

Only three of the trials (Kaplan 2001; De Felice 2007; Leong 2012) in women with dense breasts[141,150,153] and the ACRIN 6666 trial[68,163] were prospective studies. Prospective studies are more likely to have complete and consistent measurement of the key outcomes because they are defined objectively at the start of the study and collected systematically. It is worth noting in Table 10 that these trials had by far the highest recall rates (>100 recalls per 1000 examinations). Most of the other studies did not systematically report recalls after ultrasound and often reported the HHUS assessment as positive only if a biopsy was recommended, thus underestimating the true recall rate.

Table 10 on the following page summarizes the major diagnostic test results from these studies. Because the majority of the studies did not follow women with negative HHUS assessments for interval cancers, the most relevant statistics to focus on are the recall rate, the biopsy rate, the cancer detection rate, the positive predictive value of positive tests (PPV1) and the positive predictive value of biopsies performed (PPV3). As described in the background section of this report, the recall rate for mammography is typically about 100 per 1000 examinations, the biopsy rate about 10 per 1000 examinations, the cancer detection rate about 3.5 to 5 per 1000 examinations, the PPV1 about 4% and the PPV3 about 25%.

The recall rate for HHUS after normal mammography ranged from 21 to 170 per 1000 examinations with the median value across the studies of 59, lower than the typical recall rate for mammography described earlier in this report of 100 per 1,000 examinations. In the ACRIN 6666 study, HHUS recalled 186 women per 1000 examinations.[163] As noted above, all of the prospective studies (Kaplan 2001; De Felice 2007; Leong 2012; ACRIN 6666) reported recall rates greater than 100 per 1000 examinations, so these values are likely to be more accurate.[141,150,153,163]

Table 10: Test characteristics.

Study	*TP (N)*	*FP (N)*	*FN (N)*	*TN (N)*	*Sens (%)*	*Spec (%)*	*PPV1 (%)*	*NPV (%)*	*PPV3 (%)*	*Recall rate (per 1000)*	*Biopsy rate (per 1000)*	*Cancer detection rate (per 1000)*
MRI												
No studies												
HHUS												
Maestro 1998[147]	2	46	0*	302	100	86.8	4.2	100	13.3	137.1	42.9	5.7
Buchberger 2000[140]	32	241	0*	7830	100	97.0	11.7	100	9.9	33.7	39.9	3.9
Kaplan 2001[141]	5	245	0*	1612	100	86.8	2.0	100	5.4	134.3	49.4	2.7
Kolb 2002[142]	31	NR	NR	NR	-	-	-	-	10.6	-	23.9	2.5
Crystal 2003[143]	7	90	0*	1420	100	94.0	7.2	100	18.4	63.9	25.0	4.6
Leconte 2003[144]	11	NR	NR	NR	-	-	-	-	NR	NR	NR	-
Brancato 2007[149]	2	106	0*	5119	100	98.0	1.9	100	3.2	20.7	11.9	0.4
De Felice 2007[150]	12	175	0*	1567	100	90.0	6.4	100	6.4	106.6	106.6	6.4
Corsetti 2008[145]	37	412	0*	8708	100	95.5	8.2	100	5.7	49.0	9.1	4.0
Corsetti 2011[151]	32	395	8	6769	80.0	94.5	7.5	99.9	7.5	59.3	59.3	7.5
Hooley 2012[152]	3	50	0	882	100	94.6	5.7	100	5.7	56.7	56.7	3.2
Leong 2012[153]	2	22	0	117	100	84.2	8.3	100	12.5	170.2	113.5	14.2
Weigert 2012[154]	28	401	1	8217	96.6	95.3	6.5	100	6.7	49.6	48.3	3.2
Girardi 2013[155]	22	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	2.2
Parris 2013[156]	10	175	0*	5334	100	96.8	5.4	100	5.5	33.5	32.8	1.8
Berg 2012[163] R1	14	423	2	1914	87.5	81.9	3.2	99.9	6.8	185.7	88.0	5.9
ACRIN 6666 R3	9	326	14	1934	39.1	91.8	5.0	99.3	NR	85.0	NR	4.2

Study	*TP (N)*	*FP (N)*	*FN (N)*	*TN (N)*	*Sens (%)*	*Spec (%)*	*PPV1 (%)*	*NPV (%)*	*PPV3 (%)*	*Recall rate (per 1000)*	*Biopsy rate (per 1000)*	*Cancer detection rate (per 1000)*
ABUS												
Kelly 2010[157]	23	442	10	5657	69.7	92.8	4.9	99.8	30.7	75.8	12.2	3.8
Stoblen 2011[177]	0	60	0*	230	-	79.3	0	100	NR	206.9	NR	0
Giuliano 2013[53]												
- M	19	74	6	3977	76.0	98.2	20.4	99.8		22.8	NR	4.7
- M + ABUS	42	19	1	3365	97.7	99.7	80.8	99.9	15.2	15.2	15.2	12.3
DBT												
Ciatto 2013[159]												
- M	39	322	20	6913	66.1	95.5	10.8	99.7	NR	49.5	NR	5.3
- M + DBT	59	254	0*	6981	100	96.6	18.8	100	NR	42.9	NR	8.1
DBT in dense, M -	3	21	0*	1103	100	98.1	12.5	100	NR	21.3		2.7
Rose 2013[161]												
- M	56	1152	0*	12648	100	91.7	4.6	100	26.5	87.2	15.2	3.9
- M + DBT	51	467	0*	8981	100	95.1	9.8	100	24.7	54.5	10.6	5.4
Skaane 2013[55]												
- M	77	771	46	11727	62.6	93.8	9.1	99.6	NR	67.2	NR	6.1
- M + DBT	101	670	22	11828	82.1	94.6	13.1	99.8	NR	61.1	NR	8.0
Haas 2013[160]												
- M	37	0*	NR	NR	100%	NR	NR	NR	NR	120	NR	5.2
- M + DBT	35	0*	NR	NR	100%	NR	NR	NR	NR	84	NR	5.7

*0 by design. These studies do not have follow-up (or have limited follow-up) and so are unable to detect the interval cancers over the next year that represent the false negatives. Thus sensitivity and the NPV will always be 100% and the specificity will be overestimated.

Table 11: Characteristics of the screen-detected cancers.

			Mammogram					*Supplemental*		
Study	***N***	***Size, mm***	***≤ 1 cm, %***	***Lymph node negative, %***	***Stage 0 or 1, %***	***N***	***Size, mm***	***≤ 1 cm, %***	***Lymph node negative, %***	***Stage 0 or 1, %***
MRI										
No studies										
HHUS										
Maestro 1998[147]	-	-	-	-	-	2	15	0	NR	NR
Buchberger 2000[140]	142	11.2	NR	NR	NR	32	9.1	~75	NR	NR
Kaplan 2001[141]	NR	-	-	-	-	6	9	66.7	100	66.7
Kolb 2002[142]	94	-	-	-	-	31	NR	NR	NR	NR
Crystal 2003[143]	NR	13.5	NR	NR	NR	7	9.6	57.1	85.7	NR
Leconte 2003[144]	14	NR	NR	NR	NR	11	NR	NR	NR	NR
Brancato 2007[149]	5.8%	NR	NR	NR	NR	2	NR	50	NR	NR
De Felice 2007[150]	8	NR	NR	NR	NR	12	10			
Corsetti 2008[145]	166	NR	36	68	NR	37	NR	65	86	NR
Corsetti 2011[151]	20	NR	56	94	94	32	NR	84	90	90
Hooley 2012[152]	NR	-	-	-	-	3	7	100	NR	NR
Leong 2012[153]	NR	-	-	-	-	2	10	50	100	100

			Mammogram					*Supplemental*		
Study	***N***	***Size, mm***	***≤ 1 cm, %***	***Lymph node negative, %***	***Stage 0 or 1, %***	***N***	***Size, mm***	***≤ 1 cm, %***	***Lymph node negative, %***	***Stage 0 or 1, %***
Weigert 2012[154]	NR	-	-	-	-	28	19	24	9	NR
Girardi 2013[155]	NR	-	-	-	-	22	-	-	-	-
Parris 2013[156]	NR	-	-	-	-	10	9.7	NR	77	NR
Berg 2012[163] ACRIN 6666	59	NR	NR	NR	NR	32	10	NR	96	NR
ABUS										
Kelly 2010[157]	23	NR	30	NR	83	23	NR	61	NR	78
Stoblen 2011[177]	2	NR	100	100	100	0	-	-	-	-
Giuliano 2013 - M - M + ABUS	19	22.3	NR	95		42	14.3	NR	98	83
DBT										
Ciatto 2013[159]	39	13.7	NR	72	95	20	13.5	NR	80	95
Rose 2013[161]	56	16	NR	93	86	51	16	NR	88	76
Skaane 2013[55]	77	13.2	49.1	83.0	NR	30	12.8	41.3	85.2	NR
Haas 2013[160]	37	NR	NR	NR	NR	35	NR	NR	NR	NR

The biopsy rate for women having HHUS after normal mammography ranged from 12 to 114 per 1000 examinations with a median of 46. In the ACRIN 6666 study the biopsy rate was 88 per 1000 examinations the first round and about 61 per 1000 examinations the third round. The cancer detection rate varied from 0.4 to 14.2 per 1000 examinations with a median value of 3.2 per 1000 examinations. In the ACRIN 6666 trial, HHUS detected 5.9 cancers per 1000 examinations. It is not clear why the biopsy rate vary across such a wide range. Potential explanations include incomplete reporting of cyst aspirations, different thresholds for performing cyst aspirations, operator dependency in performing HHUS, and differences in the proportion of patients undergoing a first time screening HHUS compared to those with prior examinations for comparison.

There was also a wide range of estimates across the studies for the PPV1 (2.0 to 11.7%, median 6.5%) and the PPV3 (3.2 to 18.4%, median 7.1%). The heterogeneity of these results was likely due to a combination of factors. These include the study design (prospective, retrospective), the use of film or digital mammography, differences in the assessment of mammography across countries, whether a radiologist or a technician performed the HHUS, the level of experience and training of the person performing the HHUS, and differences in the populations studied (age distribution, breast cancer risk factors, time since last mammogram).

The characteristics of the cancers detected by mammography alone and of ultrasound among women with a negative mammogram are described in Table 11 on page 59. The table shows that most of the cancers detected by HHUS after negative mammography are small, node negative, early stage cancers. These are the cancers that are potentially curable by early detection before they develop into cancers with a poorer prognosis. Cancers at an early stage also require less aggressive therapy: the patient may be eligible for lumpectomy rather than mastectomy and may not require systemic chemotherapy. Thus early detection may improve both quality and quantity of life. The counter-argument is that some of these early stage cancers may not have progressed much before the next routine screening examination with mammography. Thus, they may ultimately have been detected and cured with mammographic screening alone. In addition, some proportion of these cancers may represent overdiagnosis: the identification of a cancer that would not have ever progressed to cause symptoms prior to the death of that individual woman. The identification of such cancers would lead to unnecessary labeling of the woman as someone who has cancer as well as unnecessary surgery and chemotherapy. The only way to test which of these two competing hypotheses is true would be to perform a randomized trial comparing the two approaches to breast cancer screening.

American College of Radiology Imaging Network (ACRIN) 6666 Trial

Because the ACRIN 6666 trial[68,163] was the only prospective trial performed in the United States with mostly digital mammography and one year follow-up for multiple screening rounds (high

quality), its findings are the most pertinent to the focus of this review and will therefore be described in detail below. The population studied was higher risk than that of a typical screening population, so the biopsy rate, cancer detection rate, and positive predictive values will be higher than those of a screening population. For instance, in the first round the biopsy rate based on mammography was 14.4 per 1000 examinations, the cancer detection rate was 7.5 per 1000 examinations, and the PPV3 was 31%, all of which are higher than expected for mammography in a screening population (10 per 1000, 5 per 1000, 25% respectively).

The ACRIN 6666 trial randomized 2809 high-risk women to receive both mammography (film or digital) and ultrasound in alternate order.[68] High-risk was defined by at least one of the following: a personal history of breast cancer; positive for BRCA1 or BRCA2 mutation; a lifetime risk ≥ 25%, a 5-year risk ≥ 2.5% or ≥ 1.7% with extremely dense breast tissue; prior biopsy with atypical ductal hyperplasia, atypical lobular hyperplasia, lobular carcinoma in situ or atypical papilloma; or prior mantle radiation. The study also required that the women have at least one quadrant of one breast with heterogeneously dense or extremely dense tissue on a prior mammogram. The trial did not meet the inclusion criteria for this assessment for two reasons: the study subjects are high risk rather than a general screening population and were not required to have dense breasts by BI-RADS criteria.

The women were followed for three annual cycles and upon completion of the third cycle, the women were offered additional screening with breast MRI.[163] There were 2659 women with data for analysis after the first year of follow-up. Their median age was 55 years and 93% were white. The primary risk factors for inclusion in the study were a personal history of breast cancer (53%), a lifetime risk ≥ 25% (19%), and a five-year risk ≥ 2.5% (15%). The investigators present the results of mammography alone and for the combination of mammography plus ultrasound, but not for ultrasound alone or the subgroup of women with a negative mammography assessment. When possible, we calculated the incremental results for ultrasound following negative mammography. In the first screening round, mammography detected 20 cancers (cancer detection rate 7.6 per 1000 examinations) and ultrasound detected an additional 14 cancers (5.9 per 1000 examinations).[163] There were 2 interval cancers so the sensitivity of mammography was 55.6% (20/36) and the sensitivity of ultrasound in women with negative mammograms was 87.5% (14/16). The number of recalls increased from 306 with mammography alone to 707 with mammography plus ultrasound, a 2.3-fold increase in the recall rate (from 115.1 per 1000 examinations to 265.9 per 1000). The number of breast biopsies increased from 65 to 272, a 4.2-fold increase (from 24.4 per 1000 examinations to 102.3 per 1000). The PPV3 for ultrasound in women with negative mammograms was only 6.8%.

By the third screening examination, the test characteristics changed, reflecting a reduction in prevalent cancers due to early detection, the transition to digital mammography, and improved specificity with increased experience of the radiologists and the availability of prior examinations

available for review.[163] Mammography detected 23 cancers (cancer detection rate 9.9 per 1000 examinations) and ultrasound detected an additional 9 cancers (4.2 per 1000 examinations). There were 14 interval cancers so the sensitivity of mammography was 50.0% (23/46) and the sensitivity of ultrasound in women with negative mammograms was 39.1% (9/23). The investigators did not report the recall rate and biopsy rate for round 3, but did report the numbers for the combination of rounds 2 and 3. The number of recalls increased from 453 with mammography alone to 809 with mammography plus ultrasound, a 1.8-fold increase in the recall rate (from 94.1 per 1000 examinations to 168.1 per 1000). The number of breast biopsies increased from 97 to 339, a 3.5-fold increase (from 20.1 per 1000 examinations to 70.4 per 1000). The PPV3 for ultrasound in women with negative mammograms was 7.1%.

In round 3, women were offered MRI in addition to HHUS and mammography.[163] The 612 women in the MRI sub-study had higher risk for breast cancer and were younger than those who declined participation.[178] In this group of participants, mammography alone detected 5 cancers, ultrasound detected an additional 2 cancers (sensitivity for the combination 43.8%, cancer detection rate 11.4 per 1000 examinations) and MRI detected 9 additional cancers (sensitivity 100%, incremental cancer detection rate 14.7 per 1000 examinations and combined cancer detection rate 26.1 per 1000 examinations). The 9 cancers detected by MRI only were small (median 8.5 mm) and all were lymph node negative. Both cancers seen only with HHUS (not mammography) were also diagnosed with MRI. The high cancer detection rate in the women in the MRI group reflects the high underlying risk for cancer in the women who agreed to participate in the sub-study. The recall rate was 85.0 per 1000 examinations for mammography alone, 163.4 per 1000 for the combination of mammography plus HHUS and 260.0 per 1000 for MRI. The biopsy rate was 62.1 per 1000 examinations for the combination of mammography plus HHUS and 132.3 per 1000 for the combination with MRI. The PPV3 for MRI in women with a negative mammogram was 22.4%, which is much higher than that of ultrasound.

In this high-risk population, the ACRIN 6666 study found that supplemental screening with HHUS produced a relatively high yield of cancers the first round of screening, approximately doubling the cancer detection rate, but this decreased with subsequent rounds. In order to find these cancers, the recall rate more than doubled so that one in four women (26.6%) were recalled in the first round. The number of biopsies performed increased by a factor of 4. In the first round, the combination of ultrasound plus mammography led to almost as many biopsies (10.2% of women) as women recalled with mammography alone (11.5% of women). The addition of MRI more than doubled the cancer detection rate of mammography plus ultrasound, but was associated with even an even higher recall rate and a doubling of the biopsy rate. The PPV3 for ultrasound in women with negative mammograms was very low (6.8% round 1, 7.1% rounds 2 and 3) compared to mammography alone (29.1% round 1, 38.1% rounds 2 and 3). The PPV3 for MRI in women with negative mammograms was 22.4%.

Summary: Screening HHUS of the Breast

There are no studies evaluating the impact of adding HHUS to mammographic screening among women with dense breast tissue that address the key patient-centered outcomes of breast cancer mortality and disease-free survival. The available body of evidence, focusing largely on shorter-term recall rates, biopsy rates, cancer detection rates and false positive rates, is limited by multiple factors. There were a large number of studies, but the heterogeneity of the study designs, populations, and results preclude the use of meta-analytic techniques to combine the results. The majority of the studies used film mammography, were retrospective, did not fully report the recall rate, and were not able to calculate sensitivity because women with negative mammograms were not followed for interval cancer. There is not even one prospectively designed study with one-year follow-up of HHUS in women with a negative mammogram and heterogeneously dense or extremely dense breasts. The best estimates for sensitivity and specificity come from the ACRIN 6666 trial (87.5% and 81.9% respectively) because it is the highest quality study and sensitivity and specificity are usually not influenced by the risk of the population being studied.[163] The best estimate for the incremental cancer detection rate is centered around 3-4 cancers per 1000 examinations, but the results from the three studies on the Connecticut experience were closer to 2 cancers per 1000. The results from Connecticut are more likely to be representative of routine clinical practice in the United States. The recall rates and PPV1's in these studies were greater than those of mammography, indicating that the addition of HHUS approximately more than doubles the recall rate. The recall rate doubled in the ACRIN 6666 study as well. Finally, the biopsy rates were 3-5 times higher than those of mammography, suggesting that the biopsy rate of ultrasound after negative mammography is likely to be at least four times that of mammography alone. This is the major limitation of screening ultrasound. The PPV3, which represents the percentage of biopsies that are positive for cancer, was only 7% in studies of women with dense breasts and in the high risk population in the ACRIN 6666 study. The PPV3 in mammography is approximately 25%. Thus the rate of false positive biopsies is much higher with ultrasound. Table 12 on the following page summarizes the key statistics from the three Connecticut studies (direct evidence) and the ACRIN 6666 study (high quality indirect evidence).

Table 12: Key findings from the essential studies of HHUS.

Study	*Recall rate per 1000*	*Biopsy rate per 1000*	*PPV3*	*Cancer detection rate per 1000*
Hooley 2012[152]	56.7	56.7	5.7%	3.2
Weigert 2012[154]	49.6	48.3	6.7%	3.2
Parris 2013[156]	33.5	32.8	5.5%	1.8
ACRIN 6666[163]	185.7	88.0	6.8%	5.9

The studies comparing mammography, ultrasound, and MRI in very high-risk women described in the section on MRI help with the comparative effectiveness of the three technologies. In the six studies evaluating all three technologies[62,64,66,163,167,176], mammography detected 48 cancers, ultrasound detected 53, and MRI detected 116. Ultrasound detected 19 cancers that were not detected by mammography, which represents 40% (19/48) more cancers detected. Four cancers (3%) were detected only on ultrasound. These studies suggest that the addition of HHUS would increase the cancer detection rate by about 40% (best estimate: 1.4-fold increase) more than mammography alone, but that HHUS does not increase the cancer detection rate when added to mammography plus MRI.

There are no large, well done studies in the United States that directly measure these statistics, which could serve as a reasonable estimate. There is also uncertainty about whether the early detection of these cancers by ultrasound will improve outcomes for women compared to outcomes following their detection as a lump by the women before her next mammogram (interval cancers) or when she has her next screening mammogram.

Estimates based on these data are shown in Table 13 on the following page. There is a low level of uncertainty around the PPV3 because it was fairly consistent in the literature. The cancer detection rate comes primarily from the three studies describing the experience in Connecticut. There is high uncertainty about the recall rate because of the lack of direct evidence from studies of HHUS in women with dense breast tissue and because of the heterogeneity of the findings in the studies. In all of the prospective studies the recall rate was greater than 100 per 1000 examinations.

Table 13: Estimated incremental yield of HHUS after negative digital mammography in women with dense breast tissue.

Statistic	*Digital mammography*	*Incremental yield with HHUS*	*Uncertainty*
Recall rate per 1000	128	98	High
Biopsy rate per 1000	17.8	49	Low-moderate
CDR per 1000	4.2	2-3	Low
PPV3	24%	7%	Low

These estimates suggest that HHUS would find 2-3 more cancers than those found by digital mammography alone, with a PPV3 of 7%. There would be approximately 98 additional recalls and 49 additional biopsies in order to identify these cancers.

7.3 Automated Whole Breast Ultrasound (ABUS)

Three studies[53,157,177] of ABUS evaluating approximately 9000 participants met the inclusion criteria for the assessment. The primary data are summarized in Tables 7 through 10 above. They will be described in chronological order.

Kelly and colleagues[157] recruited women from eight facilities across the United States. The investigators offered ABUS to consecutive asymptomatic women who had dense breasts. The radiologist reading the mammogram was blinded to the ABUS results and the radiologist reading the ABUS was blinded to the mammography results. Women whose compressed breast thickness at mammography was greater than 7 cm were excluded because of the limited sensitivity of ultrasound at that depth. The percentage of patients who agreed to participate at each site varied from 5% to 25%. The investigators performed 6425 ABUS examinations in 4419 women. 1434 of the examinations were diagnostic examinations because the women had a history of prior breast cancer (776/1434, 54%), breast implants (399/1434, 28%), or non-localized abnormalities such as diffuse tenderness or nodularity (159/1434, 11%). One third of the mammograms were digital and two-thirds were film. Some women at high risk preferred to alternate mammography and ABUS examinations at six-month intervals. The study followed women for one year for interval cancers. The percentage with complete follow-up was not reported, but 5089 of the women (80%) had a repeat mammogram at least one year after the original mammogram.

The ABUS examination took 5 to 10 minutes preparation time and 10 to 20 minutes for the examination.[157] The interpretation and reporting time for the radiologist was 7 to 10 minutes. The

study sample had a median age of 53 year, but included women as young as age 24 and as old as 89 years.[157] The sample included women with a personal history of breast cancer (10%), at least one first-degree relative with breast cancer (30%), and at least one second-degree relative with breast cancer (29%). These proportions are higher than in a typical screening population, suggesting that women who enrolled in the study were at higher risk for breast cancer than the general population. During the study 23 breast cancers were detected with mammography, 23 by ABUS in women with negative mammograms, and an addition 11 presented as interval cancers that were not detected by either modality.[157] One woman was diagnosed with bilateral breast cancer (56 participants diagnosed with breast cancer). The results are not presented separately for women with negative mammograms, but the statistics for ABUS after a negative mammogram can be calculated from the data presented in the tables and the results section (see Table 9 on page 52). The sensitivity of mammography plus ABUS was higher than that of mammography alone (67.6% compared to 41.1%), but the recall rate for mammography plus ABUS was almost double that of mammography alone (74.8 per 1000 compared to 32.4 per 1000). The biopsy rate was also higher with mammography plus ABUS (12.2 compared to 9.1 per 1000). The cancer detection rates were similar (3.8 compared to 3.6 per 1000).

There are many methodological concerns that limit the ability to generalize the results of this study[157] to women with dense breasts and a negative digital mammogram. The low volunteer rate (5%-25%) in this study raises concerns about spectrum bias – those who agree to participate may differ from those who do not participate in ways that impact the study results. For instance, women at higher risk for breast cancer may be more likely to volunteer for this study of additional imaging. The wide age range (down to age 27 years) also suggests that this was not a typical screening population. Two-thirds of the mammograms were film, which has much lower sensitivity than digital mammography in women with dense breasts – some of the cancer identified on ABUS would have been picked up by digital mammography.[112,179] In addition, some of the women elected to be screened with ABUS six months after the mammogram, so the cancers identified by ABUS may represent a mix of interval cancers and those missed by mammography. All of these biases would tend to increase the cancer yield of ABUS.

In the second, much smaller study, Stoblen and colleagues[177] described the results of ABUS in 304 consecutive women between the ages of 50 and 69 who were seen for routine screening mammography in Germany. The majority of the women had non-dense breasts (scattered fibroglandular densities). All subjects had digital mammography followed by ABUS. Two cases of DCIS were detected by mammography, neither of which was detected by ultrasound. The investigators reported 60 false positive assessments by ABUS (20.7% of negative mammograms) compared to 12 (4.0%) for digital mammography in the same women. Thus the false positive rate for ABUS was 207 per 1000 examinations compared to 40 per 1000 for mammography. However, it does not appear that all of the positive ultrasound findings were biopsied. In addition, no follow-up was reported other than for 2 patients with repeat examinations at 6 months, with no additional

cancers identified. The study is small and does not directly apply to women with dense breasts, but it highlights the concern about high numbers of false positive results with either automated or hand held ultrasound.

Finally, Giuliano and Giuliano[53] report on the performance of digital mammography plus ABUS in 3418 asymptomatic women with dense breasts and compared it to 4076 asymptomatic women with dense breasts screened with digital mammography in the prior year. It is unclear if consecutive women were included. The BI-RADS categories were not used to define high density, but it is likely that the women studied (mammograms with "a Wolfe classification of 50% or greater") were similar to the two high-density BI-RADS groups. The study excluded women with major risk factors for breast cancer including those with a personal or family history of breast cancer and those with a BRCA mutation. The study was performed at a single site in Florida. Two radiologists read each of mammograms and the ABUS images with final readings by consensus. There was no blinding of the radiologists, but the investigators blinded the pathologists evaluating biopsy specimens.

In the control group, the sensitivity and specificity of digital mammography alone were 76.0% and 98.2%. [53] The recall rate was 22.8 per 1000 examinations and the cancer detection rate was 4.7 per 1000. The biopsy rate was not reported. This cancer detection rate is relatively high for invasive cancer (no cases of DCIS were reported) in women with no personal or family history of breast cancer. The PPV reported is quite high for mammography, suggesting that it may be the PPV for biopsy rather than the PPV for a positive mammography assessment. The low recall rate also supports the under-reporting of recalls for positive results. In the mammography plus ABUS group, the sensitivity was 97.7% and the specificity was 99.7%. Again, the recall rate and biopsy rate are not clearly reported and both are calculated at 15.2 per 1000 examinations, suggesting that this is actually the biopsy rate. It is likely that the true recall rate was much higher and that the specificity and PPV are much lower than reported in the paper.

There are several other concerns about this study that call into question all of its results.[53] First, there were no reported cases of DCIS. In 2011, when the study was conducted, approximately 27% of all breast cancer diagnoses were DCIS.[180] Since the study reported 68 invasive breast cancers, there should have been an additional 25 cases of DCIS. Mammography is more sensitive than ultrasound at the detection of DCIS[52], so the exclusion of DCIS from the results could have a large impact on the results. It is also worrisome that the results for mammography alone were not reported for the cohort of women also examined with ABUS. It may be that digital mammography performed better in that group because the radiologists had one more year of experience with this relatively new technology. It is also remarkable that the specificity of ABUS was so high. All other reports of ultrasound consistently find a high rate of false positive studies with ultrasound with PPV1 and PPV3 being consistently lower than that of mammography. The opposite was reported in this study. The demographic characteristics of the two groups were not presented, nor compared –

if they were very different, then there should have been some adjustment for these differences. The results suggest that there were large differences: the average age for detected invasive cancers in the control group was 54 years, while in the ABUS group it was 57. If ABUS identifies cancers earlier than mammography, then the average age of detected cancers should go down, not up. Given these major concerns, as well as the non-standard breast density measurements and the lack of reporting of the results of ABUS among the women with normal mammograms, the results of this study are not useful in evaluating the appropriate role for ABUS in women with dense breasts.

Summary: Screening ABUS

None of the studies directly address the use of ABUS following negative digital mammography in a screening population of women with dense breasts. All three of the studies are of poor quality. The study of Kelly and colleagues offers the only reasonable estimates (sensitivity 67.6%, specificity 92.9%, recall rate 74.8 per 1000, biopsy rate 12.1 per 1000, cancer detection rate 3.7 per 1000), but the validity and relevance of these data are limited by concerns about spectrum bias and the use of film rather than digital mammography, which would decrease the cancer detection rate of mammography in the population of women with dense breasts. In addition, in an unreported proportion of the women, ABUS was performed between 5 and 8 months after mammography and likely found some cancers that it would not have identified at the time of mammography. Across the three studies, the recall rate varied from 5 to 207 per 1000 examinations, the biopsy rate was not reported or up to 15 per 1000 examinations, the PPV3 from not reported to 31% and the cancer detection rate ranged from 0 to 7.6 per 1000 examinations. Overall, the paucity of studies, the lack of high quality studies, and the wide range of estimates across the three studies mean that there is considerable uncertainty surrounding all of the estimates for the diagnostic test statistics for ABUS. Because of the uncertainty described above, we felt that the most reliable estimates for the test characteristics for ABUS come from the HHUS literature, but with high uncertainty. Estimates based on these data are shown in Table 14 below and are identical to those for HHUS.

Table 14: Estimated incremental yield of ABUS after negative digital mammography in women with dense breast tissue.

Statistic	***Digital mammography***	***Incremental yield with ABUS***	***Uncertainty***
Recall rate per 1000	128	98	High
Biopsy rate per 1000	17.8	49	High
CDR per 1000	4.2	2-3	High
PPV3	24%	7%	High

These estimates suggest that ABUS would find 2-3 more cancers than those found by digital mammography alone with a PPV3 of 7%. There would be approximately 98 additional recalls and 49 additional biopsies in order to identify these cancers.

7.4 Digital Breast Tomosynthesis (DBT)

Four studies, all published in 2013[55,159-162] evaluated the use of DBT in more than 34,000 participants. All four studies evaluated DBT in screening populations not restricted to women with dense breasts. The primary data are summarized in Tables 7 through 10 above. The first two studies (Skaane 2013; Ciatto 2013) compared the test characteristic of digital mammography and DBT performed in the same patients on the same day. The second two studies (Rose 2013; Haas 2013) compare two groups of patients, one screened with digital mammography alone and the other with digital mammography plus DBT. The study by Ciatto and colleagues[159] is the only study to report the results for DBT among women with dense breasts and negative mammography assessments.

Skaane and colleagues recently published initial results from a very large series of patients evaluated with both digital mammography and DBT performed on the same day.[55,162] The study evaluated DBT in 12,621 women coming in for routine screening mammography in Oslo, Norway in 2011. DBT added an average of 10 seconds per view to the time required for mammography (40 seconds total). The reading time increased from 45 seconds for mammography to 91 seconds for mammography plus DBT. The total radiation dose increased from 1.58 mGy for digital mammography to 1.95 mGy for DBT.

According to standard practice in Norway, two radiologists independently interpreted the images for each woman and the potentially positive cases were reviewed at an arbitration meeting. Follow-up is not complete, but 3 interval cancers were identified during 9 months of follow-up. These were not included in the statistics reported in the paper, but they have been counted as false negatives in the calculations performed for this assessment.

The study compared DBT to digital mammography in all women, with only limited data presented on the subset of women with dense breasts. DBT decreased false positives and false negatives. Thus DBT had higher sensitivity (82.1% compared to 62.6%, $p<0.001$) and specificity (94.6% compared to 93.8%, $p<0.001$) than digital mammography alone. The cancer detection rate increased from 6.1 to 8.0 cases per 1000 ($p=0.001$) examinations while the recall rate decreased from 67.2 per 1000 to 61.1 per 1000 ($p<0.001$). The adjusted increase in cancer detection was 40% (RR 1.40, 95% CI 1.13 to 1.71, $p<0.001$). The improvement in cancer detection was similar in women with non-dense and dense BI-RADS categories.

There are several issues that make it difficult to generalize the results of this study[55,162] to the United States. The standard of care in Norway is to have two radiologists interpret each mammogram and to have an arbitration meeting to review all positive results and decide which to call back. As noted earlier, this approach has a much lower call back rate than that observed in the United States.[115] This study evaluated DBT in all women and not just those with negative digital mammograms and dense breasts. Finally, follow-up for interval cancers was incomplete.

Ciatto and colleagues published a similar study comparing digital mammography to DBT in 7292 women coming in for routine screening mammography in Italy.[159] As in Norway, two radiologists independently interpreted the images for each woman. However, if either was positive, the woman was recalled in this study while in Norway, there was a conference to decide who should be recalled. As in the prior study, there was no follow-up, so the primary outcomes were the cancer detection rate and the recall rate. Ciatto and colleagues did publish detailed results for the two highest BI-RADS density categories.

Similar to the Norwegian study, the Italians found that compared to digital mammography, DBT had greater sensitivity (100% compared to 66.1%) and greater specificity (96.5% compared to 95.5%).[159] This translated into a higher cancer detection rate (8.1 compared to 5.3 per 1000 examinations) with a lower recall rate (42.9 compared to 49.5 per 1000 examinations). These same improvements were found in the subgroup of participants with dense breasts (sensitivity 100% compared to 62.5%; specificity 94.0% compared to 93.1%; cancer detection rate 6.6 compared to 4.1 per 1000 examinations; recall rate 65.8 compared to 72.4 per 1000 examinations).

The Ciatto 2013 publication was the only study that presented sufficient data to calculate the results of supplemental screening with DBT in women with dense breasts and a normal mammogram. In the subgroup of women with dense breasts and negative digital mammography, the sensitivity of DBT was 100%, the specificity was 98.1%, the cancer detection rate was 2.7 per 1000 examinations and the recall rate was 21.3 per 1000 examinations.[159] As in the prior study, the investigators reported that the improvement in the cancer detection rate was similar in both the dense and non-dense groups.

The primary concern with this study[159] is the lack of follow-up for interval cancers. This artificially raises the sensitivity of DBT to 100% and causes an overestimation of the specificity and negative predictive value as well. The investigators also did not report the biopsy rate in the study.
The next study on DBT[161] (Rose 2013) used a pre-post design rather than a direct comparison of the two technologies. The investigators compared the screening benchmarks for the combination of DBT and digital mammography (n=9499) to those of digital mammography alone at the same sites in Texas during the prior year (n=13,856). There was no follow-up for interval cancers, so the sensitivity, specificity and negative predictive value are overestimated. As in the prior studies, DBT had a higher cancer detection rate (5.4 compared to 4.0 per 1000 examinations) with a lower recall

rate (54.5 compared to 87.2 per 1000 examinations) and a lower biopsy rate (10.6 compared to 15.2 per 1000 examinations). In the subgroup of women with BI-RADS 3 or 4 density (summarized in Table 9 on page 52) the findings were similar. DBT had a higher cancer detection rate (5.4 compared to 3.9 per 1000 examinations) and a lower recall rate (68.8 compared to 102.8 per 1000 examinations). The biopsy rate was not reported by density.

A fourth study conducted at four sites in the United States was released online by Haas and colleagues on July 30, 2013.[160] They compared the recall rate and cancer detection rate at sites using digital mammography (n=7058) to those at sites using DBT (n=6100). All women presenting for screening mammography were included except those with breast implants or those with large breasts requiring tiled images. As in the prior studies, DBT decreased the recall rate (84 compared to 128 per 1000 examinations, p<0.01) and had a trend towards an increase in the cancer detection rate (5.7 compared to 5.2 per 1000 examinations, p=0.70). The difference in recall rate was greatest in the dense breast subgroups: among women with heterogeneously dense breasts (n=4242) the recall rate was 102 per 1000 examinations in the DBT group compared to 167 per 1000 in the mammography group (p<0.01). Among women with extremely dense breasts (n=555) the recall rate was 67 per 1000 examinations in the DBT group compared to 156 per 1000 in the mammography group (p<0.01). The cancer detection rate was not reported by breast density subgroups.

This study[160] has several major methodological concerns. First, the mammography and DBT groups were not well matched. Women in the DBT group were younger (55.8 years compared to 57.5 year, p NR), had more extremely dense breasts (5.6% versus 3.0%, p NR) and less fatty breasts (8.8% versus 13.8%, p NR), were more likely to have a personal history of breast cancer (5.5% versus 2.8%, p NR), and were more likely to have a first-degree relative with breast cancer (18.8% versus 15.9%, p NR). The investigators did not adjust for these differences in their primary analyses, but did present the results of logistic regression analyses adjusted for age, breast density, family history and personal history of breast cancer. In those analyses DBT was associated with a 35% reduction in the odds of recall (p<0.0001). Again the reduction was greatest in those with heterogeneously dense breasts (45% reduction, 95% CI 34% to 54%, p<0.001) and those with extremely dense breasts (60% reduction, 95% CI 30% to 78%, p=0.002). The investigators did not report the biopsy rates, so it is not possible to determine whether the reduction in the recall rate translated into a similar reduction in breast biopsies. They also did not report the cancer detection rates in the density subgroups, so it is not clear whether the trend towards increased cancer detection applies to the high-density subgroup. Finally, there was no follow-up for interval cancers so the sensitivity, negative predictive value and specificity cannot be calculated.

Summary: Screening DBT

Four studies[55,159,161,162] performed in over 50,000 women (34,000 with DBT) presenting for routine screening for breast cancer found that DBT increased the cancer detection rate relative to mammography while decreasing the recall rate and the biopsy rate. The results were consistent despite the different methods for interpretation used in the three different countries (two readers with arbitration conference, two readers with call back if either is positive, one reader only). The same increase in cancer detection and decrease in recall rate was seen in all studies and in both the dense and non-dense subgroups in the studies reporting those subgroups. The Italian study[159] (Ciatto 2013) is the only publication that allowed for the calculation of statistics of interest to this assessment: women with dense breasts by BI-RADS classification who have a negative digital mammogram. In that subgroup, DBT identified an additional 2.7 cancers per 1000 examinations with a recall rate of 21.3 per 1000 examinations. The estimates for sensitivity and specificity were 100% and 98.1% respectively, but these are likely overestimates because of the lack of follow-up for interval cancers. The PPV1 was 12.5%, which is more than double the PPV1 for digital mammography in dense breasts in the study and more than twice the PPV1 usually reported for digital mammography in the United States.[75,112]

There are several limitations to these results. First, DBT has been studied primarily in combination with digital mammography in all women coming in for breast cancer screening, not as supplemental screening for women with dense breasts. In addition, the data from the two US studies (Rose 2013; Haas 2013) have significant uncertainty because they studied the results in two different populations that may not be directly comparable. As described above, only one study[159] (Ciatto 2013) presented detailed results for the population relevant to this assessment: women with dense breasts and negative mammograms, and although these results were very promising, it is only one study and it was done outside of the United States.

Estimates based on these data are shown in Table 15 on the following page. There is a low-to-moderate level of uncertainty around the PPV3; while this measure was reported in only one of the four available studies (Rose, 2013), rates were comparable between DBT and digital mammography and there is no reason to expect that DBT would be inferior given that digital mammography is a component of DBT. There is greater uncertainty about rates of cancer detection and biopsy rate, as only one of the studies included results from the target population (I.e., women with dense breast tissue and negative mammography), and there were issues of study heterogeneity as well as comparability of screening populations. The greatest level of uncertainty is with recall rates, since the most rigorous studies come from outside the US where patterns of recall differ markedly.

Table 15: Estimated incremental yield of DBT after negative digital mammography in women with dense breast tissue.

Statistic	*Digital mammography*	*Incremental yield with DBT*	*Uncertainty*
Recall rate per 1000	128	20	Moderate-high
Biopsy rate per 1000	17.8	5	Moderate
CDR per 1000	4.2	1-3	Moderate
PPV3	24%	25%	Low-moderate

These estimates suggest that DBT would find 1-3 more cancers than those found by digital mammography alone with a PPV3 of 25%. There would be approximately 20 additional recalls and 5 additional biopsies in order to identify these cancers.

7.5 Summary

Mammography is the only screening test that has been shown to reduce breast cancer mortality in randomized trials.[8-12,14,18-21,25,26] However, it is not perfect. At best, the sensitivity of mammography, including digital mammography, is approximately 80%.[74,75,112] Thus for every four to five breast cancers detected on mammography, an additional interval breast cancer will be diagnosed prior to the next screening mammogram. Furthermore, to diagnose those cancers, many women will be recalled for additional imaging because of false positive assessments, and some of those women will undergo breast biopsy. Using current digital mammographic techniques in the United States, it can be estimated that for every 1000 women having a screening mammogram, approximately 100 will be recalled for additional tests, 10 will have a breast biopsy, 5 will be diagnosed with breast cancer, and 1 additional cancer will be diagnosed in the subsequent year.[74,75,112] The false positive mammography results lead to additional time lost for the women who must schedule time to come in for additional tests and adds cost to the medical system. The women may also experience unnecessary anxiety about a cancer diagnosis.

Radiologists have long known that areas of density in the breast can obscure breast cancers on film mammography leading to a false negative assessment (decreased sensitivity). Across the four categories of breast density, the sensitivity of film mammography decreases from about 85% for women in the two lowest density categories to approximately 80% for women with heterogeneously dense breast tissue, and 65% for women with extremely dense breasts.[111,112] This masking effect of breast density is one of the primary reasons that state legislatures have passed laws requiring that women be notified about their breast density if they are in one of the high density categories.

Over the past decade, however, film mammography has been replaced by digital mammography. Digital mammography has a higher dynamic range than film and greater contrast resolution allowing the display of more gradations of density when a radiologist views the image on a computer screen. One of the strengths of digital mammography is improved sensitivity for breast cancer in dense breast tissue. In the DMIST trial, which assessed women with both film and digital mammography, the sensitivity of mammography in the two high-density categories was 55% for film mammography but 70% for digital.[75] In the BCSC, a large registry of woman screened for breast cancer, the sensitivity of digital mammography was approximately 80% to 85% across all four breast density categories, with no trend towards a decrease in sensitivity with increasing breast density.[112] Thus, the risk of masking has been dramatically reduced by the widespread adoption of digital mammography.

Nonetheless, even without masking, approximately 1 in 5 cancers can still be missed by digital mammography, raising questions about the potential for benefits of additional screening, especially among women at highest risk for breast cancer. The available literature consistently has shown that all four of the advanced imaging technologies evaluated in this assessment can detect additional breast cancers in women with negative mammograms. The most convincing data on cancer detection rates come from the ACRIN 6666 trial[163]: in the third round of screening, the combination of mammography and HHUS detected 7 additional cancers (cancer detection rate 11.4 per 1000), and MRI detected an additional 9 cancers (incremental cancer detection rate 18.2 per 1000). However, the addition of MRI increased the number or recalls from 100 to 159 and the number of recommended biopsies from 38 to 81. The PPV3 of MRI in women with negative mammograms (22%) was much higher than that of HHUS in women with negative mammograms (7%). Thus the yield per biopsy of MRI was higher than HHUS. If the costs and logistics of the two were identical, MRI would be preferred as it has greater cancer detection with fewer harms from false positive biopsies. These results are similar to earlier studies that compared mammography, HHUS, and MRI in women at very high risk for breast cancer.[62,64,66,167,176] If the goal is to maximize the cancer detection rate without worrying about false positive results due to a high recall rate, then MRI is clearly the best choice. However, there is little direct evidence about the utility of MRI in the population that is the focus of this assessment: women with dense breasts and a negative mammography assessment. MRI also requires an IV, carries the risk of complications from the injection of the contrast agent, and is the most time-consuming and expensive option.

DBT, on the other hand, decreased the recall rate in the four studies considered in this assessment, particularly in women with high breast density.[55,159-162] At the same time, DBT increased the cancer detection rate by about 2 per 1000 examinations compared to digital mammography alone. One of the studies also reported that the biopsy rate decrease from 15.2 to 10.6 per 1000 examinations.[161] In the subgroup of women with dense breasts and negative mammograms, DBT identified an additional 2.7 cancers per 1000 examinations with a recall rate of 21.3 per 1000 examinations. This is an equivalent cancer detection rate to HHUS with a much lower recall rate. DBT has the

advantage of being easy to incorporate into routine mammography screening, requiring little extra time from the woman being screened.[55] However, it uses additional ionizing radiation (about the same amount again as digital mammography).[54,55] There are also technical aspects that are still under development, such as accurate biopsy techniques for abnormalities identified on DBT, but not visible on the digital mammogram.[70]

The incremental cancer detection rate of adding HHUS to mammography is likely to fall somewhere in between DBT and MRI, although there is considerable uncertainty in the data for all three technologies. The incremental cancer detection rate is about 3 per 1000 examinations for HHUS vs. mammography alone, while it is about 2 to 3 per 1000 examinations for DBT. However, DBT has a much lower recall rate and biopsy rate while HHUS markedly increases the recall rate and biopsy rate. There are far more studies on HHUS than the other technologies, but the study results vary dramatically, which introduces considerable uncertainty into the estimates of the potential impact of supplementary HHUS for women with dense breast tissue. HHUS approximately doubles the recall rate of mammography alone and quadruples the biopsy rate. HHUS has the advantage of being readily available at most breast imaging centers and not utilizing ionizing radiation. However HHUS requires substantial training and experience of the technicians and radiologists to guarantee high quality results and it involves a substantial investment in radiologists' time.

Finally, there is much less data on screening ABUS.[157,177,53] The incremental cancer detection rates ranged from 0 to 7.3 per 1000 examinations. One of the studies reported a reduction in the recall rate with ABUS[53], but the other two had substantial recall rates that were equivalent to those seen with ultrasound.[157,177] Two of the studies[157,53] have a low biopsy rate and a high PPV3 suggesting that very few women are inappropriately being referred for biopsy. ABUS also has the advantage of little operator dependency, which addresses one of the major concerns with HHUS.

Table 16 on the following page summarizes the estimates for each of the four technologies among women with dense breast tissue based on the clinical data published through mid-2013. Many of the estimates have a high degree of uncertainty and will likely change as more high quality data become available. However, they provide reasonable estimates of the clinical benefits and harms relative to each other.

Table 16: Summary of the key statistics for four supplemental screening technologies in women with dense breast tissue.

Statistic	*DM*	*MRI*	*HHUS/ABUS*	*DBT*
Recall rate	128	100	98	20
Biopsy rate	17.8	17-36	49	5
CDR	4.2	8	2-3	1-3
PPV3	24%	22%-48%	7%	25%

Table 16 highlights the low PPV3 of ultrasound compared to the other technologies, which translates into a large number of unnecessary biopsies for every cancer detected by ultrasound. The table also clearly illustrates that DBT has a much lower recall rate than the other technologies and that MRI detects the greatest number of additional breast cancers.

Thus, we know with a high degree of certainty that all forms of supplemental screening find additional breast cancers. Most of the cancers are small, lymph node negative, and thus are potentially curable. MRI finds the most cancers and DBT has the fewest false positives. HHUS results in the largest number of false positive biopsies.

The major unanswered question is whether the identification of additional cancers through supplemental screening improves outcomes for women. Some advocates of supplemental screening will argue that the majority of the cancers identified through supplemental screening are early stage cancers with an excellent prognosis following treatment. These represent the spectrum of cancers identified with mammography that led to the reduction in mortality seen with the randomized trials of screening mammography. In their view, there can be no question that patient outcomes will be improved with supplemental screening. Others will argue that many of these supplemental screen-detected cancers would have been cured when detected on physical exam or subsequent screening mammograms and that some of these cancers represent overdiagnosis, which leads to net harm for the patient. They will highlight the growing evidence for significant overdiagnosis with mammography alone. These individuals will suggest that much of the incremental cancer detection rate with HHUS (2 to 6 or more per 1000 examinations), which is much higher than the expected interval cancer rate (1 per 1000 examinations), can only represent overdiagnosis. Only large randomized trials can definitively answer this question.

8. New Evidence Following CTAF Review

8.1 Updated Search

We conducted an updated literature search of Pubmed, Embase, Cochrane and DARE, based on the search criteria employed in the CTAF review. The search timeframe spanned from April 11, 2013 to October 18, 2013, with 150 records identified. The specific timeframe reflected the gap in current literature between the CTAF report and this review. The majority of these initial results (n=108) were excluded due to inappropriate patient populations (e.g., women at high risk without assessment of breast density such as: BRCA carriers, lifetime risk > 25%, personal history of breast cancer and, recent diagnosis of breast cancer) or no outcomes of interest (e.g., inter-reader reliability, diagnosis of suspicious or palpable masses on mammography/clinical breast examinations, description of ongoing trials without results).

Any citations already considered in the CTAF report were removed. Following removal of duplicate citations and initial evaluation, full-text review was performed on six retrieved articles. A single study was found to be relevant to the scope of this review, and is discussed in detail below.

A non-randomized, single-center prospective study evaluating the cost and performance of supplemental HHUS was identified.[181] Venturini et al. recruited 40-49 year old Italian women (n=1666) to undergo screening mammography. A screening program was tailored based on each woman's lifetime risk and mammographic density. Women (n=800) with an intermediate risk of developing breast cancer (i.e., negative mammograms and BI-RADS 3 or 4 breast density as well as estimated lifetime breast cancer risk <25%) received additional screening with HHUS, as did those with breast implants (n=26). (Note: a separate cohort of women at high breast cancer risk due to genetic susceptibility or other risk factors underwent supplemental screening with HHUS and MRI irrespective of breast density; results are not presented here.)

The incremental cancer detection rate (CDR) of HHUS was 2.4 cancers per 1000 screenings, contributing to 14% of the overall CDR for mammography + HHUS. However, as with many of the other studies evaluated in this review, only patients with suspicious lesions on mammography and/or HHUS were followed after screening, so interval cancer rates could not be compared between mammography alone vs. mammography + HHUS.

The estimated cost per additional cancer detected with HHUS, including the costs of all screening tests as well as aspiration or biopsy for suspicious lesions, was €19,158 ($26,409 USD).

9. Model of Clinical and Economic Outcomes of Supplemental Screening in Women with Dense Breast Tissue

As noted in this review, published evidence on the clinical effects of supplemental screening in women with dense breast tissue is quite limited; in fact, the target population of interest (women with BI-RADS 3 or 4 breast density and a negative mammogram) has been used extensively only to evaluate HHUS. We developed a cohort model to address this gap, focusing on the clinical and economic outcomes of supplemental screening in women with dense breast tissue and a negative mammogram. Supplemental modalities considered included HHUS, ABUS, MRI, and DBT. While DBT's eventual use may be as a first-line screening test in all women and it is not yet widely available, it nevertheless represents an additional supplemental screening option for women with dense breast tissue that clinicians may wish to consider.

Information on the economic impact of any screening strategy for women with dense breast tissue is also quite limited. We nevertheless summarize the published evidence relevant to the scope of this review in the section below.

9.1 Prior Published Evidence on Costs and Cost-Effectiveness

Available evidence on the costs and cost-effectiveness of screening modalities in women with dense breast tissue is limited to four published studies. Tosteson and colleagues used modeling to evaluate the cost-effectiveness of different screening strategies using digital mammography vs. film mammography for U.S. women aged 40 and older, using data from the DMIST trial and other sources. [182] A strategy of targeted digital mammography (i.e., either for women age <50 and/or women of any age with dense breasts, film in all other women) was estimated to produce more screen-detected cases of cancer and fewer cancer-related deaths than either an all-film or all-digital strategy. Estimates of cost-effectiveness were $26,500 per quality-adjusted life year (QALY) gained for age-targeted digital mammography vs. all-film mammography and $84,500 per QALY for age- and density-targeted digital vs. all-film. A density-targeted digital strategy focused on the Medicare population (age ≥65) yielded cost-effectiveness estimates ranging from $97,000 - $257,000 per QALY gained vs. all-film, depending on assumptions regarding the test performance of digital vs. film mammography.

A second study reported actual cancer detection and costs from a series of 5,227 asymptomatic Italian women with dense breast tissue and negative mammograms who had HHUS within one month of film mammography.[149] Costs included those of HHUS, clinical examination, biopsy, and cytology, and totaled €56 ($77) per HHUS-screened woman. HHUS detected two additional cancers in this cohort (0.4 per 1,000), resulting in a cost per additional cancer detected estimate of €146,497 ($200,701). The authors hypothesize that the cancer detection rate observed in this study, which was much lower than that reported in the HHUS studies summarized in this review (range: 1.8 – 14.2 per 1,000), may have been a result of self-selection. The sample was limited to women who presented for HHUS within one month of negative mammography, which represented approximately 20% of all women screened at the study site who had dense breasts and negative mammograms. In addition, 72% of women in the study sample were age <50, which is not reflective of the age distribution of women in the general screening population or of the subset with dense breast tissue.

Data are also available from two of the three cohort studies reporting ultrasound experience following the passage of Connecticut's breast density legislation.[152,154] Hooley and colleagues estimated the cost of providing breast ultrasound to a cohort of 935 mammographically-negative women with dense breast tissue who were screened with HHUS at Yale-New Haven Hospital after passage of the law.[152,182] The incremental cancer detection rate was 3.2 per 1,000 screened. Costs, including those of HHUS, aspiration, and biopsy, totaled approximately $180,000 for the cohort, or $60,000 per additional case of cancer detected.

A larger retrospective study of HHUS screening in nearly 9,000 women with dense breast tissue and negative screened mammograms was conducted in 6 radiology practices in Connecticut for the year after passage of the breast density legislation.[154] The cancer detection rate was also 3.2 per 1,000 screened in this study. Costs were estimated for screening and biopsy based on billed charges to insurers, and totaled approximately $3.1 million, or $110,000 per additional cancer detected. Neither study compared screening costs after passage of the law to costs incurred before the law was passed.

9.2 Overview of the Cohort Model

As described above, the published literature on the clinical and economic impact of supplemental breast cancer screening modalities in women with dense breast tissue is noticeably limited. We therefore developed a cohort model to perform a population-based, one-year analysis of clinical and economic outcomes specific to New England. In the model we included all women age 40-74 except for those with certain risk factors (see "Target Population" below).

As this review has highlighted, the performance of digital mammography differs from that of film mammography among women with heterogeneously dense and extremely dense breast tissue. We first conducted baseline analyses comparing the screening performance and costs for both digital and film mammography for all women undergoing screening. Then, we used the model to compare the performance and costs of supplemental screening with each of the modalities of interest (i.e., HHUS, ABUS, MRI, and DBT) for women in BI-RADS density categories 3 or 4 who had an initial negative mammogram. For these analyses of supplemental screening, digital mammography was assumed for initial screening, as evidence indicates it is the current screening standard.

We defined the supplemental screening population as a hypothetical cohort that was stratified into different levels of underlying breast cancer risk. Specifically, we divided risk into 3 levels (low, moderate, and high) that would be based on the woman's age, breast density, and family history of breast cancer -- information likely to be available through physician-patient discussion in the primary care setting. Several more sophisticated risk assessment algorithms are available, but for modeling purposes we opted to use a simplified risk algorithm based on just these three factors to maximize the feasibility and potential generalizability of this approach (see "Overall Breast Cancer Risk" below).

We had to make several broad assumptions in designing the model that are important because they limit the ability of the model to capture the nuances of patient behavior and the many variations in clinical care patterns that occur for individual patients. For example, we assumed perfect compliance for both mammography and supplemental screening in this analysis. While it is the case that actual compliance is always less than 100%, differences across studies in the definition of the time interval within which women are considered compliant as well as considerations of what constitutes screening vs. diagnostic mammography[183] precluded our use of a uniform, widely-accepted estimate for compliance across different imaging modalities.

The model assumes that supplemental screening would occur immediately after a negative mammography result, and that one year of follow-up is available as the reference standard for both mammography and supplemental screening results. For mammography, we needed to estimate as "inputs" several important numbers based on our review of the clinical evidence, including the number of cancers detected (i.e., true positives), cancers missed (i.e., interval cancers), recalls for further testing, biopsies performed, cancer "yield" per biopsy (i.e., percentage of biopsies with positive results) and false-positive results both after biopsy and without biopsy (i.e., recalled for further testing but no biopsy recommended). We developed similar inputs for each supplemental screening strategy, but we made a simplifying assumption that all positive supplemental screening tests would result in immediate biopsy, and so did not estimate recall rates (which would equal biopsy rates in this case) or false-positive results without biopsy. As noted in this review, supplemental screening has the potential to detect both cancers missed by mammography and additional cancers that would not have presented during the interval between mammography

screenings; we therefore included both types of cancer in our estimates for each supplemental modality.

Target Population

The population we modeled included all women age 40-74 except for those with known genetic susceptibility, a personal history of breast cancer, and/or a history of mantle radiation to the chest. We calculated the initial population size based on age- and gender-specific US Census data for New England.[184] The prevalence of the risk factors noted above was estimated to total 4.7% in the general screening population, and we reduced the population size accordingly (see Table 17 below for population-based model parameters and data sources). The resulting target population size was approximately 3.1 million New England women. The distribution of BI-RADS breast density within each age band was estimated based on data from a recent BCSC publication.[185] On a population basis, approximately 46% of the New England screening population would have heterogeneously dense (1.2 million) or extremely dense (250,000) breast tissue, totaling slightly more than 1.4 million women.

Table 17: Estimates of overall screening population and target population for supplemental screening, by age.

Population/Age	***Estimate***	***Sources***
Screening Population, N		2010 U.S. Census (New England)
40-49	1,025,082	Whittemore 2004[186]
50-59	1,078,482	SEER Cancer Statistics 2010[187]
60-69	787,168	Malone, 2006[188]
70-74	243,070	
TOTAL	3,133,801	
Heterogeneously Dense, %		Kerlikowske 2013[185]
40-49	45.6	
50-59	38.6	
60-69	29.0	
70-74	26.0	
Extremely Dense, %		Kerlikowske 2013[185]
40-49	13.7	
50-59	7.5	
60-69	3.8	
70-74	3.1	
Dense Breasts & Close Family History of Breast Cancer, %	22.7	Titus-Ernsthoff, 2006[189]

Overall Breast Cancer Risk

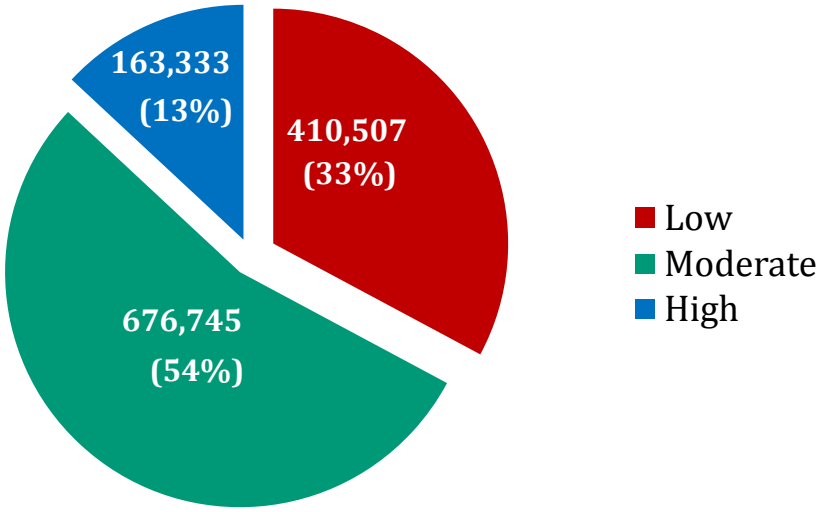
As described above, we limited risk factors for breast cancer in our assignment of risk category to age, breast density, and close family history (at least one 1st degree relative). The percentage of women with dense breast tissue and a close family history was estimated to be 22.7% based on data from a New Hampshire mammography registry study .[189] Using these three risk factors alone in the BCSC Risk Calculator,[109] we defined categories of low, moderate, and high risk as below:

- Low: BI-RADS density 3 or 4, age 40-49, no close family history (corresponds to 5-year risks generally <1.7%). Risk assumed in the model: 1% (0.2% per year)
- Moderate: BI-RADS density 3 or 4, age 40-49, with a close family history; OR BI-RADS density 3 or 4, age 50-74, no close family history (corresponds to 5-year risks generally between 1.7% and 3.0%). Risk assumed in the model: 2.5% (0.5% per year)
- High: BI-RADS density 3 or 4, age 50-74, with a close family history (corresponds to 5-year risks generally >3.0%). Risk assumed in the model: 5.0% (1.0% per year)

Support for these thresholds is available in the literature. Studies of chemoprevention generally consider a 5-year risk of approximately 1.7% to be the lower threshold for considering prophylaxis with tamoxifen or other measures to reduce breast cancer risk[123,190] while the U.S. Preventive Services Task Force's consideration of the same topic categorized women with 5-year risks >3% to be "high risk".[191]

Based on the risk categories described above, we estimate that, of all New England women with dense breast tissue and a negative digital mammogram, 33% would be low-risk, 54% moderate-risk and 13% high-risk. These proportions are displayed in Figure 2 on the following page along with the relevant estimated population sizes for each risk group.

Figure 2: Estimated numbers of New England women with dense breast tissue and negative mammography results, by level of overall breast cancer risk.



Test Diagnostic Performance

We obtained information on each test's diagnostic performance based on data identified during the systematic review. We used published data directly comparing the performance of film vs. digital mammography in the BCSC cohort[112] and applied these measures to the New England screening population; wherever possible, density-specific information was used (see Table 18 on the following page). Unfortunately, there are no available data documenting the performance of MRI in populations with the risk profile assumed for this model, and data from only a single study of DBT that lacked interval follow-up. In addition, there are multiple concerns with use of HHUS data as described in this review, including use of film mammography in most studies, lack of complete interval follow-up, and heterogeneity of study populations. We opted to estimate rates of detection of cancers missed on mammography (i.e., interval cancers) based on an extrapolation of the relative sensitivity of HHUS to MRI in similar populations (see Table 19 on page 86). Specifically, we assumed that HHUS and MRI would detect 80% and 95% of cancers respectively when all risk groups are considered. In the absence of data, we assumed that all measures of ABUS performance would be equivalent to that of HHUS, and the two modalities would differ only in terms of cost.

Table 18: Estimates of mammography performance.

Measure	*Estimate*		*Sources*
	Film	*Digital*	Kerlikowske 2011[111]
Sensitivity by BI-RADS Density, %			
Fatty	85.7	78.3	
Scattered Fibroglandular	85.1	86.6	
Heterogeneously Dense	79.3	82.1	
Extremely Dense	68.1	83.6	
Specificity by BI-RADS Density, %			
Fatty	95.4	94.7	
Scattered Fibroglandular	91.6	91.2	
Heterogeneously Dense	88.0	87.3	
Extremely Dense	89.8	88.7	
Recall Rate, per 1000 screened	93.0	100.0	
Biopsy Rate, per 1000 screened	10.6	11.0	
Cancer Yield per Biopsy, %	24.7	25.3	

BI-RADS: Breast Imaging-Reporting and Data System

Estimates of supplemental screening performance can be found in Table 19 on the following page. Information on DBT sensitivity in women with dense breast tissue and negative digital mammography is available,[159] but is overstated at 100% because of lack of interval follow-up. The rate of interval cancers in most screening populations is approximately 1 per 1,000 women screened.[112,192] Applying this rate to the Ciatto data yields a sensitivity of 75%, which was used for DBT in our model.

However, ultrasound, MRI, and DBT also detect additional cancers that would not have presented clinically during the interval between mammography screenings. Evidence from this review suggests that the rate of incremental cancer detection in relevant HHUS studies is approximately 3 per 1,000. The incremental cancer detection rate from the Ciatto study of DBT was 2.7 per 1,000. Because data from available MRI studies were representative of very high-risk women only, cancer detection rates from these studies would represent an overestimate in our target population. We therefore conservatively assumed that MRI would identify an additional 5 cases of cancer per 1,000. These studies were also assessed for information on cancer yield per biopsy. We assumed a rate of 8.5% for HHUS/ABUS based on a mean of values from the most representative studies identified. MRI's positive predictive value was 25% among women recommended for biopsy in a recent meta-analysis;[137] however, because the studies included in this review were conducted exclusively in high-risk women, we reduced our estimate to 20% to reflect MRI's potential use in a mixed-risk population. We assumed a cancer yield per biopsy of 25% for DBT based on findings from the one

DBT study that reported this value.[161] Cancer detection rates were divided by cancer yield estimates to obtain an estimate of the total number of biopsies performed for each modality.

Because the performance of screening tests has been found to improve with increasing disease prevalence,[193] we also assumed that all of these measures (i.e., detection of interval cancers, detection of additional cancers, and cancer yield per biopsy) would increase with increasing levels of breast cancer risk. We developed estimates for each risk subgroup that would equate to the overall levels described above when weighted by population size.

Definitions and estimates of the proportion of "overdiagnosed" cancers (i.e., those detected that would never have otherwise required treatment) vary substantially across studies. We estimated that between 10-30% of biopsy-detected cancers would be cases of overdiagnosis, based on the range generally reported in the literature.[91,95,194,195] We present the lower and upper boundaries of this range for each screening scenario evaluated.

Table 19: Estimates of supplemental screening performance.

	Risk Level				
Measure/Test	***Low***	***Moderate***	***High***	***Overall***	***Sources***
Sensitivity for Interval Cancers, %					
HHUS/ABUS	75.0	80.0	85.0	80.0	CTAF Review
MRI	87.5	95.0	100.0	95.0	CTAF Review
DBT	67.5	75.0	80.0	75.0	CTAF Review; Ciatto, 2013[159]
Additional Cancers Detected (per 1,000)					
HHUS/ABUS	1.5	3.5	5.0	3.0	CTAF Review
MRI	3.0	5.5	8.5	5.0	CTAF Review
DBT	1.3	3.3	4.5	2.7	CTAF Review; Ciatto, 2013[159]
Cancer Yield per Biopsy, %					
HHUS/ABUS	7.3	8.5	9.5	8.5	CTAF Review
MRI	15.0	20.0	25.0	20.0	CTAF Review; Warner, 2008[137]
DBT	20.0	25.0	30.0	25.0	CTAF Review; Rose, 2013[161]

NOTE: ABUS performance assumed to be equivalent to HHUS
HHUS: Handheld ultrasound; ABUS: Automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast

We adopted a third-party payer perspective for the model, and therefore focused on estimates of payment for screening, diagnostic imaging, and biopsy (see Table 20 on the following page). Payments were estimated primarily using the 2013 Medicare fee schedule.[196]

Table 20: Payment estimates for mammography, biopsy, and supplemental screening modalities.

Test/Procedure	*Components*	*Payment*	*Source(s)*
Screening Mammography **Film** **Digital**	Bilateral exam, computer-aided detection	 $91.87 $149.02	Medicare fee schedule
HHUS	Bilateral breast ultrasound	$100.37	Medicare fee schedule
ABUS	HHUS, +3D rendering	$183.05	Medicare fee schedule
MRI	Bilateral breast MRI, computer-aided detection	$636.23	Medicare fee schedule
DBT	Bilateral exam, computer-aided detection+additional views	$199.02	Medicare fee schedule, plus $50 additional patient fee
Diagnostic Mammography **Film** **Digital**	Unilateral exam, computer-aided detection	 $99.35 $143.58	Medicare fee schedule
Biopsy	Biopsy with ultrasound or stereotactic guidance, surgical biopsy	$923.30	Medicare fee schedule (assumes 75% percutaneous, 25% surgical)

HHUS: Handheld ultrasound; ABUS: Automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis

There is currently no separate and standardized reimbursement coding for ABUS or DBT. However, use of "add-on" codes has been reported for ABUS,[197] which are reflected in our estimates. Payment for DBT was based on the rate for digital mammography with an additional fee charged to the patient; we used an estimate of $50 based on the lower end of the reported range from facility websites.[198-200]

The costs of diagnostic workup for women recalled after positive mammography included those of a unilateral diagnostic mammogram in all and an HHUS exam in 50%, as well as biopsy in those so referred. We assumed that women presenting with an interval (i.e., "missed") cancer would

present clinically and receive both a unilateral diagnostic mammogram and a biopsy. For each supplemental screening strategy, costs of interest included those of screening, biopsy, and diagnosis of interval cancers.

9.3 Model Results

Population Estimates

As mentioned previously, 46% of the 3.1 million women age 40-74 in New England expected to undergo mammography screening would have BI-RADS density 3 or 4 (1.4 million). Of these women with dense breasts, 87% (1.25 million) would be expected to have a negative digital mammogram and therefore be candidates for supplemental screening.

Comparison of Film vs. Digital Mammography

The expected performance of film vs. digital mammography in New England is compared in Table 21 on the following page for the overall screened population as well as the subset of women with dense breast tissue. To facilitate comparisons, we present all clinical findings on a "per 1,000 women screened" basis, and costs are presented as an average per woman screened.

As shown in the table, digital mammography results in a small increase in the number of cancers detected (3.6 vs. 3.5 per 1000 for digital vs. film) and a small decrease in the number of cancers missed (0.7 vs. 0.8 per 1000) when compared to film mammography for the overall screening population. Rates of false-positive results with or without biopsy were somewhat higher for digital mammography, owing to its slightly lower specificity overall. Taking into account differences between digital and film mammography, approximately 12 biopsies would need to be performed with digital mammography in order to detect one additional cancer over film mammography (i.e., [14.8-13.6]/[3.6-3.5]). Total costs per woman screened were also higher for digital mammography ($184 vs. $120 for film), due to higher payments for the screening exam itself as well as higher rates of recall for diagnostic workup and biopsy.

Not surprisingly, recall and biopsy rates were higher in the subset of women with BI-RADS 3 or 4 breast density, as the incidence of cancer was higher with increasing breast density in the BCSC cohort. For example, cancer occurred at a rate of approximately 5 per 1000 in women with extremely dense breasts, vs. 2 per 1000 in those with fatty breasts (BI-RADS 1).

Table 21: Clinical outcomes and costs of general population screening mammography in New England: comparison of film vs. digital mammography.

Outcome (per 1,000 screened)	*Film*	*Digital*
Overall Population		
Recalls	99.0	105.1
Biopsies Performed	13.6	14.8
Cancers Detected (True Positives)	3.5	3.6
False Positive (with Biopsy)	10.1	11.2
False Positive (without Biopsy)	85.3	90.3
Cancers Missed (Interval Cancers)	0.8	0.7
Cost (per Woman Screened, $)	120	184
Women w/Dense Breast Tissue		
Recalls	120.0	128.0
Biopsies Performed	16.2	17.8
Cancers Detected (True Positives)	3.9	4.2
False Positive (with Biopsy)	12.3	13.6
False Positive (without Biopsy)	103.8	110.2
Cancers Missed (Interval Cancers)	1.1	0.9
Cost (per Woman Screened, $)	126	191

NOTE: Recalls refer to positive mammograms recalled for additional imaging and/or biopsy

Important differences in rates of cancers detected and cancers missed were apparent in women with dense breasts compared to the overall population. The increase in cancers detected, while still small, was threefold that observed in the overall cohort. The rate of cancers missed remained lower with digital mammography (0.9 vs. 1.1). Among women with dense breast tissue, the number of biopsies required with digital mammography to detect each additional cancer as compared to film mammography was much lower than that for the overall population (5 vs. 12 per 1,000 women screened).

Incremental Effects of Supplemental Screening in Women with Dense Breast Tissue and a Negative Digital Mammogram

We compared the four supplemental screening scenarios (HHUS, ABUS, MRI, and DBT) to no supplemental screening (i.e., digital mammography alone) on an overall basis as well as separately for low, moderate, and high-risk women. Results are described for each group of interest in the sections that follow.

Overall (All Risk Groups Combined)

Findings for the combined population of low-, moderate-, and high-risk women can be found in Table 22 below. As discussed previously, neither recalls nor false-positives without biopsy were estimated for these analyses, as all positive supplemental screening results were assumed to result in biopsy. We present clinical results for HHUS or ABUS together, as equivalent performance was assumed. Costs were assumed to differ, however, and are presented separately at the bottom of the table.

The addition of MRI to digital mammography detects more cancers (6.0 vs. 3.9 and 3.5 for HHUS/ABUS and DBT respectively). HHUS/ABUS would nearly quadruple the number of biopsies required over digital mammography alone, while biopsies would increase nearly threefold with MRI and less than twofold with DBT. Each of the supplemental modalities would identify nearly all of the cancers missed by mammography. MRI was the most costly strategy ($665), however, due to the higher payment rate for the test itself. Incremental costs for DBT were somewhat lower than for ABUS ($212 vs. $225 respectively) due to the lower biopsy rate estimated for DBT. Costs were lowest for HHUS ($142).

Table 22: Clinical outcomes and costs of supplemental screening in New England in all women with dense breast tissue and negative mammography: vs. digital mammography alone.

Outcome (per 1,000 screened)	***DM+HHUS/ABUS***	***DM+MRI***	***DM+DBT***	***DM Alone***
Biopsies Performed	63.2	48.4	31.8	17.8
Incremental	*45.4*	*30.6*	*14.0*	
Cancers Detected (True Positives	8.1	10.2	7.7	4.2
Incremental	*3.9*	*6.0*	*3.5*	
Adjusted for potential overdiagnosis (low)	*3.5*	*5.4*	*3.1*	
Adjusted for potential overdiagnosis (high)	*2.7*	*4.2*	*2.4*	
False Positive Biopsy	55.1	38.2	24.1	13.6
Incremental	*41.5*	*24.6*	*10.5*	
Cancers Missed (Interval Cancers)	0.2	0.1	0.3	0.9
Incremental	*(0.7)*	*(0.8)*	*(0.6)*	
Cost (per Woman Screened, $)	333/416	856	403	191
Incremental	*142/225*	*665*	*212*	

DM: Digital mammography; HHUS: Handheld ultrasound; ABUS: automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis

Low Risk

Clinical and economic outcomes of supplemental screening in low-risk (i.e., 5-year risk of 1%) women are presented in Table 23 below. In this low-risk population, HHUS (1.8) and DBT (1.5) identify a relatively small number of additional cancers, while MRI detects an incremental 3.4 per 1,000 screened. The number of biopsies required to detect this small number of cancers with MRI or HHUS is approximately five times higher than the rate for DM alone, however. DBT generates a smaller number of biopsies due to its lower false-positive rate and slightly lower rate of cancer detection. In this low-risk population, five, seven, and 14 biopsies would be required for DBT, MRI, and HHUS respectively to detect one additional cancer over digital mammography alone. As in the overall population, each of the supplemental strategies would detect nearly all of the cancers missed by mammography.

Table 23: Clinical outcomes and costs of supplemental screening in New England in women at low overall breast cancer risk with dense breast tissue and negative mammography: vs. digital mammography alone.

Outcome (per 1,000 screened)	***DM+HHUS/ABUS***	***DM+MRI***	***DM+DBT***	***DM Alone***
Biopsies Performed	31.3	28.7	13.9	6.2
Incremental	*25.1*	*22.5*	*7.7*	
Cancers Detected (True Positives)	3.4	5.0	3.1	1.6
Incremental	*1.8*	*3.4*	*1.5*	
Adjusted for potential overdiagnosis (low)	*1.6*	*3.0*	*1.3*	
Adjusted for potential overdiagnosis (high)	*1.2*	*2.4*	*1.0*	
False Positive Biopsy	27.9	23.7	10.8	4.6
Incremental	*23.3*	*19.1*	*6.2*	
Cancers Missed (Interval Cancers)	0.1	0.1	0.1	0.4
Incremental	*(0.3)*	*(0.3)*	*(0.3)*	
Cost (per Woman Screened, $)	309/391	842	391	185
Incremental	*124/206*	*657*	*206*	

DM: Digital mammography; HHUS: Handheld ultrasound; ABUS: automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis

All supplemental strategies would substantially increase screening costs compared with DM alone. Use of HHUS would increase costs by $124 (~60%) per woman screened, while the assumed greater expense for ABUS would more than double screening costs. Incremental costs with DBT would be identical to those with ABUS; while screening costs with DBT are higher than for ABUS, biopsy-

related costs would be lower. Screening costs would increase nearly fivefold compared to DM alone with the addition of MRI to DM.

Moderate Risk

Findings for patients in the moderate-risk group (5-year risk of 2.5%) can be found in Table 24 below. The higher prevalence of cancer in this subgroup is associated with higher rates of biopsy and false-positive results for all tests. HHUS/ABUS and DBT would detect an additional 4.4 and 4.1 cancers per 1,000 women screened respectively, while MRI would detect 6.5. The numbers of biopsies required to detect an additional cancer were four, five, and 12 for DBT, MRI, and HHUS respectively. There was some separation in the number of interval cancers that would have been missed by supplemental screening (0.1 for MRI vs. 0.2 and 0.3 for HHUS/ABUS and DBT respectively). Differences in cost were similar to those observed in the low-risk subgroup, although DBT becomes less costly than ABUS in this subgroup, as lower biopsy costs outweigh the additional costs of screening.

Table 24: Clinical outcomes and costs of supplemental screening in New England in women at moderate overall breast cancer risk with dense breast tissue and negative mammography: vs. digital mammography alone.

Outcome (per 1,000 screened)	***DM+HHUS/ABUS***	***DM+MRI***	***DM+DBT***	***DM Alone***
Biopsies Performed	66.8	48.1	31.7	15.5
Incremental	*51.3*	*32.6*	*16.2*	
Cancers Detected (True Positives	8.3	10.4	8.0	3.9
Incremental	*4.4*	*6.5*	*4.1*	
Adjusted for potential overdiagnosis (low)	*4.0*	*5.8*	*3.7*	
Adjusted for potential overdiagnosis (high)	*3.1*	*4.5*	*2.9*	
False Positive Biopsy	58.5	37.7	23.8	11.6
Incremental	*46.9*	*26.1*	*12.2*	
Cancers Missed (Interval Cancers)	0.2	0.1	0.3	1.1
Incremental	*(0.9)*	*(1.0)*	*(0.8)*	
Cost (per Woman Screened, $)	341/424	859	407	193
Incremental	*148/231*	*666*	*214*	

DM: Digital mammography; HHUS: Handheld ultrasound; ABUS: automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis

High Risk

Outcomes and costs of supplemental screening for women at high risk of breast cancer (5-year risk of 5%) are presented in Table 25 below. Greater than 10% of women would undergo a biopsy after screening with mammography or HHUS. The incremental rate of biopsy for HHUS would be nearly four times that of DBT in high-risk women (71.8 vs. 20.7 per 1,000 women screened).

Table 25: Clinical outcomes and costs of supplemental screening in New England in women at high overall breast cancer risk with dense breast tissue and negative mammography: vs. digital mammography alone.

Outcome (per 1,000 screened)	***DM+HHUS/ABUS***	***DM+MRI***	***DM+DBT***	***DM Alone***
Biopsies Performed	102.9	73.7	51.8	31.1
Incremental	*71.8*	*42.6*	*20.7*	
Cancers Detected (True Positives	14.7	18.5	14.1	7.9
Incremental	*6.8*	*10.6*	*6.2*	
Adjusted for potential overdiagnosis (low)	*6.1*	*9.5*	*5.6*	
Adjusted for potential overdiagnosis (high)	*4.7*	*7.4*	*4.3*	
False Positive Biopsy	88.2	55.1	37.7	23.2
Incremental	*65.0*	*31.9*	*14.5*	
Cancers Missed (Interval Cancers)	0.3	---	0.4	2.1
Incremental	*(1.8)*	*(2.1)*	*(1.7)*	
Cost (per Woman Screened, $)	366/449	875	418	199
Incremental	*167/250*	*676*	*219*	

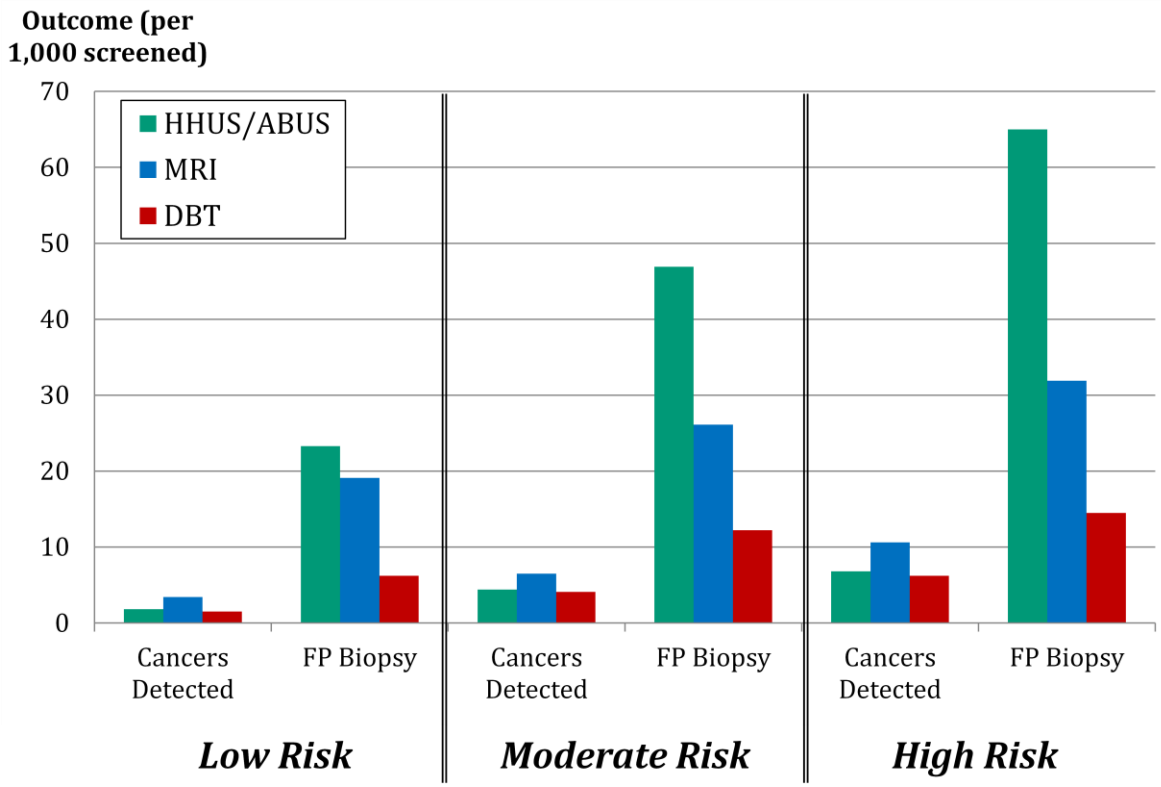
DM: Digital mammography; HHUS: Handheld ultrasound; ABUS: automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis

Estimated totals of three, four, and 11 biopsies would be required for DBT, MRI, and HHUS/ABUS to detect each additional case of cancer. MRI would correctly identify all cancers in high-risk women, although approximately one to three of the approximately 11 cancers identified have the potential to be cases of overdiagnosis. In addition, MRI would miss none of the cancers that would have been missed on mammography, while HHUS/ABUS and DBT would miss 0.3 and 0.4 cases per 1,000 respectively. Differences in false-positive rates are also magnified in the high-risk population. HHUS would produce a rate of false-positive biopsies more than twice that of MRI and nearly five times higher than that of DBT (65.0 vs. 31.9 and 14.5 per 1,000 respectively). MRI remained the most costly supplemental test strategy of the three modalities; including costs of mammography, an MRI-based strategy would cost nearly $900 per woman screened in the high-risk group.

Comparison of Risk Groups

Key incremental effects of supplemental screening (i.e., above and beyond effects of digital mammography alone) are compared for each risk group in Figure 3 below. While differences between modalities in the number of additional cancers detected remain relatively stable with increasing risk, differences in rates of false-positive biopsy become more pronounced. For example, HHUS/ABUS would produce 17.1 more false-positive biopsies per 1,000 women screened than DBT in low-risk women (23.3 vs. 6.2 per 1,000 respectively), but would generate 50.5 more per 1,000 among those in the high-risk group (65.0 vs. 14.5 per 1,000).

Figure 3: Selected incremental effects of supplemental screening, by screening modality and overall breast cancer risk.



HHUS: Handheld ultrasound; ABUS: automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis; FP: False positive

9. 4 Population Impact of Supplemental Screening

Estimates of the incremental clinical impact of supplemental screening with HHUS/ABUS, MRI, and DBT among all women with dense breast tissue in the state of New England can be found in Table 26 below.

Table 26: Population-based estimates of incremental clinical impact of supplemental screening among New England women with dense breast tissue and negative mammography results, by supplemental screening modality and overall breast cancer risk.

Outcome/Cost	*HHUS/ABUS*	*MRI*	*DBT*
Low Risk (n=410,507)			
Biopsies Performed	10,311	9,235	3,159
Cancers Detected (True Positives)	748	1,385	632
False Positive Biopsies	9,563	7,850	2,527
Cancers Missed (Interval Cancers)	44	22	57
Moderate Risk (n=676,745)			
Biopsies Performed	34,681	22,050	10,970
Cancers Detected (True Positives)	2,948	4,410	2,743
False Positive Biopsies	31,733	17,640	8,228
Cancers Missed (Interval Cancers)	145	36	181
High Risk (n=163,333)			
Biopsies Performed	11,724	6,951	3,382
Cancers Detected (True Positives)	1,114	1,738	1,015
False Positive Biopsies	10,610	5,214	2,367
Cancers Missed (Interval Cancers)	52	---	70

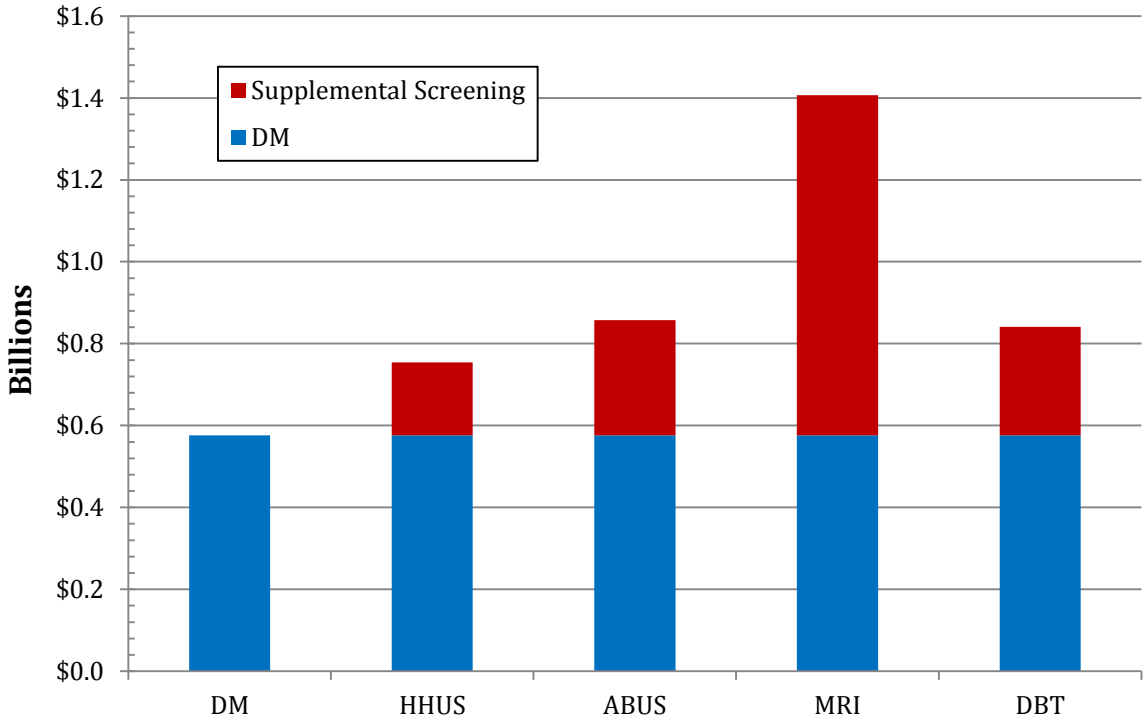
DM: Digital mammography; HHUS: Handheld ultrasound; ABUS: automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis

These results highlight the differences in clinical tradeoffs when supplemental screening strategies are used in the low- and high-risk populations. For example, the numbers of total and false-positive biopsies would be similar in these two subgroups, but supplemental screening would detect 350-400 more cancers in the high-risk population despite the fact that it is 40% the size of the low-risk group.

The estimated budgetary impact to New England of supplemental screening in all women with dense breasts and negative mammography can be found in Figure 4 on the following page. The annual cost of digital mammography screening, including costs of mammography, diagnostic workup, and biopsy, is estimated to total approximately $576 million. Supplemental screening of all women with an initial negative digital mammogram with HHUS would increase annual costs by approximately 30%, to $750 million. Use of higher-cost ABUS as the modality of choice would result

in a 49% increase in costs (to $860 million) for the same assumed clinical benefit. A similar cost increase would be seen with DBT; while assumed test costs would be higher with DBT vs. ABUS, costs of biopsy would be much lower. Finally, use of MRI results in a more than twofold increase in overall costs (to $1.4 billion) annually.

Figure 4: Costs of digital mammography and supplemental screening among New England women, by screening modality.

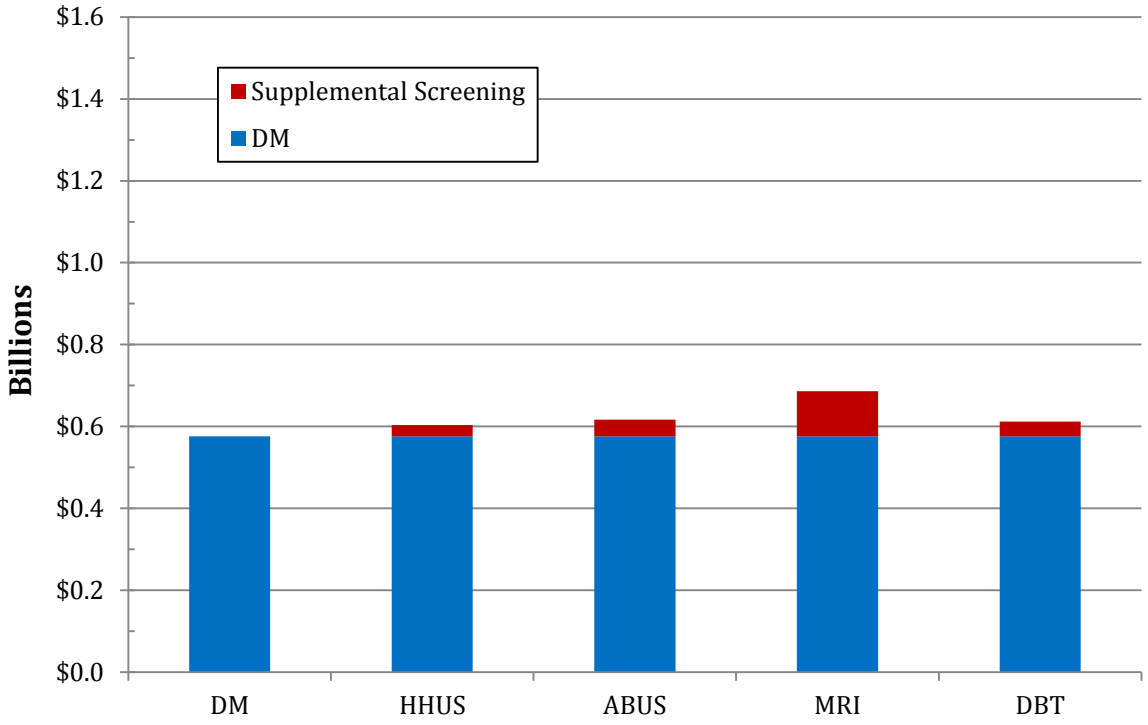


DM: Digital mammography; HHUS: Handheld ultrasound; ABUS: automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis

A substantial proportion of the additional costs of supplemental screening are generated in the low-risk population, the subgroup in which the fewest additional cancers are detected. Figure 5 on the following page shows the additional costs of supplemental screening when limited to women in the "high risk" category. If supplemental screening were limited to women age 50-74 with dense breast tissue, a family history in a first degree relative, and a negative digital mammogram (i.e., the high-risk cohort), total costs of screening would rise by a much smaller increment. However, the potential yield of additional cancers detected in this subgroup would be comparable to or better than with digital mammography alone. For example, supplemental MRI screening in high-risk women would increase costs by approximately $110 million (19%) to $690 million, and would find a total of 3,028 cases of cancer (1,290 cancers from digital mammography alone + 1,738 additional cancers from MRI). Increases in cost would be lower with the other supplemental modalities (5-

7%), but the additional cancer yield would also be lower (1,000 – 1,100 additional cancers detected over digital mammography alone). Findings such as these are important to consider in any evaluation of the tradeoffs of supplemental screening, including numbers of biopsies required, additional cancers detected and missed, and screening costs.

Figure 5: Costs of digital mammography and supplemental screening among "*high-risk*" New England women only, by screening modality.



DM: Digital mammography; HHUS: Handheld ultrasound; ABUS: automated breast ultrasound; MRI: magnetic resonance imaging; DBT: digital breast tomosynthesis

9.5 Model Limitations

We note important limitations of this cohort model. First, as required by any modeling approach, we made a number of simplifying assumptions that may not truly reflect the use of either mammography or supplemental screening in clinical practice. These included screening behaviors and clinical decisions such as perfect compliance with both types of screening as well as referral for and performance of biopsy in 100% of women with positive supplemental screening results. These assumptions likely resulted in overestimates of rates of cancer detection and cost for both mammography and supplemental screening.

Our most important assumption, however, was that each supplemental modality would identify additional cancers that would not have presented during the interval between mammography screenings, as has been demonstrated in the studies of interest for this review. However, these modalities have by and large not been studied exclusively in women with dense breast tissue and negative mammography who are at varying levels of overall breast cancer risk. It may be the case, for example, that we have overestimated the performance of HHUS/ABUS in high-risk individuals, as nearly all of the ultrasound studies evaluated in this review have been in women with breast cancer prevalence levels well below 1%. Conversely, we may also have overestimated MRI's performance in low-risk women, as the evidence base for supplemental MRI screening is currently limited to women at very high overall breast cancer risk.

We also included DBT in our analysis as a supplemental modality even though its eventual role may be to replace digital mammography as a general population screening tool and its overall evidence base is emerging, which may also have resulted in an overestimate of cost (due to the need for repeat digital mammography if not initially done in the same session as DBT). In addition, it appears that, at present, CMS will not provide additional reimbursement for DBT above and beyond existing payment for digital mammography alone, adding uncertainty to the cost of DBT to the payer and/or patient moving forward.[201]

Because the model adopted a payer perspective, we did not measure certain impacts of screening, such as potentially improved screening "throughput" with ABUS over HHUS as well as patient-time costs associated with each modality. Finally, while we attempted to provide reference figures for the number of cancers that might be "overdiagnosed" by these supplemental modalities, this did not explicitly consider the possibility that some proportion of cancers diagnosed by ultrasound or MRI would also have been diagnosed during the next round of mammography screening. This type of information will only be available through the conduct of longer-term randomized controlled trials or cohort studies comparing the benefits of supplemental screening to digital mammography alone.

10. Questions and Discussion

10.1 Introduction

During CEPAC public meetings, the Council deliberates and votes on key questions related to the systematic review of the evidence and the supplementary information presented. At the December 13, 2013 meeting, CEPAC discussed issues regarding the application of the available evidence to help patients, providers, and payers address the important questions of the benefits from supplemental screening for breast cancer in women with dense breast tissue and to support dialogue needed for successful action to improve the quality and value of health care in this population. The key questions are developed by the research team for each appraisal, with input from the CEPAC Advisory Board to ensure that the questions are framed to address the issues that are most important in applying the evidence to support clinical practice and medical policy decisions. Ex-officio CEPAC members participate fully in the discussion of the evidence but do not vote.

10.2 Summary of the Votes and Considerations for Policy

Following the evidence presentation and public comments, CEPAC voted on questions concerning the comparative clinical effectiveness and comparative value of supplemental breast cancer screening options for women with dense breast tissue. We present below the voting results along with comments reflecting the most important considerations mentioned by CEPAC members during the voting process.

Comparative Clinical Effectiveness

1. For women with dense breast tissue, is the evidence adequate to demonstrate that digital mammography offers superior diagnostic performance compared with film mammography?

   CEPAC Vote: **15 Yes** 0 No

2. For women with dense breast tissue, is the evidence adequate to demonstrate that, compared with film mammography, digital mammography substantially reduces the risk of "masking" of breast cancers?

   CEPAC Vote: **15 Yes** 0 No

3. For women with dense breast tissue with an overall "low" risk of breast cancer who have a negative screening digital mammogram, is the evidence adequate to demonstrate that supplemental screening with any technology provides more benefit than harm compared with no supplemental screening?

   CEPAC Vote: 5 Yes **10 No**

   Comments: CEPAC members who voted "no" cited the lack of direct studies in women with dense breasts and insufficient data on long-term patient outcomes, especially in women with lower levels of risk. In particular, CEPAC members emphasized their concern that supplemental screening among women with a lower prevalence of cancer would lead to a substantial increase in false positive results and unnecessary biopsies, which would in turn greatly increase patient anxiety and put women at risk of complications. Members also commented on the possibility of overdiagnosis in this population, as perhaps as many as 10%-30% of the cancers that could be detected by supplemental screening might never advance to threaten a woman's health and therefore would be unnecessarily treated as a result of supplemental detection.

   CEPAC members who voted "yes" pointed to the data on additional cancers detected and felt that the benefit of finding additional cancers outweighs the harms of false positive findings.

4. For women with dense breast tissue with an overall "moderate" risk of breast cancer who have a negative screening digital mammogram, is the evidence adequate to demonstrate that supplemental screening with any technology provides more benefit than harm compared with no supplemental screening?

   CEPAC Vote: **9 Yes** 6 No

   Comments: The trend toward more "yes" votes by CEPAC members reflects the shift in balance between benefits (additional cancers detected) and harms (anxiety and risk from false positive findings, and overdiagnosis and overtreatment). As the underlying risk of cancer increases, Council members were more likely to believe that the net health benefits of supplemental screening were positive for most women. Council members who voted "yes," however, stressed that the overall poor quality of evidence on patient outcomes made it difficult to determine precisely at what risk threshold supplemental screening would be expected to have a positive net benefit.

5. For women with dense breast tissue with an overall "high" risk of breast cancer who have a negative screening digital mammogram, is the evidence adequate to demonstrate that supplemental screening with any technology provides more benefit than harm compared with no supplemental screening?

   CEPAC Vote: **14 Yes** 0 No 1 Abstain

   Comments: In describing the rationale for their votes, CEPAC members noted the more robust data available from studies of supplemental screening among women with overall high risk of breast cancer, and commented again on the balance of benefits and harms being more likely to be positive among populations at high risk.

6. There are four options for supplemental screening reviewed in this report: hand-held ultrasound (HHUS), automated breast ultrasound (ABUS), magnetic resonance imaging (MRI), and digital breast tomosynthesis (DBT). Considering both the strength of evidence and the magnitude of potential comparative clinical benefits and harms of these four imaging modalities, if supplemental screening were to be performed for women with dense breast tissue who are at *high risk* of breast cancer, please rank in order, from highest to lowest preference, the tests you would recommend to a patient and her clinician. Health benefits and harms considered should include additional cancers detected and the possible impact on patient outcomes; false negative test results that miss critically significant cancers; false positive test results with their impact of unnecessary biopsies and anxiety; and overdiagnosis.

   CEPAC vote:

- **13 of 15 voting CEPAC members listed MRI as their first choice recommendation, and the remaining 2 members listed DBT as their first choice.**
- **9 of 15 voting CEPAC members listed ABUS as their least-preferred choice, and the remaining listed HHUS (4/15) and DBT (2/15) as their least-preferred choice.**

   Comments: CEPAC members voting for MRI noted that there was more direct evidence for the use of HHUS as a supplemental screening test, but that the evidence on MRI as a screening test among women at high risk of breast cancer suggested strongly that it would identify as many, if not more, additional cancers while producing far fewer false positive results. Though ABUS addresses many of the practical concerns related to HHUS, CEPAC members who voted for it as their least preferred option concluded that the evidence to support its use in high risk patients is much more limited, and findings to date have been largely inconsistent. The Council also noted that though the evidence for DBT is promising as a first-line screening option, there is no evidence examining its use as a supplemental screening test among women with dense breast tissue.

Comparative Value

When voting on comparative value, CEPAC was asked to assume the perspective of a state Medicaid agency or a provider organization that must make resource decisions within a fixed budget for care. While information about hypothetical budget tradeoffs are provided, CEPAC is not given prescribed boundaries or thresholds for budget impact or incremental cost-effectiveness ratios to guide its judgment of high, reasonable, or low value. For the CEPAC voting questions, comparative value is defined as the incremental cost to a public insurer for each supplemental screening option to achieve *net* health benefits, if any, in comparison to a "referent" screening option, in this case HHUS. The comparative net health benefit requires consideration of all relevant potential benefits and harms as described in the report.

7. HHUS is the lowest cost test for supplemental screening. If supplemental screening were to be performed for women with dense breast tissue who are at *high risk* of breast cancer, what is your judgment of the comparative value (high, reasonable, or low) of MRI vs. HHUS?

   CEPAC Vote: 5 High **9 Reasonable** 1 Low

   Comments: CEPAC members who voted that MRI represents "high" or "reasonable" value compared to HHUS maintained that MRI's superior balance of additional cancers detected vs. false positive results justified its higher costs and represented a reasonable use of healthcare resources if limited to the relatively small subpopulation of women at overall high risk of breast cancer. Some CEPAC members cautioned that without direct evidence on the effects of supplemental screening specifically among women with dense breast tissue or on the impact of supplemental screening on patient morbidity and mortality, the Council could be mistaken in its estimation of the value of additional testing, but that their judgment is based on the best evidence available on the effectiveness of supplemental screening with MRI for high risk women. The CEPAC member who voted that MRI represents "low" value stated that supplemental screening with MRI, even if limited to high risk women, represents a 25 percent increase in breast cancer spending with a relatively low yield in terms of demonstrated clinical benefit, and that those dollars could be better spent elsewhere in the health system.

   Note: The Council abstained from voting on the relative value of DBT and ABUS compared to HHUS due to insufficient evidence to demonstrate comparative clinical benefit between the various options.

Broader Considerations for Equity

8. Are there any considerations related to public health, equity, disparities in access or outcomes for specific patient populations, or other social values that should also be considered in medical policies related to the use of hand-held ultrasound (HHUS), automated breast ultrasound (ABUS), breast magnetic resonance imaging (MRI), or digital breast tomosynthesis (DBT)?

Comments:

- Consideration of the economic impact of supplemental screening should be broadened to consider the societal perspective, including considerations for missed work, transportation, and other costs.
- The policy community should be cautious when legislating in the area of supplemental screening, as it may have the unintended consequence of driving differential access to services and variation in practice. Some states in New England lack the capacity to sustain increased demand for public health screening, particularly in northern parts of the region.

10. 3 Roundtable Discussion and Key Policy Implications

Following CEPAC's deliberation on the evidence and subsequent voting, the Council engaged in a moderated discussion with a Roundtable composed of clinical experts, a patient advocate, regional health insurers, and provider group participants. The participants in the Roundtable discussion are shown in Appendix. The Roundtable discussion explored the implications of CEPAC's votes for clinical practice and medical policy, considered real life issues critical for developing best practice recommendations in this area, and identified potential avenues for applying the evidence to improve patient care. The main themes and recommended best practices from the conversation are summarized in the sections below.

1) Consideration of supplemental screening for women with dense breast tissue should be integrated within systems that assess their overall breast cancer risk and engage them in shared decision-making.

The Roundtable participants and CEPAC discussed the range of concerns regarding the issue of appropriate screening strategies for women based on their overall risk for breast cancer. Experts on the Roundtable noted that it is important for clinicians and women to understand that dense breast tissue conveys some increased risk for breast cancer but by itself is not a reason

to consider a woman at "high risk" of developing the disease. CEPAC members highlighted the need for more concerted efforts to develop standards or systems for clinical pathways to appropriately refer women for supplemental screening, and determine how breast density factors into those considerations. Roundtable experts from Connecticut, the only state in New England to enact a breast density notification law and mandate coverage for screening ultrasound in women with dense breasts, noted that practices vary in how they handle referrals for supplemental screening. In some practices, particularly in community settings, primary care physicians refer women with dense breast tissue for secondary screening automatically, regardless of overall risk status. In other practices, the decision to undergo additional screening is driven primarily by the patient; Roundtable participants reported that the proportion of patients requesting supplemental screening approximated 20% in some practices. Other experience at academic health centers in New England suggests that patients are not routinely sent for supplemental screening but rather are invited to consult with their primary care doctor about future screening options and help determine next steps. In these settings, women with certain risk factors are referred to a specialized breast cancer center and/or genetic counseling specialty department for further consultation.

CEPAC members stressed the importance of building systems in multi-disciplinary clinics that would be able to integrate the management of patients' questions arising from dense breast tissue notification with a reliable, efficient method for assessing their overall risk for breast cancer, and to share this information with patients. Systems should also support dialogue between patients and physicians regarding the various screening modalities available, and the patient's preference for additional screening. The ultimate goal of these systems should be to embody the principle of shared decision-making within mechanisms that would prove feasible across different practice settings.

2) Specialty societies, review groups, and others should seek to use consistent risk thresholds and assessment tools to capture overall breast cancer risk in order to avoid confusion among clinicians and patients.

CEPAC members and Roundtable panelists noted that risk assessment and stratification will never be able to identify 100% of women who will go on to develop breast cancer, and this fact should be communicated to patients. The CEPAC report noted the availability of the Gail model and the Breast Cancer Surveillance Consortium (BCSC) model for calculating breast cancer risk, both of which are sophisticated, computer-based algorithms. The CEPAC report also modeled the patient outcomes of different screening strategies based on categorizing risk according to a simplified, 3-variable version of the BCSC that could be used by primary care clinicians without direct access to computerized risk calculators. Discussion by the Roundtable

acknowledged that there is widespread variation in the use of risk assessment models across organizations and practices. Whereas 5-year risk thresholds make conceptual sense in considering supplemental screening, insurance coverage criteria for MRI are focused on determining lifetime risks high enough to warrant annual MRI. There is therefore great need for further efforts to develop robust systems for gathering breast cancer risk and applying that risk consistently to guide practice and policy decisions. Whatever approach is used to calculate risk, the Roundtable emphasized that breast cancer risk information should be readily available to clinicians and patients.

3) Physicians should adopt consistent messaging with their patients about breast density and breast cancer risk to help inform decisions for future screening.

CEPAC members and Roundtable panelists agreed that patients should be notified if they have dense breasts by their physician, but that messaging should promote a dialogue with patients and be consistent and clear in its explanation of the implications of dense breast tissue and options for additional screening. The patient advocate representative on the Roundtable cautioned that notifying a woman of her breast density status without further context can cause significant stress for the patient, but that this anxiety can be reduced through appropriate education. Communication with patients should highlight that, though dense breast tissue confers an increased risk of breast cancer, by itself it is not a reason to consider all such women at "high risk" for developing the disease. Women should also be made aware of the trade-offs involved in supplemental screening, and that though additional screening finds more cancer, it also increases the risk of false positives and unnecessary testing that can cause some women great anxiety and worry.

CEPAC members also stressed that states considering breast density notification policies should be careful to include an education component that helps patients understand the meaning of breast density, their risk, and the potential harms and benefits of supplemental screening.

4) More support is needed to help primary care physicians (PCPs) and other providers engage in discussions with their patients about breast density, risk, and options for supplemental screening.

The Roundtable and CEPAC noted the importance of developing educational materials for clinicians to help them understand the evidence on the various options for supplemental screening and provide a basis for discussions about these choices with women. Clinical experts

on the Roundtable indicated that discussion of these issues is appropriate in the primary care setting, but cautioned that many PCPs are already burdened with numerous clinical goals and a range of practice issues, limiting their capacity to have detailed conversations with patients about supplemental screening. CEPAC members agreed, however, that primary care physicians should be able to support patients in their decision-making about supplemental screening, and highlighted that this task can be made easier by coordinated provider education tools, electronic medical records (EMR), and the infrastructure provided by integrated health systems that can capture patient information on health risks at different points of entry in the health system.

In some states confronting a shortage of primary care physicians, other health personnel (e.g., physician assistants, nurse practitioners) are being trained to assume many of the roles currently played by primary care physicians. In these parts of New England, advanced training and education for clinicians on how to discuss with patients the issues around secondary screening and breast cancer risk are needed. The establishment of specialty departments or centers of excellence with automated referral processes from primary care may also be helpful in directing patients to specialists to discuss future screening options. Clinical experts also recommended that screening technologists be trained to identify patients who would benefit from further conversation with a radiologist.

Ensuring that radiologists, other specialists, and primary care clinicians share a common platform of information is critical to make certain that women receive consistent information and can participate with confidence in shared decision-making with their clinicians.

5) Greater guidance is needed to help physicians appropriately manage intervals for supplemental screening and subsequent follow-up for women with dense breast tissue.

Roundtable panelists remarked on the lack of guidance available to providers to help determine the appropriate supplemental screening intervals for women with dense breast tissue. Greater consensus on the best approaches for patient management of women with dense breasts is needed and should be reflected in clinical guidelines.

Evidence Development and Future Research Needs

CEPAC members underscored the importance of further research on supplemental screening among representative populations of women with dense breast tissue. Although data on long-term patient outcomes such as cancer-specific mortality would be ideal, it was recognized that such research would take too many years to be realistic, and that randomized trials or prospective cohort studies that follow all patients out for one year in order to capture interval cancers would be very informative.

CEPAC and Roundtable members agreed that if DBT supplants digital mammography as the primary screening test of choice, that new research will be required to evaluate the comparative benefits and harms of supplemental screening among women with dense breasts with ultrasound or MRI. If possible, studies evaluating DBT as a primary screening test could include an arm in which women with dense breasts who have a negative DBT receive supplemental screening in order to determine the incremental number of cancers detected, false positive rates, etc.

References

1. An act to add and repeal Section 123222.3 of the Health and Safety Code, relating to mammograms. In: Legislature CS, ed. *SB1538*.

2. Conn. Gen. Stat. § 38a-503 a-M, 2009).

3. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA: a cancer journal for clinicians.* Jan 2013;63(1):11-30.

4. Woloshin S, Schwartz LM, Welch HG. The risk of death by age, sex, and smoking status in the United States: putting health risks in context. *Journal of the National Cancer Institute.* Jun 18 2008;100(12):845-853.

5. Ries LAG, Melbert D, Krapcho M, et al., eds. *SEER Cancer Statistics Review, 1975-2005.* Bethesda, MD: National Cancer Institute; 2008.

6. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *The New England journal of medicine.* Oct 27 2005;353(17):1784-1792.

7. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *The New England journal of medicine.* Nov 22 2012;367(21):1998-2005.

8. Alexander FE. The Edinburgh Randomized Trial of Breast Cancer Screening. *Journal of the National Cancer Institute. Monographs.* 1997(22):31-35.

9. Andersson I, Aspegren K, Janzon L, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. *Bmj.* Oct 15 1988;297(6654):943-948.

10. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmo Mammographic Screening Program. *Journal of the National Cancer Institute. Monographs.* 1997(22):63-67.

11. Bjurstam N, Bjorneld L, Duffy SW, et al. The Gothenburg Breast Cancer Screening Trial: preliminary results on breast cancer mortality for women aged 39-49. *Journal of the National Cancer Institute. Monographs.* 1997(22):53-55.

12. Bjurstam N, Bjorneld L, Duffy SW, et al. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization. *Cancer.* Dec 1 1997;80(11):2091-2099.

13. Frisell J, Lidbrink E. The Stockholm Mammographic Screening Trial: Risks and benefits in age group 40-49 years. *Journal of the National Cancer Institute. Monographs.* 1997(22):49-51.

14. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Followup after 11 years--update of mortality results in the Stockholm mammographic screening trial. *Breast cancer research and treatment.* Sep 1997;45(3):263-270.

15. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. *Cmaj.* Nov 15 1992;147(10):1477-1488.

16. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *Cmaj.* Nov 15 1992;147(10):1459-1476.

17. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study: update on breast cancer mortality. *Journal of the National Cancer Institute. Monographs.* 1997(22):37-41.

18. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. *Journal of the National Cancer Institute.* Sep 20 2000;92(18):1490-1499.

19. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Annals of internal medicine.* Sep 3 2002;137(5 Part 1):305-312.

20. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet.* Mar 16 2002;359(9310):909-919.

21. Shapiro S. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan. *Journal of the National Cancer Institute. Monographs.* 1997(22):27-30.

22. Shapiro S, Venet W, Strax P, Venet L. Current results of the breast cancer screening randomized trial: the health insurance plan (HIP) of greater New York study. In: Miller NEDaAB, ed. *Screening for breast cancer*. Toronto Lewiston, N.Y.: Huber; 1988:3-15.

23. Tabar L, Fagerberg G, Chen HH, et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. *Cancer.* May 15 1995;75(10):2507-2517.

24. Tabar L, Fagerberg G, Duffy SW, Day NE. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit. *J Epidemiol Community Health.* Jun 1989;43(2):107-114.

25. Tabar L, Vitak B, Chen HH, et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am.* Jul 2000;38(4):625-651.

26. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet.* Dec 9 2006;368(9552):2053-2060.

27. Breast-cancer screening with mammography in women aged 40-49 years. Swedish Cancer Society and the Swedish National Board of Health and Welfare. *International journal of cancer. Journal international du cancer.* Dec 11 1996;68(6):693-699.

28. Armstrong K, Moye E, Williams S, Berlin JA, Reynolds EE. Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. *Annals of internal medicine.* Apr 3 2007;146(7):516-526.

29. Cox B. Variation in the effectiveness of breast screening by year of follow-up. *Journal of the National Cancer Institute. Monographs.* 1997(22):69-72.

30. Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. *Online J Curr Clin Trials.* Feb 25 1993;Doc No 32:[23,227 words; 195 paragraphs].

31. Glasziou PP. Meta-analysis adjusting for compliance: the example of screening for breast cancer. *Journal of clinical epidemiology.* Nov 1992;45(11):1251-1256.

32. Glasziou PP, Woodward AJ, Mahon CM. Mammographic screening trials for women aged under 50. A quality assessment and meta-analysis. *The Medical journal of Australia.* Jun 19 1995;162(12):625-629.

33. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev.* 2006(4):CD001877.

34. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet.* Jan 8 2000;355(9198):129-134.

35. Hendrick RE, Smith RA, Rutledge JH, 3rd, Smart CR. Benefit of screening mammography in women aged 40-49: a new meta-analysis of randomized controlled trials. *Journal of the National Cancer Institute. Monographs.* 1997(22):87-92.

36. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Annals of internal medicine.* Sep 3 2002;137(5 Part 1):347-360.

37. Kerlikowske K. Efficacy of screening mammography among women aged 40 to 49 years and 50 to 69 years: comparison of relative and absolute benefit. *Journal of the National Cancer Institute. Monographs.* 1997(22):79-86.

38. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography. A meta-analysis. *JAMA : the journal of the American Medical Association.* Jan 11 1995;273(2):149-154.

39. Nystrom L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet.* Apr 17 1993;341(8851):973-978.

40. Ringash J. Preventive health care, 2001 update: screening mammography among women aged 40-49 years at average risk of breast cancer. *Cmaj.* Feb 20 2001;164(4):469-476.

41. Smart CR, Hendrick RE, Rutledge JH, 3rd, Smith RA. Benefit of mammography screening in women ages 40 to 49 years. Current evidence from randomized controlled trials. *Cancer.* Apr 1 1995;75(7):1619-1626.

42. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine.* Nov 17 2009;151(10):716-726, W-236.

43. Bevers TB, Anderson BO, Bonaccio E, et al. NCCN clinical practice guidelines in oncology: breast cancer screening and diagnosis. *Journal of the National Comprehensive Cancer Network : JNCCN.* Nov 2009;7(10):1060-1096.

44. Lee CH, Dershaw DD, Kopans D, et al. Breast cancer screening with imaging: recommendations from the Society of Breast Imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *Journal of the American College of Radiology : JACR.* Jan 2010;7(1):18-27.

45. Smith RA, Brooks D, Cokkinides V, Saslow D, Brawley OW. Cancer screening in the United States, 2013: a review of current American Cancer Society guidelines, current issues in cancer screening, and new guidance on cervical cancer screening and lung cancer screening. *CA: a cancer journal for clinicians.* Mar-Apr 2013;63(2):88-105.

46. Feig SA, Yaffe MJ. Digital mammography. *Radiographics.* Jul-Aug 1998;18(4):893-901.

47. Pisano ED, Cole EB, Hemminger BM, et al. Image processing algorithms for digital mammography: a pictorial essay. *Radiographics.* Sep-Oct 2000;20(5):1479-1491.

48. Pisano ED, Zong S, Hemminger BM, et al. Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *J Digit Imaging.* Nov 1998;11(4):193-200.

49. United States Food and Drug Administration. Mammography Quality Standards Act and Program. 2013; http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/FacilityScorecard/ucm113858.htm. Accessed July 15, 2013, 2013.

50. Saslow D, Boetes C, Burke W, et al. American cancer society guidelines for breast screening with MRI as an adjunct to mammography. *CA: a cancer journal for clinicians.* Mar-Apr 2007;57(2):75-89.

51. Mahoney MC, Newell MS. Screening MR Imaging Versus Screening Ultrasound. Pros and Cons. *Magnetic Resonance Imaging Clinics of North America.* 2013.

52. Kelly KM, Richwald GA. Automated Whole-Breast Ultrasound: Advancing the Performance of Breast Cancer Screening. *Seminars in Ultrasound, CT and MRI.* 2011;32(4):273-280.

53. Giuliano V, Giuliano C. Improved breast cancer detection in asymptomatic women using 3D-automated breast ultrasound in mammographically dense breasts. *Clin Imaging.* May-Jun 2013;37(3):480-486.

54. Houssami N, Skaane P. Overview of the evidence on digital breast tomosynthesis in breast cancer detection. *Breast.* Apr 2013;22(2):101-108.

55. 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.

56. Kopans DB, Skaane P. Digital breast tomosynthesis: a better mammogram. *Radiology.* Jun 2013;267(3):968-969.

57. Drukteinis JS, Mooney BP, Flowers CI, Gatenby RA. Beyond mammography: new frontiers in breast cancer screening. *The American journal of medicine.* Jun 2013;126(6):472-479.

58. Yang WT. Emerging Techniques and Molecular Imaging in Breast Cancer. *Seminars in Ultrasound, CT and MRI.* 2011;32(4):288-299.

59. Berg WA, Gutierrez L, NessAiver MS, et al. Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer. *Radiology.* Dec 2004;233(3):830-849.

60. Hagen AI, Kvistad KA, Maehle L, et al. Sensitivity of MRI versus conventional screening in the diagnosis of BRCA-associated breast cancer in a national prospective series. *Breast.* Aug 2007;16(4):367-374.

61. Kriege M, Brekelmans CT, Boetes C, et al. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. *The New England journal of medicine.* Jul 29 2004;351(5):427-437.

62. Kuhl CK, Schrading S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* Nov 20 2005;23(33):8469-8476.

63. Leach MO, Boggis CR, Dixon AK, et al. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: a prospective multicentre cohort study (MARIBS). *Lancet.* May 21-27 2005;365(9473):1769-1778.

64. Sardanelli F, Podo F, D'Agnolo G, et al. Multicenter comparative multimodality surveillance of women at genetic-familial high risk for breast cancer (HIBCRIT study): interim results. *Radiology.* Mar 2007;242(3):698-715.

65. Tilanus-Linthorst MM, Obdeijn IM, Bartels KC, de Koning HJ, Oudkerk M. First experiences in screening women at high risk for breast cancer with MR imaging. *Breast cancer research and treatment.* Sep 2000;63(1):53-60.

66. Warner E, Plewes DB, Hill KA, et al. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. *JAMA : the journal of the American Medical Association.* Sep 15 2004;292(11):1317-1325.

67. Mann RM, Kuhl CK, Kinkel K, Boetes C. Breast MRI: guidelines from the European Society of Breast Imaging. *European radiology.* Jul 2008;18(7):1307-1318.

68. Berg WA, Blume JD, Cormack JB, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA : the journal of the American Medical Association.* May 14 2008;299(18):2151-2163.

69. Sechopoulos I. A review of breast tomosynthesis. Part I. The image acquisition process. *Med Phys.* Jan 2013;40(1):014301.

70. Viala J, Gignier P, Perret B, et al. Stereotactic vacuum-assisted biopsies on a digital breast 3D-tomosynthesis system. *Breast J.* Jan-Feb 2013;19(1):4-9.

71. *American College of Radiology: Breast imaging reporting and data system (BI-RADS).* 4th ed. Reston, VA: American College of Radiology; 2003.

72. Quality mammography standards--FDA. Final rule. *Fed Regist.* Oct 28 1997;62(208):55852-55994.

73. Linver MN, Osuch JR, Brenner RJ, Smith RA. The mammography audit: a primer for the mammography quality standards act (MQSA). *AJR Am J Roentgenol.* Jul 1995;165(1):19-25.

74. Rosenberg RD, Yankaskas BC, Abraham LA, et al. Performance benchmarks for screening mammography. *Radiology.* Oct 2006;241(1):55-66.

75. Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *The New England journal of medicine.* Oct 27 2005;353(17):1773-1783.

76. Keen JD, Keen JE. What is the point: will screening mammography save my life? *BMC Med Inform Decis Mak.* 2009;9:18.

77. Christiansen CL, Wang F, Barton MB, et al. Predicting the cumulative risk of false-positive mammograms. *Journal of the National Cancer Institute.* Oct 18 2000;92(20):1657-1666.

78. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *The New England journal of medicine.* Apr 16 1998;338(16):1089-1096.

79. Hofvind S, Thoresen S, Tretli S. The cumulative risk of a false-positive recall in the Norwegian Breast Cancer Screening Program. *Cancer.* Oct 1 2004;101(7):1501-1507.

80. Olivotto IA, Kan L, Coldman AJ. False positive rate of screening mammography. *The New England journal of medicine.* Aug 20 1998;339(8):560; author reply 563.

81. Hubbard RA, Kerlikowske K, Flowers CI, Yankaskas BC, Zhu W, Miglioretti DL. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Annals of internal medicine.* Oct 18 2011;155(8):481-492.

82. Barton MB, Moore S, Polk S, Shtatland E, Elmore JG, Fletcher SW. Increased patient concern after false-positive mammograms: clinician documentation and subsequent ambulatory visits. *J Gen Intern Med.* Mar 2001;16(3):150-156.

83. Barton MB, Morley DS, Moore S, et al. Decreasing women's anxieties after abnormal mammograms: a controlled trial. *Journal of the National Cancer Institute.* Apr 7 2004;96(7):529-538.

84. Lipkus IM, Halabi S, Strigo TS, Rimer BK. The impact of abnormal mammograms on psychosocial outcomes and subsequent screening. *Psychooncology.* Sep-Oct 2000;9(5):402-410.

85. Scaf-Klomp W, Sanderman R, van de Wiel HB, Otter R, van den Heuvel WJ. Distressed or relieved? Psychological side effects of breast cancer screening in The Netherlands. *J Epidemiol Community Health.* Dec 1997;51(6):705-710.

86. Brett J, Bankhead C, Henderson B, Watson E, Austoker J. The psychological impact of mammographic screening. A systematic review. *Psychooncology.* Nov 2005;14(11):917-938.

87. Weil JG, Hawker JI. Positive findings of mammography may lead to suicide. *Bmj.* Mar 8 1997;314(7082):754-755.

88. Woodward V, Webb C. Women's anxieties surrounding breast disorders: a systematic review of the literature. *J Adv Nurs.* Jan 2001;33(1):29-41.

89. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Annals of internal medicine.* Apr 3 2007;146(7):502-510.

90. Zahl PH, Maehlen J, Welch HG. The natural history of invasive breast cancers detected by screening mammography. *Archives of internal medicine.* Nov 24 2008;168(21):2311-2316.

91. Welch HG, Black WC. Overdiagnosis in cancer. *Journal of the National Cancer Institute.* May 5 2010;102(9):605-613.

92. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *Bmj.* 2009;339:b2587.

93. Welch HG. Overdiagnosis and mammography screening. *Bmj.* 2009;339:b1425.

94. Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. *Cancer causes & control : CCC.* Feb 2010;21(2):275-282.

95. de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiologic reviews.* Jul 2011;33(1):111-121.

96. Wu D, Perez A. A limited Review of Over Diagnosis Methods and Long Term Effects in Breast Cancer Screening. *Oncol Rev.* Sep 1 2011;5(3):143-147.

97. The benefits and harms of breast cancer screening: an independent review. *Lancet.* Nov 17 2012;380(9855):1778-1786.

98. Elmore JG, Fletcher SW. Overdiagnosis in breast cancer screening: time to tackle an underappreciated harm. *Annals of internal medicine.* Apr 3 2012;156(7):536-537.

99. Kalager M, Adami HO, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. *Annals of internal medicine.* Apr 3 2012;156(7):491-499.

100. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. *The Cochrane database of systematic reviews.* 2013;6:CD001877.

101. Boice JD, Jr. Radiation and breast carcinogenesis. *Medical and pediatric oncology.* May 2001;36(5):508-513.

102. Carmichael A, Sami AS, Dixon JM. Breast cancer risk among the survivors of atomic bomb and patients exposed to therapeutic ionising radiation. *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology.* Jun 2003;29(5):475-479.

103. de Gelder R, Draisma G, Heijnsdijk EA, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced vs prevented breast cancer deaths. *British journal of cancer.* Mar 29 2011;104(7):1214-1220.

104. Ng AK, Travis LB. Radiation therapy and breast cancer risk. *Journal of the National Comprehensive Cancer Network : JNCCN.* Nov 2009;7(10):1121-1128.

105. Ronckers CM, Erdmann CA, Land CE. Radiation and breast cancer: a review of current evidence. *Breast Cancer Res.* 2005;7(1):21-32.

106. Yaffe MJ, Mainprize JG. Risk of radiation-induced breast cancer from mammographic screening. *Radiology.* Jan 2011;258(1):98-105.

107. Hendrick RE, Pisano ED, Averbukh A, et al. Comparison of acquisition parameters and breast dose in digital mammography and screen-film mammography in the American College of Radiology Imaging Network digital mammographic imaging screening trial. *AJR. American journal of roentgenology.* Feb 2010;194(2):362-369.

108. Berrington de Gonzalez A, Reeves G. Mammographic screening before age 50 years in the UK: comparison of the radiation risks with the mortality benefits. *British journal of cancer.* Sep 5 2005;93(5):590-596.

109. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Annals of internal medicine.* Mar 4 2008;148(5):337-347.

110. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. *JAMA : the journal of the American Medical Association.* Jul 3 1996;276(1):33-38.

111. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Annals of internal medicine.* Feb 4 2003;138(3):168-175.

112. Kerlikowske K, Hubbard RA, Miglioretti DL, et al. Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. *Annals of internal medicine.* Oct 18 2011;155(8):493-502.

113. Pisano ED, Hendrick RE, Yaffe MJ, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology.* Feb 2008;246(2):376-383.

114. Vigeland E, Klaasen H, Klingen TA, Hofvind S, Skaane P. Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: the Vestfold County Study. *European radiology.* Jan 2008;18(1):183-191.

115. Hofvind S, Vacek PM, Skelly J, Weaver DL, Geller BM. Comparing screening mammography for early breast cancer detection in Vermont and Norway. *Journal of the National Cancer Institute.* Aug 6 2008;100(15):1082-1091.

116. Jensen A, Geller BM, Gard CC, et al. Performance of diagnostic mammography differs in the United States and Denmark. *International journal of cancer. Journal international du cancer.* Oct 15 2010;127(8):1905-1912.

117. Smith-Bindman R, Chu PW, Miglioretti DL, et al. Comparison of screening mammography in the United States and the United kingdom. *JAMA : the journal of the American Medical Association.* Oct 22 2003;290(16):2129-2137.

118. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer.* May 1976;37(5):2486-2492.

119. Cummings SR, Tice JA, Bauer S, et al. Prevention of breast cancer in postmenopausal women: approaches to estimating and reducing risk. *Journal of the National Cancer Institute.* Mar 18 2009;101(6):384-398.

120. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *The New England journal of medicine.* Jan 18 2007;356(3):227-236.

121. Yaghjyan L, Colditz GA, Rosner B, Tamimi RM. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to the time since the mammogram. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* Jun 2013;22(6):1110-1117.

122. Visvanathan K, Hurley P, Bantug E, et al. Use of Pharmacologic Interventions for Breast Cancer Risk Reduction: American Society of Clinical Oncology Clinical Practice Guideline. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* Jul 8 2013.

123. Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *Journal of the National Cancer Institute.* 1998;90(18):1371-1388.

124. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute.* Dec 20 1989;81(24):1879-1886.

125. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *Journal of the National Cancer Institute.* Sep 15 1999;91(18):1541-1548.

126. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *Journal of the National Cancer Institute.* Mar 7 2001;93(5):358-366.

127. Barlow WE, White E, Ballard-Barbash R, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. *Journal of the National Cancer Institute.* Sep 6 2006;98(17):1204-1214.

128. Chen J, Pee D, Ayyagari R, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *Journal of the National Cancer Institute.* Sep 6 2006;98(17):1215-1226.

129. National Comprehensive Cancer Network. Breast cancer screening and diagnosis. *NCCN Clinical Practice Guidelines in Oncology.* 2013(1.2013).

130. Solis Womens Health. Henda's Law; Supplemental screening for women with dense breast tissue and increased risk. .

131. Department of Health and Human Services Maine Center for disease control and prevention. LD 1886 Resolve Work Group final report. 2012.

132. Ind. Code § 25-22.5-13.2. 2013.

133. Federal Register, The daily journal of The United States Government. Mammography Quality and Standards Act; Regulatory Amendments.

134. H.R. 1302. Breast Density and Mammography Reporting Act of 2011. U.S. Congress. 2011.

135. D. K. Being less dense about breast cancer screening.

136. Lord SJ, Lei W, Craft P, et al. A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. *European journal of cancer.* Sep 2007;43(13):1905-1917.

137. Warner E, Messersmith H, Causer P, Eisen A, Shumak R, Plewes D. Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer. *Annals of internal medicine.* May 6 2008;148(9):671-679.

138. Gartlehner G, Thaler KJ, Chapman A, et al. Adjunct ultrasonography for breast cancer screening in women at average risk: a systematic review. *International journal of evidence-based healthcare.* Jun 2013;11(2):87-93.

139. Nothacker M, Duda V, Hahn M, et al. Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. *BMC cancer.* 2009;9:335.

140. Buchberger W, Niehoff A, Obrist P, DeKoekkoek-Doll P, Dunser M. Clinically and mammographically occult breast lesions: detection and classification with high-resolution sonography. *Semin Ultrasound CT MR.* Aug 2000;21(4):325-336.

141. Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. *Radiology.* Dec 2001;221(3):641-649.

142. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology.* Oct 2002;225(1):165-175.

143. Crystal P, Strano SD, Shcharynski S, Koretz MJ. Using sonography to screen women with mammographically dense breasts. *AJR Am J Roentgenol.* Jul 2003;181(1):177-182.

144. Leconte I, Feger C, Galant C, et al. Mammography and subsequent whole-breast sonography of nonpalpable breast cancers: the importance of radiologic breast density. *AJR. American journal of roentgenology.* Jun 2003;180(6):1675-1679.

145. Corsetti V, Houssami N, Ferrari A, et al. Breast screening with ultrasound in women with mammography-negative dense breasts: evidence on incremental cancer detection and false positives, and associated cost. *European journal of cancer.* Mar 2008;44(4):539-544.

146. Kolb TM, Lichy J, Newhouse JH. Occult cancer in women with dense breasts: detection with screening US--diagnostic yield and tumor characteristics. *Radiology.* Apr 1998;207(1):191-199.

147. Maestro C, Cazenave F, Marcy PY, Bruneton JN, Chauvel C, Bleuse A. Systematic ultrasonography in asymptomatic dense breasts. *Eur J Radiol.* Feb 1998;26(3):254-256.

148. Corsetti V, Ferrari A, Ghirardi M, et al. Role of ultrasonography in detecting mammographically occult breast carcinoma in women with dense breasts. *Radiol Med.* Apr 2006;111(3):440-448.

149. Brancato B, Bonardi R, Catarzi S, et al. Negligible advantages and excess costs of routine addition of breast ultrasonography to mammography in dense breasts. *Tumori.* Nov-Dec 2007;93(6):562-566.

150. De Felice C, Savelli S, Angeletti M, et al. Diagnostic utility of combined ultrasonography and mammography in the evaluation of women with mammographically dense breasts. *Journal of ultrasound.* Sep 2007;10(3):143-151.

151. Corsetti V, Houssami N, Ghirardi M, et al. Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: interval breast cancers at 1 year follow-up. *European journal of cancer.* May 2011;47(7):1021-1026.

152. Hooley RJ, Greenberg KL, Stackhouse RM, Geisel JL, Butler RS, Philpotts LE. Screening US in patients with mammographically dense breasts: initial experience with Connecticut Public Act 09-41. *Radiology.* Oct 2012;265(1):59-69.

153. Leong LC, Gogna A, Pant R, Ng FC, Sim LS. Supplementary breast ultrasound screening in Asian women with negative but dense mammograms-a pilot study. *Annals of the Academy of Medicine, Singapore.* Oct 2012;41(10):432-439.

154. Weigert J, Steenbergen S. The connecticut experiment: the role of ultrasound in the screening of women with dense breasts. *Breast J.* Nov-Dec 2012;18(6):517-522.

155. Girardi V, Tonegutti M, Ciatto S, Bonetti F. Breast ultrasound in 22,131 asymptomatic women with negative mammography. *Breast.* Apr 1 2013.

156. Parris T, Wakefield D, Frimmer H. Real world performance of screening breast ultrasound following enactment of Connecticut Bill 458. *Breast J.* Jan-Feb 2013;19(1):64-70.

157. Kelly KM, Dean J, Lee SJ, Comulada WS. Breast cancer detection: radiologists' performance using mammography with and without automated whole-breast ultrasound. *Eur Radiol.* Nov 2010;20(11):2557-2564.

158. Kelly KM, Dean J, Comulada WS, Lee SJ. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. *Eur Radiol.* Mar 2010;20(3):734-742.

159. 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. *The lancet oncology.* Jun 2013;14(7):583-589.

160. Haas BM, Kalra V, Geisel J, Raghu M, Durand M, Philpotts LE. Comparison of Tomosynthesis Plus Digital Mammography and Digital Mammography Alone for Breast Cancer Screening. *Radiology.* Jul 30 2013.

161. Rose SL, Tidwell AL, Bujnoch LJ, Kushwaha AC, Nordmann AS, Sexton R, Jr. Implementation of breast tomosynthesis in a routine screening practice: an observational study. *AJR Am J Roentgenol.* Jun 2013;200(6):1401-1408.

162. 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.* Apr 4 2013.

163. Berg WA, Zhang Z, Lehrer D, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA : the journal of the American Medical Association.* Apr 4 2012;307(13):1394-1404.

164. Kriege M, Brekelmans CT, Boetes C, et al. Differences between first and subsequent rounds of the MRISC breast cancer screening program for women with a familial or genetic predisposition. *Cancer.* Jun 1 2006;106(11):2318-2326.

165. Kuhl CK, Schmutzler RK, Leutner CC, et al. Breast MR imaging screening in 192 women proved or suspected to be carriers of a breast cancer susceptibility gene: preliminary results. *Radiology.* Apr 2000;215(1):267-279.

166. Lehman CD, Blume JD, Weatherall P, et al. Screening women at high risk for breast cancer with mammography and magnetic resonance imaging. *Cancer.* May 1 2005;103(9):1898-1905.

167. Lehman CD, Isaacs C, Schnall MD, et al. Cancer yield of mammography, MR, and US in high-risk women: prospective multi-institution breast cancer screening study. *Radiology.* Aug 2007;244(2):381-388.

168. Podo F, Sardanelli F, Canese R, et al. The Italian multi-centre project on evaluation of MRI and other imaging modalities in early detection of breast cancer in subjects at high genetic risk. *Journal of experimental & clinical cancer research : CR.* Sep 2002;21(3 Suppl):115-124.

169. Morris EA, Liberman L, Ballon DJ, et al. MRI of occult breast carcinoma in a high-risk population. *AJR Am J Roentgenol.* Sep 2003;181(3):619-626.

170. Port ER, Park A, Borgen PI, Morris E, Montgomery LL. Results of MRI screening for breast cancer in high-risk patients with LCIS and atypical hyperplasia. *Annals of surgical oncology.* Mar 2007;14(3):1051-1057.

171. Stoutjesdijk MJ, Boetes C, Jager GJ, et al. Magnetic resonance imaging and mammography in women with a hereditary risk of breast cancer. *Journal of the National Cancer Institute.* Jul 18 2001;93(14):1095-1102.

172. Yu J, Park A, Morris E, Liberman L, Borgen PI, King TA. MRI screening in a clinic population with a family history of breast cancer. *Annals of surgical oncology.* Feb 2008;15(2):452-461.

173. Warner E, Plewes DB, Shumak RS, et al. Comparison of breast magnetic resonance imaging, mammography, and ultrasound for surveillance of women at high risk for hereditary breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* Aug 1 2001;19(15):3524-3531.

174. Trecate G, Vergnaghi D, Manoukian S, et al. MRI in the early detection of breast cancer in women with high genetic risk. *Tumori.* Nov-Dec 2006;92(6):517-523.

175. Hartman AR, Daniel BL, Kurian AW, et al. Breast magnetic resonance image screening and ductal lavage in women at high genetic risk for breast carcinoma. *Cancer.* Feb 1 2004;100(3):479-489.

176. Kuhl C, Weigel S, Schrading S, et al. Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* Mar 20 2010;28(9):1450-1457.

177. Stoblen F, Landt S, Stelkens-Gebhardt R, Sehouli J, Rezai M, Kummel S. First evaluation of the diagnostic accuracy of an automated 3D ultrasound system in a breast screening setting. *Anticancer research.* Aug 2011;31(8):2569-2574.

178. Berg WA, Blume JD, Adams AM, et al. Reasons women at elevated risk of breast cancer refuse breast MR imaging screening: ACRIN 6666. *Radiology.* Jan 2010;254(1):79-87.

179. Pisano ED, Gatsonis CA, Yaffe MJ, et al. American College of Radiology Imaging Network digital mammographic imaging screening trial: objectives and methodology. *Radiology.* Aug 2005;236(2):404-412.

180. DeSantis C, Siegel R, Bandi P, Jemal A. Breast cancer statistics, 2011. *CA: a cancer journal for clinicians.* Nov-Dec 2011;61(6):409-418.

181. Venturini E, Losio C, Panizza P, et al. Tailored breast cancer screening program with microdose mammography, US, and MR Imaging: short-term results of a pilot study in 40-49-year-old women. *Radiology.* Aug 2013;268(2):347-355.

182. Tosteson AN, Stout NK, Fryback DG, et al. Cost-effectiveness of digital mammography breast cancer screening. *Annals of internal medicine.* Jan 1 2008;148(1):1-10.

183. Partin MR, Casey-Paal AL, Slater JS, Korn JE. Measuring mammography compliance: lessons learned from a survival analysis of screening behavior. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* Aug 1998;7(8):681-687.

184. U.S. Census Bureau, 2010: U.S. Census Bureau. New England Demographic Profile, 2010. Accessed at: http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?fpt=table, October 2013.

185. Kerlikowske K, Zhu W, Hubbard RA, et al. Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. *JAMA internal medicine.* May 13 2013;173(9):807-816.

186. Whittemore AS, Gong G, John EM, et al. Prevalence of BRCA1 mutation carriers among U.S. non-Hispanic Whites. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* Dec 2004;13(12):2078-2083.

187. SEER Cancer Statistics 2010: National Cancer Institute. Surveillance Epidemiology and End Results Database. SEER Cancer Statistics 2010. Accessed at: http://seer.cancer.gov/statfacts/html/breast.html, August 2013.

188. Malone KE, Daling JR, Doody DR, et al. Prevalence and predictors of BRCA1 and BRCA2 mutations in a population-based study of breast cancer in white and black American women ages 35 to 64 years. *Cancer research.* Aug 15 2006;66(16):8297-8308.

189. Titus-Ernstoff L, Tosteson AN, Kasales C, et al. Breast cancer risk factors in relation to breast density (United States). *Cancer causes & control : CCC.* Dec 2006;17(10):1281-1290.

190. Gail MH, Costantino JP, Bryant J, et al. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *Journal of the National Cancer Institute.* Nov 3 1999;91(21):1829-1846.

191. USPSTF, U.S. Preventive Services Task Force, Chemoprevention for Breast Cancer: Recommendations and Rationale (2002). Accessed at: http://www.uspreventiveservicestaskforce.org/3rduspstf/breastchemo/breastchemorr.pdf, August 2013.

192. NHSBSP- 2012 NHSBSP, National Collation of Breast Interval Data: Screening years 1st April 2003 - 31st March 2005 (2012). Accessed at: http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp-occasional-report1203.pdf, August 2013. .

193. Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *Journal of clinical epidemiology.* Jan 2009;62(1):5-12.

194. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmo mammographic screening trial: follow-up study. *Bmj.* Mar 25 2006;332(7543):689-692.

195. Independent UKPoBCS. The benefits and harms of breast cancer screening: an independent review. *Lancet.* Nov 17 2012;380(9855):1778-1786.

196. Centers for Medicare and Medicaid Services CfMaMSMPFSL-UT, 2013. Accessed at: http://www.cms.gov/apps/physician-fee-schedule/search/search-criteria.aspx, August 2013.

197. GE Healthcare U-S, a GE Healthcare Company. Current Coding Practices for Certain Breast Ultrasound Services. Accessed at: http://u-systems.com/healthcare/current-coding-practices-certain-breast-ultrasound-services, August 2013.

198. University of Virginia Health System UoVHSTDMAahucsb-cba--. Accessed at: http://uvahealth.com/services/breast-care/breastscreening/atproxiedcontent.2012-09-28.9643159495, August 2013.

199. South Jersey Radiology. 3D Digital Tomosynthesis Mammography. Accessed at: http://www.sjra.com/services-list/3d-digital-tomosynthesis-mammography/, August 2013.

200. Washington Radiology Associates. 3D Mammography. Accessed at: http://www.washingtonradiology.com/bca/, August 2013.

201. American College of Radiology. ACR Maintains its Coding Recommendation for Breast Tomosynthesis. Accessed at: http://www.acr.org/News-Publications/News/News-Articles/2013/Economics/20131114-ACR-Maintains-Its-Coding-Recommendation-for-Breast-Tomosynthesis, December 2013

Appendix

Search Strategies

PubMed (NLM) Search, run date 6/11/13

Search	Add to builder	Query	Items found	Time
#12	Add	Search #8 OR #10 AND English[la] Sort by: PublicationDate	909	12:17:33
#11	Add	Search #8 OR #10	1020	12:14:47
#8	Add	Search #5 OR #6 NOT (case reports[pt] OR news[pt] OR letter[pt] OR nursing journals[sb]) NOT (review[pt] NOT (random* OR systematic review* OR metaanaly* OR meta-analy*)) NOT (animals[mh] NOT humans[mh]) Sort by: PublicationDate	975	12:14:42
#10	Add	Search #9 NOT Medline[sb]	48	12:14:24
#9	Add	Search breast[ti] AND dens*[tiab] AND mammogra*[tw] AND (handheld[ti] OR hand-held[ti] OR imaging[ti] OR mri[ti] OR radiogra*[ti] OR tomosyn*[ti] OR ultraso*[ti] OR 3D[ti] OR 3-Dimension*[ti] OR three dimension*[ti])	390	12:14:24
#7	Add	Search #5 OR #6	1319	12:13:25
#6	Add	Search mammogra* AND tomosyn*	227	12:10:41
#5	Add	Search #1 AND #2 AND #3 AND #4	1134	12:10:41
#4	Add	Search magnetic resonance imaging[mh:noexp] OR (tomography, x-ray[mh] AND (tomosyn* OR 3D[tiab] OR 3-dimension*[tiab] OR three-dimension*[tiab])) OR imaging, three dimensional[mh:noexp] OR radiographic image enhancement[mh] OR ultrasonography, mammary[mh] OR ultrasonography[majr:noexp] Sort by: PublicationDate	550327	12:09:57
#3	Add	Search mammogra* Sort by: PublicationDate	29894	12:09:57
#2	Add	Search screen* Sort by: PublicationDate	478948	12:09:57
#1	Add	Search breast neoplasms[majr] OR breast neoplasms/diagnosis OR breast/pathology Sort by: PublicationDate	182235	12:09:56

Embase (Elsevier) Search , run date 6/11/13

No.	Query	Results
#9	#5 OR #6 OR #7 NOT ('case report'/exp OR letter/it OR ('review'/it OR 'short survey'/it NOT (random* OR 'systematic review' OR metaanalysis OR 'meta analysis')) OR 'conference abstract'/it) NOT ([animals]/lim NOT [humans]/lim) AND [english]/lim	**1338**
#8	#5 OR #6 OR #7	2516
#7	breast:ti AND dens*:ab,ti AND (mammogra*:ab,ti OR echomammography) AND (handheld:ti OR 'hand held':ti OR imaging:ti OR mri:ti OR radiogra*:ti OR tomosyn*:ti OR ultraso*:ti OR 3d:ti OR (3 NEXT/1 dimension*):ti OR (three NEXT/1 dimension*):ti)	503
#6	mammogra* AND tomosyn*	280
#5	#1 AND #2 AND #3 AND #4	1869
#4	'nuclear magnetic resonance'/exp OR ('tomography'/de AND (tomosyn* OR '3d' OR 3 NEXT/1 dimension* OR three NEXT/1 dimension*)) OR 'three dimensional imaging'/de OR 'image enhancement'/de OR 'echomammography'/de OR 'ultrasound'/mj	826390
#3	mammogra*	44163
#2	screen*	762860
#1	'breast tumor'/exp/mj OR 'breast tumor'/exp/dm_di OR ('breast'/exp AND (dense OR density OR densities)) OR breast NEAR/3 (dense OR density OR densities)	247494

The Cochrane Library (Wiley Online) Searches, run date 6/11/13

#1 breast:ti,ab,kw and screen* and mammogra* and magnetic resonance or tomosyn* or 3D or 3 dimension* or three dimension* or ultraso* or sonogra* or "image enhancement" or MRI (Word variations have been searched) 109

#2 MeSH descriptor: [Breast Neoplasms] explode all trees and with qualifiers: [Diagnosis - DI, Radiography - RA] 853

#3 MeSH descriptor: [Mass Screening] explode all trees 4457

#4 screen* 24418

#5 #3 or #4 24684

#6 #2 and #5 438

#7 MeSH descriptor: [Magnetic Resonance Imaging] this term only 4150

#8 MeSH descriptor: [Imaging, Three-Dimensional] this term only 534

#9 MeSH descriptor: [Radiographic Image Enhancement] explode all trees 3655

#10 MeSH descriptor: [Ultrasonography, Mammary] explode all trees 72

#11 MeSH descriptor: [Ultrasonography] this term only 790

#12 #7 or #8 or #9 or #10 or #11 8474

#13 magnetic resonance or MRI or tomosyn* or 3D or 3 dimension* or three dimension* or ultraso* or sonogra* or "image enhancement" (Word variations have been searched) 29302

#14 #6 and #12 and #13 52

#15 mammogra* and tomosyn* (Word variations have been searched) 3

#16 #1 or #14 or #15 **119**

◉ Cochrane Reviews (13) ◉ All ○ Review ○ Protocol ○ Other Reviews (25) ○ Trials (26) ○ Methods Studies (2) ○ Technology Assessments (21) ○ Economic Evaluations (31) ○ Cochrane Groups (1)

Notes:

Date of coverage:
Cochrane Database of Systematic Reviews : Issue 6 of 12, June 2013
Database of Abstracts of Reviews of Effect : Issue 2 of 4, April 2013
Cochrane Central Register of Controlled Trials (Central), Issue 5 of 12, May 2013
Methods Studies Issue 2 of 4, Apr 2013
Technology Assessments Issue 2 of 4 Apr 2013
Economic Evaluations
Cochrane Groups Issue 5 of 12, May 2013

Council Members and Roundtable Panelists

CEPAC Members *(in attendance)*	Disclosures
Ellen Andrews, PhD Executive Director, Connecticut Health Policy Project	
Robert H. Aseltine, Jr., PhD Professor, Division of Behavioral Sciences and Community Health, University of Connecticut Health Center Deputy Director, Center for Public Health and Health Policy Director, Institute for Public Health Research, University of Connecticut	Received consultancy fees in excess of $5,000 from CT State Medical Society on a project funded by United Healthcare. Contract is directly with CSMS.
Charles B. Eaton, MD, MS Professor, Family Medicine and Epidemiology, Alpert Medical School of Brown University Director, Center for Primary Care and Prevention, Director, Heart Disease Prevention Center Memorial Hospital of Rhode Island	
Teresa Fama, MD, MS Rheumatologist, Central Vermont Rheumatology	
Austin Frakt, PhD Health Economist, VA Boston Healthcare System Associate Professor, Boston University School of Medicine Associate Professor, Boston University School of Public Health	
Claudia B. Gruss, MD, FACP, FACG, CNSC (Chair) Gastroenterologist and Internist, Arbor Medical Group, LLC	Wellpoint shares held jointly with spouse in excess of $10,000.
Felix Hernandez, MD, MMM Medical Director, Surgery Division Director of Undergraduate Medical Education, Eastern Maine Medical Center	
Christopher Jones, PhD Assistant Professor, Surgery Assistant Professor, Economics Director, Global Health Economics Unit, Center for Clinical and Translational Science, University of Vermont College of Medicine	Earned $4K plus expenses for attending a multiple sclerosis Advisory Board. Earned over $5K through ForMyOdds LLC (government contracts; no corporate conflicts as yet).
William Cyrus Jordan, MD, MPH Director, Vermont Medical Society Foundation for Research and Education	Johnson & Johnson, Medtronic Inc., and Pfizer Inc. stock owned in excess of $10,000.
Toni Kaeding, MS, RN Special Projects, Green Mountain Care Board	
Stephen Kogut, PhD, MBA, RPh Professor, University of Rhode Island College of Pharmacy	
Sandhya K. Rao, MD Associate Medical Director, Massachusetts General Physicians Organization	
Cynthia N. Rosenberg, MD (ex-officio) Senior Medical Director, Harvard Pilgrim Health Care	
Jeanne Ryer, MS Director, New Hampshire Citizens Health Initiative	
Keith A. Stahl, MD Physician and Medical Director, Catholic Medical Center	
Mitchell Stein, MBA (Vice-Chair) Policy Director, Consumers for Affordable Health Care	
Robert Zavoski, MD, MPH (ex-officio) Medical Director, Connecticut Department of Social Services	

Roundtable Panelists
Catherine S. Giess, MD*
Assistant Professor of Radiology, Harvard Medical School
Attending Radiologist, Brigham and Women's Hospital
Regina Hooley, MD*
Assistant Professor of Diagnostic Radiology, Yale University School of Medicine
Ana Lourenco, MD*
Assistant Professor of Diagnostic Imaging, Alpert Medical School of Brown University
Debra Madden*
Cancer Research Advocate/Patient Advocate
Mary T. Pronovost, MSN, MD*
Assistant Clinical Professor of Surgery (Oncology), Yale University School of Medicine
Medical Director, Norma F. Pfriem Breast Care Center, Bridgeport Hospital
Thomas Simpatico, MD
Chief Medical Officer, Department of Vermont Health Access
Jeffrey Tice, MD*
Associate Professor of Medicine, University of California, San Francisco

*No conflicts of interest to disclose, defined as more than $10,000 in healthcare company stock or more than $5,000 in honoraria or consultancies during the previous year from healthcare manufacturers or insurers